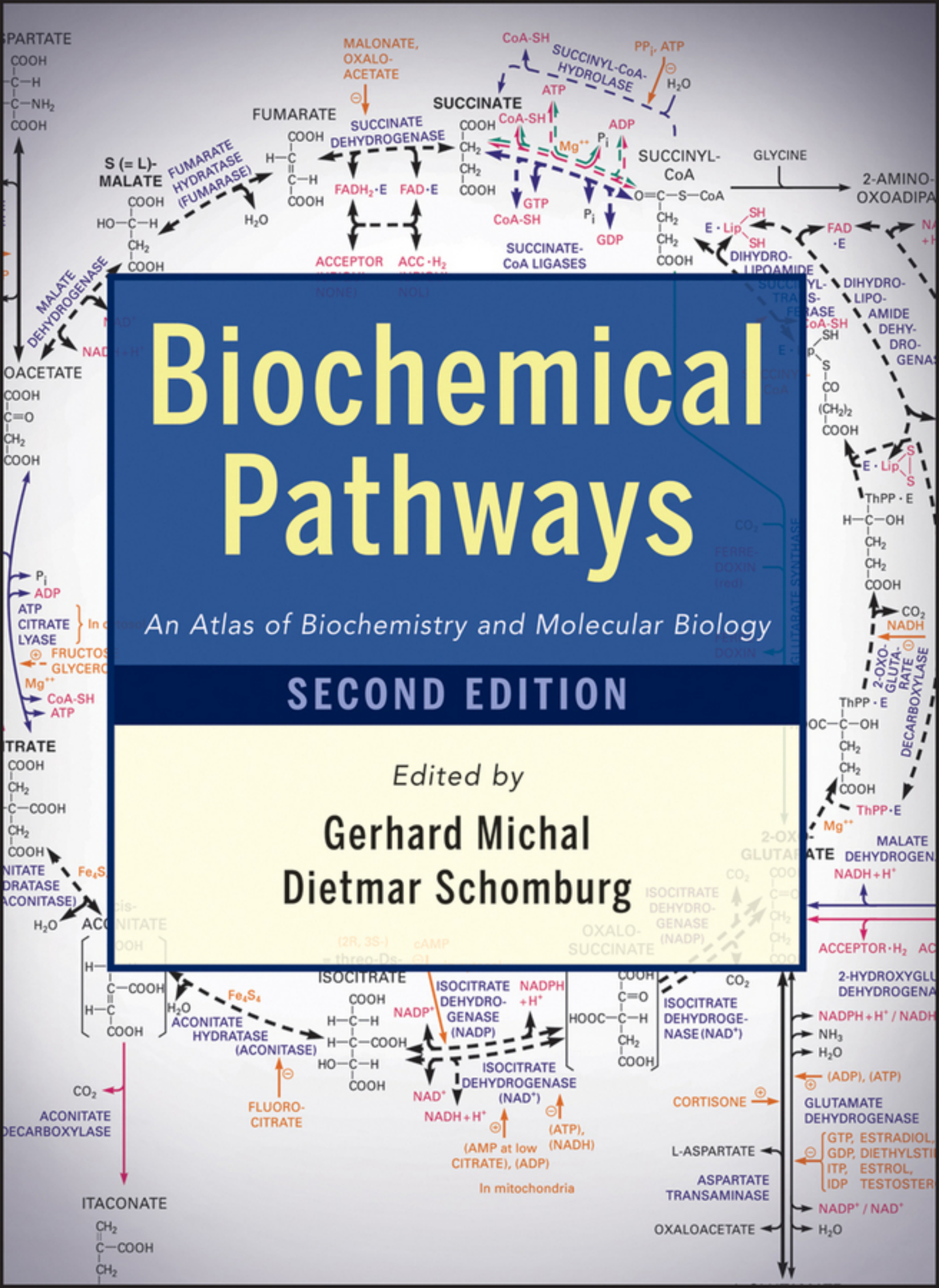

Biochemical Pathways
An Atlas of Biochemistry and Molecular Biology
SECOND EDITION
Edited by
Gerhard Michal
Dietmar Schomburg

BIOCHEMICAL PATHWAYS:
AN ATLAS OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

Biochemical Pathways Posters available from Roche Applied Science

Gerhard Michal's famous biochemical pathways posters have been a valuable resource for the global biochemistry community since 1968. Updated and revised, the *Biochemical Pathways Wallcharts* are an ideal companion to this text. Paired together, the book and wallcharts are perfect for researchers and students in biochemistry, biology, medicine, and physiology.

The wallcharts are also a great gift for anyone interested in following the myriad chemical reactions in our cells.To obtain this pair of large, detailed wall charts, contact Roche at https://www.roche-applied-science.com/techresources/index.jsp

BIOCHEMICAL PATHWAYS:

An Atlas of Biochemistry and Molecular Biology

Second Edition

Edited by

Gerhard Michal

Dietmar Schomburg

A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

Library of Congress Cataloging-in-Publication Data:

Biochemical pathways : an atlas of biochemistry and molecular biology / edited by Dietmar Schomburg, Gerhard Michal. -- 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-14684-2
1. Metabolism--Atlases. I. Schomburg, D. (Dietmar) II. Michal, Gerhard.
QP171.B685 2012
612.3'9--dc23

2011041441

Printed in the Singapore
10 9 8 7 6 5 4 3 2 1

Contents

Preface to the Second Edition

Since the publication of the first edition of 'Biochemical Pathways' in 1999 the molecular life sciences (encompassing biology, biochemistry, pharmacy and medicine) have undergone dramatic changes. With the extremely rapid development in the 'OMICS' analytical techniques (Genomics, transcriptomics, proteomics, metabolimics) we are in principle able to determine the genome of a microorganism in one day and a human genome for a couple of thousand dollars. We have also seen the advent of 'systems biology', which, based on the measured OMICS-data, aims at analysis and even prediction of biological functions by the construction of computer models. These models simulate the reaction of biological systems, including whole cells, to changes in the environment, genetic disorders or mutations.

Based on the annotation of the genome and experimental data, metabolic, regulatory and signal transduction pathways and networks are constructed and mathematically formulated. They depend entirely on our knowledge of biochemical pathways, as they are presented in this book.

As outlined in the preface to the first edition, one of us (GM) began early in the 1960s to combine an extract of the biochemical knowledge in a wall chart. The other of us (DS), towards the end of his student life saw the 'Biochemical Pathways' wall chart or 'Boehringer chart' in almost every lab working in the field of biochemistry or molecular biology. (At present, it is distributed as the 4th edition by Roche Diagnostics GmbH, Mannheim). He was impressed by the puzzle work biochemists had performed for almost one century. This presentation of important features of biochemistry was extended in the first edition of the 'Biochemical Pathways' book, which has become the standard book of reference in his and many other labs since then. In its focus on pathways and networks it is unique and was published exactly at a time when pathways, networks and systems became the focus of biochemical research. These areas have become the major fields of DS's research work in the last decade.

The fields of activities on both sides encouraged us to combine our experiences in writing and publishing the second edition of this book. The task became larger than expected on the first glance. Since the publication of the first edition our knowledge has increased tremendously. The selection of the facts to be dealt with and their condensation into a short, but legible form was no easy task. We could persuade expert authors to help us with the book. We both had a highly enjoyable cooperation and could now finally finish this work. We want to thank all authors for their contributions. In addition, Robbe Wünschiers likes to express his gratitude to Dr. Rainer Lemke for supporting the revision of the chapters.

The book not only gained one half in volume, but every sentence and every figure had to be checked and often modified. More than half of the many hundreds of figures in the book had to replaced, modified or added in this second edition.

We hope that it will help students and researchers to obtain a deeper understanding of the pathways and networks that determine biological functions.

Gerhard Michal
Dietmar Schomburg

From the Preface to the First Edition

This book is not intended to be a textbook of biochemistry in the conventional sense. There is no shortage of good biochemistry textbooks. which outline how biochemical knowledge has been gained, trace the logical and experimental developments in this field and present advances in their historical sequence.

In contrast, this book tries to condense important aspects of current knowledge. Its goal is to give concise information on the metabolic sequences in the pathways, the chemistry and enzymology of conversions, the regulation of turnover and the effect of disorders. This concentration on the sequence of facts has entailed the omission of researchers' names, experimental methods and the discussion of how results have been obtained. For information on these aspects, and for an introduction to the fundamentals of biological science, it is necessary to consult textbooks.

The scope of this book is general biochemistry, encompassing bacteria (and to some extent archaea), plants, yeasts and animals. Although a balanced representation is intended, personal interest naturally plays a role in the selection of topics. In a number of cases, the chemistry of the reactions is given in more detail, especially at metabolic key and branching points. Human metabolism, its regulation and disorders as a result of disease is a frequent topic. On the other hand, some chapters are especially devoted to bacterial metabolism.

This book grew out of my interest in metabolic interrelationships and regulation which was stimulated by my professional work at Boehringer Mannheim GmbH, Germany. Previously, this interest led me to compile the 'Biochemical Pathways'wall chart, the first edition of which appeared 40 years ago. Three more editions followed, which have been widely distributed. As a result of this experience, I developed a preference for the graphic presentation of scientific facts. In contrast to texts, illustrations allow the simultaneous display of different aspects, such as structural formulas, enzyme catalysis and its regulation, the involvement of cofactors, the occurrence of enzymes in various kingdoms of biology, etc. This form of presentation facilitates a rapid overview. A standard set of conventions is used in all illustrations (representation of formulas, symbols for proteins, the use of colors, the shape of arrows, etc. - the rare exceptions are indicated), and this assists in finding the facts quickly.

Tables have been added to provide more detailed information. They list additional properties of the system components, homologies, etc. The text plays only a supportive role. It gives a concise description of the reactions and their regulation, and puts them into the general metabolic context.

In many cases, current knowledge focuses on a limited in number of species. A rough classification of the occurrence of pathways is given by the color or the reaction arrows in the illustrations, but both generalizations and specialization are expected to be found in the future, which will necessitate modification of the picture.

The literature references have been limited in number and they usually cite recent review articles and books, if possible, from readily accessible sources. They were selected to provide more detailed information on new developments and additional references for the interested reader. There are no references to long-established biochemical facts which can be found in any textbook. I hope that this restriction will be acceptable to readers, since a complete listing of all sources for the statements presented here would take up a major portion of this volume. To compensate for the omission of such general references, a special chapter on electronic data banks and major printed sources has been added at the end of the book.

Most of all I want to thank my wife Dea, who has often encouraged me during the long time required to fiish this work. She has given me valuable advice and support in checking the text of the English edition. Without her understanding and her help this book would not have been brought to completion.

Gerhard Michal

Contributors

Helmut Burtscher, Roche Diagnostics GmbH, D-82372 Penzberg

Antje Chang, Enzymeta GmbH, D-50374 Erftstadt

Petra Dersch, Helmholtz Center for Infection Research, Dept. of Molecular Infection Biology, D-38124 Braunschweig

Julia Garbe, Institute for Microbiology, Technische Universität, D-38106 Braunschweig

Anton Haselbeck, MAB Discovery GmbH, D-82061 Neuried

Elmar Heinzle, Technical Biochemistry, Universität des Saarlandes, D-66123 Saarbrücken

Dieter Jahn, Institute for Microbiology, Technische Universität, D-38106 Braunschweig

Martina Jahn, Institute for Microbiology, Technische Universität, D-38106 Braunschweig

Wilhelm Just, Biochemistry Center, University, D-69120 Heidelberg

Horst Klima, Roche Diagnostics GmbH, D-82372 Penzberg

Klaus Klumpp, Hoffmann-La Roche Inc., Nutley NJ 07110

Gerhard Michal, Roche Diagnostics GmbH, D-82372 Penzberg, formerly Boehringer Mannheim GmbH (ret.)

Peter Müller, Helmholtz Centre for Infection Research, Dept. Gene Regulation & Differentiation, D-38124 Braunschweig

Gerhard Niederfellner, Roche Diagnostics GmbH, D-82372 Penzberg

Dieter Oesterheldt, Max-Planck Institute for Biochemistry, D-82152 Martinsried

Susanne Peifer, Technical Biochemistry, Universität des Saarlandes, D-66123 Saarbrücken

Ernst Peter Rieber, Institute for Immunology, Technische Universität, D-01011 Dresden

Stefan Ries, Roche Diagnostics GmbH, D-82372 Penzberg

Max Schobert, Institute for Microbiology, Technische Universität, D-38106 Braunschweig

Dietmar Schomburg, Institute of Biochemistry, Biotechnology & Bioinformatics, Technische Universität, D-38106 Braunschweig

Ida Schomburg, Enzymeta GmbH, D-50374 Erftstadt

Annika Steen, Institute for Microbiology, Technische Universität, D-38106 Braunschweig

Josef Wachtveitl, Institute for Physical and Theoretical Chemistry, University, D-60438 Frankfurt/M.

Röbbe Wünschiers, Biotechnology/Computational Biology, University of Applied Sciences, D-09648 Mittweida

1 Introduction and General Aspects

Gerhard Michal and Dietmar Schomburg

1.1 Organization of This Book

This book deals with the chemistry of living organisms. However, this topic cannot be considered in an isolated way, but has to be placed into a more general context. In two introductory chapters, a short outline of interconnections with neighboring sciences is given.

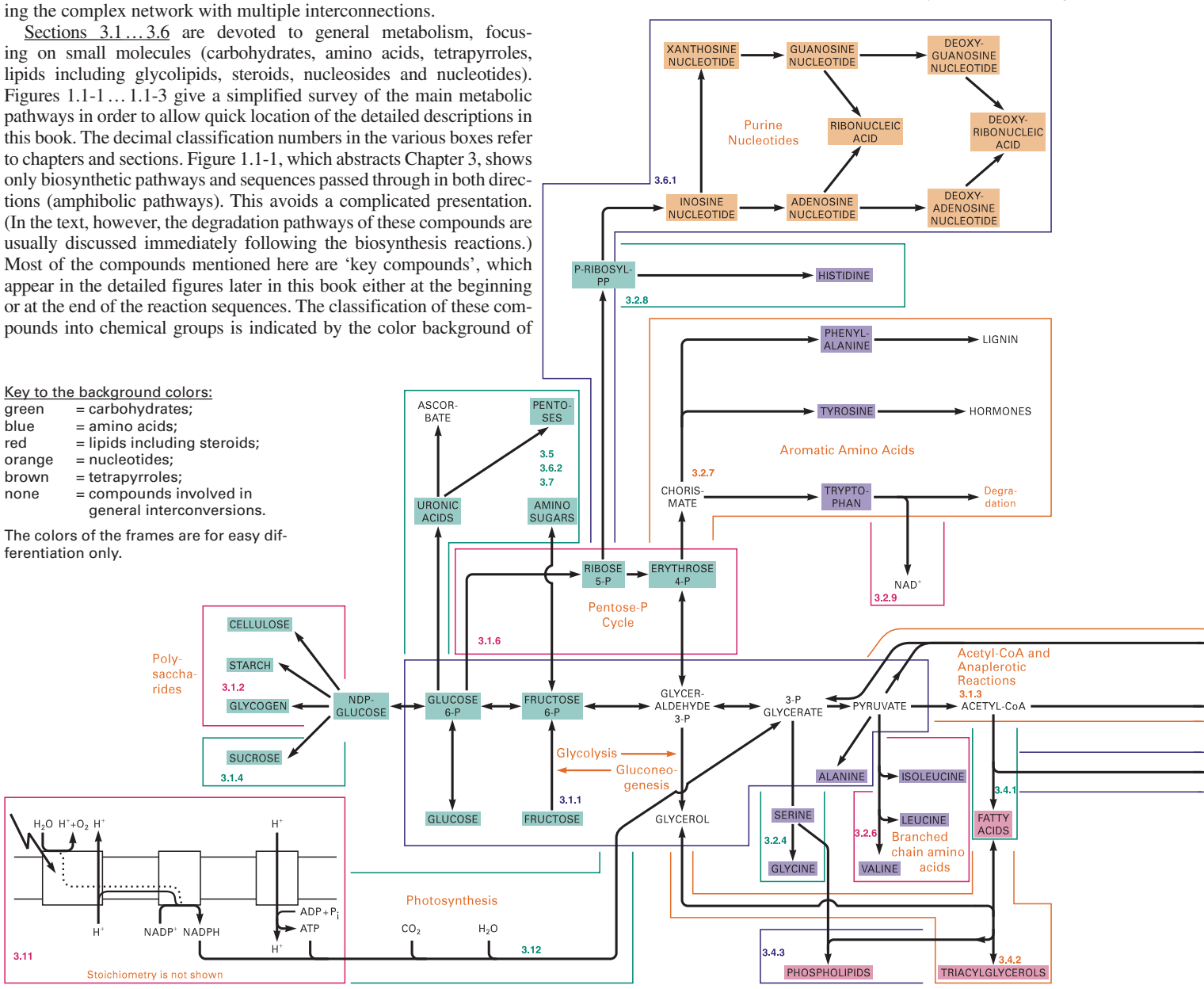
Chapter 1 deals with the organic chemistry of important components present in living organisms and with the physical chemistry of reactions.

Chapter 2 describes the overall organization of cells and their organelles as well as the structure of proteins and nucleic acids. This is followed by a discussion of enzyme function, which depends on the protein structure and regulates almost all biological processes.

The topics of **Chapter 3** are various aspects of metabolism, showing the complex network with multiple interconnections.

Sections 3.1 … 3.6 are devoted to general metabolism, focusing on small molecules (carbohydrates, amino acids, tetrapyrroles, lipids including glycolipids, steroids, nucleosides and nucleotides). Figures 1.1-1 … 1.1-3 give a simplified survey of the main metabolic pathways in order to allow quick location of the detailed descriptions in this book. The decimal classification numbers in the various boxes refer to chapters and sections. Figure 1.1-1, which abstracts Chapter 3, shows only biosynthetic pathways and sequences passed through in both directions (amphibolic pathways). This avoids a complicated presentation. (In the text, however, the degradation pathways of these compounds are usually discussed immediately following the biosynthesis reactions.) Most of the compounds mentioned here are 'key compounds', which appear in the detailed figures later in this book either at the beginning or at the end of the reaction sequences. The classification of these compounds into chemical groups is indicated by the color background of the names. Section 3.7 deals with cofactors and vitamins, which are involved in many reactions of general metabolism. Sections 3.8 and 3.9 describe the metabolism of DNA in bacteria and eukarya and the repair systems of these essential information carriers. The special metabolism of bacteria (including energy aspects), the biosynthesis and the effects of antibiotics are topics of Section 3.10. Aerobic respiration and its central role in energy turnover, as well as the photosynthetic reactions that are the source of almost all compounds in living beings, are discussed in Sections 3.11 and 3.12. Many special metabolic reactions take place in plants. These are summarized in Section 3.13.

The biosynthesis of proteins in bacteria and eukarya, and their consecutive modification, as well as the cell cycle, are discussed in **Chapter 4**. Figure 1.1-2 gives a short outline of these reactions, subdivided into bacterial reactions (left) and eukaryotic reactions (right).



Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Second Edition. Edited by Gerhard Michal and Dietmar Schomburg.

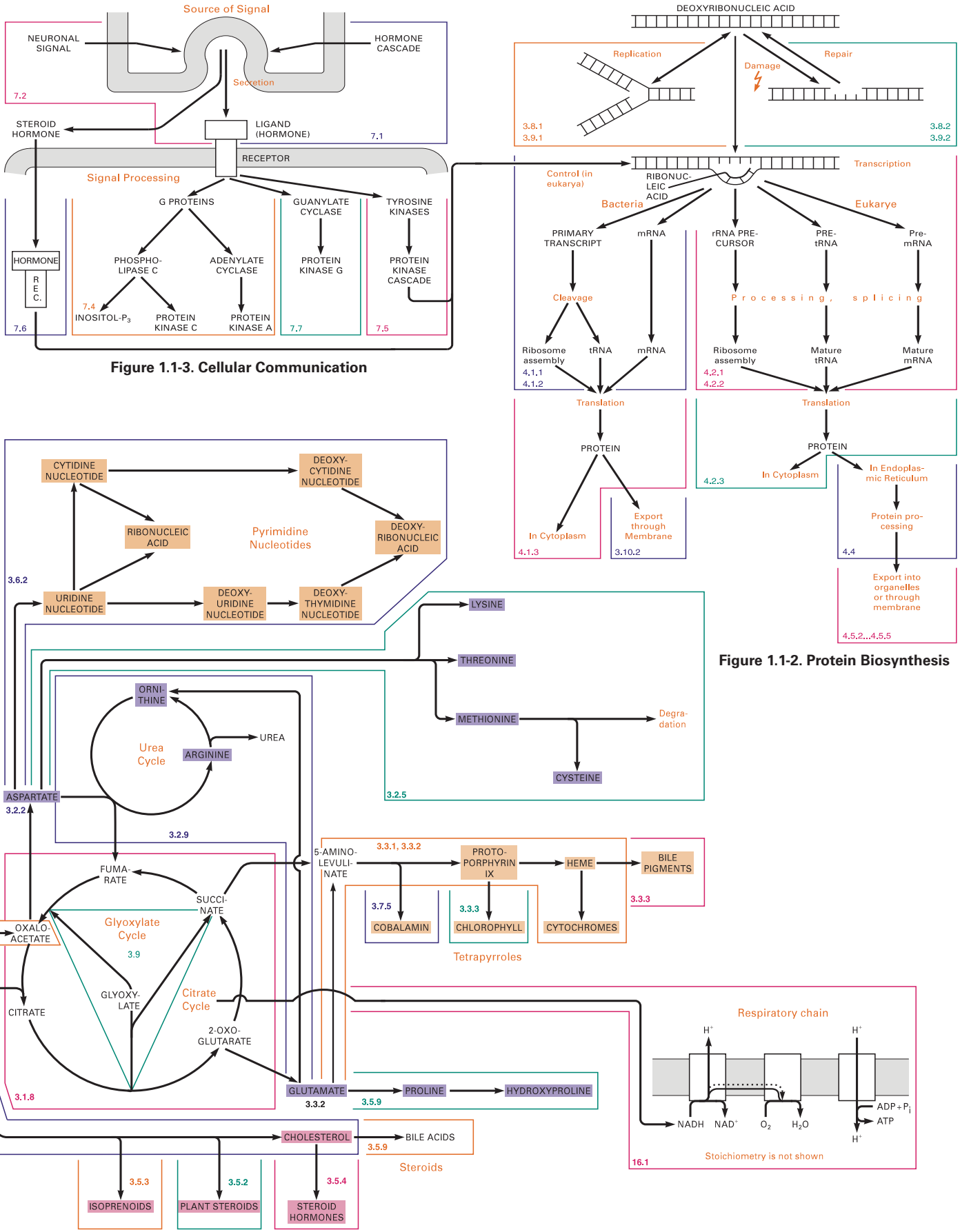


Figure 1.1-3. Cellular Communication

Figure 1.1-2. Protein Biosynthesis

Figure 1.1-1. Biosynthetic Reactions in General Metabolism

Viruses, which utilize these mechanisms in hosts, are discussed in **Chapter 5**.

Chapter 6 gives a survey of transport mechanisms through membranes and within vessels.

The topic of **Chapter 7** is cellular communication and the regulation mechanisms employed by multicellular organisms. Figure 1.1-3 briefly summarizes these multiple interconnections.

Chapter 8 deals with the defense mechanisms of higher animals and **Chapter 9** with blood coagulation.

Every presentation can only contain a selection of the present knowledge. For this reason, the final **Chapter 10,** is intended to assist in obtaining further information from electronic sources, which offer the most comprehensive collection of scientific results available today.

1.1.1 Conventions Used in This Book

1. A decimal classification system is used throughout with the following subdivisions: chapters, sections, subsections. Figures, tables, and formulas are assigned to the relevant sections, e.g., Figure 3.7.6-1.

Reactions:

2. Whenever available, the "Accepted Names" as defined by the IUMB Biochemical Nomenclature Committee are used for enzymes and substrates. The enzyme classification scheme (EC numbers) and the transporter classification scheme (TC numbers) are listed in the index.
3. Substrates of enzymatic reactions are printed in black, enzymes in blue, coenzymes in red. Regulatory effects are shown in orange. This color is also used for pathway names and for information on the location of a reaction. For numbering systems, green is used.
4. The color of the reaction arrows shows where the reaction was observed (or at least where reasonable indications for its occurrence exist): black = general pathway, blue = in animals, green = in plants and yeasts, red = in prokarya (bacteria and archaea).
5. Bold arrows indicate main pathways of metabolism.
6. Points on both ends of an arrow ↔ indicate noticeable reversibility of this reaction under biological conditions. Unless expressly noted, this type of arrow does not indicate mesomeric (resonance stabilized) states of a compound, contrary to usage in organic chemistry.
7. Double arrows ⇌ are used when the interconversion of two compounds proceeds via different reactions in each direction (e.g., for some steps of glycolysis).
8. Dashed reaction arrows show conversions with primarily catabolic (degradative) importance. Full line arrows show either mainly anabolic (biosynthetic) reactions or reactions in biological systems which are frequently passed through in both directions (amphibolic reactions).

Regulation:

9. Necessary cofactors, activating ions etc. are printed in orange next to a reaction arrow.
10. Full line orange arrows with an accompanying ⊕ or ⊖ indicate that the respective factor exerts 'fast' activation or inhibition of the reaction (by allosteric mechanisms, product inhibition etc.). Dashed arrows are used if the amount of enzyme protein is regulated, e.g., by varied expression or by changes in the degradation rate. If only one of multiple enzymes is regulated in this way, it is indicated by Roman numbers.

Enzymes and Proteins:

11. When enzyme complexes are involved, the respective components are schematically drawn in blue-lined boxes with rounded edges. This does not express the spatial structure. If possible, interacting components are drawn next to each other.
12. When a sequence of domains occurs in a protein, special symbols are used for the individual domains. They are explained next to the drawing.
13. When the peptide chain has to be shown, helices are drawn as  (e.g., in transmembrane domains), otherwise they are symbolized as .

Abbreviations and Notations:

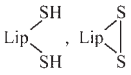
14. Organic phosphate is generally abbreviated as –P, inorganic phosphate and pyrophosphate as P_i and PP_i respectively. In drawings where the reaction mechanism is emphasized, phosphate residues are shown as $–O–PO_3^{2–}$.
15. Braces {} are used for atoms or residues which formally enter or leave during a reaction, if the molecular context is unknown.
16. While notations for genes are usually printed with small case letters (e.g., raf), the respective gene products (proteins) are written with a capitalized first letter (e.g., Raf). A number of proteins are defined by their molecular mass in kDa, e.g., p53.
17. When protein names are abbreviated, the notation frequently uses capitalized letters, e.g., cyclin dependent kinases = CDK in accordance with the literature.
18. A list of common abbreviations used throughout the book is given in 1.1.2. Less frequently used abbreviations are defined in the text.

Literature:

19. Only some recent references, primarily review articles and monographs, are listed at the end of the various sections. For more details refer to the literature quoted in these references, to electronic data banks, to review books and journals and to biochemistry textbooks.
20. Chapter 10 contains a survey on electronic data banks and a list of printed sources, which have been used frequently during the writing of this book.

1.1.2 Common Abbreviations (Other abbreviations are defined in the text)

aa	Amino acid
Acc, $AccH_2$	Acceptor, reduced acceptor (unspecified)
ACP	Acyl carrier protein
ATP, ADP, AMP, A	Adenosine tri-, di-, monophosphate, adenosine
bp, kbp	base pair (in DNA), kilobase pairs
cAMP	Cyclic AMP = adenosine 3,5-monophosphate
cGMP	Cyclic GMP = guanosine 3,5-monophosphate
CoA-SH, CoA-S-	Coenzyme A
CTP, CDP, CMP, C	Cytidine tri-, di-, monophosphate, cytidine
Cyt	Cytochrome
Da, kDa	Dalton, kilodalton (unit of molecular mass)
dATP, dADP, dAMP, dA	Deoxyadenosine tri-, di-, monophosphate, deoxyadenosine
dCTP, dCDP, dCMP, dC	Deoxycytidine tri-, di-, monophosphate, deoxycytidine
dGTP, dGDP, dGMP, dG	Deoxyguanosine tri-, di-, monophosphate, deoxyguanosine
dTTP, dTDP, dTMP, dT	Deoxythymidine tri-, di-, monophosphate, deoxythymidine
DNA	Deoxyribonucleic acid
E	Enzyme
EC number	Enzyme classification according to the IUBMB EC classification
ER	Endoplasmatic reticulum
ETF	Electron transferring flavoprotein
F_{430}	A corrinoid coenzyme (Ni)
FAD, $FADH_2$	Flavin-adenine dinucleotide, reduced flavin-adenine dinucleotide
Fd	Ferredoxin
FMN, $FMNH_2$	Flavin mononucleotide, reduced flavin mononucleotide
Fp	Flavoprotein
ΔG	Change of free energy (see 1.5.1)
G6P	Glucose 6-phosphate
GSH, GSSG	Glutathione, oxidized glutathione
GTP, GDP, GMP, G	Guanosine tri-, di-, monophosphate, guanosine
Ig	Immunoglobulin
ITP, IDP, IMP, I	Inosine tri-, di-, monophosphate, inosine
k	Velocity constant of a reaction (1.5.4)
K	Equilibrium constant of a reaction (see 1.5.1)
K_S, K_I, K_D	Dissociation constants (see 1.5.4, 7.1.2)
K_M	Michaelis constant (see 1.5.4)

kb	Kilobases (10^3 bases)
λ	Wavelength of light
	α-Lipoic acid, oxidized α-lipoic acid
NAD^+, NADH + H^+	Nicotinamide-adenine dinucleotide, reduced nicotinamide-adenine dinucleotide
$NADP^+$, NADPH + H^+	Nicotinamide-adenine dinucleotide phosphate, reduced nicotinamide-adenine dinucleotide phosphate
nt	Nucleotide
NTP, NDP, NMP, N	Any nucleotide tri-, di-, monophosphate or nucleoside
PAP	Adenosine 3′,5′-diphosphate
PAPS	3′-Phosphoadenylylsulfate
PEP	Phosphoenolpyruvate
P_i, Pp_i	Inorganic phosphate, inorganic pyrophosphate
pH	Negative decadic logarithm of the H^+ concentration
pK	Negative decadic logarithm of a dissociation constant
PQQ	Pyrroloquinoline quinone
PRPP	α-D-5-Phosphoribosylpyrophosphate
PyrP	Pyridoxal phosphate
RNA	Ribonucleic acid
mRNA, rRNA, tRNA	Messenger-, ribosomal-, transfer ribonucleic acid
R-S-S-R	Disulfide group of amino acids or peptides
S	Svedberg units (sedimentation coefficient)
SAH	S-Adenosylhomocysteine
SAM	S-Adenosylmethionine
THF	5,6,7,8-Tetrahydrofolate
THMPT	5,6,7,8-Tetrahydromethanopterin
ThPP	Thiamin pyrophosphate
UDPG	Uridine diphosphate glucose
UQ, UQH_2	Ubiquinone, reduced ubiquinone
UTP, UDP, UMP, U	Uridine tri-, di-, monophosphate, uridine

Abbreviations for amino acids are listed in Figure 1.3.2, abbreviations for sugars in Figure 4.4.1-1.

1.2 Carbohydrate Chemistry and Structure

Carbohydrate monomers are of the general formula $(CH_2O)_n$. They have the chemical structure of aldehydes or ketones with multiple hydroxyl groups (aldoses and ketoses, respectively). A common name of monomers and dimers is 'sugar'.

The large number of reactive groups, together with the stereoisomers causes a multiplicity of structures and reaction possibilities. Besides 'pure' carbohydrate monomers, oligomers (3.1.4) and polymers (3.1.2), carboxylic (3.1.5.1…2) and amino (3.1.7) derivatives, polyalcohols (3.1.5.5), deoxy sugars (3.1.5.6) etc., exist in nature.

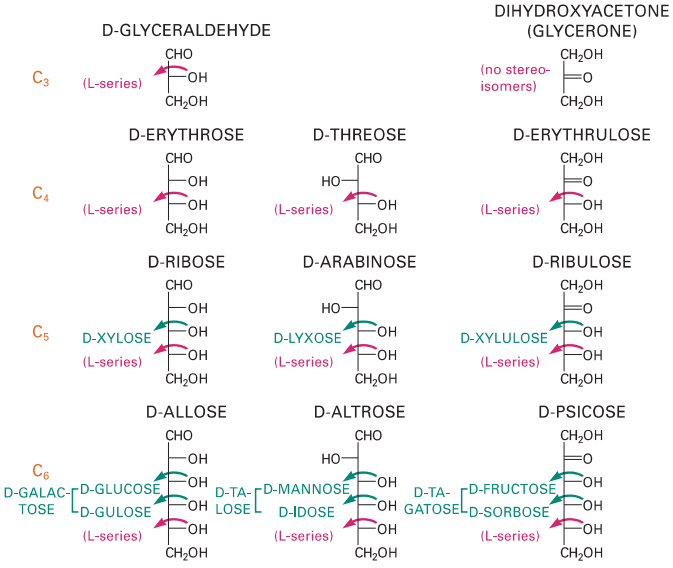


Figure 1.2-1. Nomenclature of Carbohydrates
The compounds printed in green are formally obtained by epimerization at the indicated positions. The L-enantiomers are the mirror images at the perpendicular mirror plane.

Carbohydrates are the primary products of photosynthesis (3.12.2) and function as energy storage forms (e.g., starch, glycogen, 3.1.2), as part of nucleic acid and nucleotide molecules (3.6.1, 3.6.2), in glycoproteins (4.4) and glycolipids (4.4) and as structural elements in cell walls of bacteria (3.10.1), plants (3.4) and in the exoskeleton of arthropods (3.1). They are the most abundant chemical group in the biosphere.

1.2.1 Structure and Classification

The simplest carbohydrates are the trioses (C_3 compounds) glyceraldehyde (an aldose) and dihydroxyacetone (glycerone, a ketose). Larger molecules are tetroses (C_4), pentoses (C_5), hexoses (C_6), heptoses (C_7) etc.; the C_5 and C_6 molecules are most common.

Glyceraldehyde is the smallest aldose with an asymmetric C-atom (chirality center). Therefore there are two stereoisomers (enantiomers), which cause right and left rotation of polarized light. By the Fischer convention, they are named D- and L-form, respectively. For details, see organic chemistry textbooks. Tetroses and larger carbohydrate monomers are classified (by comparison of the asymmetric center most distant to the aldehyde or keto group with D- or L-glyceraldehyde) as the D- and L-series of enantiomers (Fig. 1.2-1). With n-carbon aldoses, a total of 2^{n-2} stereoisomers exist, and with n-carbon ketoses there are 2^{n-3} stereoisomers. Epimers are stereoisomers, which differ in configuration at only one asymmetric C-atom. Most physiological sugars are of the D-configuration.

Aldopentoses, aldohexoses and ketohexoses (and higher sugars) can form cyclic structures (hemiacetals and hemiketals) by intramolecular reaction of their aldehyde or keto groups respectively with an alcohol group. This results in pyranoses (6-membered rings) and furanoses (5-membered rings, Fig. 1.2-2). In equilibrium, the cyclic structure is more prevalent as compared to the open structure. The ring closure produces another asymmetric C-atom; the respective stereoisomers are named anomers (α- and β-forms).

The nonplanar pyranose rings can assume either boat (in 2 variants) or chair conformation. The substituents extend either parallel to the perpendicular axis (axial, in Fig. 1.2-3 printed in red) or at almost right angles to it (equatorial, printed in green). The preferred

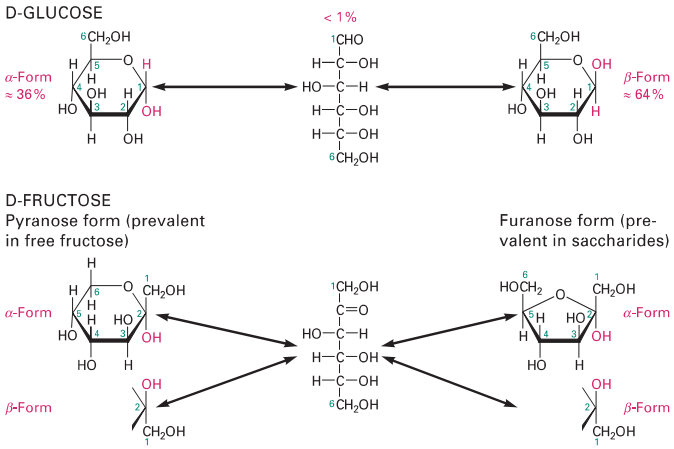


Figure 1.2-2. Ring Closure of Carbohydrates

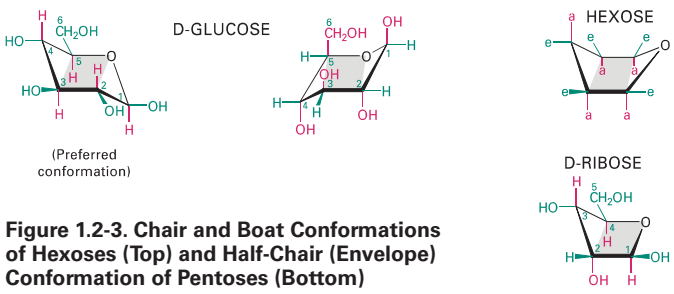


Figure 1.2-3. Chair and Boat Conformations of Hexoses (Top) and Half-Chair (Envelope) Conformation of Pentoses (Bottom)

conformation depends on spatial interference or other interactions of the substituents.

Although the bond angles of a furanose ring would permit an almost planar structure, the interference of substituents with each other causes a slight bending (puckering), e.g., to a half-chair (= envelope) structure in nucleotides and nucleic acids (Fig. 1.2-3).

The linear form of carbohydrates is usually shown as Fischer projection (ligands drawn horizontally are in front of the plane, ligands drawn vertically are behind the plane, e.g., in Fig. 1.2-1). The ring form is either drawn as Haworth formula (Fig. 1.2-2, disregarding the bent ring structure) or as boat/chair formula.

1.2.2 Glycosidic Bonds (Fig. 1.2-4)

If the hemiacetal or hemiketal hydroxyl of a sugar is condensed with an alcoholic hydroxyl of another sugar molecule, a glycosidic bond is formed and water is eliminated. Since this reaction between free sugars is endergonic ($\Delta G'_0 = 16$ kJ/mol), the sugars usually have to be activated as nucleotide derivatives (3.1.2.2) in order to be noticeably converted. Depending on the configuration at the hemiacetal/hemiketal hydroxyl (1.2.1), either α- or β-glycosides are formed. Sugar derivatives, which contain a hemiacetal or a hemiketal group (e.g., uronic acids) are also able to form glycosidic bonds.

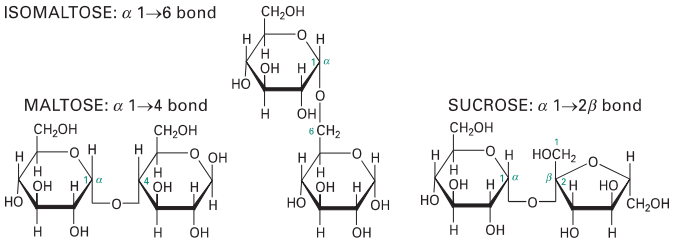


Figure 1.2-4. Examples of Glycosidic Bonds

Since sugar molecules contain several alcoholic groups, various types of bonds are possible. Frequently, 1 → 4 or 1 → 6 bonds occur. With oligo- or polysaccharides, both linear and branched structures are found. Bond formation may also take place with alcoholic, phenolic or other groups of non-sugar molecules (aglycons).

Literature:

Organic chemistry textbooks.

1.3 Amino Acid Chemistry and Structure

All amino acids present in proteins carry a carboxyl- and an amino group, hydrogen and variable side chains (R) at a single (α-)carbon atom. Thus, this C_α-atom is asymmetric (compare 1.2.1), with the exception of glycine, where R = H. Almost all of the proteinogenic amino acids occurring in nature are of the L-configuration. (The 'L' is assigned by comparison with L- and D-glyceraldehyde, which are taken as standards, Fig. 1.3-1). A number of D-amino acids are found in bacterial envelopes (3.10) and in some antibiotics (3.10).

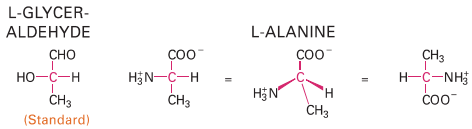


Figure 1.3-1. Asymmetric Center of Amino Acids

Unless otherwise stated, all amino acids discussed in the following sections are of the L-configuration.

Chains of amino acids form proteins and peptides. As enzymes, regulatory, mobility and structural compounds, they are the central components in all living beings. Therefore they are the topic of most of this book. Protein synthesis is described in Chapter 4. Their structure is discussed in Section 2.3, which also gives a short listing of their functions.

a) Non-polar, aliphatic amino acids. The non-polar side chains undergo hydrophobic interactions in protein structures. While the small glycine molecule allows high flexibility, the bulky proline confers enhanced rigidity to the structures.

GLYCINE (Gly, G) L-ALANINE (Ala, A) L-VALINE (Val, V) L-LEUCINE (Leu, L) L-ISOLEUCINE (Ile, I) L-PROLINE (Pro, P)

b) Polar, uncharged residues R. These functional groups are hydrophilic and can form hydrogen bonds with water or other polar compounds. Cysteine can easily be oxidized, resulting in intra- or intermolecular interconnections by disulfide bonds.

L-SERINE (Ser, S) L-THREONINE (Thr, T) L-CYSTEINE (Cys, C) L-METHIONINE (Met, M) L-ASPARAGINE (Asn, A) L-GLUTAMINE (Gln, Q)

c) Aromatic residues R. The aromatic side chains are hydrophobic, while the hydroxyl group of tyrosine and the ring nitrogen of tryptophan form hydrogen bonds, which often play a role in enzyme catalysis.

L-PHENYLALANINE (Phe, F) L-TYROSINE (Tyr, Y) L-TRYPTOPHAN (Trp, W)

d) Positively charged side chains R. The charged groups contribute in many cases to catalytic mechanisms and also influence the protein structure.

L-LYSINE (Lys, K) L-ARGININE (Arg, R) L-HISTIDINE (His, H)

e) Negatively charged side chains R. The charged groups contribute in many cases to catalytic mechanisms and are also of influence to the protein structure.

L-ASPARATE (Asp, D) L-GLUTAMATE (Glu, E)

L-PYRROLYSINE (Pyl, O) L-SELENOCYSTEINE (Sec, U)

f) Special amino acids using unusual genetic codes and found only in a small number of proteins (3.10.6.2, 4.1.3.4).

Figure 1.3-2. Amino Acids With Their 3- and 1-Letter Codes

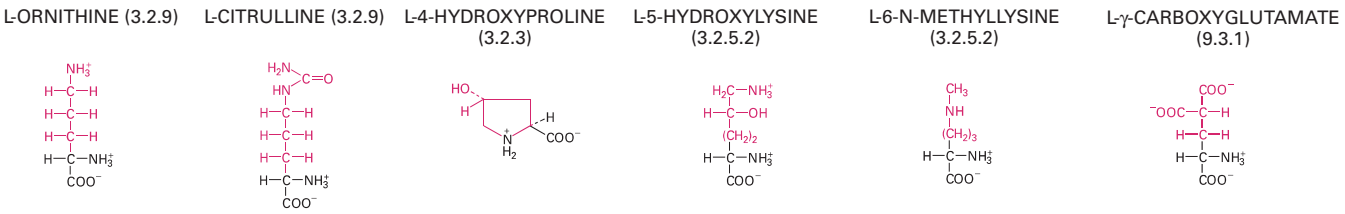


Figure 1.3-3. Some Nonstandard Amino acids

1.3.1 Structure and Classification

The individual properties of the amino acids are determined by the side chain R. This is also the criterion for amino acid classification.

There are 20 standard (classical) amino acids, which are incorporated as such into proteins, employing their own codons (4.1, 4.2). These amino acids are shown in Figure 1.3-2. Two additional amino acids, selenocysteine and pyrrolysine, are also incorporated directly by an unusual decoding procedure of mRNA (4.1). Nonstandard amino acids are produced by metabolic conversions of free amino acids (e.g., ornithine and citrulline) or by posttranslational modification of amino acids in proteins (e.g., by hydroxylation, methylation or carboxylation). Examples are given in Figure 1.3-3.

At about neutral pH, the free amino acids are 'Zwitterion' dipols with charged carboxylate (dissociation constant pK_1 = 1.82 … 2.35) and amino groups (pK_2 = 8.70 … 10.70). In seven cases, the side chains R also contain charged groups. Only the pK_α of histidine (3.2.8) is in the physiological range. In Figure 1.3-2 and 1.3-3, the charged molecules are shown, while in the rest of the book, un-ionized forms are presented for reasons of simplicity.

1.3.2 Peptide Bonds (Fig. 1.3-1)

Proteins and peptides are linear chains of amino acids connected by peptide bonds between their α-amino and carboxylate groups. Since the formation of these bonds is endergonic, the reactants have to be activated as tRNA derivatives. Details are described in 4.1.3.

The peptide bonds are rigid and planar: The carboxylate-O and the amino-H are in *trans* conformation, the C–N bond shows partially double bond characteristics. Only peptide bonds followed by proline or hydroxyproline can alternatively be *cis* (6 … 10 %). To some extent, both bonds in the backbone of the peptide chain extending from C_α can perform rotational movements (although there are still constraints on most conformations, which are shown in Ramachandran diagrams). Flexibility and constraints play a major role in the proper folding of the proteins (1.3.1).

Proteins and peptides carry charged amino- (N-) and carboxy- (C-) termini. Additional charges are contributed by the side chains. This allows analytical separation by electrophoresis. It has to be considered, however, that the pK_α of amino acids in peptides differ from those in free amino acids due to the effects of neighboring groups.

Literature:

Meister, A. *Biochemistry of the Amino Acids*. 2 Vols. Waltham, (MA):Academic Press; 1965.

Ramachandran, G.N., Sasisekharan, V. Adv. Prot. Chem. 1968;23:326–367.

Rose, G.D. *et al.* Adv. Prot. Chem. 1985;37:1–109.

Organic chemistry textbooks.

1.4 Lipid Chemistry and Structure

The common properties of lipids are their hydrophobic character and their solubility in organic solvents. Otherwise, they belong to different chemical classes. The biochemistry of most of them is described in Chapter 3.4, some other lipids are discussed in their metabolic context elsewhere (see cross-references below).

1.4.1 Fatty acids (Table 1.4-1, Fig. 1.4-1)

Fatty acids are characterized by a carboxylic group with a hydrocarbon 'tail'. The higher fatty acids are practically insoluble in water and show typical lipid properties. They serve in esterified form as triacylglycerols for energy storage or are, as glycerophospholipids, part of cellular membranes. In contrast, the short-chain fatty acids are water soluble. They act as intermediates of metabolism and are discussed in the respective chapters.

Higher fatty acids can also enter an amide bond (e.g., in ceramides). Some are precusors of other compounds (e.g., of prostaglandins, 7.4.8). Almost none of them occur in free form.

The predominant fatty acids in higher plants and animals have an even number of C atoms in the range of C_{14} … C_{20} and are unbranched. Usually, more than half of all fatty acids are unsaturated. Monounsaturated fatty acids mostly contain a *cis*-double bond between C-9 and C-10. Often additional double bonds exist towards the methyl terminus, usually with two saturated bonds in between (polyunsaturated fatty acids). Some of them cannot be synthesized in animals and have to be supplied by food intake (essential fatty acids). The notation of fatty acids is (number of C atoms) : (number of double bonds), e.g., for linoleic acid 18:2. The location of the double bonds is given as, e.g., $\Delta^{9,12}$.

Polyunsaturated fatty acids are not usually present in bacteria, but *cis*- and *trans*-monounsaturated, hydroxylated and branched fatty acids exist in many species.

Table 1.4-1. Higher Fatty Acids Frequently Occurring in Nature

Number of C atoms	Saturated	Unsaturated, Number of Double Bonds	E = Essential Fatty Acid for Humans
14	myristic acid		
16	palmitic acid		
18	stearic acid	1: oleic acid (Δ^9)	
		2: linoleic acid ($\Delta^{9,12}$)	E
		3: α-linolenic acid ($\Delta^{9,12,15}$)	E
20	arachidic acid	3: dihomo-γ-linolenic acid ($\Delta^{5,8,11}$)	
		4: arachidonic acid (AA, $\Delta^{5,8,11,14}$)	(E)[1]
		5: eicosapentaeonoic acid (EPA, $\Delta^{5,8,11,14,17}$)	
24	lignoceric acid	1: nervonic acid (Δ^{15})	

[1] can be synthesized from the essential fatty acid, linoleic acid.

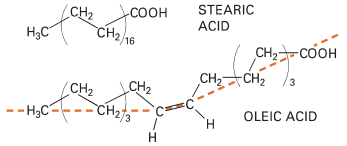


Figure 1.4-1. Structure of Saturated and Unsaturated Fatty Acids (18:0 and 18:1, showing the bend).

While saturated fatty acids tend to assume an extended shape, unsaturated fatty acids show 30° bends at their double bonds (Fig. 1.4-1). This reduces van der Waals interactions between neighboring molecules and lowers the melting point (see organic chemistry textbooks):

18:0 (70°) 18:1 (13°) 18:2 (–9°) 18:3 (–17°)

1.4.2 Acylglycerols and Derivatives (Fig. 1.4-2)

A major proportion of lipids occurring in plants and animals are triesters of glycerol (3.4.2) with higher fatty acids (triacylglycerols = triglycerides = neutral fat). In most of them, the fatty acids are different. Their type and the degree of their unsaturation determine the melting point.

Fats are solid and oils are liquid at room temperature. They are without influence on the osmotic situation in the aqueous phase due to their insolubility and do not bind water as, e.g., glycogen does. Thus, these compounds constitute an effective, convenient storage form of energy (ca.10 kg in adult humans).

Their degree of oxidation is lower than that of carbohydrates or proteins, therefore they provide higher energy during combustion: triolein yields 39.7 kJ/g. This is more than twice the value for anhydrous carbohydrates (17.5 kJ/g) or proteins (18.6 kJ/g) and about six times the energy gained from degradation of these alternative compounds in their physiological state due to their water content.

Triacylglycerols do not contain any hydrophilic groups. If, however, only one or two of the hydroxyl groups of glycerol are esterified (mono- or diacylglycerols), the remaining polar hydroxyl groups allow the formation of ordered structures at water-lipid interfaces and of lipid bilayers (1.4.8). Therefore they can act as emulsifiers, e.g., during lipid resorption from the intestine.

The remaining hydroxyl groups of mono- and diacylglycerols can also carry sugar residues. These so-called glycoglycerolipids are constituents of bacterial cell envelopes (3.10), thylakoid membranes in plants and myelin sheaths of neurons in animals. They are discussed in 3.4.

1.4.3 Waxes (Table 1.4-2, Fig. 1.4-2)

Waxes are esters of higher fatty acids with long-chain primary alcohols (wax alcohols) or sterols (Section 3.5), which are usually solid at room temperature.

They are more resistant than triacylglycerols towards oxidation, heat and hydrolysis (saponification). Frequently, they serve as protective layers, e.g., on leaves and fruits of plants or on skin, feathers and furs of animals (as secretions of specialized glands). Bees' honeycombs are also formed of waxes. In many marine animals they are the main component of lipids (for regulation of flotation and for energy storage). Fossil waxes occur in lignite and bitumen.

Table 1.4-2. Common Components of Waxes

Alcohol (primary, saturated)	Fatty Acid (saturated)
Cetyl alcohol (C_{16})	Lauric acid (C_{12})
Carnaubyl alcohol (C_{24})	Myristic Acid (C_{14})
Ceryl alcohol (C_{26})	Palmitic acid (C_{16})
Myricyl alcohol (C_{30})	Lignoceric Acid (C_{24})
	Cerotic Acid (C_{26})
	Montanic Acid (C_{28})
	Melissic Acid (C_{30})

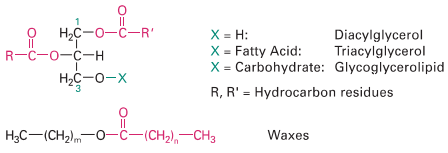
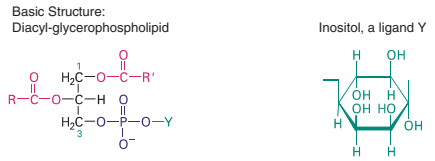


Figure 1.4-2. Structure of Acylglycerols, Glycoglycerolipids and Waxes

1.4.4 Glycerophospholipids (Phosphoglycerides, Fig. 1.4-3)

In contrast to triacylglycerols, in glycerophospholipids only two of the hydroxyl groups of glycerol are esterified with long chain fatty acids, while the group at the 3-position (according to the *sn*-numbering system) forms an ester with phosphoric acid.

All glycerophospholipids have an asymmetric C-atom in the 2-position, they occur in nature in the L-form. Most common are saturated fatty acids (C_{16} or C_{18}) at the 1- and unsaturated ones (C_{16} … C_{20}) at the 2-position. Removal of one fatty acid yields lysoglycerophospholipids.



Class	Y =	Formula of Y (F = Fatty acid)
Phosphatidic acid	H	–H
Phosphatidylethanolamine	ethanolamine	$–CH_2–CH_2–NH_2$
Phosphatidylcholine (lecithin)	choline	$–CH_2–CH_2–N(CH_3)_3^+$
Phosphatidylserine	serine	$–CH_2–CH(NH_2)–COOH$
Phosphatidylinositol	inositol	see above
Diphosphatidylglycerol (cardiolipin)	glycerol-phosphatidic acid	$–CH_2–CHOH–CH_2O–(PO_2)–$ $–O–CH_2–CHOF–CH_2O–F$

Figure 1.4-3. Classes of Glycerophospholipids

If the 3-position of glycerol carries only phosphoric acid, the compound is named phosphatidic acid. However, in most cases the phosphate group is diesterified. This extra residue ('head group', Y in Fig. 1.4-3) determines the class of the compound. These compounds are more polar than mono- or diacylglycerols and form the major part of biological membranes (1.4.8).

1.4.5 Plasmalogens (Fig. 1.4-4)

This group of compounds is related to diacylglycerophospholipids (1.4.4). Also, the head groups (Y) are similar. However, the 1-position of glycerol is not esterified, but carries an α,β-unsaturated alcohol in an ether linkage. They are major components of the CNS, brain (>10 %), heart and skeletal muscles, but little is known about their physiological role.

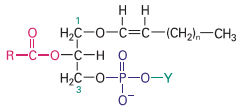
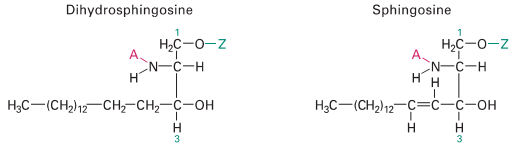


Figure 1.4-4. Structure of Plasmalogens

1.4.6 Sphingolipids (Fig. 1.4-5)

Sphingolipids are important membrane components. They are derivatives of the aminoalcohols dihydrosphingosine (C_{18}), sphingosine (C_{18} with a *trans* double bond) or their C_{16}, C_{17}, C_{19} and C_{20} homologues.

Ceramides are N-acylated sphingosines. If the hydroxyl group at C-1 is esterified with phosphocholine, phosphoethanolamine etc., sphingomyelins (sphingophospholipids) are obtained. If, alternatively, the hydroxyl group is glycosylated, glycosphingolipids (cerebrosides) result. This latter group of compounds is described in 4.4.2-3.



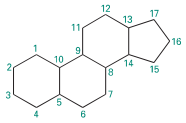
Class	A (= N-Substituent)	Z	Formula of Z
Sphingosine	H	H	–H
Ceramides	higher fatty acid	H	–H
Sphingomyelins	higher fatty acid	Phosphoethanolamine or phosphocholine	$–PO_2–O-CH_2–CH_2–OH$ $–PO_2–O–CH_2–CH_2–N(CH_3)_3^+$
Glycosphingolipids	higher fatty acid	carbohydrate chains	-carbohydrates

Figure 1.4-5. Basic Structure of Sphingolipids

1.4.7 Steroids

Steroids are derivatives of the hydrocarbon cyclopentanoperhydrophenanthrene (Fig. 1.4-6).

Figure 1.4-6. Structure of Cyclopentanoperhydrophenanthrene



Biologically important steroids carry many substituents: generally there is a hydroxy or oxo group at C-3. In addition, several methyl, hydroxy and oxo, in some cases also carboxy, groups are found. In many cases, there is a larger residue bound to C-17. Frequently, some double bonds are present. In a few cases, ring A is aromatic. Substituents below the ring system are designated α and above the ring system β (see Fig. 3.5.1-5).

Steroids are membrane components and participants as well as regulators of metabolism. A detailed description is given in Section 3.5.

1.4.8 Lipoproteins

The major function of lipoproteins is the transport of lipids. They contain non-polar lipids (triacylglycerols, cholesterol esters) in their core, surrounded by a layer of polar compounds (glycerophospholipids, cholesterol, proteins, Fig. 6.2-1). This group of compounds is discussed in context with their transport function in 6.2.

1.5 Physico-Chemical Aspects of Biochemical Processess

Some readers may be less inclined to deal with a fairly large number of mathematical formulas. However, formulas are necessary to describe biochemical processes quantitatively. Considering this, the mathematical part of this book has been concentrated into this section, while usually other chapters refer to it.

Only the most important equations required for discussion of biochemical reactions are presented. In order to facilitate their use, companion equations are given, which show the numerical values of the factors and the dimensions of the terms. For derivation of the equations, refer to physical chemistry textbooks. The units and constants used in the following paragraphs are listed in Table 1.5-1.

1.5.1 Energetics of Chemical Reactions

To each component of a system, an amount of free energy G is assigned, which is composed of the enthalpy H (internal energy + pressure * volume) and of the entropy S (measure of disorder). While the absolute values are not of importance, the change of G (ΔG) is decisive for chemical reactions:

$$\Delta G = \Delta H - T * \Delta S \qquad [1.5\text{-}1]$$

or

$$\Delta G\ [kJ * mol^{-1}] = \Delta H\ [kJ * mol^{-1}] - T\ [K] * \Delta S\ [kJ*mol^{-1} * K^{-1}]. \qquad [1.5\text{-}1a]$$

A reaction proceeds spontaneously only if ΔG is negative.

In biochemistry, ΔG of reactions are usually listed as $\Delta G'_0$, which is obtained at standard conditions of 298 K (25 °C), pH 7.0 and a reactant concentration of 1 mol/l each except for water, where the normal concentration of 55.55 mol/l and gases, where a pressure of 101.3 kPa (= 1 atm) are taken as unity and thus do not appear in the formula.

If the reactant concentrations (henceforth written as [X]) of a reaction A + B + … = Z + Y + … differ from 1 mol/l each, ΔG can be calculated by:

$$\Delta G = \Delta G'_0 + R * T * 2.303 * \log \frac{[Z] * [Y] * \ldots^{\text{(end products)}}}{[A] * [B] * \ldots^{\text{(starting comp.)}}} \qquad [1.5\text{-}2]$$

or

$$\Delta G\ [kJ * mol^{-1}] = \Delta G'_0 + 0.00831 * T * 2.303 * \log \frac{[Z] * [Y] * \ldots}{[A] * [B] * \ldots} \qquad [1.5\text{-}2a]$$

Reaction sequences can be calculated by addition of ΔG's of the individual reactions.

A reaction is at equilibrium if ΔG = 0. Then the equilibrium constant

$$K = \frac{[Z] * [Y] * \ldots^{\text{(end products)}}}{[A] * [B] * \ldots^{\text{(starting comp.)}}} \qquad [1.5\text{-}3]$$

can be calculated as follows:

$$\Delta G'_0 = -R * T * 2.303 * \log K; \qquad K = 10^{(-\Delta G / R * T * 2.303)} \qquad [1.5\text{-}4]$$

or

$$\Delta G'_0\ [kJ * mol^{-1}] = -0.00831 * T * 2.303 * \log K; \qquad K = 10^{(-\Delta G / 0.00831 * T * 2.303)} \qquad [1.5\text{-}4a]$$

Table 1.5-1. Measures and Constants (Selection)

Measure	Unit	Equivalents / Value of Constants	Equivalents in SI Basic Units
Length	meter (m)	1 mm = 10^{-3} m, 1 µm = 10^{-6} m, 1 nm = 10^{-9} m, 1 Å (Ångström) = 10^{-10} m	SI basic unit
Volume	cubic meter (m^3)	1 l (liter) = 10^{-3} m^3; 1 ml = 10^{-3} l, 1 µl = 10^{-6} l	Derived SI unit
Mass	kilogram (kg)	1 g (gram) = 10^{-3} kg; 1 mg = 10^{-3} g, 1 µg = 10^{-6} g	SI basic unit
Time (t)	second (s)	1 ms = 10^{-3} s, 1 µs = 10^{-6} s, 1 ns = 10^{-9} s, 1 ps = 10^{-12} s	SI basic unit
Temperature (T)	Kelvin (K)	0 K = –273.16 °C	SI basic unit
Quantity of matter	Mol	1 mol = 6.0221 * 10^{23} [molecules or ions] 1 mmol = 10^{-3} mol, 1 µmol = 10^{-6} mol, 1 nmol = 10^{-9} mol, 1 pmol = 10^{-12} mol This unit is also applied to photons. 1 Einstein = 1 mol photons.	SI basic unit
Electric current	Ampere (A)		SI basic unit
Force	Newton (N)	1 N = 1 [m * kg * s^{-2}]	1 N = 1 [m * kg * sec^{-2}]
Pressure	Pascal (Pa)	1 Pa = 1 N * m^{-2}, 1 kilopascal (kPa) = 10^3 Pa, 1 atm = 101.325 kPa	1 Pa = 1 [m^{-1} * kg * sec^{-2}]
Energy	Joule (J)	1 J = 1 [N * m], 1 kilojoule (kJ) = 1000 J 1 cal (calorie) = 4.181 J, 1 kcal = 4.181 kJ (Non-SI unit)	1 J = 1 [m^2 * kg * sec^{-2}]
Electric charge	Coulomb (C)	1 C = 6.241 * 10^{18} electron charges	1 C = 1 [A * sec]
Electric potential	Volt (V)	1 V = 1 [J * C^{-1}], 1 mV = 10^{-3} V	1 V = 1 [m^2 * kg * sec^{-3} * A^{-1}]
Constants	**Abbreviation**	**Equivalents / Value of Constants**	**Equivalents in SI Basic Units**
Avogadro's number	N	N = 6.0220 * 10^{23} [mol^{-1}] (see 'quantity of matter', above)	N = 6.0220 * 10^{23} [mol^{-1}]
Boltzmann's constant	k_B	k_B = 1.3807 * 10^{-23} [J * K^{-1}]	k_B = 1.3807 * 10^{-23} [m^2 * kg * sec^{-2} * K^{-1}]
Molar gas constant	R	R = N * k_B = 8.31441 [J * mol^{-1} * K^{-1}]	R = 8.31441 [m^2 * kg * sec^{-2} * mol^{-1} * K^{-1}]
Faraday's constant	F	F = 1 N electron charges = 96 484.5 [C * mol^{-1}] = 96 484.5 [J * V^{-1} * mol^{-1}]	F = 96 484.5 [A * sec * mol^{-1}]
Planck's constant	h	h = 6.6262 * 10^{-34} [J * sec]	h = 6.6262 * 10^{-34} [m^2 * kg * sec^{-1}]

In calculations in this book, usually l, g, kJ and mV are used. Since the constants have then to be expressed in these units, their numerical value changes by the factor 10^{-3} or 10^{-6} in the respective formulas.

Enzymes cannot shift the equilibrium, they only increase the reaction velocity. The kinetics of enzyme catalyzed reactions are discussed in 1.5.4.

1.5.2 Redox Reactions

Redox reactions are reactions where one compound is reduced (electron acceptor A) while its reaction partner is oxidized (electron donor B) by transfer of n electrons:

$$A_{ox}^{n+} + B_{red} = A_{red} + B_{ox}^{n+}.$$

The change of free energy during such a reaction is described by a formula, which is analogous to Eq. [1.5-2]:

$$\Delta G = \Delta G_0' + R * T * 2.303 * \log \frac{[A_{red}] * [B_{ox}^{n+}]\ \text{(end products)}}{[A_{ox}^{n+}] * [B_{red}]\ \text{(starting comp.)}} \qquad [1.5\text{-}5]$$

or

$$\Delta G\ [\text{kJ} * \text{mol}^{-1}] = \Delta G_0' + 0.00831 * T * 2.303 * \log \frac{[A_{red}] * [B_{ox}^{n+}]}{[A_{ox}^{n+}] * [B_{red}]} \qquad [1.5\text{-}5a]$$

w expresses the work gained by transferring n mol charges (= n Faraday, F) across a potential difference of $\Delta E = E_{end} - E_{begin}$

$$w = -n * F * \Delta E. \qquad [1.5\text{-}6]$$

Since a positive amount of work diminishes the free energy of the system

$$w = -n * F * \Delta E = -\Delta G \qquad [1.5\text{-}6a]$$

or

$$\Delta G\ [\text{kJ} * \text{mol}^{-1}] = n * 0.0965 * \Delta E\ [\text{mV}], \qquad [1.5\text{-}6b]$$

equation [1.5-5] can also be written as:

$$\Delta E = \Delta E_0' + \frac{R * T}{n * F} * 2.303 * \log \frac{[A_{red}] * [B_{ox}^{n+}]\ \text{(end products)}}{[A_{ox}^{n+}] * [B_{red}]\ \text{(starting comp.)}} \qquad [1.5\text{-}7]$$

or

$$\Delta E\ [\text{mV}] = \Delta E_0' + \frac{0.00831 * T}{n * 0.0965} * 2.303 * \log \frac{[A_{red}] * [B_{ox}^{n+}]}{[A_{ox}^{n+}] * [B_{red}]} \qquad [1.5\text{-}7a]$$

$\Delta E_0'$ is the difference of the redox potentials of this reaction (or the electromotive force across membranes, 1.5.3) under biochemical standard conditions (298 K = 25°C, pH 7.0 and a reactant concentration of 1 mol/l each). Only water, which is present in a concentration of 55.55 mol/l and gases, with a pressure of 1 atm are taken as unity.

Redox potentials: The reaction can be divided into two half reactions (e^- = electrons):

$$A_{red} = A_{ox}^{n+} + n\ e^- \quad \text{and} \quad B_{red} = B_{ox}^{n+} + n\ e^-$$

The zero value of the redox potential is by convention assigned to the potential of the half reaction $2\ H^+ + 2\ e^- = H_2$ at a platinum electrode at pH = 0, 298 K (25°C) and a hydrogen pressure of 101.3 kPa (= 1 atm). Thus, under the standard conditions used in biochemistry (pH = 7.0), $E_0'\ (2H^+/H_2) = -410$ mV.

Correspondingly, the half reactions can be expressed as:

$$E_A = (E_0')_A + \frac{R * T}{n * F} * 2.303 * \log \frac{[A_{ox}^{n+}]}{[A_{red}]} \qquad [1.5\text{-}8]$$

or

$$E_A\ [\text{mV}] = (E_0')_A + \frac{0.00831 * T}{n * 0.0965} * 2.303 * \log \frac{[A_{ox}^{n+}]}{[A_{red}]} \qquad [1.5\text{-}8a]$$

and analogously for B.

Various redox potentials can be combined this way: $\Delta E = E_B - E_A$ (A being the electron acceptor and B being the electron donor). The reactions proceed spontaneously only if ΔE is negative, i.e., when the potential changes to a more negative value.

Redox potentials are usually plotted with the minus values on top. A spontaneous reaction proceeds in such a plot from top to bottom (e.g., Fig. 3.12-6).

In the literature, the definition of ΔE is not uniform. In a number of textbooks it is defined in opposite order to the above: $\Delta E = E_{begin} - E_{end}$. Therefore, ΔE and ΔE_0 have to be replaced by $-\Delta E$ and $-\Delta E_0'$, respectively. This affects Eqs. [1.5-6] … [1.5-8a] and has to be considered when making comparisons.

1.5.3 Transport Through Membranes

Uncharged molecules: If an uncharged compound A is present on both sides of a permeable membrane in different concentrations, its passage through the membrane is accompanied by a change of free energy. In biochemistry, this situation occurs mostly at cellular membranes (or membranes of organelles). For import into cells, the following equation applies:

$$\Delta G = R * T * 2.303 * \log \frac{[A_{inside}]}{[A_{outside}]} \qquad [1.5\text{-}9]$$

or

$$\Delta G\ [\text{kJ} * \text{mol}^{-1}] = 0.00831 * T * 2.303 * \log \frac{[A_{inside}]}{[A_{outside}]} \qquad [1.5\text{-}9a]$$

Thus, the transport occurs spontaneously only at negative ΔG, (when $[A_{inside}] < [A_{outside}]$), i.e, from higher to lower concentrations.

Correspondingly, for export from cells, the quotient is reversed

$$\Delta G = R * T * 2.303 * \log \frac{[A_{outside}]}{[A_{inside}]} \qquad [1.5\text{-}9b]$$

Charged molecules: The situation is more complicated if there is a potential difference $\Delta\Psi$ across the membrane (e.g., by non-penetrable ions)

$$\Delta\Psi = \Psi_{inside} - \Psi_{outside} \qquad [1.5\text{-}10]$$

and the compounds passing through the membrane carry Z positive charges/molecule (or −Z negative charges/molecule). The contribution of the charges to ΔG (with the prefix of Z corresponding to the + or − charge of the ions) is expressed by:

$$\Delta G_{charge\ transport} = Z * F * \Delta\Psi \qquad [1.5\text{-}11]$$

or

$$\Delta G_{charge\ transport}\ [\text{kJ} * \text{mol}^{-1}] = Z * 0.0965 * \Delta\Psi\ [\text{mV}]. \qquad [[1.5\text{-}11a]$$

Thus, for an import process, Eq. [1.5-9] and Eq. [1.5-11] have to be combined:

$$\Delta G = R * T * 2.303 * \log \frac{[A_{inside}]}{[A_{outside}]} + Z * F * \Delta\Psi \qquad [1.5\text{-}12]$$

or

$$\Delta G\ [\text{kJ} * \text{mol}^{-1}] = 0.00831 * T * 2.303 * \log \frac{[A_{ins.}]}{[A_{outs.}]} + Z * 0.0965 * \Delta\Psi\ [\text{mV}]. \qquad [1.5\text{-}12a]$$

Correspondingly, for an export process,

$$\Delta G = R * T * 2.303 * \log \frac{[A_{outside}]}{[A_{inside}]} - Z * F * \Delta\Psi \qquad [1.5\text{-}12b]$$

The prefix of the last term in this equation is the opposite one of Eq. [1.5-12], since the membrane potential (Eq. 1.5-10) has the opposite effect on the energy situation.

An equilibrium exists if $\Delta G = 0$. Then the equilibrium potential $\Delta\Psi_0$ [mV] can be obtained by the Nernst equation:

$$\Delta\Psi_0 = -\frac{R * T}{Z * F} * 2.303 * \log \frac{[A_{inside}]}{[A_{outside}]} \qquad [1.5\text{-}13]$$

or

$$\Delta\Psi_0\ [\text{mV}] = -\frac{0.00831 * T}{Z * 0.0965} * 2.303 * \log \frac{[A_{inside}]}{[A_{outside}]} \qquad [1.5\text{-}13a]$$

An extension of this formula to the equilibrium potential of several ions is the Goldman equation (see 7.2.1).

Literature:

Physical chemistry textbooks.

1.5.4 Enzyme Kinetics

The biochemical base of enzyme catalysis is discussed in 2.4. In the following, the mathematical treatment of the kinetics is given in some more detail.

Velocity of reactions: The reaction rate v for conversion of a single compound A → product(s) (first order reaction) is proportional to the concentration of this compound [A], while for a two-compound reaction A + B → product(s) (second order reaction) it depends on the number of contacts and thus on the concentration of both components (Eq. [1.5-14] and Eq. [1.5-15]). The proportionality factor k is termed rate constant.

Eq. [1.5-15] can also be applied for the formation of a complex and Eq. [1.5-14] for the decomposition of this complex. This includes substrate-enzyme complexes (see below), ligand-receptor complexes (7.1-2), antigen-antibody complexes (8.1.4) etc.

$$v = -\frac{d[A]}{dt} = k * [A] \qquad [1.5\text{-}14]$$

$$v = -\frac{d[A]}{dt} = -\frac{d[B]}{dt} = k * [A] * [B] \qquad [1.5\text{-}15]$$

Enzyme catalyzed one-substrate reaction: The theory of the enzyme-catalyzed conversion of a single reactant (the substrate, S) is based on the assumption that the enzyme (the catalyst, E) and this substrate form a complex (ES) by a reversible reaction. This step is kinetically treated like a two-compound reaction (rate constants k_1 and k_{-1} for formation and decomposition, respectively). The complex is then converted into the product (P) with the rate constant k_2. The conversion into P is considered to be irreversible at the beginning of the reaction, when practically no product is present.

$$E + S \underset{k_{-1}}{\overset{k_1}{\leftrightarrow}} ES \xrightarrow{k_2} E + P \qquad [1.5\text{-}16]$$

Therefore, for the formation of the enzyme-substrate complex, Eq. [1.5-15] has to be applied, while for its decomposition into its components, as well as for its conversion to the products, Eq. [1.5-14] is valid. There is actually an intermediate step ES → EP before the product is released. Its rate constant is not treated as a separate entity in most discussions of kinetic behavior, but is combined with the dissociation step to k_2. This is also done in the following considerations.

Usually, the substrate is in large excess over the enzyme. In this case, after a short 'transient phase', [ES] can be considered to be sufficiently constant (steady-state assumption). Disregarding the reverse reaction by using the situation immediately after the transient phase (see above) one obtains

$$\frac{d[ES]}{dt} = 0 = k_1 * [E] * [S] - k_{-1} * [ES] - k_2 * [ES] \qquad [1.5\text{-}17]$$

If one assumes that the rate determining process is the reaction ES → E + P, the initial reaction rate v_0 can be written as a function of [ES], which is analogous to Eq. [1.5-14]

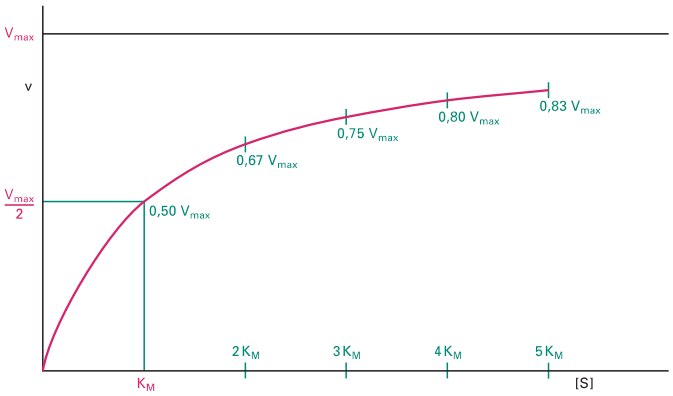


Figure 1.5-1. Reaction Velocity of an Enzyme Catalyzed Reaction
The velocity at [S] = n * K_M is shown.

$$v_0 = k_2 * [ES]. \qquad [1.5\text{-}14a]$$

By using a term for the total concentration of enzyme $[E_t] = [E] + [ES]$, by expressing the maximum reaction rate V_{max}, which is obtained when all of the enzyme is saturated with substrate ([ES] = $[E_t]$) as

$$V_{max} = k_2 * [E_t], \qquad [1.5\text{-}14b]$$

and by introducing the Michaelis constant K_M

$$K_M = \frac{k_{-1} + k_2}{k_1}, \qquad [1.5\text{-}18]$$

one obtains the so-called Michaelis-Menten equation

$$v_0 = \frac{V_{max} * [S]}{K_M + [S]}, \qquad [1.5\text{-}19]$$

which shows the dependency of the reaction rate on the substrate concentration (first-order reaction). The plot of reaction rate vs. substrate concentration is a rectangular hyperbola (Fig. 1.5-1).

These formulas describe only the forward reaction. If the reverse reaction is included, the equivalent to Eq. [1.5-19] is

$$v = \frac{\frac{(V_{max})_f * [S]}{(K_M)_f} - \frac{(V_{max})_r * [P]}{(K_M)_r}}{1 + \frac{[S]}{(K_M)_f} + \frac{[P]}{(K_M)_r}} \qquad [1.5\text{-}20]$$

where $(V_{max})_f$ and $(K_M)_f$ are identical to V_{max} and K_M in Eq. [1.5-19], while the terms $(V_{max})_r$ and $(K_M)_r = (k_{-1} + k_2)/k_{-2}$ are formed analogously for the reverse reaction.

Michaelis constant: As can be derived from Eq. [1.5-19], the Michaelis constant K_M equals the substrate concentration at half the maximal reaction rate. Most of them are in the range of 10^{-5}…10^{-1} mol/l (Fig. 10.3-3).

Instead of obtaining this value from a plot according to Figure 1.5-l, it is more convenient to use the reciprocal of the Michaelis-Menten equation, which yields a linear plot (at least in the ideal case, Lineweaver-Burk plot, Fig. 1.5-2a):

$$\frac{1}{v_0} = \frac{K_M}{V_{max} * [S]} + \frac{1}{V_{max}} \qquad [1.5\text{-}21]$$

If $1/v_0$ is plotted vs. 1/[S], then the intersections of this line with abscissa and ordinate allow the determination of K_M and V_{max}.

A disadvantage of the Lineweaver-Burk plot is the accumulation of measuring points near the ordinate (see the markings on the abscissa of Fig. 1.5-2a). Therefore other ways of plotting have been proposed. Hanes used another transformation of the Michaelis-Menten equation:

$$\frac{[S]}{v_0} = \frac{K_M}{V_{max}} + \frac{[S]}{V_{max}} \qquad [1.5\text{-}21a]$$

The plot of $[S]/v_0$ vs. [S] yields a line with the abscissa intersection $-K_M$ and the ordinate intersection K_M/V_{max}. The slope equals $1/V_{max}$ (Fig. 1.5-2 b).

Still another method, the so-called 'direct plot', has been proposed by Eisenthal and Cornish-Bowden. The Michaelis-Menten equation is rearranged as follows:

$$V_{max} = v_0 + \frac{v_0}{[S]} * K_M \qquad [1.5\text{-}21b]$$

For each individual measurement, −[S] is marked on the abscissa and v_0 on the ordinate and a line is drawn through both points. The intersection of these lines has the abscissa value K_M and the ordinate value V_{max} (Fig. 1.5-2 c).

However, the most accurate method is the statistical evaluation of the measurements. In spite of this, the Lineweaver-Burk plot will be used in the following graphical representations, since it is the best known one.

Characterization of enzyme activities: The enzyme activity is defined as the quantity of substrate turned over per time unit in the presence of a given amount of enzyme. Thus the standard dimension would be [mol * sec^{-1}] = katal. For practical reasons, usually the

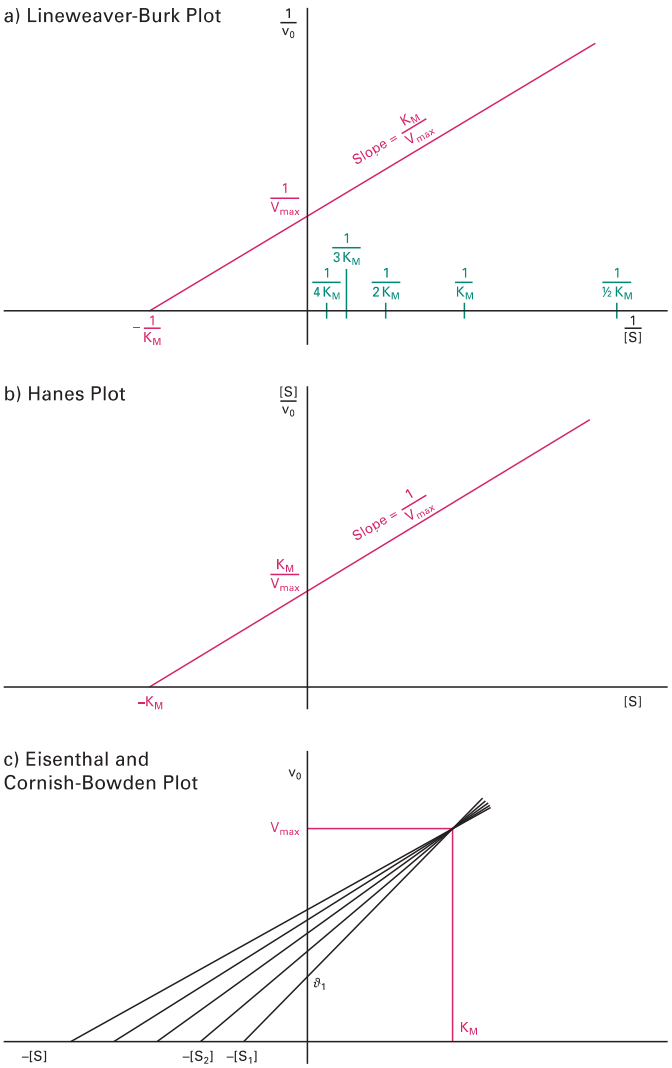


Figure 1.5-2. Linear Plots of an Enzyme Catalyzed Reaction

activity is expressed as [µmol * min^{-1}]. This term is named International Unit (U) if the measurement is performed under standard conditions (with isolated enzymes at conditions that are optimized as much as possible). The specific activity is the enzyme activity per unit of weight, e.g., per mg and is frequently used to characterize the degree of purification of isolated enzymes.

The turnover number of an enzyme is defined as the number of molecules converted by one molecule of enzyme per unit of time if the enzyme is saturated with substrate ($[E_t] = [ES]$). It is identical to the rate constant k_2 and can be calculated from Eq. [1.5-14 b] as $k_2 = V_{max}/[E_t]$. Most turnover numbers are in the range of $1 \dots 10^4$ (see Fig. 10.3-4), the value for catalase is $4 * 10^7$.

Most reactions *in vivo* proceed below the saturation limit of the enzyme, frequently at $[S] = 0.01 \dots 1\ K_M$. By the combination of Eq. [1.5-17], Eq. [1.5-18] and Eq. [1.5-14 a] one obtains

$$v_0 = \frac{k_2}{K_M} * [E] * [S] \qquad [1.5\text{-}22].$$

At low substrate concentration, only a small portion of the enzyme forms an enzyme-substrate complex and $[E] \approx [E_t]$ = constant. The term k_2/K_M indicates how often a contact of enzyme and substrate leads to a reaction and is therefore a measure of the catalytic efficiency. It has an upper limit of ca. 10^9 [l * mol^{-1} * sec^{-1}], when practically every contact leads to a reaction, and the reaction rate is determined by the diffusion speed. The value for catalase ($4 * 10^8$) is one of the highest observed.

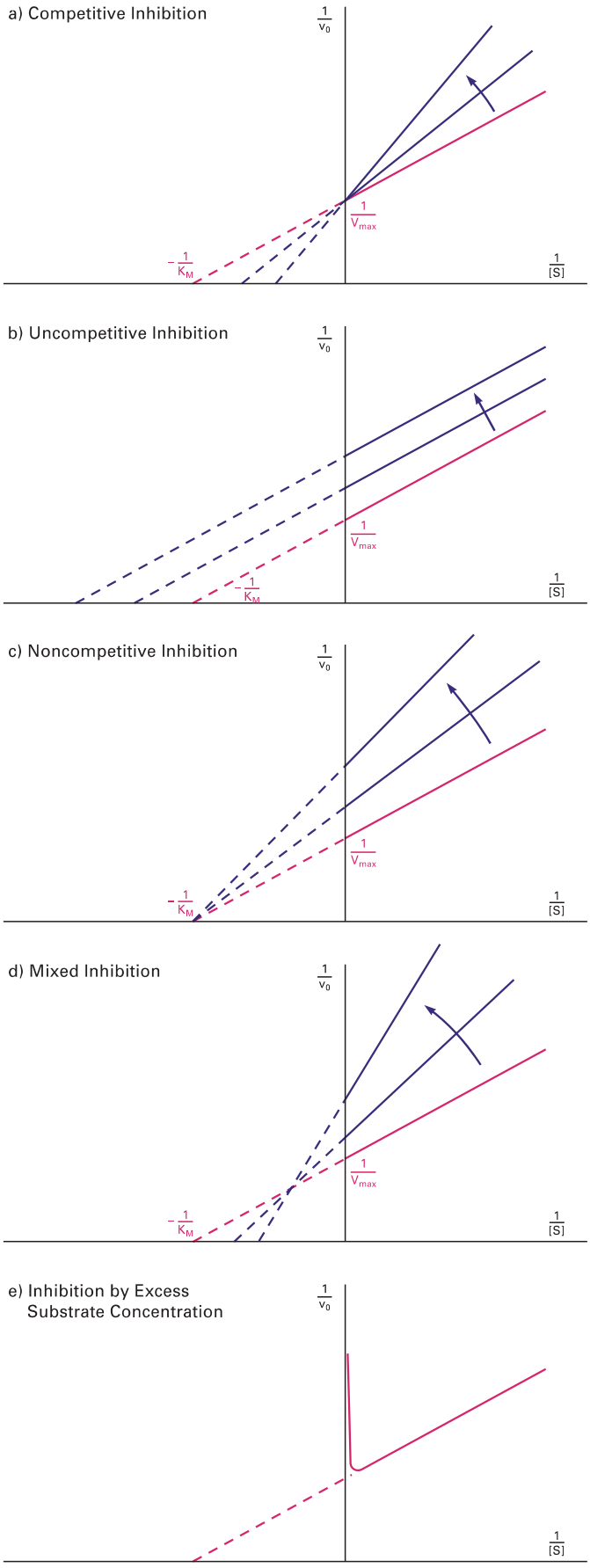


Figure 1.5-3. Lineweaver-Burk Plots of Inhibited Reactions
Red = uninhibited reaction, blue = inhibited reaction, arrow = shift of the plot at increasing inhibitor concentrations.

Inhibition: The mathematical treatment of an inhibited reaction depends on the mechanism of the inhibition. The general principles of inhibition are described in 2.5.2.

Competitive inhibition: The inhibitor competes with the substrate for reversible binding to the active site of the enzyme. The enzyme-substrate and the enzyme-inhibitor complexes are formed with the dissociation constants K_S and K_I, respectively.

$$K_s = \frac{[E] * [S]}{[ES]} \qquad [1.5\text{-}23]$$

$$K_I = \frac{[E] * [I]}{[EI]} \qquad [1.5\text{-}23a]$$

This results in the equation

$$\frac{1}{v_0} = \frac{K_M}{V_{max} * [S]} * \left(1 + \frac{I}{K_I}\right) + \frac{1}{V_{max}} \qquad [1.5\text{-}24]$$

In the Lineweaver-Burk plot, lines obtained at different inhibitor concentration intersect at the ordinate (Fig. 1.5-3a).

Uncompetitive inhibition: The inhibitor reacts reversibly only with the enzyme substrate-complex, but does not affect its formation. The dissociation constant is K'_I.

$$K'_I = \frac{[ES] * [I]}{[ESI]} \qquad [1.5\text{-}25]$$

This yields the equation

$$\frac{1}{v_0} = \frac{K_M}{V_{max} * [S]} + \frac{1}{V_{max}} \left(1 + \frac{[I]}{K'_I}\right) \qquad [1.5\text{-}26]$$

In the Lineweaver-Burk plot, parallel lines are obtained at different inhibitor concentrations (Fig. 1.5-3b).

Noncompetitive and mixed inhibition: If the inhibitor binds both to the enzyme and to the enzyme-substrate complex according to Eqs. [1.5-23a] and [1.5-25] and prevents formation of the product, the following equation results

$$\frac{1}{v_0} = \frac{K_M}{V_{max} * [S]} * \left(1 + \frac{[I]}{K_I}\right) + \frac{1}{V_{max}} * \left(1 + \frac{[I]}{K'_I}\right) \qquad [1.5\text{-}27]$$

If the affinities of the inhibitor to the enzyme and to the enzyme-substrate complex are equal ($K_I = K'_I$), then the lines obtained at different inhibitor concentrations intersect in the Lineweaver-Burk plot at the negative abscissa (K_M remains unchanged, noncompetitive inhibition, Fig. 1.5-3c). Otherwise, they intersect in the second quadrant (left of the ordinate, mixed inhibition, Fig. 1.5-3d).

Inhibition by excessive substrate concentrations. If the reaction velocity decreases at very high substrate concentrations, this results in a Lineweaver-Burk curve bent upwards near the ordinate. This situation is mostly observed in *in vitro* experiments (Fig. 1.5-3e).

Two-substrate reactions: The formulas describing the kinetics are derived from the same assumptions as for one-substrate reactions. Their form depends on the reaction sequence. They involve separate Michaelis constants for the turnover of each substrate.

The Cleland nomenclature system uses the following expressions for the number of the substrates and products of the reaction; 1 – Uni, 2 – Bi, 3 – Ter, 4 – Quad. The substrates are named A, B, C…, the products P, Q, R… and the enzyme species (original state, intermediates and final state) E, F, G… If all components have to combine before the reaction takes place, this is called a sequential reaction. This may take place in an ordered way or at random. If, however, one component leaves the enzyme before the other enters, it is a ping-pong reaction. The mechanisms are schematically drawn in Figure 1.5-4.

The formula for an ordered sequential Bi-Bi reaction is

$$\frac{1}{v_0} = \frac{1}{V_{max}} + \frac{(K_M)_A}{V_{max} * [A]} + \frac{(K_M)_B}{V_{max} * [B]} + \frac{(K_M)_{AB}}{V_{max} * [A] * [B]} \qquad [1.5\text{-}28]$$

The general formula for random sequential reactions is very complicated.

A ping-pong Bi-Bi reaction is described by

$$\frac{1}{v_0} = \frac{1}{V_{max}} + \frac{(K_M)_A}{V_{max} * [A]} * \frac{(K_M)_B}{V_{max} * [B]} \qquad [1.5\text{-}29]$$

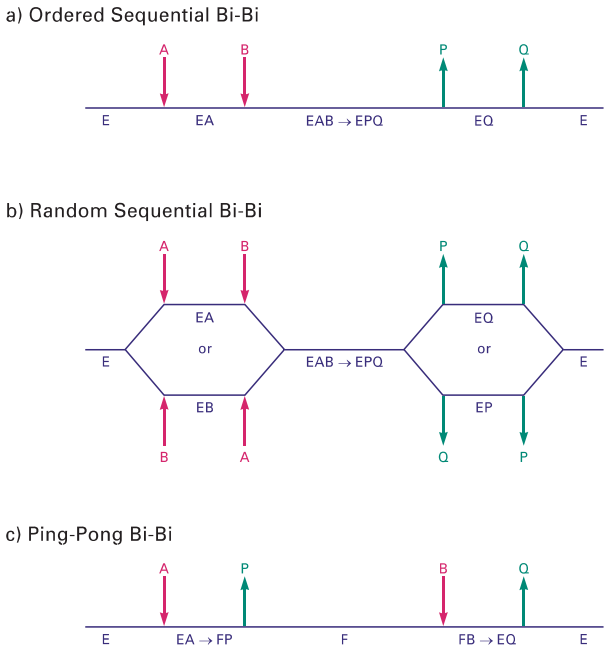


Figure 1.5-4. Types of Two-Substrate-Two-Product (Bi-Bi) Reactions
The enzyme is represented by the horizontal line.

If in second order reactions the concentration of one of the substrates is very much above the respective Michaelis constant, then the terms containing this concentration in Eqs. [1.5-28] and [1.5-29] are practically zero and the equations become identical with Eq. [15.1-21], allowing the same evaluation as with a first order reaction.

If a series of measurements are made in which one substrate is varied while the other is kept constant, then one obtains Lineweaver-Burk plots that formally resemble those obtained with inhibited reactions. However, increasing concentrations of the second substrate shift the lines in the other direction (Fig. 1.5-5). Ordered sequential

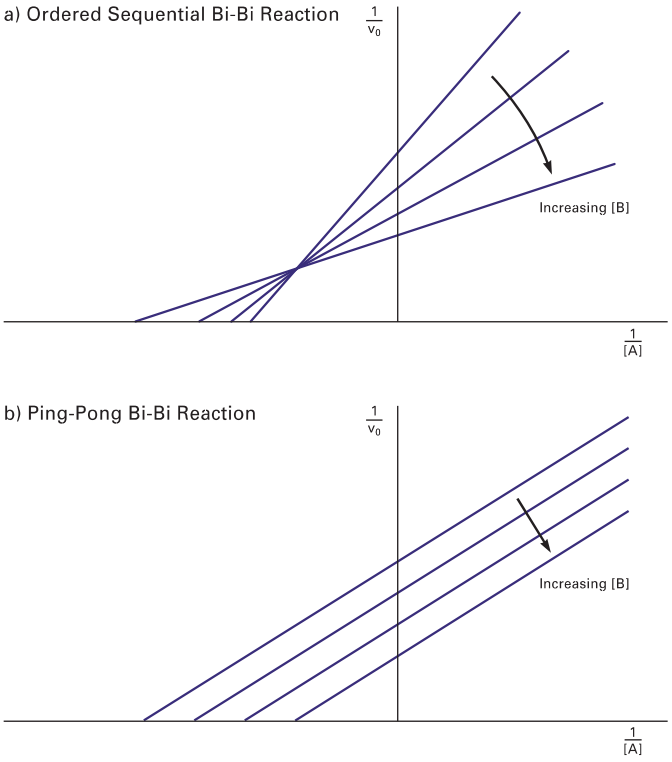


Figure 1.5-5. Lineweaver-Burk Plots of Two-Substrate Reactions
Arrow = shift of the plot when the concentration of the other substrate is raised.

mechanisms yield a series of lines, which intersect left of the ordinate (above or below the abscissa), while ping-pong mechanisms yield parallel lines.

Dependence of reactions on temperature and activation energy: A more refined consideration of the reaction sequence Eq. [1.5-16] shows that only collisions of the reactants above a certain energy level will lead to the formation of complexes, e.g., ES and EP. Also, the reaction ES → EP requires an initial energy input. Thus, the reaction has to cross 'energy hills', which represent metastable states (Fig. 2.4-1). They are called transition complexes $X^{\neq}$ and can either return to the original components or progress towards the products of the reaction, quickly achieving equilibrium in both cases. Among the 'energy hills' mentioned above, the highest one represents obviously the rate determining step of the reaction and has to be the one considered further. (It takes the place of [ES] in the previous equations.) Thus, the equilibrium for formation of this complex can be described analogously to Eq. [1.5-3] by

$$K = \frac{[X^*]}{[A] * [B]} \qquad [1.5\text{-}30]$$

The energy required for its formation is called activation energy $\Delta G^{\neq}$, which can be calculated from this equilibrium by applying Eq. [1.5-4] as

$$\Delta G^{\neq} = -R * T * 2.303 * \log \frac{[X^*]}{[A] * [B]} \qquad [1.5\text{-}31]$$

According to Eq. [1.5-14] the reaction rate for formation of the product(s) from this complex is expressed by $v_0 = k * [X^{\neq}]$. When combining this equation with Eq. [1.5-31], one obtains the following formula for the temperature and $\Delta G^{\neq}$ dependence:

$$v_0 = \text{const.} * [A] * [B] * 10^{-\Delta G^{\neq} / 2{,}303 * R * T} \qquad [1.5\text{-}32]$$

The increase of the reaction rate with rising temperature is limited, however. When the enzyme becomes thermally denatured, the rate drops (Fig. 2.4-4).

Fractal enzyme kinetics: The above considerations assume 'ideal' conditions; purified enzymes, low concentrations, free movement of the reactants. However, *in vivo* the situation is different. Based on a power-law derivation it has been shown that, e.g., restrictions in space require the introduction of non-integer powers > 1 to the concentration terms in Eq. [1.5-17]:

$$\frac{d\,[ES]}{dt} = \alpha_1 * [E]^g * [S]^h - \alpha_{-1} * [ES] - \alpha_2 * [ES] \qquad [1.5\text{-}33]$$

The consecutive equations change analogously. This system is called fractal kinetics. Its main implications are:

- K_M is dependent on the enzyme concentration; it decreases with increasing enzyme concentration.
- The plot of enzyme activity vs. substrate concentration has a tendency towards a sigmoid shape even with monomeric enzymes.
- The velocity of the reaction increases if the movements are, e.g., restricted to surface interfaces (e.g., 3.4.3.2) or to one dimension (e.g., by sliding along nucleic acid strands, 3.8.1.2, 4.2.3.2 or by 'substrate channeling', 3.2.7.1).
- In sequences of reactions, the flux responses are faster and the accumulation of intermediates is lower as compared to the Michealis-Menten assumption.

In some respects, fractal kinetics resemble allosteric situations (2.5.2). Velocity calculations according to this theory have a tendency to yield higher values as according to the Michaelis-Menten theory, which represents a borderline case of a more general treatment, but is still of value for understanding the basic principles of enzyme catalysis.

Literature:

Cornish-Bowden, A., Wharton, C.W. *Enzyme Kinetics*. Oxford: IRL Press; 1988.

Dixon, M., Webb, E.C. *Enzymes.* 3rd Ed. Waltham (MA): Academic Press; 1979.

Fersht, A. *Structure and Mechanism in Protein Science*. New York: W.H. Freeman & Company; 1999.

Freeman (1985). Savageau, M.A. J. Theor. Biol. 1995;176:115–124.

Sigman, D.S., Boyer, P.D. (Eds.). *The Enzymes.* 3rd Ed. Vols. 19 and 20. Waltham (MA): Academic Press; 1990 and 1992.

2 The Cell and Its Contents

Gerhard Michal and Dietmar Schomburg

This chapter presents selected information on the structure and organization of living organisms and their major components to serve as a background for the biochemical text of this book. For more details, refer to biology textbooks.

2.1 Classification of Living Organisms

Life is associated with a number of characteristics such as propagation, metabolism, response to environmental influences, and evolution. Cells are the basic unit of organization for all living beings. Whereas unicellular organisms exist as separate entities, the various cells of multicellular organisms fulfill different functions, and the organism depends on mutual cellular interaction.

There are several systems of classification of living organisms. From a phylogenetic viewpoint, the classification into the three domains; bacteria, archaea and eukarya (which are further subdivided) appears most justified (Table 2.1-1). When common aspects of eubacteria and archaea are discussed, the term prokarya is used.

The metabolic reactions in this book are indicated by colored arrows. Since frequently the occurrence of the reactions is known only for a few species and also in order to prevent an 'overloading' of the figures with too much detail, the arrow colors have been combined into (black) general metabolism, (red) bacteria and archaea, (green) plants, fungi and protists, (blue) animals.

Living organisms exhibit a high degree of order. The sum of all endogenous life processes results in a steady decrease of free energy (1.5.1). Therefore, life can only be kept up by an energy input from the environment, either as light energy or by uptake of oxidizable compounds. Another essential requirement of life is the availability of an adequate carbon source. Living beings can be classified according to the mode of energy uptake and the carbon source used (Table 2.1-2).

During the oxidation of compounds, electrons are released, which have to be taken up by a terminal electron acceptor. Energy wise, oxygen is most favorable (3.11): previous to its appearance in the primeval atmosphere, living organisms had to use other acceptors. This is still the case in oxygen-free habitats (Table 2.1-3).

Literature:

Fox, GE. *et al.*: Science 1980;209:457–463.

Holt, JG. *et al.*: *Bergey's Manual of Descriptive Bacteriology* (9th Ed.). Williams and Wilkins (1994).

Margulis, L., Schwartz, K.V.: *Five Kingdoms*. 2nd. Ed. Freeman (1987).

Woese, CR. *et al.*: Structure of Cells. Proc. Natl. Acad. Sci. USA 1990;87:4576–4579.

2.2 Structure of Cells

2.2.1 Prokaryotic Cells (Fig. 2.2-1)

The genetic information is stored in a single, circular double helix of deoxyribonucleic acid (DNA, 2.6.4). It is located in the central portion of the cell in a densely packed form (nucleoid), but without

Table 2.1-1. Some Typical Properties of Living Organisms (Exceptions exist)

Domains	Bacteria	Archaea	Eukarya			
Kingdoms	Bacteria	Archaea	Protists[1]	Plants	Fungi	Animals
Nucleus	no (common term **prokarya**)		Yes			
Genome	circular, ca.10^6 … 5 * 10^7 kb, extra plasmids		linear, 10^7 … > 10^{11} kb, organized in several chromosomes			
RNA polymerase	one type	several types				
Starting amino acid for translation	formylmethionine	methionine				
Reproduction	binary scission		asexual/sexual			
Cellular organization	unicellular (some are aggregated)		mostly unicellular	multicellular	uni-/multicellular	multicellular
Nutrition (Table 2.1-2)	chemoorganotrophic, photoautotrophic or photoheterotrophic	chemolithotrophic, photoautotrophic or chemoorganotrophic.	chemoorganotrophic or photoautotrophic	photoautotrophic	chemoheterotrophic including saprobiontic	chemoheterotrophic
Size of cells	average 1 … 5 µm, wide variation		average 10 … 100 µm, wide variation			
Cell membranes	rigid, contain peptidoglycans	rigid, without peptidoglycans	rigid or soft	rigid, contain cellulose and lignin	rigid, contain chitin	soft, lipid bilayer only
Internal membranes	no		yes, they enclose organelles /vesicles			

[1] Algae, protozoa, fungi-related. The exact demarcation is under discussion.

Table 2.1-2. Sources of Carbon and of Energy

	Phototrophy (Energy input by light)	Chemotrophy (Energy provided by oxidizable compounds from the environment)
Autotrophy (only CO_2 needed as carbon source)	green plants, some protists, photosynthesizing bacteria (3.12)	Prokarya (mainly archaea). Oxidation of inorganic material (chemolithotrophy, 3.10.7)
Heterotrophy (organic compounds needed as carbon source)	some prokarya	All animals and fungi, non-green plants, many protists and prokarya. Oxidation of organic material (chemoorganothrophy). Included are saprobionts (use decaying organic material) and parasites (feed from living beings)

Table 2.1-3. Terminal Electron Acceptors for Oxidation Reactions

	Atmospheric Oxygen		
	Not required		Required
Energy obtained by	anaerobic respiration (3.10.6)	fermentation (3.10.5)	aerobic respiration (3.11)
Electron acceptors	oxidized external compounds (mostly inorganic)	internally generated compounds	atmospheric O_2
Organisms	anaerobes: part of archaea and bacteria		all other organisms
	facultative anaerobes (bacteria)		

Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Second Edition. Edited by Gerhard Michal and Dietmar Schomburg.

a special separation from the rest of the cell. Its replication and the translation of the information into protein structures are described in 3.8 and 4.1.

In prokarya, frequently a number of plasmids may occur, which also consist of circular DNA and replicate independently of the main DNA. They carry only a few genes. Although plasmids are not usually essential for survival, they are involved in DNA transfer during conjugation, provide resistance to antibiotics, etc. Some plasmids can be reversibly integrated into the main DNA (episomes). Similar properties are exhibited by DNA viruses and retroviruses (5.2, 5.4). The translocation of genetic material is not discussed here.

The cytoplasm is a semifluid, concentrated solution of proteins, metabolites, nucleotides, salts, etc. It also contains several thousand ribosome particles involved in translation (4.1.2). It is the site of most metabolic reactions and exchanges material in a controlled way with the environment (3.10.1 … 4).

Prokaryotic cells are surrounded by an envelope (3.10.1, Fig. 2.2-2). It not only has an enclosing and protective function. Rather, a number of metabolic reactions take place at transmembrane proteins (e.g., respiration and ATP synthesis) or at membrane associated proteins. In bacteria, the sequence of membrane components from the interior outwards is:

- The plasma membrane, a lipid bilayer with embedded proteins
- The rigid cell wall, which in the case of bacteria consists of either multiple layers (Gram positive bacteria) or a single layer (Gram negative bacteria) of peptidoglycans (murein).
- An additional outer membrane (only in Gram negative bacteria).
- An additional gelatinous capsule is frequently superimposed on the cell wall. It consists mainly of polysaccharides (polymerized glucose, rhamnose, uronic acids etc.). There may also be mucus layers.

Extensions of the cell envelope are pili and flagella, which provide for cellular contact, conjugation, propulsion, etc.

The composition of an *E. coli* cell by weight is H_2O about 70 %, protein 15 %, DNA 1 %, RNA 6 %, polysaccharides 3 %, lipids 2 % (both are mainly present in the envelope), small organic molecules 1 %, inorganic molecules 1 %.

Mycoplasms are a group of bacteria which lack a cell wall. Among them are the smallest self-reproducing organisms (0.10 … 0.25 µm diameter).

Archaea differ from bacteria by

- A different composition and arrangement of rRNAs
- Differences in the RNA polymerase and in the translation mechanism (Table 2.1-1)
- Different composition of the cellular envelope. e.g., murein (3.10.1) is absent, acylglycerols are replaced by branched chain glycerol ethers (3.4.3.3)
- Unusual pathways of metabolism and habitats (methanogens, 3.10.6.2, halobacteria, 3.12.1, thermophiles, etc.)

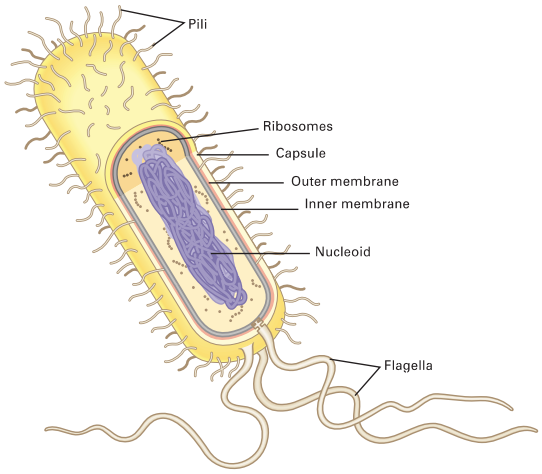


Figure 2.2-1: General Structure of a Bacterial Cell
The colors are for easy differentiation only. After Campbell, N.A.: Biology 4th Ed. Benjamin/Cummings 1996.

2.2.2 General Characteristics of Eukaryotic Cells (Fig. 2.2-3)

Compared to prokaryotic cells, eukaryotic cells exhibit a much more complicated structure. Inside the plasma membrane there are the nucleus and the cytoplasm, which encompass the fluid cytosol and many organelles. These are compartments enclosed by individual membranes, which are devoted to specific functions.

Nucleus: All eukaryotic cells show the presence of a separate nucleus, which contains the major portion of the genetic material of the cell. (The rest is present in mitochondria and chloroplasts, see below.) The nuclear DNA is organized in a number of chromosomes. Each double helix of chromosomal DNA (2.6.3) can be present once (in haploid organisms) or twice (in diploid organisms). During cell division (4.3), the condensed chromosomes arrange themselves separately. Otherwise they are combined with proteins as a ball of chromatin with an elaborate fine structure (2.6.4).

The number of chromosomes present in the various species differs widely (from 4 to >500; humans have 46 in the diploid set). While bacterial genomes contain $< 10^6 \ldots 5 * 10^7$ bp, the diploid set of eukaryotic DNA varies between ca. 10^7 bp (some fungi) and $> 10^{11}$ bp (lungfish, some algae). The diploid human genome contains ca. $5.8 * 10^9$ bp ($2 * 2{,}900{,}000$ bp).

In addition to the DNA the nucleus also contains the nuclear matrix, which is composed of the enzymes and factors required for DNA replication, DNA repair, transcription, and processing of the transcription products (3.9, Chapter 4).

The nucleus is surrounded by a double membrane of lipid bilayers with integrated proteins. Nuclear pores (4.5.3, ca. 125 nm diameter) span the nuclear membrane and enable the transport of proteins, rRNA etc. The inner surface of the nuclear membrane is covered by nuclear lamina, a net of protein fibers which stabilizes the structure

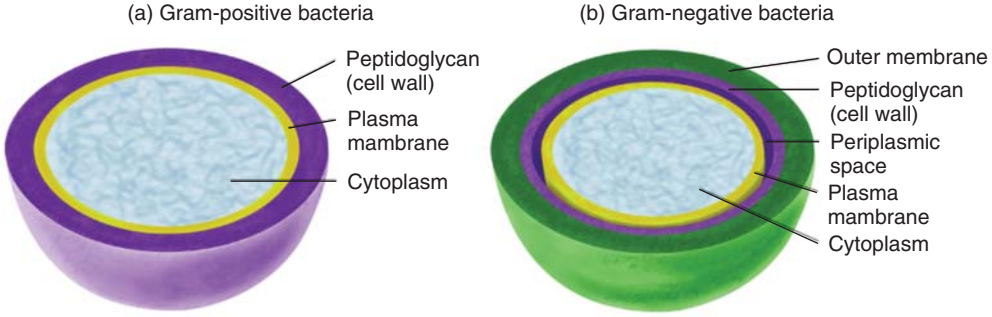


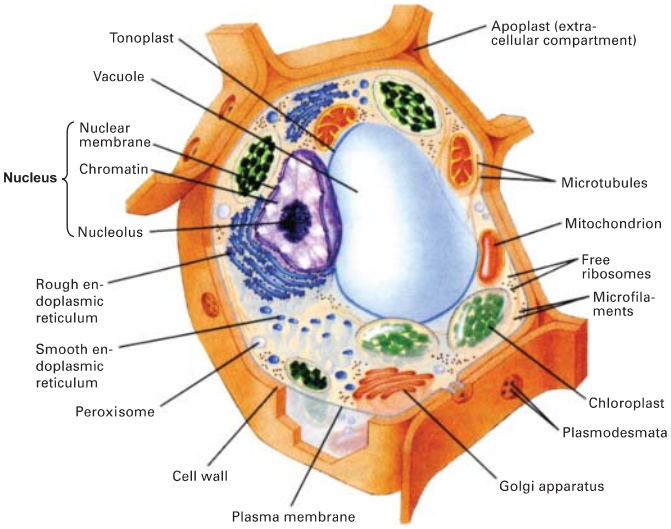
Figure 2.2-2. Structure of the Bacterial Envelope
After Voet, D. and J.G, Pratt, C.W.: Fundamentals of Biochemistry. John Wiley & Sons 2002, Figure 8-14.

and provides attachment points for the chromatin (2.6.4). During cell division, the nuclear membrane dissolves.

Cytosol: Although in eukarya many cytosolic functions have been taken over by specific organelles (see below), a large number of metabolic reactions still take place here. This includes glycolysis (3.1.1) and the synthesis of cytosolic proteins (4.2) and fatty acids (3.4.1.1).

Endoplasmic reticulum (ER, Fig. 2.2-4, Section 4.4.1): This is a labyrinth of membranes, which frequently encompasses half of the total amount of membranes. It consists of a system of sac- and tube-like structures, which locally expand into cisterns. Its internal lumen is connected with the intermembrane space of the nuclear membrane.

Part of the ER is studded on the outside with ribosomes (rough ER), which take part in protein synthesis (4.2). The proteins thus formed enter the ER lumen and are processed mainly by glycosylation (4.4.2). They leave the ER via vesicles (4.5.2). The rough ER is also the site of membrane-phospholipid synthesis, thus providing membrane material for the departing vesicles (3.4.3, 4.4.1).



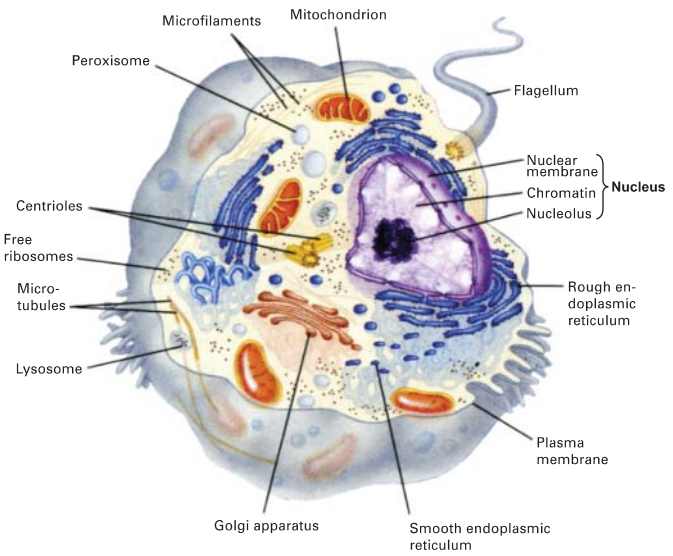


Figure 2.2-3. General Structure of a Plant Cell (Top) and an Animal Cell (Below)
The colors are for easy differentiation only. After Campbell, N.A.: Biology 4th Ed. Benjamin/Cummings 1996.

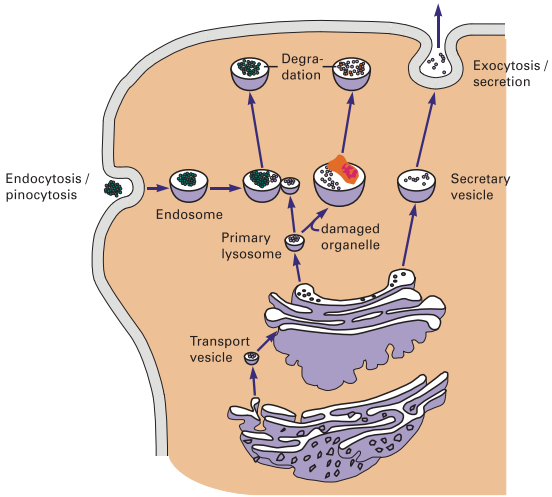


Figure 2.2-4. Interrelationship of Rough Endoplasmic Reticulum, Golgi Apparatus, Lysosomes, Endosomes and Transport Vesicles

The other part of the ER is free of ribosomes (smooth ER). Enzymes of the smooth ER are involved in the synthesis of fatty acids (desaturation, 3.4.1.3), phospholipids (3.4.3), steroids, especially steroid hormones (3.5.4), and other lipids. They are located on the outside of this organelle. The smooth ER plays a role in detoxification by hydroxylation reactions. Its equivalent in muscles (sarcoplasmic reticulum) stores Ca^{2+} ions and is thus involved in the many reactions regulated by this ion (7.4.4). In liver, the smooth ER is also the storage site for glycogen (3.1.2.2).

Golgi apparatus (4.5.2): This organelle consists of stacks of flattened membrane sacs, which are especially numerous in plants (dictyosomes). The properties on both sides of an individual sac and the sacs of the whole stack differ (details in 4.4.2). Their main function is the further processing and sorting of proteins and their export to the final targets. In most cases, these are secreted or membrane proteins. In addition, the Golgi apparatus also produces polysaccharides (e.g., hyaluronic acid, 2.9.2), glycosphingolipids (3.4.4.1) etc.

Lysosomes: Lysosomes are vesicles of 0.1 … 0.8 µm diameter, which are enclosed by a lipid bilayer. They are formed by budding from the *trans* side of the Golgi apparatus (4.4.2). These organelles are filled with many enzymes for polysaccharide, lipid, protein and nucleic acid degradation. They fuse with endosomes containing internalized lipoproteins (6.2.4) or phagocytosed nutrients (6.1.5) and hydrolyze these compounds (Fig. 2.2-4). They act also on intracellular material to be removed and even contribute to the apoptosis (programmed cell death, 7.6, 8.2.5) of their own cell. Lysosomes of special cells (e.g., macrophages) destroy, in the same way, bacteria or viruses as a defense mechanism (8.2.8).

The degradative enzymes of lysosomes exhibit an activity optimum at pH = 5, identical to the pH of the lysosome lumen. It is kept constant by continuous proton pumping into the lysosome lumen. The pH difference to the usual cytosolic pH of 7.0 is a safety measure, since after accidental leakage of some lysosomes, the released enzymes are almost inactive at cytosolic pH. Only a cumulated release from many lysosomes is deleterious to the cell.

Insufficient activity of lysosomal enzymes is the reason for many diseases. Examples are mucopolysaccharidoses (2.9.2), gangliosidoses (3.4.4) and glycogen storage diseases (e.g., Pompe's disease, 3.1.2.5).

Peroxisomes: This is another example of how enclosure into an organelle allows reactions to take place which would otherwise be deleterious to the rest of the cell. Peroxisomes (ca. 0.5 µm diameter) are surrounded by a single membrane. They are generated from components of the cytosol and do not bud from other membranes. The main task of these organelles is the performance of monooxygenase (hydroxylase) or oxidase reactions, which produce hydrogen peroxide (H_2O_2).

This dangerous compound is immediately destroyed by catalase within this enclosed space:

$X\text{-}H + O_2 + H_2O = X\text{-}OH + H_2O_2$ or

$X\text{-}CH_2OH + O_2 = X\text{-}CHO + H_2O_2$, followed by

$2\ H_2O_2 = 2\ H_2O + O_2$.

Reactions of this kind are the oxidative degradation of fatty acids (3.4.1.5), of alcohols etc. In glyoxisomes of plants, the reactions of the glyoxylate cycle (3.1.9.1) take place.

Mitochondria (Fig. 2.2-5): In a typical eukaryotic cell, there are in the order of 2000 of these organelles, which are often of ellipsoidal shape (length ca. 1 … 10 µm). They have a smooth outer membrane and a highly folded inner membrane with numerous invaginations (cristae), which contain most of the membrane-bound enzymes of mitochondrial metabolism. The internal area contains the mitochondrial matrix, while the intermembrane space is the narrow area between both membranes. Since protons are permanently pumped out through the inner membrane, the matrix is more alkaline than the intermembrane space and the cytoplasm (pH 8 vs. ≈ pH 7).

Mitochondria are the site of respiration and ATP synthesis (3.11.4.5), but also of many other central reactions of metabolism, e.g., citrate cycle (3.1.8), fatty acid oxidation (3.4.1.5), glutamine formation (3.2.2.1), and part of the pathway leading to steroid hormones (3.5.5.1, 3.5.8.1). The latter sequence, as well as the initiation of gluconeogenesis (3.1.3.5), are examples of how the site of reactions frequently changes from one organelle to another within a single pathway.

Besides chloroplasts (2.2.3), mitochondria are the only organelles which are equipped with their own (circular) DNA, RNA and ribosomes and thus can perform their own protein synthesis. The components and the mechanism resemble more the bacterial than the eukaryotic system. However, less than 10 % of the mitochondrial proteins are generated by this means, the rest is encoded by nuclear DNA and imported (4.5). Mitochondria reproduce by binary scission similar to bacteria. Their membranes do not exchange material with the rest of the cellular membrane system; the membrane proteins are produced on internal or cytosolic ribosomes. The membranes contain many transport systems resembling those of bacteria. These and other arguments are the base of the endosymbiont theory, which assumes that mitochondria have originated from ingested bacteria (e.g., aerobically living heterotrophs). They provide effective energy production, but are dependent on the host in many aspects.

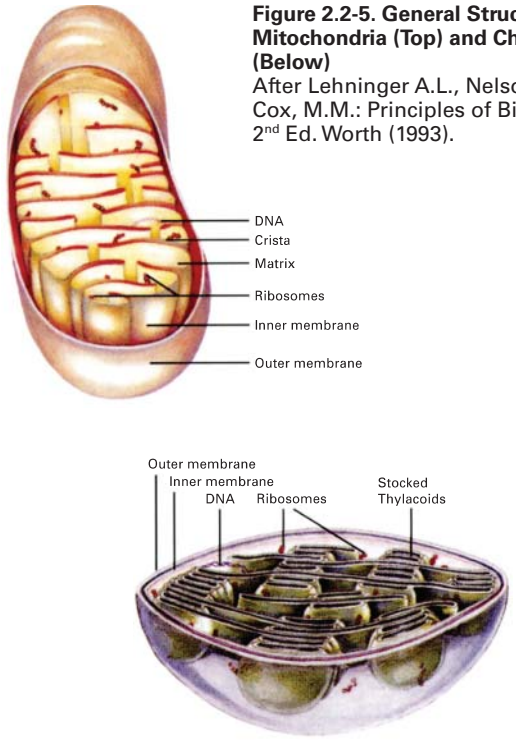


Figure 2.2-5. General Structure of Mitochondria (Top) and Chloroplasts (Below)
After Lehninger A.L., Nelson, D.L., Cox, M.M.: Principles of Biochemistry 2nd Ed. Worth (1993).

Cytoskeleton: The internal cytoskeleton is a general component of eukaryotic cells, but is of special importance in animal cells. Here it is the primary factor for maintenance of the external structure, which is achieved in plant cells by means of the rigid cell wall (2.2.3). The cytoskeleton also provides anchoring points for the organelles and for some enzymes and even contributes to movements either by its own formation and degradation or by motor proteins moving along the filaments.

The cytoskeleton is composed of microtubules, actin filaments and intermediary filaments (Table 2.2-1). More details on their composition, formation and degradation are presented in 6.1.6.

Table 2.2-1. Components of the Cytoskeleton

	Microtubules	Microfilaments	Intermediary Filaments
Components	α/β-tubulin	actin	keratin, vimetin, desmin etc.
Shape	hollow tube of 13 tubulin sequences	2 strings of actin monomers, coiled	variable, frequently coiled coil of α helices
Diameter	25 nm	7 nm	8 … 12 nm
Examples of function	chromosome separation (4.3.5), axonal transport (7.2.6), other movements	muscle contraction (7.4.5), movement of the cytoplasm, cell division (furrow ingression, 4.3.5)	structure stabilization, anchoring of organelles

The hollow microtubules (Fig. 6.1-3) frequently originate from the centrosome, which is located close to the nucleus (4.3.5). The moving parts of cilia and flagella are circular arrangements of tubulin. Attached dynein molecules cause movements of the tubules relative to each other and thus curvature of these cellular annexes (6.1.6).

Actin filaments associate with myosin. The heads of the myosin molecules can 'walk along' the string of actin monomers. The best known action of actin filaments is the muscle contraction (7.4.5), but they are also involved in many other, mostly movement, functions. E.g., the cytoplasmic movement in some algae is caused by the transport of organelles along actin filaments by myosin.

Intermediary filaments are composed of variable units, depending on the particular cell type. In general they have a structural function. They are more long-lasting than the other components of the cytoskeleton, but are also subject to rearrangements. Vimetin occurs in, e.g., endothelial cells and adipocytes and anchors the nucleus and the fat droplets. Desmin filaments keep the Z disks of muscle cells in place (7.4.5). Neurofilaments reinforce the long axons of neuronal cells. The nuclear lamina (see above) consists likewise of intermediary filaments.

2.2.3 Special Structures of Plant Cells (Fig. 2.2-3)

Plant cell walls: Cell walls are an essential factor distinguishing plant and animal cells. They provide stability and prevent an expansion of the cell beyond its fixed size. These secondary walls consist mainly of cellulose (3.1.2.2), hemicelluloses (3.1.6.3) and lignin (3.13.1.2). During an early phase of formation, the primary walls contain considerable amounts of pectate (3.1.5.6) and glycoproteins (4.4.1). On the inside, the cell contents are enclosed by the usual lipid bilayer membrane. Plasmodesmata interconnect neighboring cells and enable the transport of water and metabolites.

Chloroplasts (Fig. 2.2-5): These are lens-shaped organelles of about 1 … 5 µm diameter, which occur only in photosynthesizing green plants. They are enclosed by a double membrane with a thin intermembrane space in between. The interior contains the fluid stroma and a third membrane system surrounding the thylakoid space. This system has the shape of interconnected flat disks, which, in most cases, are stacked on top of each other (grana). The thylakoid membranes are the site of photosynthesis (3.12.1). In photosynthesizing bacteria, their role is taken over by the cytoplasmic membrane.

Due to the permanent pumping of protons into the thylakoid space, its interior is much more acidic than the stroma (pH = 5 vs. pH = 8).

While some chloroplast proteins originate from their own protein synthesis system, the majority are nuclear encoded and are imported by a special mechanism into the stroma and the thylakoid space (4.5).

Analogous to mitochondria (2.2.2), chloroplasts resemble bacteria in several aspects (protein synthesis, membrane structure, reproduction etc.). Likewise the endosymbiont theory refers to them. Possible precursors of chloroplasts could be photoautotrophic cyanobacteria.

Vacuoles: Although vacuoles also occur in other kingdoms (e.g., protists), they are most prominent in plants. Young plant cells contain several vacuoles, which originate in vesicles released from the endoplasmic reticulum and the Golgi apparatus. In mature plant cells they combine to form a single central vacuole, which is enclosed by a membrane (tonoplast).

The central vacuole can occupy up to 90 % of the total cell volume and is a storage space for inorganic salts, saccharose, proteins, pigments and waste. The accumulation of these compounds causes the inflow of water by osmosis and keeps up the internal pressure (turgor).

2.2.4 Special Structures of Animal Cells

Animal cell membrane and extracellular matrix (Fig. 2.2-6): The actual cell membrane consists only of the lipid bilayer. However, it is covered by a complex, gelatinous extracellular matrix, which is formed by glycoproteins secreted from the cell. Its main component is collagen (2.3.1). The collagen fibers are enclosed by a network of carbohydrate rich proteoglycans (2.9.1.2). Tasks of the extracellular matrix are the interconnection and anchoring of the cells, the support of the structure and to a certain extent also the localization of cells (e.g., during embryonic development).

Depending on their organ-specific structure, they allow the passage of fluids and dissolved material at varying degrees (e.g., filtration effects in the glomerulus of the kidney). The components of the extracellular matrix are connected via linker proteins (mostly fibronectin) with membrane-spanning proteins (integrins). The integrins, in turn, are associated at the interior side of the membrane with microfilaments of the cytoskeleton.

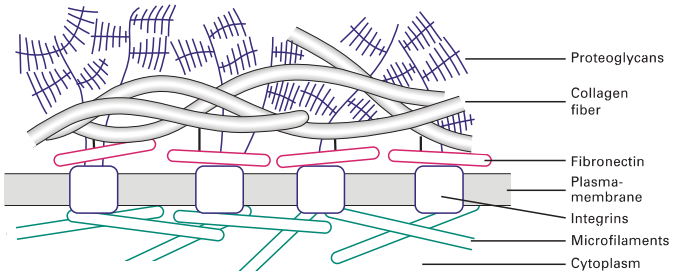


Figure 2.2-6. Structure Elements of the Extracellular Matrix

Interconnection between cells (Fig. 2.2-7): Neighboring cells (especially epithelial cells) are interconnected in several different ways. Tight junctions firmly attach the cells to each other and prevent any fluid passage between them. Point desmosomes (also frequently occurring in plants) are firm interconnections of neighboring cells via a network of glycoprotein filaments. They are reinforced on the inside of the cells by filaments of the cytoskeleton. Small channels known as gap junctions pass through both membranes of neighboring cells. They allow the exchange of small molecules and ions (7.2.3) and represent an equivalent of plant plasmodesmata (2.2.3). Pores, channels, transporters and receptors involved in material and signal passage through individual membranes are described elsewhere (7.1.2, 7.2.3, 6.1).

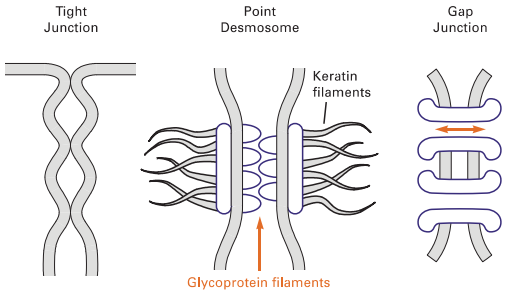


Figure 2.2-7. Types of Contact between Cells

Literature:

de Duve, C. *A Guided Tour of the Living Cell*. Scientific American Books, 1984.

Textbooks of biology, e.g., Campbell, N.A. *Biology* 4th Ed. The Benjamin/Cummings Publ. Co, 1996.

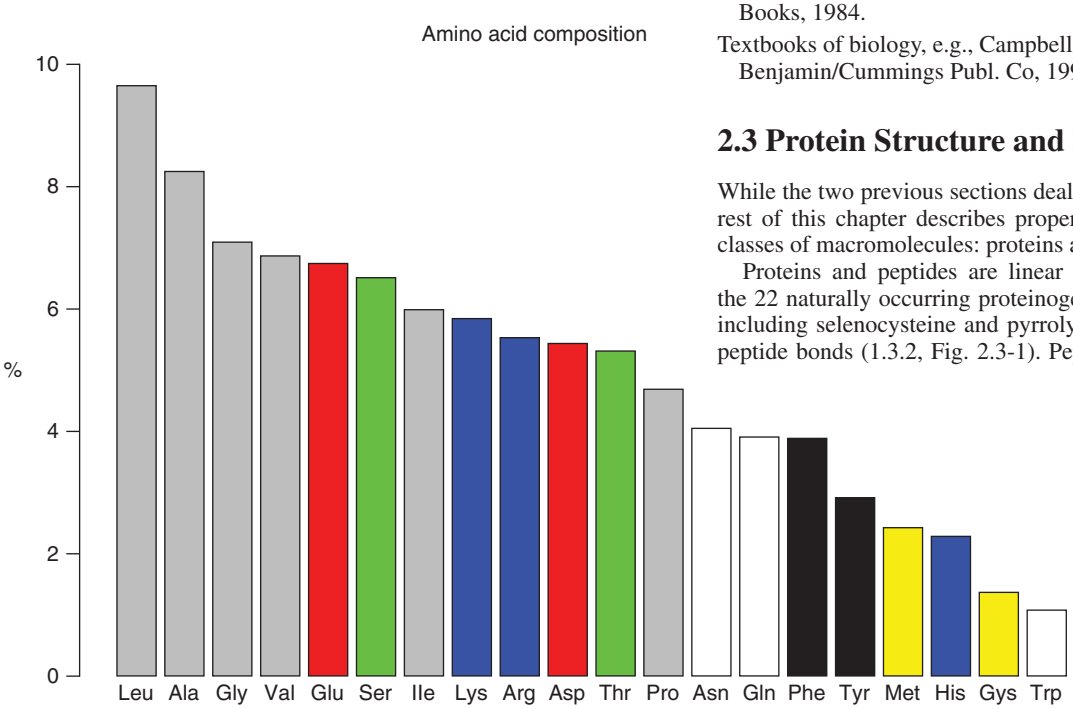


Figure 2.3-1. Amino Acid Composition of Proteins
(gray = aliphatic, red = acidic, green = small hydroxy, blue = basic, black = aromatic, white = amide, yellow = sulfur). After UniProtKB/Swiss-Prot protein knowledgebase (see references).

2.3 Protein Structure and Function

While the two previous sections dealt with the cellular structures, the rest of this chapter describes properties of the two most important classes of macromolecules: proteins and nucleic acids.

Proteins and peptides are linear polymers which are built from the 22 naturally occurring proteinogenic L-amino acids (1.3 and 3.2, including selenocysteine and pyrrolysine, 4.1.3.4) and are linked by peptide bonds (1.3.2, Fig. 2.3-1). Peptides containing D-amino acids

are less frequent. They occur mostly in microorganisms. Chains of up to about 50 amino acid residues are usually named 'oligo-(or poly-) peptides', larger ones are termed 'proteins'. However, the expression 'peptides' is frequently also used for the whole class of compounds containing peptide bonds. The average protein sequence length listed in UniProtKB (Chapter 10) is 352 amino acids, the longest sequence has 35,213 amino acids.

The average amino acid distribution in proteins is far from even, some amino acids occur much more frequently than others. Leucine is found nine times as often as tryptophan (Fig. 2.3-1).

All polypeptides and proteins are synthesized by ribosomal synthesis, using mRNA as the source of information for the sequence of amino acids (4.1.2, 4.2.2). However, a number of small peptides are formed by a non-ribosomal sequence of specific enzyme reactions, e.g., glutathione (3.2.5.7), the penicillin precursor peptide (3.10.8) and many peptide antibiotics.

After formation of the peptide chain, additional post-ribosomal modifications of the amino acid components may take place, e.g., hydroxylation (3.2.3, 3.2.5.2), carboxylation (3.2.2.2, 3.7.13), methylation (e.g., of glutamate during chemotactic mechanisms), acylation (acetylation, myristoylation, palmitoylation, 4.4), phosphorylation (2.5.2, 7.5), glycosylation (4.4.2) and formation of disulfide bonds between the -SH groups of cysteine (4.5.1).

In a number of cases, a long peptide is cleaved after synthesis, e.g., to be activated for a special function. Examples are hormones (7.1.3, 7.1.5), digestive enzymes (2.5.1) and blood coagulation factors (9.2 … 5).

2.3.1 Levels of Organization

The amino acid sequence contains all the information necessary to determine the three-dimensional structure of the protein which is assumed to be the thermodynamically most stable one among the kinetically possible ways of folding (4.5). Several levels of structural organization can be distinguished:

Primary structure (Fig. 2.3-2): This is the amino acid sequence. The CO–NH bond is fairly rigid and usually assumes the *trans* configuration. To a limited extent, rotations are possible around both other bonds of the peptide backbone (torsion angles Φ and Ψ). Steric restrictions are caused by the side chains of the amino acids and the oxygen atom of the CO group.

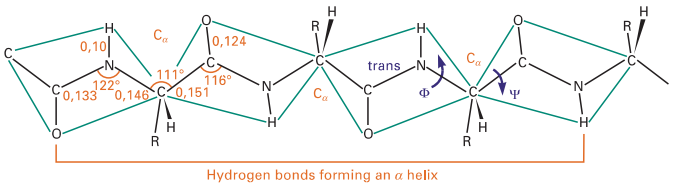


Figure 2.3-2. Primary Structure of a Polypeptide Chain in Extended Conformation
The distances (in nm = 10 Å) and the bond angles are shown. The green quadrangles indicate the rigid structure of the peptide bonds. Ψ and Φ are the angles characterizing the rotation around the Cα-CO and N-Cα bonds (in this figure Ψ = Φ = 180°). The formation of hydrogen bonds leading to α helices is indicated.

Secondary structure (Fig. 2.3-3): These are regular arrangements of the backbone of the polypeptide chain, which are stabilized by hydrogen bonds between amide and carboxyl groups of the peptide and by sterically favorable backbone torsion angles. Frequently occurring secondary structures are the α-helix and the β-pleated sheet (Fig. 2.3-3). In globular proteins, a variable portion (often below 40 %, sometimes up to 100 %) is arranged in less regular loop structures, which are frequently found on the outside of water-soluble proteins.

The α-helix is a rod-like structure, which is formed by hydrogen bonds between a CO group and the fourth following NH group of the sequence (shown in Fig. 2.3-2). Helices formed by hydrogen bonds between other CO–NH pairs are rare. There are also van der Waals bonds across the helix. The side chains of the amino acids point to the outside. All α helices occurring in nature are right-handed (i.e., when looking along the axis of the helix from the N to the C terminus, the chain turns clockwise). One turn is formed by 3.6 residues, the pitch amounts to 0.54 nm. α-Helices occur both in globular and in fibrous proteins and are found in almost all transmembrane sections.

The β-pleated sheet forms slightly twisted planes. This structure is also stabilized by hydrogen bonds, however, between different strands. Between 2 … 15 strands may arrange themselves in parallel or antiparallel directions. The side chains of consecutive amino acids extend sideways in opposite directions.

Tertiary structure is a term which refers to larger arrangements, describing the location (x,y,z coordinates) of all atoms in space. Additional levels are frequently used:

Supersecondary structures are combinations of secondary structure elements, such as the coiled-coil α-helices (e.g., in many fibrillar proteins) or barrels (in globular proteins) or nucleotide-binding structures. A number of structural motifs occur in these superfolds.

Two or more α-helices can associate into an antiparallel arrangement, which allows interactions of their side chains (αα motif). This is an essential element of fibrous proteins (e.g., in α-keratin, Fig. 2.3-7 and in the segments of spectrin, which is present in erythrocytes), but does also occur in globular proteins (e.g., in hemoglobin, Fig. 2.3-6).

A number of common motifs occurring in globular proteins are caused by the turns of the peptide backbone at the end of β-pleated sheets (Fig. 2.3-4): Antiparallel β-pleated sheets frequently form β-loops (also called hairpin loops), which are stabilized by hydrogen bonds between a CO group and the third NH group of the sequence. A series of such turns can form a β-meander. Another arrangement of turns is the so-called 'Greek key'. Parallel β-pleated sheets require crossover connections, which often assume right-handed helix structures (βαβ motif). These loops and crossover connections frequently occur on the outside of globular proteins.

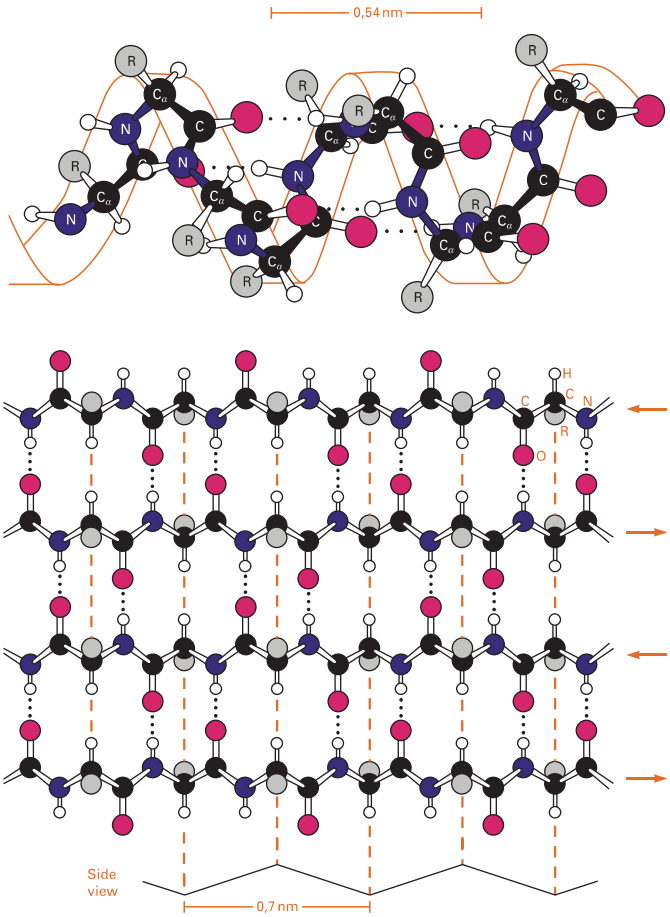


Figure 2.3-3. α-Helix (Top) and Antiparallel β-Pleated Sheet Structure (Bottom)
C = black, O = red, N = blue, H = white, side chains = grey.

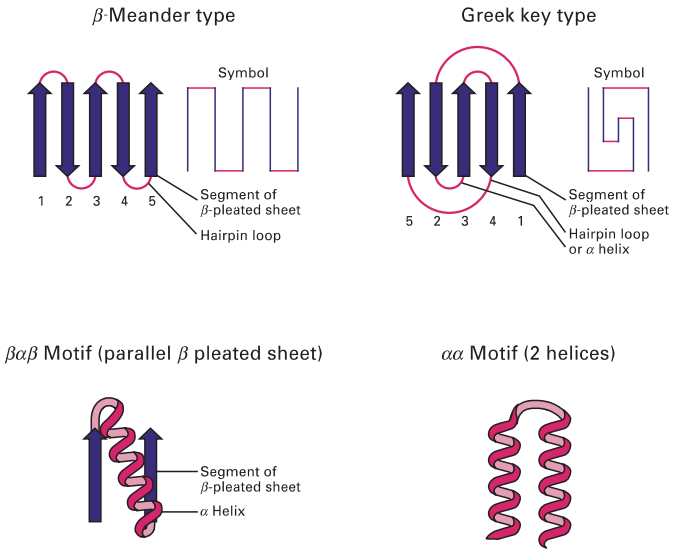


Figure 2.3-4. Elements of Supersecondary Protein Structures
βαβ Units (antiparallel β-pleated sheet).

The 'plane' of β-pleated sheets in globular proteins is always slightly twisted in a right-handed direction. A large β-pleated sheet can roll up to a β-barrel structure (e.g., triosephosphate isomerase). Another possibility is a saddle-shaped arrangement (e.g., bovine carboxypeptidase A, Fig. 2.3-5).

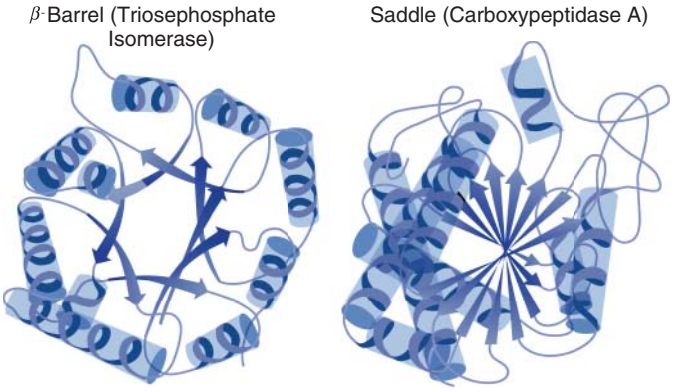


Figure 2.3-5. Examples of Supersecondary Structures

Structural domains are structurally independent globular arrangements of ca. 100 … 200 amino acid residues. They are connected with each other by flexible polypeptide segments (e.g., in immunoglobulins, Fig. 8.1-5) and mostly assume their fold even when separated from the rest of the chain.

Globular proteins are the result of the folding of complete single polypeptide chains into quite compact structures.

Globular proteins are usually arranged with most of the hydrophobic amino acid residues buried in the interior and hydrophilic residues facing the outside. There are exceptions: internal polar groups can be required for the protein function (e.g., in the active center of enzymes, 2.4.1). Transmembrane protein sections are usually structured as an α-helix composed of 19 … 20 amino acids with nonpolar side chains. A number of transmembrane passes can be circularly arranged with hydrophobic residues facing the lipid membrane and polar residues forming a channel, which permits the passage of polar or charged molecules (6.1.2). The size of the pore and the electric charges of its lining determine its specificity. Often the passage is regulated by controlled movements of the protein. Still more complicated are transporters, which interact with the cargo and often require an input of energy (6.1.4).

Fibrous proteins frequently consist of several helices turned around each other as a 'coiled coil'.

α-Keratin is the principal component of the epidermal layer and also of hair, feathers etc. Regularly spaced nonpolar residues in both α-helices interact noncovalently. They form a hydrophobic contact strip and cause both helices to wind in a left-handed mode around each other (Fig. 2.3-7) as the primary keratin element.

Collagen is the major component of the connective tissue of the skin and of blood vessels. In collagen type I, three individual helices [left-handed (!), 3.3 amino acids/turn, pitch 1 nm] wind around each other, forming a right-handed superhelix (pitch 8.6 nm, total length ca. 300 nm, Fig. 2.3-8). Its formation is also caused by a regular amino acid arrangement within the individual helices: Every third amino acid is glycine, which points towards the center of the superhelix. Its amino groups exhibit strong hydrogen bonds with a CO group of another chain. The other amino acids are mostly proline and hydroxyproline, which contribute to the rigidity due to their inflexible bond structure (1.3.2). Individual superhelices associate to fibers.

Generally, protein domains having more than 30 % of their sequence in common will adopt the same overall fold. However, even proteins with lower sequence similarities can belong to the same structural family.

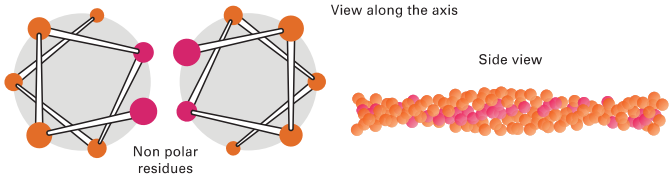


Figure 2.3-7. Structure of α-Keratin

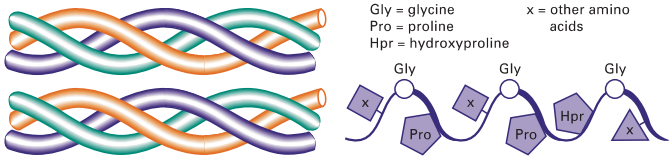


Figure 2.3-8. Structure of Collagen

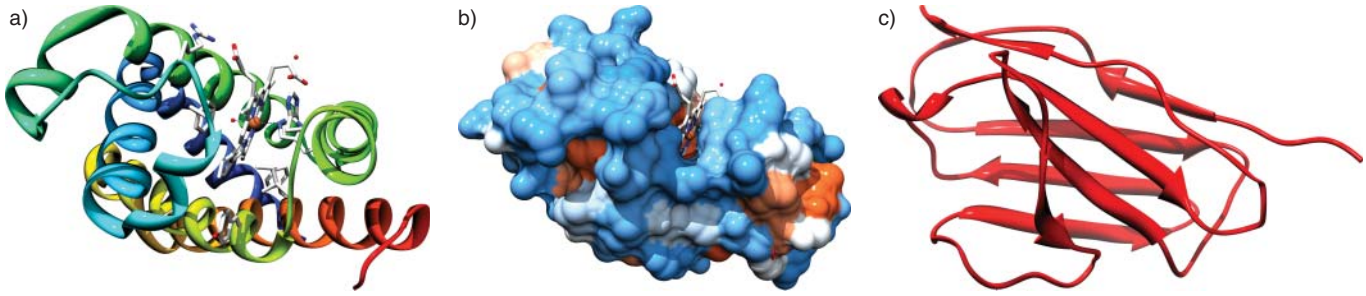


Figure 2.3-6. Tertiary Structures of Proteins Showing Different Aspects of Secondary Structures
a) Hemoglobin main chain (with central heme group), b) Hemoglobin surface. c) *E.coli* cytotoxic necrotisizing factor type 1.

Quaternary structure: This term is used for the association of two or more polypeptide chains into a multi-subunit aggregate. This arrangement can contribute to

- fast and effective transfer of the substrate from one active center of enzymes to the next, e.g., in
  - fatty acid synthase (3.4.1.1)
  - biotin dependent CO_2 transfer (Fig. 3.4.1-2)
- allosteric regulation (2.5.2), e.g., of
  - hemoglobin (Fig. 3.3-9).

2.3.2 Protein Function

Proteins and peptides play crucial roles in virtually all processes occurring in living cells and organisms. Their functional diversity is represented by the examples in Table 2.3-1.

Table 2.3-1. Biological Function of Proteins

Function	Section[1]	Examples
Enzyme catalysis	2.4, 2.5	lactate dehydrogenase, trypsin, DNA polymerases
Transport	3.3.2.2, 6.1, 6.2	hemoglobin, serum albumin, membrane transporters
Storage		ovalbumin, egg-white protein, ferritin
Motion	7.4.5, 6.1.6	myosin, actin, tubulin, flagellar proteins
Structural and mechanical support	2.2.2, 5.1.2	collagen, elastin, keratin, viral coat proteins
Defense	8.1, 8.2	antibodies, complement factors, blood clotting factors, protease inhibitors
Signal transduction	7.1 … 7	receptors, ion channels, rhodopsin, G proteins, signalling cascade proteins
Control of growth, differentiation and metabolism	4.1.2, 4.2.2, 4.3, 7.1 … 7, 8.1.8	repressor proteins, growth factors, cytokines, bone morphogenic proteins, peptide hormones, cell adhesion proteins
Toxins	3.4.3.2, 7.4.1	snake venoms, cholera toxin

[1] Only those sections are given, where the principle of action is presented. Additional information is given in the sections dealing with the individual proteins.

Literature:

Kleinkauf, H., von Döhren, H. Eur. J. Biochem. 1996;236:335–351.

Kyte, J. *Structure in Protein Chemistry*. Garland (1994).

Matthews, B.W. Ann. Rev. Biochem. 1993;62:139–160.

(various authors). *Methods in Enzymology*, especially Vols. 82, 85, 91, 117, 130, 131, 144, 145. Academic Press.

UniProtKB/Swiss-Prot protein knowledgebase release 2010_06 statistics.Website:http://expasy.org/sprot/relnotes/relstat.html.

2.4 Enzymes

Enzymes are proteins which catalyze the chemical modification of other compounds (substrates) without being consumed themselves and without changing the equilibrium of the reaction.

- They accelerate the rate of reactions of their substrate by factors up to ca. 10^{15} as compared to the uncatalyzed reaction.
- They often show substrate specificity by limiting their action to one substance or a small number of structurally related substances
- They also show reaction specificity by catalyzing only one of several possible ways of substrate conversion.

Whereas reaction specificity usually is absolute, the degree of substrate specificity may vary among different enzymes.

2.4.1 Catalytic Mechanism (For quantitative aspects, see 1.5.4)

Every conversion of a molecule into another one requires at first an input of energy (activation energy $\Delta G^{\ddagger}$), which leads to a transition state. From this state, either the conversion to the product or a return to the original molecule takes place.

Single-substrate reaction: In an enzyme-catalyzed reaction involving a single substrate (S), this molecule combines with the enzyme (E) to form an enzyme-substrate complex (ES), followed by conversion to the product (P) which is then released. The enzyme is then free to enter the next catalytic cycle:

$$E + S \leftrightarrow ES \leftrightarrow EP \leftrightarrow E + P.$$

The interaction of the enzyme and the substrate facilitates the formation of the transition complex. Therefore, a lower amount of free energy $\Delta G^{\neq}$ is required for its formation and thus for the consecutive conversion of the substrate to the product (Fig. 2.4-1). This results in an increase of the reaction velocity. The same is true for the reverse reaction. Since the equilibrium constant K is determined by the overall reaction rate constants k_f and k_r for the forward and reverse reactions, respectively

$$K = \frac{k_f}{k_r} \qquad [2.4\text{-}1]$$

and enzymatic catalysis increases both rate constants by the same factor, their ratio remains constant and the equilibrium does not change.

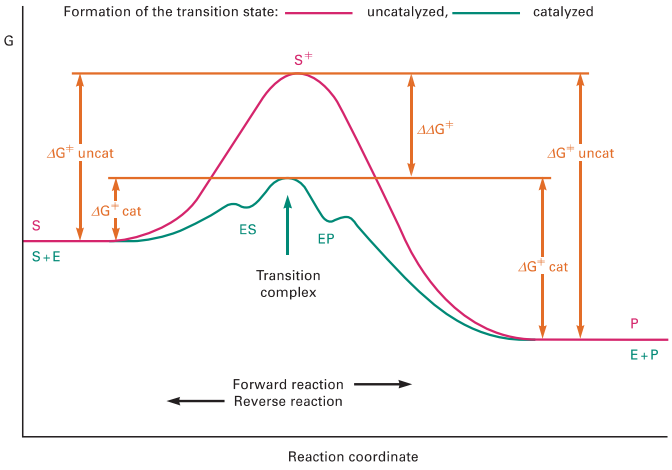


Figure 2.4-1. Changes of the Free Energy $\Delta G^{\ddagger}$ During the Course of the Reaction S → P (For Details, see 1.5.4)

There is an exponential relationship between lowering of the activation energy and increase of the reaction rate (Eq. [1.5-32]). 'Good' enzymes, which noticeably decrease the activation energy, elevate the substrate turnover tremendously.

If the activation energy $\Delta G^{\neq}$ [kJ*mol^{-1}] necessary for an uncatalyzed reaction is lowered by $\Delta\Delta G^{\neq}$ [kJ*mol^{-1}] in the presence of the enzyme, an increase of the reaction rate by the factor $e^{\Delta\Delta G^{\neq}/RT}$ takes place. At ambient temperature, a lowering by $\Delta\Delta G^{\neq}$ = 5.7 [kJ/mol] increases the rate constant tenfold, a lowering by $\Delta\Delta G^{\neq}$ = 34 [kJ/mol] increases the rate a millionfold.

Two-substrate reactions: These reactions proceed analogously via enzyme-substrate complex formation. There are several possible ways of interaction between enzyme E and the substrates A and B (Details in 1.5.4):

- They may form a ternary intermediate complex (EAB). After the catalyzed reaction has taken place (in the case of two products EAB → EPQ), the products leave the complex. Formation and decomposition of the complex can take place either in definite order (ordered sequential mechanism) or at random (random

sequential mechanism). This is the preferred process when a group or an ion is directly transferred from one substrate to another.

Examples for the ordered mechanism are most NAD^+, NADH, $NADP^+$ or NADPH dependent dehydrogenase reactions (3.7.9.2) where the first contact to the enzyme is made by the coenzyme. A random mechanism is performed by some kinases.

- The substrate A may contact the enzyme E (or its prosthetic group) and modify it in some way (designated as F). Then it leaves as product P, before the second substrate B reacts with the modified enzyme, reestablishes its primary state and leaves as product Q. Thus, this ping-pong-mechanism is a sequence of binary intermediates.

$$E + A \leftrightarrow EA,\ EA \leftrightarrow FP,\ FP \leftrightarrow F + P$$
$$F + B \leftrightarrow FB,\ FB \leftrightarrow EQ,\ EQ \leftrightarrow E + Q$$

Examples for this reaction mechanism are chymotrypsin (Fig. 4.5.6-1) and the transaminases (Fig. 3.2.2-3). In the latter case, the cofactor pyridoxal phosphate, which exchanges amino and oxo groups, represents the temporarily modified part of the enzyme.

Catalytic center: The part of the enzyme structure involved in the binding of the substrate and in catalysis is called the catalytic center (active site). It usually encompasses only a small part of the enzyme and consists of a few amino acids, which are often present on non-contagious locations of the protein chain (and in some cases even on different protein chains of multimeric enzymes). Prosthetic groups or cofactors (see below) may also be present in the catalytic center.

The catalytic center is a three-dimensional unit, which performs interactions with definite structures of the substrate

- for recognition, providing the specificity and
- for catalysis.

Usually, the catalytic center has a cave-like structure which permits the exclusion of water (unless it is a reaction partner), but contains some polar amino acids involved in the catalytic mechanism. In a number of cases (e.g., carboxypeptidase A, 4.5.6.1, some dehydrogenases) a movement of amino acid residues seals the center after the substrate has entered, while in other cases, structural movements are part of the catalytic process or are required to allow access to the active center (some lipases). Movements during allosteric regulation play a special role (2.5.2).

Typical catalytic mechanisms: The enzymes may facilitate the formation of the transition complex in various ways, e.g., by

- general acid and base catalysis (proton donation and abstraction)
- covalent bond formation (the bonds must not be too strong, therefore the reacting group must be highly polarizable, e.g., imidazoles or thiols)
- redox effects (by metal ions, enzyme residues or cofactors)
- electrostatic effects (by metal ions or amino acid residues, enhanced by a mainly non-polar environment)
- orientation and proximity effects (moving the reactants into favorable positions and diminishing their relative motions)
- straining the bound substrate(s) into a transition-state like configuration

Examples for such mechanisms are found in various places in this book.

Cofactors: While in a number of reactions the enzyme protein exerts the catalytic activity on its own, in other cases, cofactors are needed for catalysis (Table 2.4-1). They can be

- Coenzymes, which are transiently associated with the enzyme and frequently serve as a carrier for atoms or electrons or effect the molecule transfer to other enzymes (e.g., NAD^+, coenzyme A).
- Prosthetic groups, which are firmly bound to the enzyme (e.g., biotin). The inactive enzyme protein is then termed apoenzyme which, upon combining with the prosthetic group, forms the enzymatically active holoenzyme.

Examples of the contribution of enzyme amino acid residues to the catalytic mechanism are shown in Figs. 4.5.6-1 and 4.5.6-2 (chymotrypsin and pepsin), for the involvement of metal atoms in Figure 4.5.6-3 (carboxypeptidase A) and for a non-covalently bound prosthetic group in Figure 4.2-3, 4.4-2 and 4.4-3 (pyridoxal phosphate). The structure of the intermediate complex of L-lactate and NAD^+ with the catalytic center of L-lactate dehydrogenase is shown in Figure 2.4-2.

Effects of pH: In most cases, positively or negatively charged amino acid side chains are involved in substrate binding and catalysis. Therefore, most enzymes have a characteristic pH value at which their activity is maximal (Fig. 2.4-3). Above and below this pH the activity declines.

Effects of temperature: At higher temperatures all chemical reactions occur with an increased reaction rate. This is also true for

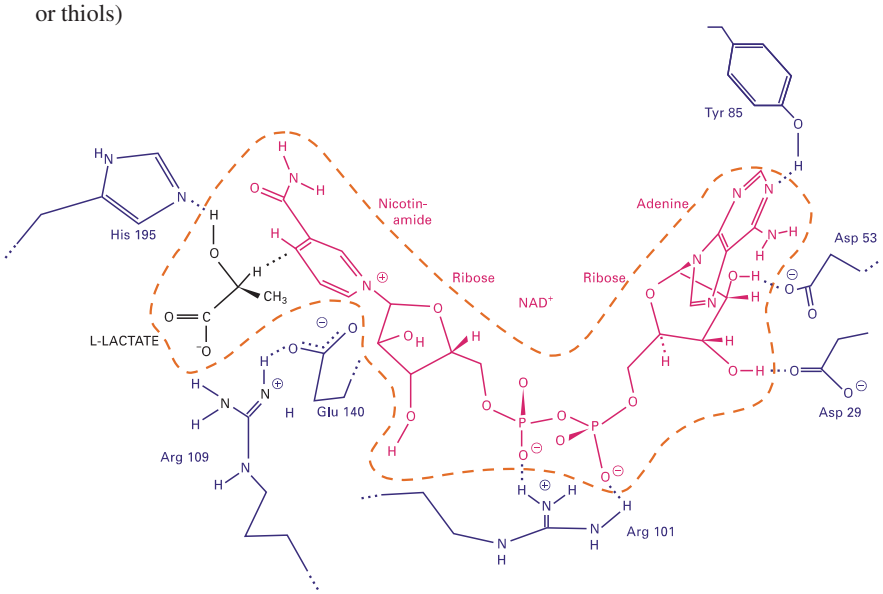


Figure 2.4-2. Catalytic Center of L-Lactate Dehydrogenase
(Blue: Enzyme, Red: NAD+, Black: L-Lactate)

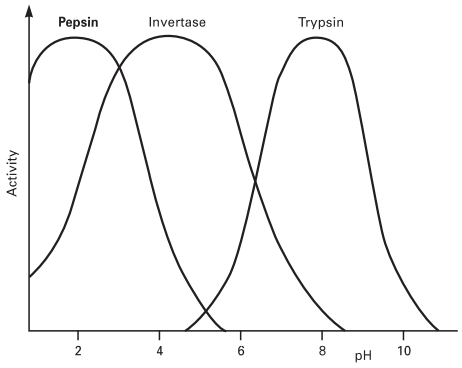


Figure 2.4-3. pH Dependence of an Enzyme-Catalyzed Reaction (Schematically)

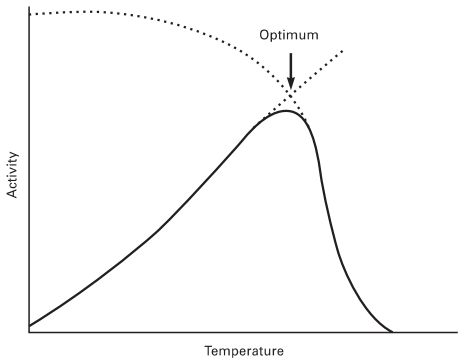


Figure 2.4-4. Temperature Dependence of an Enzyme-Catalyzed Reaction
(Schematically)

Table 2.4-1. Examples of Cofactors in Enzyme-Catalyzed Reactions

Coenzymes or Prosthetic Groups	Function (Examples)	Section
Metal ions (e.g., Mg^{++}, Mn^{++}, Zn^{++}, Cu^{++}, Co^{++}, Fe^{++})	participation in many enzyme-catalyzed reactions	4.5.6.5
Nicotinamide nucleotides (NAD^+, NADH, $NADP^+$, NADPH)	redox reactions, hydrogen transfer	3.7.9.2
Flavin nucleotides (FAD, FMN)	redox reactions, electron and hydrogen transfer	3.7.3.2
Quinones (e.g., ubiquinone, phylloquinone, pyrroloquinoline quinone)	redox reactions, electron and hydrogen transfer	3.2.7.2, 3.10
Coenzyme A (CoA)	acyl group transfer	3.7.7.2
Pyridoxal phosphate (PyrP)	aminoacyl group interconversions, decarboxylation	3.7.4.2
Thiamine pyrophosphate (ThPP)	group transfer and decarboxylation	3.7.2.2
Biotin	CO_2 transfer	3.7.8.2
S-Adenosyl methionine (SAM)	methyl group transfer	3.7.5.4
Deoxyadenosyl cobalamine (Coenzyme B_{12})	methyl group transfer	3.7.5.2
Lipoate (Lip)	transfer and oxidation of carbonyl compounds	3.7.14.1
Nucleoside diphosphates (UDP, CDP, ADP, GDP, dTDP)	glycosyl group transfer	3.1.2.2, 3.1.4.1, etc.

enzyme-catalyzed reactions (Eq. [1.5-32]). Beyond an optimum temperature, however, denaturation of the enzyme protein results in a quick decrease of the enzyme activity (Fig. 2.4-4). For many mammalian enzymes, this optimum temperature is about 40 … 50 °C. On the other hand there are enzymes of thermophilic archaea, which are still active at >100 °C.

2.4.2 Isoenzymes

Within a single given species, variants of an enzyme may be present. This may be caused by the occurrence of multiple gene loci coding for distinct versions of the enzyme protein, by the existence of multiple alleles at a single gene locus, or by different splicing positions. Such variations are designated isoenzymes if they represent multiple forms of the same enzyme activity present in the same species but are encoded by independent genes. Multiple forms of enzymes resulting from different post-translational modifications are not true isoenzymes.

Well-known examples of isoenzymes are animal lactate dehydrogenases (e.g., bovine, porcine, or rabbit) which are formed from random association of two types of subunits (H- and M-types) into tetramers, resulting in five discernible species HHHH, HHHM, HHMM, HMMM, MMMM. In mammalian creatine kinases, two different subunit types (B- and M-type) can associate to give three forms of the dimeric enzyme. Isoenzymes may have different kinetic or regulatory properties.

2.4.3 Multienzyme Complexes

Inside the cell, enzymes frequently catalyze multistep anabolic or catabolic sequences where the product of one enzyme becomes the substrate of the next. In many cases, the individual enzymes are dissolved in the cytosol without physical association, and the substrates which are turned over diffuse more or less rapidly from one enzyme molecule to the next. However, there are also higher levels of organization in multi-enzyme systems:

- Two or more enzyme activities may reside within one polypeptide chain which, therefore comprises more than one active site (multifunctional enzymes), e.g., some enzymes of the purine or pyrimidine synthesis in vertebrates (3.6.1.1, 3.6..2.1).
- Several individual enzymes may be associated non-covalently or even covalently to give a multienzyme complex, such as the pyruvate (3.1.3.1) and the 2-oxoglutarate dehydrogenase complexes (3.1.8.1), the biotin-containing carboxylases (3.7.8.2) or the polyketide synthases (3.10.8). The fatty acid synthase of mammals (3.4.1.1) is a homodimer of two cooperating multifunctional enzyme chains.
- The most highly organized enzyme systems are associated with supramolecular structures, such as ribosomes (4.1.3.2, 4.2.3.1), the proteasome (4.5.6) or the components of the respiratory chain which spans the inner membrane of mitochondria (3.11.4).

2.4.4 Reaction Rate (For a Detailed Discussion, see 1.5.4)

The dependence of the initial reaction rate v_0 of an enzyme reaction on the concentration of a single substrate [S] is given by the Michaelis-Menten equation (Eq. [1.5-19]):

$$v_0 = \frac{V_{max} * [S]}{K_M + [S]}$$

that represents a hyperbolic curve (Fig. 1.5-1). V_{max} is the maximum velocity obtained with a fixed amount of enzyme, when the enzyme is saturated with substrate. K_M, the Michaelis constant, equals the substrate concentration at which $v = 0.5\ V_{max}$. A similar plot is obtained for a two-substrate reaction when the concentration of one substrate is kept constant and only the other is varied. Since *in vivo* the substrate concentrations are frequently in the range of the K_M value, this can serve regulatory effects (Fig. 3.1-4).

2.4.5 Classification of Enzymes

A Nomenclature Committee set up by the International Union of Biochemistry and Molecular Biology (IUBMB) established a system for the classification of enzymes, and continuously classifies new enzymes. Presently there are 4151 classified enzymes. This system provides unambiguous enzyme names based on the nature of the catalyzed reaction and classifies the enzymes according to the catalyzed reactions. The application of this EC list has improved greatly the terminology of the scientific literature. Due to the continuous efforts of some researchers this is the only largest class of biological molecules where a systematic nomenclature system exists. As far as possible, this system has been used for the enzymes covered in this book. In general, the EC numbers are listed in the Index.

The EC list is based on a numerical classification system with classes, subclasses, sub-subclasses and individual enzyme names. Besides the systematic names, for most enzymes 'accepted names' are admitted, which are easier to use. The main arrangement of the EC-numbers is as follows:

1. Oxidoreductases (1138 classified enzymes)
   1.1 Acting on the CH-OH group of donors
      1.1.1 With NAD^+ or $NADP^+$ as acceptor
         1.1.1.1 Alcohol dehydrogenase
         1.1.1.2 Alcohol dehydrogenase ($NADP^+$)
         .....

   1.1 Acting on the CH-OH group of donors, 1.2 Acting on the aldehyde or oxo group of donors, 1.3 Acting on the CH-CH group of donors, 1.4 Acting on the $CH\text{-}NH_2$ group of donors, 1.5 Acting on the CH-NH group of donors, 1.6 Acting on NADH or NADPH, 1.7 Acting on other nitrogenous compounds as donors, 1.8 Acting on a sulfur group of donors, 1.9 Acting on a heme group of donors, 1.10 Acting on diphenols and related substances as donors, 1.11 Acting on a peroxide as acceptor, 1.12 Acting on hydrogen as donor, 1.13 Acting on single donors with O_2 as oxidant and incorporation of oxygen into the substrate (oxygenases), 1.14 Acting on paired donors, with O_2 as oxidant and incorporation or reduction of oxygen. The oxygen incorporated need not be derived from O_2, 1.15 Acting on superoxide as acceptor, 1.16 Oxidizing metal ions, 1.17 Acting on CH or CH_2 groups, 1.18 Acting on iron-sulfur proteins as donors, 1.19 Acting on reduced flavodoxin as donor, 1.20 Acting on phosphorus or arsenic in donors, 1.21 Acting on X-H and Y-H to form an X-Y bond, 1.97 Other oxidoreductases.

2. Transferases (1186 classified enzymes)
   2.1 Transferring one-carbon groups, 2.2 Transferring aldehyde or ketonic groups, 2.3 Acyltransferases, 2.4 Glycosyltransferases, 2.5 Transferring alkyl or aryl groups, other than methyl groups,

2.6 Transferring nitrogenous groups, 2.7 Transferring phosphorus-containing groups, 2.8 Transferring sulfur-containing groups, 2.9 Transferring selenium-containing groups.

3. Hydrolases (1128 classified enzymes)
   3.1 Acting on ester bonds, 3.2 Glycosidases, 3.3 Acting on ether bonds, 3.4 Acting on peptide bonds (peptidases), 3.5 Acting on carbon-nitrogen bonds other than peptide bonds, 3.6 Acting on acid anhydrides, 3.7 Acting on carbon-carbon bonds, 3.8 Acting on halide bonds, 3.9 Acting on phosphorus-nitrogen bonds, 3.10 Acting on sulfur-nitrogen bonds, 3.11 Acting on carbon-phosphorus bonds, 3.12 Acting on sulfur-sulfur bonds, 3.13 Acting on carbon-sulfur bonds.
4. Lyases (391 classified enzymes)
   4.1 Carbon-carbon lyases, 4.2 Carbon-oxygen lyases, 4.3 Carbon-nitrogen lyases, 4.4 Carbon-sulfur lyases, 4.5 Carbon-halide lyases, 4.6 Phosphorus-oxygen lyases, 4.99 Other lyases.
5. Isomerases (165 classified enzymes)
   5.1 Racemases and epimerases, 5.2 *Cis-trans* isomerases, 5.3 Intramolecular oxidoreductases, 5.4 Intramolecular transferases (mutases), 5.5 Intramolecular lyases, 5.99 Other isomerases.
6. Ligases (143 classified enzymes)
   6.1 Forming carbon-oxygen bonds, 6.2 Forming carbon-sulfur bonds, 6.3 Forming carbon-nitrogen bonds, 6.4 Forming carbon-carbon bonds, 6.5 Forming phosphoric ester bonds, 6.6 Forming nitrogen—metal bonds.

A similar system has been developed for membrane transport proteins which are classified by the Transport Commission (TC) system (see 6.1.1.2). The transport systems are classified by the mechanism of transport. The individual TC-number, e.g., 1.A.1.2.10, indicates category – class - superfamily – family – protein.

Literature:

Fersht, A. *Enzyme Structure and Mechanism*. 2nd Ed. Freeman, 1983.

International Union of Biochemistry and Molecular Biology. Nomenclature Committee. Enzyme Nomenclature, 1992. Academic Press. Website: http://www.chem.qmul.ac.uk/iubmb/enzyme.

Sigman, D.S., Boyer, P.D. (Eds.). *The Enzymes* 3rd Ed. Vols. XIX and XX. Academic Press, 1990, 1992.

(Various authors). *Methods in Enzymology*. More than 250 volumes, especially Vols. 63, 64, 87, 249. Academic Press.

2.5 Regulation of the Enzyme Activity

Since enzymes catalyze almost all metabolic reactions of living beings, it is of utmost importance that their activities be strictly regulated according to the needs of the cell or organism. Regulation of enzyme activity occurs at several levels. Characteristically, control mechanisms regulating the enzyme synthesis at the transcriptional or translational level are slow (often with response times of hours or even days), whereas rapid regulatory mechanisms act directly on the enzyme molecules.

2.5.1 Regulation of the Quantity of Enzymes

Regulation of enzyme synthesis and degradation: The biosynthesis of some enzymes is constitutive, i.e., they are formed independently of the environmental or metabolic conditions of the cell. However, for most enzymes (as well as for other proteins) the production is regulated:

- The rate of gene expression (and thus, the amount of enzyme protein synthesized) can be modulated at the transcription level.

  A well-known example is the repression of the lac operon in *E. coli* by the lac repressor (4.1.2.2), and the induction of the biosynthesis of β-galactosidase, galactose permease, and thiogalactoside transacetylase by lactose or isopropyl-thiogalactoside acting as inducers.
- Regulation of the rate of protein synthesis at the translation level, e.g., by modification of the activity of initiation factors. This occurs mostly on eukarya (4.2.4).
- The length of mRNA half-life is another parameter which influences the amount of the synthesized protein (4.2.5). It is usually very short in bacteria, but varies greatly in eukarya.
- In addition to the regulation of protein synthesis, the amount of enzyme present is also be influenced by the rate of degradation (or turnover) of the enzyme protein.

Zymogen activation: Some enzymes are synthesized as inactive precursors (zymogens or proenzymes) which have to be processed into their active form by limited proteolysis. Examples are digestive enzymes such as chymotrypsinogen or trypsinogen (Fig. 2.5-1). They are synthesized in mammalian pancreas as zymogens. After hormone-controlled release into the small intestine they are irreversibly processed by trypsin or enteropeptidase, respectively, to become active proteases (Fig. 7.1-11). Another example is given by the blood clotting enzymes.

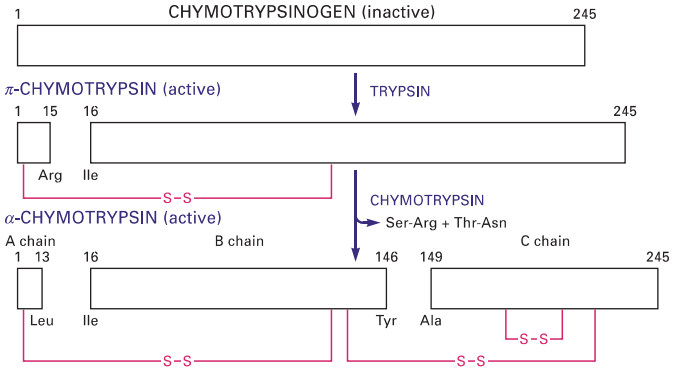


Figure 2.5-1. Conversion of Chymotrypsinogen into Chymotrypsin

2.5.2 Regulation of the Activity of Enzymes (For a Mathematical Treatment see 1.5.4)

Regulation depending on substrate concentration: In the simplest case, the dependency of the reaction rate on the substrate concentration according to the Michaelis-Menten equation (Eq. [1.5-19], Fig. 1.5-1) is a means of controlling the substrate throughput.

Since *in vivo*, the substrate concentrations frequently are in the range of the K_M value, any variation in substrate concentration will result in a corresponding change of the reaction rate (e.g., if the substrate concentration rises, the conversion will be faster). An example is glucokinase (Fig. 3.1.1-4).

Competitive inhibition (Fig. 2.5-2 a): Competitive inhibitors react reversibly with the normal substrate binding site of the enzyme to form an enzyme-inhibitor complex which prevents substrate binding and conversion to the reaction products. Therefore, competitive inhibitors are usually substances with some structural similarity to the substrate. Competitive inhibition can be reversed by increasing the substrate concentration. Formally, the inhibitor increases the apparent K_M value of the substrate, while V_{max} remains constant (Fig. 1.5-3 a).

Noncompetitive inhibition (Fig. 2.5-2 b): This reaction type is found when the inhibition cannot be overcome by increased substrate concentration. Presumably, the inhibitor binds to an enzyme locus different from the substrate binding site, both on the free enzyme and on the enzyme-substrate complex and prevents the turnover to the products. This is equivalent to a deactivation of a portion of the enzyme. Formally, the inhibitor decreases V_{max}, while the K_M value of the substrate remains constant (Fig. 1.5-3c). If, however, the inhibitor binds to the free enzyme and to the enzyme-substrate complex with different affinities, this is called a **mixed inhibition:** V_{max} is decreased, while K_M is elevated (Fig. 1.5-3d).

Uncompetitive inhibition (Fig. 2.5-2c): There are also cases where both K_M and V_{max} are affected in a different way: both V_{max} and K_M decrease. One possible reason is that the inhibitor binds only to the enzyme-substrate complex, but does not affect its formation.

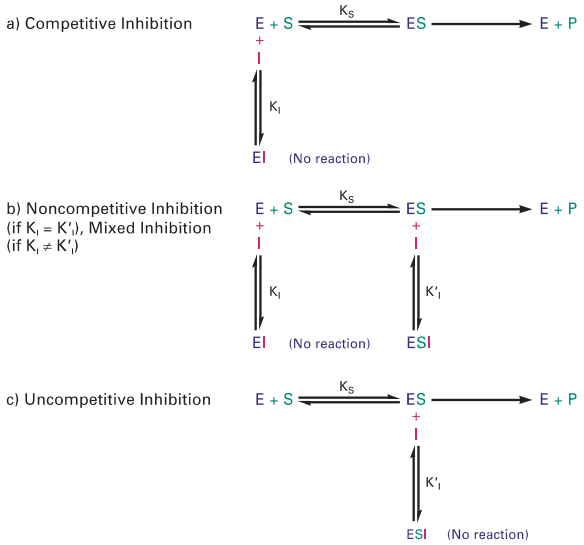


Figure 2.5-2. Different Inhibition Mechanisms of Enzyme-Catalyzed Reactions
Their influence on reaction rates shown in Figure 1.5-3.

Allosteric regulation: Many oligomeric enzymes are regulated by inhibition or activation due to interaction with an allosteric effector (an inhibitor or activator) at a binding site separate from the active center of the enzyme (negative or positive cooperative regulation). The regulator substance changes activity through the induction of conformational changes in the enzyme which may be concerted (Monod-Changeux-Wyman symmetry model) or sequential due to 'induced fit' (Koshland-Nemethy-Filmer sequential model). Allosteric regulation may also be exerted by the substrates or products of the reaction themselves (homotropic regulation). Allosteric enzymes usually do not obey Michaelis-Menten kinetics, but show sigmoid (S-shaped) substrate-saturation curves (blue curve in Fig. 2.5-3).

The theory of allosteric mechanisms actually deals with the binding of ligands (S) to oligomeric proteins, no matter whether there is a consecutive reaction or not. Therefore, this discussion applies to the oxygen binding to hemoglobin (6.3.2) as well as to the substrate binding to an enzyme. Since enzymatic reactions require the previous formation of the enzyme-substrate complex (2.4.1), allosterism is a means of regulating the reaction rate.

The protein subunits may exist in two possible states: R (relaxed) or T (tight) with different binding affinities to the ligand. In the following discussion, tetrameric proteins are assumed, which are very common among allosteric proteins.

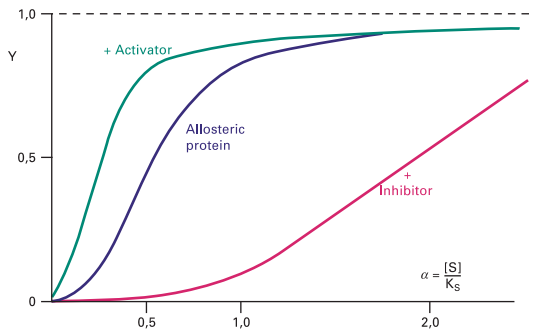


Figure 2.5-3. Saturation Curves of Allosteric Proteins; Influence of Activators and Inhibitors
The Degree of saturation Y is proportional to the reaction rate in case of allosterically regulated enzymes.

Symmetry model: In the case of homotropic regulation (the regulation is performed by the ligand S itself), the following assumptions are made:

- T- as well as R-forms can bind the ligand. However, they possess different binding affinities. The T-form has the lower affinity. Thus, low or high binding affinity only depends on whether the subunit is in the T or R-state.
- All four subunits of the tetramer have to be either in the T or in the R-state (T_4 or R_4, respectively).
- The T-state tetramers are in equilibrium with those R-state tetramers which have bound the same number of ligands.
- Due to close coupling between the subunits, all of them shift simultaneously from the T to the R-state (and vice versa, concerted allosteric transition (Fig. 2.5-4).
- The more of the 4 subunits have bound a ligand, the more likely is the common transition from the T to the R-state.

Therefore, the binding of a ligand facilitates the binding of the next ligand. This leads to a steep increase of the saturation curve at medium and high ligand concentrations, causing a sigmoid shape. The amount of cooperativity (interaction between the subunits) is expressed by the degree of upward curvature, that is, by the deviation from a simple hyperbolic curve.

If the allosteric regulation is performed by activators or inhibitors (heterotropic regulation), it is assumed that activators bind exclusively or at least preferentially to the R-state. Their action can be considered to be similar to the action of an additional substrate molecule: They promote the transition from the T to the R-state and thus enhance substrate binding. Inhibitors, however, bind to the T-state. They decrease the tendency of transition from the T to the R-state and thus antagonize the effect of bound substrate on the binding of additional substrate.

Therefore, activators shift the saturation curve to the left (to lower substrate concentrations) and inhibitors to the right (to higher substrate

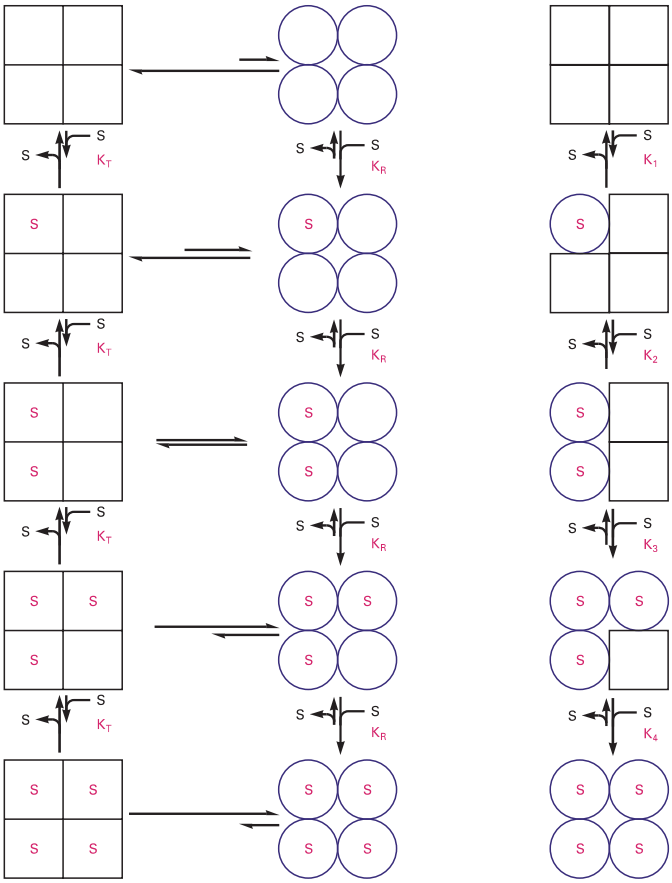


Figure 2.5-4. Interconversions between T and R-states According to the Symmetry Model (Left) and to the Sequential Model (Right) of Allosteric Reactions

concentrations, Fig. 2.5-3). The same effect can be seen with the reaction velocity plots of allosterically controlled enzyme reactions.

The efficiencies of allosteric inhibitors (I) or activators (A) depend (among other factors) on their concentrations in relation to their dissociation constants, $\beta = [I]/K_I$ and $\gamma = [A]/K_A$, while the concentration of the ligand is expressed by $\alpha = [S]/K_S$. If L is the ratio of T and R-states in absence of ligand, $L = [T_0] / [R_0]$, then the degree of saturation Y of an n-meric protein (for a tetrameric protein n = 4) by the ligand is expressed by the formula

$$Y_s = \frac{\alpha * (1 + \alpha)^{n-1}}{(1 + \alpha)^n + \dfrac{L * (1 + \beta)^n}{(1 + \gamma)^n}}$$

In Figure 2.5-3, the respective data are: n = 4, L = 1000, for the inhibited curve $\beta = 2.5$, $\gamma = 0$, for the activated curve $\gamma = 100$, $\beta = 0$.

Sequential model: This model also assumes T and R-states of the allosteric protein, but with important differences:

- Only complexes of the ligand L with the R-state (RL) exist.
- Hybrids of T and R-states are possible [e.g., $T_3(RL)$, $T_2(RL)_2$, $T(RL)_3$].
- Ligand binding induces the conformational change of the particular subunit ('induced fit') from the T to the R-state.
- The transition of a subunit from the T to the R-state leads to a cooperative interaction with the neighboring subunit, which facilitates its ligand binding and thus its transition from the T to the R-state (Fig. 2.5-4).
- The more subunits are already in the R-state, the more interactions with the remaining T-state subunits take place. This gradually increases the ease of their transition into the R-state.

The strength of the interaction between the subunits determines the degree of cooperativity. In case of very strong coupling, the transition from T to R-states for all 4 subunits takes place almost simultaneously, which corresponds to the assumptions of the symmetry model.

Although both models can be fairly well adapted to many actual situations, they represent two contrasting assumptions. In general, the reactions of allosteric proteins contain elements of both models.

Covalent modification: Another method of reversible regulation of enzyme activities involves covalent modifications by attachment or removal of special groups. Two possibilities exist:

- The enzyme is normally inactive, but is interconverted to an active form by the addition of a group. Examples are
  - phosphorylation (e.g., glycogen phosphorylase, 3.1.2.4)
  - adenylylation (e.g., *E. coli* glutamine synthetase, 3.2.2.1)
  - NAD ribosylation (7.4.1).
- The enzyme is normally active, but is inactivated by the addition of a group.

  An example is the phosphorylation of glycogen synthase (3.1.2.4).

These modification reactions are catalyzed by more or less specific, regulatory enzymes which themselves may be subject to various control mechanisms.

There are cases where enzymes which catalyze the interconversion of two compounds in different directions are both regulated by such covalent modifications.

An example is the formation and degradation of glycogen in animals. (Fig. 2.5-5. This figure is simplified; more details are given in Fig. 3.1.2-5.) Glycogen degradation is effected by phosphorylase. This enzyme is activated by phosphorylation in a reaction catalyzed by phosphorylase kinase (change from phosphorylase b into phosphorylase a). Glycogen synthase, which effects glycogen formation, is deactivated by phosphorylation. This reaction is also catalyzed by phosphorylase kinase. Phosphorylase kinase, in turn, is activated in a multistep reaction by hormones (glucagon, epinephrine).

If the metabolic situation changes, the regulation system is reversed. The phosphorylated glycogen synthase is dephosphorylated and thus activated by protein phosphatase 1. The phosphorylated phosphorylase a is dephosphorylated and deactivated by the same enzyme protein phosphatase 1, which is controlled by the hormone insulin in a multistep mechanism.

Thus each of the regulating enzymes, phosphorylase kinase and protein phosphatase 1, influence both glycogen synthesis and degradation, albeit in opposite directions.

Cascade mechanisms: The effect of a regulated change in the activity of an enzyme is sometimes amplified through a cascade mechanism: the first enzyme activates a second enzyme, this, in turn, a third one and so on. This causes a potentiation of the regulatory effect and very rapidly produces a large amount of the active form of the last enzyme. Activation may proceed via proteolysis (e.g., in blood coagulation, 9.3) or by phosphorylation/dephosphorylation mechanisms (e.g., in the regulation of glycogen synthesis and degradation mentioned above, regarding the intermediate steps between hormone action and activation of phosphorylase kinase and protein phosphatase, Fig. 3.1.2-5).

2.5.3 Site of Regulation

In multistep metabolic pathways, the product of a late (or the last) step frequently acts as an inhibitor of the first committed step in this pathway ('feedback inhibition'). This way, the end product of a pathway controls its own synthesis and prevents useless accumulation of intermediates. All of the regulation mechanisms mentioned above can be applied.

Examples are the regulation of glycolysis (Fig. 3.1.1-3), pyruvate dehydrogenase (Fig. 3.1.3-2), citrate cycle (Fig. 3.1.8-3), fatty acid synthesis (Fig. 3.4.1-5), cholesterol synthesis (Fig. 3.5.1-7), bile acid synthesis (Fig. 3.5.9-3), the repression of genes coding for the enzymes of lactose and tryptophan synthesis (Figs. 4.1.2-2 and 4.1.2-5) etc.

Since many metabolic pathways are branched, complex regulatory patterns such as multiple inhibition, sequential inhibition, or the presence of isoenzymes with different regulatory properties can be observed.

Examples for such mechanisms are, e.g., the regulation of the threonine, methionine and lysine biosynthesis (Fig. 3.2.5-1), aromatic amino acid biosynthesis (Fig. 3.2.7-2), purine nucleotide biosynthesis (Fig. 3.6.1-5), ribonucleoside-diphosphate reductase (Fig. 3.6.1-7), etc.

Literature:

Lipscomb, W.N. Chemtracts-Biochemistry and Mol. Biol. 1990;2:1–15.

(Various authors). *Methods in Enzymology.* More than 250 volumes, especially Vol. 63, p. 383-515, Vol. 64, p. 139-192. Academic Press, 1979/1980.

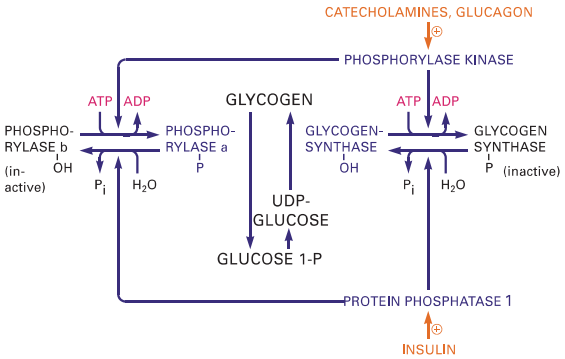


Figure 2.5-5. Regulation of Glycogen Synthesis and Degradation (Simplified)

2.6 Nucleic Acid Structure

2.6.1 Components of Nucleic Acids

Nucleic acids are sequences of nucleotides (3.6) interconnected by phosphate diester bonds. Thus they contain two structural features: the information-carrying bases and a sugar-phosphate 'backbone'.

In ribonucleic acid (RNA), the sugar is ribose, in deoxyribonucleic acid (DNA), it is deoxyribose (Fig. 2.6-1).

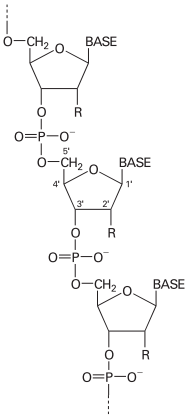


Figure 2.6-1. General Structure of Nucleic Acids
(R= –OH in RNA, R = H in DNA)

The sugar residues of the backbone are linked at their 5′ and 3′ positions by a phosphate diester bridge. This gives the nucleic acid strand a distinct polarity (5′ and 3′ end). Nucleic acids form unbranched chains of greatly varying length (from a few building blocks up to millions).

The bases are purines (adenine, guanine, 3.6.1) or pyrimidines (cytidine, thymine in DNA, cytidine, uracil in RNA, 3.6.2). They can interact via hydrogen bonds with each other according to distinct rules (Watson-Crick pairing rules, Fig. 2.6-2):

- adenine only pairs with thymine (in DNA) or uracil (in RNA), interacting via 2 hydrogen bonds
- guanine only pairs with cytidine, interacting via 3 hydrogen bonds.

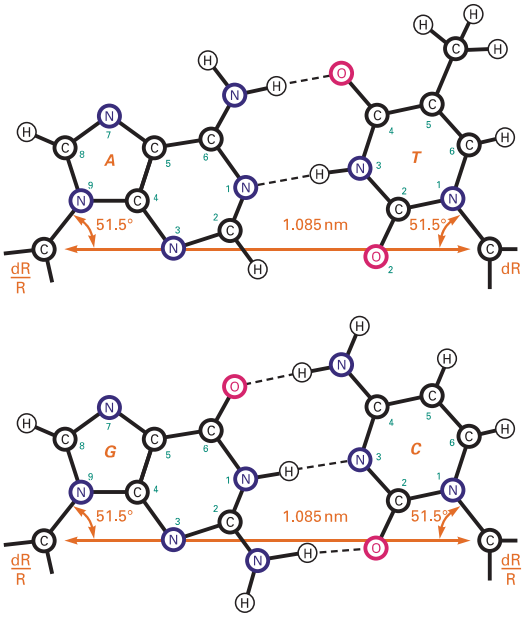


Figure 2.6-2. Watson-Crick Base Pairing Rules
(Double bonds not shown)

Since the distances of the bases and the angles of the base-sugar bonds are exactly equal for both Watson-Crick pairs, an exchange of bases does not affect the backbone structure of the double helix (see below).

In some cases, non-Watson-Crick base pairs occur, e.g., of the Hoogsteen type (Fig. 2.6-3). They are found, e.g., in tRNA (4.1.1.2, 4.2.1.1) or in ‘wobble base pairing’ during transcription (2.7.2).

The bases are the elements of the information carried by nucleic acids (analogously to ‘bits’ in computer science), but they are also indispensable for the structural and catalytic features of nucleic acids.

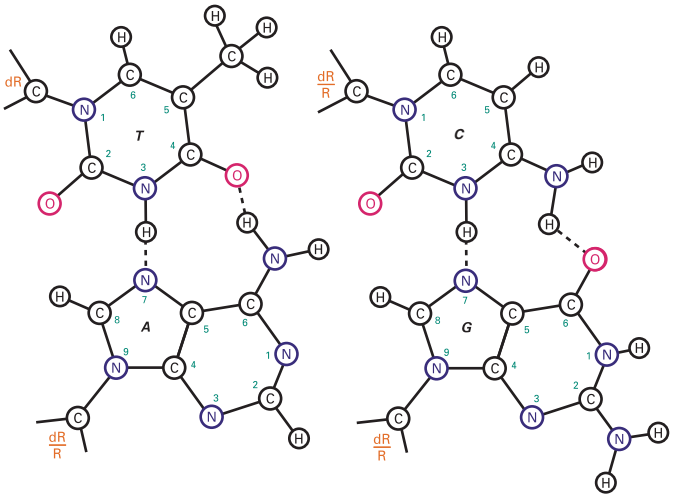


Figure 2.6-3. Non-Watson-Crick Base Paires
(Hoogsteen Pairs; double bonds not shown)

DNA serves almost exclusively for information storage. RNA molecules have many different functions (4.1.1, 4.1.3, 4.2.1 4.2.2):

- structural and catalytic roles in ribosomes (rRNA)
- information transfer (mRNA)
- translation of the nucleic acid information into the protein information (tRNA)
- control of gene expression (miRNA, siRNA, snoRNA, snRNA)
- catalytic roles during posttranscriptional processing of other RNAs (as ‘ribozymes’, e.g., in spliceosomes)
- information storage in RNA viruses (5.3, 5.4)

2.6.2 Properties of RNA Chains

Likely due to the presence of the 2′-OH group, steric inhibition prevents RNAs from assuming a DNA-like B-type double helix conformation (see below). They are mainly single-stranded with intrastrand base-pairing forming stems, loops etc., which resemble the less favored A-type DNA helices (see below). The same A-type helix structure is assumed by the short DNA-RNA hybrids, which are formed during transcription (4.1.1, 4.2.1) initiation of DNA replication (3.8.1, 3.9.1) and reverse transcription of retrovirus RNA (5.4.1).

Depending on their different tasks, the properties of RNA molecules vary greatly:

- tRNAs have a cloverleaf arrangement of loops, which are folded into an ‘L’ shape (Fig. 4.1.3.1). They consist of 60 … 100 nucleotides.
- mRNAs are mainly linear, but hairpin loops have functional properties. Examples are the termination of transcription (Fig. 4.1.2-5) or the regulation of gene translation, e.g., by binding of regulatory proteins (4.2.4) or in polycistronic mRNA. mRNAs vary greatly in size.
- rRNAs of several sizes occur in ribosomes (Tables 4.1.3-1 and 4.2.3-1). They contain a large number of stem loops. Although variations in the base sequences have occurred, many aspects of the secondary structure have been conserved during evolution.

2.6.3 Properties of DNA Chains

DNA occurs mainly as a double helix of two antiparallel strands with complementarily paired bases. The polar sugar-phosphate backbone is located on the outside and the hydrophobic bases are stacked in the central area with their planes parallel to each other, almost perpendicular to the long axis of the helix. This stacking allows hydrophobic interactions, which are responsible for the stability of the double helix in the first place. The helix is usually ‘right-handed’ (i.e., when looking along the axis of the helix, the strands are coiled clockwise) with

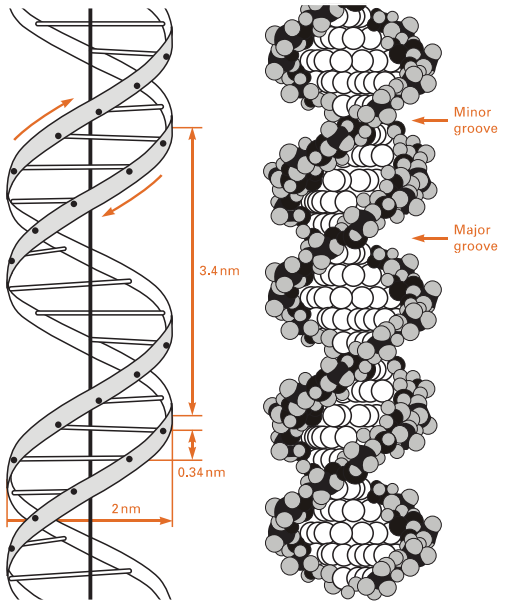


Figure 2.6-4. Structure of the B-Type DNA Double Helix

a major and a minor groove. Both grooves are wide enough to allow proteins to come into contact with the bases (Fig. 2.6-4).

The DNA double helix is plectonemically coiled, i.e. the helices can only be separated by unwinding the coils.

Several helix forms have been described (Table 2.6-1). Some interconversions of these helix forms are possible depending on the concentration and types of salts present. The natural form of DNA is the B type helix, which is also the most prevalent form in solution. In this helix, the major and minor grooves are most pronounced (2.2 nm and 1.2 nm wide, respectively). Adjacent base pairs are rotated by 34.5 to 35.5°. This twist angle can change depending on the sequence which may result in kinking of the double helix. This kinking can also be caused by other properties of the DNA or by proteins. The twist can be superimposed on the double helix, resulting in supercoiling. A helix can become positively (twist in the direction of winding) or negatively supercoiled. Supercoiling results in a more compact structure of DNA. This is very important in DNA packaging.

Table 2.6-1. Different Forms of the DNA Double Helix

	B Type	A Type	Z Type
Diameter (nm)	ca. 2	ca. 2.6	ca. 1.8
Turning mode	right handed	right handed	left handed
Base pairs/turn	10.4 … 10.65	11	12
Pitch/base pair (nm)	0.34	0.26	0.37
Pitch/turn (nm)	3.4	2.8	4.5
Occurrence	most common	bacterial spores, DNA-RNA hybrids	in alternating purine-pyrimidine sequences, during torsional stress or when stabilized by supercoiling, proteins, methylation etc.

Torsional stress due to supercoiling can be overcome by the formation of DNA structures other than the B-form. E.g., negative supercoiling is a strong driving force for the stabilization of Z-DNA. During transcription, positive supercoils are formed in front of the transcription apparatus and negative supercoils behind it. These supercoils are controlled by enzymes that wind and unwind the DNA double helix (3.9.1.4).

2.6.4 Compaction Levels of DNA Chains

The genetic material of all organisms and viruses exists in the form of tightly packaged nucleoprotein. A high degree of compaction is necessary in order to store the long DNA molecules within cells.

The circular DNA of *E. coli* is ca. 1600 µm long and has to be placed into a cell of 1 … 3 µm length. In humans, the diploid DNA has a linear equivalent (contour length, sum of all 46 chromosomes) of 2 m, which has to be packed into a nucleus of 10 µm diameter. The DNA of the individual chromosomes has contour lengths of 16 … 82 mm.

Bacterial genomes: They contain several DNA binding proteins (20 % by mass), some of which are small and highly basic. They condense the DNA and wrap it into a bead-like structure. This nucleoid is kept together by RNA and protein which forms the core of the condensed chromosome. The rest of the DNA exists as a series of highly twisted or supercoiled loops.

Eukaryotic genomes: They are arranged in several chromosomes which contain chromatin composed of

- a double helix of DNA (e.g., 51 * 10^6 … 245 * 10^6 bp in the various human chromosomes)
- protein (about the same quantity by weight as DNA):
  - histones (small basic proteins, most abundant, see below)
  - nuclear scaffold proteins (acidic non-histone proteins, see below)
  - some enzymes (DNA polymerases, DNA topoisomerase II etc.),
- nuclear RNA. Much of the RNA in chromatin (<10 % of DNA mass) consists of nascent chains still associated with the template DNA.
- a small lipid fraction

Since DNA cannot be directly packed into the final dense structure of chromatin, a hierarchy of organization levels is necessary (Fig. 2.6-6, Table 2.6-2).

Table 2.6-2. Condensation Levels of Eukaryotic DNA

Level	Total base pairs/unit	Size (nm)	Condensation ratio
DNA	10.5/turn	2.6	1
Nucleosome filament	150 … 240/nucleosome	10 (diameter)	7
Chromatin filament (solenoid)	900 … 1500/turn	30 (diameter)	35 … 50
Looped Domains	20,000 … 100,000/loop	150 … 300 (length of loop)	1,700 … 2,500
Only during mitosis: Miniband	360,000 … 1,800,000	840	10,000 … 12,000
Chromosome arm		840 (diameter)	

Specific proteins contribute to these steps:

- Histones are small basic proteins (11 … 21 kDa), which are involved in the first packaging level of DNA. In most eukarya there are 5 types: H1, H2A, H2B, H3 and H4. They are rich in the positively charged amino acids lysine and arginine, binding tightly to the negatively charged DNA. The amino acid sequences of histones have been extremely conserved during evolution (e.g., in histone H4 only 2 out of 102 amino acids differ between pea seedlings and calf thymus, and even these exchanges are conserved, Val→Ile and Lys→Arg). There are many posttranslational modifications (methylations, acetylations, phosphorylations). Some of them are reversible and seem to be connected to the cell cycle.
- For the higher compaction mechanisms, a set of highly conserved non-histone (ribonucleo-)proteins, the so-called nuclear scaffold proteins (or nuclear matrix) are extremely important. Their organization and function are still not fully understood.

  The nuclear scaffold proteins provide a three-dimensional structural framework for DNA, and even seem to contribute to tissue-specific gene expression. They also contain topoisomerases, which allow unwinding of DNA supercoils. Eukaryotic DNA has many scaffold-attached regions (SARs) or matrix-associated regions (MARs) of 200 bp length, mostly A/T rich and containing special sequences like topoisomerase cleavage sites.

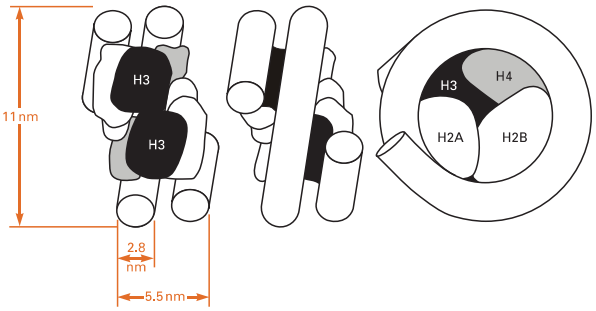


Figure 2.6-5. Structure of the Nucleosome Core

Organization levels: The fundamental organizational unit of the eukaryotic chromosome is a histone-DNA complex, the nucleosome (Fig. 2.6-5). Folding of chromosomal DNA into core nucleosomes results in a sevenfold compaction in length. The nucleosomes are linked by a short stretch of 'linker' DNA (normally 30 … 60 nt). Histone H1 is bound to this region. The binding of histone H1 increases the supercoiling of DNA and plays a major role in higher order structure and in chromatin condensation. A continuous string of nucleosomes forms a 10 nm filament.

At the center of this nucleosome core particle, there is an octameric complex of histone proteins: a central tetramer composed of two molecules each of H3 and H4 is flanked on either side by a dimer of H2A/H2B. This core self-assembles in the presence of the protein nucleoplasmin. It resembles a short cylinder. It binds 146 bp of DNA into 1.65 left-handed turns around the outside, followed by linker DNA. Thus, up to 240 bp of DNA are organized per nucleosome. In the nucleated erythrocytes of birds, fish and amphibians the histone H1 variant H5 can take the role of H1 when the chromatin is inactive. This seems to be necessary for very dense packaging.

At the next higher level of compaction the 10 nm fiber of nucleosomes forms a left-handed hollow helix of 6 nucleosomes per turn, the 30 nm chromatin filaments. They have the shape of a solenoid with an 11 nm pitch per turn. This aggregation is apparently effected by the histone H1 molecules, which polymerize and form a band in the center of the helix. The degree of compaction is between 35 and 50.

The solenoids are compacted further to form looped DNA domains. Each loop is 150 … 300 nm long and contains about 50 solenoid turns. Loop domains are thought to be the basic unit of higher order DNA structure in all eukaryotic cells. In the interphase nucleus (Fig. 4.3-5), this is the maximum compaction level of euchromatin, which represents the transcriptionally active part of DNA. The loops are attached to the nuclear matrix (netlike nuclear lamina) at sites where origins of replication exist. Active genes seem to be located close to these regions. An interphase nucleus contains ca. 50,000 loop domains.

The most highly condensed eukaryotic DNA known is in the mammalian sperm nucleus, which is about six times more condensed than the mitotic chromosome (see below). The packaging system differs from normal cells (e.g., histones are replaced by protamines).

Mitosis (4.3.5): The degree of compaction varies throughout the cell cycle. In the course of mitosis the loop domains of euchromatin are further packaged to form a 250 nm fiber, which coils to form the arms of a metaphase chromosome. One layer of the coil (diameter ca. 840 nm) consists of 18 loops and is also called a miniband. The packaging ratio (length of DNA/length of the unit containing it) is now up to 10,000 … 12,000.

Centromeres and telomeres are two special regions of eukaryotic chromosomes. Centromeres are *cis*-acting genetic loci, which are essential for proper segregation of chromosomes during mitosis and meiosis and are made up of very large regions of repeated sequences. Telomeres are specialized DNA-protein complexes that form the ends of linear chromosomes (3.9.1.5). They are important for keeping the integrity of individual chromosomes.

All of the vital processes of DNA metabolism must deal with the topological complexity of the chromosome. The chromatin structure is involved in the regulation of replication, transcription, repair and recombination. The topological organization of DNA is both cell and tissue-specific and DNA can take many forms of higher order structure necessary for the expression of only the appropriate tissue-specific genes.

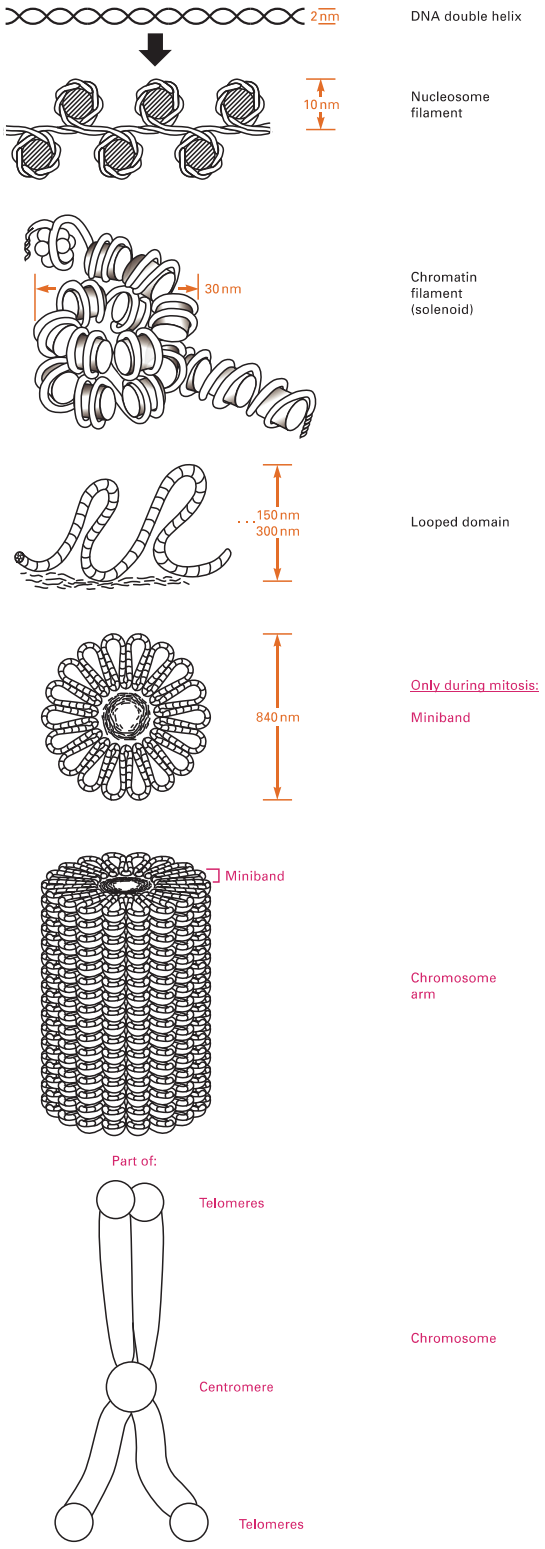


Figure 2.6-6. Organization Levels of Eukaryotic Chromosomes

It should be kept in mind that many of the compaction mechanisms described above take place at the same time in different parts of the same DNA molecule. Simultaneously, this DNA molecule is involved in many processes, such as transcription, replication and repair (4.1.1, 3.9, 4.2). It is quite impressive how such widely divergent actions can result in almost perfectly ordered processes.

Literature:

Auffinger, L. *et al*. J. Mol. Biol. 1997;273:54–62.

Benbow, R. M. Sci. Progress Oxford 1992;76:425–450.

Dickerson, R. Nucl. Acids Res. 1998;26:1906.

Getzenberg, R. H. *et al.* J. Cell. Biochem. 1991;47:289–299.

Gutell, R.R. *et al*. Progr. Nucl. Acid Res. Mol. Biol. 1985;32:155–216.

Luger, K. *et al.* Nature 1997;389:251–260.

Smith, M. M.: Curr. Opinion in Cell Biology 1991;3:429–437.

2.7 Genetic Code and the Flow of Information

This section deals with the general principles of protein synthesis in all kingdoms of biology (archaea, bacteria and eukarya).

2.7.1 From DNA to RNA

Deoxyribonucleic acid stores the genetic information encoded into a sequence of deoxyribonucleotides (A,C,G,T). Both strands of the double stranded DNA are linked by specific base pairing according to the Watson-Crick rules (A=T and G≡C, 2.6.1). Thus, each strand contains the complete genetic information. Specific base pairing is not only a safety measure to enable repair of damages (3.8.2 and 3.9.2), but also the tool to transmit the information during synthesis of more DNA (replication, 3.8.1 and 3.9.1) or of ribonucleic acid (transcription, 4.1.1 and 4.2.1).

During the transcription process, selected sequences of the DNA are copied into messenger RNA (mRNA). Other DNA sequences are transcribed into RNA with other functions. One of the two DNA strands acts as a template; the other (not involved in this procedure) carries the same information sequence as the synthesized RNA. Thus, the latter one is called the 'coding', 'sense' or (+) strand, while the template DNA strand has the sequence of a complementary antisense (–). In RNA, thymidine nucleotides are replaced by uridine nucleotides.

2.7.2 From Nucleic Acids to Proteins – The Genetic Code

The information for each polypeptide stored in a DNA sequence is named a gene (one gene – one polypeptide hypothesis). After transcription of this DNA sequence into mRNA, the information has to be translated into an amino acid sequence during protein synthesis (translation). A group of three consecutive nucleotides (codon or triplet) represents one amino acid. The compilation of these relationships is the genetic code (Fig. 2.7-3).

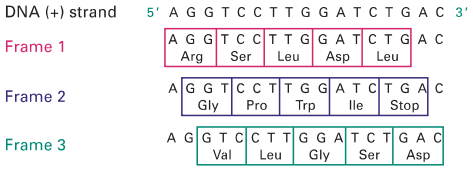


Figure 2.7-1. Example of Different Reading Frames

There are 20 'classical' amino acids, for which codons exist. There are two 'special' amino acids: Selenocysteine (Sec) uses the codon UGA and pyrrolysine (Pyl) the codon UAG. Both codons are otherwise stop codons (4.1.3.4). The triplet codons follow each other in 5′→3′ direction without interruption.

Therefore, when starting from a different nucleotide, one gets completely different triplets coding for different amino acids. This requires a clearly recognizable starting point for translation. On the other hand, the same nucleotide sequence can code for more than one polypeptide, e.g., in some viruses (frameshift, e.g., 2.7-1). The other DNA strand can also be used as a template; it is read in the opposite direction (e.g., in some viruses).

The translation of the nucleic acid information into the amino acid sequence of proteins needs a 'translator' or 'adaptor' system, which is competent in both 'languages'. For each amino acid there are one or more tRNA molecules that can read the nucleic acid language via base pairing. Their specificity towards amino acids is effected by specific enzymes charging them with the cognate amino acid. Thus, the flow of information is as shown in Figure 2.7-2 (only reverse transcriptase is an exception, 5.4).

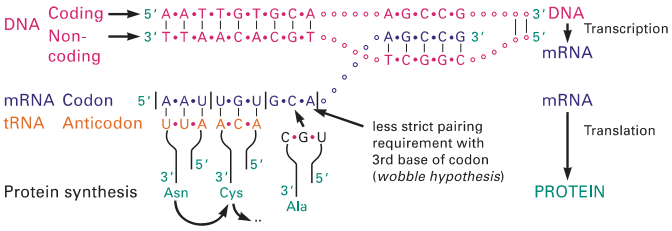


Figure 2.7-2. Information Transfer During Protein Synthesis

Since there are $4^3 = 64$ possible nucleotide triplets, but only 22 amino acids (including Sec and Pyl) plus some stop codons (see below), most amino acids are encoded by several nucleotide triplets (synonyms, at maximum 6 in the cases of Arg, Leu and Ser). Most variants occur in the 3rd position of the code ('wobble hypothesis'). The code is therefore called 'degenerate' or 'redundant'. The relationship between triplets and encoded amino acids is shown in Figure 2.7-3.

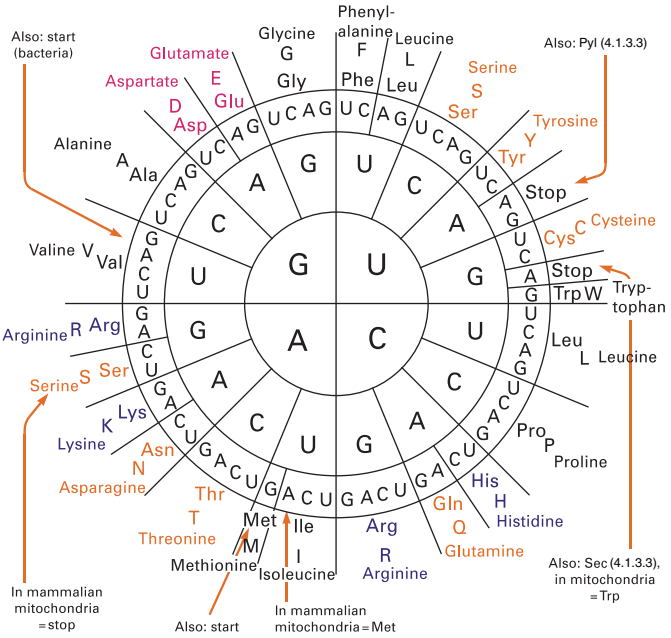


Figure 2.7-3. Genetic Code

The triplet sequence is read from the center outwards. The mRNA nucleotide terminology is shown. For DNA nucleotide sequences, replace U with T. The amino acids are given with their full names and with their three- and one-letter abbreviations. Basic amino acids are shown in blue, acidic amino acids in red, neutral amino acids in black and amino acids with uncharged, polar residues in orange. For details see 1.3.1.

The triplet sequence is read from the center outwards. The mRNA nucleotide terminology is shown. For DNA nucleotide sequences, replace U with T. The amino acids are given with their full names and with their three- and one-letter abbreviations. Basic amino acids are shown in blue, acidic amino acids in red, neutral amino acids in black and amino acids with uncharged, polar residues in orange. For details see 1.3.1.

Although the code has been assumed to be universal among living beings, several exceptions have been found, for example in mitochondria (Table 2.7-1). Furthermore, in mycoplasma, UGA codes for Trp, while in some ciliated protozoa, the normal 'stop' codons UAG and UAA specify Gln and UGA specifies Cys instead. Yet another exception exists in some *Trypanosoma*, *Paramecia* and *Tetrahymena*, where the 'stop'-codon UGA encodes Gln (Table 2.7-2).

Table 2.7-1 Codon Differences in Mitochondria

	UGA	AUA	CU(A,C,G,U)	AG(A,G)	CGG
General code	stop/Sec	Ile	Leu	Arg	Arg
Mitochondria					
Mammals	Trp	Met/start		stop	
Drosophila	Trp	Met/start		Ser (AGA only)	
Protozoans	Trp				
Higher plants					Trp
S. cerevisiae	Trp	Met/start	Thr		

Table 2.7-2 Codons with Special Functions

	Start codons	Stop codons
General code	AUG (codes also for Met)	UAG, UAA, UGA('amber, ochre, opal')
Eukarya	AUG (codes also for Met), CUG (rare), ACG (rare), GUG (rare)	UAG, UAA, UGA
Bacteria	AUG, GUG, UUG (rare)	UAG, UAA, UGA
Mitochondria	AUA (codes also for Met), AUG	AGA, AGG (mammals)

2.7.3 Influence of Errors

Unrepaired changes in the DNA sequences (e.g., mutations) lead to variations in the expressed proteins, frequently with dramatic effects on the organisms and their progeny. Therefore, many repair systems exist to remove these damages (3.8.2, 3.9.2).

- Insertions or deletions of single nucleotides cause frameshifts (2.7.2). From this site on, the translated protein exhibits a different amino acid sequence. Also, stop codons might be either newly formed or removed.
- Modifications of single nucleotides influence only the particular triplet and thus may change a single amino acid (missense or nonsense mutations). However, since many of the amino acids are encoded by several triplets (degenerate code, see above), such changes do not always take place (silent mutations).

Literature:

Crick, F.H. J. Mol Biol. 1968;38:367–79.

Khorana, H.G. in: *Nobel Lectures in Molecular Biology* 1933–1975. Elsevier, 1977; p. 303–331.

Knight, R. D., Freeland, S. J. and Landweber, L. F. Nat Rev Genet 2001;2:49–58.

Nirenberg, M., in: *Nobel Lectures in Molecular Biology* 1933–1975. Elsevier, 1977; p. 335–360.

2.8 Polymeric Carbohydrates

The basic chemistry and structure of carbohydrates has been discussed in chapter 1.2. This chapter will give a short overview on the cellular and extracellular role of polymeric carbohydrates. Actually, most of the carbohydrates in nature occur in the form of polysaccharides, also called glycans. These are homo- or heteropolymers of medium to high molecular weight. They are either storage forms of monosaccharides and in this role serve as energy reservoirs or are structural elements of plants, animals or bacteria. An overview of the different structures and roles of carbohydrates is given in table 2.8-1.

2.8.1 Polymeric Carbohydrates in Energy Storage

The most important forms of energy-storing polysaccharides are starch in plants and glycogen in animals and bacteria. They form an energy reserve that is less compact and contains less energy per weight than the triglycerides, but they can be quickly mobilized when glucose is needed.

Starch contains in general two different forms of glycan, 10–30 % amylose and 70–90 % amylopectin. Both are homopolymers of glucose that are connected via α1→4 glycosidic bonds. Whereas amylose is a linear polymer, forming a helical structure, amylopectin is branched every 24-30 residues by α1→6 glycosidic linkages (Figs. 2.8-1, 2.8-2 and 2.8-3).

Glycogen, the energy-storage polysaccharide found in animals is very similar to amylopectin in structure but is branched every 8-12 glucose residues. In human liver hepatocytes, glycogen can compose up to 8 weight-% after a meal. Only the 100–120 g glycogen stored in the liver can be made accessible to other parts of the body.

Under some conditions, bacteria accumulate relatively large amounts of polyglucose compounds with properties similar to those of animal glycogen.

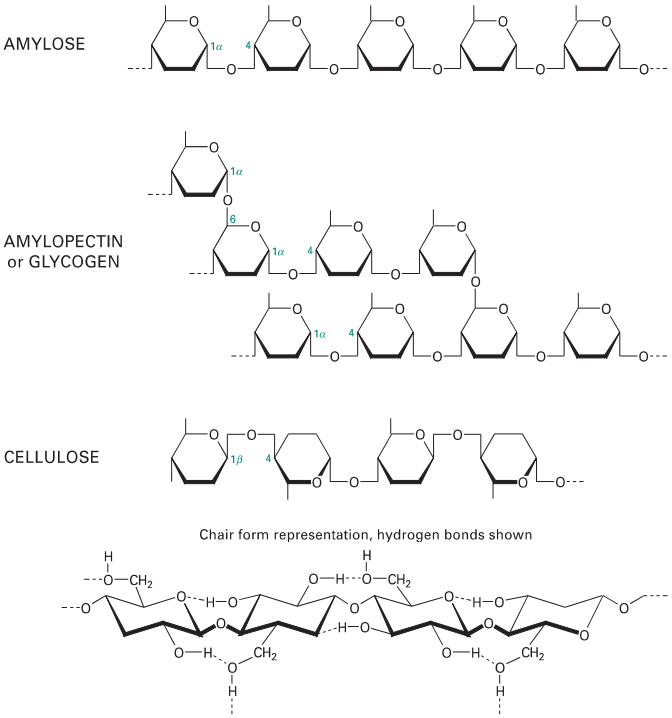


Figure 2.8-1. Chemical Structure of Amylose, Amylopectin Glycogen and Cellulose

Table 2.8-1. Structures and Roles of Polymeric Carbohydrates

Name	Polymerisation	Connection	Number of monomers	Biological Role
Amylose	Homopolymer	Linear α1→4 Glc	50-5000	Energy storage in plants
Amylopectin	Homopolymer	Linear α1→4 Glc, with α1→6 branches every 24-30 residues	Up to 1,000,000	Energy storage in plants
Glycogen	Homopolymer	Linear α1→4 Glc, with α1→6 branches every 8-12 residues	Up to 50,000	Energy storage in animal and bacterial cells
Cellulose	Homopolymer	β1→4 Glc	Up to 15,000	Structure element of plants
Chitin	Homopolymer	β1→4 GlcNAc	Large	Structure elements in a number of animal families.
Gycosaminoglycans	Acidic Heteropolymers	Different	Up to 100,000	Extracellular matrix of connective tissue and skin, joints, etc.

2.8.2 Polymeric Carbohydrates as Structural Elements

Cellulose is by far the most common biological compound. More than 50 % of the carbon in the biosphere is bound in cellulose. More than 10^{12} tons of cellulose per year are produced by plants and degraded mostly by fungi and bacteria. It gives strength to the cell wall of plants and can resist a pressure of up to 20 bar. Cellulose consists of long chains of glucose monomers connected by β(1→4) glycosidic bonds. Thus, every other glucose unit is "rotated". Due to this arrangement, the chains become interconnected by hydrogen bonds resulting in layers, which in turn are stacked by hydrophobic interactions between the layers (Figs. 2.8-1 and 2.8-4).

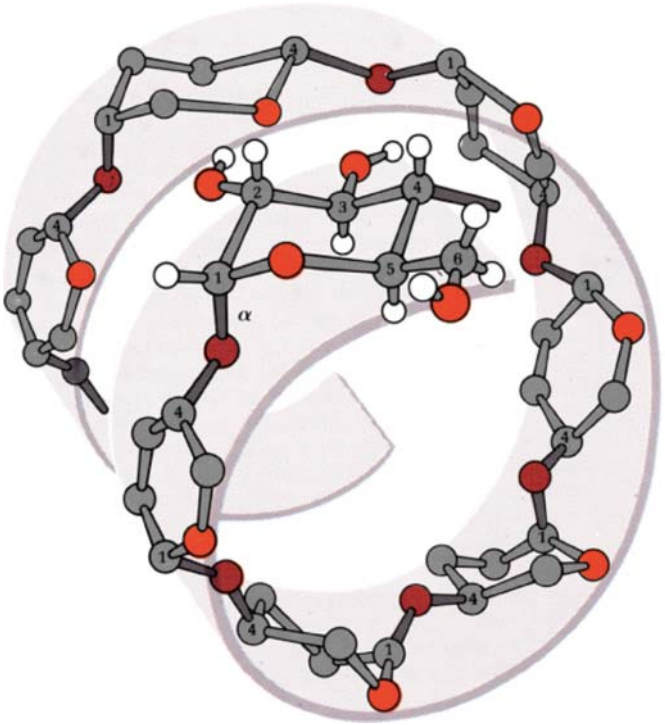


Figure 2.8-2. Model of the 3D Structure of Amylose
After Voet, D. and J.G, Pratt, C.W.: Fundamentals of Biochemistry. John Wiley & Sons 2002, Figure 8-10.

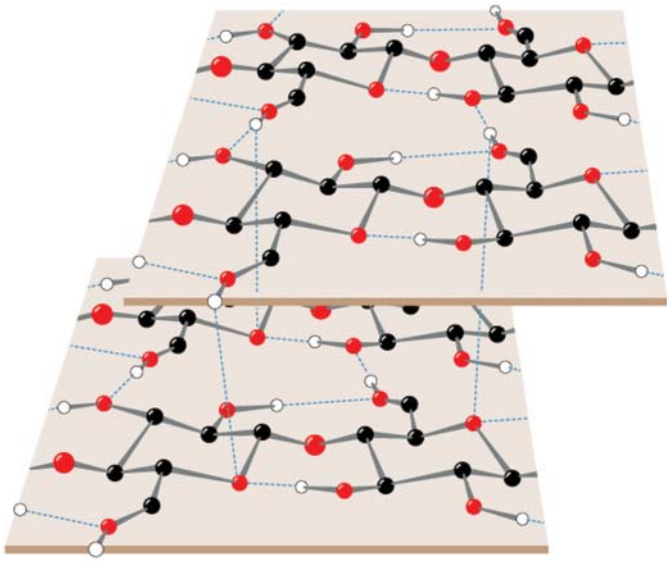


Figure 2.8-4. Model of the 3D Structure of Cellulose
After Voet, D. and J.G, Pratt, C.W.: Fundamentals of Biochemistry. John Wiley & Sons 2002, Figure 8-9.

Chitin has a very similar structure and function as cellulose. It is the main component of the cell walls of fungi, the exoskeletons of arthropods such as crustaceans and insects, the radulas of mollusks and the beaks of cephalopods, including squid and octopuses. It is the second most common polysacharide, consisting of long chains of N-acetylglucosamine linked by β-1,4-glycosidic bonds. Thus it differs from cellulose only in the replacement of the 2-OH groups by acetamido groups.

2.9 Glycosylated Proteins and Peptides

Glycosylated derivatives of proteins and peptides play a major role in cellular metabolism and recognition (Table 2.9-1). Many of them are constituents of cell membranes or even structural elements of cell walls, while others are present in the intracellular or extracellular space.

Table 2.9-1 Classification of Glycosylated Proteins

	Glycoproteins	Proteoglycans	Peptidoglycans
Occurrence	Animals, plants and fungi	Animals	Bacteria
Saccharide chain	Short oligosaccharide chains of different units (frequently branched, frequently terminating with sialic acid, n = 2…20) are covalently linked to protein (N- and O-glycosylation, 4.4).	Long sequences of many repeating disaccharide units containing amino sugars and frequently uronic acids (average of n = 80, called glycosaminoglycans or mucopolysaccharides) are bound to the protein through 'core' carbohydrate structures analogous to glycoproteins	Long unbranched chains of repeating disaccharide units consisting of N-acetylglucosamine and acetylmuramic acid
Peptide chain	Many variations	Core protein, up to 3 * 10^5 Da, forming together with the attached sacccharide chains a 'bottle brush' structure; additional link proteins. These structures are bound noncovalenty to hyaluronic acid.	Short chains of amino acids interconnecting the saccharide chains.
Total mol. mass	Varying, mostly < 3 * 10^5 Da	2 * 10^3 … > 10^7 Da	Very large
Examples	Almost all exported proteins (e.g., secreted immunoglobulins and blood group substances, transport proteins, proteohormones), mucin, Cell membrane components (e.g., MHC compounds, receptors, bound immunoglobulins, fibronectin), many cytoplasmic enzymes and other proteins.	Constituents of the extracellular matrix, e.g., chondroitin sulfate, dermatan sulfate, heparan sulfate, heparin, keratan sulfate, hyaluronic acid (as backbone of many proteoglycans).	Bacterial cell walls

The structures of the carbohydrate chains are generally determined by enzyme action and not by genetic matrices. The glycan chains are bound to asparagine and sometimes to the ε-group of lysine (N-glycosylation) or to serine or threonine, occasionally also to 5-hydroxylysine, hydroxyproline, histidine or tyrosine (O-glycosylation). They increase the solubility of proteins and protect against proteolysis.

In eukarya, the N-glycosylation of glycoproteins and proteoglycans begins in the endoplasmic reticulum (4.4.1) and is continued in the Golgi apparatus (4.4.2), while the O-glycosylation usually starts in the Golgi. The transport between the various cellular compartments and to the cell surface proceeds via vesicles (4.5.2). In some cases, specific carbohydrate sequences 'target' the glycoprotein to its destination (e.g., terminal mannose 6-phosphate guides to lysosomes (4.5.2.1). In its absence due to a genetic defect, the glycoprotein is secreted: I-cell disease).

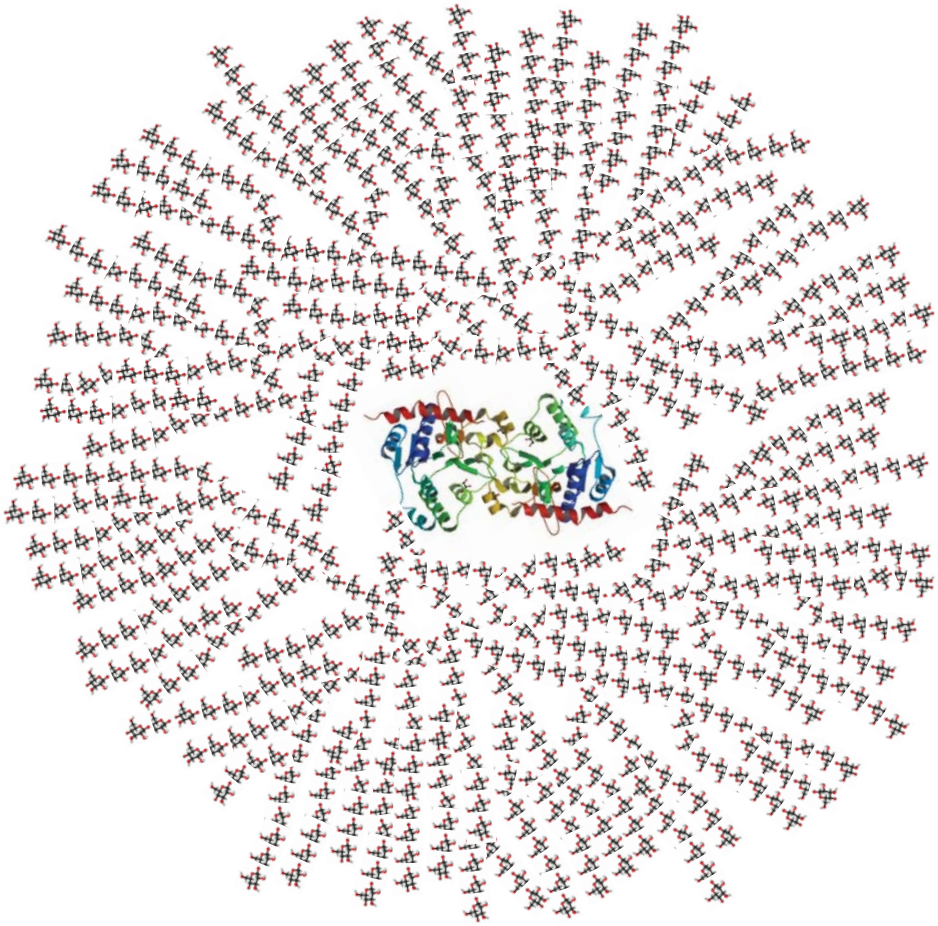


Figure 2.8-3. Model of the Glycogen Structure with a Central Protein Molecule
After commons.wikimedia.org

2.9.1 Glycoproteins (Syntheses in 4.4.2.1)

The carbohydrate content varies within a wide range (from almost 0 to 85 %). Both linear and branched carbohydrate chains occur. Frequently, the glycosidic groups are essential for protein function, e.g., in cell-cell interactions (cellular attachment, immunological reactions, formation of neuronal interconnections). Many soluble plasma glycoproteins carry terminal sialic acid (N-acetylneuraminic acid, 3.1.7). Degradation starts after removal of the neuraminic acid by neuroaminidase. The protein is recognized by an asialoglycoprotein receptor of the liver, internalized and degraded.

Mucins are very large (ca. $10^7 \ldots 10^8$ Da) glycoproteins occurring in animals. The protein backbone carries a huge number of mostly O-linked glycosidic chains. The saccharide residues are often sulfated, also sialic acid is present. Mucins protect epithelial layers (e.g., of the stomach).

In plant cell walls, hydroxyproline-rich, rod-shaped glycoproteins carry mostly side chains of arabinose tetrasaccharide units (3.1.6.3).

2.9.2 Proteoglycans (Synthesis in 4.4.2.1)

Proteoglycans form a considerable part of the cell surface. They are secreted into the extracellular space, where they frequently become non-covalently linked to hyaluronic acid and become part of the extracellular matrix. Proteoglycans contain a large number of binding sites for binding of other proteins and for mediation of specific cell-cell interactions and cell-cell recognition.

Many proteoglycans are important components of the extracellular connecting tissue, including heparin as an anticoagulant, hyaluronic acid as a component of the synovial fluid (lubricant in body joints), chondroitin sulfate in connective tissues, cartilage and tendons, dermatan sulfate playing a role in coagulation, cardiovascular disease, carcinogenesis, infection, wound repair, and fibrosis, and keratan sulfate, being found in the cornea, cartilage, and bone. Because of their high hydrophilicity they are surrounded by water molecules and can provide viscosity and lubrication in joints.

These compounds consist of a ‘core’ protein with many attached carbohydrate side chains (Fig. 2.9-1). Their length increases from the N to the C terminus of the protein. In longer chains, there is usually a short sequence of sugars, amino sugars and sialic acid (N-acetylneuraminic acid) closest to the protein, followed by many repeating disaccharide units containing amino sugars (glycosaminoglycans,

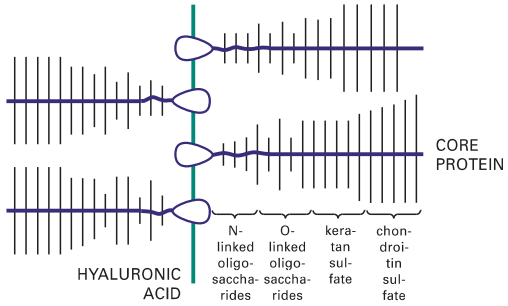


Figure 2.9-1 Structure Model of Cartilage Proteoglycan
(The areas of the various side chains are overlapping)

mucopolysaccharides, 2.9.5 and Fig. 2.9-2). The carbohydrate content can be up to 95%. The covalent binding of the glycosidic side chains to the 'core' protein takes place in the Golgi apparatus of fibroblasts, chondroblasts (cartilage), osteoblasts (bone) and similar cells.

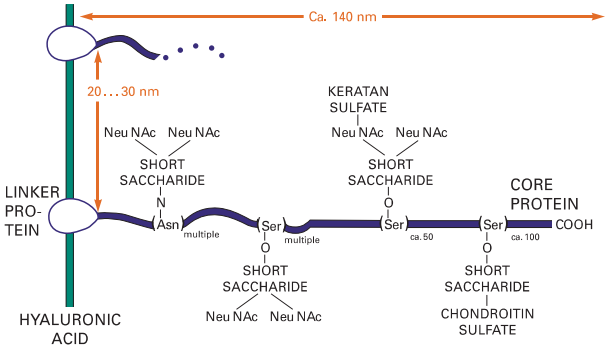


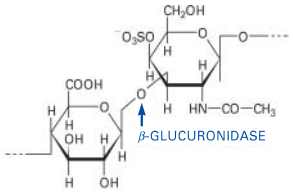
Figure 2.9-2 Schematized Structure of Proteoglycans

Degradation Diseases and Mucopolysaccharidoses: The saccharide chains of glycoproteins and proteoglycans are degraded in lysosomes. Lack of enzymes due to genetic reasons is the cause of various inherited lysosomal storage diseases. They are usually characterized by accumulation of the undegraded compounds in tissues and urine, skeletal deformations etc.

Since the specificities of the enzymes usually refer only to the structure of the particular bond, a single enzyme defect can cause various diseases (e.g., Morquio syndrome and G_{M1} gangliosidosis). Typical glycoprotein degradation diseases are fucosidosis and aspartylglucosaminuria (failure to cleave the N-acetylglucosamine-asparagine bond). Mucopolysaccharidoses are mentioned below.

Repeating Units of Glycosaminoglycans as Components of Proteoglycans. The attack site of degradative enzymes is shown by an arrow.

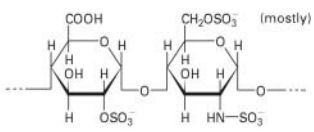
CHONDROITIN 4-SULFATE
[D-glucuronate (β1 → 3) D-GalNAc (4-sulfate) (β1 → 4)]$_n$



The sulfate group may also be at 6-position. These units are attached to a central sequence of Ser-O-Xyl-Gal-Gal-. Molecular mass = 5 * 10^3…5 * 10^4 Da. The biosynthesis proceeds via the consecutive action of the enzymes with EC-numbers 2.4.2.26, 2.4.1.133…135, (2.4.1.174 and 2.4.1.135)$_n$, 2.8.2.5 or 2.8.2.17. Large quantities are present in cartilage, aorta, connective tissue, bone, and skin.

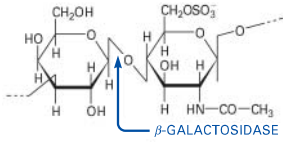
Lack of β-glucuronidase for degradation = mucopolysaccharidosis VII.

HEPARIN
[D-glucuronate (2-sulfate) (α1 → 4) N-sulfo-D-glucosamine (6-sulfate) (α1 → 4)]$_n$ (mostly)

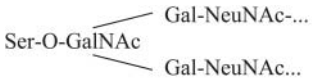


6 * 10^3…2.5 * 10^4 Da, contains also some D-Gal and D-xylose. Large quantities are found in mast cells, lung and liver. Heparin is an inhibitor of blood coagulation (9.3.4) by promoting a complex formation between active proteases (IIa, IXa, Xa) and antithrombin III.

KERATAN SULFATE
[D-Gal (β1→4) D-GlcNAc(6-sulfate) (β1→3)]$_n$



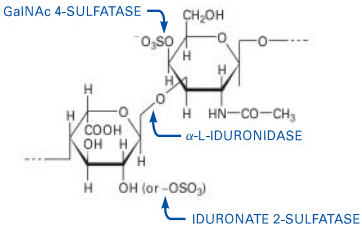
The units are attached to a central sequence of



5 * 10^3…2 * 10^4 Da, also contains D-mannose and L-fucose. Large quantities occur in cartilage, cornea, vertebrate disks.

Lack of β-galactosidase for degradation = Morquio type B syndrome (mucopolysaccharidosis IV). Deficiency in fucosidase = fucosidosis.

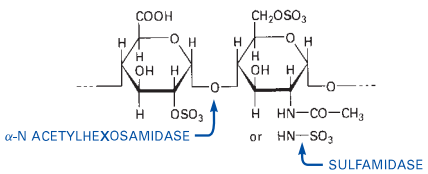
DERMATAN SULFATE
Same structure as chondroitin sulfate + about 20% of structure
[L-iduronate (α1 → 3) D-GalNAc (4-sulfate) (β1 → 4)]$_n$



1.5 * 10^4…4 * 10^4 Da, also contains some D-Gal, D-glucuronate, D-xylose. To some extent, sulfate is attached to the 2-position of L-iduronate. Found in skin, heart, and blood vessels.

Lack of α-L-iduronidase for degradation = Hurler's syndrome (mucopolysaccharidosis I). Lack of iduronate 2-sulfatase = Hunter's syndrome (mucopolysaccharidosis II). Lack of N-GalNAc 4-sulfatase = Maroteaux-Lamy's syndrome (mucopolysaccharidosis VI). Dermatan sulfate also inhibits blood coagulation (9.3.4).

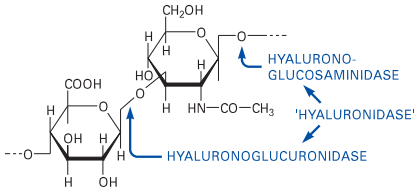
HEPARAN SULFATE
Similar structure to heparin, but with more N-acetyl groups and less O- or N-attached sulfate groups



5 * 10^3…1.2 * 10^4 Da. Ubiquitously present at cell surfaces and laminae.

Lack of sulfamidase for degradation = Sanfilippo's syndrome A. Lack of α-N-acetylhexosamidase = Sanfilippo's syndrome B (mucopolysaccharidosis III). Like heparin, it inhibits blood coagulation (9.3.4).

HYALURONIC ACID
[D-glucuronate (β1 → 3) GlcNAc (β1 → 4)]n



4 * 10^3…8 * 10^6 Da. Large quantities in connective tissues, skin, cartilage, synovial fluid. Hyaluronic acid forms the backbone of many proteoglycans. E.g., in the proteoglycan of cartilage, core proteins (which bear side chains of keratan sulfate and chondroitin sulfate) are non-covalently attached to hyaluronic acid with their N-terminus. The binding site is stabilized with linker proteins (Fig. 2.9-2).

Degradation takes place by hyaluronidases (a mixture of the enzymes with EC-numbers 3.2.1.35 and 36), which occur in many animal tissues, snake and insect toxins and in bacteria.

2.9.3 Peptidoglycans (Fig. 2.9-3)

Murein is a most important component of practically all bacterial cell walls. The glycosidic chains consists of repeating N-acetyl muramic acid – N-acetylglucosamine disaccharides MurNA (β1→ 4) GlcNAc. Each N-acetyl muramic acid is connected to a short (4- to 5-residue) amino acid chain, containing D-alanine, D-glutamic acid, *meso*-diaminopimelic acid and D-alanine in the case of *Escherichia coli* (a Gram negative bacterium) or L-alanine, D-glutamine, L-lysine, and D-alanine in the case of *Staphylococcus aureus* (a Gram positive bacterium). The amino acid side chains are cross-linked to a strong and rigid structure. The biosynthesis and function of murein is described in 3.10.1

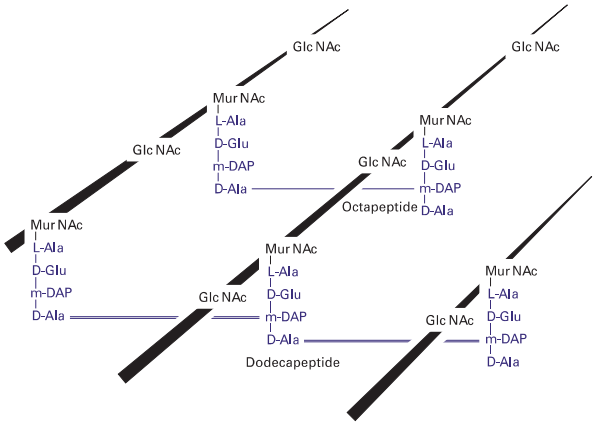


Figure 2.9-3 Structure of Murein (*Escherichia coli*)

Literature:

Montreuil, J., Vliegenthart, J.F.G., Schachter, H. Glycoproteins. *New Comprehensive Biochemistry* Vol. 29a. Elsevier, 1995.

Ruoslahti, E. Ann. Rev. of Cell Biol. 1988;4:229–255.

2.10 Lipid Aggregates and Membranes

Compounds, which contain hydrophobic (aliphatic or aromatic hydrocarbons) and hydrophilic regions (charged or polar alcoholic or phenolic groups) are called amphiphilic molecules. In particular, glycerophospholipids, sphingolipids, glycoglycerolipids, mono- and diacylglycerols, but also alkali salts of fatty acids (soaps) show this property. Molecules of this kind arrange in unique ways when in contact with an aqueous phase (Fig. 2.10-1):

- They can associate to spherical micelles with only hydrophilic groups located on the outside of a monomolecular layer. The interior of larger micelles can be filled with 'neutral' lipids lacking polar groups (e.g., triacylglycerols and cholesterol esters), yielding mixed micelles (detergent effect).

  This arrangement is mostly made by 'single-tailed' molecules, e.g., monoacylglycerol, soaps and artificial detergents and occurs only above a critical micelle concentration (cmc, < 1 µmol/l for biological lipids), otherwise the molecules cannot shield their hydrophobic tails from water.

- A feature of enormous importance in biology is the formation of lipid bilayers of about 6 nm thickness, which face aqueous phases on both sides. This is the basic arrangement of all cellular membranes including intracellular ones.

  Best suited for membrane formation are compounds with space filling hydrophobic areas (otherwise they would form micelles), such as glycerphospholipids and sphingolipids. In addition, the membranes usually contain many other compounds, primarily proteins (which contribute to the membrane function as well as to transport and metabolism) and cholesterol. The membrane constituents can move laterally within the membrane (fluid mosaic model).

Characteristics of membranes: The fluidity of the membranes depends on the lipid interior of the bilayer: it increases, when the order of arrangement decreases (e.g., caused by bent unsaturated fatty acids). Biological membranes undergo a phase change at a transition temperature (usually between 10…40° C), above which lateral movements of membrane components can take place more easily, although the basic structure is still kept up. Cholesterol (which by itself does not form bilayers) widens the transition range.

The lipids of the membranes are asymmetrically distributed. The 'flip-flop' exchange rate to the other membrane side is very low ($t_{1/2}$ = days, much higher in bacteria). In human erythrocytes, the distribution is as follows:

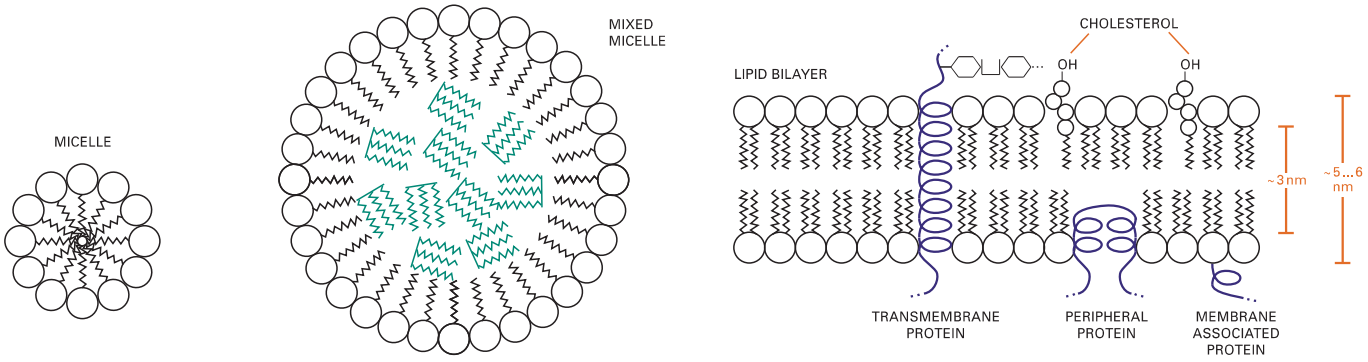


Figure 2.10-1. Structure of Lipid Aggregates

mol-% of total lipids	total	outer leaflet	inner leaflet
Phosphatidylcholine	30	22	28
Phosphatidylethanolamine	31	7	24
Phosphatidylserine	29	20	29
Sphingomyelin	25	21	24

Transmembrane proteins (integral proteins) span the membrane with α-helices of about 19 amino acids length. These domains contain mostly hydrophobic amino acids, which interact with the membrane lipids. If the protein contains several transmembrane domains, they can arrange in parallel columns in the membrane, which is hydrophobic on the outside and may form a hydrophilic channel at the inside. This enables the passage of hydrophilic compounds through the membrane, e.g., for the import of metabolites or ions (7.2.3, 6.1.1). Peripheral proteins are attached by hydrophobic interactions with the membrane lipids or by electrostatic or hydrogen bonds with integral proteins (e.g., liver cytochrome b_5). They can be removed under relatively mild experimental conditions. Membrane-associated proteins are bound to the membrane by lipid anchors, such as phosphatidylinositol (4.4.1.4) or isoprenoids (3.5.3.4).

Lipoproteins: For transport in the circulation system, the hydrophobic lipid aggregates are covered by a protein layer. The various lipoproteins are dealt with in Section 6.2.

Literature:

Dawidowicz, E.A. Ann. Rev. of Biochem. 1987;56:43–61.

Vance, D.E., Vance, J.E. (Eds.). *Biochemistry of Lipids, Lipoproteins and Membranes. New Comprehensive Biochemistry* Vol. 31. Elsevier, 1996.

3 Metabolism

3.1 Carbohydrate Metabolism and Citrate Cycle

Röbbe Wünschiers

3.1.1 Glycolysis and Gluconeogenesis

3.1.1.1 Glycolysis (Fig. 3.1.1-1)

Glycolysis (Embden-Meyerhof pathway) is the conversion of glucose (or, in a wider sense, of hexoses) to pyruvate. In consecutive metabolic steps, pyruvate is oxidized in the citrate cycle (3.1.8) or, with insufficient oxygen supply, converted to lactate (e.g., in animals, some microorganisms, and plant roots) or to ethanol (e.g., in yeast and plant roots) in order to reconstitute NAD^+, which is required for further progress of glycolysis.

Glycolysis is a key pathway of metabolism. It takes place in almost all living cells and

- supplies energy (in the form of 2 mol ATP per 1 mol glucose metabolized)
- supplies reducing equivalents (in the form of 2 mol NADH per 1 mol glucose), which yield additional ATP under aerobic conditions (3.11) or are consumed by reductions (e.g., under anaerobic conditions)
- converts carbohydrates into compounds which undergo terminal oxidation (acetyl-CoA, 3.1.8) or are used for biosynthesis (e.g., glycerol, acetyl-CoA).

$$\text{glucose} + 2\ ADP + 2\ P_i + 2\ NAD^+ = 2\ \text{pyruvate} + 2\ ATP + 2\ H_2O + 2\ NADH + 2\ H^+ \qquad \Delta G'_0 = -84\ \text{kJ/mol}$$

$$\text{glucose} + 2\ ADP + 2\ P_i = 2\ \text{lactate} - 2\ ATP + 2\ H_2O + 2\ H^+ \qquad \Delta G'_0 = -136\ \text{kJ/mol}$$

$$\text{glucose} + 2\ ADP + 2\ P_i = 2\ \text{ethanol} + 2\ ATP + 2\ H_2O + 2\ CO_2 \qquad \Delta G'_0 = -174\ \text{kJ/mol}$$

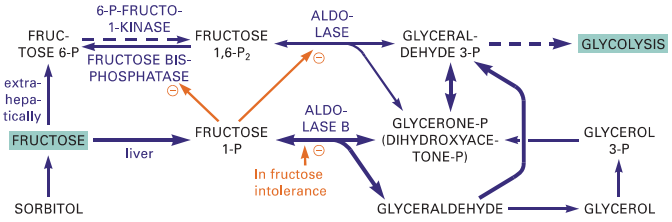
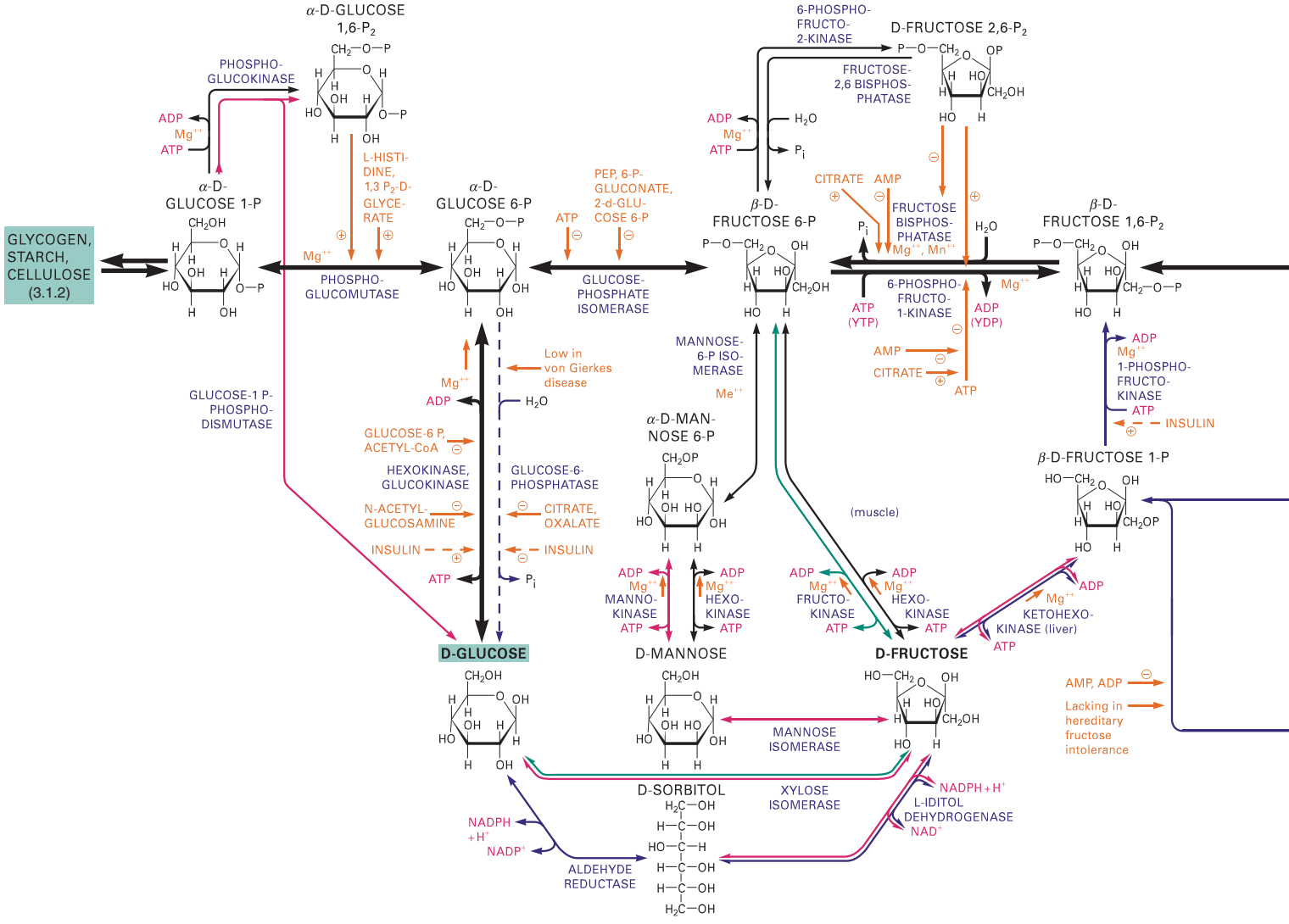


Figure 3.1.1-2. Fructose Metabolism



Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Second Edition. Edited by Gerhard Michal and Dietmar Schomburg.

The sequence of the glycolysis reactions and the formation of some important compounds for biosynthesis (glycerol, alanine etc.) are shown in Figure 3.1.1-1.

The enzymes are present in the cytosol, but are partially bound to structures (cellular membranes, in eukarya also to the mitochondrial outer membrane and the cytoskeleton). In plants, however, some of the glycolytic reactions also occur in the Calvin cycle (Fig. 3.12-8), which takes place in the chloroplasts of plant cells. Many enzymes are associated with each other or are interconnected by a common product/substrate (substrate channeling). Glycolysis shares many intermediates with the pentose phosphate cycle (3.1.6.1) and both pathways often interact. In plants, there is communication between both pathways across the plastide envelope.

Glucose directly enters the main pathway of glycolysis (Fig. 3.1.1-1) by phosphorylation (3.1.1.2). In contrast, mannose is first phosphorylated and then isomerized to fructose 6-P (its biosynthesis is discussed in 3.1.5.6). The interconversion of galactose to glucose and vice versa (via 1-phosphates and UDP-derivatives) by epimerization and transferase reactions takes place during both galactose catabolism and anabolism (3.1.4.2).

Fructose is converted in liver, kidney and intestinal mucosa by fructokinase to fructose 1-P (Fig. 3.1.1-2). This reaction is independent of hormones (therefore diabetics can tolerate fructose). In liver and kidney, this compound is cleaved by fructose bisphosphate aldolase B to glycerone-P (dihydroxyacetone-P, which is a member of the main path of glycolysis) and to glyceraldehyde (which enters the main path by phosphorylation to its 3-phosphate). In other tissues (e.g., muscle), phosphorylation of fructose to fructose 6-P takes place at a low rate. In hereditary fructose intolerance, fructose bisphosphate aldolase B is lacking in liver. The accumulated fructose 1-P inhibits fructose 1,6-bisphosphatase and fructose bisphosphate aldolase and therefore disturbs glycolysis and gluconeogenesis.

Sorbitol is oxidized to fructose in various tissues. Therefore, its metabolism is also hormone-independent.

Cofactors for the phosphoglucomutase and phosphoglyceromutase reactions are glucose-1,6-bisphosphate and 2,3-bisphosphoglycerate, respectively. Both are formed by kinase reactions from glucose 1-phosphate and 3-phosphoglycerate. In the mutase reaction, they confer a phosphate moiety (via phosphorylation of the enzyme) to the substrate. Thus, the cofactor is turned into the product, while the substrate is converted into the cofactor.

Isomerase and epimerase reactions usually involve the formation of enediolate intermediates by removal of a proton by a basic group at the enzyme. The reprotonation yields a different product.

Glyceraldehyde-3-P dehydrogenase uses an oxidation reaction to obtain a high-energy phosphate bond (which in consecutive reactions enables ATP formation):

D-glyceraldehyde 3-P + $NAD(P)^+$ + P_i = 1,3-P_2 D-glycerate + NAD(P)H

$\Delta G'_0$ = + 7.5 kJ/mol (actual ΔG_{phys} almost 0).

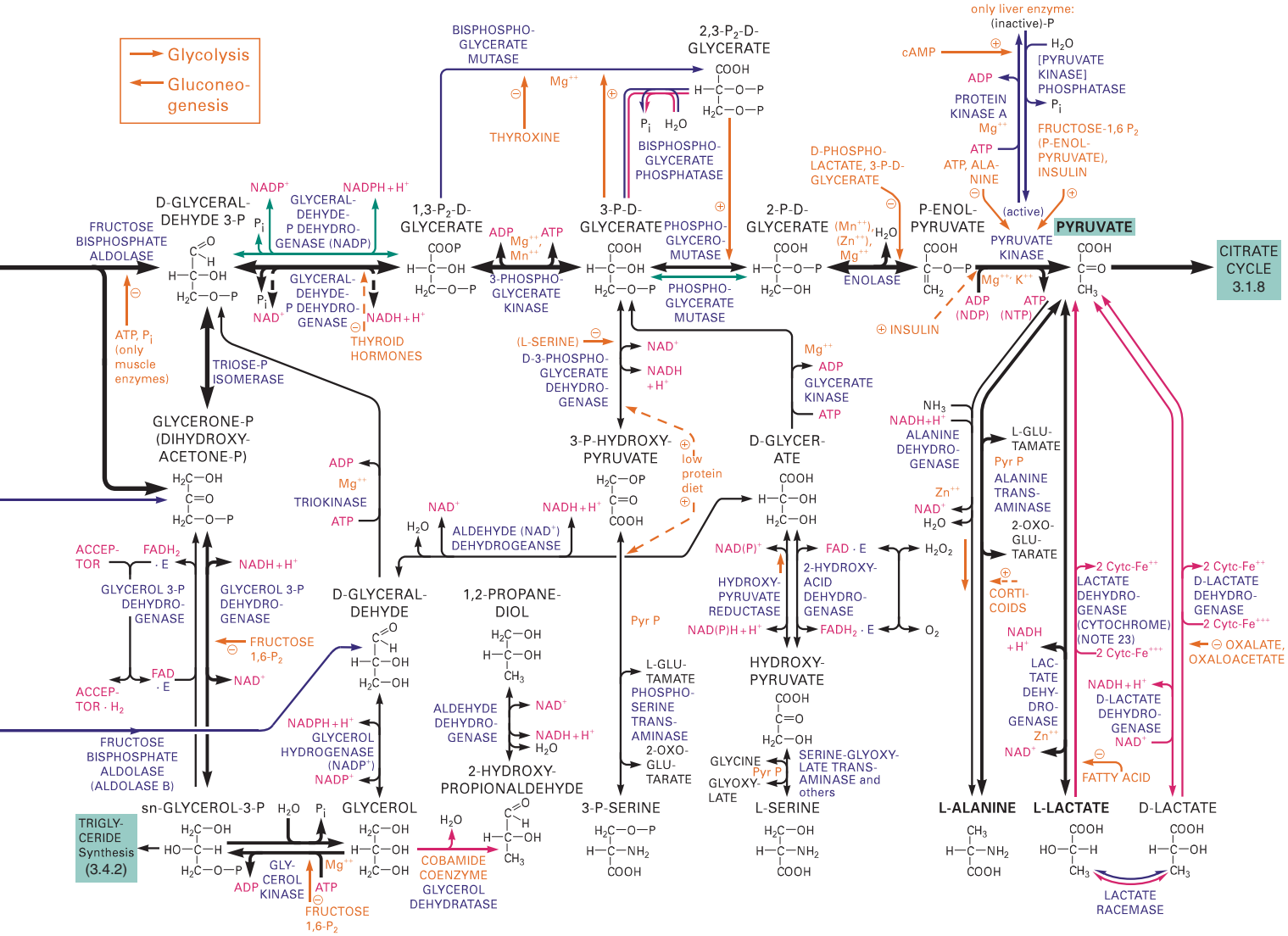


Figure 3.1.1-1. Glycolysis and Gluconeogenesis

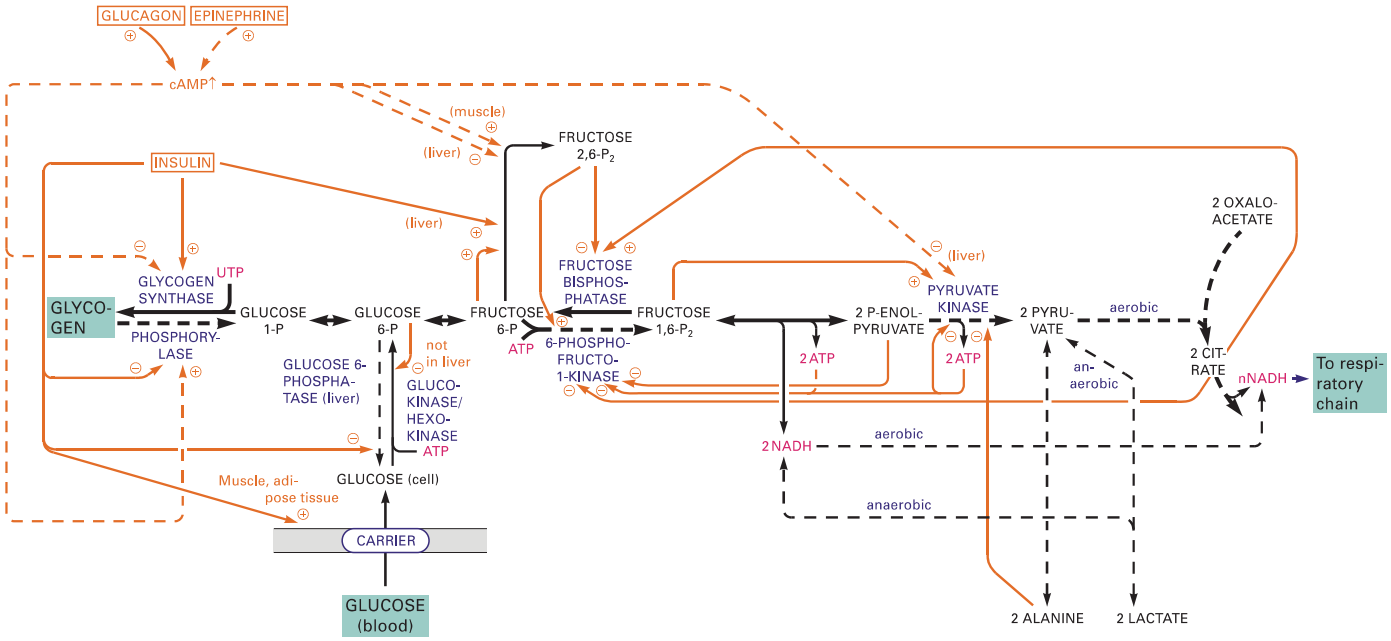


Figure 3.1.1-3. General Regulation Steps of Glycolysis (in Animals)
(Dashed orange arrows: control by glucagon/epinephrine, full orange arrows: control by insulin).

3.1.1.2 Regulation Steps in Glycolysis

It is obvious that this key metabolic pathway has to be strictly regulated. In animals, this occurs at many levels (Fig. 3.1.1-3). Under physiological conditions, the major regulated enzymes glucokinase (or hexokinase), phosphofructokinase and pyruvate kinase have restricted flow rates, the substrates pile up to some extent and the reactions are far from equilibrium (ΔG_{phys} strongly negative), while the other reactions are almost at equilibrium (ΔG_{phys} close to zero) and thus reversible.

Phosphorylation of glucose takes place by the following reaction:

glucose + ATP = glucose 6-P + ADP

The reaction is catalyzed in liver by glucokinase (K_M ca. 8 mmol/l). The enzyme is not saturated by its substrate under physiological conditions. Thus, a higher glucose concentration increases the rate of phosphorylation (Fig. 3.1.1-4). The reaction is not inhibited by its end products ADP and glucose 6-P. The phosphorylation is practically irreversible, but is counteracted *in vivo* by glucose 6-phosphatase (see below).

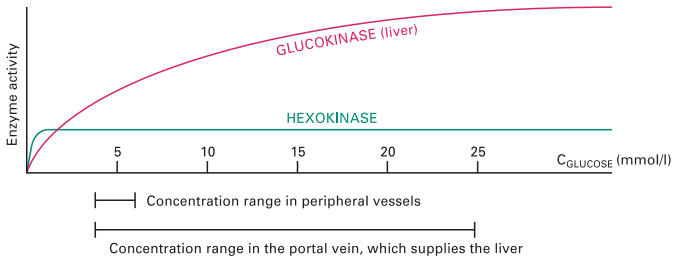


Figure 3.1.1-4. Phosphorylation of Glucose

In other organs, the reaction is catalyzed by hexokinase (K_M ca. 0.1 mmol/l). This enzyme is saturated by substrate under physiological conditions (Fig. 3.1.1-4). It is inhibited by its end products ADP and glucose 6-P. This achieves a steady supply rate for glycolysis intermediates. The enzyme is bound in a regulated way to the mitochondrial outer membrane (e.g., in the brain at the receptor porin → direct coupling to the energy state of the cell).

Plants contain multiple forms of hexokinase, which differ in cellular location and affinity for particular hexose substrates. While some isoenzymes phosphorylate a range of hexoses, others are specific for glucose (glucokinase) or fructose (fructokinase).

Dephosphorylation: Glucose 6-P is transferred to the endoplasmic reticulum of liver, kidney and intestine in a carrier-dependent way and is hydrolyzed there. The glucose formed returns to the cytosol and is secreted into the bloodstream. This causes elevation of low blood glucose levels. The glucose 6-phosphatase is inhibited by insulin.

The reaction of 6-phosphofructo-1-kinase (PFK, tetrameric, 4 * 85 kDa) is the committing step (first unambiguous step) in glycolysis. All earlier reactions can also lead to other pathways. Thus, it is the principal regulator of the flux of carbon through glycolysis and is regulated by a number of different effectors and mechanisms (Fig. 3.1.1-5). They also encompass fructose-1,6-bisphosphatase (FBPase, the corresponding reaction of gluconeogenesis) and prevent in this way the futile cycle of simultaneous occurrence of glycolysis and gluconeogenesis.

- PFK is allosterically inhibited by high ATP levels, this is counteracted by elevation of AMP levels (sensing of the energy supply, Pasteur-effect).
- PFK is allosterically inhibited by low pH (sensing of acidification due to lactate formation, 3.1.1.5).
- PFK is inhibited by high citrate concentrations (by increasing the ATP effect, indicating sufficient material for biosynthesis or for oxidation in the citrate cycle).
- The concentration of fructose 2,6–P_2, an activator of PFK and inhibitor of FBPase activities in liver, is regulated by fructose 6-P (indicating sufficient substrate) and by hormones (sensing the general metabolic situation). This effector is synthesized and degraded by the enzymes 6-phosphofructo-2-kinase and fructose-2,6-bisphosphatase, respectively, which are located on the same polypeptide chain.
  - Fructose 6-P activates 6-phosphofructo-2-kinase and inhibits fructose-2,6-bisphosphatase, thus increasing the fructose 2,6-P_2 level and promoting PFK activity (feed-forward activation).
  - The bifunctional peptide itself is regulated by phosphorylation and dephosphorylation (Fig. 3.1.1-5). In liver, the dephosphorylated state leads to an elevated level of fructose 2,6-P_2, thus to PFK activation and to enhanced glycolysis. This situation is achieved by insulin, which allosterically inhibits phosphorylation. Phosphorylation (which is activated by glucagon) reverses the situation and decreases the fructose 2,6-P_2 level. PFK is inhibited and FBPase is activated, the reaction is shifted from glycolysis to gluconeogenesis. In muscle a different isoenzyme exists, which responds the opposite way to phosphorylation and dephosphorylation.

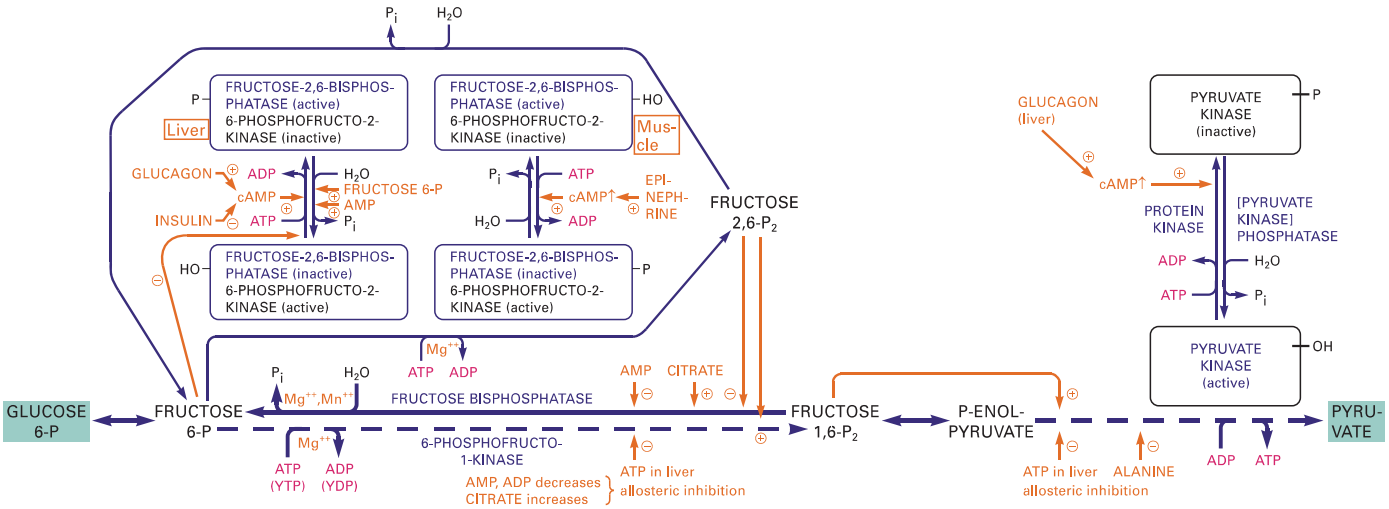


Figure 3.1.1-5. Detailed Regulation Mechanisms of 6-Phosphofructo 1-Kinase, Fructose Bisphosphatase and Pyruvate Kinase (in Animals)

In plants, PFK is neither regulated by fructose 2,6-P_2 nor by hormones. The situation is more complex due to the presence of plastid and cytosolic isoforms of PFK. Generally, the principal variable controlling plant PFK activity is the phosphoenolpyruvate/phosphate ratio, which links the activity of the enzyme to the status of the energy-conserving lower end of gycolysis.

Pyruvate kinase (tetrameric, 4 * 57 kDa) is inhibited by its reaction products ATP and alanine (formed from pyruvate by transamination, Fig. 3.1.1-5). In animals and microorganisms, but not in plants, it is feed-forward activated by fructose 1,6-P_2. In addition, the liver isoenzyme (but not the muscle isoenzyme) is regulated via its phosphorylation state: Glucagon promotes the inactivation by phosphorylation (Fig. 3.1.1-5). This reduces specifically the glucose consumption in liver and leaves the consumption in other organs unaffected.

Induction of enzymes: In addition to the 'fast' regulation by allosteric mechanisms, phosphorylation and dephosphorylation, the quantity of enzymes is also under hormonal and carbohydrate substrate control by gene expression regulation. The counteracting effects of hormones on various enzymes are shown in Table 3.1.1-1.

Since plants are anchored to the ground and cannot escape an unfavorable environment, they have a very flexible metabolism. Three bypass reactions in plant glycolysis circumvent the use of ATPs. These are:

a) pyrophosphate-dependent phosphofructokinase catalyzing

fructose 6-P + PP_i ↔ fructose 1,6-P_2 + P_i

b) non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase catalyzing

glyceraldehyde 3-P + $NADP^+$ → 3-P glycerate + $NADP^+$ + H^+

c) phosphoenolpyruvate phosphatase catalyzing

phosphoenolpyruvate → pyruvate + P_i

These bypasses may provide the plant with an ability to withstand phosphate starvation.

3.1.1.3 Gluconeogenesis

Gluconeogenesis is the formation of glucose (or its phosphates, which are further converted to poly- or oligosaccharides, 3.1.2 and 3.1.4) from non-carbohydrate sources. Examples are amino acids (alanine is first converted into pyruvate, other amino acids are converted into citrate cycle intermediates, 3.1.8) and glycerol, in plants and some microorganisms also fatty acids (via the glyoxylate cycle, 3.1.9.1). In animals, this takes place in the liver and to a minor extent in the kidney cortex. Its major purpose is to keep up a sufficient glucose supply to brain and muscles. In plants, gluconeogenesis is important for, e.g., the conversion of lipids to sugar in oil seeds.

Most reaction steps are a reversal of the respective glycolysis reactions, with the exception of those reactions which are highly regulated in glycolysis (3.1.1.2) and thus usually far from equilibrium. They are replaced by alternative reactions, which are energetically more favorable in the direction of glucose synthesis (glucose-6-phosphatase instead of glucokinase/hexokinase; hexose diphosphatase instead of phosphofructokinase). The complicated reactions for the circumvention of pyruvate kinase and their regulation are discussed in 3.1.3.5.

While the pathway glucose → pyruvate yields 2 ATP (3.1.1.1), the reverse conversion consumes 6 ATP in order to make it thermodynamically feasible.

2 pyruvate + 4 ATP + 2 GTP + 2 NADH + 6 H_2O =
glucose + 4 ADP + 2 GDP + 6 P_i + 2 NAD^+ + 2 H^+ $\quad \Delta G'_0 = -37.7$ kJ/mol

Table 3.1.1-1. Modulation of Enzyme Expression by Insulin, Catecholamines / cAMP and Glucocorticoids (L = in liver, A = in adipose tissue)

	Insulin induces (# only in presence of glucose)				Insulin represses
	Glycolysis enzymes:			**Other enzymes:**	
Catecholamines repress (via cAMP)	glucokinase	(3.1.1.2)	L	GLUT4 transporter (3.1.1.4) A	
	phosphofructokinase	(3.1.1.2) #	L A	acetyl-CoA carboxylase (3.4.1.1) L A	
	fructose-6-P 2-kinase	(3.1.1.2) #	L	fatty acid synthase (3.4.1.1) L A	
	pyruvate kinase	(3.1.1.2)	L	lipoprotein lipase (6.2.2) A	
Catecholamines induce (via cAMP)					**Gluconeogenesis enzymes:**
					pyruvate carboxylase (3.1.3.4) L
					PEP carboxykinase (3.1.3.4) L
Glucocorticoids induce					glucose 6-phosphatase (3.1.1.2) L
					fructose-1,6-bisphosphatase (3.1.1.2) L
					pyruvate carboxylase (3.1.3.4) L
					PEP carboxykinase (3.1.3.4) L
Glucagon induces (via cAMP)					fructose 1,6-bisphosphatase (3.1.1.2) L
					PEP carboxykinase (3.1.3.4) L

Regulation of gluconeogenesis: In addition to the inhibition of the fructose bisphosphatase reaction by fructose 2,6-P_2 (3.1.1.2), this enzyme is also inhibited by AMP and activated by citrate. Further, the expression of enzymes is regulated by hormones (Table 3.1-1).

3.1.1.4 Resorption of Glucose (Fig. 3.1.1-6)

The passage of glucose through most cell membranes proceeds via the regulated transport proteins GLUT1…5 with 12 transmembrane domains each. Glucose transport is provided by conformation changes of the protein.

- The uptake of glucose from the intestine into the epithelial cells takes place through the Na^+-glucose symporter (Table 6.1-3). Glucose passes from these cells into the bloodstream via GLUT5.
- GLUT1 and 3 are almost ubiquitous in mammalian cells. With a low K_M (ca. 1 mmol/l) they provide the basic glucose supply.
- Liver and pancreatic β-cells possess also the GLUT2 protein with the high K_M of ca. 20 mmol/l, well above the physiological glucose blood levels (4 … 7 mmol/l). Thus glucose uptake is proportional to the blood level and high only at elevated glucose concentrations.
- GLUT4 with an intermediate K_M of 8 mmol/l provides glucose uptake into muscles and adipose tissue. The number of transport proteins at the cellular membrane increases greatly in presence of insulin, indicating sufficient glucose supply. This takes place by reversible transfer of GLUT4 from internal vesicles to the cytoplasmic membrane.

3.1.1.5 Response of Animal Organs to High and Low Glucose Levels (Fig. 3.1.1-7)

The various organs respond differently to varying blood glucose concentrations, depending on their function. Both glycolysis and storage

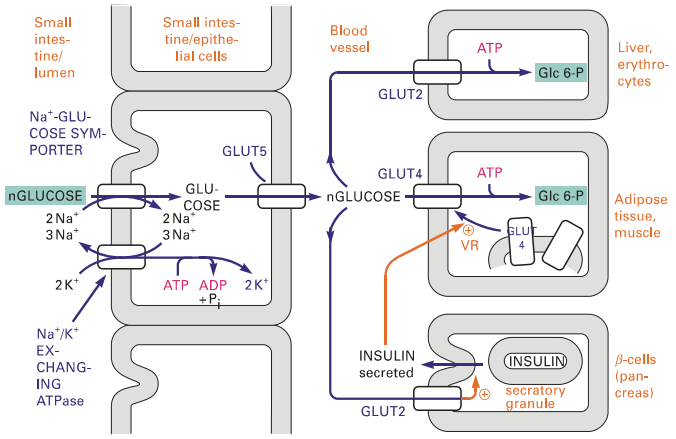


Figure 3.1.1-6. Glucose Resorption and Transport

as glycogen play a role. The mechanisms of glycogen formation, degradation and their hormonal control are described in detail in 3.1.2.

- Blood contains a small share of glucose and functions only as a transport medium between organs.
- Liver is the central storage organ for glycogen and acts as a buffer during the intervals of food uptake. It can release glucose for use by other organs.
- Muscles store a considerable amount of glycogen, but can use it only for their own purposes.
- All other organs consume glucose taken up from blood. Some of them can also use other compounds for their metabolism with the notable exception of the brain.

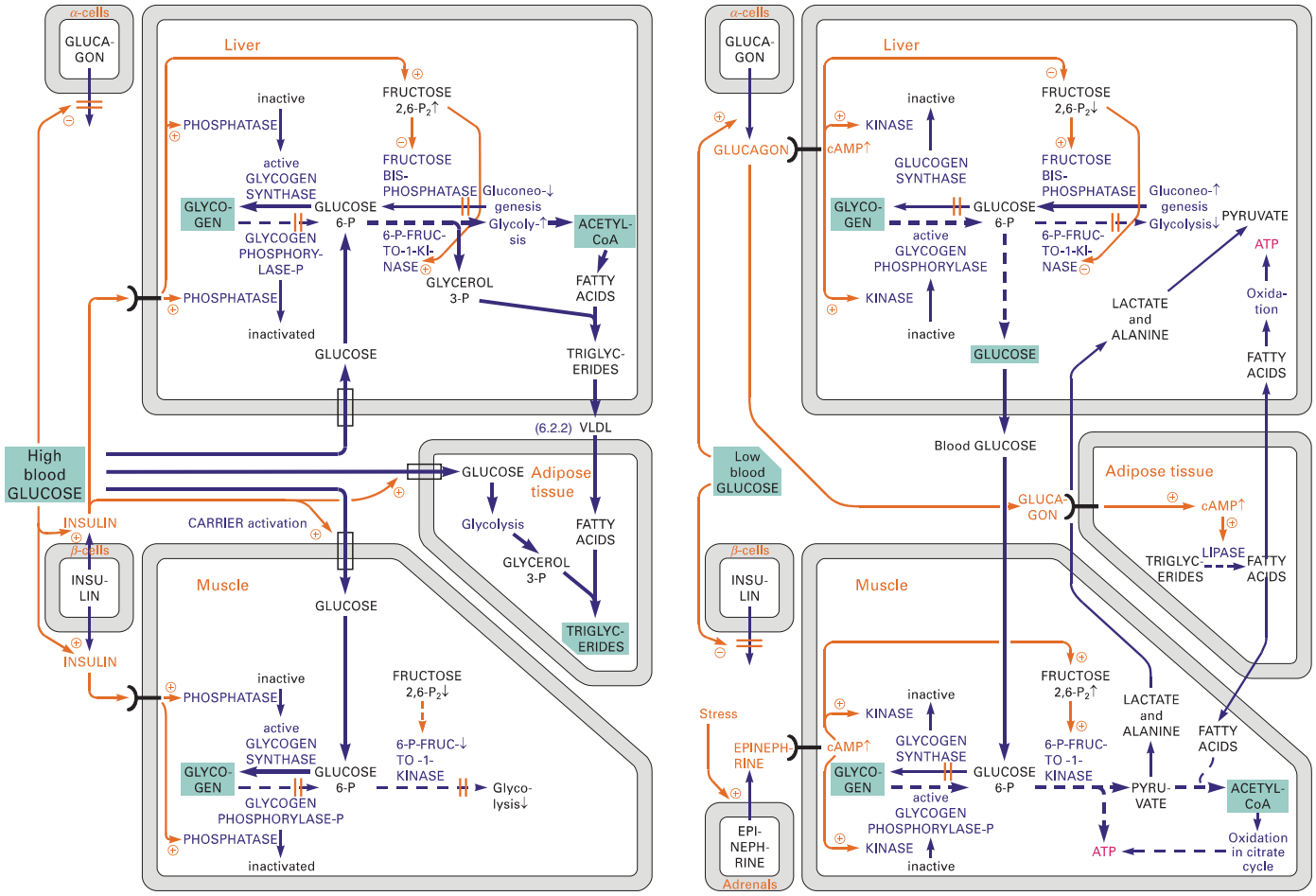


Figure 3.1.1-7. Reactions at High (Left) and Low (Right) Glucose Levels

Effects after glucose intake: At high glucose levels, storage reactions and conversions for biosynthetic purposes prevail.

- In liver, both the glucose uptake by GLUT2 (3.1.1.4) and its phosphorylation by glucokinase (3.1.1.2) respond to the high glucose concentration and remove a great portion of glucose from the bloodstream by forming glucose 6-P. The antagonizing reaction by glucose 6-phosphatase is inhibited.
- In the β cells of Langerhans islets in the pancreas increased glucose 6-P formation is the primary signal for insulin release (7.1.3).
- Glucose uptake in muscles and adipose tissue is increased by insulin, which acts on the GLUT4 transporter (see above). In these and in other organs, glucose phosphorylation is performed by hexokinase.
- Glycogen synthesis from glucose 6-P is activated in liver and muscle, while glycogen degradation is inhibited.
- Glycolysis from glucose 6-P is activated in liver and adipose tissue. This enables the formation of triglycerides as energy storage forms and of amino acids for biosynthetic reactions.
- Gluconeogenesis is generally inhibited.

Effects of starvation or sudden energy demand: At low blood glucose levels, glycogen synthesis is inhibited, while glycogen degradation and glucose release from liver is activated.

- In liver, the glucose 6-P formed from glycogen is dephosphorylated and glucose is released into the blood in order to keep up the physiological glucose level. This is essential especially for brain/nerves (daily consumption of an adult human 150 g), adrenal medulla and erythrocytes. The glycolysis steps beyond glucose 6-P are inhibited in liver.
- In muscle during exercise, glucose uptake and endogenous glycogenolysis take place to meet the ATP demand. The glucose 6-P formed from glycogen is not dephosphorylated to glucose, but rather passes through glycolysis to pyruvate and further on to lactate. The lactate formed is carried via the bloodstream to the liver, where it is reconverted to glucose ('Cori cycle'). Besides being reduced to lactate, pyruvate can also be transaminated to alanine (3.2.2.3), which likewise is transported to the liver and serves as source for gluconeogenesis.
- After prolonged glucose starvation, muscle and other organs degrade proteins to amino acids (3.2), which are used for gluconeogenesis or oxidized directly.
- In adipose tissue, triglycerides are degraded (3.4.2.2) and supply fatty acids and glycerol to liver and muscle. There they can be used for gluconeogenesis or are directly oxidized.

Literature:

Hers, H.G., Van Schaftingen, E. Biochem. J. 1982;206:1–12.

Lemaigre, F.P., Rousseau, G.G. Biochem. J. 1994;303:1–14.

Mueckler, M. Eur. J. Biochem. 1994;219:713–725.

Pilkis, S.J. *et al.* Ann. Rev. of Biochem. 1995;64:799–835.

Pilkis, S.J., Granner, D.K. Ann. Rev. of Physiol. 1992;52:885–909.

Silverman, M. Ann. Rev. of Biochem. 1991;60:757–794.

(Various authors). *The Enzymes of Glycolysis*. Phil. Trans. R. Soc. London 1981;293:1–214.

3.1.2 Polysaccharide Metabolism

3.1.2.1 Important Polysaccharides

Glycogen in heterotrophic organisms (animals, fungi, bacteria) and **starch** in photoautotrophs (plants, see Table 2.1-2) are polymeric storage forms for carbohydrates, which decrease the osmotic pressure. Glycogen is essential for glucose homeostasis as a readily available source.

Cellulose is a structure-forming component of plant cell walls, but also occurs in marine invertebrates. The amount synthesized in plants is estimated to be 10^{12} t/year. It is the compound present in largest quantities in the biosphere. Structure and function of these compounds are discussed in 2.8 and below.

Next to starch and sucrose, fructans are the most important metabolizable storage carbohydrates in plants.

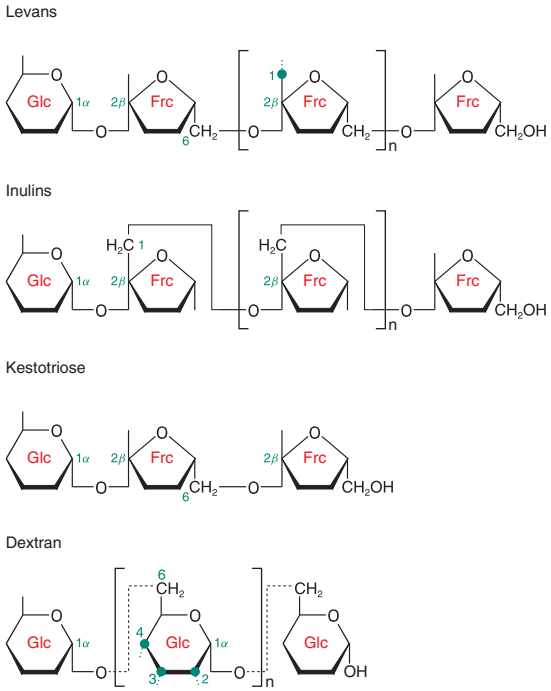


Figure 3.1.2-1. Fructans and Dextran
Glc = glucose, Frc = fructose, n = number of monomers, ...• = branching points

Fructans are linear or branched polyfructose chains (with predominantly β-fructofuranosyl-fructose linkages), which are attached to a single fructose molecule at various positions. This is used for classification (Fig. 3.1.2-1):

- Levans show mostly β2→6 linkages between fructose monomers with some β2→1 branches.
- Inulins show exclusively β2→1 linkages between fructose monomers.
- Kesto-n-oses (n denotes the number of monomers) are short chain fructans with both fructose-glucose and fructose-fructose linkages of the levan and inulin type.

Bacterial fructans are generally long-chained with molecular weights up to 10^7. Plant levans consist of approx. 200, plant inulins of about 50 monomers. In plant cells, fructans accumulate in the vacuole. Oligo- and polysaccharides of the fructan class are also synthesized by a large number of fungi. Inulins are considered as prebiotic in human nutrition and are used as fat substitute by the food industry.

Dextran is a branched glucan polymer of several thousand mostly α1→6 and α1→3 (branches) linked glucose units. Also α1→2 and α1→4 branching linkages occur. Dextran is mainly produced by lactic acid bacteria under anaerobic conditions and serves as water-binding carbohydrate protecting them from their environment. Dextran is widely used for chromatographic separation of macromolecules (Sephadex®) and as blood plasma substitute.

3.1.2.2 Biosynthesis of Polysaccharides (for structures, see 2.8)

The formation of the glycosidic bonds in polysaccharides proceeds in most cases via an activated incoming sugar:

glucose 6-P = glucose 1-P

glucose 1-P + UTP (ATP, GTP) = UDP (ADP, GDP)-D-glucose + PP_i
$\Delta G'_0$ = ca. 0 kJ/mol

$PP_i + H_2O = 2\ P_i$ $\Delta G'_0 = -33.5$ kJ/mol (drives the previous reaction to the right)

UDP (ADP, GDP)-D-glucose + $polysaccharide_n$
= $polysaccharide_{n+1}$ + UDP (ADP, GDP).

UTP is used in glycogen and in bacterial cellulose synthesis, ATP in starch synthesis, GTP and UTP in cellulose synthesis, respectively.

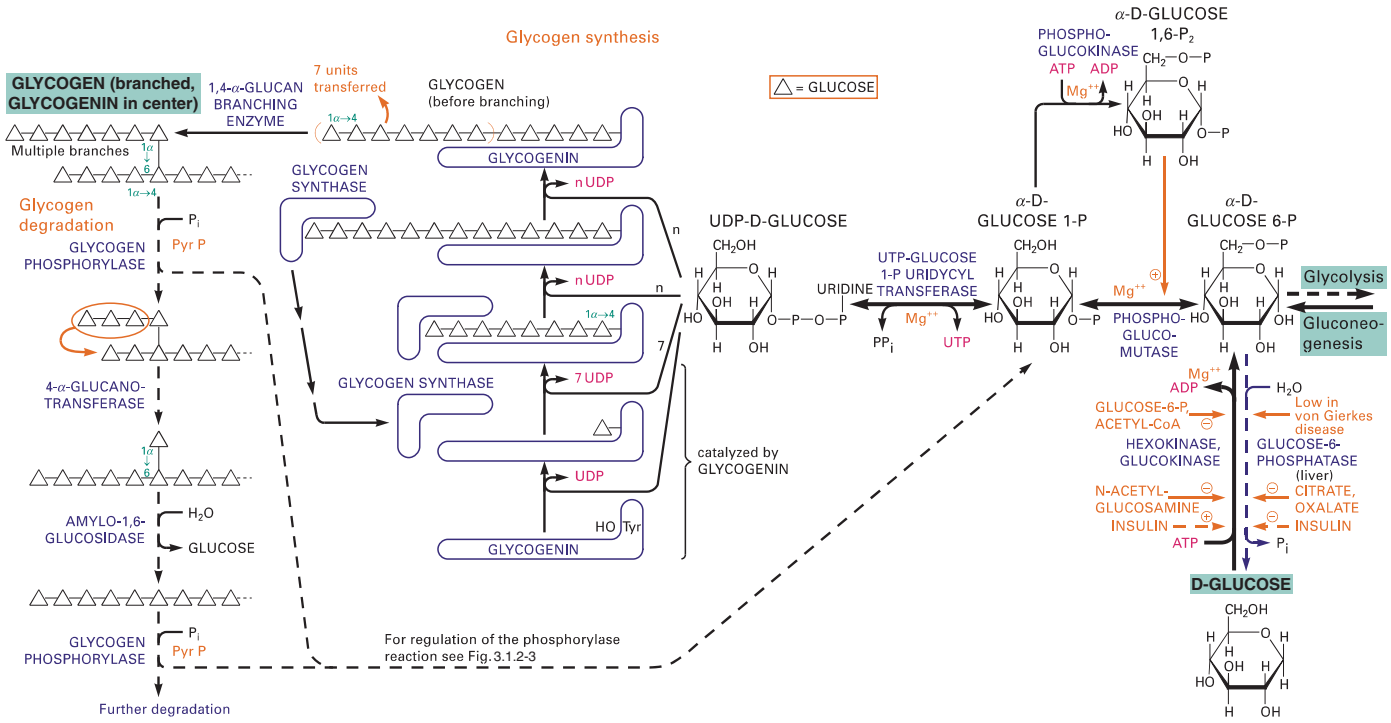


Figure 3.1.2-2. Glycogen Synthesis and Degradation

Glycogen synthesis (Fig. 3.1.2-2) in vertebrate liver and muscles starts, when the enzyme glycogenin (37 kDa) catalyzes the glucosylation of its own tyrosine residue and the extension of the chain to 8 glucose units. Then glycogen synthase takes over and enlarges the molecule as long as the synthase remains in firm contact with glycogenin (which remains in the center of the glycogen, Fig. 2.8-3). The branching mechanism during glycogen synthesis is catalyzed by 1,4-α-glucan branching enzyme, which transfers a terminal fragment of ca. 7 glucose residues to the C-6 hydroxyl of a glucose in the same or in another chain (Fig. 3.1.2-2). This enlarges the number of non-reducing ends, which are the attack points for degradation by phosphorylase.

Starch synthesis takes place in plant chloroplasts (3.12.2) from fructose 6-P, which is converted to glucose 6-P and further on to glucose 1-P and to ADP-glucose (Fig. 3.1.2-3). The latter reaction is regulated: low P_i and high 3-phosphoglycerate in chloroplasts promotes it. This situation

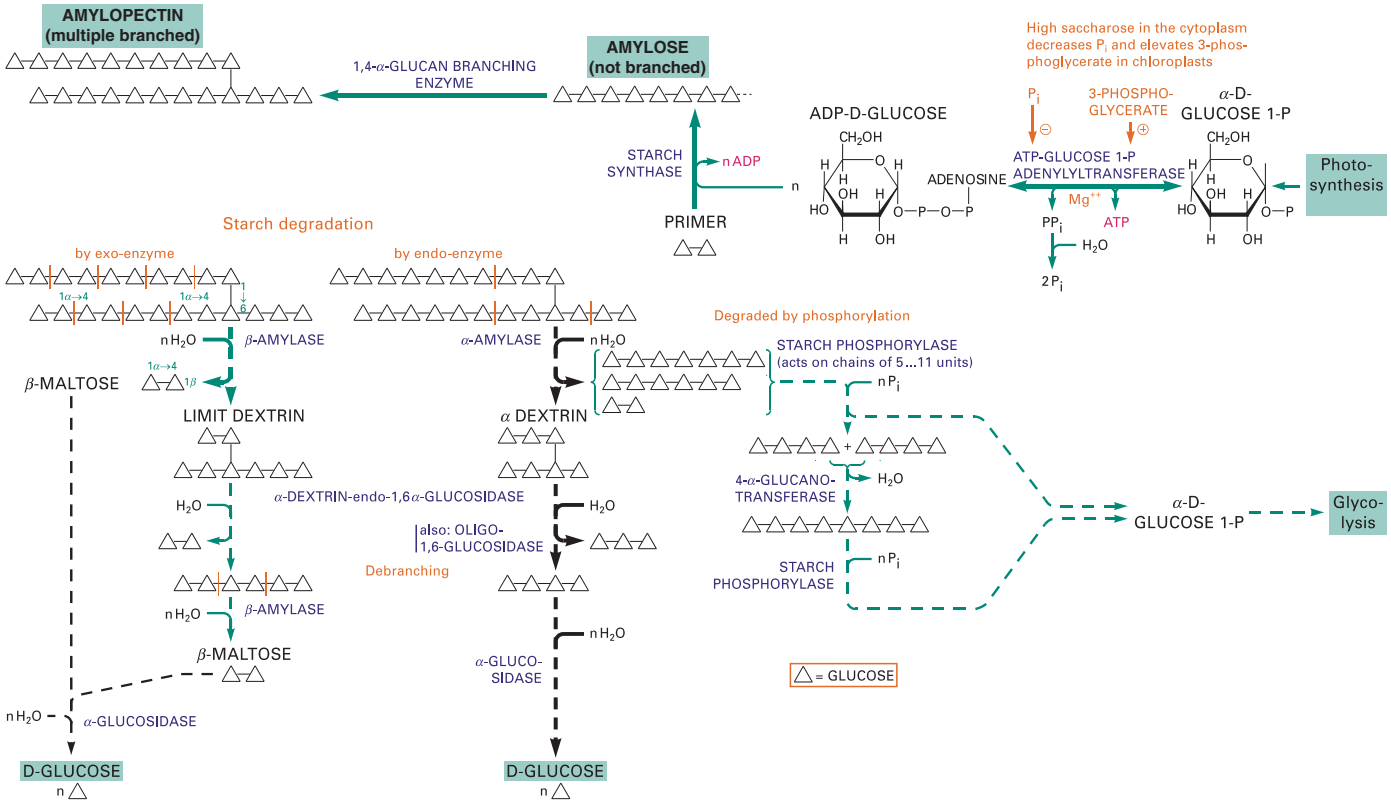


Figure 3.1.2-3. Starch Synthesis and Degradation

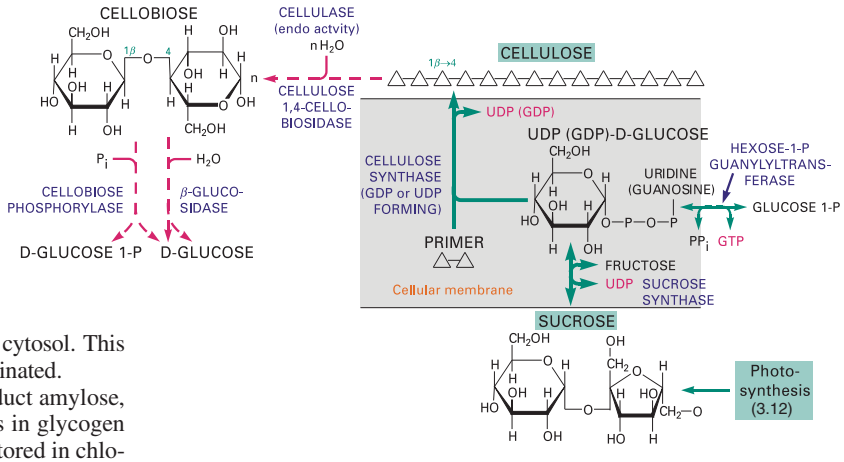


Figure 3.1.2-4. Cellulose Synthesis and Degradation

occurs, if there is a high concentration of sucrose in the cytosol. This way, the synthesis rate of both starch and sucrose is coordinated.

The action of starch synthase results in the linear product amylose, which is thereafter branched to amylopectin similarly as in glycogen synthesis, but at different chain lengths. The product is stored in chloroplast or in leukoplast granules (in heterotrophic plant organs).

Cellulose synthesis from UDP-D-glucose or GDP-D-glucose is catalyzed by cellulose synthases, which are bound to the plasma membrane (Fig. 3.1.2-4).

Frequently, the product of photosynthesis, sucrose (3.12.2), is converted at the membrane by sucrose synthase (3.1.4.1) and provides the activated glucose for cellulose synthesis at the location of use.

sucrose + UDP = UDP-glucose + fructose

The synthesis of dextrans and fructans also starts from sucrose, but by direct transfer of hexose units without intermediate formation of nucleotide derivatives. Dental plaques consist of dextrans.

3.1.2.3 Catabolism of Polysaccharides

Animal glycogen degradation (Fig. 3.1.2-2) starts at the non-reducing ends and is initiated by the action of phosphorylase (dimeric, 2 * 97 kDa). This reaction determines the hydrolysis rate and is strictly regulated (for details see 3.1.2.4).

$$\text{glycogen}_n + P_i = \text{glycogen}_{n-1} + \text{glucose 1-P} \qquad \Delta G'_0 = +3.1 \text{ kJ/mol}$$

Since, however, the physiological concentration ratio P_i/glucose 1-P is 30 … >100, the reaction is driven to the right (ΔG_{phys} ca. − 6 kJ/mol). It yields a phosphorylated product without requiring ATP. The reaction requires pyridoxal phosphate, which acts first as a proton donor and then as a proton acceptor (acid-base catalysis).

When the phosphorylase reaction has shortened the outer glycogen chains to a length of about 4 glucose units, a transfer reaction by 4-α-glucanotransferase takes place. The remaining single α1 → 6 bound glucose is hydrolytically removed by amylo-1,6-glucosidase, which is another function of the same peptide chain.

The degradation of starch can take place hydrolytically (in general) or by the starch phosphorylase reaction (analogous to glucagon, but only in plants). See Figure 3.1.2-3.

α-Amylase (in saliva and pancreatic secretions of animals and in plants, fungi, bacteria) is an endo-enzyme (randomly acting on inner bonds) and produces α-maltose, maltotriose and α-dextrin, which contains many α1 → 6 bonds. The latter compound is debranched by hydrolytic removal of the α1 → 6 bound glucoses. The linear oligosaccharides are thereafter degraded to glucose. Alternatively, starch phosphorylase converts them to glucose 1-P, similarly to above.

β-Amylase (in germinating plant seeds / malt) is an exo-enzyme, removing β-maltose units from the non-reducing ends (the α-bond in starch is inverted to the β-configuration). The reaction is interrupted when 1 → 6 bonds are reached; dextrin remains. Then, debranching (similar to above) has to take place.

$$\text{starch}_n + H_2O = \text{starch}_{n-2} + \beta\text{-maltose}$$

Cellulose is mostly cleaved to cellobiose by cellulase (endo-β-1,4-glucanase) (Fig. 3.1.2-4). It occurs in free-living and symbiontic bacteria, protozoa, fungi and some insects, e.g., higher termites.

cellulose + n H_2O = cellobiose

3.1.2.4 Regulation of Glycogen Metabolism in Mammals (Fig. 3.1.2-5)

The central importance of glycogen for glucose metabolism requires a tight control of its synthesis and degradation. Both phosphorylase and glycogen synthase are hormonally regulated via phosphorylation and dephosphorylation cascades; for obvious reasons in opposite directions. Details of such cascade mechanisms are described in 7.3 and 7.5. Generally,

- phosphorylation of the various enzymes is initiated by, e.g., epinephrine (in muscle) and glucagon (in liver) and favors glycogen degradation.
- dephosphoryation is initiated by insulin and favors glycogen synthesis.
- allosteric mechanisms provide another level of regulation.

Phosphorylation: Phosphorylase b, the usually inactive phosphate free form, is activated by phosphorylation at serine in position 14 to phosphorylase a (for allosteric interconversions see below). The activating phosphorylase kinase itself is activated by phosphorylation, catalyzed by protein kinase A, which is cAMP dependent and thus hormone controlled (7.4.2, e.g., by epinephrine).

Phosphorylase kinase also requires Ca^{++} for activation. In the phosphorylated state, already a moderate Ca^{++} elevation is sufficient. The elevation is sensed by calmodulin (7.4.4), which constitutes the δ subunit of the phosphorylase kinase $(\alpha\beta\gamma\delta)_4$. This way, phosphorylase kinase integrates stimulatory effects by hormones (via phosphorylation) and neuronal impulses (via Ca^{++} response).

Glycogen synthase exists in dephosphorylated and (9-fold) phosphorylated form. In contrast to phosphorylase, the dephosphorylated form (glycogen synthase a) is generally active, while the phosphorylated form (glycogen synthase b) requires a high level of glucose 6-P for activity (operates therefore only at high glucose supply). The phosphorylation is performed by the cAMP dependent protein kinase A and some other kinases.

Dephosphorylation: The phosphorylase system is deactivated by dephosphorylation either of phosphorylase a or of phosphorylase kinase. In both cases, the reaction is catalyzed by phosphatase 1 (PP1). The same enzyme also removes the phosphates from glycogen synthase and activates it.

The catalytic subunit of PP1 (37 kDa) obtains the affinity to glycogen (and thus to its protein substrates, which are associated with glycogen) by combination with the G-subunit ('glycogen binding', 160 kDa). Phosphorylation of the G-subunit by an insulin-dependent and cAMP-dependent kinase enables the association and promotes the phosphatase activity. If, however, the G-subunit becomes phosphorylated at another site by the cAMP dependent protein kinase A, the association of the catalytic and the G-subunits is prevented and the catalytic subunit remains inactive.

Additionally, the activity of PP1 is prevented by the inhibitor 1. This, however, only takes place if the inhibitor was phosphorylated by the cAMP dependent protein kinase A.

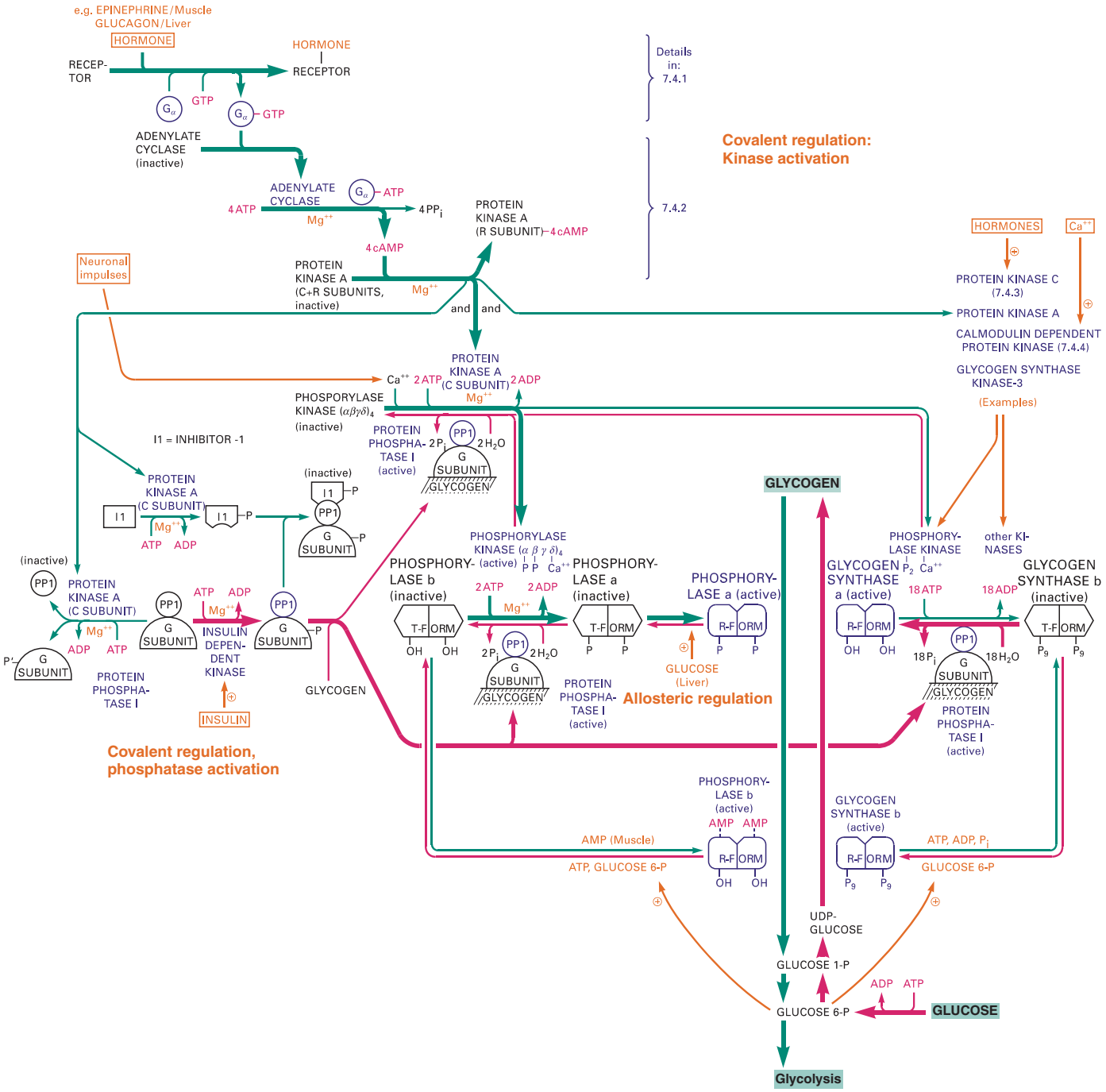


Figure 3.1.2-5. Regulation of Glycogen Synthesis and Degradation in Animals (contrary to the arrow colors in other figures, red and green arrows indicate reactions and regulation mechanisms leading to glycogen synthesis and degradation, respectively)

Allosteric mechanisms: Both phosphorylase a and phosphorylase b exist in active R- (relaxed) and inactive T-forms (tense, see 2.5.2).

In liver, glucose binding causes a shift of phosphorylase a from the R- to the T-form. This exposes the bound phosphate and enables inactivating dephosphorylation to yield phosphorylase b. Thus phosphorylase acts as a glucose sensor and prevents glycogenolysis if abundant glucose is available. Thereafter, the formed phosphorylase b releases the phosphatase, which is now free to act on glycogen synthase and activate it for formation of glycogen.

In the resting muscle, phosphorylase b prevails in the inactive T-form. There are two mechanisms of activation when the muscle requires energy supply:

- Hormones cause phosphorylation of phosphorylase b to a as described above.
- When AMP is bound (indicating low energy supply), phosphorylase b is converted from the T-form into the active R-form. This is counteracted by ATP and glucose 6-P (indicating sufficient energy and glucose supply).

3.1.2.5 Medical Aspects

Many diseases are caused by inheritable defects of the enzymes involved in glycogen metabolism (Table 3.1.2-1). Except in disease IX, glycogen is either elevated or of abnormal structure.

Literature:

Alonso, M.D. *et al*. FASEB J. 1995;9:1126–1137.

Browner, M.F., Fletterick, R.J. Trends in Biochem. Sci. 1992;17:66–71.

Table 3.1.2-1. Hereditary Glycogen Storage Diseases

Type	Name	Enzyme Deficiency	Tissue
0		glycogen synthase	
I	von Gierke's disease	glucose 6-phosphatase	liver, kidney
II	Pompe's disease	α-1,4 glucosidase	lysosomes
III	Forbes-Cori's disease	alpha-1,6-glucosidase	general
IV	Andersen's disease	1,4-α-glucan branching enzyme	liver, general?
V	McArdle's disease	glycogen phosphorylase	muscle
VI	Hers' disease	glycogen phosphorylase	liver
VII	Tarui's disease	phosphofructokinase	muscle
VIII	now classified as VI	phosphorylase b kinase	liver
IX		phosphorylase kinase, PHKA2	muscle
X	now classified as VI	phosphorylase kinase	liver
XI	glucose transporter, GLUT2	Fanconi-Bickel syndrome	liver, kidney

Hers, H.G. *et al.* in Scriver, C.R. *et al*. (Eds.) *The Metabolic Basis of Inherited Disease* 6th Ed. McGraw-Hill, 1989, p. 425–452.

Krebs, E.G. Angew. Chem. 1993;105:1173–1180.

Larner, J. Adv. in Enzymol. 1990;63:173–231.

Lomako, J. *et al.* FASEB J. 1993;7:1386–1393.

Roach, P.J. *et al.* Adv. Enzyme Regul. 1991;31:101–120.

3.1.3 Pyruvate Turnover and Acetyl-Coenzyme A

3.1.3.1 Pyruvate Oxidation (Fig. 3.1.3-1)

Pyruvate oxidation is common to all aerobic organisms. By action of the pyruvate dehydrogenase enzyme complex, pyruvate is converted to acetyl-CoA (3.1.3.3), the activated form of acetate. In eukarya, pyruvate is transported into the mitochondria, where this reaction takes place.

Under anaerobic conditions, pyruvate is converted into reduced compounds, such as lactate (e.g., in animal muscles) or ethanol (e.g., by yeast).

The oxidative decarboxylation is catalyzed by the multi-enzyme complex pyruvate dehydrogenase (lipoamide) consisting of pyruvate dehydrogenase (E_1), dihydrolipoamide acetyltransferase (E_2) and dihydrolipoamide dehydrogenase (E_3) subunits. They catalyze the following reactions (ThPP = thiamine pyrophosphate, Lip = lipoamide, for mechanism see 3.7.2 and Fig. 3.1.3-4):

E_1: pyruvate + ThPP-E_1 + H^+ = hydroxyethyl-ThPP-E_1 + CO_2

hydroxyethyl-ThPP-E_1 + lipoamide-E_2 = acetyl-dihydrolipoamide-E_2 + ThPP-E_1

E_2: acetyl-dihydrolipoamide-E_2 + CoA-SH = acetyl-CoA+dihydrolipoamide-E_2

E_3: dihydrolipoamide-E_2 + NAD^+ = lipoamide-E_2 + NADH + H^+

In summary, pyruvate is oxidized and decarboxylated to acetyl-CoA, CO_2, and NADH.

In the enzyme complex from *E. coli* (4600 kDa), 8 trimers of E_2, 12 dimers of E_1 and 6 dimers of E_3 subunits are arranged in highly symmetrical cubical order. The multienzyme complex from eukarya (8400 kDa) contains a 'nucleus' of 60 E_2 monomeric subunits surrounded by 30 E_1 dimers and 6 E_3 dimers, as well as 1 … 3 copies of pyruvate dehydrogenase kinase and -phosphatase. In both cases, lipoic acid (3.7.14.1) is bound to the ε-amino group of an E_2-lysine residue (hence 'lipoamide'). This 'arm' moves the attached acetyl group from E_1 to E_3. This enhances the reaction speed, coordinates the regulation of the reactions and avoids side reactions.

3.1.3.2 Regulation of Pyruvate Dehydrogenase Activity (Fig. 3.1.2-2)

In eukarya, E_2 and E_3 are competitively inhibited by the products of their reaction, acetyl-CoA and NADH, respectively. The eukaryotic E_1-subunit is inactivated by covalent phosphorylation at a serine moiety and activated by dephosphorylation. Acetyl-CoA and NADH promote the inactivating reaction. The modifying enzymes are attached to the E_2-'nucleus' of the multienzyme complex.

At high product concentrations, the reaction course of the E_2 and E_3 enzymes can be reversed. The E_1-catalyzed reaction, however, is irreversible.

In plants, ammonia inhibits pyruvate oxidation by stimulating the kinase. This effect may underlie the observed inhibition of the citrate cycle in illuminated leaves, since ammonia is produced within the mitochondrium during operation of the photorespiratory cycle. Eventually, starch production is initiated.

3.1.3.3 Acetyl-Coenzyme A (Acetyl-CoA)

Acetyl-CoA is an example of an 'energy-rich' thioester bond ($\Delta G'_0$ for hydrolysis = −31.5 kJ/mol). This is similar to the change of free energy during hydrolysis of high energy phosphate bonds of ATP ($\Delta G'_0$ = −30.5 kJ/mol for the γ-bond, −32.2 kJ/mol for the β-bond). Thus, acetyl-CoA can be used for ATP formation in acetate fermentations (3.10.5).

Acetyl-CoA plays a central role in metabolism. It is the common degradation product not only of carbohydrates, but also of fatty acids and of ketogenic amino acids (lysine and leucine, as well as of parts of the carbon skeleton of isoleucine, phenylalanine, tyrosine, tryptophan and threonine). For details, see the respective pages.

Acetyl-CoA may enter the citrate cycle (3.1.8) for degradation, but it can also be the origin of synthesis of fatty acids (3.4.1.1) and of cholesterol (3.5.1.1).

3.1.3.4 Anaplerotic Reactions (Fig. 3.1.3-1)

If members of the citrate cycle are used for biosyntheses, insufficient oxaloacetate is available for the reaction with acetyl-CoA. Anaplerotic ('filling up') reactions are required. The carboxylation of pyruvate requires an energy source, while this is not necessary for the conversion of the 'energy-rich' phosphoenolpyruvate.

- Pyruvate carboxylase reaction (in liver and kidney):

  pyruvate + ATP + CO_2 = oxaloacetate + ADP + P_i.

  The enzyme uses biotin (3.7.8) as a prosthetic group, which is attached by its valerate side chain to an ε-amino group of lysine, forming a mobile 'arm'. The enzyme is strongly activated allosterically by acetyl-CoA (for restarting the citrate cycle) and inhibited by high levels of nucleoside triphosphates (indicating sufficient energy supply) or insulin.

- 'Malic enzyme' reaction (frequent in eukarya, including plants and in prokarya):

  pyruvate + NADPH + CO_2 = malate + $NADP^+$.

  The energy for carboxylation is provided by NADPH oxidation. The reverse reaction in the cytosol is used to supply NADPH. Plant mitochondria contain relatively large amounts of NAD-linked malic enzyme.

- Phosphoenolpyruvate carboxykinase reaction (in heart and skeletal muscle):

  phosphoenolpyruvate + CO_2 + GDP = oxaloacetate + GTP.

  The reverse reaction is used in gluconeogenesis (3.1.3.5). In plants, the guanosine nucleotides are replaced by adenosine nucleotides. Bacteria use similar reactions (with phosphate, see Fig. 3.1.3-1) mainly for oxaloacetate formation during fermentations (3.10.5).

- Phosphoenolpyruvate carboxylase reaction (in higher plants, yeast, bacteria). This enzyme is also part of the CO_2 pumping mechanism (3.12.2):

  phosphoenolpyruvate + CO_2 + H_2O = oxaloacetate + P_i.

- In bacterial fermentations, pyruvate carboxylation can also be achieved by CO_2 transfer from other compounds via carboxyltransferases (3.10.5).

3.1.3.5 Initiation of Gluconeogenesis (Fig. 3.1.3-1)

The pyruvate kinase reaction is highly exergonic in the direction from phosphoenolpyruvate (PEP) to pyruvate ($\Delta G'_0$ = −23 kJ/mol) and under *in-vivo* conditions irreversible. Therefore, pyruvate cannot be

used directly for gluconeogenesis. The pyruvate carboxylase reaction, energized by ATP hydrolysis, initiates a bypass reaction and forms the 'energy-rich' compound oxaloacetate. In the PEP carboxykinase reaction, it is decarboxylated and accepts concomitantly a phosphate group from GTP (in animals) or ATP (in plants), yielding PEP. Insulin inhibits and glucagon activates both animal enzymes pyruvate carboxylase and PEP carboxykinase. The standard $\Delta G'_0$ of the overall reaction is 0.9 kJ/mol, but under physiological conditions ΔG_{phys} amounts to ca. −25 kJ/ mol, making it irreversible.

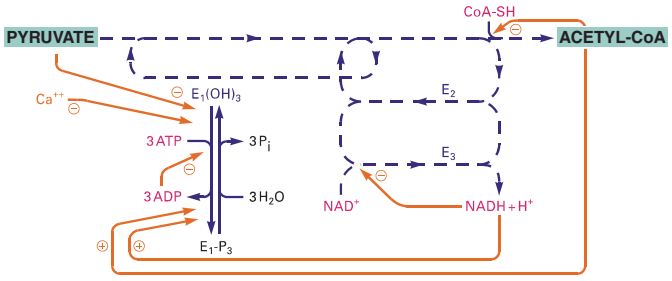


Figure 3.1.3-2. Regulation of the Reactions Catalyzed by Pyruvate Dehydrogenase Subunits

The carboxylation of pyruvate takes place only in mitochondria, while the conversion to PEP can occur either in mitochondria or in the cytosol (species dependent; in humans in both compartments). The further steps of gluconeogenesis (3.1.1.3) generally take place in the cytosol. Therefore, either oxaloacetate or PEP has to leave the mitochondria in order to enter this pathway. While PEP can be transported across the mitochondrial membrane, in animal tissues oxaloacetate has to be exported via the malate shuttle (Fig. 3.1.3-3) or alternatively via the aspartate shuttle (Fig. 3.11-2). The selected route depends on the cytosolic NADH requirements.

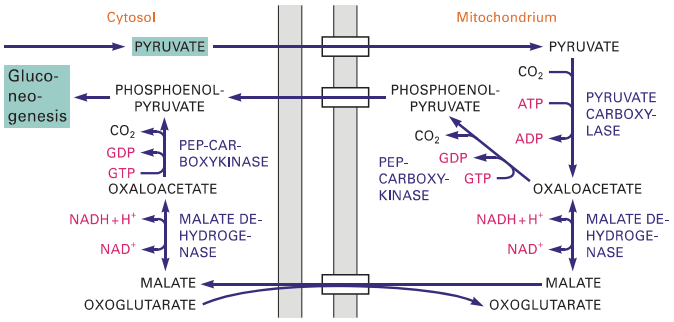


Figure 3.1.3-3. Transfer of Compounds Through the Mitochondrial Membrane by the Malate Shuttle

3.1.3.6 Alcoholic Fermentation

Under anaerobic conditions, yeast and a number of bacteria, but also some higher plants convert pyruvate to acetaldehyde by the action of pyruvate decarboxylase (Fig. 3.10.5-2). The first step is identical with the E_1 reaction of pyruvate dehydrogenase (3.1.3.1). Instead of transferring the acyl group of hydroxyethyl-ThPP to lipoamide, however, the group is eliminated. (In Fig. 3.1.3-4, the actually existing ionized forms are drawn, while in the main figure the non-ionized forms are shown.) Acetaldehyde is afterwards reduced to ethanol, thus regenerating NAD^+. This is similar to lactate formation in animals under anaerobic conditions.

In other bacterial species, different mechanisms exist for fermentative pyruvate turnover (e.g., pyruvate: ferredoxin oxidoreductase or pyruvate formate-lyase; the acetyl-CoA formed is converted via acetaldehyde to ethanol, 3.10.5).

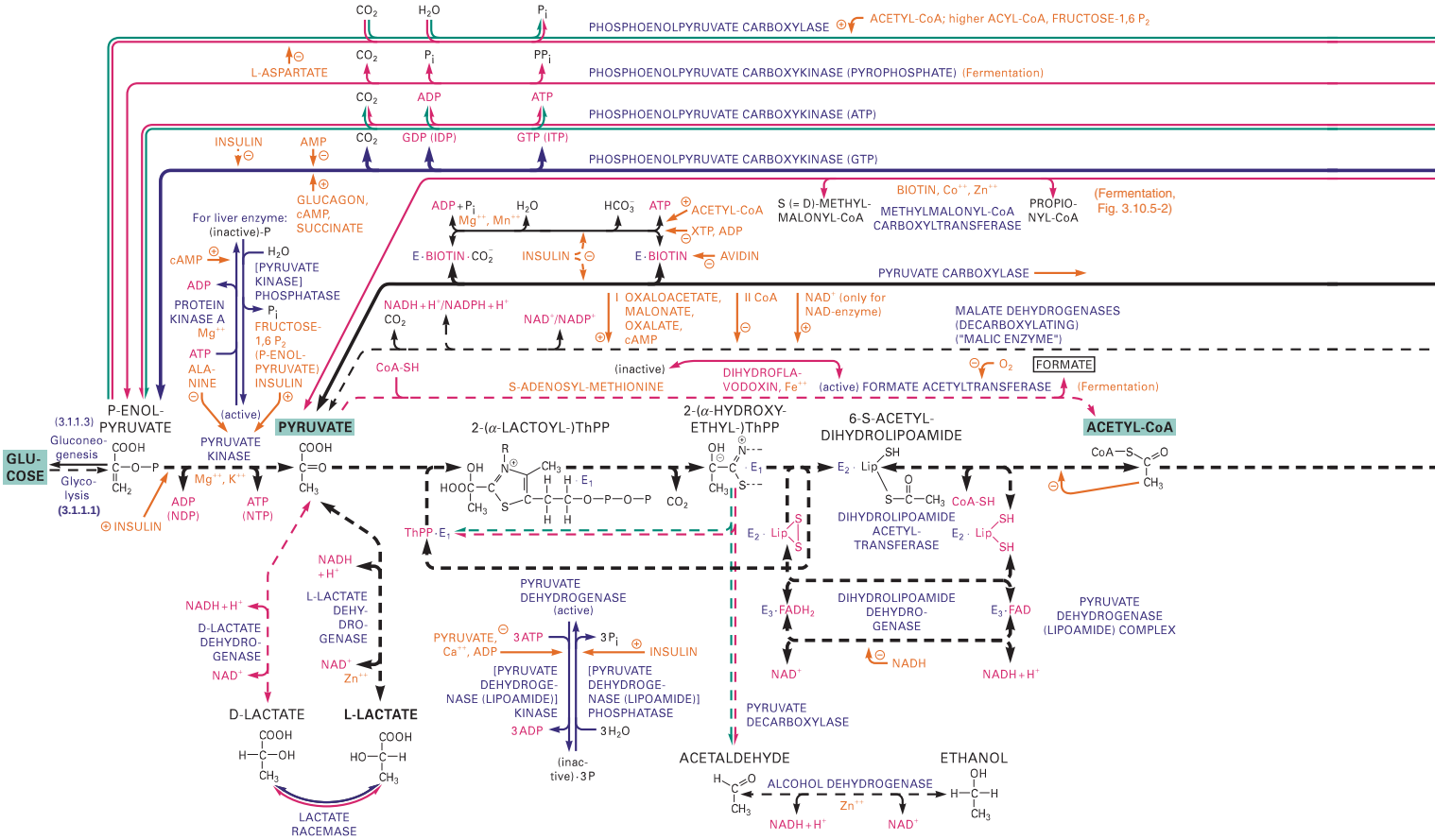


Figure 3.1.3-1. Reactions of Pyruvate and Phosphoenolpyruvate

Literature:

Attwood, P.V. Int. J. Biochem. Cell Biol. 1995;27:231–249.

Mallevi, A. *et al.* Curr. Opin. Struct. Biol. 1992;2:877–887.

Patel, M.S., Roche, T.E. FASEB J. 1990;4:3224–3233.

3.1.4 Di- and Oligosaccharides

3.1.4.1 Sucrose (Fig. 3.1.4-1)

Besides starch, sucrose is an important product of plant photosynthesis (3.12.2). The primarily formed triosephosphates are exported from the chloroplasts to the cytosol, converted to UDP-glucose and fructose 6-P and condensed to sucrose 6-P by sucrose-P synthase (Fig. 3.12-8). The phosphate is then removed by sucrose-P phosphatase.

The regulation of sucrose synthesis takes place at the fructose 1,6-bisphosphatase step by the concentration of the inhibitory fructose 2,6-P_2 (compare gluconeogenesis, 3.1.1.3). An additional regulation point is the sucrose-P synthase, which is activated by the substrate precursor glucose 1-P and inhibited by the product P_i. Additionally, this enzyme activity is decreased by phosphorylation and enhanced by dephosphorylation (similarly to glycogen synthase, Fig. 3.1.2-5). The sucrose concentration, in turn, regulates the starch synthesis (3.1.2.2).

Sucrose is a transport form of carbohydrates, as well as the precursor of starch (in cells distant from the site of photosynthesis) and of cellulose. The nucleotide sugars, which are necessary for their synthesis, are formed by the reversible enzyme sucrose synthase:

sucrose + UDP = UDP-D-glucose + fructose.

Other synthesis reactions take place by direct transglycosylations, e.g., the formation of fructans (in plants, 3.1.2.2) and of dextrans (in bacteria and yeast). Synthesis of the raffinose family starts with an isomerization of UDP-D-glucose to UDP-D-galactose, its condensation with myo-inositol to galactinol, followed by a transglycosylation reaction with sucrose.

Sucrose in food is cleaved in the intestine by α-glucosidase or by β-fructofuranosidase (invertase).

3.1.4.2 Lactose (Fig. 3.1.4-2)

Lactose synthesis takes place in the mammary gland of mammals. UDP-D-glucose is epimerized to UDP-D-galactose. Then, it is condensed with glucose.

The enzyme lactose synthase consists of the subunits galactosyl transferase (which preferably condenses UDP-D-galactose with N-acetylglucosamine, e.g., in synthesis of complex glycoproteins, 4.4.2.1) and α-lactalbumin (which changes the specificity, so that the transferase accepts glucose). During pregnancy, the transferase biosynthesis is induced by insulin, cortisol and prolactin, while the lactalbumin biosynthesis is inhibited by progesterone. Shortly before birth, the progesterone concentration decreases and this inhibition ceases. Lactose synthesis starts.

For catabolism, lactose is hydrolyzed in the intestine by β-galactosidase (lactase). After resorption, it is phosphorylated in the liver to galactose 1-P. Galactose 1-P reacts with UDP-D-glucose to yield glucose 1-P and UDP-D-galactose. This reaction is catalyzed by UDP-D-glucose-hexose-1-P uridylyltransferase. UDP-D-galactose is epimerized afterwards to UDP-D-glucose.

While in infants β-galactosidase is generally present, it exists only at a low level in the adult black and oriental population (lactose intolerance). In hereditary galactosemia, the uridylyltransferase has low activity in liver and erythrocytes, causing an increase of galactose 1-P. This compound inhibits phosphoglucomutase, glucose-6-phosphatase and glucose-6-P dehydrogenase, causing serious disturbances in glucose metabolism.

In bacteria, β-galactosidase is an inducible enzyme (for regulation see 4.1.2).

3.1.4.3 Other Glycosides

Many other α- and β-saccharides are synthesized by reaction of nucleotide sugars with other sugars or aglycons. A large number have been found in plants, yeast etc. Also, the synthesis of homo- or heteropolymeric glycosides usually starts from nucleotide sugars (e.g., 3.1.1.2, 4.4.1, 4.4.2).

3.1.5 Metabolism of Hexose Derivatives

3.1.5.1 Uronic acids (Fig. 3.1.5-1)

Uronic acids are derivatives of hexoses, in which the hydroxyl group at C-6 is oxidized to a carboxyl group. This primary oxidation takes place with UDP-glucose or GDP-mannose (not with the free sugar). By 4′or 5′epimerizations (1.2.1), UDP-D-glucuronate is converted into UDP-D-galacturonate or into UDP-L-iduronate, respectively. Uronic acids (as well as aldonic acids, 3.1.5.2) and their derivatives have a strong tendency to form internal esters (lactones).

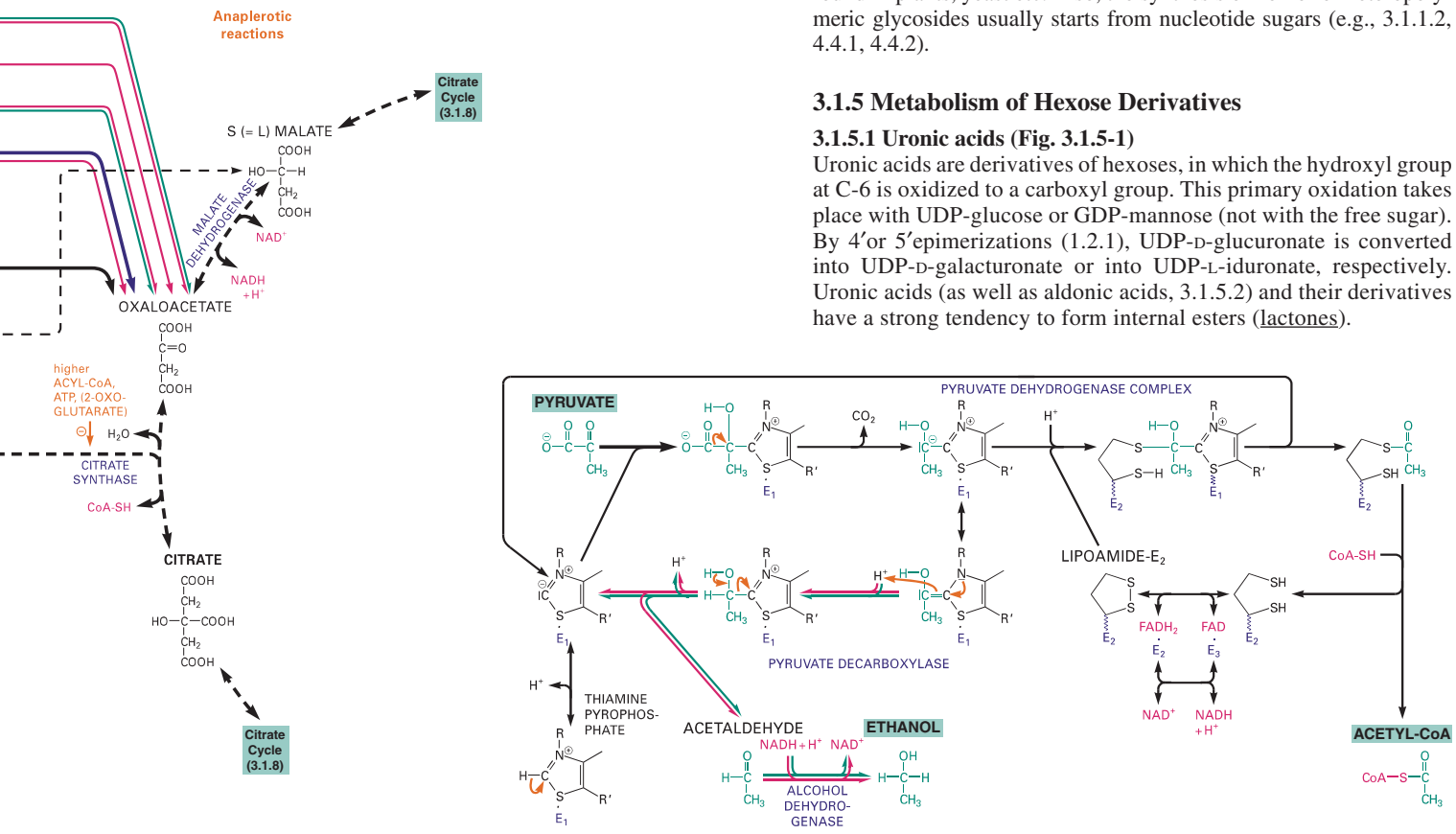


Figure 3.1.3-4. Reaction Mechanisms of Pyruvate Dehydrogenase and Decarboxylase

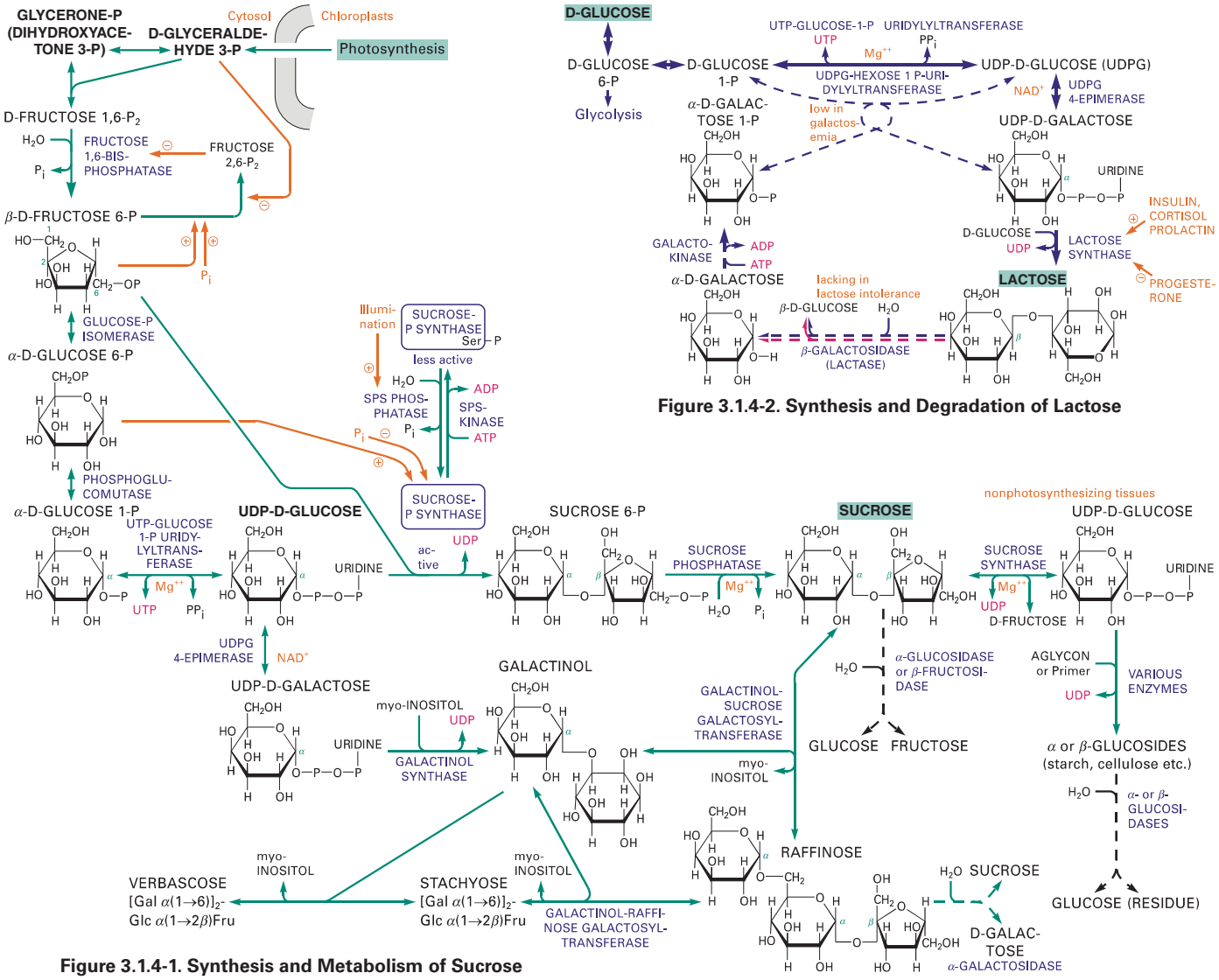


Figure 3.1.4-2. Synthesis and Degradation of Lactose

Figure 3.1.4-1. Synthesis and Metabolism of Sucrose

UDP-glucuronate reacts with many aglycons (primary amines, alcohols, carbonic acids) to glucuronides. Analogous reactions take place with other UDP-uronates. In animals, glucuronidation is of importance for excretion in urine. Uronates are important components of proteoglycans (2.9.2). Alginate (the polymerization product of D-mannuronate and L-guluronate) occurs in the cell walls of brown algae and is used in food industry as viscous gum.

Additional oxidation of D-glucuronate at C-1 to the carboxyl level yields the dicarboxylic D-glucaric acid.

3.1.5.2 Aldonic Acids (Fig. 3.1.5-1)

If the oxidation of D-glucose to the carboxylic function occurs at the C-1 aldehyde group, D-gluconate results. This reaction may take place either with free glucose or with its 6-phosphate derivative. The product of the latter reaction, 6-P-gluconate, may enter the pentose phosphate cycle (3.1.6.1). In bacteria, it is an intermediate of the Entner-Doudoroff pathway (3.1.5.3).

Another pathway leading to aldonic acids is the reduction of the C-1 hemiacetal group of D-glucuronate to the hydroxyl function, which yields L-gulonate. Here, the change to the L-form is not a conversion to an enantiomer (1.2.1), but the conventional numbering of this compound starts at the opposite end. Oxidation of L-gulonate or its lactone results in 3-dehydro-L-gulonate or ascorbate (vitamin C, 3.7.10), respectively.

3.1.5.3 Entner-Doudoroff Pathway (Fig. 3.1.5-1, see also 3.10.5)

The Entner-Doudoroff pathway describes an alternate series of reactions that catabolize glucose to pyruvate using a set of enzymes different from those used in either glycolysis or the pentose phosphate pathway. This pathway, also named 2-dehydro-(or keto-) 3-deoxy-6-P-gluconate pathway, is restricted to prokarya (e.g., *Zymomonas*) and catalyzes the degradation of gluconate and glucose. After phosphorylation, 6-P-gluconate is dehydrated to 2-dehydro-3-deoxy-6-P-gluconate and cleaved into pyruvate and glyceraldehyde 3-P (which is then also converted to pyruvate); further reactions lead to ethanol.

Via several intermediate steps, glucuronate and galacturonate can also yield 2-dehydro-3-deoxy-6-P-gluconate and are degraded the same way afterwards.

3.1.5.4 Inositol (Fig. 3.1.5-1)

The most prominent naturally occurring inositol, *myo*-inositol, is formed from glucose 1-phosphate in a NAD^+-catalyzed oxidation/reduction reaction. It is a cyclic alcohol with 6 hydroxyl groups, one of many stereoisomers. In the phosphorylated form, it plays an important role in intracellular signal transfer (7.4.4) and is also present in phospholipids (3.4.3) and in glycolipid anchors of proteins (4.4.1.4). In plants, inositol phosphates are present in large quantities; part of

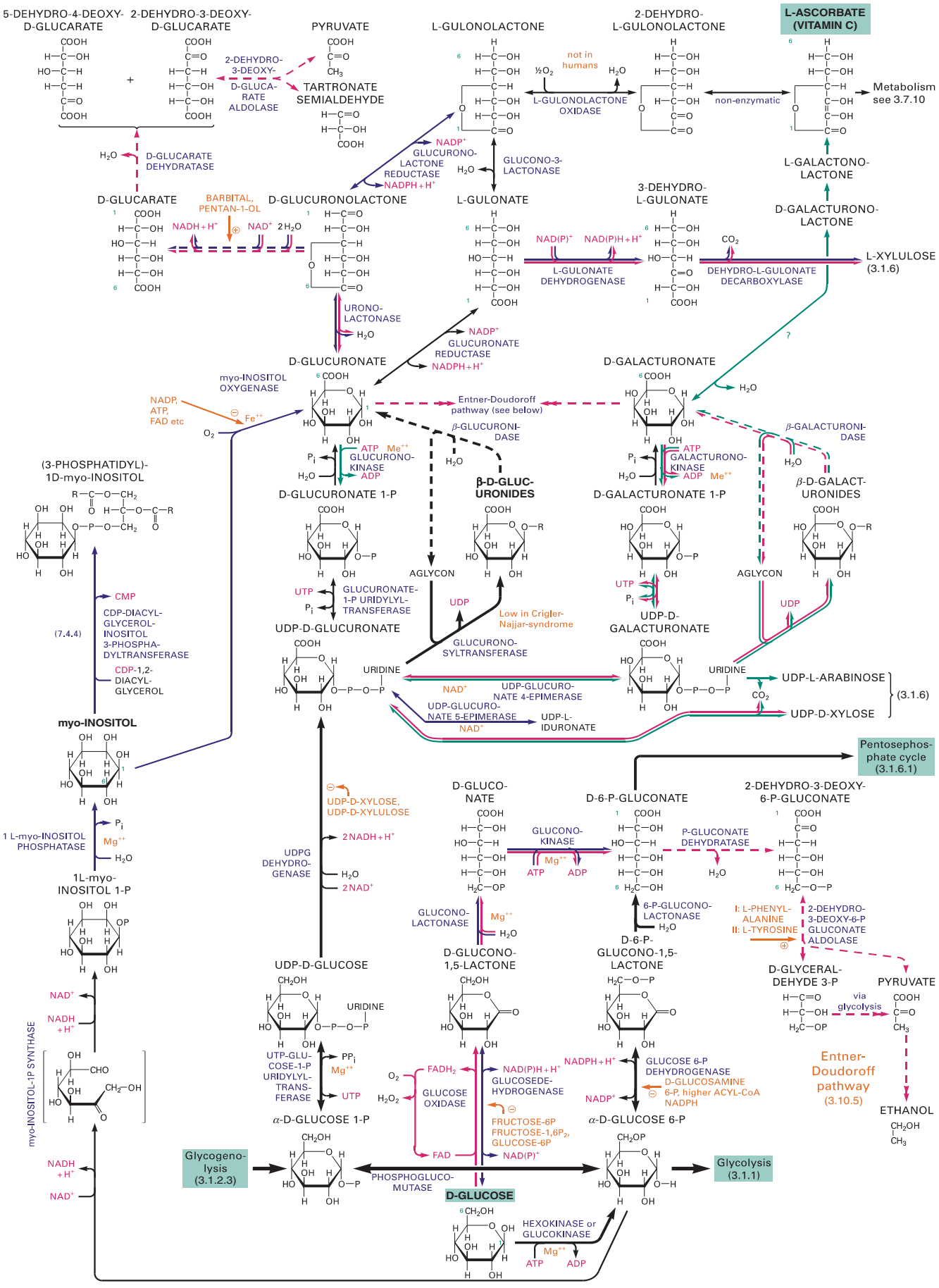


Figure 3.1.5-1. Acidic Hexose Derivatives and Inositol

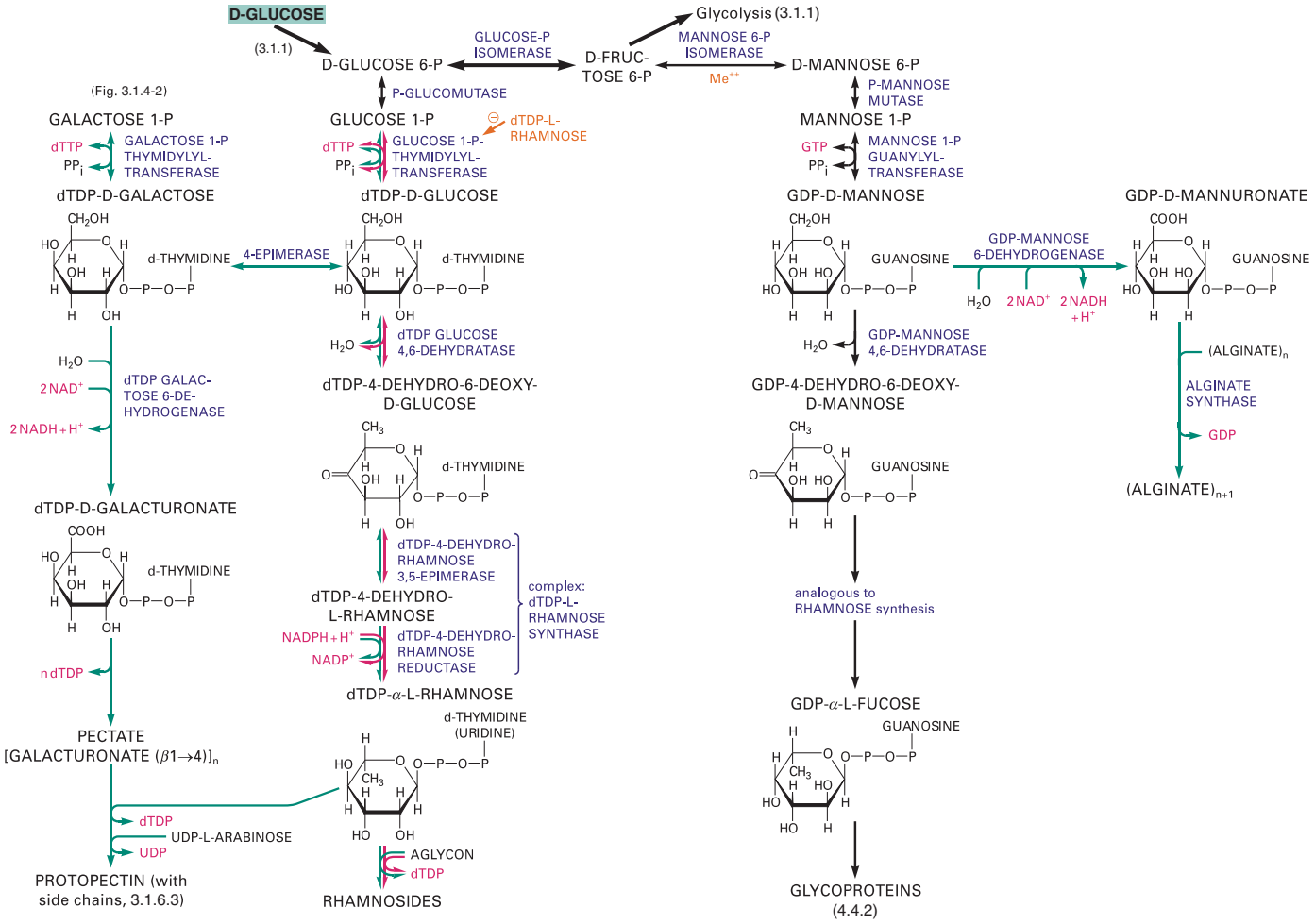


Figure 3.1.5-2. Mannose and Deoxy Hexoses

them may have a storage function. For degradation, myo-inositol can be oxidized to glucuronate.

3.1.5.5 Hexitols

Many aldoses and ketoses can be reduced by NAD^+ or $NADP^+$ dependent reactions to the corresponding linear alcohols. Hexitols (C_6, e.g., sorbitol, Fig. 3.1-1), as well as pentitols (C_5, Fig. 3.1.6-3) frequently occur in plants. In human nutrition, they are used as food additives due to their sweet taste; their metabolism in humans starts with reconversion to the respective sugars.

3.1.5.6 Mannose and Deoxy Hexoses (Fig. 3.1.5-2)

Isomerization of fructose 6-P by the respective enzymes results in either glucose 6-P (3.1.1.1) or in the other epimer, mannose 6-P. Conversion to mannose 1-P and further on to GDP-mannose yields the activated sugar as a precursor of glycoproteins and glycolipids (4.4.1, 4.4.2) as well as of mannuronates (e.g., alginate, 3.1.5.1).

The biosyntheses of L-rhamnose (6-deoxy-L-mannose) and L-fucose (6-deoxy-L-galactose) proceed via dehydration, epimerization and consecutive reduction of dTDP-glucose and of GDP-mannose, respectively (Fig. 3.1.5-2). L-Rhamnose combines with polymeric galacturonate (pectate) to form rhamnogalacturonan, which is essential for the formation of primary plant cell walls (3.1.6.3). L-Rhamnose is also present in bacterial cell walls (3.10.1). L-Fucose is an important component of many glycoproteins (4.4.1 and 4.4.2).

3.1.6 Pentose Metabolism

Pentoses are essential parts of nucleic acids and nucleotides, glycoproteins, plant cell walls etc. They are usually generated via hexose intermediates. The formation of deoxyribose (present in DNA) occurs by reduction of ribonucleotides and is described in 3.6.1.4.

3.1.6.1 Pentose Phosphate Cycle (Fig. 3.1.6-1)

The pentose phosphate cycle is a pathway of glucose turnover alternative to glycolysis and occurs in most species. Its major function is the production of reducing equivalents (in form of NADPH) and of pentoses and tetroses for biosynthetic reactions (nucleoside and amino acid syntheses) in variable ratios (see below). It has been estimated that more than 10 % of glucose is shuttled through the pentose phosphate cycle.

The enzymes are present in the cytosol and plant plastids. In humans, major activities are found in liver, adipose tissue, lactating mammary glands, adrenal cortex, testis, thyroid gland and erythrocytes.

Glucose 6-P is converted into ribulose 5-P by two dehydrogenase reactions (yielding 2 NADPH) and a decarboxylation step. The initial glucose-6-P dehydrogenase is the regulated enzyme. Its activity depends on the $NADP^+$ concentration, NADPH inhibits. Isomerization and epimerization reactions of ribulose 5-P yield ribose 5-P and xylulose 5-P, respectively.

Ribose 5-P can be used for biosynthetic purposes. Otherwise, C_2-units ('active glycolaldehyde') are moved by transketolase (TK) and a C_3-unit by transaldolase (TA), resulting in C_4 and C_7 intermediates and finally in fructose 6-P (glucose 6-P) and glyceraldehyde 3-P. The latter compound can either enter glycolysis or be reconverted into glucose 6-P. The stoichiometry for this reconversion is

$$6 \text{ glucose 6-P} + 12\ NADP^+ + 7\ H_2O = 12\ NADPH + 12\ H^+ + 6\ CO_2 + 5 \text{ glucose 6-P} + P_i.$$

Thus, the pentose cycle can meet different requirements of metabolism. If there is excessive demand for pentoses, the TK and TA reactions of the pentose cycle can run backwards.

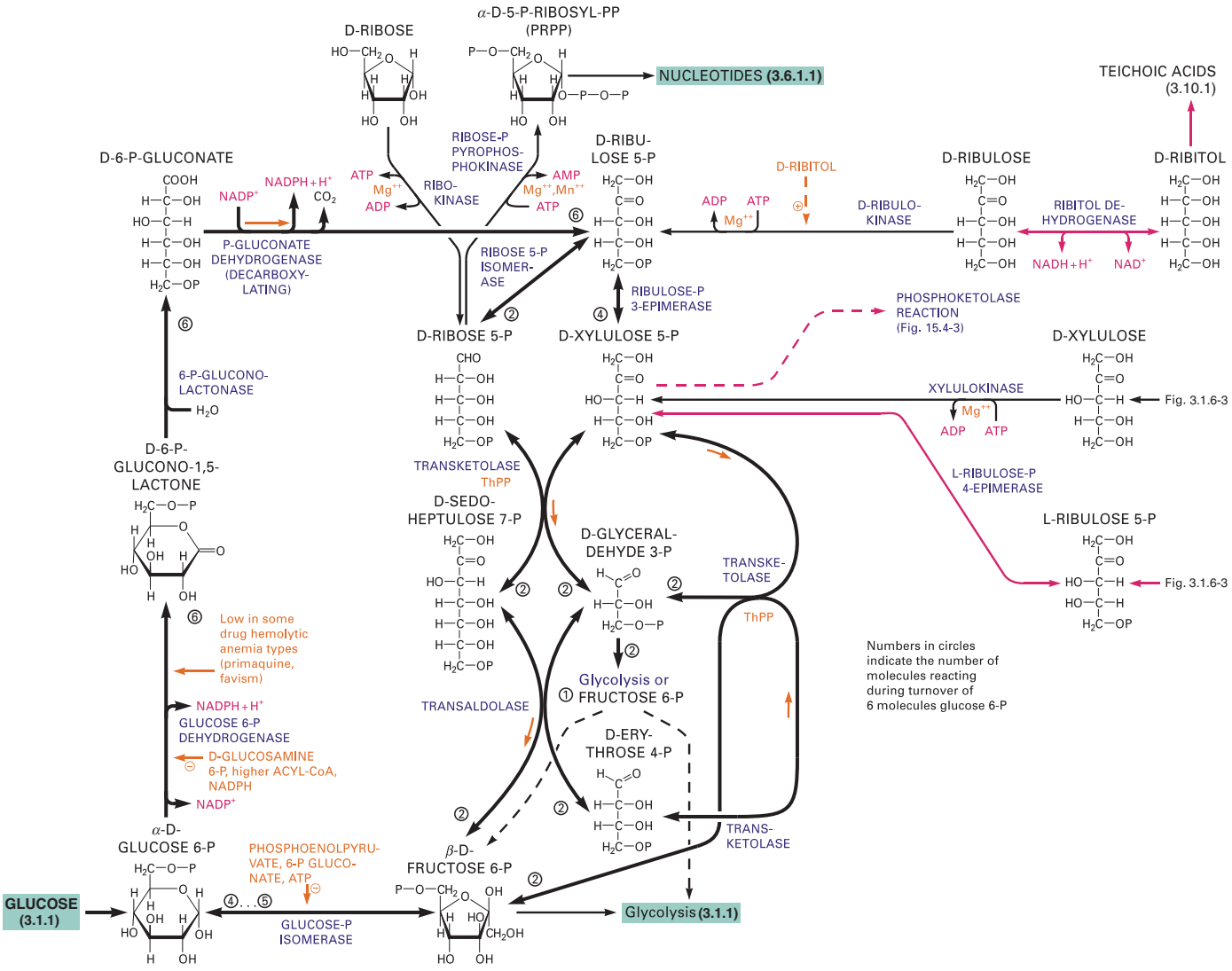


Figure 3.1.6-1. Pentose Phosphate Cycle

The TK reaction requires thiamine pyrophosphate (ThPP) as a coenzyme. It reacts with the keto moiety of the substrate (xylulose 5-P or sedoheptulose 7-P) similarly to the mechanism of the pyruvate decarboxylase reaction (Fig. 3.1.3-4). Following cleavage, the remaining activated intermediate (in this case the 1,2-dihydroxyethyl residue) is transferred. For details, see 3.7.2.2.

TA does not require a coenzyme. It starts the reaction by forming a Schiff base between a ε-lysine group of the enzyme with the keto moiety of the substrate (sedoheptulose 7-P). This leads to an aldol cleavage, releasing erythrose 4-P. The remaining activated residue is then transferred to the acceptor glyceraldehyde 3-P, resulting in fructose 1,6-P_2. In the otherwise analogous fructose-bisphosphate aldolase reaction (3.1.1.1) the activated residue (dihydroxyacetone-P) is released after protonation (Fig. 3.1.6-2).

In humans, deficiency of glucose-6-P dehydrogenase (and thus of NADPH which is essential for glutathione regeneration, 3.2.5.7) causes hemolytic anemia after administration of some drugs (e.g., primaquine), which produce elevated H_2O_2 levels.

In bacteria, the phosphoketolase cleavage of xylulose 5-P leads to lactate and ethanol (Fig. 3.10.5-3). The reduction of D-ribulose yields ribitol, which is an essential component of teichoic acids (3.10.1).

3.1.6.2 Other Decarboxylation Reactions (Fig. 3.6-3)

Pentoses can be also formed from UDP-derivatives of uronic acids (UDP-glucuronate, UDP-galacturonate) by decarboxylation. This pathway is especially prominent in plants; the products (e.g., L-arabinose, D-xylose) occur in their cell walls (3.1.6.3).

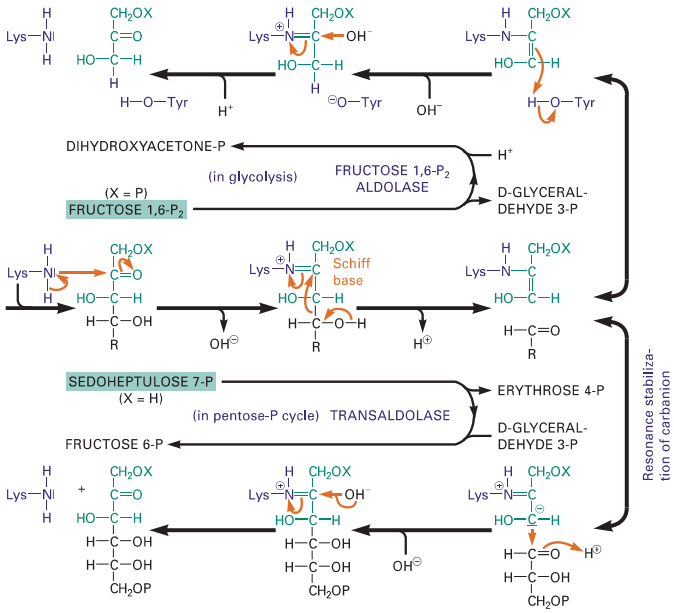


Figure 3.1.6-2. Aldol Cleavage Reactions of Carbohydrates

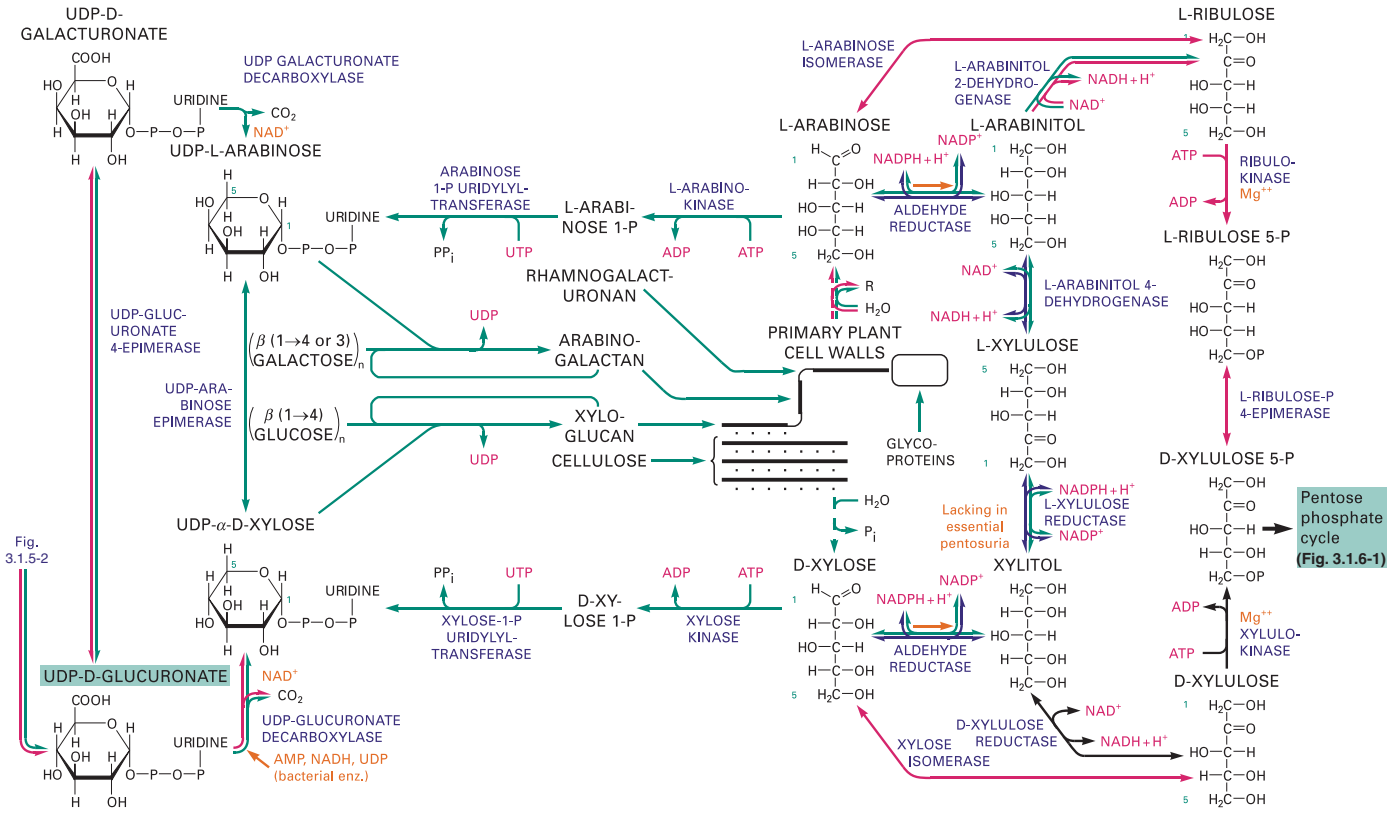


Figure 3.1.6-3. Formation of C_5 Compounds by Decarboxylation and Cell Wall Synthesis in Plants

The oxidation products of L-gulonate and its lactone (3-dehydro-L-gulonate and of L-ascorbate = vitamin C, respectively) yield upon decarboxylation C_5 compounds (L-xylulose, L-xylonate and L-lyxonate) or are metabolized differently (3.7.10).

3.1.6.3 Plant Cell Walls

Primary plant cell walls show a wide variation of composition. In a typical example (*Sycamore* maple), they consist of roughly similar quantities of rhamnogalacturonan (3.1.5.6), arabinogalactan, xyloglucan, cellulose (3.1.2.2) and glycoproteins (4.4.2.1). In fungi and in the animal kingdom (arthropods), cellulose is replaced by chitin (linear polymer of N-acetylglucosamine, 3.1.7.1).

Cell wall synthesis (reactions are schematically shown in Fig. 3.1.6-3, structures in Figs. 3.1.6-4 and 3.1.6-5): After cell division, the synthesis of primary cell walls starts at a middle lamella consisting of protopectin (rhamnogalacturonan, Fig. 3.1.5-2). This structure contains many negative charges due to its galacturonic acid content. It binds Ca^{++} and Mg^{++} and is highly hydrated. Hemicelluloses (xylans, xyloglucan, derivatives of mannose, galactose, fucose etc.) and arabinogalactan are produced by dictyosomes (plant Golgi apparatus) and become attached by covalent bonds. Glycoproteins (with up to 30% hydroxyproline) are located between the hemicellulose molecules. Then the cellulose fibers are synthesized by enzymes located at the plasma membrane (3.1.2.2). The single straight β-glucose chains of cellulose are combined by hydrogen bonds in overlapping fashion, forming a fiber bundle of about 70 … 100 neighboring chains with a partially crystalline structure. This bundle gets attached to the hemicelluloses by many hydrogen bonds and provides the tensile strength. Plasmodesmata (plasma-membrane lined channels of about 5 nm diameter) cross the cell wall and allow movement of fluids and metabolites.

The secondary cell wall is formed from the primary one by thickening due to deposition of additional cellulose layers and of lignin (3.13.1.2), which fills the spaces between the fibers (analogously to reinforced concrete). Lignin synthesis originates at erythrose 4-P from the pentose phosphate cycle.

Degradation of wood: The complex structure requires many enzymes. They are mostly of bacterial or fungal origin, among them cellulase (EC 3.2.1.4) for cellulose; polygalacturonase (EC 3.2.1.15)

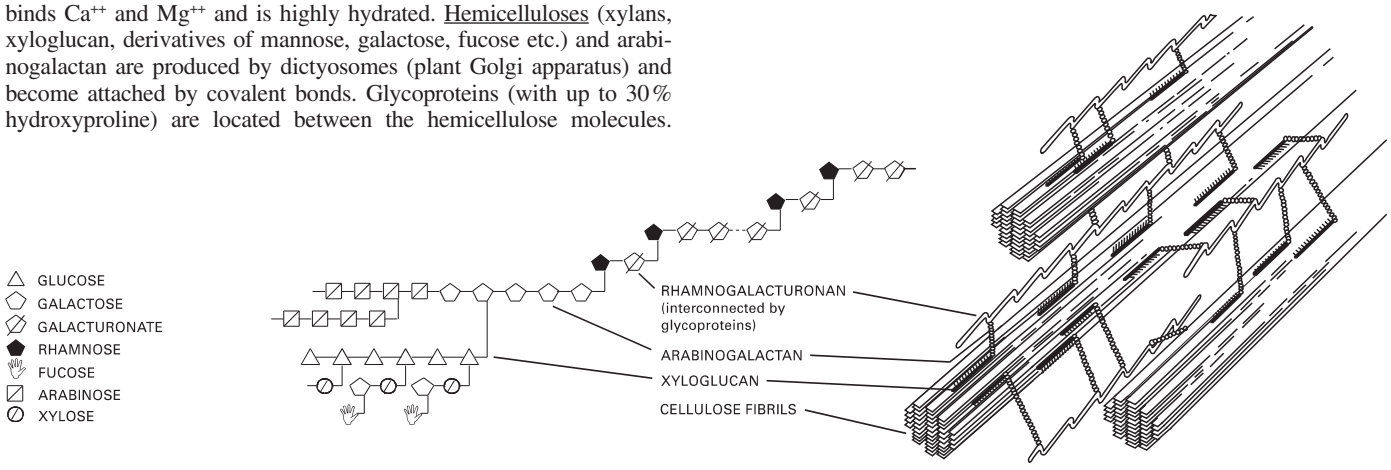


Figure 3.1.6-4. Components of Plant Cell Walls

Figure 3.1.6-5. A Structure Model of Primary Cell Walls in Plants (modified from Strasburger: *Botanik*).

and α- and β-rhamnosidases (EC 3.2.1.40 and 43) for pectins; arabinogalactan-endo-galactosidases (EC 3.2.1.89 and 90) and arabinan-endo-arabinosidase (EC 3.2.1.99) for arabinosides; xylan xylosidase (EC 3.2.1.37 and 72) for xylans. A lignin degrading enzyme is lignostilbene α,β-dioxygenase (EC 1.13.11.43).

3.1.6.4 Pentose Metabolism in Humans

Dietary L-arabinose in humans is metabolized by intestinal bacteria via L-ribulose to D-xylulose 5-phosphate, which is part of the pentose phosphate cycle (3.1.6.1). D-xylose is also converted to D-xylulose 5-phosphate. High pentose content of food, as well as a low inherited enzyme activity lead to pentosuria, a condition where unusually high concentrations of xylulose is found in the urine.

3.1.7 Amino Sugars

Amino sugars are sugar derivatives, in which a hydroxyl group has been replaced by an amino group. Frequently, they are acetylated. Other prominent derivatives are, e.g., N-acetylneuraminic acid (sialic acid) and N-acetylmuramic acid. They play an important role in glycoproteins and glycolipids, as components of cell walls and of exoskeletons (see below).

3.1.7.1 Biosynthesis (Fig. 3.1.7-1)

The biosynthesis of amino sugars starts from D-fructose 6-P by a transamination reaction with glutamine, yielding D-glucosamine 6-P. Apparently, a Schiff base is the intermediate. Thereafter, acetylation by acetyl-CoA takes place, followed by a mutase reaction to yield N-acetylglucosamine 1-P, and by the formation of the UDP-derivative. UDP-N-acetyl-D-glucosamine is the substrate for epimerization reactions at the 4-position (yielding UDP-N-acetylgalactosamine) or at the 2-position (yielding UDP-N-acetylmannosamine, which, however, is immediately hydrolyzed to N-acetyl-D-mannosamine). Acetylation reactions can also take place with unphosphorylated amino sugars.

Both UDP-N-acetylglucosamine and UDP-N-acetylgalactosamine take part in the synthesis of many glycoproteins and glycolipids in animals and plants (4.4.1, 4.4.2).

The linear polymerization of UDP-N-acetylglucosamine by chitin synthase yields chitin [$(\beta 1 \rightarrow 4\text{-glucosamine})_n$], which is structurally very similar to cellulose (3.1.2.2). The interaction of neighboring molecules and therefore the strength of the structure even surpasses cellulose. Chitin is a major component of the exoskeleton of invertebrates (crustaceans, insects), it also occurs in the cell walls of many fungi and algae. The chitin fibers provide strength to the elastic framework of glycoproteins or glycans. Deacetylation of chitin yields chitosan, which also occurs in cell walls of fungi.

The condensation of UDP-N-acetylglucosamine with phosphoenolpyruvate and the consecutive reduction results in N-acetylmuramate, the starting compound for the formation of murein (3.10.1), which is an essential component of bacterial cell envelopes.

The synthesis takes place at undecaprenyl anchors by adding amino acids (part of them in the D-configuration, which renders them resistant to common proteases) and N-acetylglucosamine. Both amino sugar derivatives alternate in the backbone chain, while the amino acids interconnect these chains (3.10.1).

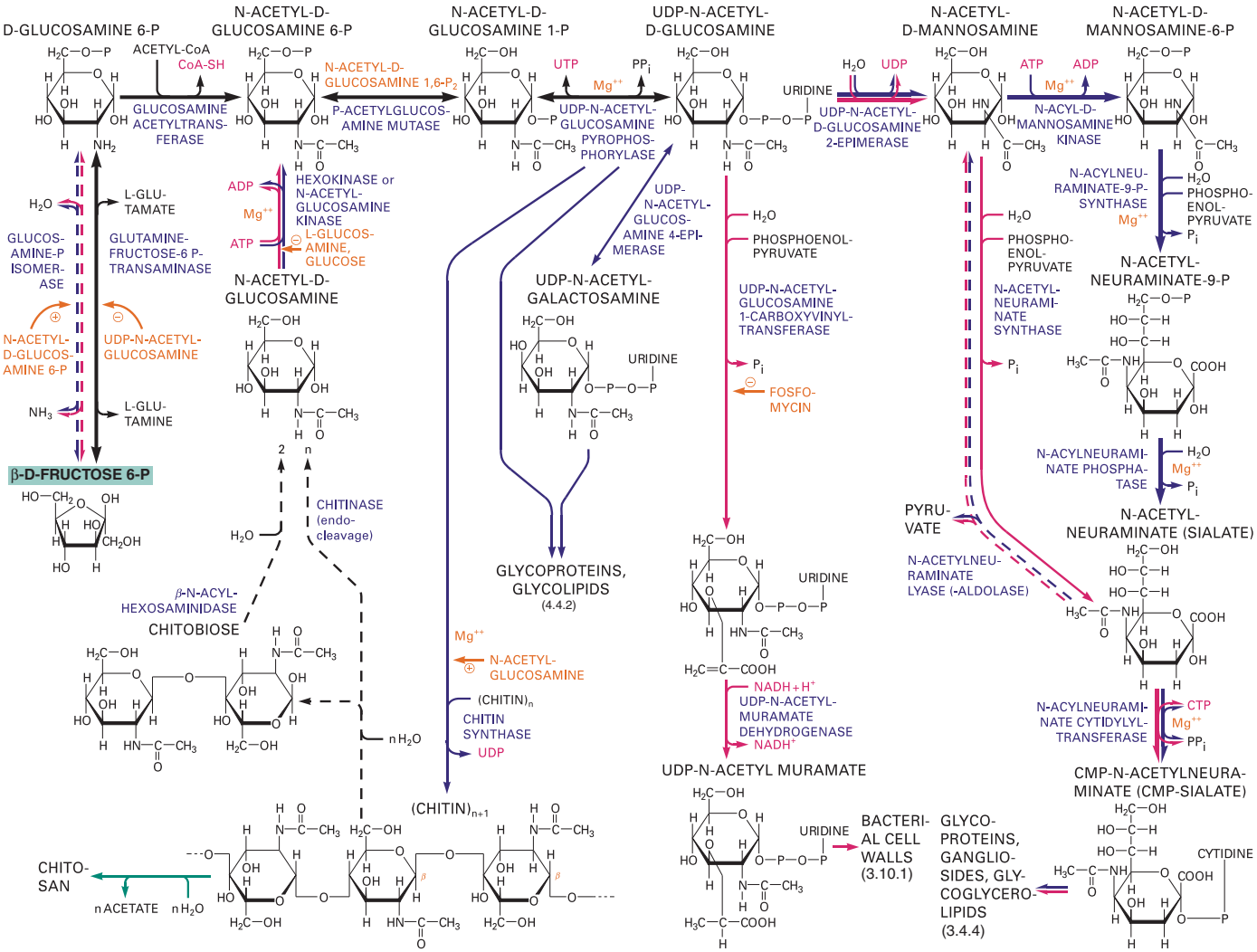


Figure 3.1.7-1. Metabolism of Amino Sugars

Either N-acetyl-D-mannosamine or its 6-phosphate can be condensed with phosphoenolpyruvate to yield N-acetylneuraminate (sialate) or its 6-phosphate, respectively, which is thereafter dephosphorylated. The ring form of this compound includes the 3 added C-atoms. Sialate is an essential component of gangliosides (3.4.4.1). For introduction into these compounds, it is activated as CMP-derivative.

3.1.7.2 Catabolism

Murein is split by lysozyme between the N-acetylmuramate and N-acetylglucosamine residues, yielding disaccharides. Lysozyme is present in tears, sneeze mucus and other body secretions as well as in bird eggs (protective function).

The degradation of chitin can take place by chitinase, an endo (β 1→ 4) glucosaminidase, which performs random hydrolysis, but also by lysozyme. Remaining chitobiose is further cleaved by β-n-acyl-hexosaminidase.

Terminal neuraminic acid in glycoproteins and glycolipids is removed by neuraminidase (4.4.2.1). This is frequently the first step of degradation.

Glucosamine 6-P can be reconverted into fructose 6-P by an isomerase reaction, NH_3 is liberated.

Literature to 3.1.4 … 3.1.7:

Heldt, H.W. *Plant Biochemistry and Molecular Biology*. Oxford University Press, 1998.

Huber, S.C. *et al*. Int. Reviews of Cytology 1994;149:47–98.

Kretchmer, M. Scientific American 1972;227 (4):74–78.

Smith, H.S. *et al*. Plant Physiology 1995;107:673–677.

Ruoslahti, E. Ann. Rev. of Cell Biol. 1988;4:229–255.

Varner, J.E., Lin, L.S. Cell 1989;56:231–239.

(General) Textbooks of organic chemistry.

Heldt, H.W. *Plant Biochemistry and Molecular Biology*. Oxford University Press, 1998.

Lewis, N.G., Yamamoto, L. Ann. Rev. of Plant Physiol. and Plant Mol. Biol. 1990;41:455–496.

Raven, P.H. *et al*. *Biology of plants*. 5th Ed. Worth, 1992.

Strasburger,E. *Lehrbuch der Botanik für Hochschulen* 36. Aufl. Gustav Fischer Verlag, 2008.

Vaseghi, S. *et al*. Metab. Eng. 1999;2:128–40.

Kumagai, I. *et al.* J. Biol. Chem. 1992;267:4608–4612.

Muzzarelli, R.A. *Chitin*. Pergamon Press, 1977.

3.1.8 Citrate Cycle

The citrate cycle (Krebs cycle, tricarboxylic acid cycle) plays the central role in metabolism both of eukarya and most prokarya (Fig. 3.1.8-1). It is the major site of oxidation of carbon chains from carbohydrates, fatty acids (both entering via acetyl-CoA) and many amino acids to CO_2 and water. It supplies NADH as essential substrate for the oxidative phosphorylation in the respiratory chain (3.11.4) and thus has a major role in energy metabolism. It also provides intermediates for the synthesis of amino acids (3.2.2.2, 3.2.2.4) and porphyrins (3.3.1). For the anaplerotic reactions, which are required to replace cycle compounds removed for biosynthesis, see 3.1.3.4.

The cycle operates only under aerobic conditions, since it requires the quick reconstitution of NAD^+ from NADH by the respiratory chain (3.11.4).

$$\text{acetyl-CoA} + 3\ NAD^+ + FAD + GDP\ (\text{or } ADP) + P_i + 2\ H_2O = 2\ CO_2 + 3\ NADH + FADH_2 + GTP\ (\text{or } ATP) + 2\ H^+ + CoA$$

$$\Delta G'_0 = -41\ \text{kJ/mol}$$

In green sulfur bacteria (*Chlorobiaceae*), some other bacteria and archaea, the 'reductive citrate cycle' is operative, which yields acetyl-CoA by autotrophic CO_2 fixation under anaerobic conditions. It resembles the citrate cycle running backwards and substitutes the Calvin cycle. Reduced ferredoxin serves as the main reductant for CO_2 fixation. It is speculated that inorganic catalysis of analogous reactions might have played a role in the origin of life on earth.

3.1.8.1 Reaction Sequence (Fig. 3.1.8-2)

In eukarya, the enzymes of the cycle are located in the mitochondrial matrix, only succinate dehydrogenase is part of the inner mitochondrial membrane. Some of these enzymes also exist in the cytosol, fulfilling other tasks, e.g., as members of the urea cycle, 3.2.9.1.

Citrate is formed by a condensation reaction of acetyl-CoA and oxaloacetate, catalyzed by citrate synthase. The structure of this enzyme favors the initial condensation reaction to citryl-CoA by an ordered sequential mechanism (1.5.4): an 'open form' binds oxaloacetate; the conformation change to the 'closed' form generates the acetyl-CoA binding site. After condensation, hydrolytic removal of coenzyme A takes place.

Acetyl-CoA cannot pass the mitochondrial membrane, it is converted by citrate synthase into citrate, which is exported via an antiport carrier and is cleaved in the cytosol by the ATP-energized citrate lyase. While the acetyl-CoA is metabolized, oxaloacetate is reduced to malate or is converted into pyruvate, which returns to the mitochondrion by a carrier. Another citrate lyase reaction occurs in citrate fermentation (Fig. 3.10.5-3).

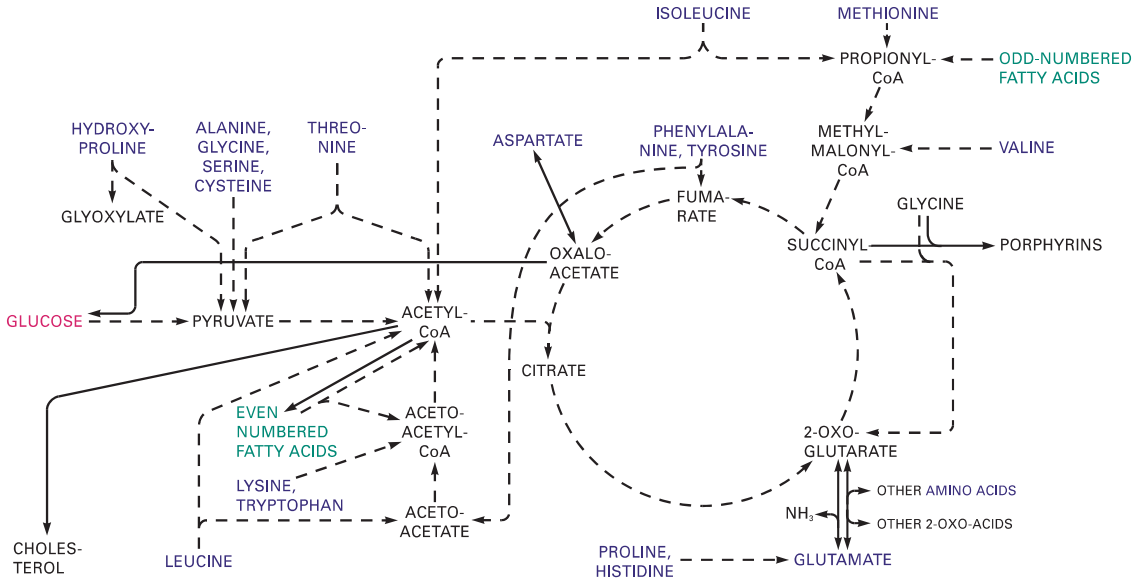


Figure 3.1.8-1. Interrelations of the Citrate Cycle With Other Metabolic Pathways (blue: amino acids, red: carbohydrates, green: lipids)

Cis-aconitate and oxalosuccinate are labile intermediates of the aconitate hydratase (aconitase) and the isocitrate dehydrogenase reactions, respectively. They remain bound to the enzymes. The NAD^+-dependent mitochondrial form of isocitrate dehydrogenase is a member of the citrate cycle (in mammals), while the $NADP^+$-dependent form also occurs in cytosol and is responsible for NADPH and oxoglutarate supply (e.g., for conversion to glutamate). The 2-oxoglutarate dehydrogenase enzyme complex is very similar in its structure to the pyruvate dehydrogenase complex (3.1.3.1) and uses the same coenzymes. However, no regulation of its activity by phosphorylation and dephosphorylation takes place. Succinyl-CoA ligase (-synthetase) uses the energy of succinyl-CoA hydrolysis for the formation of a high-energy phosphate bond (GTP in mammals, ATP in plants and bacteria). The membrane-bound succinate dehydrogenase is a member of the electron transport chain and feeds protons and electrons into the quinone pool (3.11.4). The resulting fumarate is hydrated and oxidized, yielding oxaloacetate, whose condensation with acetyl CoA yields citrate and closes the cycle.

The interconversion of malate and oxaloacetate is also used for the initiation of gluconeogenesis to circumvent the irreversible pyruvate kinase reaction (3.1.3.5). It is a means for transport through membranes (Figs. 3.1.3-3, 3.4.1-1, 3.11-2, 3.12-10).

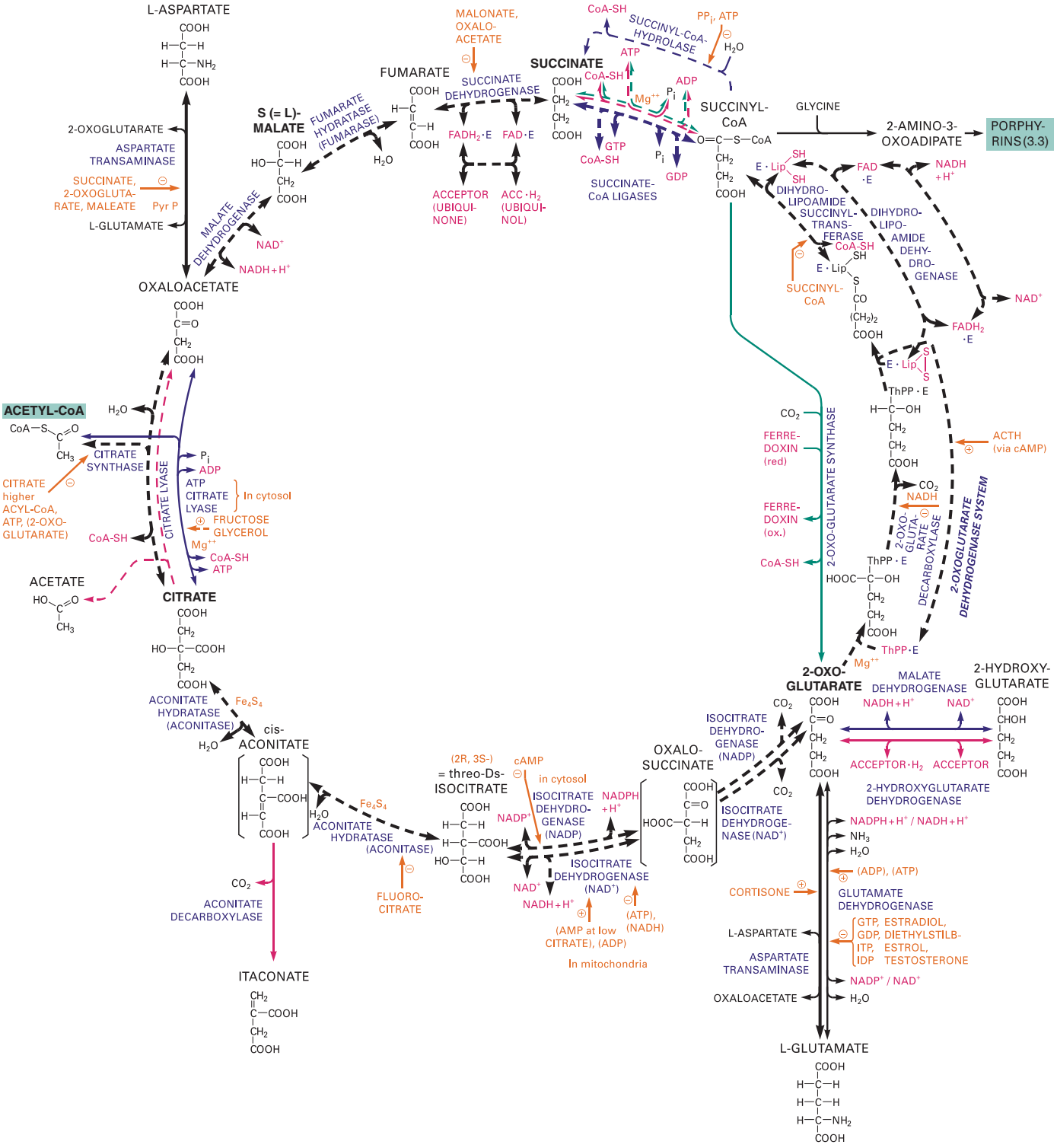


Figure 3.1.8-2. Citrate Cycle

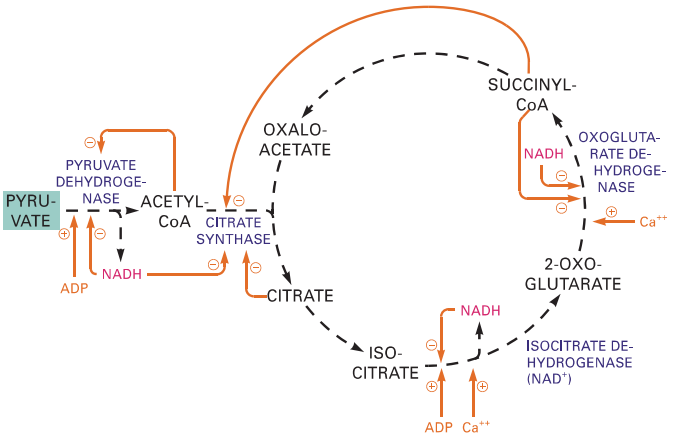


Figure 3.1.8-3. Regulation of the Citrate Cycle

3.1.8.2 Regulatory Mechanisms (Fig. 3.1.8-3)

The citrate cycle in muscle is regulated in a relatively simple way by substrates and products. The regulated enzymes are citrate synthase, isocitrate dehydrogenase and oxoglutarate dehydrogenase. For the regulation of pyruvate dehydrogenase, which provides the 'fuel' acetyl-CoA, see 3.1.3.2.

As described for glycolysis (3.1.1.2), the ΔG_{phys} of the regulated enzymes is strongly negative. Major mechanisms are the inhibitions by NADH (control by the redox state) and by ATP, the activations by ADP (control by the energy state of the cell) and by Ca^{++} (in muscles, as a result of muscular activity). Some feedback inhibitions by products also exist. Since usually the *in vivo* concentrations of acetyl-CoA and oxaloacetate do not saturate the citrate synthase, this enzyme is additionally regulated by the availability of the substrates.

The quantity of circulating compounds changes when compounds are removed for biosynthesis or when citric cycle intermediates are generated by catabolic reactions (e.g., by transaminations of glutamate or aspartate (3.2.2.2, 3.2.2.4), formation of succinyl-CoA by degradation of odd-numbered fatty acids (3.4.1.5), or of branched-chain amino acids, 3.2.6.2). This is taken care of by control of anaplerotic reactions (3.1.3.4), prevalently in liver.

In bacteria (*E. coli*), the isocitrate dehydrogenase becomes inactivated by phosphorylation of the active site serine-113. Citric acid cycle regulation in plants is not well understood. Presumably, turnover rates depend on the rates of NADH reoxidation and ATP utilization.

3.1.8.3 Energy Balance

The chemical combustion of pyruvate yields $\Delta G'_0 = -1145$ kJ/mol, while the conversion of pyruvate to acetyl-CoA ($\Delta G'_0 = -34$ kJ/mol) and the oxidation of acetyl-CoA in the citrate cycle (see above) combined, decrease the $\Delta G'_0$ only by −75 kJ/mol. Thus, the major part of the free energy of the oxidation reaction (ca. 94 %) is stored in NADH, $FADH_2$ and GTP (or ATP). 4 NADH are formed in the pyruvate-, isocitrate-, 2-oxoglutarate- and malate dehydrogenase reactions, 1 $FADH_2$ in the succinate dehydrogenase reaction, 1 GTP (in animals) or ATP (in plants or microorganisms) in the succinate-CoA ligase (= succinyl-CoA synthetase) reaction.

In the respiratory chain (3.11.4), additional ATP is produced by oxidation of NADH (2.5 cytosolic ATP/1 NADH) and $FADH_2$ (1.5 cytosolic ATP/1 $FADH_2$), yielding a maximum of 4 * 2.5 + 1.5 + 1 = 12.5 mol ATP/mol pyruvate. The free energy conserved in these bonds are $\Delta G'_0 = 12.5 \times 30.5$ kJ = 381 kJ, resulting in an energy yield of about 33 %. For conversion of glucose, the ATP and NADH yield of glycolysis has to be added (3.1.1.1).

Frequently in the literature, a maximum yield of 3 ATP/1 NADH and a total yield of 15 mol ATP/mol pyruvate (or 38 mol ATP/mol glucose) are stated. This figure refers to the ATP obtained within the mitochondria, while the figures above also consider the energy requirement of ATP export into the cytosol (3.11.3). All of these values are only approximate, since in the living cell the actual conditions differ from the standard conditions and the efficiency of the oxidative phosphorylation varies according to the metabolic situation.

3.1.9 Glyoxylate Metabolism

3.1.9.1 Glyoxylate Cycle (Fig. 3.1.9-1)

Enzymes of the glyoxylate cycle have been found in bacteria, plants, insects and vertebrates including humans. The major purpose of this cycle in plants is the conversion of acetyl-CoA (from fat degradation) into succinate and further into malate for gluconeogenesis (see 3.1.1.3). The specific enzymes of this cycle occur in glyoxisomes, a special type of peroxisomes mostly present in tissues, where fat conversion to carbohydrates is important (e.g., in seed during germination).

Acetyl-CoA combines with oxaloacetate and is converted into isocitrate analogously to the citrate cycle reactions. The consecutive cleavage reaction by isocitrate lyase yields succinate and glyoxylate. The latter compound condenses with another acetyl-CoA to malate. This reaction is catalyzed by malate synthase, which resembles citrate synthase. Malate is oxidized to oxaloacetate, which can start another round of the glyoxylate cycle. Succinate, however, passes over to the mitochondria and is converted into oxaloacetate by the citric acid cycle enzymes. These compounds leave the mitochondria via translocator mechanisms and initiate the gluconeogenesis pathway in the cytosol (3.1.3.5). The overall reaction is

$$2\ \text{acetyl-CoA} + 2\ NAD^+ + FAD = \text{oxaloacetate} + 2\ NADH + FADH_2 + 2\ H^+.$$

Some authors assume, that malate from the glyoxisomes enters the cytosol, while the succinate, which is imported into the mitochondria, is converted into oxaloacetate, transaminated into aspartate, transported via a shuttle mechanism to the glyoxisomes and reconverted into oxaloacetate.

Similar reactions also take place in the cytoplasm of bacteria. Their purpose is mainly the consumption of acetate, which is taken up from the medium and converted into acetyl-CoA before entering this pathway. In *Escherichia coli*, isocitrate dehydrogenase (which competes with isocitrate lyase for the substrate in the same compartment) is inactivated by phosphorylation and activated by dephosphorylation of the enzyme (3.1.8.2). The modifying bifunctional kinase/phosphatase

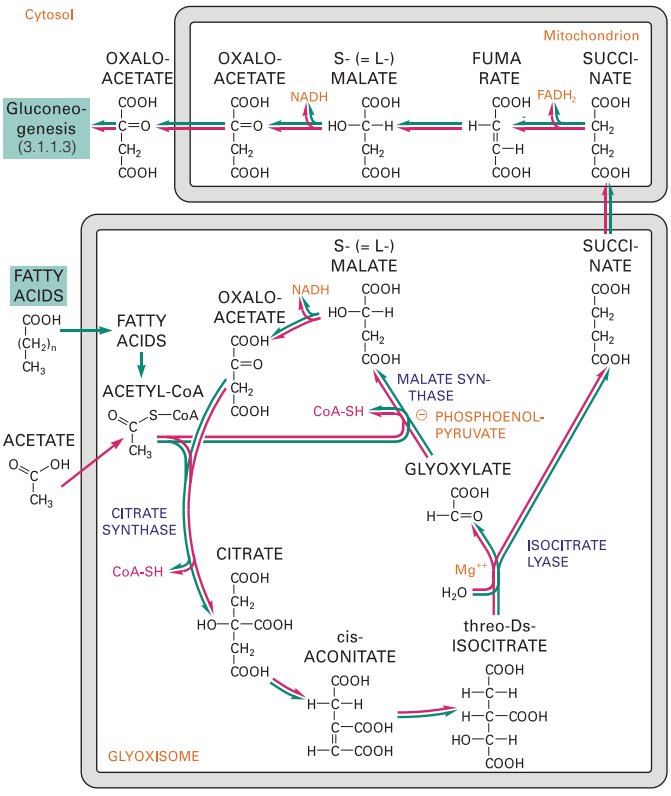


Figure 3.1.9-1. Glyoxylate Cycle

is allosterically regulated by intermediates of the citrate cycle, which also allosterically regulate the isocitrate lyase in the opposite direction. These mechanisms determine the share of isocitrate being converted by the citrate and glyoxylate cycles.

3.1.9.2 Other Glyoxylate Reactions

In photosynthesizing plants, the oxygenase side activity of rubisco yields 2-phosphoglycolate (Fig. 3.12-8), which after dephosphorylation leaves the chloroplasts via a translocator mechanism and enters the peroxisomes (Fig. 3.1.9-2). There it is oxidized to glyoxylate, the H_2O_2 formed is decomposed by catalase. A transamination reaction yields glycine. The complicated condensation, oxidation and partial decarboxylation to serine takes place in mitochondria by the multienzyme complex glycine cleavage enzyme and glycine hydroxymethyltransferase (3.2.4.2). Serine is converted via further steps into phosphoglycerate, which reenters the Calvin cycle (3.12.2). However, this 'photorespiration' detour causes a loss of energy (1 decarboxylation, 1 dephosphorylation).

The oxidation of glyoxylate to oxalate and further on to formate and CO_2 occurs in plants (as a side activity of the peroxisomes) and in bacteria (Fig. 3.1.9-3).

The location of the complete glyoxylate and H_2O_2 metabolism in glyoxisomes/peroxisomes protects the rest of the cell from these highly toxic compounds.

Literature to 3.1.8 … 3.1.9:

Perham, R.N. Biochemistry 1991;30:8501–8512.

Remington, S.J. Curr. Top. Cell Reg. 1992;33:202–229.

Wächtershäuser, G. Proc. Natl. Acad. Sci (USA) 1990;87:200–204.

Walsh, C. *Enzymatic Reaction Mechanisms*. Freeman, 1979.

Douce, R. Biochem. Soc. Transactions 1994;22:184–188.

Gerhardt, B. Progress Lipid Res. 1992;31:417–446.

Husic, D.W. *et al. CRC Crit. Revs. in Plant Sciences* Vol. 1, 1987, p. 45–100.

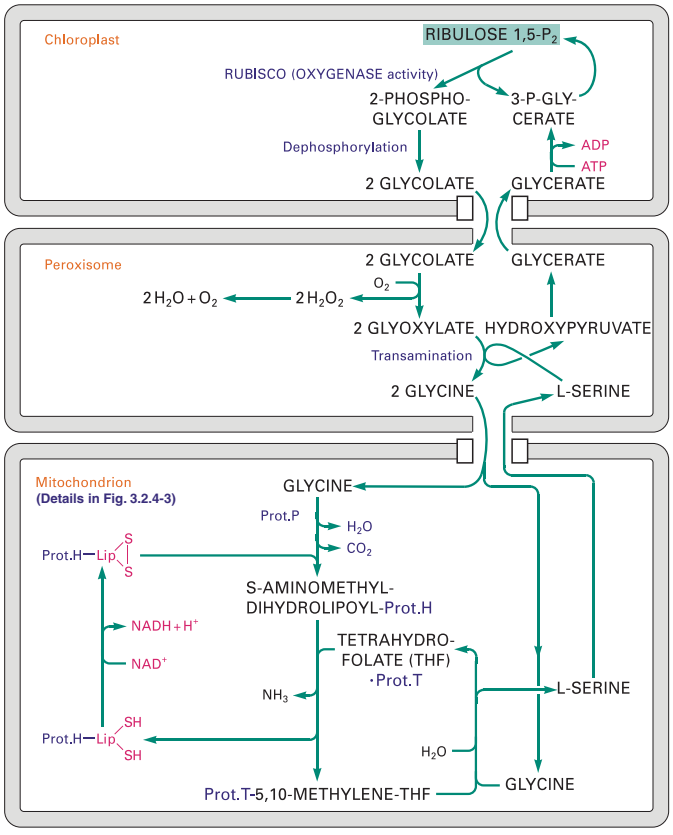


Figure 3.1.9-2. Recycling of Photorespiration Products

Holms, W.H. Curr. Top. Cell. Regul. 1986;28:69–106.

Moore, T.S. *Lipid Metabolism in Plants*. CRC Press, 1993.

Popov, V. N. *et al*. J. Evol. Biochem. Phys. 2005;41:631–639.

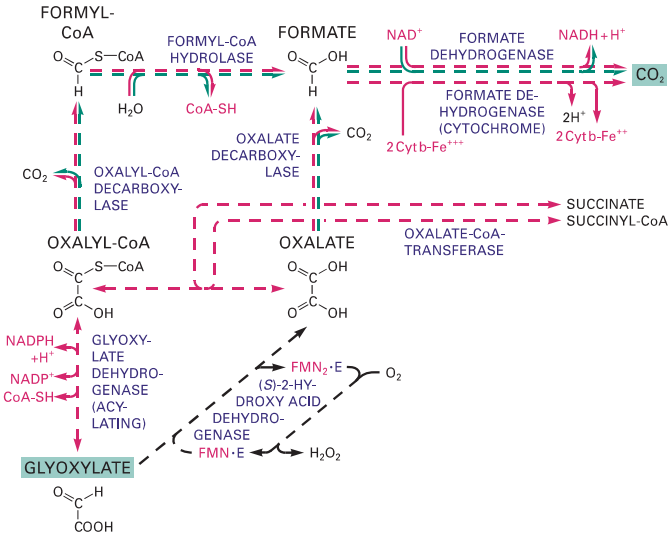


Figure 3.1.9-3. Oxidation of Glyoxylate

3.2 Amino Acids and Derivatives

Röbbe Wünschiers

3.2.1 Nitrogen Fixation and Metabolism

Most of the nitrogen in the biosphere is bound in amino acids and nucleotides (which are formed from amino acids). Since only a small number of bacteria are able to convert atmospheric nitrogen into utilizable compounds (primarily NH_3/NH_4^+), plants, especially, had to adapt themselves to a limited nitrogen supply.

Nitrogen fixation is performed by bacteria living as symbionts with plants (e.g., *Rhizobium*) as well as by some free-living bacteria (e.g., *Azotobacter*, *Cyanobacteria*). About 7 … 11 × 10^7 t/year are converted in this way (as compared with 3 × 10^7 t / year by industrial means). Cleavage of the excessively stable triple bond of N_2 ($\Delta G'_0$ = 945 kJ/mol) by nitrogenase requires a high energy input in the form of ATP and reducing equivalents:

$$N_2 + 8\,H^+ + 8\,e^- + 16\,ATP + 16\,H_2O = 2\,NH_3 + H_2 + 16\,ADP + 16\,P_i.$$

Nitrogenase is composed of 2 parts:

- a homodimeric Fe protein (2 * 32 kDa) with 2 ATP binding sites and a 4Fe-4S cluster located between both subunits (dinitrogenase-reductase)
- a MoFe protein [220 kDa, $(\alpha\beta)_2$, actual dinitrogenase]. Each αβ heterodimer contains two 4Fe-4S 'P'-clusters linked with each other (located between the α and β subunits) and a 4Fe-3S•Mo•3Fe-3S complex of cubane structure linked by 3 S bridges (located in a cavity of the α subunit). In some organisms, Mo can be replaced by V or Fe.

An electron liberated by oxidative or photosynthetic reactions (3.12.1) is transferred via ferredoxin to the Fe-protein and reduces it. In some species, the reductant ferredoxin is replaced by flavodoxin. After a conformation change, which is energized by hydrolysis of 2 ATP, the electron passes on to the MoFe protein. After eight such rounds, the Fe-Mo complex of the MoFe protein is reduced. It is then able to reduce N_2 to 2 NH_3 (Fig. 3.2.1-1) and in a side reaction 2 H^+ to H_2, which then partially counteracts the first step of N_2 reduction. In fact, more than 25 % of the electrons used by nitrogenase are allocated for proton reduction. Most bacteria recycle the produced hydrogen by action of uptake hydrogenases. Anaerobic or at least microaerobic

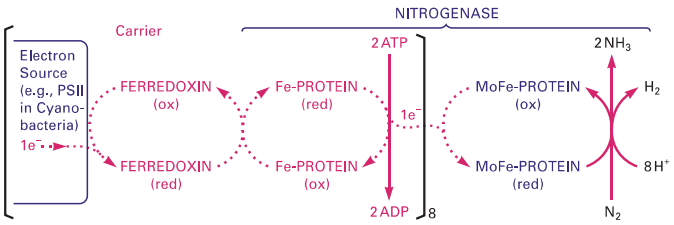


Figure 3.2.1-1. Mechanism of the Nitrogenase Reaction
(Dotted connections indicate electron transfer reactions)

conditions are required. The protection against O_2 in *Rhizobium* takes place through the symbiontic synthesis of leghemoglobin by the host plant, which binds O_2 with high affinity and releases it in limited quantities to the bacterial membrane (location of the respiratory chain), thus avoiding interference with the nitrogenase activity, which is located in the cytoplasm.

Circulation of nitrogen: Ammonia obtained by the nitrogenase reaction or by degradation procedures is introduced into amino acids by glutamate-ammonia ligase (glutamine synthase) or in smaller quantities by glutamate dehydrogenase (3.2.2.2). Both reactions take place in all living beings. Other amino acids are obtained from glutamine and glutamate mostly by transamination reactions (3.2.2). The nucleotide-N is contributed by amino acids (3.6.1.1, 3.6.2.1).

Many soil bacteria derive metabolic energy from the oxidation of ammonia to NO_2 and NO_3^- (nitrification, Fig. 3.10.7-1), while facultative anaerobic bacteria and plants can reduce them again (nitrate ammoniafication, Fig. 3.10.6-1). Alternatively, the reduction leads to nitrogen (denitrification, 3.10.6).

Essential amino acids: While plants and bacteria are able to synthesize all amino acids (3.2.2 … 3.2.9), mammals are unable to synthesize some of them and have to obtain them by food intake (Table 3.2.1-1).

Table 3.2.1-1. Essential and Non-essential Amino Acids for Humans
(histidine is considered to be semi-essential because the body does not always require it)

Essential:	histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, valine
Non-essential:	alanine, arginine, asparagine, aspartate, cysteine, glutamate, glutamine, glycine, proline, serine, tyrosine

Surplus amino acids cannot be stored, they are degraded. In most terrestrial vertebrates, amino-N is converted to urea and excreted (3.2.9.1). Urea can be easily cleaved by bacteria, resulting in ammonia. Ammonia is also the terminal product of amino acid degradation in other species.

Literature:

Deng, H., Hoffmann, R. Angew. Chem Int. Ed. 1993;32:1062–1065.
Kim, J., Rees, D.C. Biochemistry 1994;33:389–397.
Mylona, P. *et al*. The Plant Cell 1995;7:869–885.

3.2.2 Glutamate, Glutamine, Alanine, Aspartate, Asparagine and Ammonia Turnover

All amino acids of this group are connected via transaminations to the citrate cycle or to pyruvate. Since these compounds can be used for gluconeogenesis (3.1.1.3), these amino acids are termed glucogenic. Glutamine is the primary entrance gate for ammonia into bacterial and plant metabolism. In animals, which obtain amino acids by food intake, glutamate plays the central role in amino acid interconversions.

3.2.2.1 Glutamine Metabolism (Fig. 3.2.2-1)

Glutamine synthesis: The ubiquitous glutamate-ammonia ligase (glutamine synthetase) performs the ATP-energized ammonia binding reaction:

Glutamate + ATP + NH_3 = glutamine + ADP+ P_i

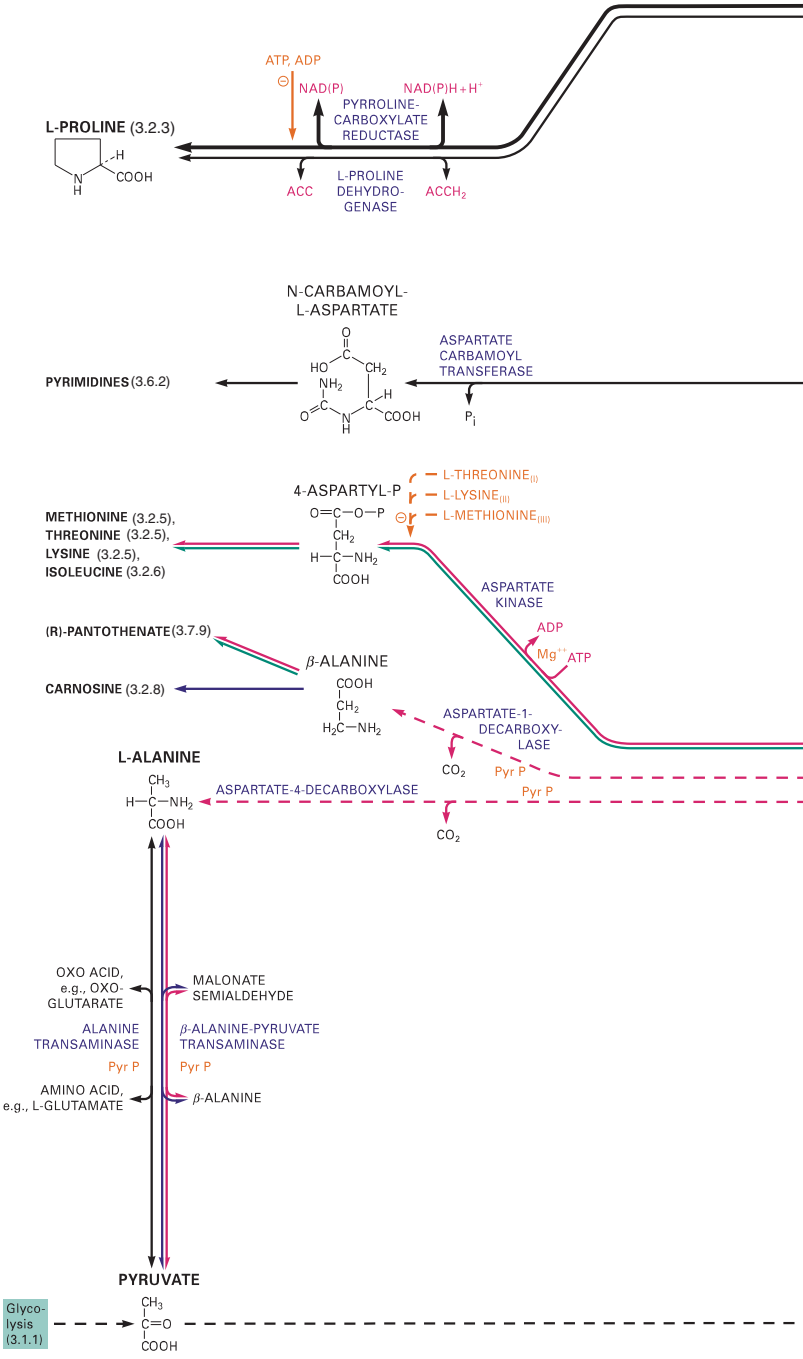
The *E. coli* enzyme consists of 12 identical subunits. Due to its central role in nitrogen metabolism, it is regulated at three levels (Fig. 3.2.2-2, p. 61):

① Many end products of biosynthesis inhibit the ligase allosterically in a cumulative feedback fashion (each end product causes partial inhibition).

② The sensitivity to allosteric inhibitors is increased by the reversible adenylylation of tyrosine-397. This effect becomes the more pronounced, the more the subunits are adenylylated.

③ The adenylylation level, in turn, is regulated by controlled uridylylation of an auxiliary protein P_{II}, which associates with the adenylyl transferase.

In mammals, the enzyme is present in the mitochondria of all organs; in liver, however, it is only present in the small portion



of paravenous cells. Its regulation is much simpler: it is only activated by oxoglutarate. By action of this enzyme, free ammonia from degradation procedures is bound. The resulting glutamine is the major transport form of amino groups between organs. Free ammonia is toxic beyond moderate concentrations, therefore its blood concentration is kept low (< ca. 60 µmol/l, except in portal vein blood).

In plant chloroplasts, the enzyme mostly acts to recover ammonia, which is liberated at the glycine oxidation step of photorespiration (3.1.9.2) in the neighboring mitochondria. It also occurs in leukoplasts of roots, where it binds the ammonia taken up from the soil or obtained by reduction of nitrate.

Glutamine conversions and degradation: Glutamine can be directly incorporated into proteins (4.1.3, 4.3.2.3). The amino group of glutamine can be transferred by transaminase reactions (3.2.2.5) to other moieties for synthesis of amino acids (3.2.2 … 3.2.9), amino sugars (3.1.7.1), nucleotides (3.6.1, 3.6.2), NAD (3.7.9.1) etc. The glutamate formation by glutamate synthase is described below. In animals, liberation of ammonia by glutaminase occurs in several organs.

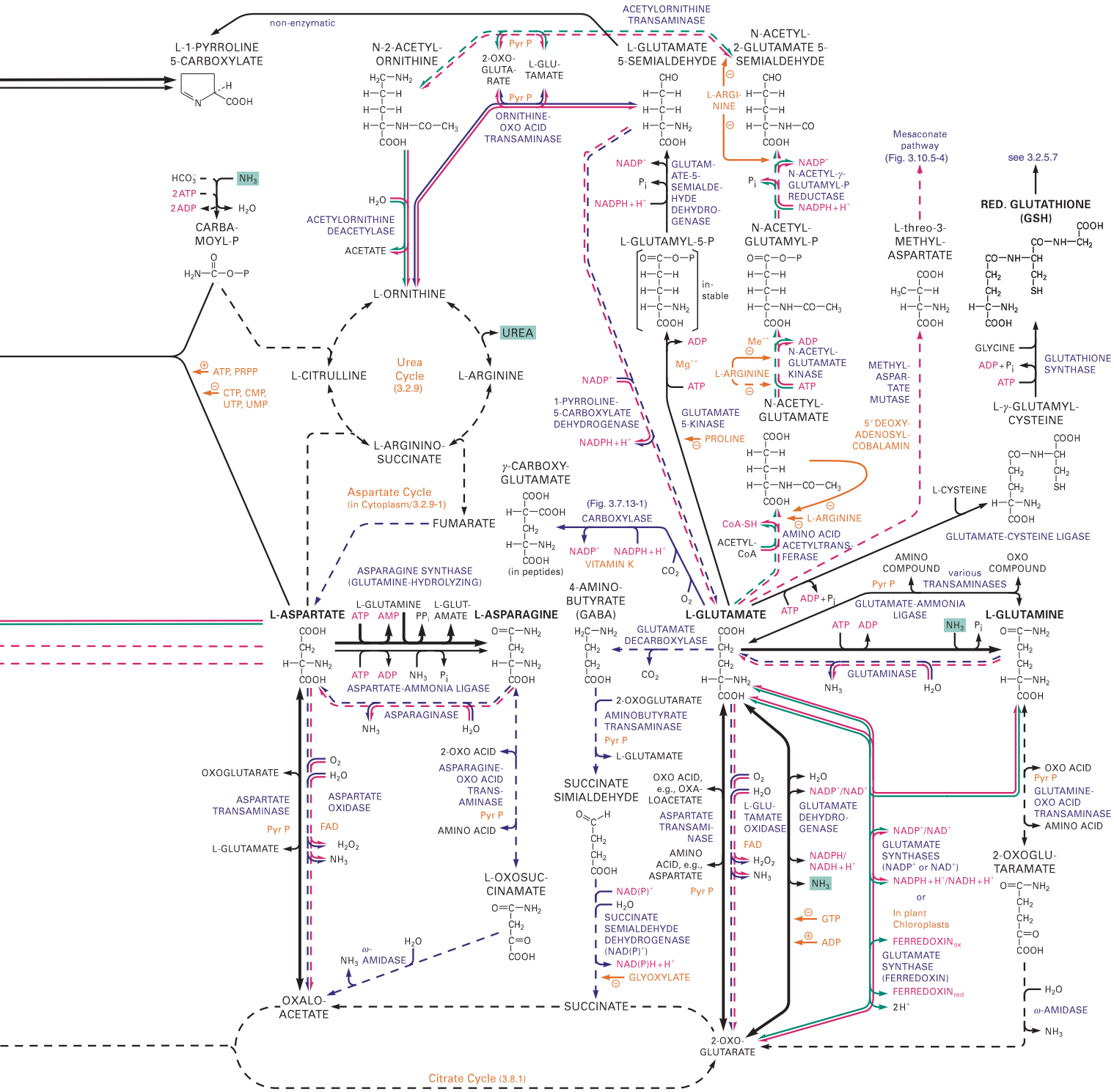


Figure 3.2.2-2. Metabolism of Glutamate, Aspartate, and Related Compounds

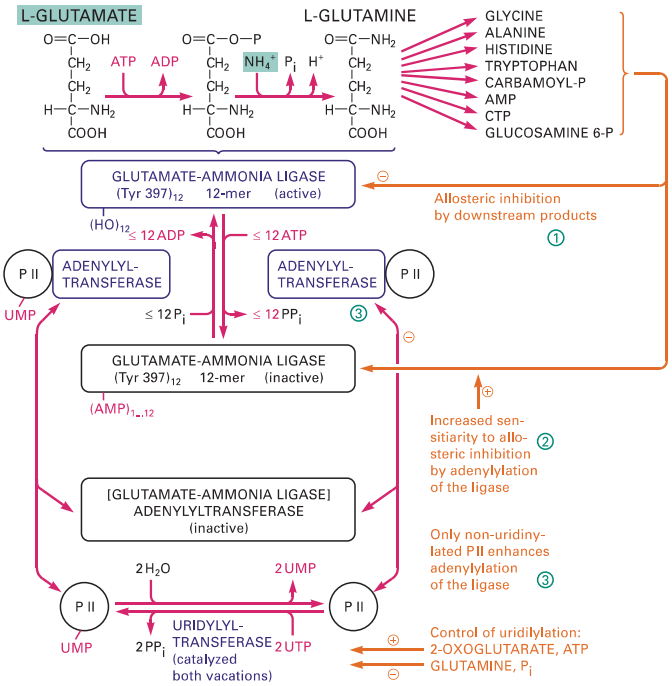


Figure 3.2.2-1. Regulation of Glutamate-Ammonia Ligase in *E. coli*

The resulting glutamate is oxidized there for energy supply or enters biosynthetic reactions.

3.2.2.2 Glutamate Metabolism (Fig. 3.2.2-2)

Glutamate synthesis: In bacteria and in plant chloroplasts, the main pathway to glutamate starts from 2-oxoglutarate and uses glutamine (3.2.2.1) as the amino source. It is catalyzed by glutamate synthase in a reduction-transamination reaction.

The major glutamate production in animals, however, takes place by transamination between 2-oxoglutarate and amino acids to be catabolized (Fig. 3.2.2-3). This proceeds mainly in the liver, which is the central organ for amino acid interconversions and acts also as a 'buffer' after resorption.

For removal of free ammonia and simultaneously for glutamate biosynthesis, liver mitochondria use the glutamate dehydrogenase reaction (operating in the direction of glutamate formation).

Glutamate conversions: Besides being incorporated into proteins and peptides (e.g., glutathione, 3.2.5.7) and taking part in glutamine formation, glutamate acts as donor of amino groups for biosynthetic reactions in numerous transamination reactions.

Additional glutamate reactions are:

- Glutamate, as well as its decarboxylation product 4-aminobutyrate (GABA) act in the central nervous system as neuro transmitters (Table 7.2-2). The 4-aminobutyrate degradation leads finally to succinate; it also occurs in bacteria.
- Condensation of glutamate with acetyl-CoA yields N-acetylglutamate, the initial compound for ornithine and arginine synthesis and an activator of carbamoyl synthesis (urea cycle, 3.2.9).
- Phosphorylation and consecutive reduction leads to glutamate 5-semialdehyde, which is the precursor of pathways to both proline (3.2.3) and to ornithine. Ornithine, in turn, can be converted into arginine (3.2.9.1).
- Posttranslational γ-carboxylation of glutamate in coagulation factors is essential for their activity (9.3.1). This reaction requires the presence of vitamin K (phylloquinone, 3.7.13).

Glutamate degradation: Depending on the metabolic state, glutamate dehydrogenase in liver mitochondria (see above) also may operate in the catabolic direction by converting glutamate into 2-oxoglutarate. Ammonia, which is liberated this way, either enters the urea cycle (3.2.9.1) or is directly excreted (species-dependent). The other product, 2-oxoglutarate:

- is oxidized in the citrate cycle (3.1.8.1) or
- enables by transamination the conversion of many other amino acids into their respective oxo acids.
  - The oxo acids are then oxidized (Fig. 3.1.8-1).
  - Alternatively, these oxo acids can enter the gluconeogenesis pathway. Glucogenic amino acids are: Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile*, Met, Phe*, Pro, Ser, Thr*, Trp*, Tyr*, Val. Amino acids marked with * are also ketogenic; upon degradation they yield ketone bodies (acetyl-CoA, acetoacetate, 3.4.1.7). Only Lys and Leu are strictly ketogenic.

Thus, glutamate dehydrogenase plays a central role in amino acid metabolism. In vertebrates it is allosterically regulated by the energy situation. It is inhibited by GTP (some of which is formed in the citrate cycle, 3.1.8.1) and ATP, but activated by GDP and ADP. In a number of organisms, the enzyme uses both NAD^+ and $NADP^+$.

Glutamate, as well as a number of other amino acids, can also be degraded by amino acid oxidases. Some of them are fairly unspecific. In humans, they occur in the endoplasmic reticulum of liver and kidney.

3.2.2.3 Alanine Metabolism (Fig. 3.2.2-2)

Alanine is an essential component of bacterial murein walls (3.10.1). In animals, alanine is an important transport metabolite for amino groups besides glutamine. Pyruvate, which is abundantly generated in muscles during exercise, accepts the amino groups from glutamate by transamination and passes them on to the liver, where the reverse reaction reconstitutes pyruvate (used for gluconeogenesis, Fig. 3.1.1-7) and glutamate (glucose-alanine cycle).

$$\text{pyruvate} + \text{glutamate} \underset{(liver)}{\overset{(muscle)}{\rightleftharpoons}} \text{alanine} + \text{oxoglutarate}$$

3.2.2.4 Aspartate and Asparagine Metabolism (Fig. 3.2.2-2)

Aspartate is connected by a transaminase reaction with the citrate cycle component, oxaloacetate. This reaction is used for aspartate biosynthesis as well as for degradation.

Oxaloacetate + glutamate = aspartate + oxoglutarate

The mitochondrial and cytosolic isoenzymes of aspartate (ASAT=GOT) and alanine transaminases (ALAT=GPT, see above 3.2.2.3) are of importance for diagnosis of liver damage.

Aspartate is the starting point of important biosynthetic pathways:

- The condensation with carbamoyl-P is the initiation reaction for pyrimidine biosynthesis (3.6.2.1).
- The pathway to inosine monophosphate and later to adenosine monophosphate involves 2 condensation steps with aspartate (3.6.1.1, 3.6.1.2).
- The condensation with citrulline leads to the member of the urea cycle, argininosuccinate, which is a precursor of arginine (3.2.9.1).
- Aspartate phosphorylation, yielding aspartyl-P leads in bacteria and plants to the biosynthesis of the essential amino acids methionine, threonine, lysine and isoleucine (3.2.5.2 … 4; 3.2.6.1).
- Decarboxylation (in bacteria) yields β-alanine, which is used in pantothenate (3.7.7) and carnosine biosynthesis (3.2.8).

Asparagine results from the asparagine-ammonia ligase reaction with glutamine as an amino group donor. Its degradation takes place in the cytosol by the asparaginase reaction (which proceeds analogously to the glutaminase reaction) or in mitochondria by transamidation and oxidation.

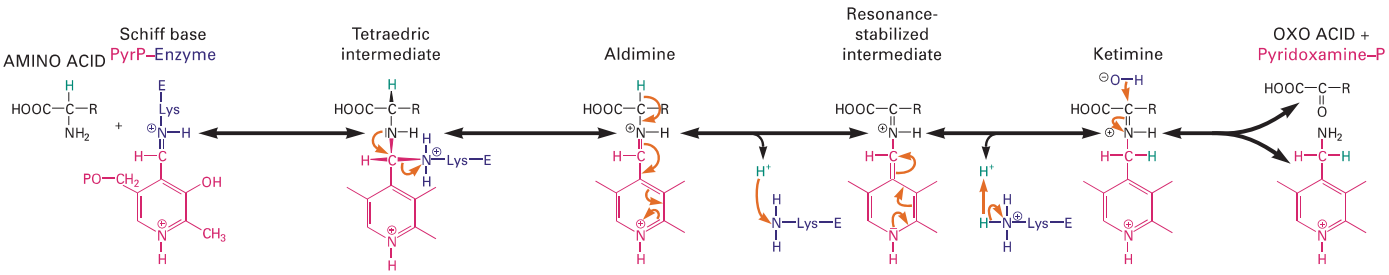


Figure 3.2.2-3. Mechanism of Transamination Reactions

3.2.2.5 Transamination Reactions (Fig. 3.2.2-3)

Transamination is an important step for synthesis and degradation of many amino acids, as well as of other amino compounds (e.g., D-glucosamine, 3.1.7.1). It takes place by pyridoxal-P dependent transfer between amino (in some cases amido, e.g., glutamine) and oxo compounds:

$$R_1\text{–}NH_2 + R_2\text{=}O \rightleftharpoons R_1\text{=}O + R_2\text{–}NH_2$$

The catalyzed reaction starts by nucleophilic attack of the substrate amino group on the Schiff base structure, which exists between pyridoxal phosphate and an ε-lysine group of the enzyme, and replaces it by an 'aldimine' Schiff base between the substrate and pyridoxal phosphate. Abstraction of a proton by the Lys–NH_2 group of the enzyme yields a resonance-stabilized intermediate. The consecutive protonation results in a 'ketimine' Schiff base (tautomerization). Hydrolysis leads to release of the oxo acid. After the binding of another oxo acid, the reaction proceeds in the reverse direction to yield the respective amino acid (Ping-Pong Bi-Bi-mechanism, 1.5.4).

In the aldimine structure, not only the C_α–H bond, but also the C_α–COOH bond, the C_α–R bond and the next intra-R bond are labilized, since cleavage of each of them leads to the resonance-stabilized carbanion. The pyridine ring acts as an electron sink. This allows several reaction variants (for details see 3.7.4.2):

- Cleavage of C_α–H bond: removal of amino group, yielding an oxo group; the consecutive reversal reaction results in transamination.
- Cleavage of C_α–COOH bond: decarboxylation (e.g., aspartate 1- or 4-decarboxylase, glutamate and histidine decarboxylases).
- Cleavages regarding R (α,β and β,γ eliminations): e.g., serine dehydratase (Fig. 3.2.4-2) and kynureninase (Fig. 3.2.7-4).

Literature:

Bender, D.A. *Amino Acid Metabolism*. 2nd Ed. Wiley, 1985.
Hayashi, H. *et al.* Ann. Rev. of Biochem. 1990;59:87–110.
Liaw, S.-H., Eisenberg, D. Biochemistry 1994;33:675–681.
Meister, A. *Biochemistry of the Amino Acids*. 2nd Ed. Academic Press, 1965.
Umbarger, H.E. Ann. Rev. of Biochem. 1978;47:522–606.

3.2.3 Proline and Hydroxyproline (Fig. 3.2.3-1, next page)

Biosynthesis: Glutamate is phosphorylated and consecutively reduced to glutamate semialdehyde (see also Fig. 3.2.2-2). This compound cyclizes non-enzymatically to L-1-pyrroline-5-carboxylate. A NADH or NADPH dependent reduction leads to the nonessential amino acid proline.

In proteins, the rigid ring structure of proline does not allow rotation at the carboxylate-C–N bond and impedes the formation of α-helices. Therefore, proline puts many constraints on the protein structure. On the other hand, amino acid proline bonds can assume both *trans*- and *cis*-configurations (1.3.2). The isomerization by peptidyl-proline-*cis*-*trans*-isomerases (PPI) is frequently the rate-determining step in protein folding procedures (4.5.1.1). This enzyme is an essential subunit of proline-4-hydroxylase in eukarya.

4-Hydroxyproline (Hyp) is produced posttranslationally by procollagen-proline 4-dioxygenase (proline 4-hydroxylase), which is an intermolecular dioxygenase containing Fe^{2+}. Besides proline, Hyp is a major component of collagen (2.3.1) and stabilizes it by formation of hydrogen bonds. Small amounts of 3-hydroxyproline and 5-hydroxylysine (3.2.5.2) are also present.

Proline-4-dioxygenase binds in ordered sequence, Fe^{2+}, the cosubstrate 2-oxoglutarate, O_2 and a peptide containing the sequence X-Pro-Gly (preferably with X = Pro). O_2 is activated by interaction with the bound Fe^{2+} and performs a nucleophilic addition to oxoglutarate. The complex hydroxylates proline; then the products leave the enzyme. Ascorbate is oxidized in substoichiometric amounts and apparently prevents Fe from being in the oxidized state after the reaction cycle has ended. A similar situation exists with procollagen-lysine 5-dioxygenase (3.2.5.2). Diminished hydroxylation of proline due to a lack of ascorbate prevents the proper formation of collagen fibers; their melting temperature is diminished (scurvy).

Degradation of proline starts with the oxidation of 1-pyrroline-5-carboxylate, finally yielding glutamate and oxoglutarate. Apparently, glutamate 5-semialdehyde is an intermediate. Degradation of hydroxyproline takes place analogously, leading to hydroxyglutamate. After cleavage, transamination results in pyruvate and glycine.

Bacterial enzymes can perform racemization of L-proline and L-hydroxyproline to D-proline and D-*allo*-4-hydroxyproline, respectively. Their degradation takes place by D-amino acid oxidase.

Literature:

Adams, E., Frank, L. Ann. Rev. of Biochem. 1980;49:1005–1061.
Counts, D.F. *et al.* Proc. Natl. Acad. Sci. (USA) 1978;75:2145–2149.
Myllylä, R. *et al.* Biochim. Biophys. Res. Commun. 1978;83:441–448.

3.2.4 Serine and Glycine

Serine and glycine are nonessential amino acids. They are derived from 3-phosphoglycerate. In plants and bacteria, serine can be directly converted into cysteine, while in animals, the essential amino acid methionine is needed as reaction partner for cysteine biosynthesis (3.2.5.5). The interconversion of serine and glycine requires the participation of the tetrahydrofolate C_1-transfer system (3.7.6.2). Glycine is an inhibitory neurotransmitter (Table 7.2-2).

3.2.4.1 Serine Metabolism (Fig. 3.2.4-1, p. 64)

The **biosynthesis** of serine starts from 3-phosphoglycerate and proceeds by a sequence of dehydrogenase (to 3-P-hydroxypyruvate), transaminase (to 3-P-serine) and phosphatase reactions. In animals, the controlled enzyme is the third in this pathway (phosphoserine phosphatase) instead of the usual initial one.

Besides its occurrence in proteins, serine is a component of glycerophospholipids (3.4.3.2) and the origin of sphingosine and ceramide biosynthesis (3.4.3.4). The interconversion with glycine proceeds in both directions (3.2.4.2). In plants and bacteria, serine can be acetylated and thereafter the acetyl group exchanged with sulfide, resulting in cysteine. Animals, however, synthesize cysteine from the essential amino acid methionine (see 3.2.5.4) via the intermediate cystathionine.

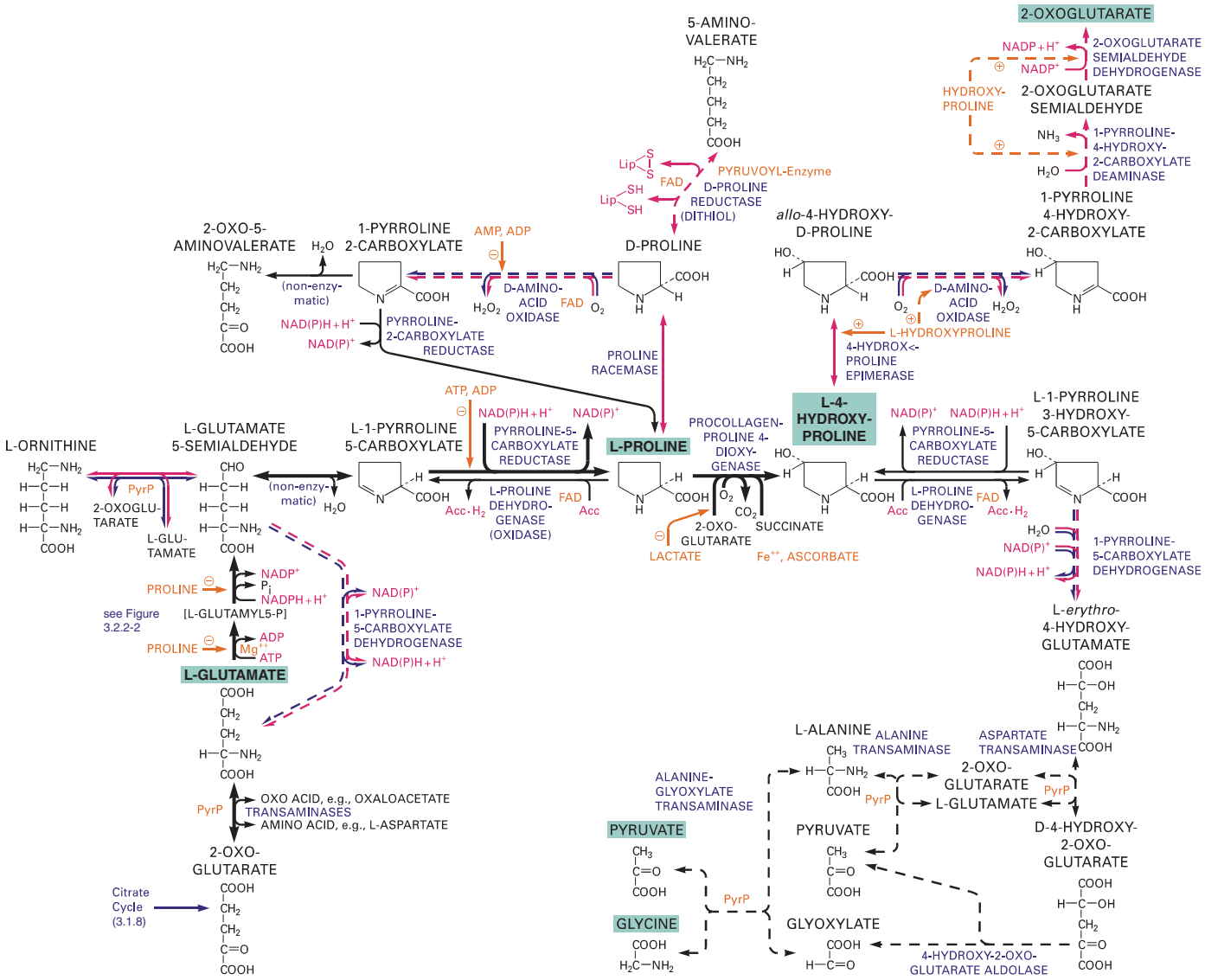


Figure 3.2.3-1. Proline and Hydroxyproline Metabolism

Serine degradation: The major route for serine catabolism is the serine dehydratase reaction, leading to pyruvate (Fig. 3.2.4-2). This way, serine (and glycine after interconversion to serine) can enter the gluconeogenesis pathway (3.1.3.5).

The dehydratase and transaminase (3.2.2.5) mechanisms are related, both require pyridoxal phosphate. In both cases, an aldimine structure is formed. In the transaminase reaction, after removal of the α-hydrogen of the amino acid, there is a protonation of the C-4′ atom of pyridoxal phosphate (Fig. 3.2.2-3). In the dehydratase reaction, however, β-elimination of the hydroxyl group from serine takes place.

Reconversion of serine to 3-phosphoglycerate via the non-phosphorylated compounds hydroxypyruvate and glycerate is another way of serine utilization. In liver and kidney of animals, it is used for gluconeogenesis. In plants, it is part of the photorespiration sequence (3.1.9.2, 3.12.2).

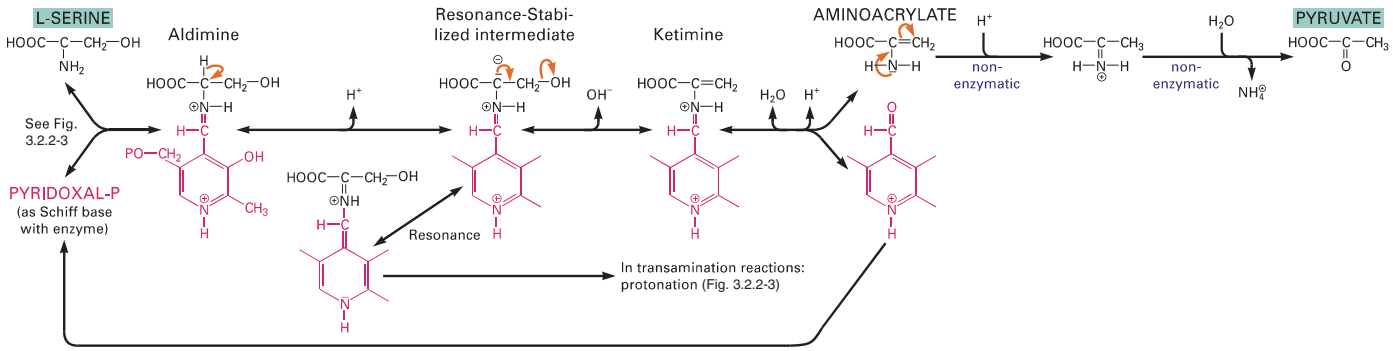


Figure 3.2.4-2. Mechanism of Serine Dehydratase

3.2.4.2 Glycine Metabolism (Fig. 3.2.4-1)

Both glycine synthesis and catabolism proceed mainly by interconversion with serine. This involves a C_1 group transfer by tetrahydrofolate (THF, 3.7.6.2).

For glycine formation, a methylene group is moved from serine to THF by the pyridoxal dependent enzyme glycine hydroxymethyltransferase (also named serine hydroxymethyltransferase), yielding 5,10-methylene-THF and releasing glycine. The folate coenzyme passes on this C_1 moiety to various acceptors for biosynthetic purposes.

In the opposite direction, a molecule of glycine is at first converted by the mitochondrial glycine cleavage system to 5,10-methylene-THF and CO_2 (Fig. 3.2.4-3). This system is a multi-enzyme complex, which resembles the pyruvate dehydrogenase complex (3.1.3.1) and consists of the components

- glycine dehydrogenase (decarboxylating) (P protein, contains pyridoxal-P)
- aminomethyltransferase (H protein, contains an 'arm' of lipoic acid for transport between P and T proteins)

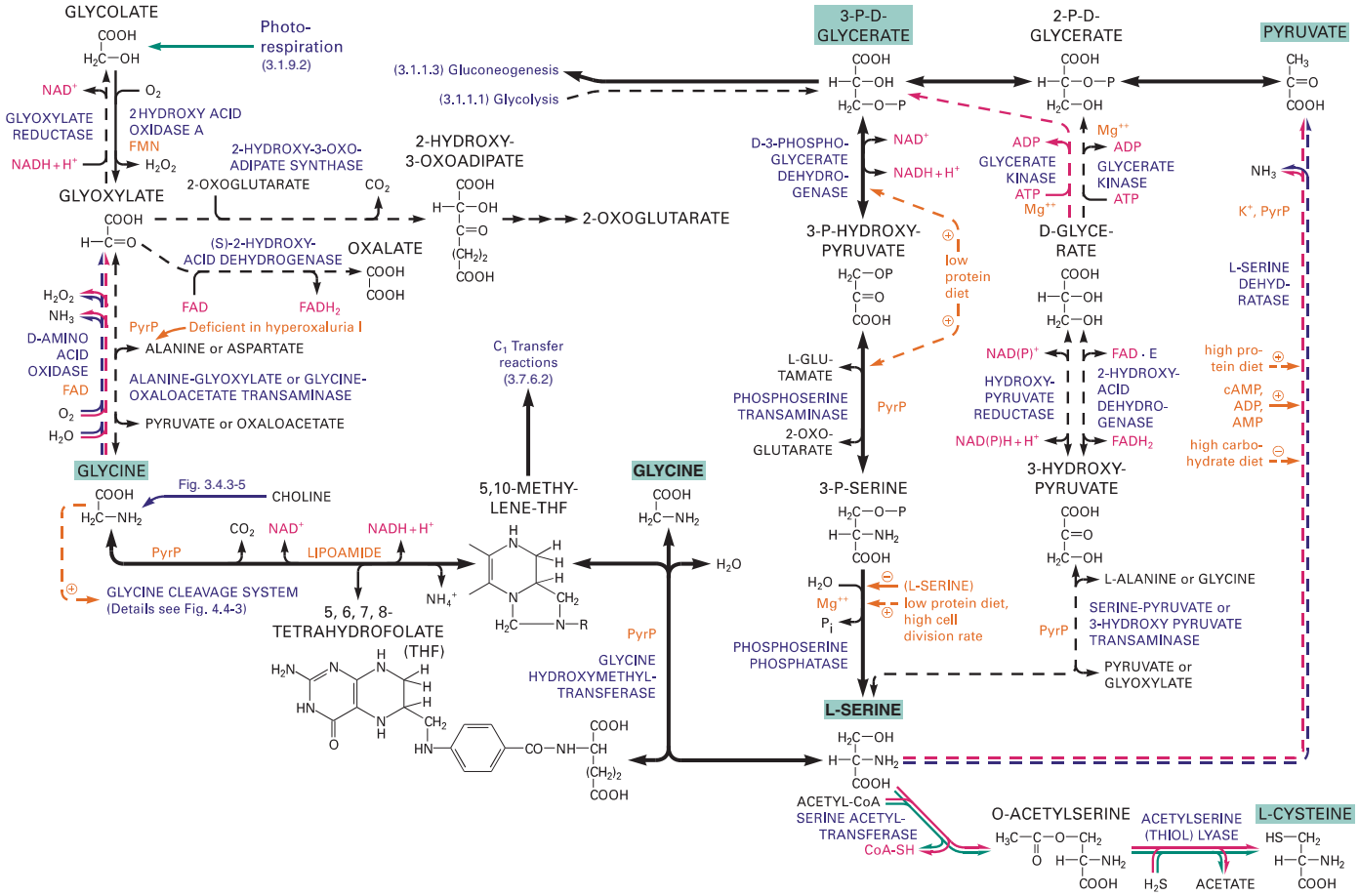


Figure 3.2.4-1. Serine and Glycine Metabolism

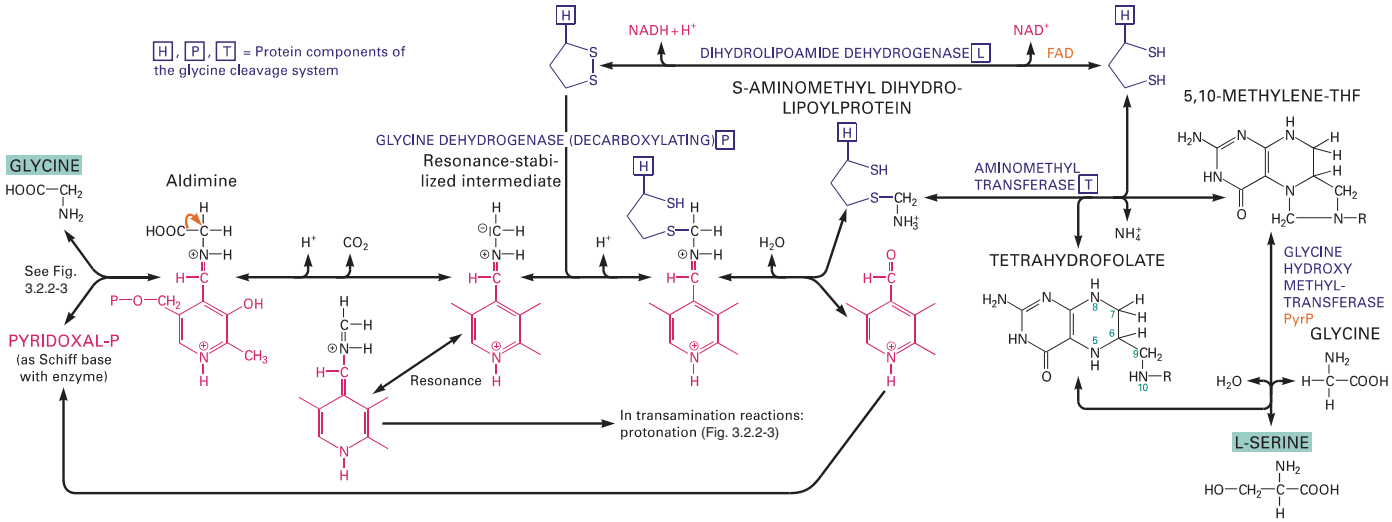


Figure 3.2.4-3. Synthesis of Serine from Glycine (Reversible)

- aminomethyltransferase (T protein, contains tetrahydrofolate)
- dihydrolipoyl dehydrogenase (L protein)

The initial decarboxylation reaction is due to the labilization of the C_α–COOH bond in the aldimine structure (see transaminase mechanism, 3.2.2.5).

One of the various possible reactions of the 5,10-methylene THF product is the C_1 transfer to a second glycine molecule by glycine hydroxymethyltransferase, resulting in serine formation. Thus, this enzyme catalyzes, together with the glycine cleavage system, the reversible reaction:

$$2\ \text{glycine} + H_2O + NAD^+ \rightleftharpoons \text{serine} + CO_2 + NH_4^+ + NADH.$$

Glycine hydroxymethyltransferase can also react in absence of tetrahydrofolate. This way it catalyzes the cleavage of threonine, resulting in glycine and acetaldehyde ('threonine aldolase', Fig. 3.2.5-2).

Another method of glycine synthesis is by transamination of glyoxylate, which originates from glycolate. It is part of the photorespiration sequence in plants (3.1.9.2, 3.12.2), but also occurs in non-photosynthetic organisms, e.g., yeast, in gluconeogenesis.

Literature:

Rapoport, S. *et al.* Eur. J. Biochem. 1980;108:449–455.

Snell, K. Adv. Enzyme Regul. 1984;22:325–400 and 1990;30:13–32.

Umbarger, H.E. Ann. Rev. of Biochem. 1978;47:533–606.

3.2.5 Lysine, Threonine, Methionine, Cysteine, and Sulfur Metabolism

This group of amino acids is formed from aspartate. Only in fungi, lysine originates from 2-oxoglutarate. Although cysteine is the only one out of this group which is not considered essential for mammals, it still requires the essential amino acid methionine for its synthesis. Precursors of isoleucine are both threonine and pyruvate. Its biosynthesis is discussed in 3.2.6.1.

3.2.5.1 Common Steps of Biosynthesis and Their Regulation (Fig. 3.2.5-1)

The **biosynthesis** of lysine (in bacteria and most plants), threonine and methionine starts with the phosphorylation of aspartate. As usual, the first committed step of the pathway is the point of action for regulatory mechanisms. Since the pathways for synthesis of the individual amino acids branch later, each amino acid has to perform its own control. In various organisms, different systems have been realized. They are excellent examples of the various possibilities for control of metabolism.

In *E. coli*, there are 3 aspartate kinases. Each of them is feedback inhibited by one of the amino acids (multiple enzyme control). Additionally, feedback inhibition of the respective enzymes after branch points takes place. Each of the aspartate kinases I and II (which are regulated by threonine and methionine) exists together with the respective homoserine dehydrogenase as a bifunctional enzyme on a single peptide chain.

The operon encompassing the three structural genes coding for aspartate kinase I / homoserine dehydrogenase I, homoserine kinase and threonine synthase is controlled by a single promoter-operator locus. Bivalent repression takes place by threonine and isoleucine. The repression of aspartate kinase II activity by methionine is also enhanced by isoleucine.

In *Rhodopseudomonas spheroides*, however, only aspartate semialdehyde (the last common intermediate) inhibits the single aspartate kinase, while in *R. capsulata* and others, a synergistic inhibition of the aspartokinase by lysine and threonine takes place (cooperative feedback control). Still other variants have been found in other organisms.

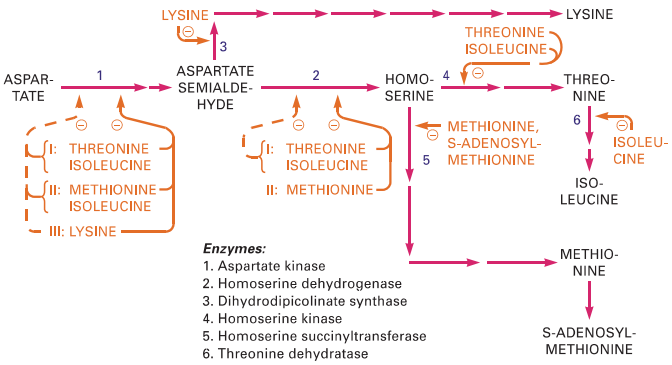


Figure 3.2.5-1. Regulation of Threonine, Methionine and Lysine Biosynthesis in *E. coli*
Feedback inhibition is indicated by solid arrows, repression of enzyme synthesis by dashed arrows. I, II and III indicate different enzymes with individual regulation.

3.2.5.2 Lysine Metabolism (Fig. 3.2.5-2, p. 67–68)

Lysine biosynthesis in bacteria: The individual pathway begins with a condensation reaction of aspartate semialdehyde with pyruvate. The product cyclizes immediately. After reduction, the non-cyclic form is stabilized by succinylation or acetylation. Further reactions lead to diaminopimelate, which is also a component of bacterial cell walls (3.10.1). By decarboxylation, lysine is obtained. In some bacteria, a reductive amination leads directly from Δ^1-piperidine 2,6-dicarboxyate to diaminopimelate.

Lysine biosynthesis in fungi: These organisms use a completely different pathway for biosynthesis, which starts from 2-oxoglutarate and acetyl-CoA. The first steps resemble the beginning of the citric acid cycle, employing homologous (= homo-) compounds. The oxoadipate formed undergoes a transamination reaction, followed by the reduction to the semialdehyde (possibly via an adenylylated intermediate). The ε-amino group is not introduced by transamination, but rather in a sequence involving reduction, formation of the covalently linked intermediate saccharopine and oxidative cleavage to yield L-lysine. This pathway is regulated by feedback inhibition of the initial condensation reaction.

Biological role of lysine: Lysine is involved in various mechanisms of enzyme catalysis (e.g., transaminases, 3.2.2.5). Biotin (3.7.8) and lipoic acid (3.7.14.1) are bound to lysine residues of enzymes. In procollagen, lysine is 5-hydroxylated in the endoplasmic reticulum (analogous to proline, 3.2.3). After its glycosylation, final folding to the collagen triple helix takes place (2.3.1). A similar hydroxylation of trimethylated lysine is an intermediate step in the synthesis of carnitine, which is important for uptake of fatty acids (3.4.1.4).

Lysine degradation: There are several degradation mechanisms. The one prevalent in mammalian liver starts with reactions which are essentially the reversal of the biosynthesis reactions in fungi (see above). Likely, the roles of NAD^+ and $NADP^+$ are exchanged. Oxidative decarboxylation of 2-oxoadipate results in glutaryl-CoA (the homologue of succinyl-CoA), which after a second decarboxylation yields crotonyl-CoA. This is a member of the fatty acid degradation pathway (3.4.1.5), which leads to acetyl-CoA. Thus, lysine is a ketogenic amino acid.

3.2.5.3 Threonine Metabolism (Fig. 3.2.5-2, p. 67–68)

Threonine biosynthesis: The individual part of the pathway is the isomerization of homoserine to threonine. This reaction takes place by the formation of a phosphate ester with the hydroxyl group and a consecutive β,γ-elimination reaction requiring pyridoxal phosphate. (For other examples of pyridoxal-P catalyzed reactions, see Figs. 3.2.2-3, 3.2.4-2 and 3.2.4-3). The kinase reaction is competitively inhibited by threonine.

Threonine conversions and degradation: By action of threonine dehydratase, threonine is converted to 2-oxobutyrate by a reaction analogous to the serine dehydratase reaction (Fig. 3.2.4-2). 2-Oxobutyrate acts as a precursor of isoleucine (3.2.6.1) or is degraded (3.2.5.4).

For degradation, threonine is cleaved directly or after oxidation, yielding glycine (which is converted via serine to pyruvate, 3.2.4.2) and acetaldehyde or acetate, which are consecutively converted into acetyl-CoA.

3.2.5.4 Methionine Metabolism (Fig. 3.2.5-2)

Methionine biosynthesis: The γ-hydroxyl group of homoserine is activated by succinyl-, acetyl- or phosphate groups (in higher plants). In bacteria, this reaction is synergistically inhibited by methionine and S-adenosylmethionine. Then, in all organisms a transsulfuration with cysteine takes place. The resulting cystathionine is cleaved to yield homocysteine. Several variants of this sequence exist, see footnote to Figure 3.2.5-2. Homocysteine is then methylated by methyltetrahydrofolate (3.7.6.2, in most cases), resulting in methionine. Various diseases are caused by defects of these enzymes.

Biological roles: Methionine (in bacteria: formylmethionine) is the starting amino acid in protein biosynthesis (4.1.3.3, 4.2.3.2). Homocysteine is likely a risk factor for arteriosclerosis.

S-Adenylylation of methionine leads to S-adenosyl-L-methionine (SAM) with a positively charged sulfur atom, which activates the neighboring methyl group. This compound is most important as methyl group donor in transfer reactions (e.g., 3.2.9.1, 3.4.3.2, 3.8.3.4).

Thus, by the formation of methionine and this activation reaction, the moderate methylation ability of methyltetrahydrofolate (3.7.6.2) is converted into a more reactive mode. The adenylylation reaction is feedback-inhibited by its products S-adenosyl-L-methionine, PP_i and the cleavage product P_i.

Decarboxylation of SAM yields S-adenosylmethioninamine, which enters some biosynthetic reactions (3.2.9.3). A conversion to ethane occurs in fruits.

After transfer of the methyl group, the resulting S-adenosylhomomocysteine (SAH) is deadenylated to homocysteine, which enters another methylation-demethylation cycle or is degraded.

Methionine degradation: The demethylation product homocysteine undergoes a condensation reaction with serine, yielding cystathionine. After releasing cysteine (which is the biosynthesis reaction of this amino acid in animals), 2-oxobutyrate is oxidatively decarboxylated, resulting in propionyl-CoA. Then, carboxylation by the biotin dependent enzyme propionyl-CoA carboxylase takes place, which resembles the acetyl-CoA carboxylase reaction. After epimerization, a mutase reaction converts L-methylmalonyl-CoA into the citric acid cycle component succinyl-CoA.

The same final steps of this sequence take place during catabolism of threonine, valine and isoleucine (3.2.6.2) and of odd-numbered fatty acids (3.4.1.5). The mechanism of biotin-dependent carboxylations is dealt with in 3.7.8.2. The mutase reaction is one of the 2 mammalian reactions employing a vitamin B_{12} derivative (3.7.5.2). Deficiencies of propionyl-CoA-carboxylase lead to an increase in propionate in blood. An increase of methylmalonate takes place, if the mutase enzyme or the coenzyme B_{12} biosynthesis is defective.

In bacteria, propionate fermentation proceeds via essentially the same sequence in the opposite direction. The CO_2 released from methylmalonyl-CoA is transferred by a biotinyl protein to pyruvate, yielding propionyl-CoA and oxaloacetate (Fig. 3.10.5-2).

Possibly an additional catabolic pathway exists in mammals, which proceeds via transamination of methionine, decarboxylation and liberation of methanethiol. Its oxidation leads to CO_2 and sulfate.

methionine → oxo acid → 3-methylthiopropionate → methanethiol → $CO_2 + SO_4^{2-}$

3.2.5.5 Cysteine Metabolism

Cysteine biosynthesis (Fig. 3.2.5-2): Bacteria and plants are able to convert L-serine into L-cysteine via acetylserine. The acetyl group is directly exchanged with H_2S, which is provided by reduction of sulfate (Fig. 3.2.5-4). Animals, however, require homocysteine (from methionine degradation) as source of sulfur and produce cysteine via the condensation product cystathionine.

Biological role of cysteine: The thiol group of cysteine takes part in a number of enzyme reaction mechanisms (e.g., glutathione reductase, 3.2.5.7). It also forms Fe-S centers in electron transfer proteins involved in, e.g., respiration and photosynthesis (3.11.4, 3.12.1). The oxidation of cysteine to the disulfide cystine (Cys-S-S-Cys) plays an essential role in formation and maintenance of the secondary structure of proteins (2.3.1, 4.5.1.3) and enzyme activity regulation (e.g., Calvin cycle enzymes via thioredoxin). On the other hand, cysteine oxidation by atmospheric oxygen is a frequent cause of protein inactivation.

Cysteine conversions and degradation (Fig. 3.2.5-3, p. 68): Oxidation of the -SH group and decarboxylation result in taurine, which is a conjugation partner of bile acids (3.5.9) and is present in retina, brain, lymphocytes etc. It may have a detoxifying and membrane protecting effect. The final product of other cysteine degradation pathways is pyruvate. The release of sulfur can proceed in several ways (as H_2S, SO_3^{2-} or SCN^-).

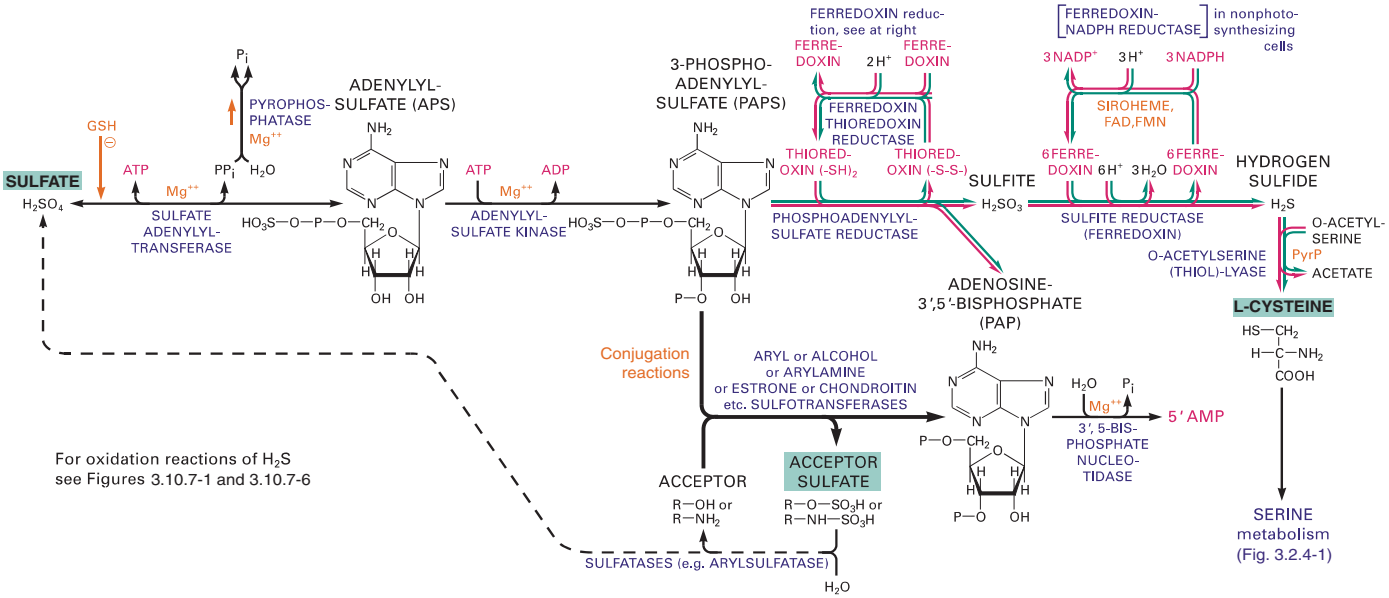
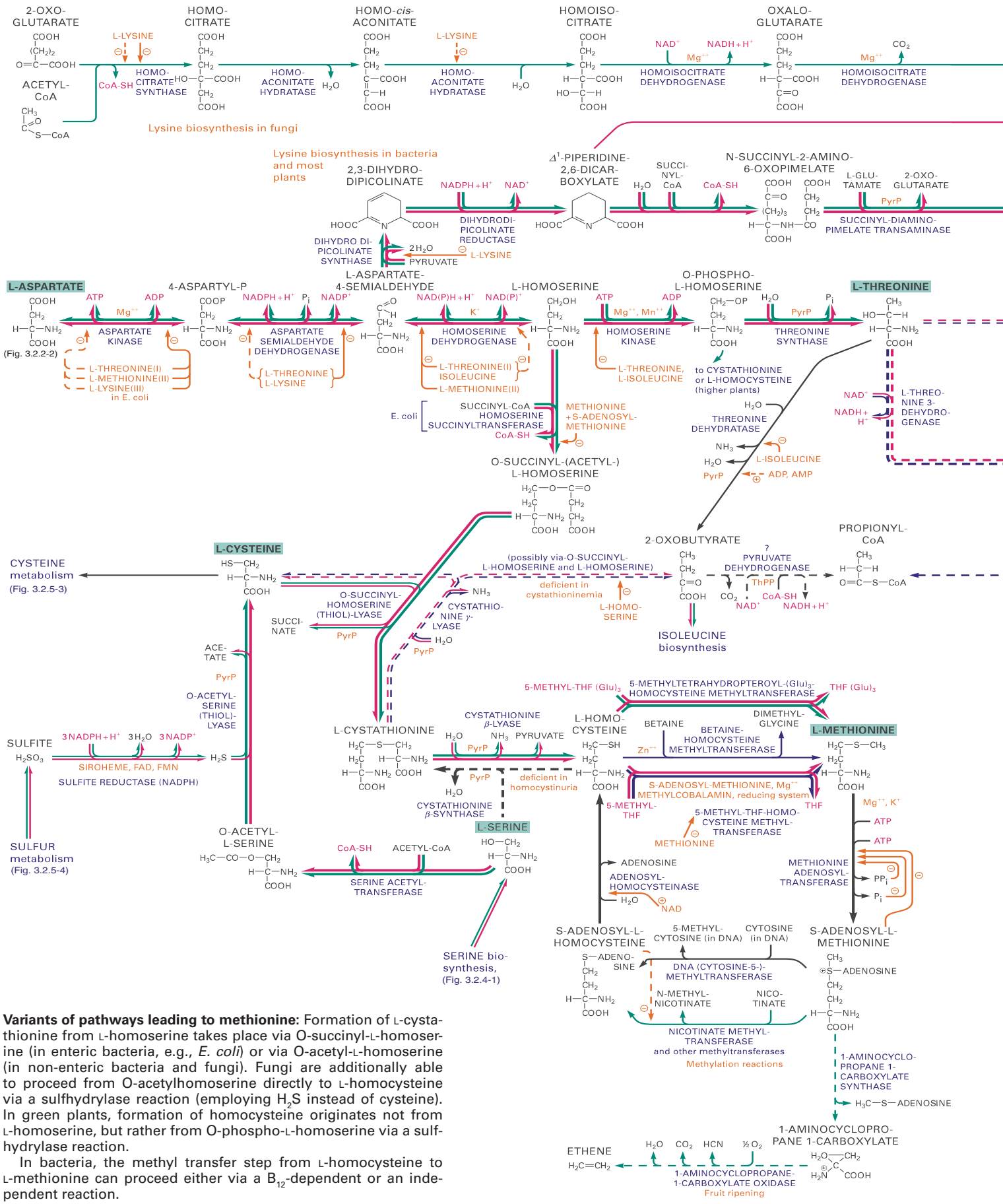


Figure 3.2.5-4. Sulfur Metabolism

(continuation on p. 69)



Variants of pathways leading to methionine: Formation of L-cystathionine from L-homoserine takes place via O-succinyl-L-homoserine (in enteric bacteria, e.g., *E. coli*) or via O-acetyl-L-homoserine (in non-enteric bacteria and fungi). Fungi are additionally able to proceed from O-acetylhomoserine directly to L-homocysteine via a sulfhydrylase reaction (employing H_2S instead of cysteine). In green plants, formation of homocysteine originates not from L-homoserine, but rather from O-phospho-L-homoserine via a sulfhydrylase reaction.

In bacteria, the methyl transfer step from L-homocysteine to L-methionine can proceed either via a B_{12}-dependent or an independent reaction.

Figure 3.2.5-2. Lysine, Threonine, Methionine, and Cysteine Metabolism

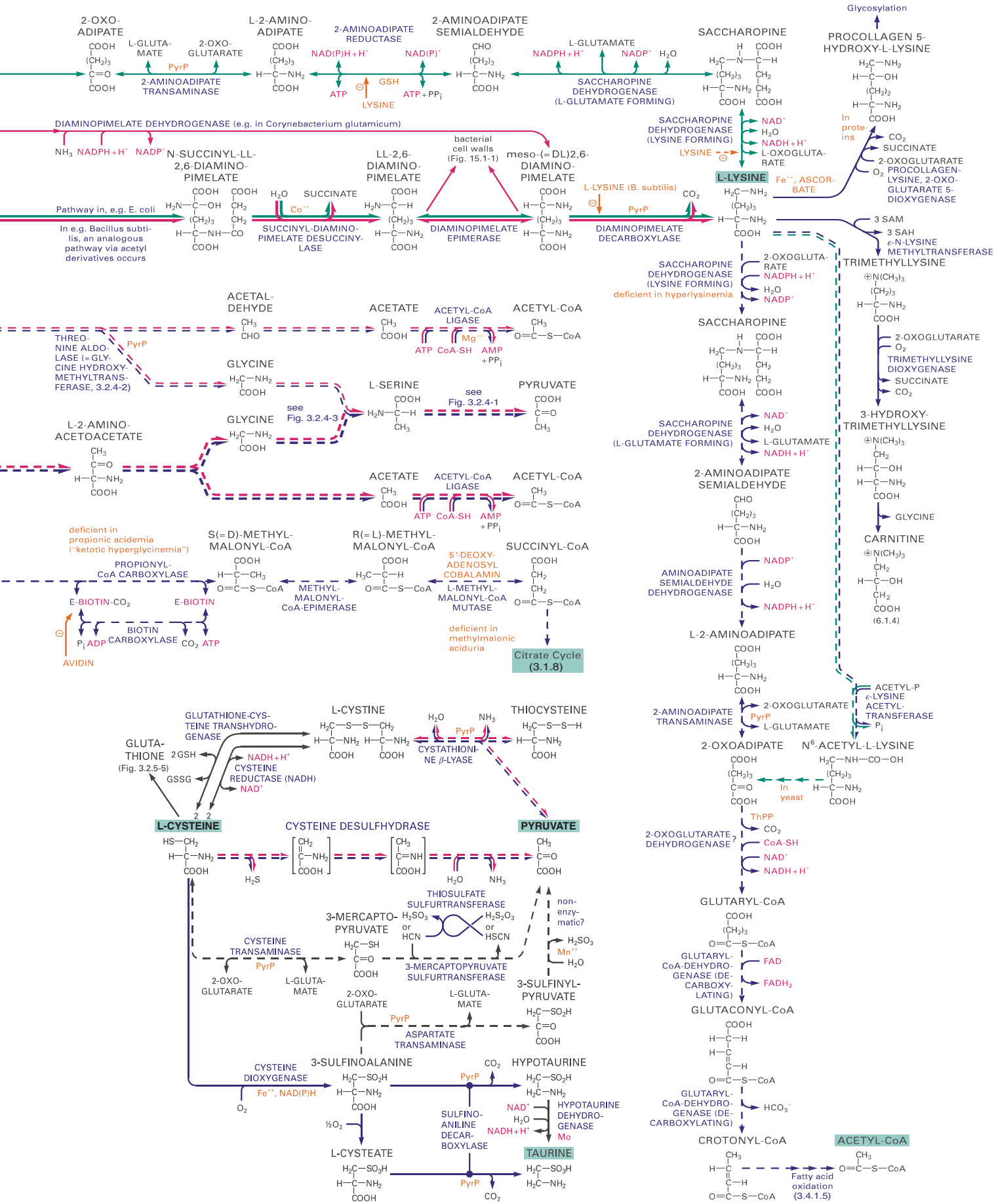


Figure 3.2.5-3. Conversions and Degradation of Cysteine

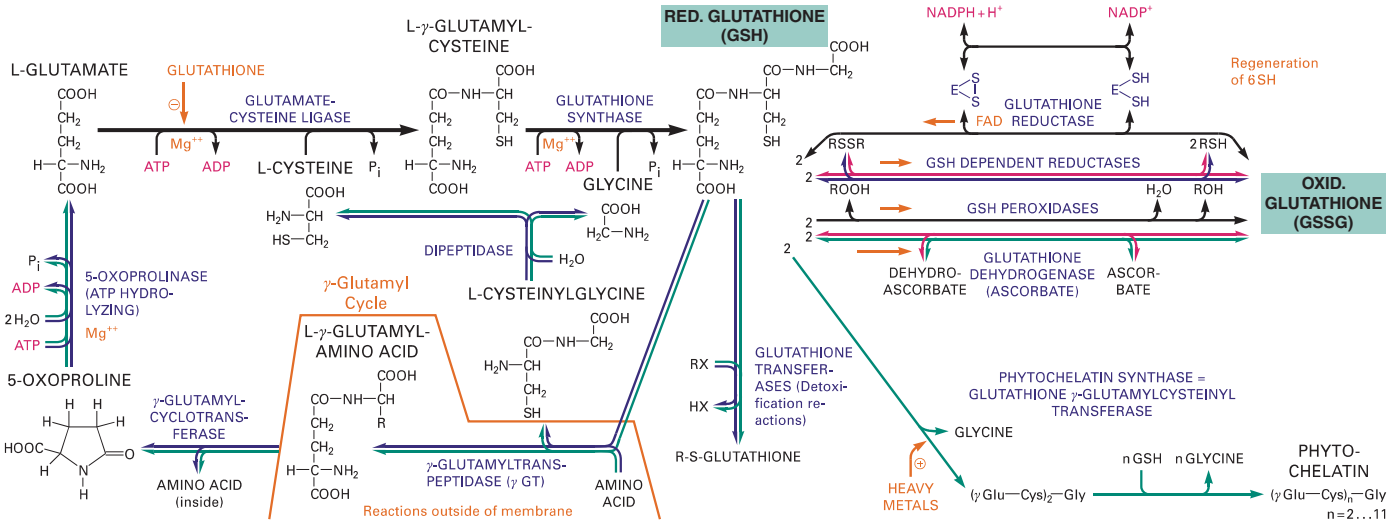


Figure 3.2.5-5. Glutathione Metabolism

3.2.5.6 Sulfur Metabolism (Fig. 3.2.5-4, p. 66)

In the biosphere, sulfur plays a role

- in oxidized form, primarily as bound -OSO_3H in sulfated glycoproteins (4.4.2.1) and glycolipids (3.4.4) and as a conjugation partner for excretion (e.g., 3.1.5.6 … 3.1.5.8)
- in reduced form as -SH and -S- in cysteine, cystine and methionine (also in proteins), glutathione and redox centers (3.11.4, 3.12.1)
- Anaerobic bacteria use reduction and oxidation of sulfur compounds for anaerobic respiration and for chemolithotrophy. The energy aspect of these reactions is dealt with in 3.10.6 and 3.10.7. Slow reduction to the -SH level can be performed by plants.

Metabolism of sulfate: For conjugation as well as for reduction reactions, sulfate has to be activated in an ATP-dependent reaction to adenylylsulfate (APS). Since $\Delta G'_0$ for hydrolysis of the sulfate-phosphate bond (–71 kJ/mol) is much larger than that for the α-β pyrophosphate bond in ATP (–32.2 kJ/mol), the reaction has to be 'pulled' by hydrolysis of the liberated pyrophosphate and by additional phosphorylation of APS to 3′-phosphoadenylylsulfate (PAPS). PAPS introduces sulfate residues in many compounds; the resulting adenosine 3′,5′-diphosphate (PAP) is then hydrolyzed to yield 5′-AMP.

For conversion to sulfite in photosynthesizing plants, reduced ferredoxin (from photosynthesis, 3.12 or regenerated by NADPH) reduces thioredoxin (ca. 100 amino acids, containing the sequence Cys-Gly-Pro-Cys) from the -S-S- to the $(-SH)_2$ state. Catalyzed by PAPS reductase, PAPS is then converted into free sulfite and PAP. The further reduction to sulfide is catalyzed by sulfite reductase, an enzyme containing siroheme (3.7.5.3), which closely resembles the nitrite reductase. The reduction equivalents are supplied by ferredoxin. Also yeast and a number of bacteria use PAPS as an intermediate. Animals are unable to reduce sulfate.

Another pathway in plants, which is under discussion, starts directly from APS (3.10.6) and proceeds via reduction of a glutathione thiosulfate intermediate to the disulfide level (not shown).

Metabolism of sulfide: Fixation of sulfide occurs in plants and bacteria by reaction with acetylserine, yielding cysteine (3.2.5.5). The further reactions of sulfur in amino acids are dealt with in 3.2.5.4, 3.2.5.5 and 3.2.5.7.

3.2.5.7 Glutathione Metabolism (Fig. 3.2.5-5)

The tripeptide glutathione (GSH, γ-Glu-Cys-Gly) is synthesized by specific enzymes and is intracellularly present at high concentrations in animals (ca. 2 … 5 mmol/l, more than 99 % in reduced form). Glutathione is an important reductant (E'_0 = –230 mV). The oxidized form (GSSG) resulting from these reactions is reduced again in a NADPH-dependent reaction by glutathione reductase (see below). GSH shows many detoxifying and cytoprotective effects. Major reaction types are:

- Removal of H_2O_2 (directly or indirectly) by peroxidase reactions (3.2.5.8)
- Stabilization of the redox state of peptides (e.g., insulin) and proteins by the protein-disulfide reductase reaction

  $$2\ GSH + R_1Cys–S–S–Cys\text{-}R_2 \rightleftharpoons GSSG + R_1–Cys–SH + R_2–Cys–SH$$

  In some cases, the reduction of proteins is performed by thioredoxin instead, e.g., for activation of Calvin cycle enzymes (3.12.2).
- Reduction of ribonocleotides (e.g., in some bacteria, 3.6.1.4)
- Cellular import of amino acids by the γ-glutamyl cycle.

  A noticeable portion of the intracellularly synthesized glutathione is exported, in animals mostly from the liver and kidney. At the outside of the cellular membrane of animals and yeast, GSH transfers its glutamate moiety to amino acids to be imported, catalyzed by the membrane bound enzyme γ-glutamyltranspeptidase (γ-GT). After membrane passage, the dipeptide is hydrolyzed and the temporarily formed 5-oxoproline is reconverted to glutamate. This reaction is energized by ATP hydrolysis.
- Conjugation reactions, mainly for neutralizing toxic compounds (xenobiotics), but also for biosynthesis (Fig. 7.4-15).

  The thiol group of GSH reacts with C=C bonds, carbonyl groups, sulfates, etc. This is catalyzed by various, mostly unspecific glutathione transferases. In plants, the conjugates are then actively imported into vacuoles.
- Formation of phytochelatins in plants. As a detoxification reaction, phytochelatins form complexes with heavy metals, which are then transported into vacuoles.

The reduction of GSSG in order to regenerate GSH takes place by glutathione reductase. The reaction mechanism involves the initial reduction of a Cys-S-S-Cys bond in the enzyme by NADPH, formation of a charge-transfer complex of one Cys-S- with FAD, nucleophilic attack of the other Cys-SH on GSSG and cleavage of the mixed disulfide. Other disulfide oxidoreductases (e.g., dihydrolipoamide dehydrogenase, 3.1.3.1) react in an analogous mode.

Lack of NADPH leads to insufficient reduction of GSSG. This is especially manifest in erythrocytes, where NADPH formation can take place only by the glucose-6-phosphate dehydrogenase reaction (3.1.6.1) since mitochondria are absent. An inherited diminished activity of this enzyme (X-chromosomal, occurring in more than 100 million persons) is further decreased by some antimalarial drugs and by fava bean (*Vicia faba*) alkaloids. This leads to hemolytic anemia.

Literature:

Cooper, A.J.L. Ann. Rev. of Biochem. 1983;52:187–222.

Meister, A., Anderson, M.E. Ann. Rev. of Biochem. 1983;52:711–760.

Meister, A. J. Biol. Chem. 1988;263:17205–17208.

Schwenn, J.D. Z. Naturforsch. 1994;49C:531–539.

Umbarger, H.E. Ann. Rev. of Biochem. 1978;47:533–606.

3.2.5.8 Reactive Oxygen Species, Damage, and Protection Mechanisms (Fig. 3.2.5-6)

Molecular oxygen, the most important element of aerobic life, can enter into some side reactions, which release dangerous reactive oxygen species (ROS) or reactive oxygen intermediates (ROI). Therefore, appropriate protection mechanisms are essential. On the other hand, ROS can be used for defense and for regulatory reactions. In the following, selected reactions are shown.

Superoxide radical ($O_2^{\bullet-}$): This radical is formed by one-electron transfer to O_2 ($\Delta E_0' = -330$ mV). $O_2^{\bullet-}$ is generated by side reactions of some oxidases (e.g., xanthine oxidase, 3.6.1.6), of photosystem I, of electron carriers in the respiratory chain (mainly ubiquinone), or by transfer from semiquinone intermediates (① e.g., during monooxygenase reactions in the ER).

$O_2^{\bullet-}$ is also produced by high-energy irradiation. The aggressive compound acts not only as a reductant, but also sometimes as an oxidant (e.g., on sulfite). Superoxide dismutase (SOD) converts it into the similarly toxic hydrogen peroxide H_2O_2 (see below). The same reaction also takes place spontaneously.

The superoxide radical seems to be the most toxic ROS. It shows a certain chemical selectivity, e.g., for reactions with unsaturated phospholipids in biomembranes leading to lipid peroxidation. It is also involved in reactions causing, e.g., ischemia-reperfusion injury (restoration of tissue blood circulation after temporary restriction in blood supply).

Activated macrophages generate $O_2^{\bullet-}$ during inflammation by the cytochrome b_{558} containing NADPH oxidase ② and convert it by SOD to H_2O_2. This compound is used by myeloperoxidase to generate the aggressive hypochlorite (HOCl) for defense purposes ③.

$$NADPH + 2\ O_2 \rightarrow 2\ O_2^{\bullet-} + NADP^+ + H^+$$

$$2\ O_2^{\bullet-} + 2\ H^+ \rightarrow H_2O_2 + O_2$$

$$H_2O_2 + Cl^- \rightarrow HOCl + OH^-$$

In addition, $O_2^{\bullet-}$ reacts rapidly with NO (7.8.2) yielding peroxynitrite ④.

$$O_2^{\bullet-} + NO^{\bullet} \rightarrow O{=}N{-}O{-}O^-$$

This compound is highly reactive and produces, e.g., peroxides. At present, it is not exactly known if this reaction plays a role in NO metabolism.

Hydrogen peroxide (H_2O_2) is a reaction product of various oxidases (e.g., in peroxisomes, 2.2.2 and in the ER, 4.4.1) and a product of the SOD reaction ⑤.

H_2O_2 toxicity may be partly caused by direct inactivation of enzymes and (mostly reversible) of hemoproteins. H_2O_2 also can give rise to the very deleterious hydroxyl radical $HO^{\bullet}$ by reaction with semiquinones (⑥, see below). It is questionable, whether a metal ion dependent cleavage of H_2O_2 to form hydroxyl radicals (Fenton reaction) takes place *in vivo*, since free heavy metal ions are practically absent under these conditions.

The hydroxyl radical ($HO^{\bullet}$) is the most reactive, short lived and unspecific ROS. It reacts with nearly all biomolecules and may thereby start radical chain reactions ⑦.

$$HO^{\bullet} + \ldots{-}HCH{-}\ldots \rightarrow \ldots{-}HC^{\bullet}{-}\ldots + H_2O$$

Its main source is probably the reaction of H_2O_2 with semiquinones and other reductants. This may explain the toxicity of quinone compounds of foreign (xenobiotic) origin.

Peroxide radicals (ROO•) arise by the very fast, spontaneous reaction of oxygen with organic radicals (…–HC•–…, ⑧, which have been previously formed by hydroxyl or superoxide radicals in an attack reaction (see above). The peroxide radicals are intermediates in the chain reactions of lipid peroxidation (of membranes, LDL, etc.).

$$\ldots{-}HC^{\bullet}{-}\ldots + O_2 \rightarrow \ldots{-}HCOO^{\bullet}{-}\ldots$$

Singlet oxygen ($^1O_2$): This is the only physiologically relevant 'excited state' of molecular oxygen. Singlet oxygen may be produced during photosynthesis or non-enzymatically by reactions of hypochlorite ⑨ and peroxynitrite (see above).

$$HO{-}X + H_2O_2 \rightarrow H^+ + X^- + H_2O + {}^1O_2 \text{ (e.g., X = Cl)}$$

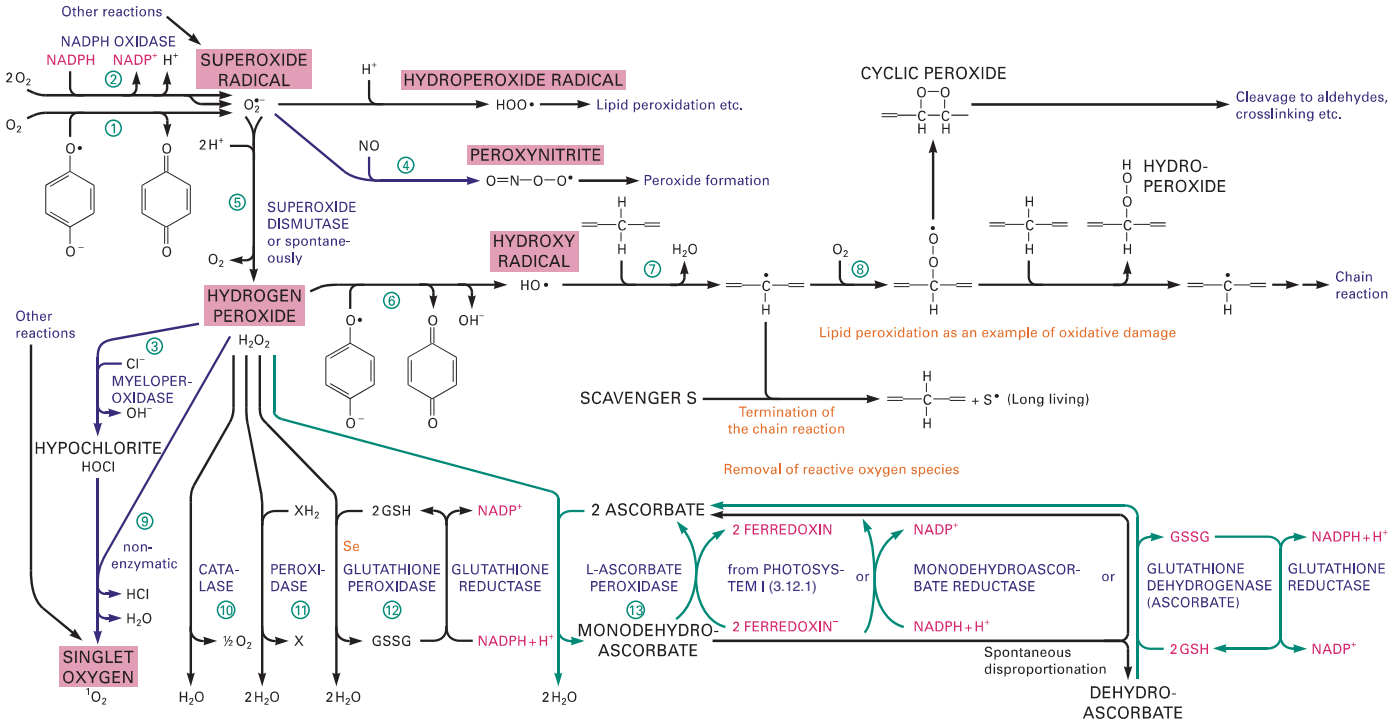


Figure 3.2.5-6. Generation of Reactive Oxygen Species (Red Background) and Their Removal

In oxidation reactions, singlet oxygen directly accepts 2 electrons with the normal antiparallel spin, while 'regular' oxygen would require 2 electrons with the uncommon parallel spin. Thus it is usually restricted to consecutive 1-electron transfer reactions (e.g., those catalyzed by flavin coenzymes, 3.7.3.2), while $^1O_2$ can react either way. This oxygen species is, e.g., involved in photosensitization processes (illumination of FMN, retinal, porphyrins etc.). This can be the cause of dermatoses.

Regulatory functions of reactive oxygen species: Some ROS appear to be involved in signal transduction, e.g., by oxidative activation of transcription factor NF-κB or by modulation of tyrosine phosphorylation cascades (7.5.2).

Antioxidant defenses: In all aerobic living organisms, both $O_2^{\bullet-}$ and H_2O_2 must be removed for protection. Various superoxide dismutases (SOD) disproportionate $O_2^{\bullet-}$ but still yield H_2O_2 as in the spontaneous reaction ⑤. This compound is destroyed by the ubiquitously occurring catalase ⑩ and by various peroxidases ⑪ including glutathione peroxidase ⑫ (in animals, it also removes organic peroxides) and ascorbate peroxidase ⑬ (in plants). The resulting monodehydroascorbate can be recycled by ferredoxin or by NADPH, while dehydroascorbate is recycled by GSH. Since NADPH is required for the reconversion of GSSG into GSH, an adequate supply is essential. The critical situation in erythrocytes has been described in 3.2.5.7.

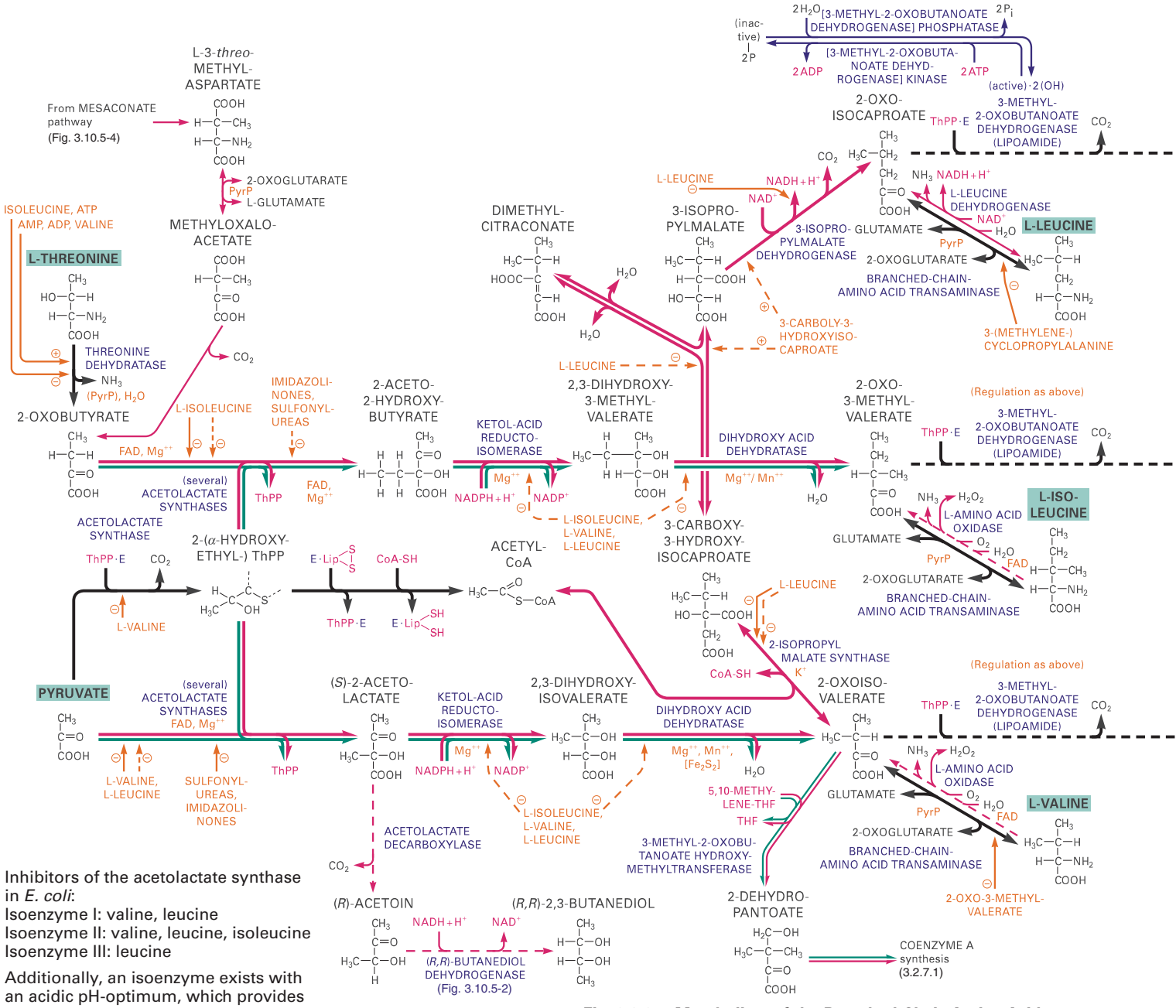
Molecular scavengers, such as α-tocopherol (vitamin E, 3.7.12), interrupt the sequence of radical transfers in peroxidative chain reactions by accepting the radical function themselves. These long-living radicals then react with themselves or with physiological antioxidants (ascorbate, uric acid). The singlet oxygen radical is preferably trapped by β-carotene (3.5.3.2), ascorbate, GSH and somewhat less by α-tocopherol.

Literature:

Finkel, T. Current Opinion in Cell Biology. 1998;10:248–253.

Fridovich, I. Ann. Rev. of Biochem. 1995;64:97–112.

Rice-Evans, C.A., Burdon, R.H. (Eds.) *Free Radical Damage and its Control. New Comprehensive Biochemistry* Vol. 28. Elsevier, 1994.



Inhibitors of the acetolactate synthase in *E. coli*:
Isoenzyme I: valine, leucine
Isoenzyme II: valine, leucine, isoleucine
Isoenzyme III: leucine

Additionally, an isoenzyme exists with an acidic pH-optimum, which provides acetolacetate for the butanediol fermentation

Fig. 3.2.6-1. Metabolism of the Branched-Chain Amino Acids

3.2.6 Leucine, Isoleucine and Valine

The essential amino acids L-leucine, L-isoleucine and L-valine are commonly named 'branched-chain amino acids'. Their carbon skeletons originate from pyruvate (in case of isoleucine additionally from threonine). Their biosynthesis pathways are closely interrelated.

Leucine is a strictly ketogenic amino acid (the degradation products are acetoacetate and acetyl-CoA), isoleucine is ketogenic (acetyl-CoA) as well as glucogenic (succinyl-CoA), while valine is completely glucogenic.

3.2.6.1 Biosynthetic Reactions (Fig. 3.2.6-1)

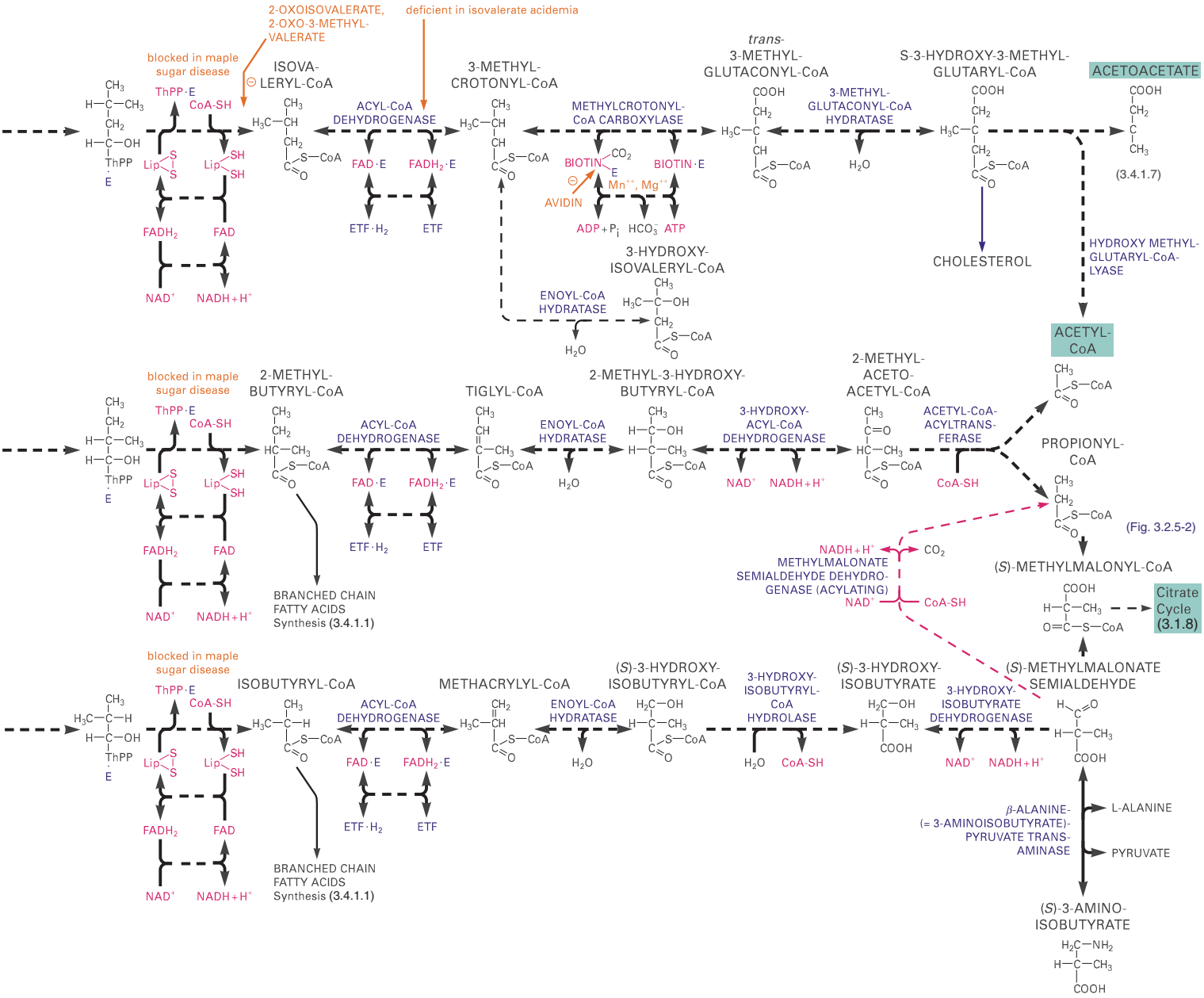
The **synthesis pathways** for valine and isoleucine in plants and bacteria begin with the condensation of a 2-oxo acid with 2-hydroxyethyl-ThPP (thiamine diphosphate), an intermediate of the pyruvate dehydrogenase reaction (3.1.3.1). This is followed by a reductive alkyl group migration reaction which results in dihydroxy acids. Dehydratation leads to the oxo acid precursors, which are converted to the amino acid by transamination.

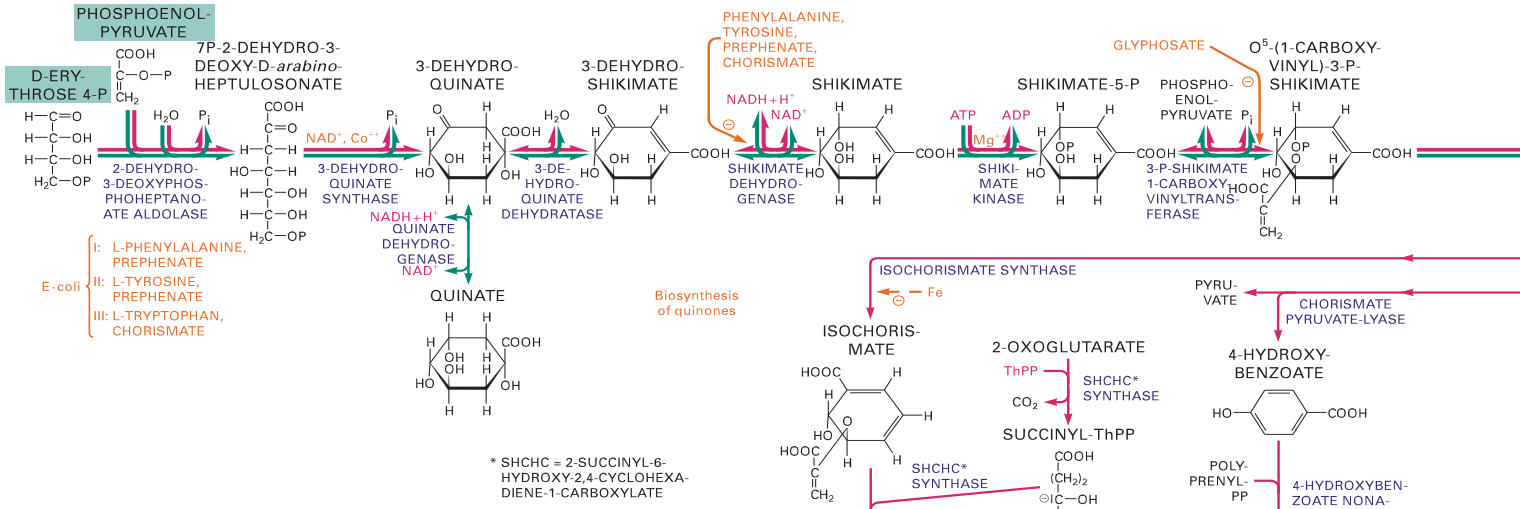
The leucine biosynthesis starts from the oxo acid precursor of valine and resembles the initial 3 steps of the citrate cycle (3.1.8.1).

In green plants, the biosyntheses of valine and isoleucine occur in the chloroplasts, while in yeasts, they are located in the mitochondria. The specific part of leucine biosynthesis in yeasts takes place either in the mitochondria or in the cytosol.

In some bacteria, L-3-threo-methylaspartate, an intermediate of the mesaconate pathway (Fig. 3.10.5-4), can also be converted into 2-oxobutyrate and act as a precursor of isoleucine. Acetolactate is also an intermediate during mixed-acid fermentation (Fig. 3.10.5-2), while 2-oxoisovalerate is the starting compound for coenzyme A synthesis (3.7.7.1).

Regulation: There are many antagonisms of the three amino acids in feedback regulation, control of expression and active transport. Isoleucine acts as a negative effector on threonine dehydratase, while valine antagonizes its effect. Control of the following biosynthesis reactions is complicated, since the pathways leading to valine





and isoleucine (and the beginning of leucine biosynthesis) proceed through identical steps. Frequently, multiple enzyme control takes place by individual end-product inhibition of isoenzymes (e.g., up to six acetolactate synthase isoenzymes in Enterobacteria). The regulation of gene expression varies between species. In *E. coli*, the genes for threonine dehydratase, dihydroxyacid dehydratase and branched-chain-amino-acid transaminase are located in a single operon, which is multivalently repressed by the 3 branched-chain amino acids. The genes for the initial acetolactate synthase isoenzymes are individually regulated.

The first specific enzyme of leucine biosynthesis, 2-isopropylmalate synthase, is feedback-inhibited by leucine, a feature, which generally occurs after branch points. The enzymes for the individual part of the leucine biosynthesis are encoded by a single operon, which is coordinately repressed or derepressed over a 1000-fold range.

In animals, these amino acids are preferably taken up by peripheral organs (skeletal and cardiac muscle, kidney) instead of by the liver. This is effected by the tissue distribution of the specific L-amino acid transporter.

3.2.6.2 Degradation of Branched-Chain Amino Acids (Fig. 3.2.6-1)

The initial catabolic steps for all 3 amino acids are catalyzed by the same mitochondrial enzymes. The sequence begins with reconversion to the respective oxo acids. Then a decarboxylation-dehydrogenation reaction similar to the pyruvate dehydrogenase reaction (3.1.3.1) follows. In isoleucine and valine catabolism, the resulting CoA derivatives undergo β-oxidation reactions analogously to the respective compounds during fatty acid degradation.

In isoleucine degradation, the resulting 2-methylacetoacetyl-CoA is cleaved to yield acetyl-CoA and propionyl-CoA. The latter compound is converted via methylmalonyl-CoA into succinyl-CoA as described in methionine catabolism (Fig. 3.2.5-2).

3-Hydroxyisobutyryl-CoA formed from valine, however, loses its CoA-group, is then oxidized to methylmalonate semialdehyde and then further on to methylmalonate, which is converted into the CoA derivative. The degradation steps end with succinyl-CoA. The 'CoA detour' in valine degradation is required, since the carbon chain is too short for regular β-oxidation. Some bacteria convert methylmalonate semialdehyde into propionyl-CoA instead. The intermediates of catabolism, isobutyryl-CoA and 2-methylbutyryl-CoA can also be used for the synthesis of branched-chain fatty acids (3.4.1.1).

While the first steps of leucine catabolism are analogous to both other amino acids, a biotin-dependent carboxylation step (3.7.8.2, analogous to acetyl-CoA carboxylase, 3.4.1.1) precedes the hydratase reaction. The final compound, 3-hydroxy-3-methylglutaryl-CoA is cleaved in mitochondria to acetyl-CoA and acetoacetate. It can be also used for cholesterol biosynthesis (3.5.1.1).

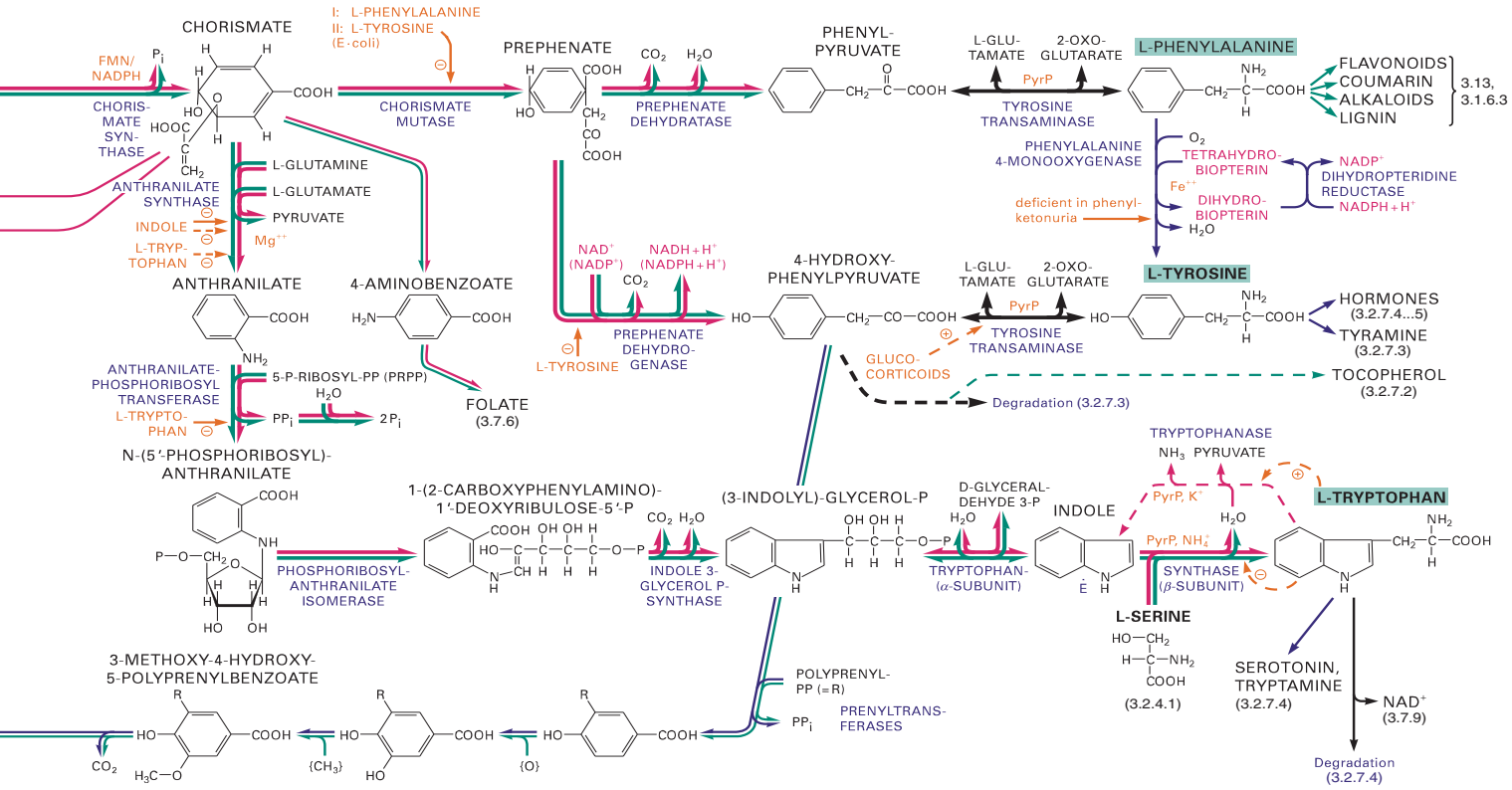


Figure 3.2.7-1. Biosynthesis of Aromatic Amino Acids and of Quinone Cofactors

In mammals, the first specific enzyme for catabolism, 3-methyl-2-oxobutanoate dehydrogenase (branched-chain 2-oxoacid dehydrogenase), is enzymatically inactivated by phosphorylation and activated by dephosphorylation. In resting muscle, the enzyme is usually in the inactivated state. Thus, liberated branched-chain amino acids are not catabolized, but released into the bloodstream. During exercise, the enzyme is activated and the amino acids are degraded and oxidized for energy production in the organ itself.

Defects in the initial decarboxylation-dehydrogenation enzyme system cause maple-sugar disease. This can occur with any of the 3 enzymes involved. There is an elevation in blood of all 3 branched chain amino acids and of the respective 2-oxo acids. Also 2-hydroxy acids are found, which arise by reduction of the oxo acids. Serious disturbances of the CNS and early death can be the consequences, if not treated by an appropriate diet.

Literature:

Harris, R.A. *et al.* J. Nutr. 1994;124:1499S–1502S.

Kishore, G.M., Shah, D.M. Ann. Rev. of Biochem. 1978;57:627–663.

Patel, M.S., Harris, R.A. FASEB J. 1995;9:1164–1172.

Umbarger, H.E. Ann. Rev. of Biochem. 1978;47:533–606.

3.2.7 Phenylalanine, Tyrosine, Tryptophan and Derivatives

The amino acids of this group are commonly named 'aromatic amino acids'. Tryptophan and phenylalanine are essential for mammals. This is indirectly also valid for tyrosine, since it is formed from phenylalanine (bacteria, however, can bypass the phenylalanine step). The first part of biosynthesis is common to all three compounds.

NAD, NADP and a number of important biogenic amines and quinone compounds are derived from these amino acids or from their precursors. Tryptophan is frequently the limiting amino acid for protein biosynthesis.

The degradation of all these amino acids yields ketogenic acetoacetate. Tyrosine and phenylalanine also produce glucogenic fumarate. Tryptophan yields alanine, which can be transaminated to glucogenic pyruvate.

3.2.7.1 Biosynthesis of Aromatic Amino Acids (Fig. 3.2.7-1)

Common ('shikimate') pathway: The biosynthesis starts with phosphoenolpyruvate (PEP) and erythrose 4-P (from the pentose-P cycle, 3.1.6.1). Their condensation product cyclizes in a reaction requiring only catalytic amounts of NAD. Formation of two double bonds and introduction of another molecule phosphoenolpyruvate yield chorismate. The latter reaction is inhibited competitively to PEP by glyphosate ($(OH)_2OP\text{-}CH_2\text{-}NH\text{-}CH_2\text{-}COOH$, phosphonomethyl-amino-acetic acid), a widely used weed killer (Round-up®). Interestingly, other PEP-reactions are not inhibited.

Chorismate is a branch point for different pathways leading to phenylalanine and tyrosine, to tryptophan, ubiquinone, phylloquinone and menaquinone (3.2.7.2), to 4-aminobenzoate (3.7.6.1) and to plant secondary metabolites (3.13).

Phenylalanine and tyrosine biosynthesis: A mutase reaction yields prephenate, which undergoes either a dehydratase/decarboxylase or a reduction/decarboxylase reaction (in some bacteria) resulting in aromatization. The oxo analogues of phenylalanine and tyrosine obtained in this way are then converted into the respective amino acids by transamination. A variant is the arogenate pathway in plants (3.13.1.1).

Mammals obtain tyrosine from phenylalanine by a hydroxylation reaction catalyzed by phenylalanine 4-monooxygenase, which uses tetrahydrobiopterin as a hydrogen donor (3.7.6.3). The reduced tetrahydrobioperin is regenerated by NAD(P)H, catalyzed by dihydrofolate reductase.

The conversion into tyrosine is also the initial step of phenylalanine degradation (3.2.7.3). Disturbances of this reaction lead to phenylketonuria.

Tryptophan biosynthesis: Removal of the phosphoenolpyruvate moiety from chorismate and amination yields the aromatic compound anthranilate. Condensation with 5-P-ribosyl-PP (compare 3.6.1.1, an activated form of ribose 5-P) supplies two carbons to the 5-membered tryptophan ring, the rest of the ribose moiety is then cleaved off and replaced by L-serine.

In *E. coli*, these two final reactions are performed by the α- (29 kDa) and β- (43 kDa) subunits of the tetrameric enzyme tryptophan synthase ($\alpha_2\beta_2$). The combination of the (individually active) subunits greatly increases the reaction rate. This is due to direct transfer of the indole intermediate between the α- and β-subunits through a 'tunnel' in the enzyme ('channeling effect'), which prevents loss of this lipophilic compound through membranes. Serine forms a Schiff base with pyridoxal phosphate present in the β-subunit, which is then dehydrated and nucleophilically attacked by indole.

Regulation: The regulatory systems differ greatly among species. Two examples are schematically shown in Figure 3.2.7-2. In *Bacillus subtilis*, each amino acid inhibits the first enzyme of its individual pathway after the respective branch point. Prephenate and chorismate, which are the end products of the common part of the pathway, inhibit its initial reaction (sequential control). Some of these enzymes are bifunctional.

A different system exists in *E. coli*: the initial reaction of the common pathway is catalyzed by 3 isoenzymes; each of which is inhibited by an amino acid endproduct (multiple enzyme control). After the first branch point, anthranilate synthase is inhibited by tryptophan, while the 2 bifunctional isoenzymes catalyzing the reactions to phenylalanine and tyrosine are correspondingly inhibited by these amino acids. After branching of these pathways, there is another inhibition by the respective amino acids (compare Fig. 3.2.5-1).

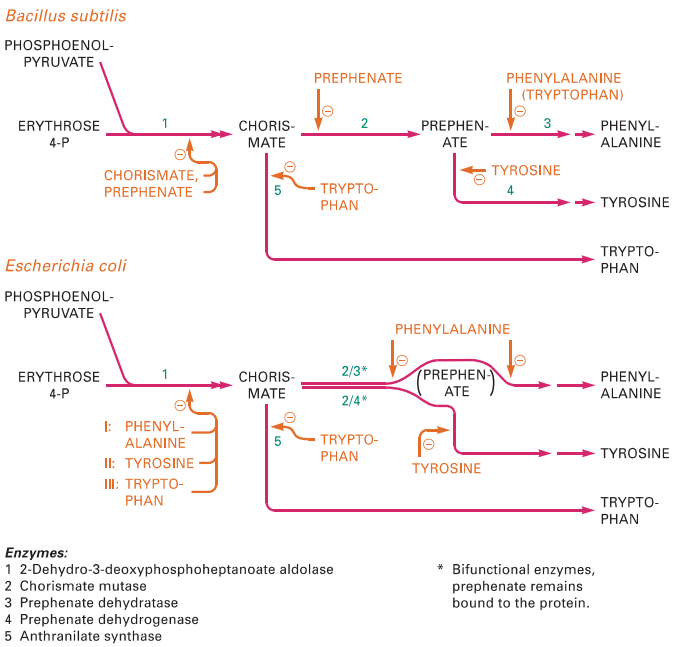
In *Neurospora crassa*, the inhibition pattern of the three isoenzymes catalyzing the first enzyme of the common pathway is analogous to *E. coli*. Anthranilate synthase is an individual enzyme. The other enzymes of the common pathway and of the anthranilate pathway (after the branch point) form multi-enzyme complexes, which allow catalytic facilitation by direct transfer to the next enzymatic function. Enzyme associations or multifunctional enzymes occur also in other organisms.

3.2.7.2 Biosynthesis of Quinone Cofactors

The quinone cofactors play an essential role in hydrogen transfer reactions, e.g., during aerobic and anaerobic respiration (3.11.4; 3.10.6) and photosynthesis (3.12.1). They form the 'quinone pool' in membranes. Phylloquinone is the cofactor for glutamate carboxylation, which is necessary for coagulation reactions (9.3.1).

Menaquinone and phylloquinone biosynthesis in bacteria (Fig. 3.2.7-1): By action of a mutase, chorismate is converted into isochorismate. A thiamine-PP requiring reaction performs the decarboxylation of oxoglutarate. Then a condensation with isochorismate and the removal of pyruvate takes place. After dehydratation (causing aromatization), the dihydroxynaphtoate structure is formed. It reacts with unsaturated polyprenyl-PP of various lengths (up to n = 15 isoprene units, Fig. 3.5.3-2), followed by methylation of the quinol and oxidation to menaquinone (vitamin K_2). The plant product phylloquinone is synthesized similarly. It differs in having a phytyl side chain of four isoprene units with a single double bond (vitamin K_1).

Ubiquinone biosynthesis (Fig. 3.2.7-1): In bacteria, the reaction starts from chorismate. After the removal of pyruvate, the 4-hydroxybenzoate formed is prenylated and decarboxylated. This is followed by hydroxylation and methylation reactions. In yeast and in animals, the starting compound for biosynthesis is 4-hydroxyphenylpyruvate (which can be obtained from tyrosine by transamination). The attachment of a polyprenyl chain (6 … 10 isoprene units, Fig. 3.5.3-2) is followed by hydroxylation, methylation and decarboxylation. The last steps are identical to the pathway in bacteria.



Enzymes:
1 2-Dehydro-3-deoxyphosphoheptanoate aldolase
2 Chorismate mutase
3 Prephenate dehydratase
4 Prephenate dehydrogenase
5 Anthranilate synthase

* Bifunctional enzymes, prephenate remains bound to the protein.

Figure 3.2.7-2. Regulation of Aromatic Amino Acid Biosynthesis

3.2.7.3 Derivatives and Degradation of Aromatic Amino Acids

Phenylalanine derivatives: The most important compounds directly derived from phenylalanine are found in plants. They are described in 3.13.1.

Tyrosine derivatives (Fig. 3.2.7-3): The product of tyrosine hydroxylation by tyrosine 3-monooxygenase is dihydroxyphenylalanine (DOPA). This reaction resembles the conversion of phenylalanine to tyrosine and also employs tetrahydrobiopterin. The result of DOPA decarboxylation is dopamine. This compound is taken up by neurons and acts as a neurotransmitter (Table 7.2-2). An insufficient supply leads to Parkinson's syndrome. By a second hydroxylation at the 3-position with the participation of ascorbate, the important hormones norepinephrine (noradrenaline, 3.2.7.4) and further on epinephrine (adrenaline) are generated. Iodination of tyrosine leads to thyroid hormones (3.2.7.5).

In mammalian melanocytes, tyrosine is oxidized to DOPA and further on to dopaquinone, which after cyclization and oxidative condensation reactions yield eumelanins (dark pigments of the skin). This reaction is activated by irradiation. If the enzyme is deficient, albinism occurs. Similar DOPA oxidations in plants (by phenoloxidases) lead to darkening of cut fruits or branches. Decarboxylation of tyrosine without previous hydroxylation (e.g., by intestinal bacteria) yields tyramine, which elevates blood pressure.

Phenylalanine and tyrosine degradation (Fig. 3.2.7-3): Fermenting bacteria usually degrade both compounds by deamination and shortening of the side chain. In animals, however, the degradation of phenylalanine starts with its conversion to tyrosine (3.2.7.1).

Disturbances of the monooxygenase system lead to phenylketonuria, a very frequent hereditary disease in humans (frequency ca. 1:10,000). Transamination to phenylpyruvate and consecutive reduction to phenyllactate or other reactions take place. The metabolites disturb neuronal development. The disease is treated by a low-phenylalanine diet.

For degradation, tyrosine undergoes transamination, followed by a decarboxylating dioxygenase reaction, yielding homogentisate. This enzyme requires the presence of ascorbate (or of another reducing agent). A second dioxygenation opens up the aromatic ring. The final products are fumarate and acetoacetate.

In plants, homogentisate is the precursor of plastoquinone and of tocopherol (vitamin E, 3.7.12). Their formation begins by the attachment of a polyprenyl-PP residue (for plastoquinone) or of a phytyl-PP residue (for tocopherol), followed by methylations and in case of tocopherol by a cyclization reaction.

Tryptophan derivatives (Fig. 3.2.7-4): This amino acid is an important source of biogenic amines. Via 5-hydroxylation and decarboxylation, serotonin is generated, which ubiquitously occurs in animals (as a neurotransmitter, Table 7.2-2) and in plants. Its formation in the brain is regulated by the tryptophan concentration (due to a high K_M of tryptophan hydroxylase, 2.5.2). The first step of serotonin catabolism is an oxidation by monoamine oxidase type A. The end product of serotonin degradation is 5-hydroxyindoleacetate, which is secreted in urine.

After its synthesis in the CNS and in gastrointestinal cells of higher animals, serotonin is interneuronally transported to vesicles at the nerve endings or it is released into the bloodstream, where it is transported by platelets to receptors, which are located on smooth muscle, endothelial and epithelial cells, neurons and platelets. See Table 7.2-2 for neuronal effects. In addition, serotonin relaxes the smooth muscles of blood vessels and of the gastrointestinal tract.

In the pineal gland, serotonin is converted by acetylation and methylation into melatonin. This amine is synthesized and secreted in pronounced diurnal rhythm and acts as an antagonist to melanotropin (Fig. 7.1-4). In mammals it antagonizes the function of the thyroid gland and the secretion of luteinizing hormone (7.1.5) and generally lowers metabolism ('sleep hormone'). Catabolism of melatonin takes place by hydroxylation and excretion in urine.

By direct decarboxylation of tryptophan without previous hydroxylation, tryptamine is obtained. Further oxidation of the side chain leads to indole acetate (auxin), which is a growth factor for plants. Alternatively, its biosynthesis can proceed via indolepyruvate. As well as in plants, tryptamine formation takes place in bacteria, in the kidney and the liver.

Tryptophan is the precursor of NAD and NADP (3.7.9.1). The pathway for synthesis of these coenzymes branches off from tryptophan catabolism.

Tryptophan degradation (Fig. 3.2.7-4): The major catabolic pathway in mammalian liver begins with the oxidative cleavage of the 5-membered ring and the consecutive liberation of formate (which is taken up by tetrahydrofolate, 3.7.6.2), resulting in kynurenine. After hydroxylation, a shortening of the side chain and release of alanine by action of kynureninase takes place.

The cytoplasmic kynureninase reaction is a β,γ-elimination requiring pyridoxal phosphate. If this cofactor is present only at low levels, the substrates are rather transaminated to kynurenate and xanthurenate (mostly by a mitochondrial transaminase which is less affected by

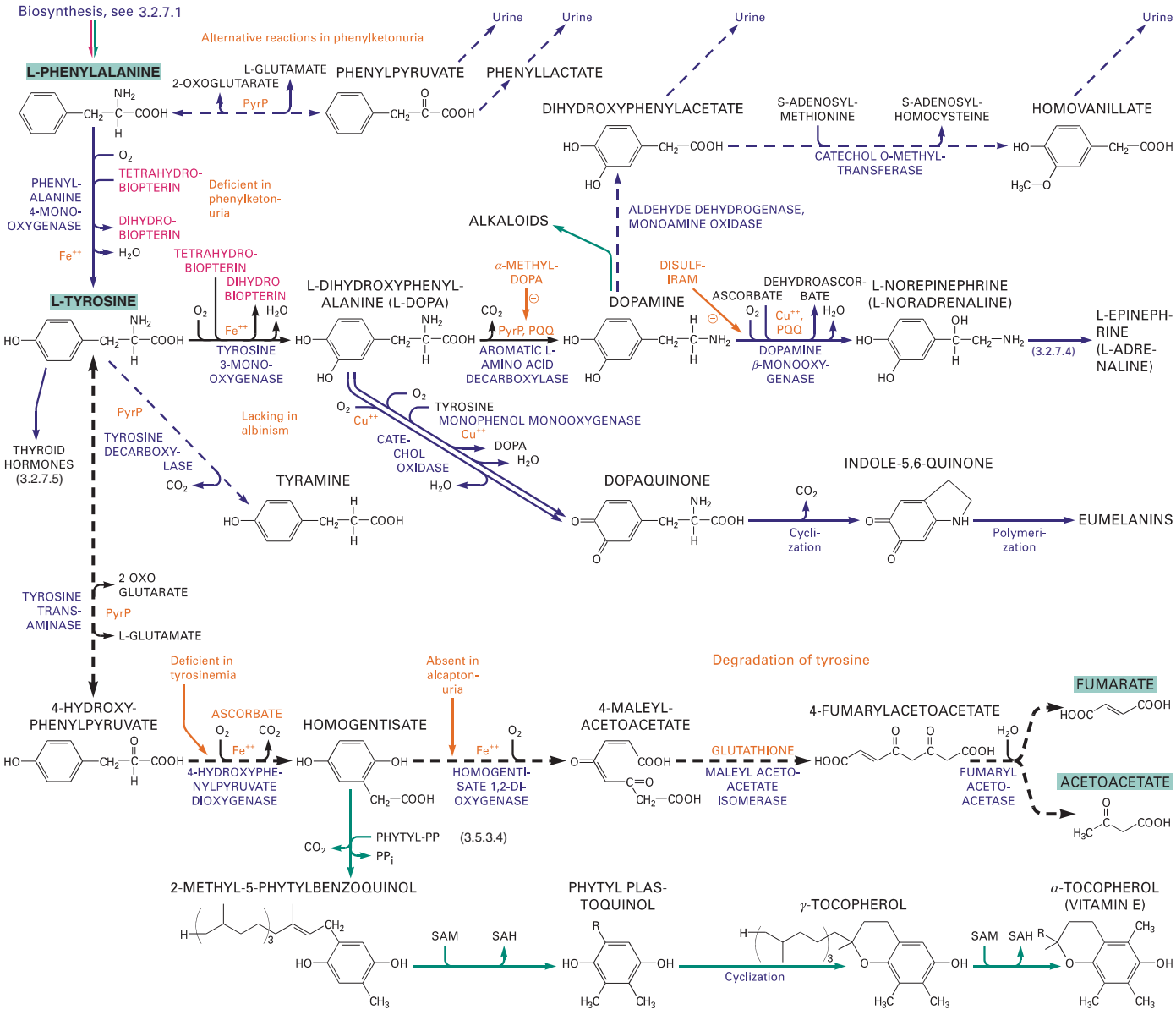


Figure 3.2.7-3. Derivatives and Degradation of Phenylalanine and Tyrosine

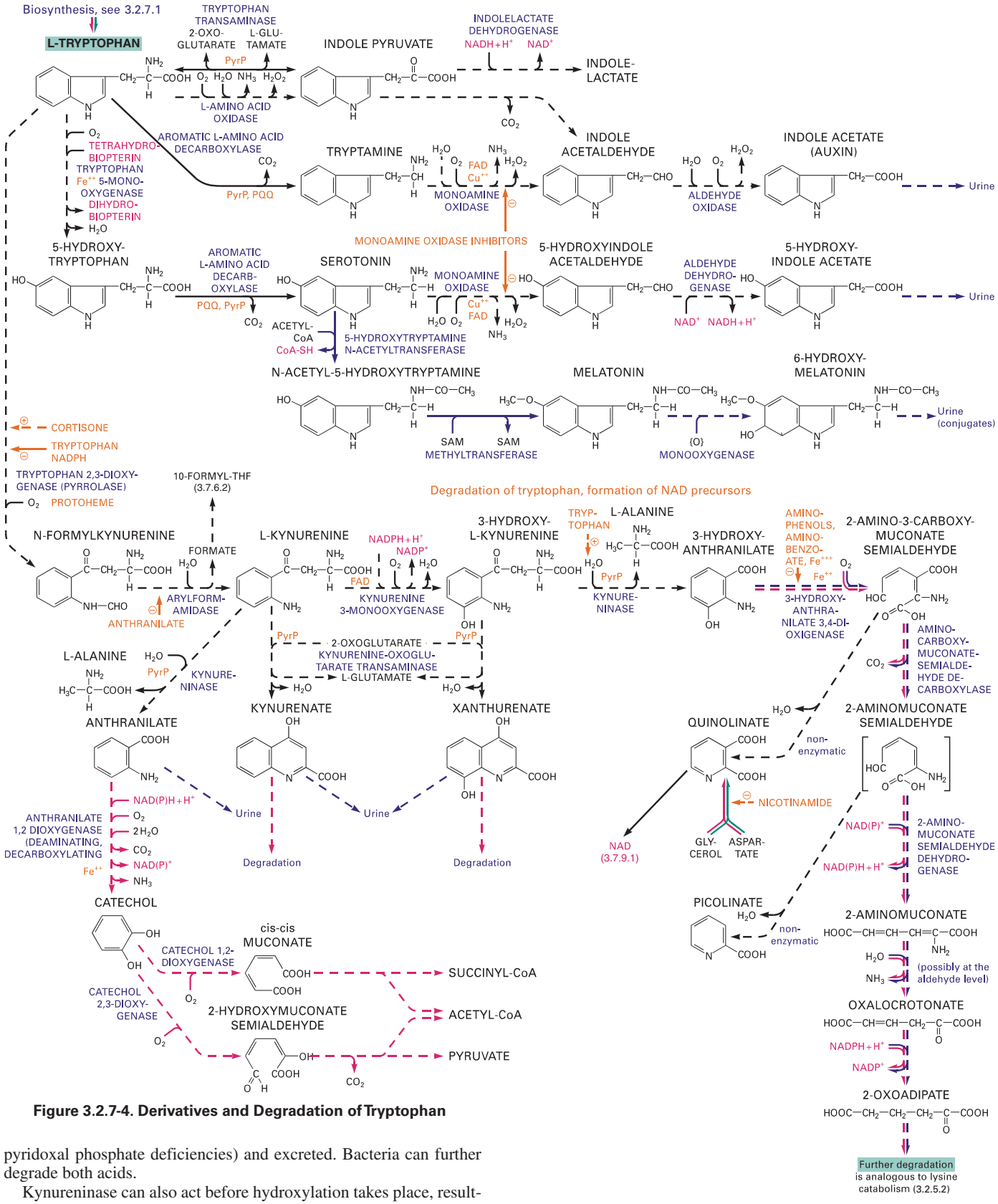


Figure 3.2.7-4. Derivatives and Degradation of Tryptophan

pyridoxal phosphate deficiencies) and excreted. Bacteria can further degrade both acids.

Kynureninase can also act before hydroxylation takes place, resulting in anthranilate. Its conversion to catechol, ring opening and further degradation is a common way of tryptophan catabolism in bacteria. Animals, however, excrete anthranilate in conjugated form in urine.

After the kynureninase step in mammalian metabolism, the action of 3-hydroxyanthranilate 3,4-dioxygenase causes release of alanine and ring opening. The resulting semialdehyde can spontaneously cyclize to quinolinate, which is the initial compound for NAD and NADP biosynthesis (3.7.9.1). The same cyclization reaction can also take place after the β-carboxyl group has been removed, resulting in picolinate. The main degradation pathway leads to crotonyl-CoA (which is a member of the fatty acid degradation pathway) and further on to acetyl-CoA. The steps after 2-oxoadipate are also identical with lysine catabolism (3.2.5.2).

Literature:

Barker, H.A. Ann. Rev. of Biochem. 1981;50:23–40.

Bender, D.A. *Amino Acid Metabolism.* 2nd Ed. Wiley, 1985.

Crawford, I.P., Stauffer, G.V. Ann. Rev. of Biochem. 1980;49:163–195.

Eisensmith, R.C., Woo, S.L.C. Mol. Biol. Med. 1991;8:3–18.

Kaufman, S. Adv. in Enzymology 1993;67:77–264.

Kishore, G.M., Shah. D.M. Ann. Rev. of Biochem. 1988;57:527–663.

Pau, P. *et al.* Trends in Biochem. Sci. 1997;22:22–27.

Umbarger, H.E. Ann. Rev. of Biochem. 1978;47:533–606.

3.2.7.4 Catecholamines

Catecholamines is the common term for the important hormones norepinephrine (noradrenaline) and epinephrine (adrenaline). These compounds are produced by the adrenal medulla and by postganglionic nerve terminals, which are members of the adrenergic nervous system of vertebrates. While in some species norepinephrine is the major product, humans, dogs and some other species produce mostly epinephrine. These compounds are also found in lower animals and plants.

Biosynthesis (Fig. 3.2.7-5): Hydroxylation of tyrosine and consecutive decarboxylation yields dopamine (3.2.7.3). This compound is transferred by a specific carrier from the cytosol into chromaffin granules. Dopamine β-monooxygenase, a PQQ (3.10.8) and Cu^{++} dependent enzyme, catalyzes the hydroxylation to norepinephrine. A consecutive methyl transfer from S-adenosylmethionine, catalyzed by phenylethanolamine N-methyltransferase, leads to epinephrine. The catecholamines are stored in the granules and are released upon arrival of signals from the sympathetic nervous system.

Regulation of the biosynthesis (Details in Fig. 7.1-3): Norepinephrine inhibits the tyrosine monooxigenase, epinephrine inhibits the methyltransferase by allosteric mechanisms. Neuronal signals induce the expression of both biosynthetic hydroxylases, while glucocorticoids induce the expression of the methyltransferase.

Hormone effects: The hormones act on α- and β-receptors, which are present in different organs. Although stimulation of both receptor types often causes opposite effects, they cooperate in response to physical or emotional stress situations by enhancing the energy supply and directing it to the sites where it is primarily needed (by influencing the circulation, increasing the heart output etc.). For details, see 7.1.4 and Table 7.2-2.

The antihypertensive drugs α-methyltyrosine and α-methyldopa are metabolized to the α-methyl analogue of epinephrine (side chain methylated), inhibit the hormone synthesis and antagonize the hormone action at the receptor.

Degradation (Fig. 3.2.7-5): The degradation starts by methylation and oxidation reactions, catalyzed by catechol-O-methyltransferase (COMT) and monoamine oxidase (MAO), respectively. The products are excreted in urine, some of them in conjugated form.

3.2.7.5 Thyroid Hormones

Thyroid hormones (T_3 = 3,5,3′-triiodothyronine and T_4 = thyroxine = 3,5,3′,5′-tetraiodothyronine) are the other important hormone group in vertebrates derived from tyrosine.

Biosynthesis (Fig. 3.2.7-6): The formation of the thyroid hormones takes place in the follicles of the thyroid gland. The epithelial cells synthesize thyreoglobulin (dimeric, 2 * 330 kDa, 2 * 72 tyrosyl residues).

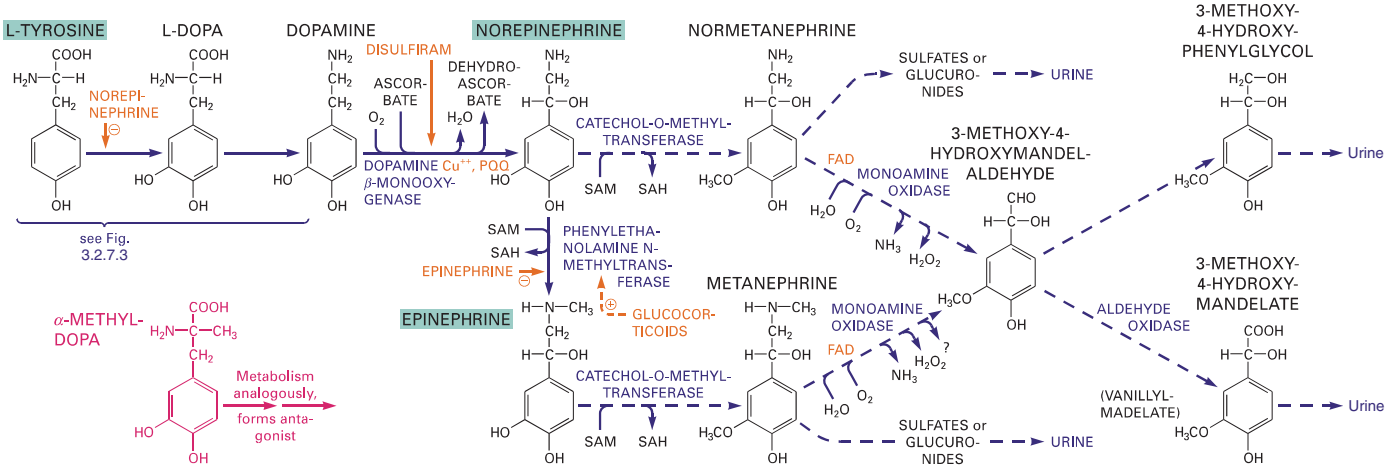


Figure 3.2.7-5. Biosynthesis and Degradation of Catecholamines

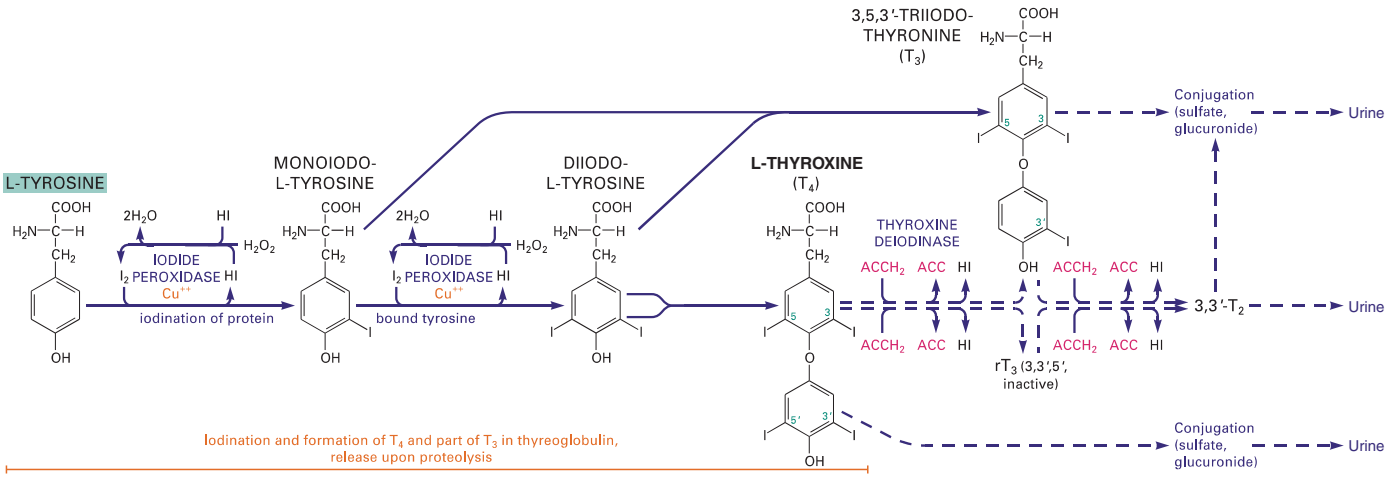
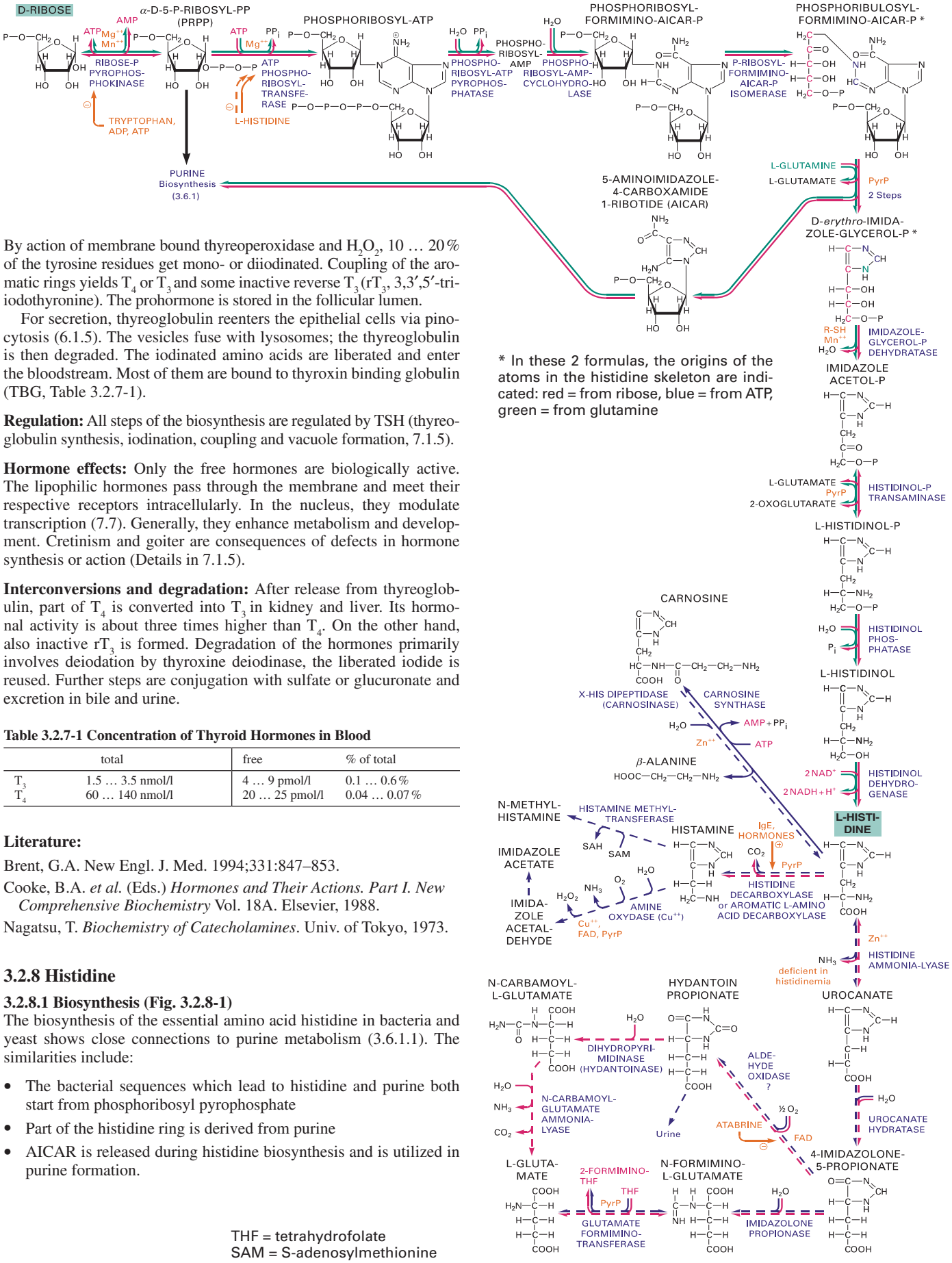


Figure 3.2.7-6. Biosynthesis and Degradation of Thyroid Hormones



* In these 2 formulas, the origins of the atoms in the histidine skeleton are indicated: red = from ribose, blue = from ATP, green = from glutamine

Figure 3.2.8-1. Biosynthesis and Degradation of Histidine

THF = tetrahydrofolate
SAM = S-adenosylmethionine
SAH = S-adenosylhomocysteine

By action of membrane bound thyreoperoxidase and H_2O_2, 10 … 20 % of the tyrosine residues get mono- or diiodinated. Coupling of the aromatic rings yields T_4 or T_3 and some inactive reverse T_3 (rT_3, 3,3′,5′-triiodothyronine). The prohormone is stored in the follicular lumen.

For secretion, thyreoglobulin reenters the epithelial cells via pinocytosis (6.1.5). The vesicles fuse with lysosomes; the thyreoglobulin is then degraded. The iodinated amino acids are liberated and enter the bloodstream. Most of them are bound to thyroxin binding globulin (TBG, Table 3.2.7-1).

Regulation: All steps of the biosynthesis are regulated by TSH (thyreoglobulin synthesis, iodination, coupling and vacuole formation, 7.1.5).

Hormone effects: Only the free hormones are biologically active. The lipophilic hormones pass through the membrane and meet their respective receptors intracellularly. In the nucleus, they modulate transcription (7.7). Generally, they enhance metabolism and development. Cretinism and goiter are consequences of defects in hormone synthesis or action (Details in 7.1.5).

Interconversions and degradation: After release from thyreoglobulin, part of T_4 is converted into T_3 in kidney and liver. Its hormonal activity is about three times higher than T_4. On the other hand, also inactive rT_3 is formed. Degradation of the hormones primarily involves deiodation by thyroxine deiodinase, the liberated iodide is reused. Further steps are conjugation with sulfate or glucuronate and excretion in bile and urine.

Table 3.2.7-1 Concentration of Thyroid Hormones in Blood

	total	free	% of total
T_3	1.5 … 3.5 nmol/l	4 … 9 pmol/l	0.1 … 0.6 %
T_4	60 … 140 nmol/l	20 … 25 pmol/l	0.04 … 0.07 %

Literature:

Brent, G.A. New Engl. J. Med. 1994;331:847–853.

Cooke, B.A. *et al.* (Eds.) *Hormones and Their Actions. Part I. New Comprehensive Biochemistry* Vol. 18A. Elsevier, 1988.

Nagatsu, T. *Biochemistry of Catecholamines*. Univ. of Tokyo, 1973.

3.2.8 Histidine

3.2.8.1 Biosynthesis (Fig. 3.2.8-1)

The biosynthesis of the essential amino acid histidine in bacteria and yeast shows close connections to purine metabolism (3.6.1.1). The similarities include:

- The bacterial sequences which lead to histidine and purine both start from phosphoribosyl pyrophosphate
- Part of the histidine ring is derived from purine
- AICAR is released during histidine biosynthesis and is utilized in purine formation.

This connection and the presence of the catalytically active imidazole group in these compounds might be a hint for a phylogenetic development from the 'RNA world' to the 'protein world' and from ribozyme (e.g., 4.1.3.2) to enzyme catalysis. The pathway in higher plants has not been elucidated in all details.

The activated compound phosphoribosyl-PP reacts with the purine ring of ATP, which is thereafter opened. After isomerization, cleavage of this molecule releases 5′P-ribosyl-5-amino-4-imidazole carboxamide (AICAR). This is also an intermediate of purine formation (3.6.1.1) and can be converted with a few steps into ATP, thus recovering this compound, which has been previously utilized for histidine biosynthesis. During the same reaction step, glutamine adds an amino group to the rest of the molecule, then closure of the imidazole ring takes place. After 4 more steps, histidine is obtained. Histidine inhibits the first individual step of the pathway, ATP-phosphoribosyltransferase.

3.2.8.2 Interconversions and Degradation (Fig. 3.2.8-1)

An important product obtained by histidine decarboxylation in animals is the biogenic amine histamine. The reaction proceeds in mast cells and in many tissue cells. Histamine is released from mast cells in IgE-mediated hypersensitivity of the immediate type (8.3.1, allergic reactions). It acts on specific receptors (Table 7.2-2). This leads to bronchoconstriction and to NO release, which in turn results in vasodilatation (7.8.2). Histamine also stimulates the secretion of HCl into the stomach. Histamine formation as well as degradation is hormone or cytokine controlled. Histamine is also present in stinging nettles, bee venom, secretion of mosquitoes etc. The development of antihistamines is an important issue in pharmacology.

The dipeptide carnosine (β-alanylhistidine) occurs in large quantities in muscles. It is synthesized in an ATP-dependent reaction from histidine and β-alanine. It provides histidine for metabolism if its supply by food is insufficient. Another muscle component is 3-methylhistidine, which is present in actin and myosin (7.4.5). The methylation is a post-transcriptional reaction.

Histidine degradation starts with a non-oxidative deamination resulting in urocanate. This is followed by hydratation (which saturates two double bonds!) and hydrolytic opening of the ring. The resulting formiminoglutamate transfers its formimino group to tetrahydrofolate (3.7.6.2), glutamate is obtained. Histidinemia results from deficiency of histidine ammonia-lyase. In the case of folate deficiency, formiminogluatamate appears in urine.

Literature:

Umbarger, H.E. Ann. Rev. of Biochem. 1978;47:533–606.

3.2.9 Urea Cycle, Arginine and Associated Reactions

In animals, ammonia, which is liberated by the degradation of amino acids and which is not incorporated into other compounds, has to be excreted.

Animals living in water can easily do this with the free compound (ammonotelic animals). This requires the simultaneous, constant excretion of water, which is not possible in terrestrial animals. Due to the toxicity of the free compound, these animals convert ammonia into nontoxic urea in the liver (urotelic animals). Birds, terrestrial reptiles and many insects, however, excrete uric acid (3.6.1.6, uricotelic animals).

3.2.9.1 Urea Cycle (Krebs-Henseleit Cycle, Fig. 3.2.9-1, next page)

The urea cycle is not only a means to remove ammonia from the body, but CO_2 as well. In animals, plants and bacteria, it is also the source of the proteinogenic amino acid arginine, of polyamines and guanidino compounds (3.2.9.2).

In animals, the first steps of the urea cycle proceed in the mitochondria of the liver. While ammonia release due to degradation reactions already takes place there, a major part of CO_2 has to pass from the cytosol into the mitochondria, where it is converted into bicarbonate by carbonate dehydratase (carboanhydrase). The ATP-driven carbamoyl-P synthesis by carbamoyl synthase I (ammonia) from NH_3 and HCO_3^- requires N-acetylglutamate (3.2.2.2) as an allosteric activator.

The concentration of N-acetylglutamate parallels the glutamate concentration, thus it apparently acts as a sensor of glutamate increase resulting from elevated protein degradation or supply from food.

The glutamine hydrolyzing carbamoyl synthase II occurs in the cytosol and is involved in pyrimidine biosynthesis (3.6.2.1). In bacteria, this is the only type of carbamoyl synthase.

The condensation reaction with ornithine (biosynthesis see 3.2.2.2) yields citrulline (both amino acids are not proteinogenic). Citrulline leaves the mitochondria and condenses with aspartate in an ATP-dependent reaction to argininosuccinate. Cleavage of this compound yields the proteinogenic amino acid arginine and fumarate. The latter compound is converted via malate to oxaloacetate. Although these reactions are identical to citrate cycle reactions, they take place in the cytosol. Oxaloacetate is reconverted by transamination into aspartate (aspartate cycle). Finally, arginase cleaves arginine, liberates urea and reconstitutes ornithine, which returns to the mitochondria.

While carbamoyl-P synthase and ornithine carbamoyl transferase are associated with the inner membrane of liver mitochondria, the other enzymes of the cycle are associated with the outside of the outer membrane. Thus they are in close proximity (forming a metabolon), which allows direct substrate passage from one enzyme to the next (substrate channeling).

The urea cycle is energized by the hydrolysis of 4 energy-rich phosphate bonds (3 ATP → 2 ADP + AMP). However, they are reconstituted during oxidative phosphorylation, using the NADHs formed by the glutamate dehydrogenase (which releases ammonia) and malate dehydrogenase reactions.

The expression of all enzymes of the urea cycle in animals is increased during protein-rich diets. Also in bacteria, the expression of all enzymes of arginine synthesis is co-regulated by a common mechanism (regulon, 4.1.2.2).

Minor activities of the urea cycle enzymes are also found in other organs besides the liver. Apparently their major task is the biosynthesis of arginine or the provision of ornithine for synthesis of polyamines (3.2.9.3).

Deficiencies of all enzymes of the urea cycle have been observed (Table 3.2.9-1, combined homozygotic frequencies of these diseases ca. 1/25,000). They lead to increased blood concentration of their particular substrates, but also of their precursors in the cycle as far back as ammonia. The primary effect of hyperammonemia is brain damage.

Degradation of arginine, as well as of the other urea cycle intermediates to glutamate occurs via glutamate semialdehyde by reversal of their biosynthesis (except the phosphorylation step).

Table 3.2.9-1. Diseases Caused by Deficiencies in Urea Cycle Enzymes

Deficient Enzyme	Disease
carbamoyl-P-synthase (ammonia)	hyperammonemia I
ornithine carbamoyltransferase	hyperammonemia II
argininosuccinate synthase	citrullinuria
argininosuccinate lyase	argininosuccinic aciduria
arginase	argininemia
N-acetylglutamate synthase (forms activator)	N-acetylglutamate synthase deficiency

Literature:

Batshaw, M.L. Ann. Neurol. 1994;35:133–141.

Grisolia, S. *et al. The urea cycle.* Wiley, 1976.

Meijer, A.J. *et al.* Physiol. Rev. 1990;70:701–748.

3.2.9.2 Phosphagens (Phosphocreatine and Phosphoarginine, Fig. 3.2.9-1)

In vertebrates, the guanidino residue of arginine can be transferred to glycine by action of glycine amidinotransferase. This reaction yields guanidinoacetate and ornithine. Guanidinoacetate is further

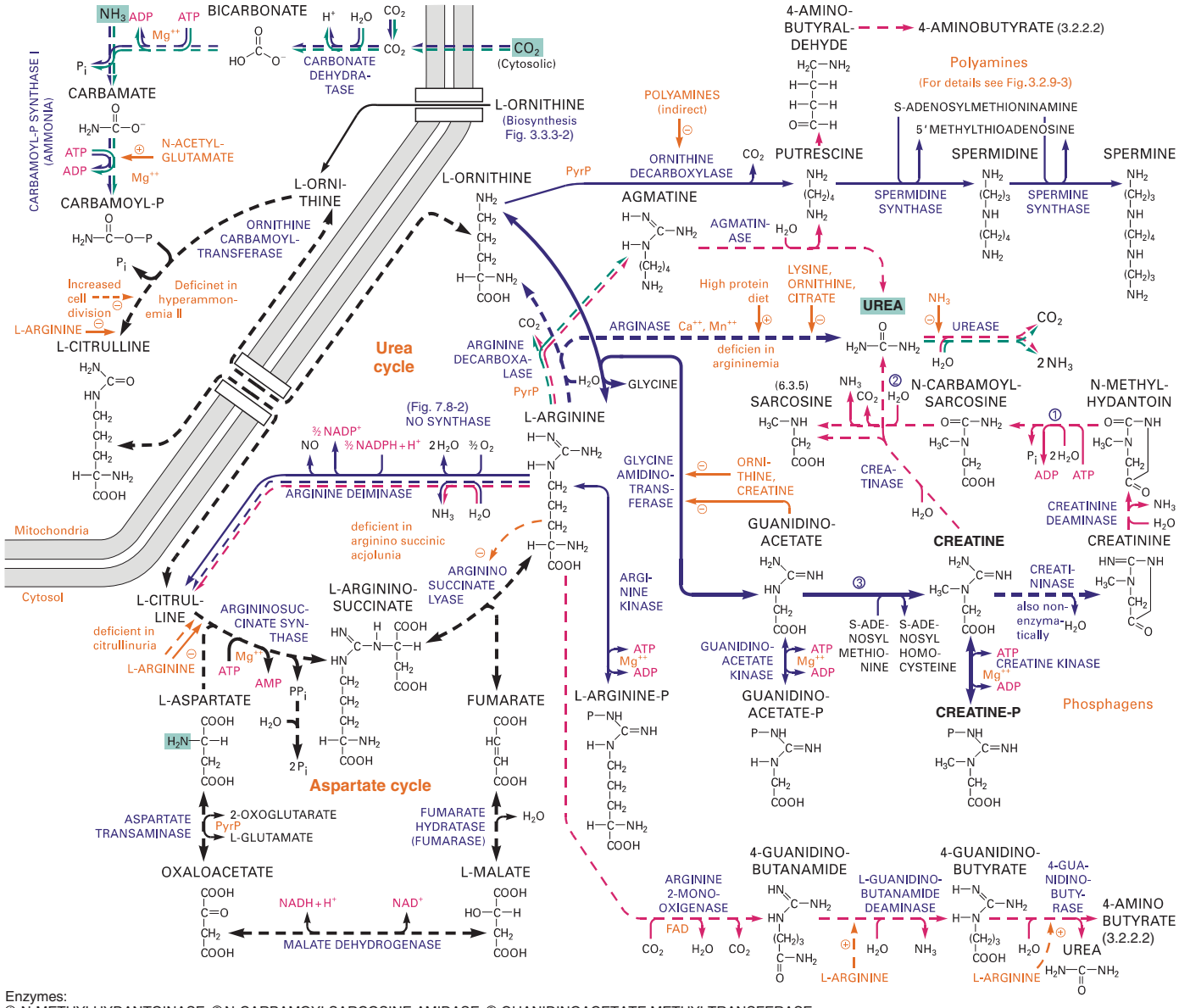


Figure 3.2.9-1. Urea and Aspartate Cycles, Polyamine and Creatine Synthesis

methylated to creatine. These reactions take place in liver, kidney and brain. Via the bloodstream, a considerable amount of creatine is transported to the skeletal muscles. Phosphorylation of creatine results in the high energy compound (3.6.1.3) phosphocreatine. The formation of phosphocreatine is endergonic:

$$\text{creatine} + \text{ATP} \rightleftharpoons \text{phosphocreatine} + \text{ADP} \qquad \Delta G'_0 = +12.6 \text{ kJ/mol}$$

$$K = \text{Conc}_{\text{phosphocreatine}} * \text{Conc}_{\text{ADP}} / \text{Conc}_{\text{creatine}} * \text{Conc}_{\text{ATP}} = 6.18 * 10^{-3}$$

Due to this equilibrium, phosphorylation of creatine can only take place when there is ample supply of ATP. On the other hand, quick reconstitution of ATP is possible when high energy demand causes the dephosphorylation of ATP to ADP. Correspondingly, creatine phosphorylation takes place inside the mitochondria, phosphate transfer to ADP proceeds in the cytosol. This way, phosphocreatine acts as a 'buffer' and enlarges the amount of quickly available energy (especially in skeletal muscle, but also in heart and brain, less in other tissues). The same function in invertebrates is exerted by phosphoarginine.

The lability of phosphocreatine is caused by the competition of the electron withdrawing imino and phosphoryl groups, which decrease the resonance stabilization of the C-NH-P bridge. This effect surpasses the analogous one in pyrophosphate bonds (e.g., in ATP), as can be seen by the $\Delta G'_0$ for hydroysis

$$\text{phosphocreatine} \rightarrow \text{creatine} + P_i \qquad \Delta G'_0 = -43.1 \text{ kJ/mol}$$

$$\text{ATP} \rightarrow \text{ADP} + P_i \qquad \Delta G'_0 = -30.5 \text{ kJ/mol}$$

For excretion in urine, creatine is converted to creatinine.

Literature:

Bessman, S.P., Carpenter, C.L. Ann. Rev. of Biochem. 1985;54:831–862.

3.2.9.3 Polyamines (Fig. 3.2.9-2)

Decarboxylation of ornithine results in the diamine putrescine. In bacteria, this compound can also be formed from arginine by a two-step reaction. Inhibition of ornithine decarboxylase takes place by tight, noncovalent binding of specific proteins (antizymes) and indirectly by polyamines. In mammals (not in yeast or bacteria), ornithine decarboxylase has a high turnover rate ($t_{1/2}$ = 10 … 30 min). Its expression is enhanced and the enzyme degradation is delayed by growth factors (NGF, EGF, PDGF), insulin, corticosteroids and testosterone.

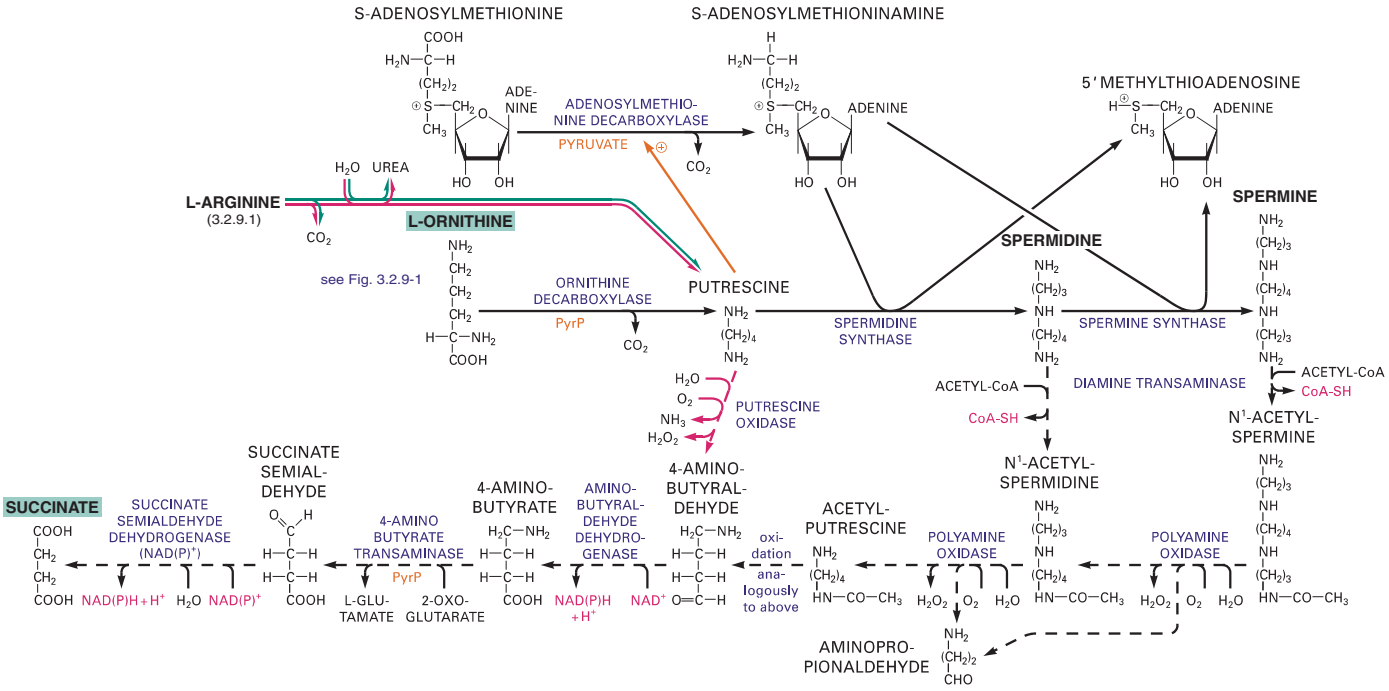


Figure 3.2.9-2. Polyamine Metabolism

The addition of one and two aminopropyl residues leads to spermidine and spermine, respectively. These polyamines (except spermine) occur ubiquitously; especially in quickly dividing cells. They are essential growth factors. Their highest concentration is found at the end of the G_1 phase. Possibly their biological function is the stabilization of DNA by association of their amino groups with the phosphate residues of DNA. They also have an enhancing effect on RNA synthesis and on tRNA and ribosome stability.

S-adenosylmethioninamine is the donor of the aminopropyl groups. This compound is generated by decarboxylation of S-adenosylmethionine; the decarboxylase contains the unusual cofactor pyruvate instead of pyridoxal phosphate. While in the transfer reactions from S-adenosylmethionine the methyl group is transferred (3.2.5.4), the spermidine and spermine synthase reactions cleave the other sulfonium-carbon bond and transfer the aminopropyl group (Fig. 3.2.9-2), yielding 5′-methylthioadenosine instead of S-adenosylhomocysteine.

Degradation of polyamines: In animals and yeast, spermine and spermidine are acetylated to their N^1 derivatives and then oxidized at the secondary amino group by polyamine oxidase, releasing aminopropionaldehyde and finally acetylputrescine. This compound is converted into 4-aminobutyrate and further on via succinate semialdehyde to succinyl-CoA, ending up in the citrate cycle. In some bacteria, instead of aminopropionaldehyde, 1,3-diaminopropane is released.

Literature:

Canellakis, E.S. *et al*. Biosci. Rep. 1985;5:189–204.

Large, P.J. FEMS Microbiol. Rev. 1992;8:249–262.

Tabor, C.W., Tabor, H. Ann. Rev. of Biochemistry 1984;53:749–790.

Verdine, G.L. Cell 1994;76:198–206.

3.3 Tetrapyrroles

Martina Jahn and Dieter Jahn

Tetrapyrroles consist of four pyrrole derivatives usually connected by methine (–CH=) bridges (For corrinoids see below.) Circular closed and open chain tetrapyrroles are known (Fig. 3.3-1). A nomenclature system has been published (see Literature, p. 86).

Cyclic tetrapyrroles are characterized by the reduction state of the ring system, the ring substituents and the metal chelated in the center of the ring. In general, the pyrrole moieties are substituted by methyl, vinyl, acetate or propionate residues. The true porphyrins, such as hemes, are compounds with a system of conjugated double bonds. They are essential cofactors of enzyme complexes. Iron-containing heme as prosthetic group of hemoglobin is responsible for the transport of oxygen and carbon dioxide through the cardiovascular system (for details, see 6.3). Cytochromes are involved in electron transport during respiration and photosynthesis (3.11.4, 3.12.1). They function also as sensor molecules for oxygen, carbon dioxide and nitrogen oxides. Finally, they are found in many enzymes including catalases (3.2.5.8), peroxidases (3.2.5.8) and cytochrome P_{450}.

Magnesium binding chlorophylls and bacteriochlorophylls are the green light collecting pigments of the photosynthetic apparatus in bacteria and plants (3.12.1). Pheophytins are the corresponding compounds without a central magnesium atom.

Tetrapyrroles with variations in the basic porphyrin structure include compounds with more reduced ring systems such as the chlorins and bacteriochlorins of the chlorophyll and bacteriochlorophyll classes, respectively. The isobacteriochlorin heme d_1 is the cofactor of the dissimilatory cytochrome cd_1 nitrite reductase, while the greenish siroheme is found in enzymes for assimilatory reduction of nitrite and sulfite (3.10.6). The yellow nickel tetrapyrrole coenzyme F_{430} is the cofactor of methyl-coenzyme M reductase involved in methanogenesis by archaea (3.10.6.2).

Another variation of the basic porphyrin scheme is the contracted ring system of the corrinoids. The pink cobalt-containing cobalamin (vitamin B_{12}, 3.7.5) and its derivatives are the most complex tetrapyrroles. They serve as cofactors for enzymes that catalyze methyl transfer reactions or radical-based rearrangements.

The linear tetrapyrroles are all derived from cyclic precursors. They do not chelate metal ions. Phycobilines serve as chromophoric photoreceptors in the light harvesting systems of bacteria and algae (3.12.1). Moreover, light-responsive regulatory proteins including the phytochromes utilize linear phytochromobilines.

3.3.1 Pathways for the Biosynthesis of Tetrapyrroles (Fig. 3.3-2)

The obvious structural similarity of all tetrapyrroles implies a common biosynthetic pattern. The steps from aminolevulinic acid to the first cyclic intermediate uroporphyrinogen III are common for all tetrapyrroles. Thereafter, the two major biosynthetic routes split.

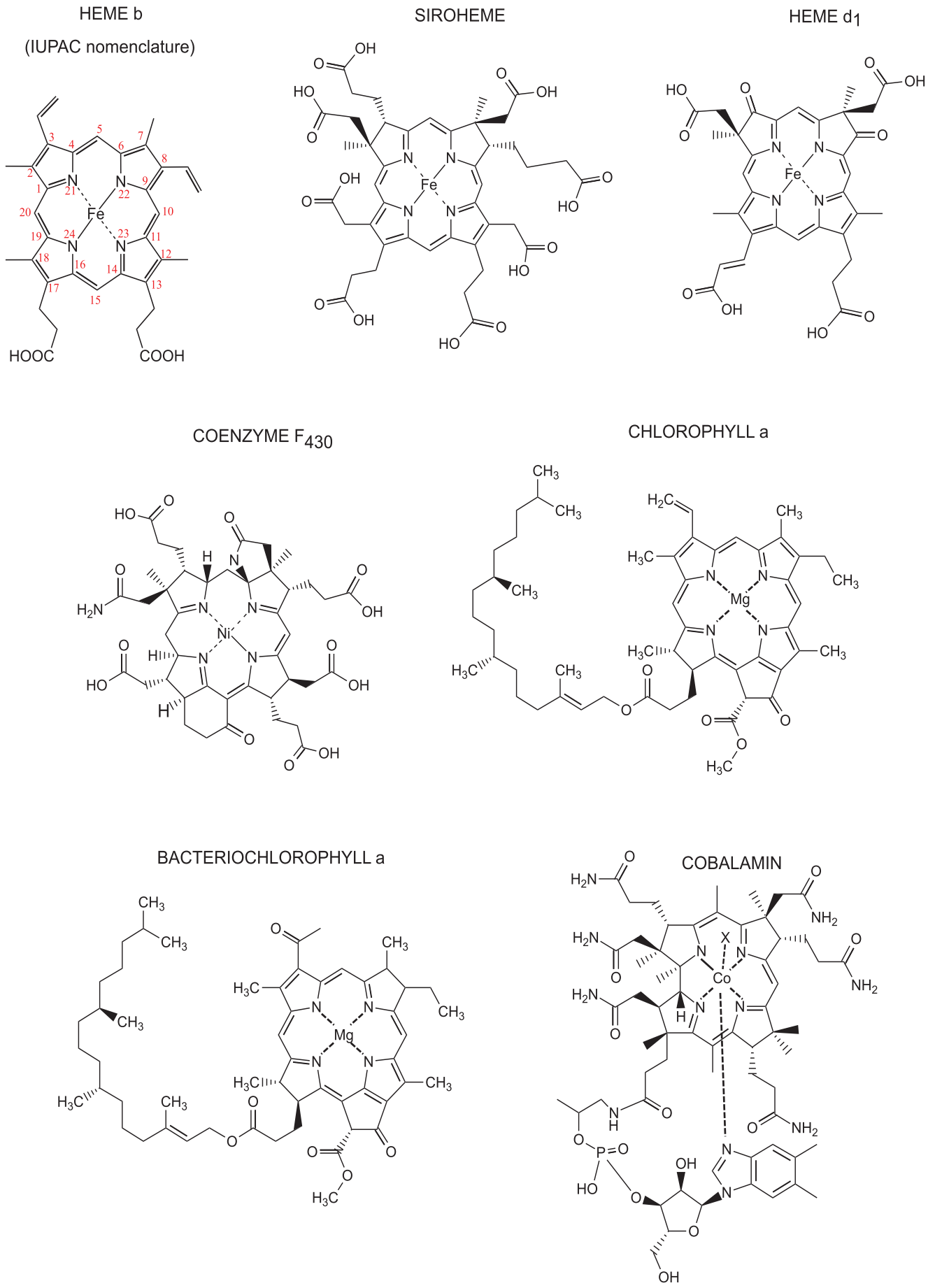


Figure 3.3-1. Important Representatives of Naturally Occurring Tetrapyrroles

The pathway via coproporphyrinogen III leads to hemes, chlorophylls and bacteriochlorophylls. Conversion into precorrin-2 yields furthermore siroheme, heme d_1, coenzyme F_{430} and cobalamin. In a few bacteria and archaea precorrin-2 is also the precursor for heme formation via a yet unknown pathway. Prokarya can form all tetrapyrroles mentioned above while eukarya are limited to heme, chlorophyll, siroheme and biline synthesis.

3.3.1.1 Biosynthesis of 5-Aminolevulinic Acid (ALA, Fig. 3.3-3)

ALA is the general precursor of all known tetrapyrroles and carries all carbon and nitrogen atoms required for the synthesis of the tetrapyrrole macrocycle. In nature two completely independent routes for the formation of ALA exist.

Animals, fungi and the α-group of the proteobacteria utilize the one-step condensation of succinyl-CoA and glycine with the release of carbon dioxide, known as 'Shemin-pathway'. The reaction is catalyzed by the pyridoxal-5-phosphate dependent 5-aminolevulinate synthase. Catalysis starts with binding of glycine to pyridoxal-5-phosphate and the formation of an external aldimine. Removal of the pro-*R* proton of glycine is followed by the attachment of succinyl-CoA to the glycine with the generation of an α-amino-β-keto adipate intermediate. Subsequently, an active site histidine residue as general base promotes protonation and decarboxylation with the release of CO_2. Finally, ALA is formed as a bound external aldimine via a quinonoid intermediate prior to its release and the regeneration of the internal aldimine between an active site lysine and pyridoxal-5-phosphate dependent. The expression of the ALAS genes in vertebrates is regulated by a variety of transcriptional and posttranscriptional mechanisms. The gene for the housekeeping isoform in humans (*ALAS1*) is expressed in all human cell types and negatively regulated by the presence of heme. The *ALAS2* gene is exclusively expressed in erythroid cells. Its transcription does not respond to heme, however, the translation of the corresponding mRNA is regulated by iron via an iron response element (IRE). Furthermore, protein stablility, transport and enzyme activity are also subject to heme regulation. Bacterial ALAS genes are regulated by heme and oxygen tension.

In plants, archaea and most bacteria ALA is formed from the C5-skeleton of glutamate via the so-called 'C5-pathway'. Glutamate is activated via an ester bond to $tRNA^{Glu}$, catalyzed by glutamyl-tRNA synthetase (ligase). This enzyme is also utilized in protein biosynthesis. Glutamyl-tRNA reductase (GluTR) catalyzes the NADPH-dependent reduction of tRNA-bound glutamate to glutamate-1-semialdehyde (GSA). The nucleophilic attack of a GluTR cysteine residue on the α-carbonyl group of tRNA-bound glutamate results in an enzyme-substrate thioester intermediate. By hydride transfer from NADPH, GSA is produced. GluTR forms a tight complex with glutamate-1-semialdehyde-2,1-aminomutase (GSAM) which converts the labile GSA in a pyridoxal-5-phosphate dependent reaction into ALA by transferring the amino group of GSA from C2 to C1.

3.3.1.2 Conversion of 5-Aminolevulinic Acid into Uroporphyrinogen III (Figs. 3.3-3 and 3.3-4)

The first macrocyclic intermediate of tetrapyrrole biosynthesis, uroporphyrinogen III, is formed from eight molecules of ALA in three consecutive enzymatic steps.

Porphobilinogen synthase catalyzes the asymmetric condensation of two ALA molecules to the pyrrole derivative porphobilinogen (PBG). The enzyme utilizes various metal ions (zinc, magnesium, sodium) for activity. Two substrate molecules of ALA are bound via Schiff bases to two lysine residues of the enzyme, followed by inter-substrate C–C bond formation between C3 of substrate ALA 1 and C4 of substrate ALA 2, before the C–N bond between the two substrate molecules is formed. The final deprotonation of substrate ALA 2 yields the reaction product PBG.

Porphobilinogen deaminase (PBGD) catalyzes the condensation of four PBG molecules to the instable linear tetrapyrrole pre-uroporphyrinogen (1-hydroxmethylbilane). The enzyme utilizes a unique dipyrromethane cofactor. This cofactor is composed of two PBG substrate molecules and gets assembled on the enzyme posttranslationally. During catalysis the cofactor serves as a primer for the sequential polymerization of the following four PGB substrate molecules.

Uroporphyrinogen III synthase (UROS) catalyzes the cyclization of the linear tetrapyrrole pre-uroporphyrinogen under inversion of ring D to form the asymmetric cyclic tetrapyrrole uroporphyrinogen III. In absence of this enzyme, pre-uroporphyrinogen cyclizes non-enzymatically without inversion of ring D to the symmetric toxic dead-end product uroporphyrinogen I. This indicates that PBGD and UROS have to act in a highly coordinated fashion. Consequently, UROS was also named 'co-synthase'. The proposed mechanism starts with rearrangement of ring A resulting in a first azafulvene intermediate which in turn attacks the substituted α-position of ring D to yield a spirocyclic pyrrolenine intermediate. This one collapses via bond breakage between rings C and D leading to another azafulvene intermediate on ring C. Now the enzyme orientates ring D such that its free α-position comes in close proximity to the ring C azafulvene. After attack of the azafulvene, deprotonation and bond rearrangement, the product uroporphyringen III is formed.

3.3.1.3 Conversion of Uroporphyrinogen III into Protoheme (Fig. 3.3-3)

The transformation of uroporphyrinogen III into protoheme needs four consecutive enzymatic steps. During these reactions the side chains of the tetrapyrroles are modified, the oxidation state of the ring system is changed and iron gets inserted.

At first uroporphyrinogen III decarboxylase converts four acetate side chains into methyl groups to yield coproporphyrinogen III. The lack of cofactors is unique among decarboxylases. Structural studies suggest a reaction mechanism in which decarboxylation at ring D takes place in the active site of one subunit while the following decarboxylations at rings A, B, and C occur in the active side of the other subunit. This mechanism avoids the otherwise necessary 180° flipping of the substrate within a single active site.

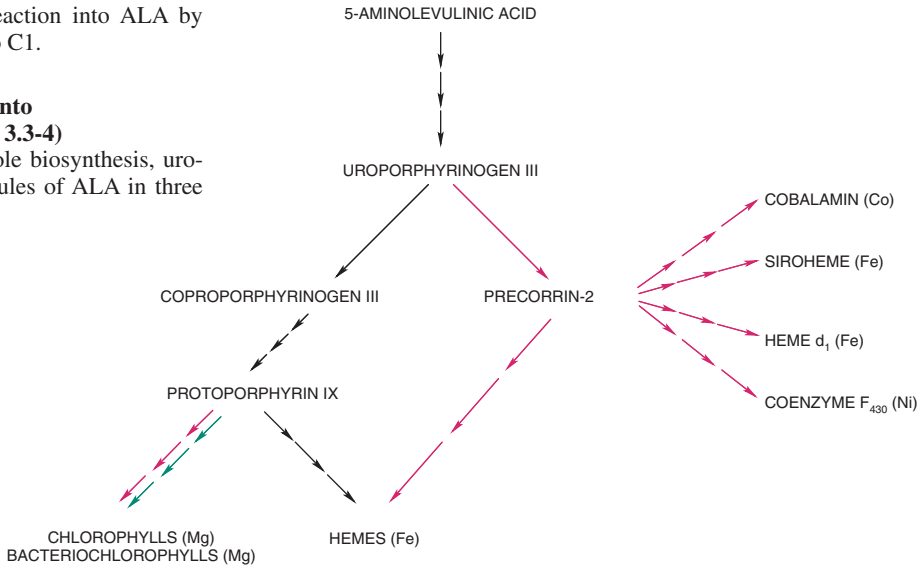


Figure 3.3-2. Biosynthetic Pathways for Various Tetrapyrroles

In the next step, coproporphyrinogen III oxidase catalyzes the oxidative decarboxylation of the propionate side chains on rings A and B to the corresponding vinyl groups, resulting in protoporpyhrinogen IX. In nature two structurally not related enzymes are used for this catalytic step. Eukarya and a few bacteria utilize oxygen-dependent HemF while most bacteria employ the oxygen-independent radical enzyme HemN (Fig. 3.3-5). This enzyme receives electrons from an electron donor, such as reduced flavodoxin or ferredoxin and transfers them via a [4Fe-4S] cluster to S-adenosyl-L-methionine (SAM). After homolytic cleavage of SAM, a highly reactive 5′-deoxyadenosyl radical is formed, which abstracts the pro-*S*-hydrogen atom from the β–C atom of the substrate propionate side chain resulting in a substrate radical. The decarboxylation is completed by the elimination of carbon dioxide and an electron.

The following step of heme biosynthesis, the six electron oxidation of colorless protoporphyrinogen IX to red colored protoporphyrin IX, is catalyzed by protoporphyrinogen IX oxidase. Different structurally-unrelated enzymes catalyze the reaction. Again eukarya and a few bacteria use an oxygen-dependent FAD-dependent enzyme.

Protoporphyrin IX is the branching point for the pathways leading to heme and to chlorophyll (3.3.4).

3.3.1.4 Porphyrias Caused by Enzyme Defects in Humans

Partial deficiency of the enzymes of heme biosynthesis gives rise to pathway blockade and increased levels of metabolites, which in turn results in characteristic clinical features. These deficiencies are called porphyrias. They are usually genetically determined and inherited over generations. Most porphyrias are associated with skin lesions due to photosensitisation of light-absorbing porphyrins and with neuronal manifestations. Mutation of the gene for erythroid specific ALA synthase 2 causes X-linked sideroblastic anaemia. Defects in porphobilinogen synthase (ALA dehydratase) cause ALA dehydratase porphyria, characterized by a highly variable symptomatology, ranging from failure to thrive in infants to the development of a polyneuropathy in adults. Common congenital diseases are the acute intermittent porphyria caused by porphobilinogen deaminase failure and the nonacute erythropoietic porphyria associated with a deficiency of uroporphyrinogen III synthase, which prevents the formation of

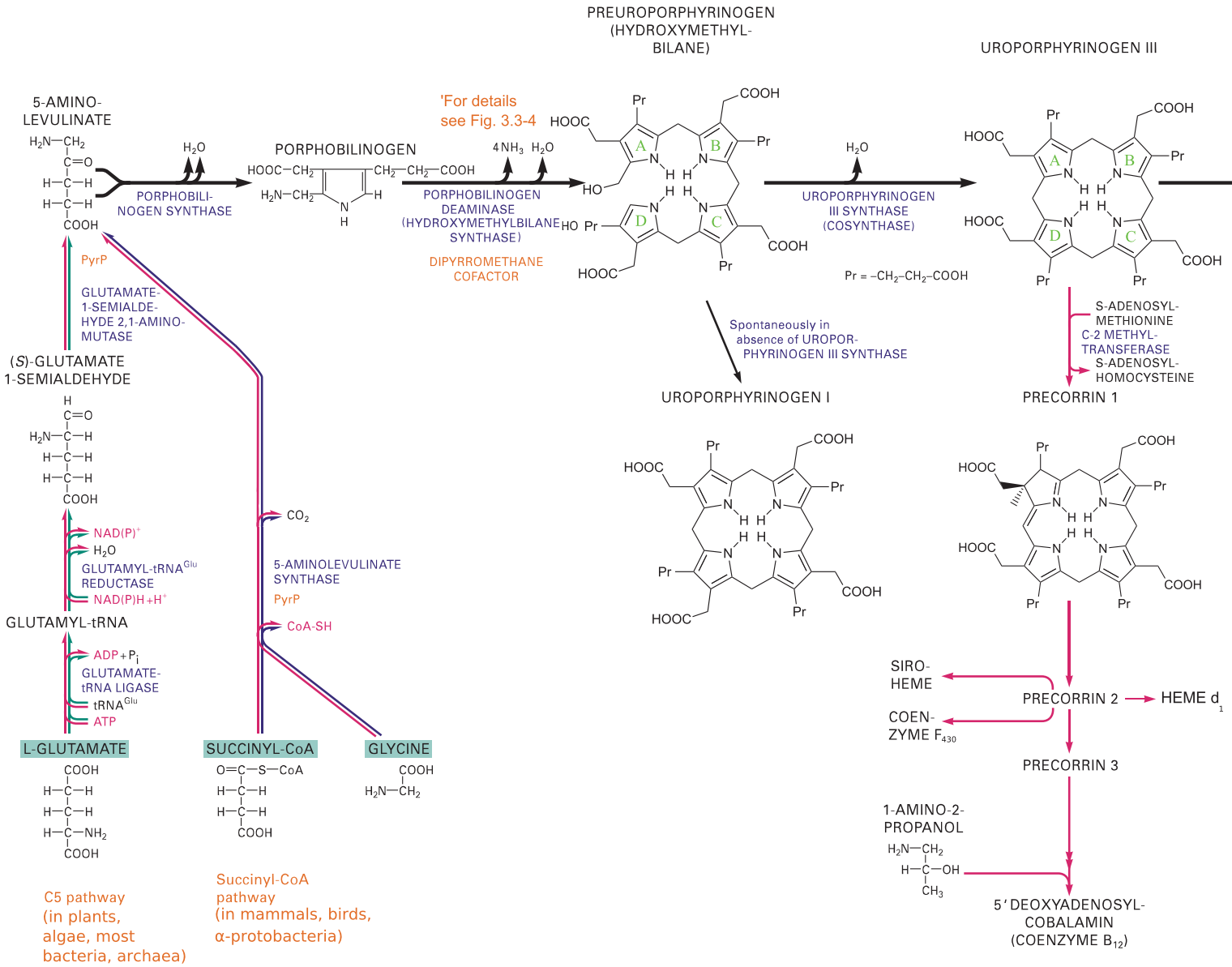


Figure 3.3-3. Pathway for the Biosynthesis of Heme

uroporphyrinogen III. Consequently, nonenzymatic conversion of pre-uroporphyrinogen to uroporphyrinogen I and its consecutive decarboxylation to coproporphyrinogen I take place, which causes red colored urine after oxidation to coproporphyrin I. Nonacute porphyria cutanea tarda affects uroporphyrinogen decarboxylase. Acute hereditary coproporphyria and variegate porphyria are caused by enzyme defects in coproporpyhrinogen III and protoporphyrinogen IX oxidases, respectively. They are characterized by abdominal pain, tachycardia, hypertension, peripheral neuropathy etc. Acute porphyrias are usually treated by intravenous application of human hemin or in some cases of 5-aminolevulinic acid.

Literature:

Heinemann, I.U. *et al.* Arch. Biochem. Biophys. 2008;474:238–251.

IUPAC Nomenclature of Tetrapyrroles (www version). http://www.chem.qmul.ac.uk/iupac/tetrapyrrole.

Jahn, D., Heinz, D.W. in Warren, M.J., Smith, A.G. (Eds.) *Tetrapyrroles: Birth, Life and Death.* Landes Bioscience, Austin, 2009. (This book is an extensive general compilation of tetrapyrrole metabolism).

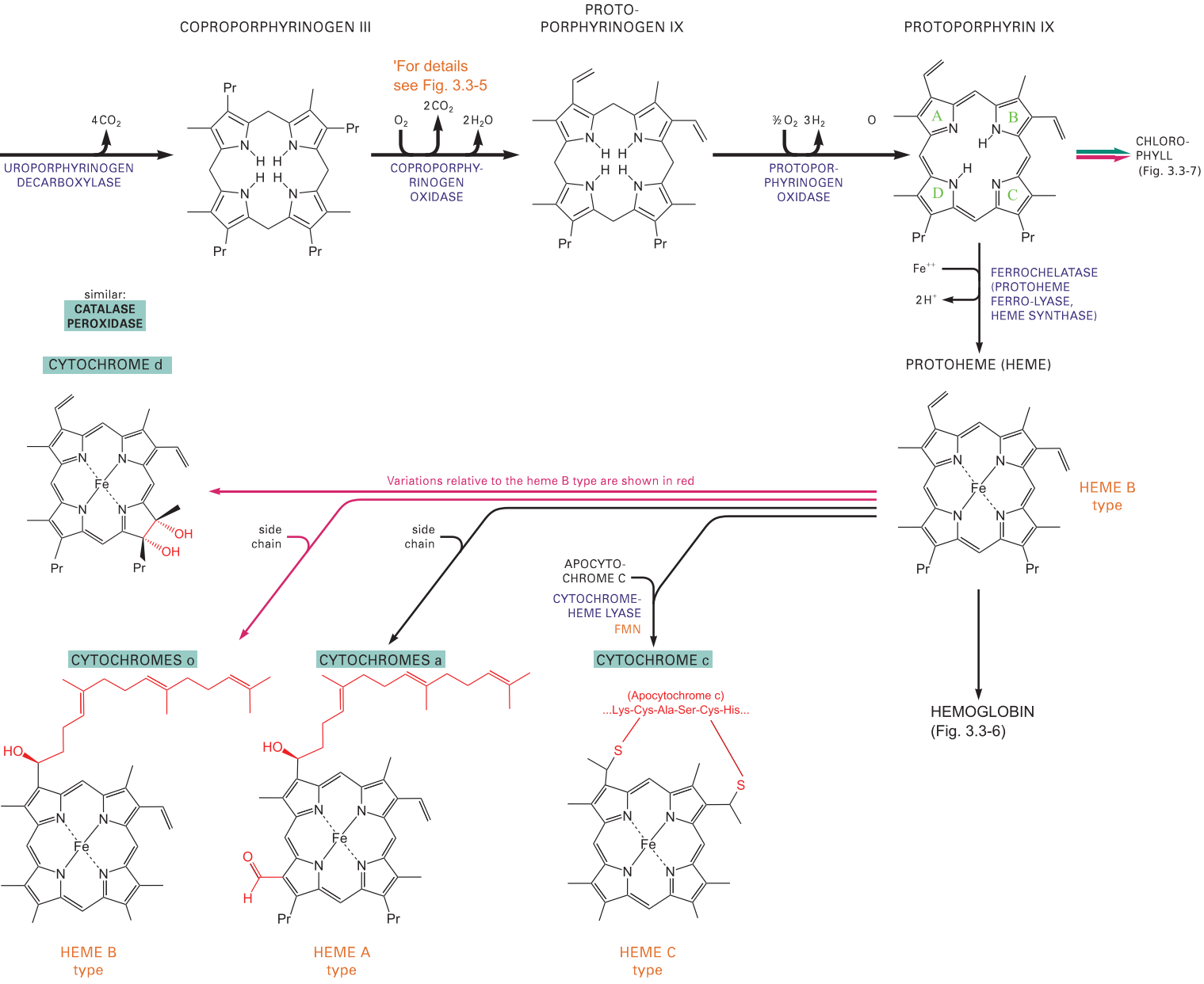
Layer, G., *et al. Biosynthesis of heme and vitamin B12*. Elsevier, 2009.

Layer, G. *et al*. Protein Sci. 2010;19:1137–1161.

3.3.2 Heme and Cytochrome Biosynthesis

3.3.2.1 Introduction of Iron (Fig. 3.3-3)

The last step of heme biosynthesis, the insertion of ferrous iron into protoporphyrinogen IX, is catalyzed by ferrochelatase. Eukaryotic ferrochelatases are membrane-associated homodimers. Investigated bacterial ferrochelatases are soluble, monomeric enzymes. Eukaryotic ferrochelatases, with the exception of plant enzymes, contain one [2Fe-2S] cluster of yet unknown function per subunit. Bacterial ferrochelatases are found with and without the cluster. No further cofactors are observed. During catalysis the enzyme causes a distortion of the porphyrin ring to a nonplanar saddle conformation, which facilitates iron chelation. Although ferrochelatases from different species share little amino acid



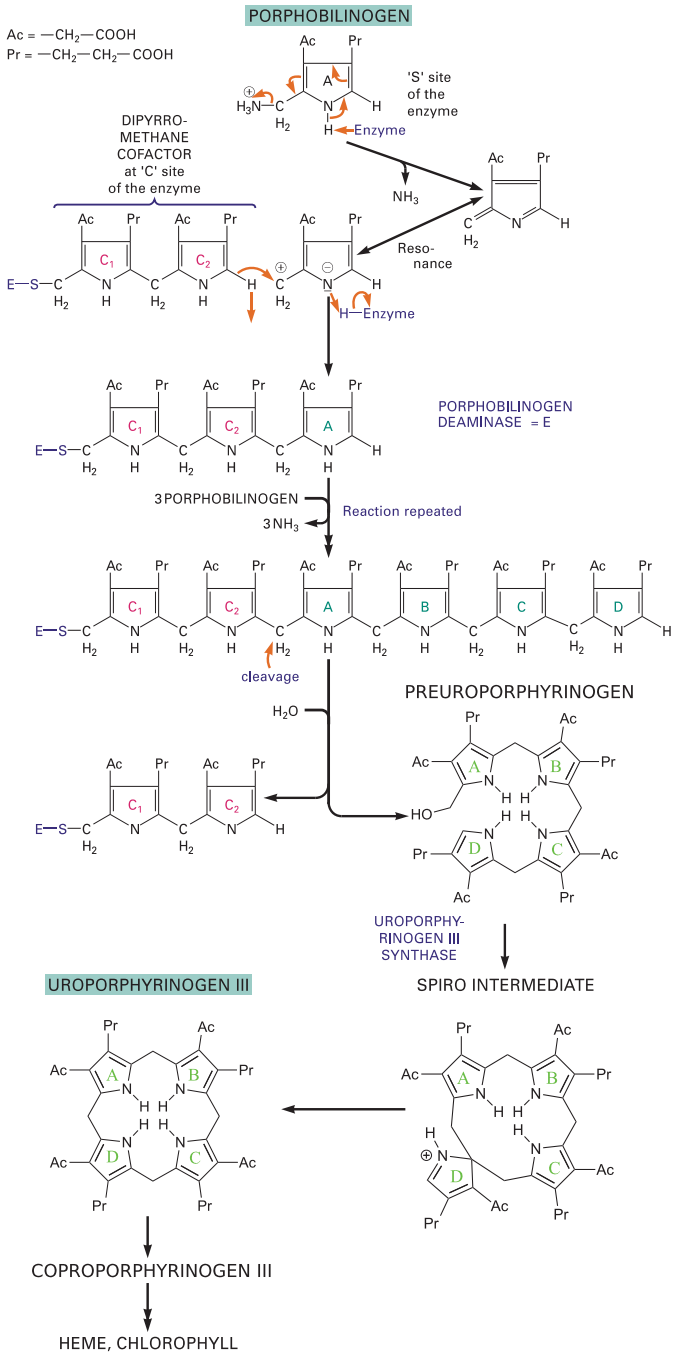


Figure 3.3-4. Mechanisms for the Enzymes Converting Porphobilinogen into Uroporphyrinogen III

sequence identity, the solved crystal structures of human, yeast and bacterial enzymes revealed a high degree of structural conservation.

3.3.2.2 Cytochromes Carrying Different Types of Heme (Fig. 3.3-3)

Cytochromes are hemoproteins taking part in redox reactions by reversible transitions between the Fe^{2+} and Fe^{3+} states. Cytochromes are members of electron transfer chains (respiratory chain, 3.11.4, photosynthesis, 3.12.1) and participate in enzymatic reactions like cytochrome P450 (3.5.4.1, 7.4.8). They occur in many variants in almost all organisms.

Cytochromes are usually named after the bound type of heme and the peak wavelength of the light absorption band α. Various hemes are grouped by the substitutions at the porphyrin ring (Fig. 3.3-3, Table 3.3-1). The ring substitutions of B-type cytochromes are identical with protoporphyrin IX and with heme in hemoglobin. Heme A is synthesized from heme B with heme O as a stable intermediate. Heme O is derived from heme B by substitution of the vinyl group at carbon 2 with a hydroxyethyl farnesyl side chain. Heme A further differs from heme O by carrying a formyl group instead of a methyl group on carbon 8 of the porphyrin ring. The two biosynthetic steps are catalyzed by heme O and heme A synthases, respectively. The heme A synthase is dependent on molecular oxygen. It is thought that the methyl side chain on carbon 8 of heme O is subjected to two consecutive hydroxylations catalyzed by heme A synthase. Heme D, the cofactor of various catalases, is a *cis*-hydroxychlorin γ-spirolactame. It carries an extra ring on porphyrin ring III. It should be formed autocatalytically from heme B during binding to the catalase.

Some cytochromes are transmembrane proteins with several membrane-spanning helices. In some cases, 2 different types of hemes are bound to a single protein chain (e.g., cytochrome bd, cytochrome bo, Fig. 3.11-1). Other cytochromes are peripheral proteins, which act as 'electron shuttles'. A number of cytochromes are shown in Table 3.3-2.

Table 3.3-1. Types of Cytochromes

Heme Type	Substitution at Position 3	8	18
B	vinyl	vinyl	methyl
O	2-farnesyl-1-hydroxyethyl	vinyl	methyl
A	2-farnesyl-1-hydroxyethyl	vinyl	formyl
C	1-mercaptoethyl	1-mercaptoethyl	methyl

Table 3.3-2. Examples of Cytochromes

Heme Type	Cytochromes (Occurrence)	Section
A	a, a_3 (respiratory chain in mitochondria) a_1 (terminal oxidase in bacteria)	3.11.4.4
B	$B_L = b_{566}$, $b_H = b_{560(562)}$ (respiratory chain in mitochondria) b_{556}, b_{558}, b_{595} (bd complex), b_{562} (bo complex) (respiratory chain in bacteria) b_L, b_H (photosynthesis in plants + cyanobacteria) b_L, b_H (photosynthesis in purple bacteria) b_2 (yeast lactate dehydrogenase) b_5 (microsomes in animals) P450 (monooxygenases)	3.11.4.3 3.11.4.5 3.12.1 3.12.1 3.5.4 3.5.4.1
C	c, c_1 (respiratory chain in mitochondria) f (photosynthesis in plants + cyanobacteria) c_1, c_2 (photosynthesis in purple bacteria) c_5 (bacterial nitrate assimilation)	3.11.4.3 3.12.1 3.12.1 3.10.6
O	o (bacteria)	3.11.4.5
D	d (bacteria) Similar: bacterial catalase	3.11.4.5

Literature:

Brzezinski, P., Gennis, R.B. *J Bioenerg Biomembr* 2008;40:521–531.

Gennis R.B. *Frontiers in Bioscience* 2004;9:581–591.

Hershey, J.W.B. Ann. Rev. of Biochem. 1991;60:717–755.

Jordan, P.M. *Biosynthesis of Tetrapyrroles. New Comprehensive Biochemistry* Vol. 19. Elsevier, 1991.

Junemann S. *Biochim Biophys Acta* 1997;1321:107–127.

Unden G., Duennwald P. *The aerobic and anaerobic respiratory chain of Escherichia coli and Salmonella enterica: Enzymes and Energetics*. In Boeck, A., Curtiss III, R. *et al.* (eds.), ASM Press, Washington, D.C.,2008.

E. coli & Salmonella database. htttp://www.ecosal.org.

3.3.3 Linear Tetrapyrroles (Fig. 3.3-6)

Bilins are open-chain tetrapyrroles. They all derived from the region-specific cleavage of a cyclic heme (iron-protoporphrin IX) molecule performed by a heme oxygenase. They are widespread in nature with multiple functions.

3.3.3.1 Heme Oxidation in Humans and Bile Pigment Formation (Fig. 3.3-6)

Erythrocytes hosting hemoglobin as oxygen carrier have a limited lifetime of about 120 days in humans. They are degraded in the spleen and bone marrow. The amino acids originating from globin are usually reutilized, while the degradation products of heme (with exception of the released iron) are excreted.

The heme ring is cleaved by membrane-bound heme oxygenase (decyclizing) between the pyrroles A and B and the carbon of the methine bridge is liberated as CO. The enzyme requires molecular oxygen and a total of seven electrons provided by NADPH via a cytochrome P450 system (compare 3.5.4.1). It produces biliverdin IXα, CO and free iron. Biliverdin is excreted by reptiles and birds, while in mammals it is converted to bilirubin IXα by NAD(P)H-dependent biliverdin IXα reductase. This lipophilic compound forms a complex with albumin for the transport in blood. After entering the liver, the propionate side groups are esterified with glucuronic acid in order to increase the solubility. Finally, bilirubin mono- and diglucuronides are excreted via bile into the intestine. There, anaerobic intestinal microorganisms remove the glucuronide groups reducing bilirubin to urobilinogen and stercobilinogen. Some of the former is reabsorbed from the intestine, converted in the kidney to urobilin and excreted in the urine. The major portion of the heme degradation products appear as stercobilin in stool and cause its dark color. Bilirubin shows an effective antioxidant activity and removes peroxide radicals (3.2.5.8) in blood and in membranes.

Medical aspects: If a high amount of bilirubin is present in blood, deposition of this compound in the skin and in the eye leads to jaundice. This may be caused by elevated red cell destruction, obstruction of the bile duct or diminished glucuronidation in the liver (especially in premature babies, where the enzyme activity has not been fully developed).

3.3.3.2 Plant and Bacterial Linear Tetrapyrroles (Fig. 3.3-6)

In photosynthetic organisms bilins have two major roles. Bound to the protein moiety of the photoreceptor phytochrome they can function as sensors of light intensity, quality, direction and duration. Associated to phycobiliproteins, phycobilins also harvest light and transfer this energy to the photosynthetic reaction centres (3.12.1). Various bilins are employed as chromophores. All are derived from biliverdin IXα.

Currently, three different biosynthetic directions are described:

1. Phytochromobilin synthase converts biliverdin IXα into 3Z-phytochromobilin.
2. 3Z-phycocyanobilin: ferredoxin oxidoreductase produces 3Z-phycocyanobilin with $18^1,18^2$-dihydrobiliverdin as intermediate.
3. 15,16-dihydrobiliverdin : ferredoxin oxidoreductase forms 15,16-dihydrobiliverdin, before 3Z-phycoerythrobilin : ferredoxin oxidoreductase converts this into 3Z-phycoerythrobilin.

Phytochromes are a highly diverse family of receptors. The protein parts vary from approximately 750–1500 amino acid residues in length. All possess a conserved N-terminal photosensory domain including a bilin covalently bound to a cysteine residue of a small-molecule-ligand (GAF) domain. The C-terminal regulatory domain is highly variable. Many plants use 3Z-phytochromobilin as chromophore, while cyanobacteria utilize 3Z-phycocyanobilin for their phytochromes. Surprisingly, several non-photosynthetic bacteria, including *Pseudomonas aeruginosa* and *Deinococcus radiodurans*, also possess phytochrome like receptors with biliverdin IXa as chromophore. Their function is only poorly understood.

Several bilins also function as photosynthetic light harvesting chromophores when bound to special proteins. They are present in several photosynthetic bacteria and some algae. Phycoerythrins with bound phycoerythrobilines and sometimes phycourobilines show an absorption maximum at 565–575 nm, phycoerythrocyanin with one bound phycoviolobilin and two phycocyanobilines at 574 nm, phycocyanins with phycocyanobilines and phycourobilines at 615–640 nm and *allo*-phycocyanin with attached phycocyanobilins at 650–655 nm.

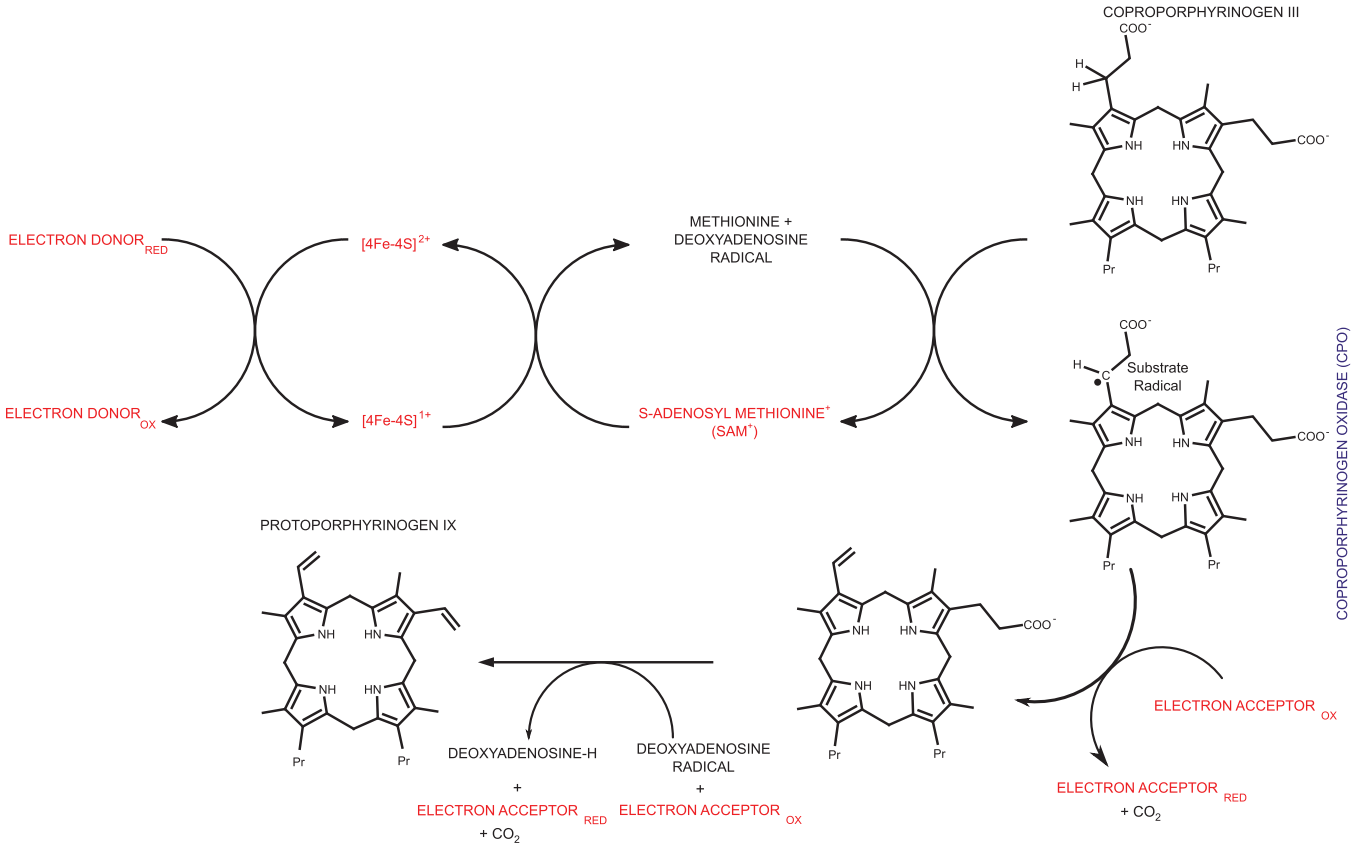


Figure 3.3-5. Mechanism of the Radical SAM Enzyme Coproporphyrinogen III Deydrogenase

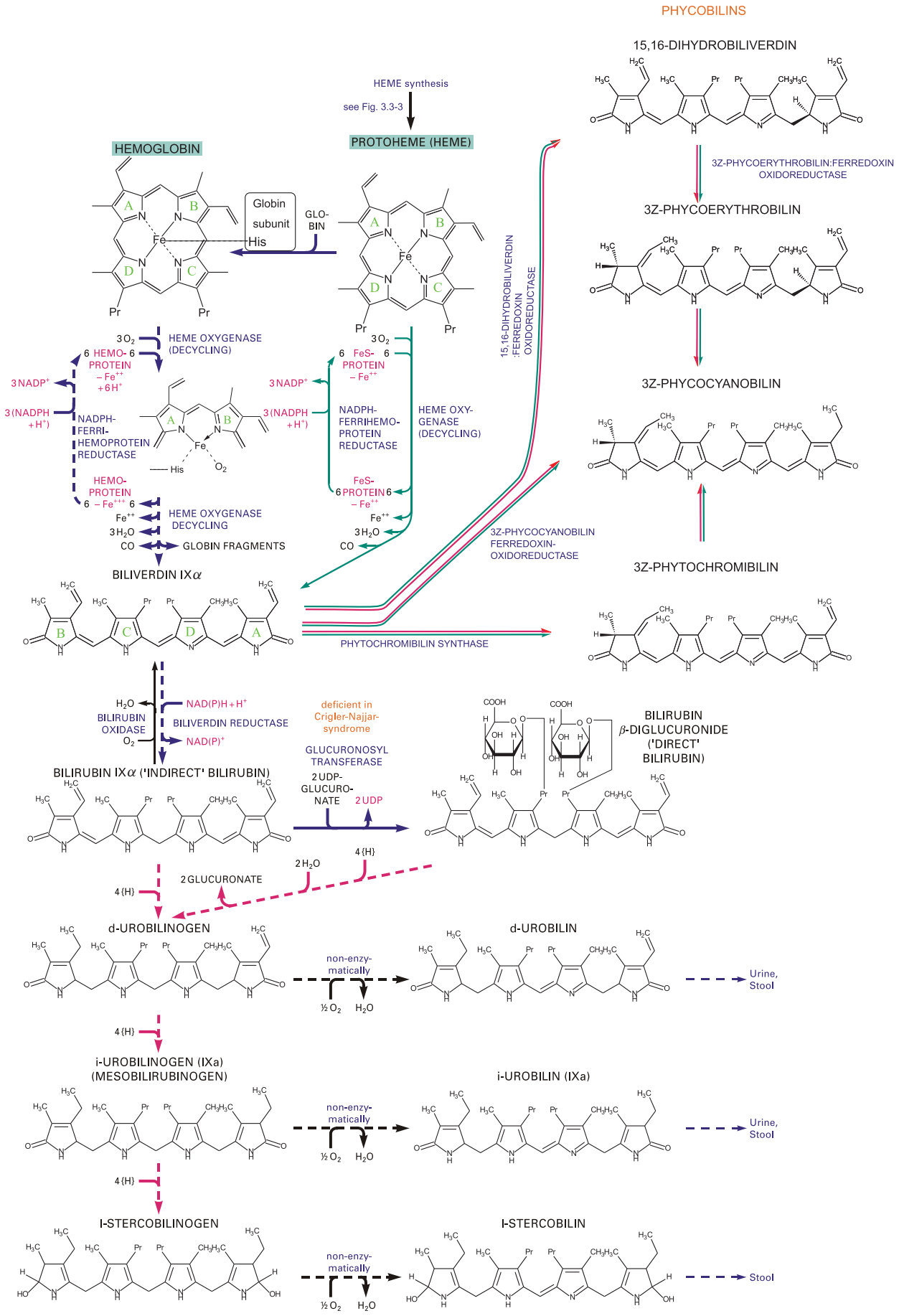


Figure 3.3-6. Bilin Formation via Hemoglobin Degradation

Literature:

Frankenberg-Dinkel, N. and Terry, M.J, in Warren, M.J., Smith, A.G. (Eds.): *Tetrapyrroles: Birth, Life and Death.* Landes Bioscience, Austin, 2009.

3.3.4 Biosynthesis of Chlorophylls (Fig. 3.3-7)

Chlorophyll and bacteriochlorophyll molecules are the most abundant pigments on Earth, with at least 10^9 tons synthesized per year. These molecules are the basis for photosynthesis, the conversion of light energy in chemical energy, which provides the driving force for life on Earth (3.12). Two major classes of these (bacterio)chlorines exist, the chlorophylls and the bacteriochlorophylls. They are characterized by their central metal ion magnesium and the fifth ring being localized on the third ring of the tetrapyrrole. Over 10 different (bacterio)chlorophylls are known, which differ in the degree of ring reduction and the nature of the ring substituents.

Up to the formation of protoporphyrin IX chlorophyll biosynthesis follows the path for heme formation (Fig. 3.3-3). Then, the first committed step of (bacterio)chlorophyll biosynthesis is the ATP-dependent insertion of a magnesium ion into protoporphyrin IX by magnesium protoporphyrin IX chelatase to form magnesium protoporphyrin IX. Next, a *S*-adenosyl-L-methionine (SAM)-dependent magnesium protoporphyrin IX methyltransferase translocates a methyl group to the propionate side chain of ring C to form magnesium protoporphyrin IX monomethylester and *S*-adenosyl-L-homocysteine.

An oxidative cyclization is required to create the fifth ring of chlorophyll. This conversion of magnesium protoporphyrin IX monomethylester into protochlorophyllide is catalyzed by two different, structurally unrelated Mg protoporphyrin IX monomethyl ester cyclases. In higher plants, an oxygen-dependent enzyme activity has been described. The direct integration of atmospheric oxygen into the 13-methyl propionate side chain to create the corresponding keto group has been shown using isotopic labelling experiments. It was postulated that a hydroxyl and a keto derivative are reasonable intermediates during the conversion to protochlorophyllide. A completely different mechanism for the oxidative cyclization reaction as compared to the one found in higher plants and aerobic bacteria was suggested for the enzyme found in *Rhodobacter sphaeroides*. Isotopic labelling experiments using $H_2{}^{18}O$ indicated, that the oxygen atom in the fifth ring of bacteriochlorophyll is derived from water and not from O_2. Reduction of protochlorophyllide *a* to chlorophyllide is a central step in the biosynthesis of chlorophylls and bacteriochlorophylls. Again, two evolutionarily unrelated enzymes are capable of catalyzing the stereospecific two-electron reduction of the C17-C18 double bond of protochlorophyllide. The first is the light-dependent protochlorophyllide oxidoreductase (POR). The substrate protochlorophyllide bound in the active site of this monomeric enzyme absorbs

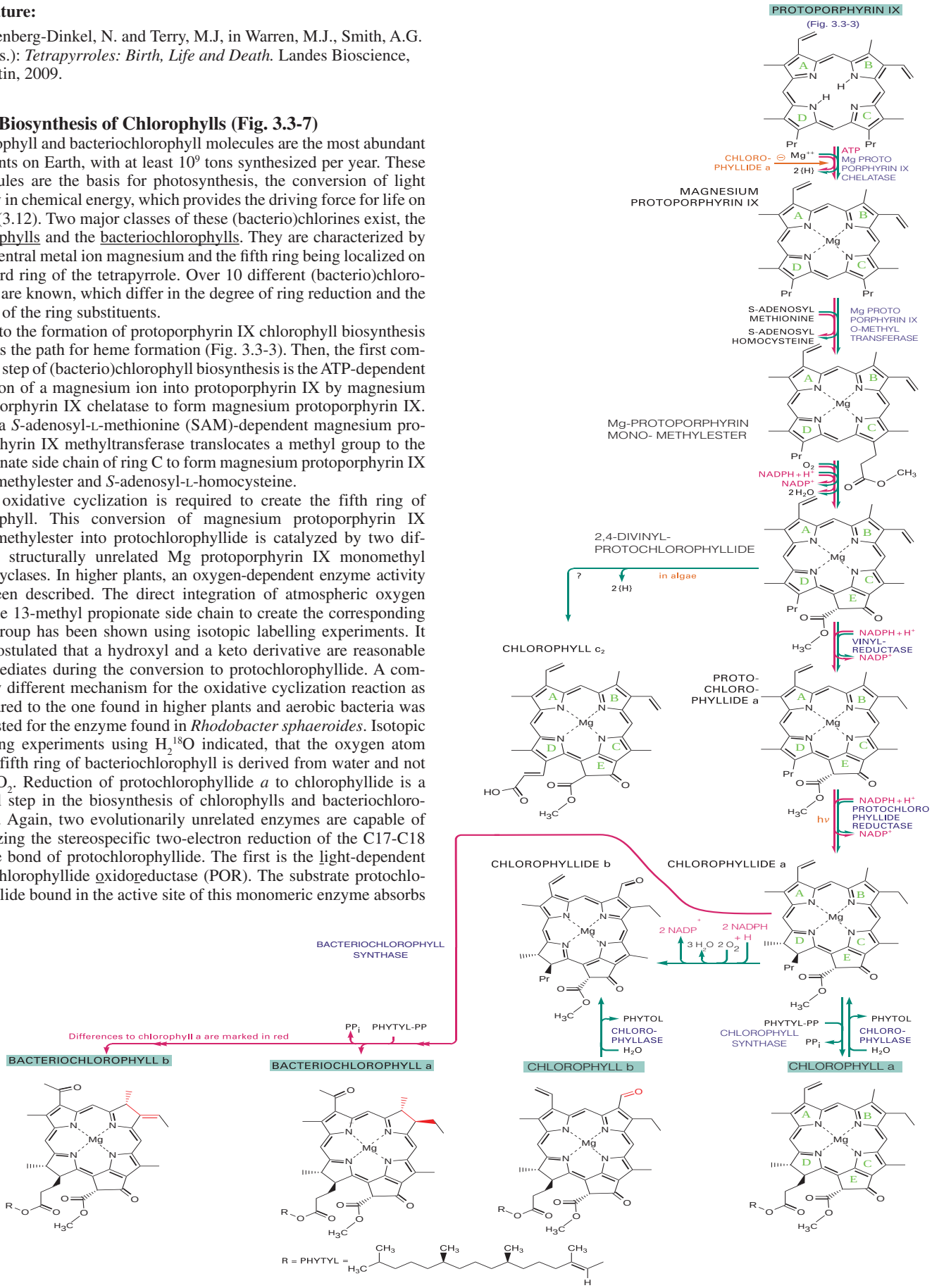
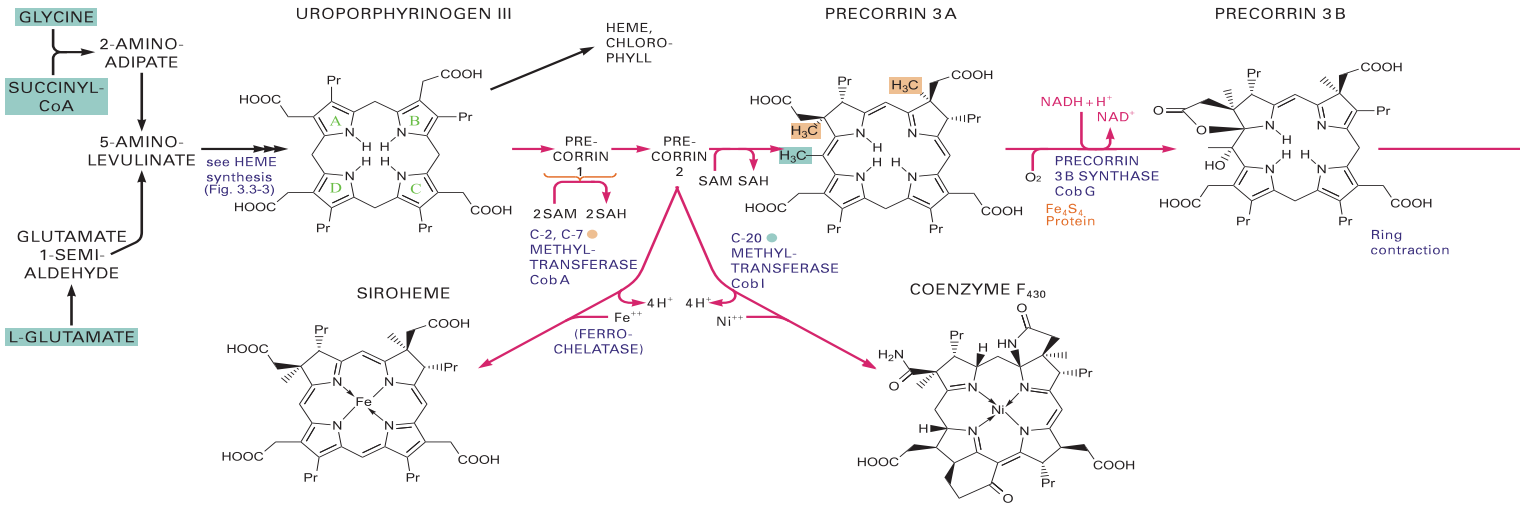


Figure 3.3-7. Biosynthesis of Chlorophylls and Bacteriochlorophylls



light energy in order to drive the NADPH dependent reduction. Light dependency of this catalysis prevents angiosperms from synthesizing chlorophyll in the dark. By contrast anoxygenic-photosynthetic bacteria make use of a different ATP-dependent protochlorophyllide reducing system, which is termed the light-independent, dark operative protochlorophyllide oxidoreductase (DPOR). A mechanism with ATP-dependent electron transfer via two intersubunit [4Fe-4S] clusters to the substrate similar to nitrogenase catalysis was demonstrated. Light dependency of POR catalysis prevents angiosperms from synthesizing chlorophyll in the dark. Gymnosperms, mosses, ferns, algae and cyanobacteria possess both POR and DPOR.

Bacteriochlorophylls differ from chlorophylls in the nature of the substituents on rings A and B. Three steps are required to form bacteriochlorophyllide a. The first two are interchangeable. One step includes the hydration of the 3-vinyl group of ring A. ^{18}O-labelling experiments showed that the carbonyl oxygen of the acetyl group at position 3 of ring A is derived from water, indicating a hydratase type mechanism rather than an oxygenase-like catalysis. Reduction of the C7-C8 double bound of ring B is performed by chlorine reductase. Next the hydroxyethyl group on ring A of 3-hydroxyethyl bacteriochlorophyllide a is oxidized by a dehydrogenase to the corresponding acetyl group. Biochemical and structural information is lacking for the enzymes performing these late steps of bacteriochlorophyll biosynthesis.

Due to the incorporation of a long-chain alcohol residue, chlorophylls and bacteriochlorophyll are highly hydrophobic molecules. During this process phytol, a C_{20} isoprenoid alcohol, forms an ester bond to the proprionate side chain of ring D. The corresponding enzymes are called bacteriochlorophyll (BchG)- and chlorophyll synthases (ChlG). Higher plants and green algae also contain chlorophyll b, where the methyl group in position 3 has been converted into a formyl group (–CHO) by an O_2 and $NADP^+$ dependent reaction. Chlorophyll c occurs in many algae. It is not a chlorin, but a porphyrin with an unreduced ring D. It is probably synthesized directly from Mg-divinylphaeoporphyrin a_5.

Pheophytins are Mg-free chlorophyll-like compounds. They occur in the photosystem II of plants (pheophytin a) or in the photosynthetic reaction center of bacteria (pheophytins a, b or g). It is assumed that they are generated from the respective chlorophylls by loss of the metal atom.

Literature:

Grimm, B. *et al. Chlorophylls: Biochemistry, Biophysics, Functions and Applications.* Springer, 2006.

Heyes, D.J. and Hunter, C.N, in Warren, M.J., Smith, A.G. (Eds.): *Tetrapyrroles: Birth, Life and Death.* Landes Bioscience, Austin, 2009.

3.3.5 Biosynthesis of Cobalamins (Vitamin B_{12} Derivatives, Fig. 3.3-8)

Vitamin B_{12} and its derivatives represent some of the most complex small molecules found in biochemistry. They contain a corrin ring, in which one of the methine bridges is missing. The central cobalt atom is coordinated with six ligands. Four of them are pyrrols of the corrin ring, to which is attached a nucleotide with the unusual base dimethylbenzimidazole. The metal is bound by the nitrogens of the pyrrole rings and a nitrogen of the base. The axial sixth ligand for the cobalt can be 5′deoxyadenosine in 5′deoxyadenosylcobalamin (coenzyme B_{12}), a methyl group in methylcobalamin, cyanide in vitamin B_{12} (cyanocobalamin, actually an artifact from the extraction procedure), hydroxyl or water. Other forms of naturally occurring cobalamins are norcobalamin with ethanolamine instead of the aminopropanol linker and pseudocobalamin with adenine replacing the dimenthylbenzimidazole base.

Coenzyme B_{12} is biosynthesized exclusively by microorganisms in about 30 enzyme catalyzed steps. The first reactions of the pathway up to uroporphyrinogen III are common to all tetrapyrroles (3.3.1). For the further steps of cobalamin biosynthesis two distinct pathways exist; an aerobic and an anaerobic one. Their major differences are associated with ring contraction and cobalt insertion mechanisms. However, major parts have been highly conserved in evolution. Figure 3.3-8 shows the steps in *Pseudomonas denitrificans*.

The major further steps of cobalamin biosynthesis are:

1. Addition of eight SAM-derived methyl groups to the periphery of the tetrapyrrole ring structure.
2. Extrusion of the methylated C20 carbon as acetic acid or acetaldehyde, resulting in ring contraction.
3. Cobalt chelation and reduction from Co^{++} level to the Co^{+} level.
4. Decarboxylation of the ring C acetate side chain.
5. Amidation of six acidic side chains.
6. Adenosylation of the cobalt to generate the upper axial ligand.
7. Attachment of aminopropanol (3.3.1.1) to the ring D propionate side chain.
8. Formation of the lower nucleotide loop by attachment of α- ribazole.

These compounds are essential participants in several rearrangement and methyl transfer reactions. The needs of mammals (in humans < 10 µg/day) are met by the B_{12} production of intestinal bacteria or from animal food sources. Resorption and further metabolism, as well as their biochemical role are described in 3.7.5.

3.3.6 Siroheme Biosynthesis

The iron-containing isobacteriochlorin siroheme (Fig. 3.3-1) is not a real heme. It shares the initial steps of biosynthesis with the

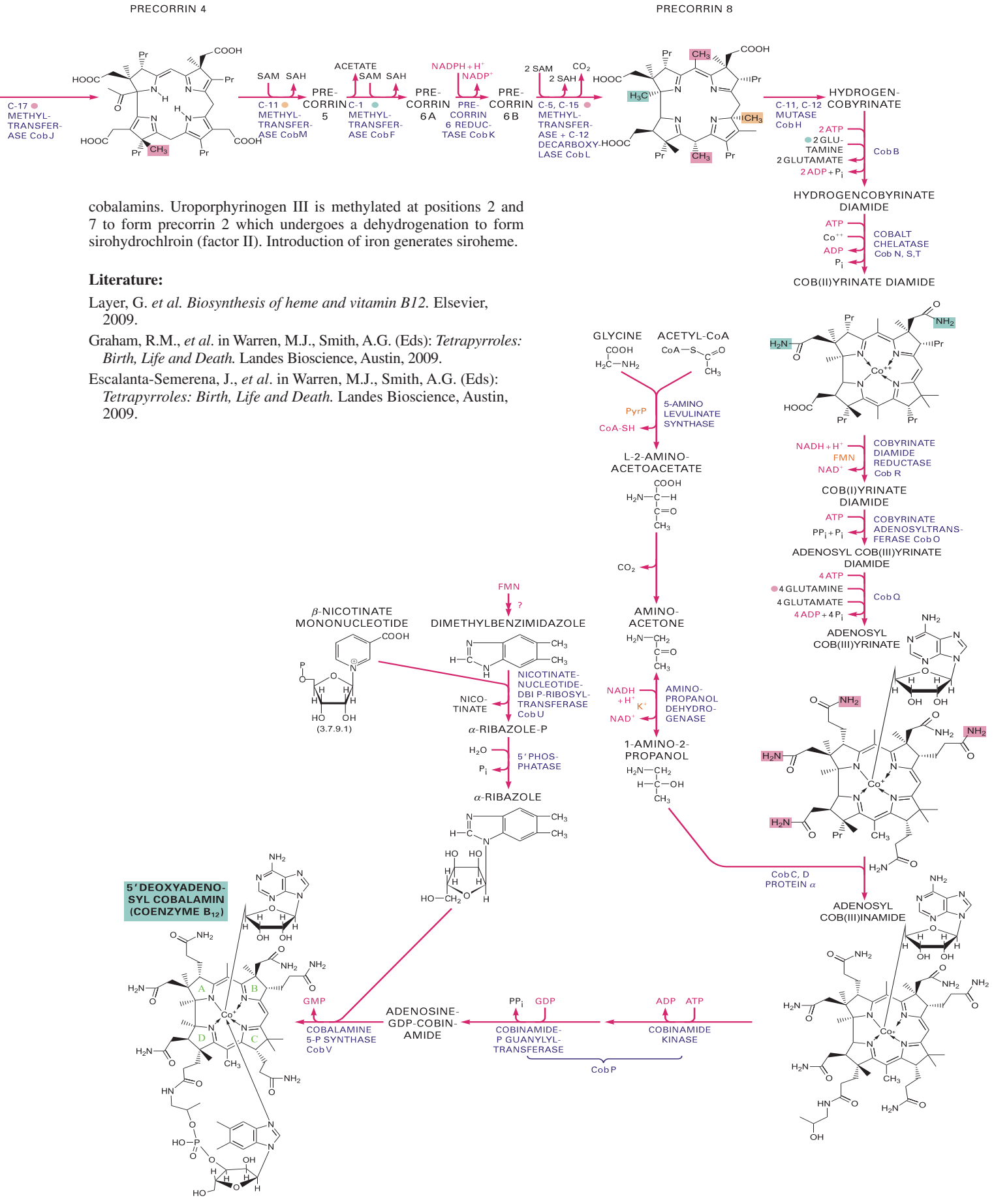


Figure 3.3-8. Biosynthesis of Coenzyme B_{12}

cobalamins. Uroporphyrinogen III is methylated at positions 2 and 7 to form precorrin 2 which undergoes a dehydrogenation to form sirohydrochlroin (factor II). Introduction of iron generates siroheme.

Literature:

Layer, G. *et al. Biosynthesis of heme and vitamin B12.* Elsevier, 2009.

Graham, R.M., *et al*. in Warren, M.J., Smith, A.G. (Eds): *Tetrapyrroles: Birth, Life and Death.* Landes Bioscience, Austin, 2009.

Escalanta-Semerena, J., *et al*. in Warren, M.J., Smith, A.G. (Eds): *Tetrapyrroles: Birth, Life and Death.* Landes Bioscience, Austin, 2009.

3.4 Lipids and Glycolipids

Röbbe Wünschiers

The chemistry and structure of major lipids is presented in Section 1.4.

3.4.1 Fatty Acids and Acyl-CoA

Most fatty acids of biological importance have a chain length in the range of C_{14} … C_{20}. Only a small portion of them occur as free fatty acids in nature. In esterified form they are essential constituents of membranes (e.g., as phospholipids, 3.4.3 and glycolipids, 3.4.4). A smaller portion of amide derivatives are also formed (e.g., ceramides). Biosynthesis and degradation of fatty acids proceed by addition or removal of C_2 units, however in different ways (Table 3.4.1-1).

Table 3.4.1-1. Differences Between Fatty Acid Biosynthesis and Degradation Mechanisms in Eukarya

	Biosynthesis	Degradation
Location	cytosol (in plants: stroma of chloroplasts)	mitochondrial matrix (in plants: glyoxisomes)
Enzyme	multienzyme complex	individual enzymes, partly associated
Fatty acid chain is bound to	acyl carrier protein	coenzyme A
Added / removed C_2 unit	malonyl-CoA – CO_2	acetyl-CoA
Configuration of the hydroxyacyl derivative	D	L
Redox coenzyme	NADPH (yeast: + $FADH_2$)	NAD^+ and FAD

3.4.1.1 Biosynthesis of Fatty Acids

The primary sources for the biosynthesis of fatty acids in animals are carbohydrates, followed by a number of amino acids. Their degradation yields acetyl-coenzyme A, which is present in mitochondria (3.1.3.1, 3.2.2 … 3.2.7). This is the starting compound of fatty acid biosynthesis, which proceeds in the cytosol of many organs (with preference in the liver) according to the formula

$$8\ \text{acetyl-CoA} + 14\ \text{NADPH} + 7\ \text{ATP} = \text{palmitate} + 14\ \text{NADP}^+ + 8\ \text{CoA} + 6\ \text{H}_2\text{O} + 7\ \text{ADP} + 7\ \text{P}_i.$$

Since acetyl-CoA cannot pass through the mitochondrial membrane into the cytosol, it has to be transferred there via the citrate/pyruvate shuttle (Fig. 3.4.1-1), where ATP is required to regenerate the high-energy bond of acetyl-CoA in the cytosol. The carboxylation step of acetyl-CoA to malonyl-CoA in the first committed and rate regulating step of fatty acid biosynthesis is performed by the multienzyme complex acetyl-CoA carboxylase (ACC). Biotin functions as a CO_2 carrier (3.7.8.2).

The overall reaction of ACC comprises two distinct half-reactions, the carboxylation of biotin with hydrogen carbonate and the subsequent transfer of CO_2 group from carboxybiotin to acetyl-CoA, resulting in malonyl-CoA. For this two-step reaction, ACC requires two different protein sub-assemblies, biotin carboxylase (49 kDa), and transcarboxylase (35 + 33 kDa). A third protein, biotin carboxyl carrier protein (BCCP, 22 kDa, Fig. 3.4.1-2), carries the essential biotin cofactor covalently bound to a lysine residue located 35 residues from the carboxyl terminus. In mammals, fungi, and plant cytosols, all three ACC components are part of one polypeptide chain (210 … 265 kDa).

In *E. coli*, acetyl-CoA carboxylase comprises four distinct proteins (Fig. 3.4.1-2): biotin carboxyl carrier protein (BCCP; AccB), biotin carboxylase (AccC) and two proteins catalyzing the carboxyltransferase partial reaction (AccA and AccD). The consecutive reactions of fatty acid synthesis are catalyzed in fungi and animals by the cytoplasmic multienzyme complex fatty acid synthase (240 … 275 kDa), which consists of two (yeast: α, β) or one (vertebrates) multifunctional polypeptide chains. The yeast enzyme has the structure $\alpha_6\beta_6$, the animal enzyme is a homodimer (Fig. 3.4.1-3). A 10 kDa acyl carrier protein (ACP) component acts as the central carrier of the growing acyl chain, using a 4′-phosphopantetheine moiety as a 'swinging arm' (2 nm length, 9.7.2) to move the acyl intermediates to the various active sites.

In plant plastids and bacteria, the reaction steps are catalyzed by individual enzymes, which are closely associated. Here, a preferred alternative is the use of acetyl-CoA instead of acetyl-S-enzyme for condensation with malonyl-ACP in the synthase reaction. Out of the three plant oxoacyl synthases, enzyme III catalyzes the condensation of acetyl-CoA with malonyl-ACP, enzyme I the elongation up to C_{16} and enzyme II the final step to C_{18}.

Reaction sequence (Fig. 3.4.1-4): Specific acyl transferases (one in animals, two in yeast) transfer an acetyl residue from acetyl-CoA to a cysteine-SH of the 3-oxoacyl synthase component ('peripheral SH group') ① and a malonyl residue to the 'central' SH group of the ACP component ⑯ of the multienzyme system. Condensation with an acetoacetyl chain (bound to the central SH group) occurs with simultaneous decarboxylation, thereby energetically favoring chain elongation ②. This step is irreversible. Then the resulting C_4 residue is reduced by NADPH at the oxoacyl reductase component ③, dehydrated to a 2,3-desaturated intermediate ④ and reduced again by the enoyl reductase ⑤ (also using NADPH as a reductant; NADH in *E. coli*). The yeast enzyme contains flavin. Finally, the butyryl chain is transferred to the 'peripheral' SH group ①. Condensation

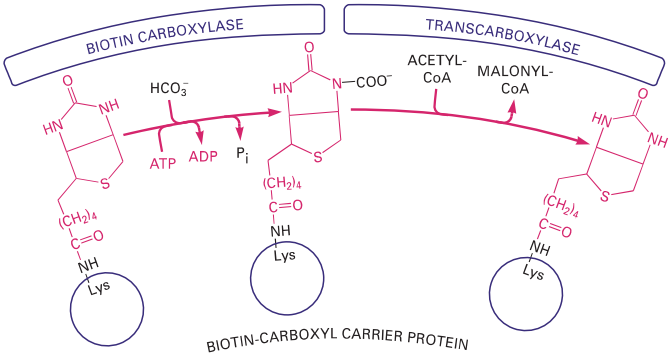


Figure 3.4.1-2. Carboxylation of Acetyl-CoA in *E. coli*

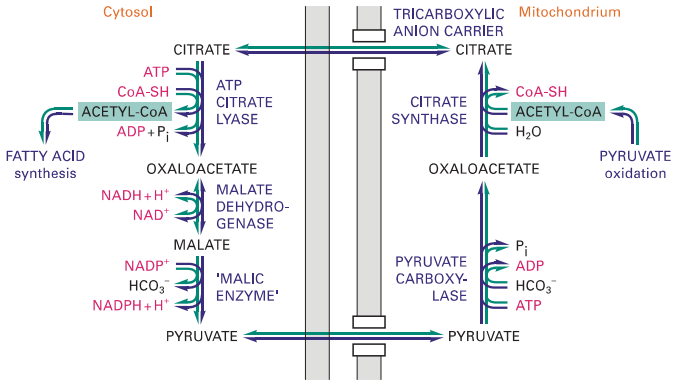


Figure 3.4.1-1. Citrate-Pyruvate Shuttle for Acetyl-CoA Export from Mitochondria

Green and blue arcs: Functional interaction variants of the acyl-ACP moiety with the enzyme domains. X = domain without enzymatic function. Z = NADPH binding sites. The numbers correspond to Figure 3.4.1-4.

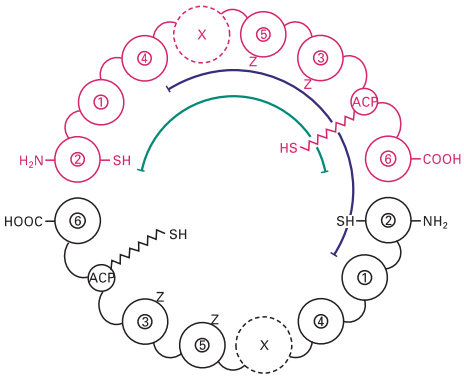
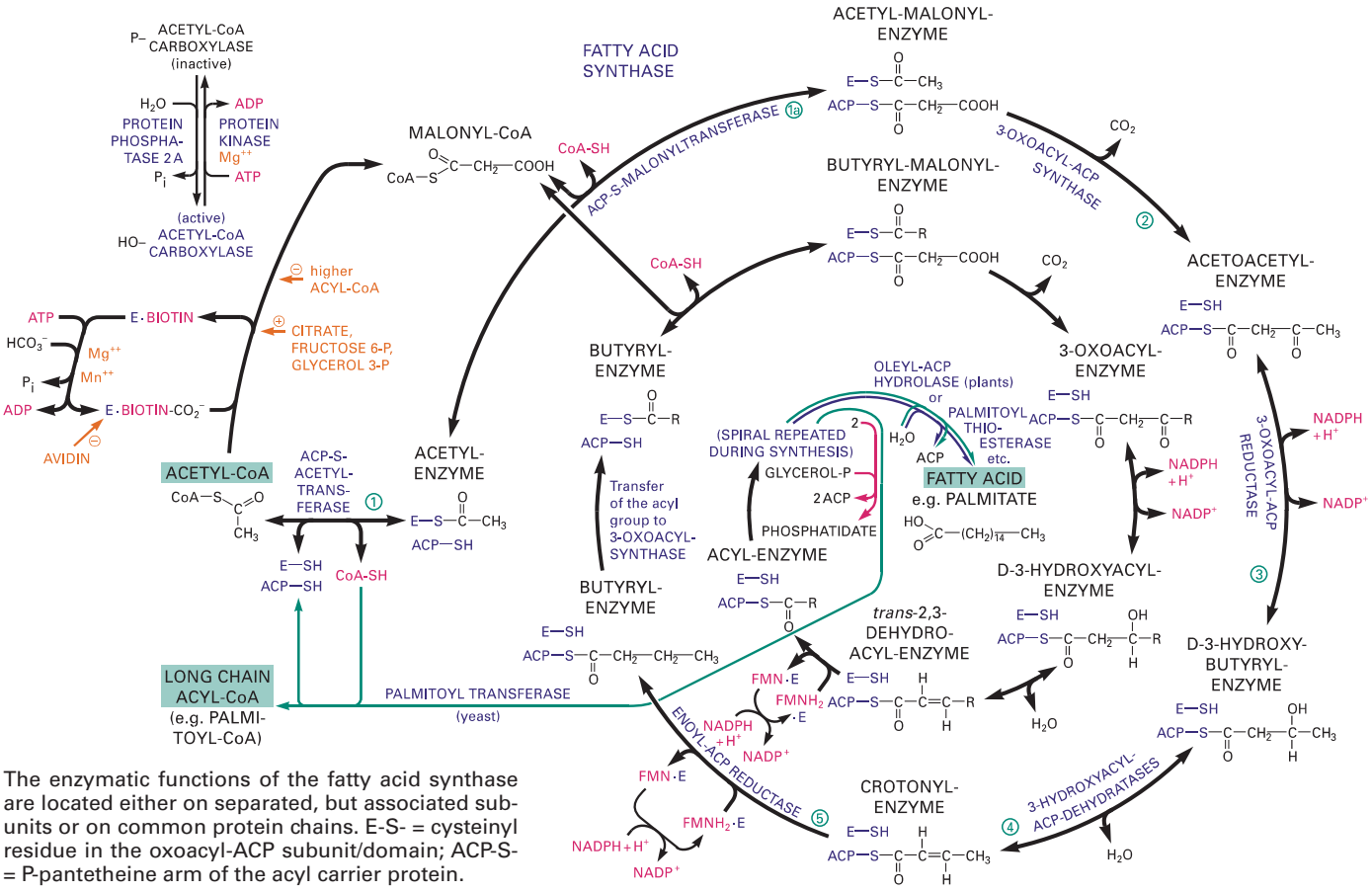


Figure 3.4.1-3. Schematic Drawing of the Animal Fatty Acid Synthase Dimer (modified after Joshi and Smith)



The enzymatic functions of the fatty acid synthase are located either on separated, but associated subunits or on common protein chains. E-S- = cysteinyl residue in the oxoacyl-ACP subunit/domain; ACP-S- = P-pantetheine arm of the acyl carrier protein.

Figure 3.4.1-4. Steps of Fatty Acid Synthesis

with another malonyl residue initiates another round of chain elongation. The fatty acid synthesis is terminated in vertebrates by hydrolysis of the long-chain (mainly C_{16}) acyl residue by palmitoyl thioesterase ⑥ to give free palmitic acid. About half of the NADPH required for the reduction steps is provided by the 'malic enzyme' reaction (Fig. 3.4.1-1), the rest is supplied by the initial reactions of the pentose phosphate cycle (3.1.6.1).

A similar reaction takes place in plants. The enzyme prefers oleyl-ACP. In yeast, a specific palmitoyl transferase transfers the product to CoA, yielding palmitoyl- (or stearoyl-) CoA. In bacteria, long-chain acyl residues are directly transferred to glycerophosphate for the synthesis of membrane phosphatidates.

Biosynthesis of fatty acids of medium chain length: Fatty acids of C_6 … C_{10} length are present especially in milk fat. In lactating mammary glands of ruminants, their presence is due to the broad specificity of acyl transferase which already removes medium-chain-length fatty acids from the complex. In contrast, the mammary glands of non-ruminants and the preen glands of birds contain a second 29-kDa thioesterase which releases C_8 … C_{14} fatty acids.

Biosynthesis of odd-numbered and branched fatty acids: Due to the addition of C_2 units during synthesis, fatty acids with an even number of carbon atoms are most abundant in nature. However, fatty acid synthase can also utilize propionyl-CoA instead of acetyl-CoA as a primer, generating saturated fatty acids with odd carbon-numbers (mainly C_{15} and C_{17}). If malonyl-CoA is replaced by methylmalonyl-CoA, methyl-branched fatty acids are formed (e.g., in sebaceous glands, where they can be the major products).

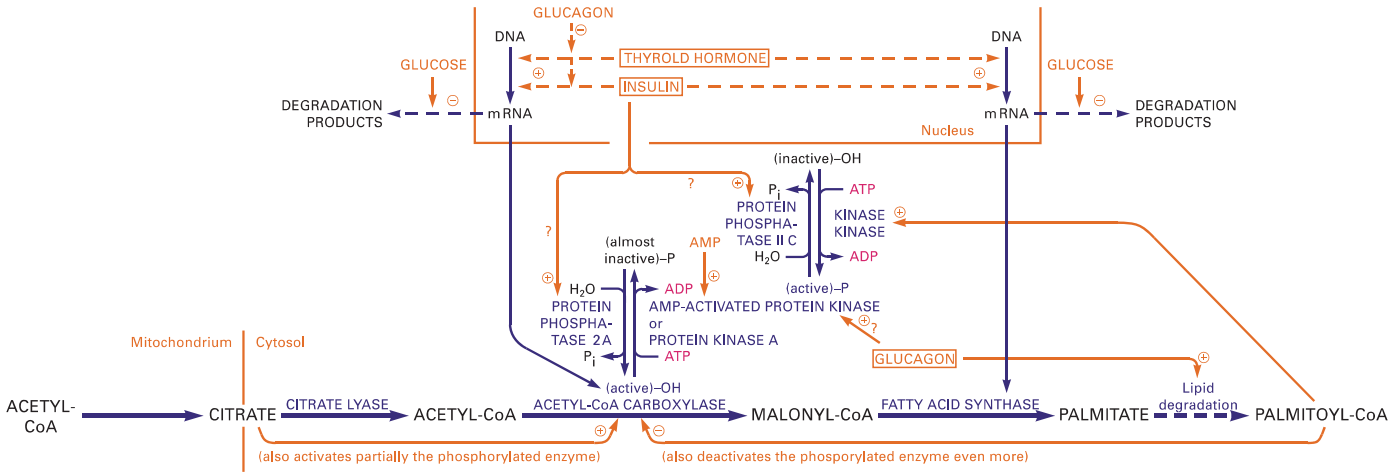


Figure 3.4.1-5. Control Mechanisms for Acetyl-CoA Carboxylase in Liver

3.4.1.2 Regulation of Fatty Acid Synthesis (Fig. 3.4.1-5)

The carboxylation of acetyl-CoA is the committed step of fatty acid biosynthesis (for definition, see 3.1.1.1) and the point of action for regulatory mechanisms.

In vertebrates, acetyl-CoA carboxylase exists as a phosphorylated, enzymatically inactive homodimer (450 kDa) and as a dephosphorylated, active, filamentous polymer (up to 10^4 kDa). The interconversion is allosterically stimulated by citrate (feed-forward activation, indicating sufficient supply of the precursor in the cytosol) and inhibited by palmitoyl-CoA (feed-back inhibition, palmitoyl-CoA is elevated during starvation due to degradation of fats). Citrate can partially counteract the loss of activity by phosphorylation, while palmitoyl-CoA enhances the phosphorylation effect.

Inactivating phosphorylations at different serine residues are catalyzed by an AMP-dependent protein kinase (which, in turn, is regulated by activating phosphorylation and by inactivating dephosphorylation and appears to be the most important kinase), by acetyl-CoA carboxylase kinase type 2 (in mammary glands) and by cAMP dependent protein kinase A (7.4.2). The reactivating dephosphorylation of the carboxylase is catalyzed by protein phosphatase 2A. The interconversions are under the control of hormones, although the mechanisms are not known in detail.

- In well-fed states with ample supply of glucose, insulin is secreted (Fig. 3.1.1-7). This causes a low degree of phosphorylation of the enzyme. The activity of the carboxylase and the rate of fatty acid synthesis are high. In liver, fatty acid synthase expression is stimulated by insulin. Transcription factors mediating the stimulatory effect of insulin include upstream stimulatory factors (USFs) and sterol response element binding protein 1 (SREBP-1). In fat cells, expression of SREBP-1 and of fatty acid synthase is inhibited by leptin, a hormone involved in regulation of food intake and fat metabolism.
- During fasting and in sudden energy demand, glucagon and epinephrine, respectively, are secreted (Fig. 3.1.1-7). In response to these hormones, the degree of inactivating phosphorylation is high (in liver and in adipocytes, respectively) and the rate of fatty acid synthesis is low or nil, since at first, the glucose requirements of various cells (primarily brain and erythrocytes) have to be satisfied.
- Fatty acid synthesis is also low in diabetes, where there is a low insulin level.

In mammals, two isoforms of the ACC enzyme exist, Acc1 and Acc2. Acc1 is expressed mainly in lipogenic tissues (e.g., liver, lactating mammary gland, adipose tissue) and catalyzes the rate-determining and committed step in the biosynthesis of long-chain fatty acids. In contrast, Acc2 is expressed mainly in the heart and skeletal muscle, and its malonyl-CoA product functions as a potent inhibitor of fatty acid oxidation in these tissues.

The major portion of fatty acid synthesis and consecutive esterification takes place in the liver (for transport of lipids to other organs, see 6.2.2. There are interconnections with the regulation of cholesterol synthesis (3.5.1.4). The *de novo* synthesis of fatty acids in plants occurs in the plastids. Regulation of chloroplast fatty acid synthesis is mediated by the response of acetyl-CoA carboxylase to the redox state of the plastid, which ensures that the carbon metabolism is linked to the energy status.

Long-term adaptive control of fatty acid synthesis occurs through changes in the rates of synthesis (at the transcription level) and degradation of the enzymes participating (acetyl-CoA carboxylase and fatty acid synthase). The formation of fatty acid synthase- and malic enzyme-mRNAs is enhanced by insulin and thyroid hormones and decreased by glucagon and cAMP. Polyunsaturated (but not saturated) fatty acids decrease the rate of transcription. High glucose levels stabilize the mRNAs.

In *E. coli*, citrate has no effect on regulation. The activity of the carboxylase is controlled by guanine nucleotides.

3.4.1.3 Fatty Acid Desaturation and Chain Elongation (Fig. 3.4.1-7)

Desaturation of fatty acids takes place at preferred sites (Fig. 3.4.1-6). The numbering system can start at either end.

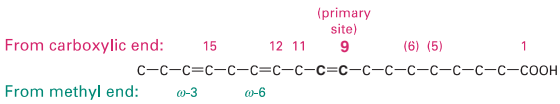


Figure 3.4.1-6. Common Sites of Desaturation (Example: C_{18} Acid)

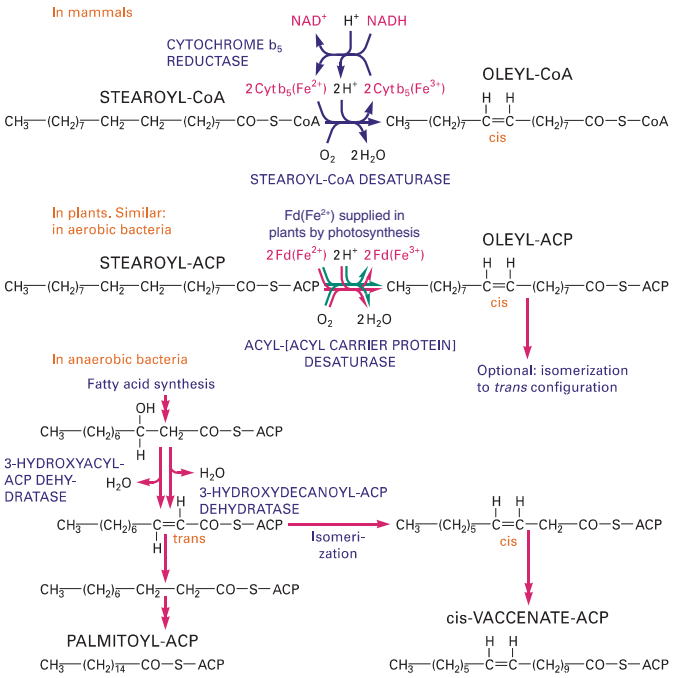


Figure 3.4.1-7. Mechanisms of Desaturation

- In mammals, desaturation takes place with the CoA derivatives of preformed long chain fatty acids (C_{14} … C_{18}), preferably at the 9-10 bond of stearoyl-CoA, yielding the *cis*-compound oleyl-CoA. This is an oxidative reaction in the endoplasmic reticulum of the liver, catalyzed by stearoyl-CoA desaturase (Δ^9 -desaturase, 53 kDa, a non-heme iron protein, highly sensitive to cyanide). The enzyme complex also contains cytochrome b_5 (16.7 kDa) and cytochrome-b_5 reductase (43 kDa). The expression of the enzymes depends greatly on the nutritional state.
- Mammals are unable to introduce double bonds beyond the Δ^9 position and have to obtain polyunsaturated fatty acids by food intake from plant sources (essential fatty acids: linoleic and linolenic acids, 18:2 and 18:3, respectively).
- However, they can form double bonds closer to the carboxyl end, e.g., at 5-6 or 6-7 bonds of linoleic and linolenic acids. The enzymatic mechanism is similar to introduction of the first double bond (see above). Additionally, chain elongation may take place. One of the products is arachidonic acid (7.4.8).
- In seeds of the model plant *A. thaliana*, eight fatty acid species are predominantly found: 16:0, 18:0, 18:1, 18:2 and 18:3 (polyunsaturated fatty acids, PUFAs), 20:0, 20:1 and 22:1 (very long chain fatty acids, VLCFAs).
- In plant plastides, the most common initial desaturation takes place at the 9-10 bond by stearoyl-ACP desaturase, which is a soluble homodimer. A similar reaction is performed by bacteria under aerobic conditions. After reduction by ferredoxin, the desaturase binds oxygen by a diiron-cluster, which then attacks the fatty acid, removes hydrogen and forms the double bond. Thereafter, other enzymes desaturate at definite distances from the methyl end, e.g., at ω-6 and ω-3, resulting in cis-polyunsaturated fatty acids.
- Starting from 9,10-unsaturated fatty acids, bacteria can form 9,10-cyclopropane fatty acids by introduction of a methylene residue from S-adenosylmethionine. This modification increases the oxygen stability of the fatty acid, preserving the rigidity found at double bonds.

- In anaerobic bacteria, unsaturated fatty acids are produced by a variation of the fatty acid synthase system (3.4.1.1). A different 3-hydroxydecanoyl-ACP-dehydratase catalyzes an alternative dehydratase reaction at the C_{10} level of fatty acid synthesis. The reaction product is not the *trans*-2,3-, but rather the *cis*-3,4-unsaturated acid. Immediately thereafter the synthase reaction takes place, which is catalyzed by a special 3-oxoacyl-ACP synthase. The final result is *cis*-vaccenate (18 : 1, cis-Δ^{11}), a major fatty acid of microorganisms.
- Bacteria can also produce *trans*-unsaturated acids (TFA) by isomerization of preformed *cis*-unsaturated fatty acids. This is apparently a measure to control membrane fluidity. This proceeds, e.g., in the gastrointestinal tract of ruminants. TFAs are consecutively found in milk and in beef fat. (They also occur in chemically hydrogenated plant oils.) TFAs show adverse effects on blood lipids, especially on the low-density lipoprotein (LDL): high-density lipoprotein (HDL) cholesterol ratio but their entire influence on metabolic processes in mammals has not been unambiguously resolved.

Importance of unsaturated fatty acids: Unsaturated fatty acids are essential for membrane fluidity. Polyunsaturated fatty acids are precursors of biologically active molecules, e.g., eicosanoids (7.4.8). Their controlled release from membrane phospholipids may even initiate control functions by modulation of transcription via specific receptors and response elements, of protein kinase C isoforms (7.4.3), of guanylate and adenylate cyclases etc. In humans, tissue availability of polyunsaturated fatty acids is of major importance for health, and it depends on both dietary intake and metabolic turnover. A lack causes severe disturbances (weight loss, change in lipid composition and in organ function, death). If saturated fatty acids prevail in food over polyunsaturated ones, atherosclerosis and coronary heart disease may result.

Elongation: In animals, the *de novo*-synthesis in the cytosol usually ends with palmitate. In the endoplasmic reticulum, further condensations of malonyl-CoA with long chain acyl-CoA take place, followed by reactions analogous to the reaction sequence above (3.4.1.1). This leads to C_{18} and higher acids. There are two different condensation enzymes for saturated and unsaturated fatty acids.

3.4.1.4 Transport and Activation of Fatty Acids (Fig. 3.4.1-8)

Fatty acids are transported between animal organs either as fatty acids complexed with serum albumin or in the form of triacylglycerols associated with lipoproteins (6.2.2). These are hydrolyzed outside of the cells by membrane-bound lipoprotein lipase to yield free fatty acids. Once long chain fatty acids have entered the cells, they either diffuse or are transported to mitochondria, peroxisomes or the endoplasmic reticulum where they are activated by formation of their CoA thioesters for consecutive conversions. Fatty acid binding proteins (FASBPs, 14 … 15 kDa) in the cytosol of various animal tissues may promote the cellular uptake and transport of fatty acids and target them to specific metabolic pathways.

Activation of fatty acids by ATP dependent fatty-acid-CoA ligases (acyl-CoA synthetases, Fig. 3.4.1-8), a group of enzymes with different chain length specificities, occurs via the intermediate acyl adenylate. Hydrolysis of the released diphosphate by pyrophosphatase drives the reaction. The acyl residue is then transferred to the SH group of coenzyme A to give an acyl-CoA thioester which serves as a substrate for β-oxidation (3.4.1.5) or for triglyceride biosynthesis (3.4.2.1). In mammals, long-chain acyl-CoA synthetases (ACSL) are required for fatty acid degradation, phospholipid remodeling, and the production of long acyl-CoA esters regulating multiple physiological processes.

They can be classified into five sub-families based on the chain length of their preferred acyl groups:

- acyl-CoA synthetase short-chain (ACSS), C_2 to C_4;
- acyl-CoA synthetase medium-chain (ACSM), C_4 to C_{12};
- acyl-CoA synthetase long-chain (ACSL), C_{12} to C_{20};
- acyl-CoA synthetase bubblegum-related (ACSBG, based on the fatty acid accumulation in the *Drosophila melanogaster* mutant *bubblegum*), C_{14} to C_{24};
- acyl-CoA synthetase very long-chain (solute carrier family 27A, SLC27A), C_{18} to C_{26}.

For **degradation** via the mitochondrial β-oxidation pathway, the acyl group of cytosolic acyl-CoA compounds passes the mitochondrial membrane bound to carnitine (3.2.5.2) via a carnitine shuttle operated by carnitine-acylcarnitine translocase (Fig. 3.4.1-8) and is afterwards transferred to CoA again.

Mammalian carnitine acyltransferases exhibit different chain length specificities: two isoenzymes of carnitine O-palmitoyltransferase (CPT-I and CPT-II) present on both sides of the mitochondrial inner membrane react specifically with long-chain acyl-CoA compounds, while carnitine O-acetyltransferase I found in microsomes, peroxisomes and mitochondria accepts C_2 … C_{10} residues from acyl-CoA. A carnitine O-octanoyltransferase is also found.

3.4.1.5 Fatty Acid Oxidation (Figs. 3.4.1-9 and 3.4.1-10)

The overall sequence of reactions for fatty acid oxidation is similar to a reversal of the fatty acid synthesis. During each reaction cycle, the acyl chain is shortened by removal of a two-carbon unit as acetyl-CoA (β-oxidation). However, there are some differences between the degradation and synthesis pathways, as shown in Table 3.4.1-1.

β-Oxidation in animals starts with the acyl-CoA dehydrogenase reaction. The removed hydrogens are taken up by FAD, which is reoxidized by the electron transferring flavoprotein, ETF (a FAD-linked matrix protein). ETF transfers the reducing equivalents via the ETF-ubiquinone dehydrogenase to the ubiquinone pool of the respiratory chain (3.11.4). The other reducing step of the fatty acid oxidation, the 3-hydroxyacyl-CoA dehydrogenase reaction yields NADH, which is reoxidized by the NADH dehydrogenase complex of the respiratory chain (3.11.4.1). The liberation of acetyl-CoA is performed by the acetyl-CoA C-acyltransferase ('general' 3-oxoacyl-CoA thiolase). Only the final cleavage of acetoacetyl-CoA is catalyzed by acetyl-CoA C-acetyltransferase (acetoacetyl thiolase). However in liver, a portion of acetoacetyl-CoA is used for the formation of cholesterol (3.5.1.1) and of ketone bodies (3.4.1.7).

There are multiple enzymes for each of the constituent steps of fatty acid degradation which vary in their chain-length specificity.

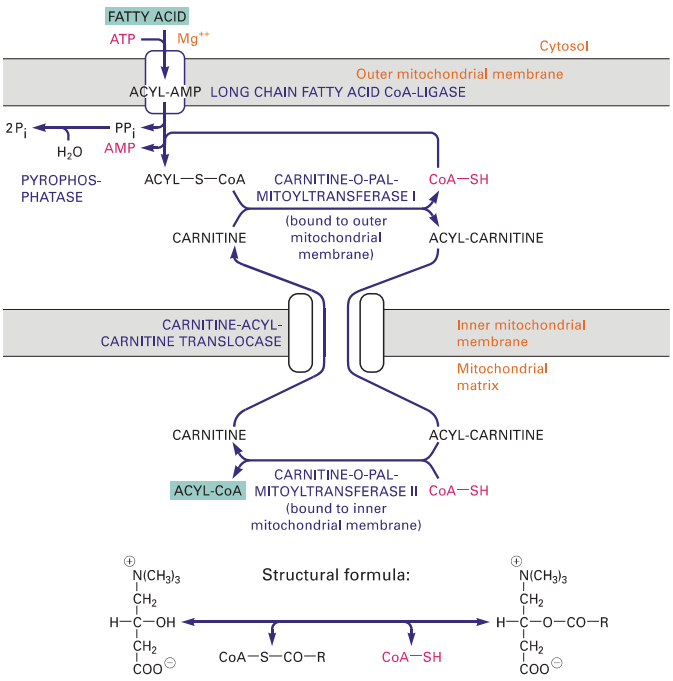


Figure 3.4.1-8. Activation of Fatty Acids and Carnitine Shuttle
(Example: long chain fatty acids)

In eukarya, the enzymes which turn over long chain substrates are bound to the inner mitochondrial membrane. The enoyl-CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase and acetyl-CoA C-acyltransferase activities are combined into a trifunctional enzyme with the first two activities located on the α-subunit and the third on the β-subunit. The enzymes with specificity for shorter chain length are present in the mitochondrial matrix. They can be isolated as separate enzymes but may be organized as a multienzyme complex providing 'substrate channeling'. This question has not completely been solved. A model of the possible arrangement is shown in Figure 3.4.1-10.

Odd-numbered fatty acids yield propionyl-CoA as the product of the final thiolysis step. Its conversion into succinyl-CoA is shown in Figure 3.2.5-2.

During the degradation of unsaturated fatty acids, an isomerase catalyzes the conversion of a *cis*-Δ^3 into a *trans*-Δ^2 double bond (compare Fig. 3.4.1-7), which is a regular member of the β-oxidation sequence. Further enzymes provide the reduction of additional double bonds in polyunsaturated fatty acids.

Medical aspects of fatty acid oxidation: Most common are deficiencies of carnitine O-palmitoyltransferase (CPT II) and of long-chain 3-hydroxyacyl-CoA and acyl-CoA dehydrogenases (ca 1:10,000 for homozygotes). They result in non-ketotic hypoglycemia, muscle breakdown, cardiomyopathy etc.

Regulation of fatty acid oxidation in mammals: The rate of fatty acid oxidation is primarily a function of the plasma concentration of unesterified fatty acids which is regulated by the action of hormones like insulin and glucagon (7.1.3) on the hormone-sensitive lipase (3.4.2.2). In the liver, the rate of mitochondrial β-oxidation depends on the entry of acyl groups into the mitochondria. This process is regulated by modulation of the carnitine O-palmitoyltransferase I activity (Fig. 3.4.1-8), which is inhibited by malonyl-CoA. In addition, there is feedback inhibition of all of the fatty acid oxidation enzymes by the products.

If there is ample supply of glucose, acetyl-CoA and malonyl-CoA are formed and used for fatty acid synthesis. The counteracting fatty acid oxidation is switched off by malonyl-CoA. This inhibition ceases when the glucose level is low (see Fig. 3.1.1-7). Fatty acids are oxidized for the specific requirements of the organ and for the production of ketone bodies (3.4.1.7).

Initiation of fatty acid degradation by an oxidase reaction: In animals, there is also considerable fatty acid degradation in the peroxisomes of the liver and of other organs. In contrast to the mitochondrial system, the initial step of the degradation is catalyzed by the FAD-containing acyl-CoA oxidase and produces H_2O_2. This toxic compound is decomposed by catalase (3.2.5.8).

$$\text{acyl-CoA} + O_2 = \text{trans-2,3-dehydroacyl-CoA} + H_2O_2$$
$$2\ H_2O_2 = H_2O + O_2$$

The further reactions leading to acetyl-CoA correspond to the mitochondrial sequence. The peroxisomal degradation system shortens only long fatty acids and does not result in acetoacetate. Peroxisomes do not possess a mechanism for the reoxidation of NADH which is produced by the 3-hydroxyacyl-CoA dehydrogenase reaction. They have to transfer the reduction equivalents into the cytosolic space.

Fatty acid oxidation in plants is an important part of the mobilization of storage oil and takes place in glyoxysomes where β-oxidation and part of the glyoxylate cycle (3.1.9.1) occurs. Glyoxysomes are structurally similar but metabolically different from the more ubiquitous peroxisomes. Glyoxysomes comprise two enzymes that are

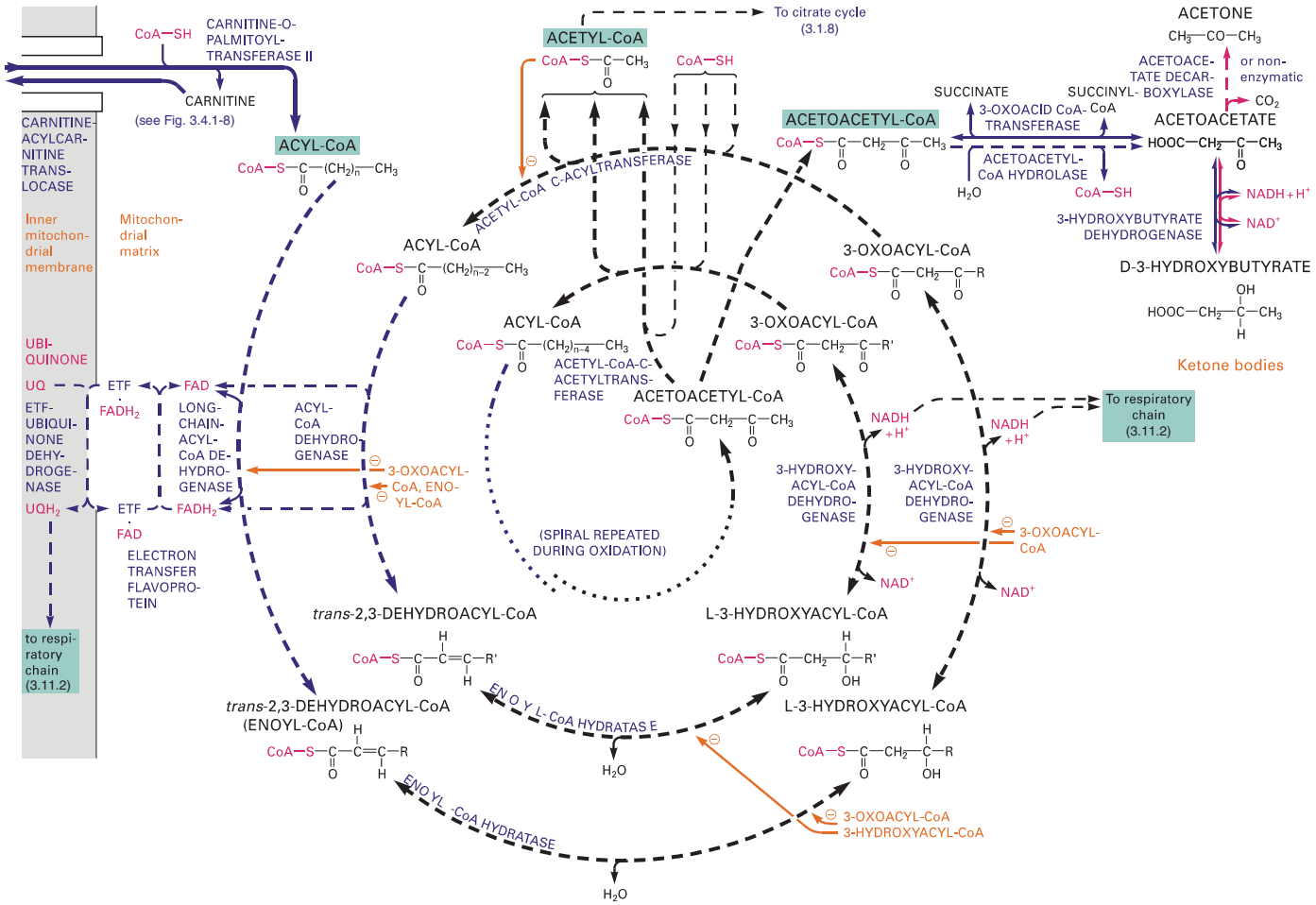


Figure 3.4.1-9. β-Oxidation of Fatty Acids. There are several enzymes catalyzing each step that differ in their chain-length specificity. A similar acyl-CoA dehydrogenase occurs in bacteria. In plants, an oxidase reaction takes its place.

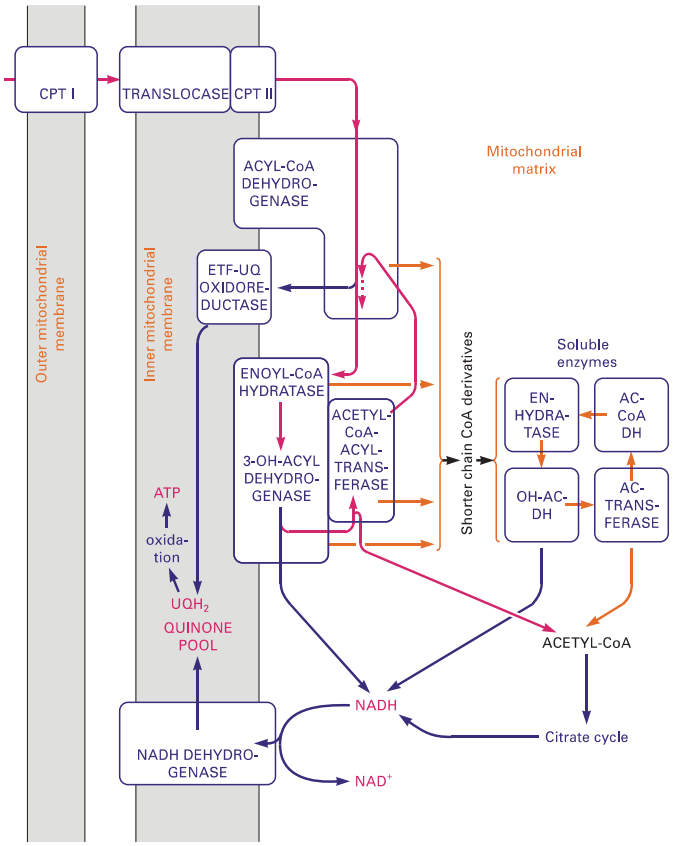


Figure 3.4.1-10. Possible Arrangement of the Enzymes for β-Oxidation in Mammals (modified after Eaton *et al.*)
Red arrows: reactions of long chain acyl derivatives. Orange arrows: reactions of medium and short chain acyl derivatives. Blue arrows: Transfer of hydrogen.

unique to the glyoxylate cycle, malate synthase (MLS) and isocitrate lyase (ICL). β-oxidation converts fatty acids to acetyl-CoA, which is converted into 4-carbon compounds via the glyoxylate cycle. These 4-carbon molecules are subsequently transported into the mitochondria, where they can either be used as respiration substrates, or are converted into malate and are transported to the cytosol for gluconeogenesis.

The fatty acid oxidation in *E. coli* proceeds via the same steps as in animal mitochondria. On the other hand, however, *E. coli* can convert acetyl-CoA into oxaloacetate via the glyoxylate cycle in a way similar to plants.

3.4.1.6 Energy Yield of the Fatty Acid Oxidation

Oxidation of fatty acid provides a major source of energy, especially for the heart and skeletal muscle of mammals. The complete degradation of palmitoyl-CoA by β-oxidation proceeds as follows:

$$\text{palmitoyl-CoA} + 7\ \text{FAD} + 7\ \text{NAD}^+ + 7\ \text{CoA} + 7\ \text{H}_2\text{O} =$$
$$8\ \text{Acetyl-CoA} + 7\ \text{FADH}_2 + 7\ \text{NADH} + 7\ \text{H}^+ \qquad \Delta G'_0 = -9797\ \text{kJ/mol}.$$

Considering the ATP yield of the citrate cycle (3.1.8) and the respiratory chain (3.11.3; using the ratios 2.5 ATP/1 NADH; 1.5 ATP/1 FADH), a total of up to 108 ATP is formed. Since two energy-rich phosphate bonds are consumed in the activation of palmitate, the net yield is 106 mol ATP / mol palmitate. This corresponds to a total energy conservation of about 33 % under standard conditions.

3.4.1.7 Ketone Bodies

The acetoacetyl-CoA formed in fatty acid degradation is normally thiolyzed to acetyl-CoA (3.1.3.3). In liver, a major portion of acetoacetyl-CoA is converted into 3-hydroxy-3-methylglutaryl-CoA, a precursor of cholesterol and steroids (3.5.1.1). Some acetoacetyl-CoA is hydrolyzed, yielding the ketone bodies acetoacetate and its reduction product 2-hydroxybutyrate.

When in mammals the carbohydrate supply is limited (e.g., in starvation), the mobilization of depot lipids and as a consequence the fatty acid degradation increases (3.1.1.5). More ketone bodies are formed by the liver. As a substitute for glucose, these compounds deliver the life-sustaining energy supply especially to the brain. Also muscle cells can use ketone bodies. For utilization they are converted to the CoA-thioesters through a CoA transfer from succinyl-CoA.

Under extreme conditions, but also due to dysregulation in diabetes (7.1.3), the ketone body level in blood increases from normally < 3 mg/100 ml up to 90 mg/100 ml (ketosis). By decarboxylation of acetoacetate, acetone is produced, which cannot be reutilized. Excessive urinary excretion of ketone bodies takes place.

Literature:

Ascherio, A. Atherosclerosis Supplements. 2006;7:25–27.

Brownski, R.W., Danton, R.M. in Boyer, P.D., Krebs, E.G. *The Enzymes.* Vol. XVIII. Academic Press, 1987, p. 123–146.

Eaton, S. *et al.* Biochem. J. 1996;329:345–357.

Graham, I. A. Annual Review of Plant Biology 2008;59:115.

Haunerland, N. H., and Spencer, F. Progress in Lipid Research 2004;43:328–349.

Joshi, A.K. *et al.* Biochemistry 1998;37:2515–2523.

Neidhardt, F.C. *et al.* (Eds.) *Escherichia coli and Salmonella*. ASM Press, 1996.

Rawsthorne, S. Progress in Lipid Research 2002;41:182–196.

Riezman, H. Cell 2007;130:587–588.

Soupene, E. and Kuypers, F. A. Experimental Biology and Medicine 2008;233:507–521.

Tong, L. Cellular and Molecular Life Sciences (CMLS) 2005;62:1784–1803.

Vance, D.E., Vance, J.E. (Eds.): *Biochemistry of Lipids, Lipoproteins and Membranes. New Comprehensive Biochemistry* Vol. 31. Elsevier, 1996.

3.4.2 Triacylglycerols (Triglycerides)

Triacylglycerols (triesters of fatty acids with glycerol) are highly concentrated stores of metabolic energy due to their nearly anhydrous storage form as well as to the high yield of free energy on oxidation (38 kJ/g). This is about three times the energy per gram as compared to glycogen, which is stored in hydrated form. They are also osmotically inert and without influence on the pH.

E.g., an average human male (70 kg) stores about 11 kg of triacylglycerols with an energy content of about 4×10^5 kJ, while glycogen reserves (3.1.2) add less than 1 % to this value. The total muscle protein, which can only partially be turned over, has a total energy content of ca. 25 % of the triacylglycerol value.

Although in mammals most tissues are able to produce triacylglycerols, the major site of their accumulation is the cytoplasm of adipose cells. Both the synthesis of triacylglycerols as well as their hydrolysis to fatty acids serving as fuel molecules for other organs takes place there.

Within all cell types, triacylglycerols are stored as 'lipid droplets' ('fat globules', 'oil bodies', 'adiposomes') enveloped by a monolayer of phospholipids and hydrophobic proteins (e.g., perilipins in adipose tissue or oleosins in seeds).

3.4.2.1 Biosynthesis of Triacylglycerols (Lipogenesis, Fig. 3.4.2-1)

In mammals, triacylglycerols are transported by chylomicrons and VLDL to the cellular membrane of adipose cells and to muscle cells. There they are hydrolyzed by lipoprotein lipase and resorbed as free fatty acids (6.2.2). Other cells mainly extract free fatty acids from blood, which are present there in albumin-bound form. Inside of the cell, the fatty acids are converted into the acyl-CoA thioesters by fatty acid-CoA ligase (3.4.1.4). Glycerol-3-phosphate, the reaction partner for esterification, is provided by reduction of glycerone-P (dihydroxyacetone-P), an intermediate of glycolysis (Fig. 3.1.1-1).

Contrary to adipose tissue, liver, kidney, intestinal mucosa and the lactating mammary gland are additionally able to phosphorylate free glycerol, thus enabling the utilization of this compound (Fig. 3.1.1-1).

The acyl residues are linked to the *sn*-1 position of *sn*-glycerol 3-phosphate by mitochondrial or microsomal glycerol-3-phosphate O-acyltransferases to form acylglycerol 3-P (lysophosphatidate) and then to the *sn*-2 position by microsomal 1-acylglycerol 3-phosphate O-acyltransferase resulting in the formation of 1,2-diacylglycerol 3-P (l-phosphatidate). Usually, the *sn*-1 position is occupied by saturated acyl residues, while at *sn*-2 unsaturated residues predominate (probably due to the substrate specifity of the transferases).

Dephosphorylation catalyzed by phosphatidate phosphatase results in 1,2-diacylglycerols. Phosphatidate phosphatase-1 (PAP1) enzymes play a key role in glycerolipid synthesis by the conversion of phosphatidate to diacylglycerol, the immediate precursor of triacylglycerol, phosphatidylcholine, and phosphatidylethanolamine. PAP1 activity in mammals is determined by the lipin family of proteins (lipin-1, lipin-2 and lipin-3), comprising distinct tissue expression patterns. 1,2-diacylglycerols are also the precursors of phosphatides (3.4.3). They may also arise from 1- or 2-monoacylglycerols, especially in the intestinal mucosa. The consecutive formation of triacylglycerols in various mammalian tissues catalyzed by diacylglycerol O-acyltransferase takes place in the microsomal fraction. They are then stored as triacylglycerol droplets or incorporated into secretory products, lipoproteins and milk, dependent on the tissue.

Regulation of triacylglycerol synthesis occurs in mammals at three steps. The glycerol-3P and diacylglycerol acyltransferases are regulated by hormonally controlled phosphorylation-dephosphorylation mechanisms. In addition, insulin increases the supply of glycerol 3-P.

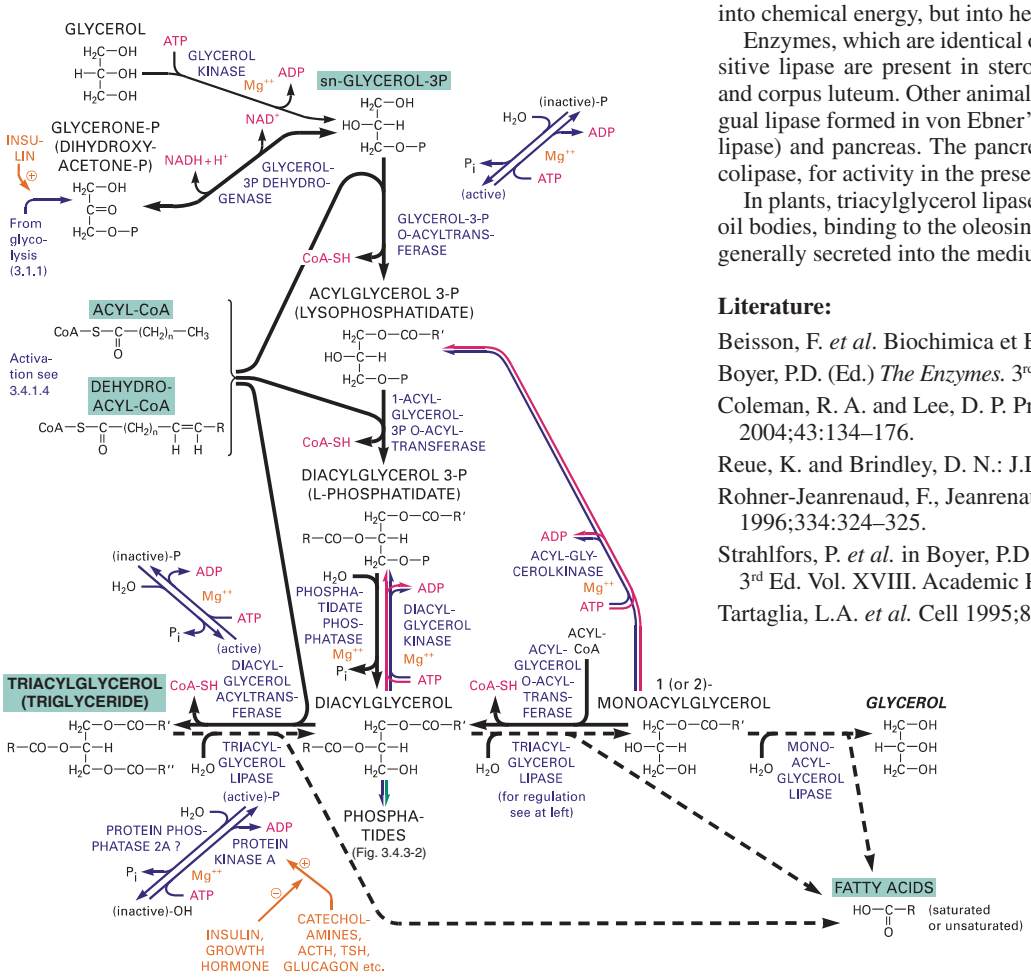


Figure 3.4.2-1. Formation and Degradation of Triacylglycerols
(The regulation by phosphorylation refers to mammalian enzymes).

During lipogenesis, adipose tissue secretes the peptide hormone leptin (*ob* gene product). It binds to a receptor in the hypothalamus. This diminishes the release of the neuropeptide Y (which expresses the feeling of hunger, releases insulin and glucocorticoids) and thus controls the uptake of food. In obesity of humans, apparently the receptor is defective.

3.4.2.2 Mobilization of Triacylglycerols (Lipolysis, Fig. 3.4.2-1)

The mobilization of triacylglycerols by hydrolysis to yield free fatty acids is catalyzed by lipases. In mammals, a hormone-sensitive lipase exists in adipose tissue which is interconverted by a phosphorylation-dephosphorylation cycle between an active (lipase A) and an inactive (lipase B) form. The phosphorylation is performed by protein kinase A, which is activated by catecholamines, glucagon and other hormones via cAMP (7.4.2). Insulin and growth hormone counteract this effect, possibly by activating the cAMP degrading phosphodiesterase (Fig. 7.4-2). This lipase is responsible for hydrolysis of the esters in position 1 and 3 of triacylglycerol. It also hydrolyzes cholesterol esters (3.5.1, 6.2.4). A second enzyme, 2-monoacylglycerol lipase, catalyzes the hydrolysis of the remaining ester to yield free fatty acids (which are released to supply other organs with fuel) and glycerol (which is shuttled back to the liver, phosphorylated and oxidized to glycerone-P = dihydroxyacetone-P for use in oxidation or gluconeogenesis).

The hormone-sensitive lipase also occurs in brown adipose tissue, which is rich in mitochondria. In these cells, the resulting free fatty acids are not released, but undergo β-oxidation (3.4.1.5). The fatty acids also bind to and activate an uncoupling protein in the mitochondrial inner membrane. This protein 'uncouples' the respiratory chain and the ATP synthesis, since it enables the return of protons to the mitochondrial matrix, bypassing the ATP synthase reaction. The energy of the electrochemical potential gradient is thus not converted into chemical energy, but into heat (thermogenesis, 3.11.4.5).

Enzymes, which are identical or closely related to the hormone sensitive lipase are present in steroid producing cells, e.g., in adrenals and corpus luteum. Other animal lipases are present in the mouth (lingual lipase formed in von Ebner's gland), in milk (bile salt-stimulated lipase) and pancreas. The pancreatic lipase needs a protein cofactor, colipase, for activity in the presence of bile salts.

In plants, triacylglycerol lipase appears to associate transiently with oil bodies, binding to the oleosin protein coating. Bacterial lipases are generally secreted into the medium.

Literature:

Beisson, F. *et al*. Biochimica et Biophysica Acta 2001;1531:47–58.

Boyer, P.D. (Ed.) *The Enzymes.* 3rd Ed. Vol. XVI. Academic Press, 1983.

Coleman, R. A. and Lee, D. P. Progress in Lipid Research 2004;43:134–176.

Reue, K. and Brindley, D. N.: J.Lipid Res. 2008;49:2493–2503.

Rohner-Jeanrenaud, F., Jeanrenaud, B. New Engl. J. Med. 1996;334:324–325.

Strahlfors, P. *et al.* in Boyer, P.D., Krebs, E.G. (Eds.): *The Enzymes.* 3rd Ed. Vol. XVIII. Academic Press, 1987, p. 147–177.

Tartaglia, L.A. *et al.* Cell 1995;83:1263–1271.

3.4.3 Phospholipids

Phospholipids are important building blocks of cell membranes. They are derived from phosphate esters of glycerol (glycerophospholipids), glycerol ethers (plasmalogens) or sphingosine (sphingophospholipids), as shown in Figure 3.4.3-1. Their basic structures and their chemistry are dealt with in 1.4.4 … 1.4.6.

They show an amphipathic behavior, which is caused by the presence of

- a hydrophilic 'head group' (esterified with the phosphate residue) and
- a hydrophobic 'lipid tail' (long hydrocarbon chains of fatty acids, ethers or sphingosine).

3.4.3.1 Occurrence of Phospholipids

Whereas *E. coli* possesses a very simple phospholipid composition, consisting primarily of phosphatidylethanolamine, phosphatidylglycerol and cardiolipin, the phospholipid composition of eukaryotic membranes is more complex and varies in different tissues (Table 3.4.3-1).

In addition to phospholipids, cholesterol (3.5.1.3) and many glycolipids (3.4.4) including glycoglycerolipids are present in eukaryotic membranes. The contribution of the various components to the membrane properties is discussed in 2.2.4.

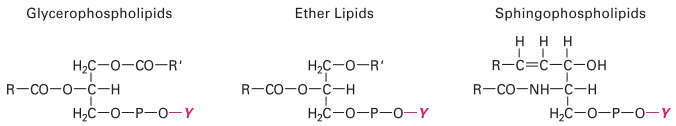


Figure 3.4.3-1. Basic Structures of Phospholipids (Y = Head Group)

3.4.3.2 Glycerophospholipids (Phosphoglycerides)

Biosynthesis (Fig. 3.4.3-2): For the phospholipid-producing condensation reaction, either the glyceryl moiety or the head group has to be activated. Bacteria use the first way, while animals and plants are able to use both pathways.

Bacteria: In *E. coli*, diacylglycerol 3-phosphate (L-phosphatidate, 3.4.2.1) is converted to a mixture of CDP-1,2-diacylglycerol and dCDP-1,2-diacylglycerol by a single enzyme, phosphatidate cytidyltransferase (CDP-diacylglycerol synthase). This compound then reacts with either serine or glycerol-3-P to form phosphatidylserine (PS) or 3-phosphatidyl-glycerol-1-P, respectively.

PS is only a minor membrane constituent of *E. coli*, since it is rapidly decarboxylated by PS decarboxylase to yield phosphatidylethanolamine. PS decarboxylase (36 kDa) uses a pyruvoyl moiety as the prosthetic group. In this way, the ethanolamine moiety is produced *de novo*. 3-Phosphatidylglycerol-1-P, on the other hand, is

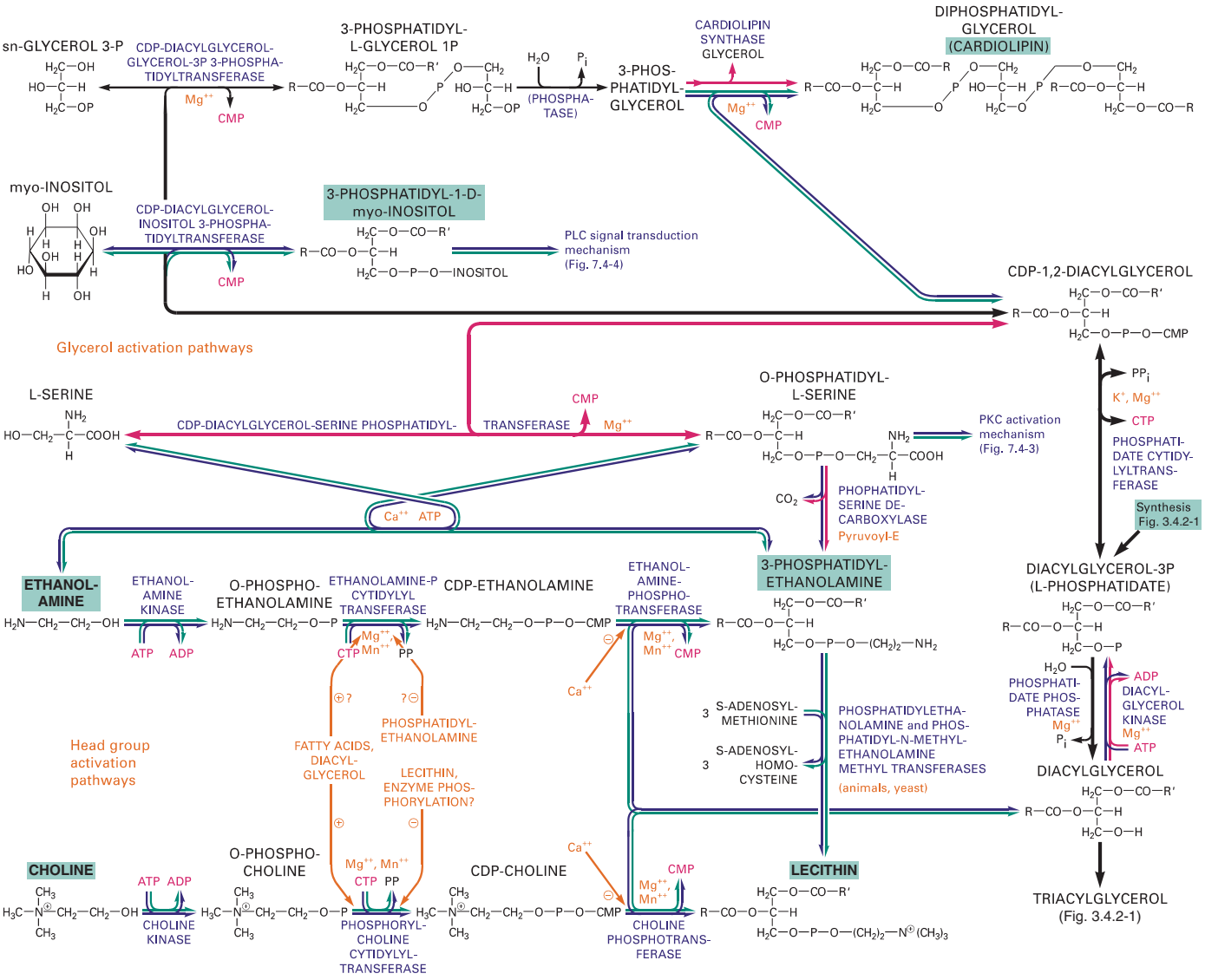


Figure 3.4.3-2. Biosynthesis of Glycerophospholipids

Table 3.4.3-1. Distribution of Phospholipids (PL) in Various Species (Harwood, 1980)

Phospholipid	'Head group' (Fig. 3.4.3-1) Y =	% of Total PL in Animals			% of Total PL in Plants		% of Total PL in Bacteria (ca.)
		Bovine Liver	Bovine Brain	Bovine Heart	Mitochondrial Inner Membrane	Plasma Membrane	
Glycerophospholipids:							
Phosphatidate	H	2	1	2	0	0	0
Phosphatidylcholine (lecithin, PC)	choline	54	29	24	27	32	0
Phosphatidylethanolamine (PE)	ethanolamine	9	12	17	29	46	75
Phosphatidylinositol (PI)	inositol	8	6	4	0	19	0
Phosphatidylserine (PS)	serine	4	17	2	25	0	trace
3-Phosphatidylglycerol (PG)	glycerol	trace	trace	trace	0	0	15 … 20
Cardiolipin (diphosphatidylglycerol)	phosphatidylglycerol	4	1	9	20	0	5 … 10
Ether lipids:							
Choline plasmalogen	choline	2	trace	18			
Ethanolamine plasmalogen	ethanolamine	4	21	11			
Sphingophospholipids:							
Sphingomyelin	choline	6	13	12			
Sphingoethanolamine	ethanolamine	6	13	12			

dephosphorylated to 3-phosphatidylglycerol (PG). At least two different phosphatases exist which may catalyze this reaction. By condensation of two PG molecules and release of glycerol, cardiolipin (1,3-bis(sn-3′-phosphatidyl)-sn-glycerol) is obtained..

Animals, plants and yeast: L-phosphatidate is hydrolyzed to 1,2-diacylglycerol (DG), catalyzed by two phosphatidate phosphatases (a cytoplasmic form bound to the endoplasmatic reticulum, and an enzyme bound to the plasma membrane). DG can react with CDP-ethanolamine (formed from ethanolamine by phosphorylation, followed by transfer of a CMP moiety from CTP) to yield phosphatidylethanolamine (PE). However, PE can also be formed by decarboxylation of phosphatidylserine, by an exchange reaction of ethanolamine with PS, or by re-acylation of lyso-PE.

Diacylglycerol can also react with CDP-choline to produce phosphatidylcholine (lecithin, PC) which, especially in liver, can also be formed from PE by methylation. The regulated step for PC (and likely) PE synthesis is the cytidylyl transferase reaction. The active enzyme is reversibly removed from the membrane and in this way inactivated by the final product PC or by lack of the substrate DG.

In eukarya, phosphatidylserine (PS) is not formed through the pathway of glycerol activation (except in yeast), but by a base-exchange reaction from PE or PC. The catalyzing phosphatidylserine synthase requires Ca^{2+} and ATP and is located in the microsomes. Together with the decarboxylation step of PS this accounts for the net synthesis of ethanolamine or choline:

PC + serine = PS + choline PE + serine = PS + ethanolamine

PS = PE + CO_2 PS = PE + CO_2

PE + 3 SAM = PC + 3 SAH

(SAM = S-adenosylmethionine,
SAH = S-adenosylhomocysteine)

Alternatively to the conversion into diacylglycerol (3.4.2.1), phosphatidate in mammals (as in *E. coli*) reacts with CTP to yield CDP-1,2-diacylglycerol. This activated compound then condenses with inositol to yield phosphatidylinositol or with glycerol-3-P, resulting in phosphatidylglycerol. The consecutive reactions leading to cardiolipin proceed differently to bacteria and employ the condensation with another molecule of CDP-diacylglycerol. Cardiolipin is a major phospholipid in metabolically active cells of heart and skeletal muscle.

Remodelling of fatty acid substituents: Once a phospholipid is synthesized, the fatty acid substituents in the *sn*-1 or *sn*-2 positions can be remodelled via deacylation-reacylation reactions catalyzed by phospholipases A_1 and A_2, acyl-CoA: lysophospholipid 1-(or 2-) acyltransferases, and CoA dependent and independent transacylation systems. Their substrate specificity regarding acyl chain length and saturation is of crucial importance for the fatty acyl pattern of phospholipids.

Degradation of glycerophospholipids (Fig. 3.4.3-3): The hydrolysis of phospholipids is catalyzed by phospholipases (PL) which are classified according to their positional specificity (Fig. 3.4.3-3 and Table 3.4.3-2). Lysophospholipids, the products of phospholipase A_1 and A_2 reactions, are further hydrolyzed by lysophospholipases.

Phospholipases are abundant enzymes which have been found in animals, bacteria, and plants. Their activity is restricted to the interface between the aqueous and lipid phases so that usual enzyme kinetics do not apply. In plants, phospholipase D (PLD) is most abundant, and several PLD genes exist (e.g., *A. thaliana* 12, *Oryza sativa /* rice 17). They are classified based on their lipid-binding domains.

Metabolic role of glycerophospholipids and phospholipases: The role of phospholipids in eukaryotic membranes is not just a structural one. They take part in many metabolic processes.

Phosphatidylinositol (PI) is a key component in the PLC-PKC signal transduction mechanism (7.4.3). After two phosphorylations, it is cleaved upon arrival of hormonal signals. Both cleavage products act as second messengers: inositolphosphates release calcium, diacylglycerol is an activator of protein kinase C. Phosphatidylserine is involved in this reaction. Eukaryotic PI-PLCs are classified into β, γ, δ, ε, η, and ζ subfamilies. While mammalian cells contain all six isoforms, plants comprise only one class of PLCs (PLCζ).

Release of arachidonic acid from phosphatidylinositol by PLA_2 and other phospholipases, e.g., during an inflammatory process leads to the production of prostaglandins and leukotrienes (7.4.8).

The interaction of Ca^{++} with phospholipids is of importance for membrane fusion processes. Phosphatidylinositol functions also as a membrane anchor for proteins (4.4.1.4).

With respect to foreign cells, phospholipases can be mediators of infectivity (secreted bacterial phospholipases) or act as toxins on membranes (e.g., in snake venoms).

3.4.3.3 Ether Lipids

Ether lipids of various configuration occur in membranes of higher and lower animals, of anaerobic prokarya (mostly archaea) and of some higher plants (as minor components). A high concentration of ether lipids is found in tumors. These compounds apparently decrease the permeability and lower the melting temperature of membranes. The platelet activating factor (PAF, see below) is involved in many regulatory functions.

Table 3.4.3-2 Phospholipases

Enzyme	Source	Properties, Mol. Mass
Phospholipase A_1 (PL A_1)	bacteria mammals	mostly membrane-bound in lysosomes, various forms known
Phospholipase A_2 (PL A_2)	snake venoms (type I): cobras and kraits mammalian pancreas (type I) snake venoms (type II): rattlesnakes and vipers mammalian tissues and synovial fluid (type II) Bee, lizard (type III) mammalian tissues (type IV)	secreted, ca. 13–15 kDa secreted, ca. 13–15 kDa secreted, ca. 13–15 kDa, (see 17.4.8) secreted, ca. 13–15 kDa secreted, 16–18 kDa cytoplasmic; ca. 85 kDa; regulated by Ca^{++} concentration (see 7.4.8)
	mammals (heart muscle)	cytoplasmic; ca. 40 kDa; no requirement for Ca^{++}
	mammals (PAF acetyl hydrolase)	blood, ca. 60 kDa; no requirement for Ca^{++}
Lysophospholipase (LPL)	microorganisms, bee venom, mammalian tissues	variable in size and pH optimum
Phospholipase C (PL C)	bacteria (e.g., *Bacillus cereus*), mammals	PC- and PI-specific enzymes known, many isoenzymes, role in signal transduction (17.4.3)
Phospholipase D (PL D)	plants, microorganisms, mammals	functions as transphosphatidylase with various alcohol acceptors and in signal transduction

Some fucose derivatives have blood group specificity (4.4.3.1).

Biosynthesis (Fig. 3.4.3-4): The initial reaction is an unusual exchange reaction of a fatty alcohol with the acyl group of 1-acyldihydroxyacetone-P. Some details are still unknown. Then a NADPH-dependent reduction of the oxo group at C-2, acylation and dephosphorylation follow, resulting in alkylacylglycerol. Analogous to the reaction with diacylglycerol, CDP-ethanolamine and CDP-choline add the respective residues to yield plasmanylethanolamine and plasmanylcholine. A 1-2 double bond can be introduced into plasmanylethanolamine by action of the plasmanylethanolamine desaturase, resulting in plasmalogen.

The exchange of the long fatty acid (frequently arachidonate) esterified at the C-2 site of plasmanylcholine with an acetyl residue yields the platelet activating factor (PAF). This reaction takes place mainly in leukocytes and endothelial cells and is induced by thrombin, antigens, bradykinin, ATP etc. PAF is involved in the pathogenesis of asthma, hypertension, allergies and hypersensitivity, inflammation etc. and acts on specific receptors in a concentration of 10^{-10} mol/l (Table 3.4.3-3). The further mechanism in unknown. The compound is inactivated by removal of the acetyl (by a group of enzymes called PAF acetylhydrolases) or of the phosphocholine residue. Also a hydroxylation at the α-C atom of the long-chain alcohol, followed by hydrolytic removal of the formed aldehyde moiety takes place. This reaction is catalyzed by glyceryl-ether monooxygenase, a tetrahydrobiopterin dependent enzyme.

Table 3.4.3-3 Selected Cellular Functions Influenced by the Platelet Activating Factor (PAF). See 8.2, 9.4.

Bronchoconstriction ↑	Aggregation of neutrophils and platelets ↑
Pulmonary edema ↑	Degranulation of platelets, mast cells etc. ↑
Pulmonary blood pressure ↑	Chemotaxis of neutrophils ↑
But: systemic blood pressure ↓	Activity of
Heart rate ↑	– Protein kinase C ↑
Hypersensitivity response ↑	– G-protein receptor kinases ↑
Vascular permeability ↑	– Protein-tyrosine kinase ↑
Phosphoinositide turnover ↑	– PL-C ↑, leads to Ca^{++} uptake ↑

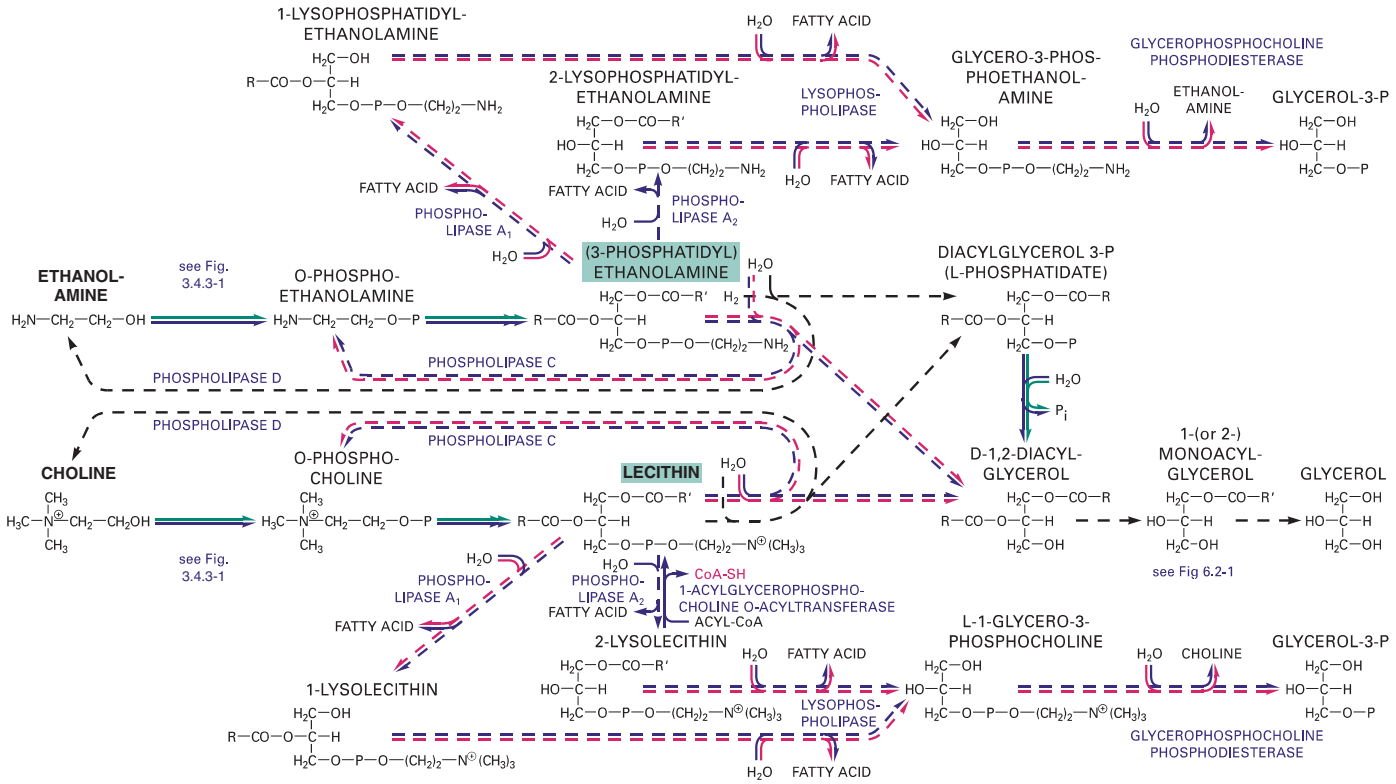


Figure 3.4.3-3. Degradation of Glycerophospholipids

3.4.3.4 Sphingophospholipids

While a large number of glycosphingolipids exists in various organisms (3.4.4.1), the only major sphingophospholipid is sphingomyelin in higher animals. It is present in membranes, lung surfactant, lipoproteins (especially LDL, 6.2.2) and atherosclerotic plaques.

Biosynthesis (Fig. 3.4.3-4): The biosynthesis up to ceramide is common for both glycosphingolipids and sphingophospholipids and takes place in the endoplasmic reticulum. It starts with the condensation of palmitoyl-CoA and serine to give 3-dehydrosphinganine (3-ketosphinganine) in a pyridoxal-P dependent reaction (β-replacement, 3.7.4.2). The product is then reduced to sphinganine, acylated, and desaturated to yield ceramide. This compound is also the precursor of glycosphingolipids (3.4.4.1).

Sphingomyelin is the most abundant nuclear sphingolipid and occurs in the nuclear envelope and intranuclear sites. It is synthesized in the Golgi apparatus of liver and in plasma membranes of the nervous system (oligodendrites and myelin membranes) by transfer of phosphorylcholine from phosphatidylcholine (lecithin) to ceramide, liberating diacylglycerol. Degradation begins with the removal of the head group (as choline or phosphocholine). This step is regulated by, e.g., tumor necrosis factor (TNF-R1), interferon-γ, nerve growth factor (NGF) etc. Depending on the location of this process (lysosomal or associated to the cellular membrane), the ceramide formed initiates signal cascades (7.5.3) leading to proliferation or to apoptosis. Many details are still unknown.

Niemann-Pick patients have a deficiency of the lysosomal sphingomyelin phosphodiesterase. Further degradation steps involve deacylation to sphingosine (after which the whole group of compounds has been named) and its cleavage.

Literature:

Harwood, J.L., Sanchez, J. Progr. in Lipid Res. 1994;33:1–202.

Hornick, C. A. *Medical Biochemistry* (Fourth Edition). San Diego:Academic Press, 2002, p. 401–427.

Kent, C. Ann. Rev. Biochem. 1995;64:315–343.

Ledeen, R. W., Wu, G. J. Lipid Res. 2008;49:1176–1186.

Munnik, T. and Testerink, C. J. Lipid Res. 2009:50:S260-S265. Epub.2008:Dec.20.

Snyder, J. Biochim. Biophys. Acta 1995;1254:231–249.

Taguchi, H., Armarego, W. Med. Res. Rev. 1998;18:43–89.

Testi, R. Trends in Biochem. Sci. 1996;21:468–471.

Vance, D.E., Vance, J.E. (Eds.) *Biochemistry of Lipids, Lipoproteins and Membranes*, New Comprehensive Biochemistry Vol. 31, Elsevier, 1996.

3.4.3.5 Choline, Betaine, Sarcosine (Fig. 3.4.3-5)

Choline metabolism: In plants and to a limited extent in animals, phosphatidylcholine is synthesized from phosphatidylethanolamine by methylation. Its hydrolysis yields choline (3.4.3.2). Many microorganisms (e.g., *E. coli*) cannot synthesize choline.

In addition to its role as a component of membrane phospholipids (3.4.3.2 … 3.4.3.4), choline plays an additional important role in the form of acetylcholine. This is an important neurotransmitter (7.2.3, Table 7.2-2). It is synthesized in the cells of the cholinergic system by choline acetyltransferase and is stored in presynaptic vesicles. Upon neuronal excitation, acetylcholine is released into the synaptic cleft, acts on postsynaptic membrane receptors and is inactivated within milliseconds by acetylcholine esterase, which is present in neighborhood of the receptors. Choline is thereafter taken up by the presynaptic neurons and recycled into acetylcholine.

Trimethylammonium compounds (e.g., physostigmin) inhibit the enzyme reversibly and thus enhance temporarily the excitatory acetylcholine effects, while most organophosphorus compounds (insecticides, e.g., E605, nerve gases, e.g., sarin, tabun) act irreversibly and cause cardiac arrest and respiratory paralysis. Acetylcholinesterase inhibitors (e.g., physostigmine, tacrine, donepezil) are being tested for treatment of Alzheimer's disease.

Betaine metabolism: Betaine (= glycine betaine) is an atypical amino acid, since it contains a quaternary amino group. It is synthesized from choline. Betaine is nonproteinogenic, but is important as an osmoprotectant. Analogous derivatives of other amino acids are also called 'betaines'.

Biosynthesis: In animals, betaine is synthesized by a two-step oxidation of choline. This takes place in the mitochondrial matrix by the membrane-bound choline dehydrogenase (CDH, which is coupled to the respiratory chain, 3.11.4) and the soluble betaine aldehyde dehydrogenase

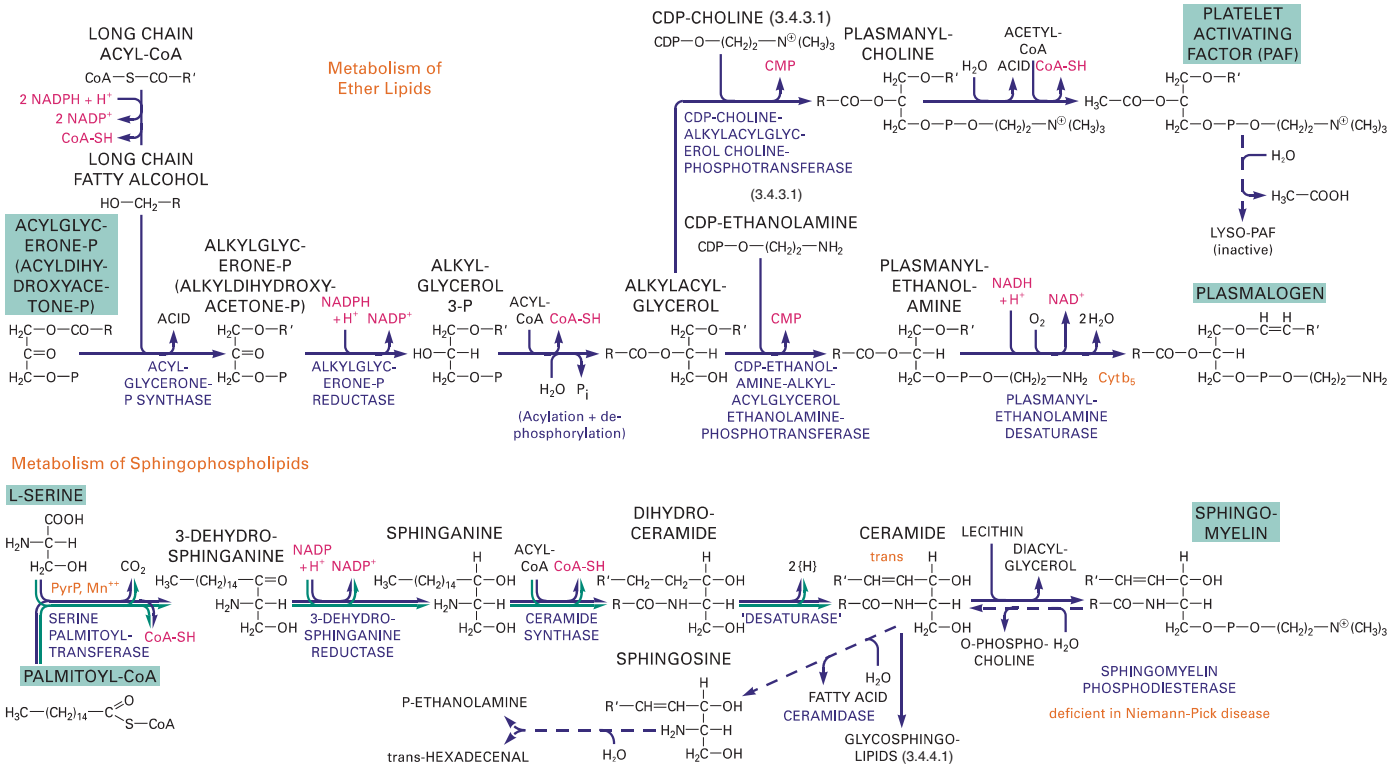


Figure 3.4.3-4. Metabolism of Ether Lipids and Sphingophospholipids

(BADH). In the chloroplast stroma of spinach, the first oxidation step is catalyzed by the soluble ferredoxin dependent choline monooxygenase. In microorganisms, the first step is performed by CDH or via a soluble, H_2O_2 generating oxidase. Both enzymes are flavoproteins.

Osmoprotection: In plants and microorganisms, environmental changes such as salinity, cold or drought lead to osmotic stress. Many organisms can respond with the uptake or biosynthesis of osmoprotectants. These are substances with no adverse effects on the cell in concentrations of up to ca. 1 mol/l, which maintain the turgor of the cell in equilibrium with the environment (compatible compounds). Glycine betaine and proline betaine are two major organic osmoprotectants that accumulate in various plant species in response to environmental stresses (drought, salinity, extreme temperatures, UV radiation, heavy metals), Others are proline, sucrose, polyols, trehalose, alanine betaine, choline O-sulfate, hydroxyproline betaine, and pipecolate betaine. In animals, osmoprotection of the kidney tubules against the increasing concentration of urine is of special importance.

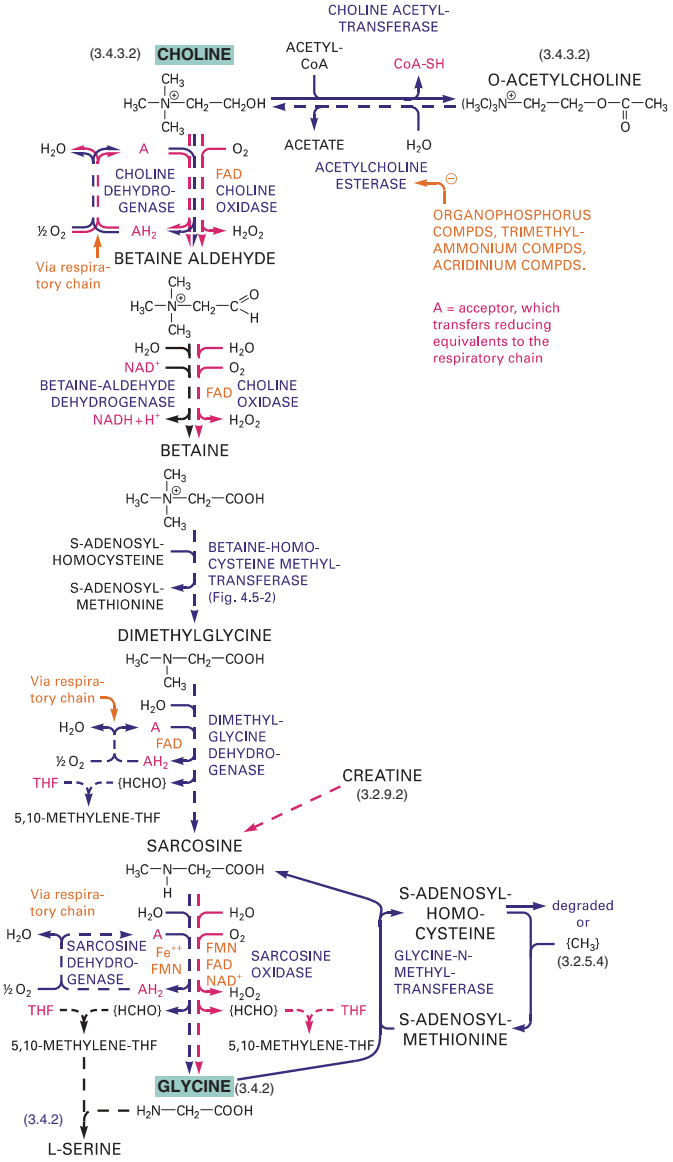


Figure 3.4.3-5. Metabolism of Choline, Betaine and Sarcosine

Sarcosine metabolism: Betaine degradation leads to sarcosine and is apparently restricted to animals (mainly in liver and kidney), except for some microorganisms.

The first step involves a methyl transfer to homocysteine. This is the only way of methionine synthesis (3.2.5.4) in the absence of methyltetrahydrofolate (3.7.6.2) and coenzyme B_{12} (3.7.5.2) and thus of major metabolic importance. The following two oxidative demethylation steps proceed via sarcosine (N-methylglycine) to glycine (3.2.4.2) and are catalyzed in the mitochondrial matrix by dimethylglycine dehydrogenase and sarcosine dehydrogenase, which are connected to the respiratory chain. The methyl groups of betaine are oxidized and bound as methylene residues to tetrahydrofolate. Methylenetetrahydrofolate, in turn, may contribute to the conversion of glycine to serine by glycine hydroxymethyltransferase (3.2.4.2).

Some bacteria can obtain sarcosine by degradation of creatine (3.2.9.2) and oxidize it by sarcosine oxidase to glycine, formaldehyde and H_2O_2. Other anaerobic bacteria reduce betaine, sarcosine and glycine via reductases containing selenocysteine (4.1.3.4) to trimethylamine, methylamine or ammonia, respectively. These reactions are coupled to energy conservation.

Sarcosine synthesis from glycine: Glycine N-methyltransferase (GNMT), an enzyme found in animals (liver, kidney and exocrine cells), catalyzes the S-adenosylmethione (SAM) dependent methylation (3.2.5.4) of glycine to sarcosine. This is the only example of an 'unimportant' methylation (no key metabolite synthesis or regulatory effect on the target takes place). The demethylation of SAM, however, enables the use of its sulfur for the synthesis of, e.g., acidic glycosphingolipids (sulfatides, 3.4.4.1) or glycosaminoglycan sulfates (2.9.2), and of the C-backbone for gluconeogenesis (3.1.1.3).

Literature:

Ashraf, M. and Foolad, M. R. Environmental and Experimental Botany 2007;59:206–216.

Burg, M.B. Am. J. Physiol. 1995;268:F983–F996.

Doraiswamy, P. M. CNS. Drugs 2002;16:811–824.

Meyer, M. *et al.* Eur. J. Biochem. 1995;234:184–191.

Reuber, B.E. *et al.* J. Biol. Chem. 1997;272:6766–6776.

Trossat, C. *et al.* Plant Physiol. 1997;113:1457–1461.

Yeo, E.J., Wagner, C. Proc. Natl. Acad. Sci. USA 1994;91:210–214.

3.4.4 Glycolipids

Similar to many glycosylated proteins (2.9), glycolipids are members of the glycocalyx layer covering the cellular membrane. The lipid part forms the anchor in the membrane (Fig. 3.4.4-1), to which the (frequently antigenically reactive) glycosidic chain is bound. Three different types are distinguished (Table 3.4.4-1).

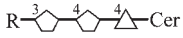
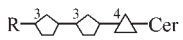
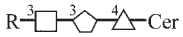
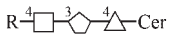
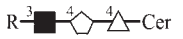
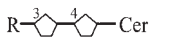
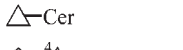
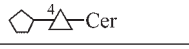
3.4.4.1 Glycosphingolipids

Glycosphingolipids occur in vertebrates and lower animals, as minor components in plants and (only rarely) in bacteria. Many diverse structures exist. They are species- and organ- specific. Some of them vary with age. They are classified by the type of the three sugars next to the lipid moiety ('root structure', Table 3.4.4-2). Frequently, they carry acidic groups (sulfate, uronic acids, sialic acid, etc.). In membranes they tend to aggregate.

Table 3.4.4-1. Types of Glycolipids

	Glycosphingolipids	Glycoglycerolipids	Glycosylphosphopolyprenols
Membrane anchor (lipid)	ceramide or other sphingosine derivatives	mono- and diacylglycerol, glycerol ethers and phosphatidic acid	isoprenoids, e.g., dolichol-P, ficaprenolundecaprenol-PP
Lipid-sugar bond	O-glycosidic	O-glycosidic or phosphate ester	phosphate ester
Function	membrane constituent	membrane constituent	precursor of glycosylated proteins
Site of synthesis	Golgi apparatus, bacterial membranes	Golgi apparatus, chloroplasts, bacterial membranes	endoplasmic reticulum, bacterial membranes

Table 3.4.4-2. Series of Glycosphingolipids (for symbols, see Fig. 3.4.4-2)

Series	Abbr.	Root Structure	Present in	Examples	Derivatives
Globo	Gb		mammalian extraneural tissue (major glycolipid)	R= ■—: main component of *human* erythrocyte membrane (= globoside), R = ■—■—: Forssman antigen (in some *mammals*, also in *bacteria*)	Gangliosides
Isoglobo	iGb				
Lacto	Lc		erythrocytes (non-human)	R = ⬠—: precursor of ABO blood groups and Le antigens (not in *humans*)	gangliosides, sulfate derivatives
Neolacto	nLc		erythrocytes (human, animals), spleen	R = ⬠—: precursor of ABO blood groups and Le^x antigens (*humans* and *others*). R = ⬠⬠—: P antigen	gangliosides, sulfate derivatives
Ganglio	Gg		mostly in brain gangliosides, also extraneural		gangliosides, sulfate derivatives
Gala	Ga		animals, fungi	in *mammals* only R = H- or ■—■—	esters (C_{16}, C_{18})
Arthro	Ar		arthropods	R = □— or ■—⬠	
Mollu	Ml		mollusks, plants	R = □— or –xylose–	
(Glucosylceramide)			animals, plants, fungi	intermediates of synthesis for globo, lacto, ganglio series	esters
(Lactosylceramide)			animals, plants, fungi		esters

Some fucose derivatives have blood group specificity (4.4.3.1).

Gangliosides are glycosphingolipids, which also contain sialic acid (N-acetylneuraminic acid or its 9-O acetyl derivative in humans, additional derivatives in other mammals, fish, birds, etc.). They are found primarily in the central nervous system, but also in other organs of vertebrates at the outer side of the plasma membrane (Table 3.4.4-3).

Table 3.4.4-3. Gangliosides in Mammals

Series	Occurrence, e.g., in	Examples
Ganglio	brain grey matter	$sialyl_{1...4}$ gangliotetraosyl ceramide
	brain white matter	$sialyl_1$ gangliotetraosyl ceramide
	erythrocytes (rat)	$sialyl_2$ gangliotetraosyl ceramide
Neolacto	erythrocytes (man, dog, pig)	$sialyl_{1...2}$ lactosyl ceramide
	peripheral nerves	$sialyl_{1...2}$ lactosyl ceramide
		$sialyl_{1...2}$ lactotetraose ceramide

They mediate interactions between cells and with the extracellular matrix, cell-cell communication, immunological functions, leukocyte differentiation, oncogenic reactions, etc. During embryogenesis and thereafter, the ganglioside pattern changes. Gangliosides (and other glycosphingolipids) influence cell differentiation and the development of the neuronal network. Gangliosides of other series occur in lower animals.

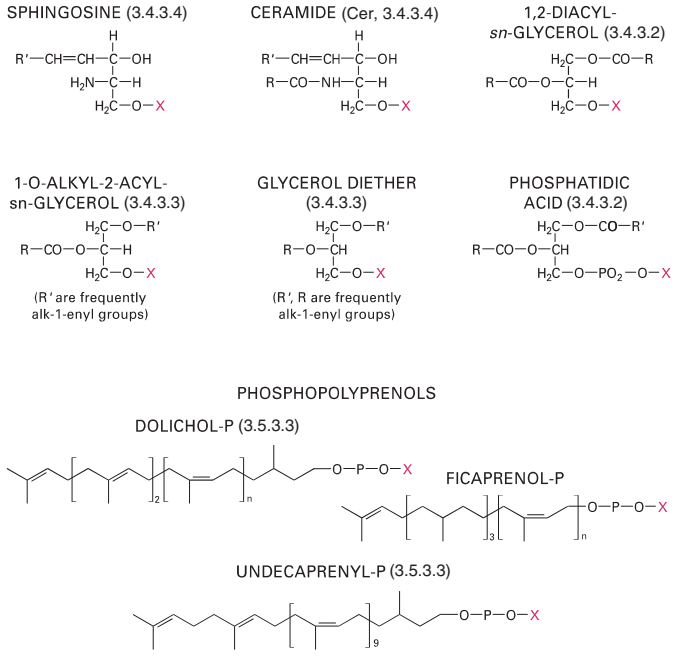


Figure 3.4.4-1. Structure of Common Membrane Anchors
(The carbohydrate chains are bound at -X)

Biosynthesis: Ceramide is synthesized in the endoplasmic reticulum (3.4.3.4). The primary glycosylation steps, which take place mainly in the Golgi apparatus, are described in 4.4.2. They are analogous to the biosynthesis of glycosylated proteins by ligand transfer from XDP-sugars, XDP-aminosugars or CMP-sialic acid to ceramide. Higher gangliosides (Fig. 3.4.4-2) are formed by addition of more sialyl and glycoside groups to the basic structures shown in Fig. 3.4.4-1. The nomenclature system of Svennerholm is used.

Degradation: The lysosomes are the site of glycolipid degradation. Exoglycosidases sequentially remove single sugar moieties from the end. Special proteins facilitate interaction of the enzymes with their lipid substrate. Disturbances of degradation are most striking with gangliosides (Fig. 3.4.4-3).

Medical aspects: Deficiencies in the degradation process lead to various inherited diseases, which are manifested by severe neurodegenerative defects. They are characterized by accumulation of those glycolipids, which are substrates of the defective enzyme. Tay-Sachs and Sandhoff diseases are two forms where a deficiency in the enzyme β-hexosaminidase leads to the accumulation of ganglioside G_{M2}. In Tay-Sachs, the α subunit of this enzyme and in Sandhoff the β subunit is affected. Some other diseases are caused by a defect in activator proteins for β-hexosaminidase. In Gaucher's disease, the terminal glucose residue cannot be removed due to a defect in β-glucosidase. Fabry's disease originates in a defect in α-galactosidase A, while deficiency of β-galactosidase leads to G_{M1}gangliosidosis. General decrease of gangliosides is observed in Creutzfeldt-Jacobs and Bovine Spongiform Encephalopathy (BSE) that are caused by prions.

3.4.4.2 Glycoglycerolipids (Table 3.4.4-4)

Glycoglycerolipids occur in all kingdoms of living beings. Their structure is species specific. Contrary to glycosphingolipids, they play a major role in bacteria (e.g., polymeric structures are members of the bacterial cell wall, 3.10.1). In plants, they are specifically enriched in the thylakoid membrane of chloroplasts, while in mammals, they are major constituents of the myelin sheath of neurons and occur in mucous layers and in germ cells.

Biosynthesis and biodegradation: The lipophilic part is synthesized in mammals and in plants at the membrane of the endoplasmic reticulum. See 3.4.2.1 for glycerol esters and 3.4.3.3 for glycerol ethers.

With glycolipid-anchored proteins in animals, the further steps take place in the endoplasmic reticulum (4.4.1). However, the transfer of

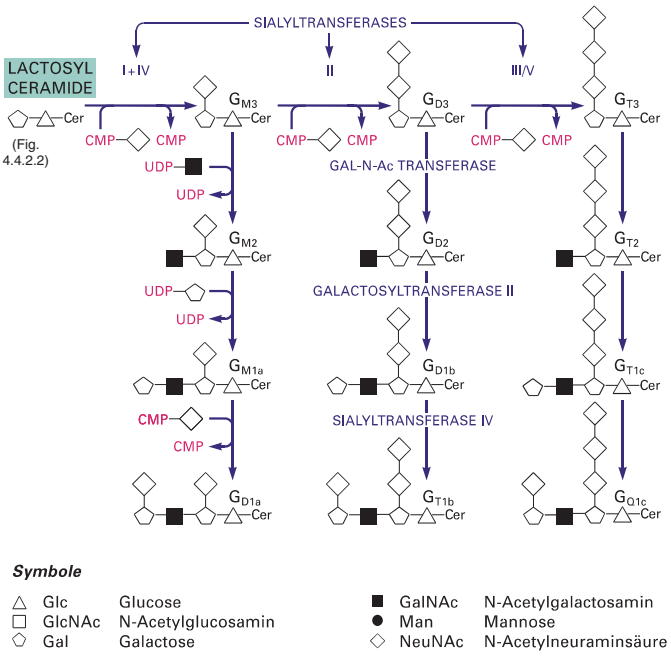


Figure 3.4.4-2. Synthesis of Higher Gangliosides (Examples only, many variants exist. For basic structures, see Fig.4.4.2-2)

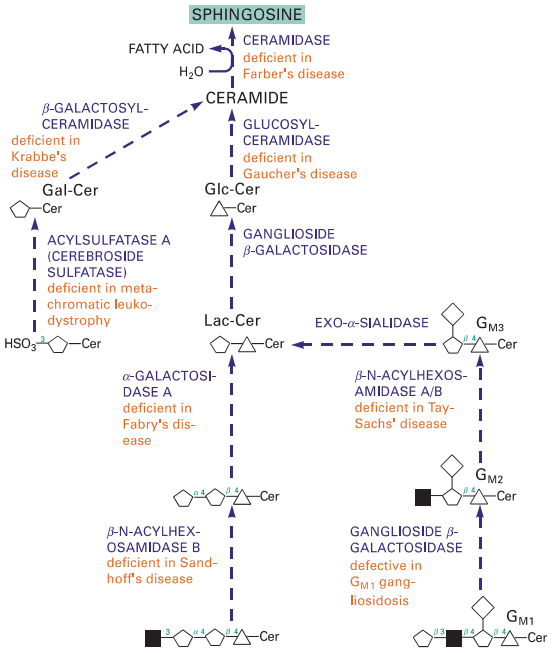


Figure 3.4.4-3. Degradation of Glycosphingolipids

sugars and sulfate groups in animals usually occurs in the Golgi apparatus (4.4.2), while in plants, this is performed preferably in the chloroplast envelope. The assembled compounds are later integrated into the thylakoid membranes. In bacteria, both biosynthesis and sugar transfer are located in the cell membrane. The reactions follow the same principles as with glycosphingolipids (3.4.4.1), using activated XDP-sugars. For bacterial macromolecules, see 3.10.1.

Biodegradation in mammals takes place by a stepwise removal of the terminal residues by specific lysosomal enzymes. Lack of sulfatase causes metachromatic leukodystrophia.

3.4.4.3 Glycosylphosphopolyprenols

These compounds function as carriers of glycosyl groups. The synthesis of phosphorylated proteins in the endoplasmic reticulum of eukarya uses dolichol-P. Undecaprenyl-P functions as a carrier in the formation of bacterial cell wall murein as well as in biosynthesis of the O-antigen part of lipopolysaccharide (Gram negative bacteria) and of teichoic acids (Gram positive bacteria). Both polyprenols are membrane anchors for a growing glycosidic chain. They have *tri-trans* or *di-trans*, *poly-cis* configuration. Their biosynthesis is described in 3.5.3, while their function as carriers of glycosidic chains is dealt with in 4.4.1 and 3.10.1.

Literature:

Bell, R.M., *et al.* (Eds.) *Advances in Lipid Research* Vols. 25 and 26. *Sphingolipids*. Academic Press, 1993.

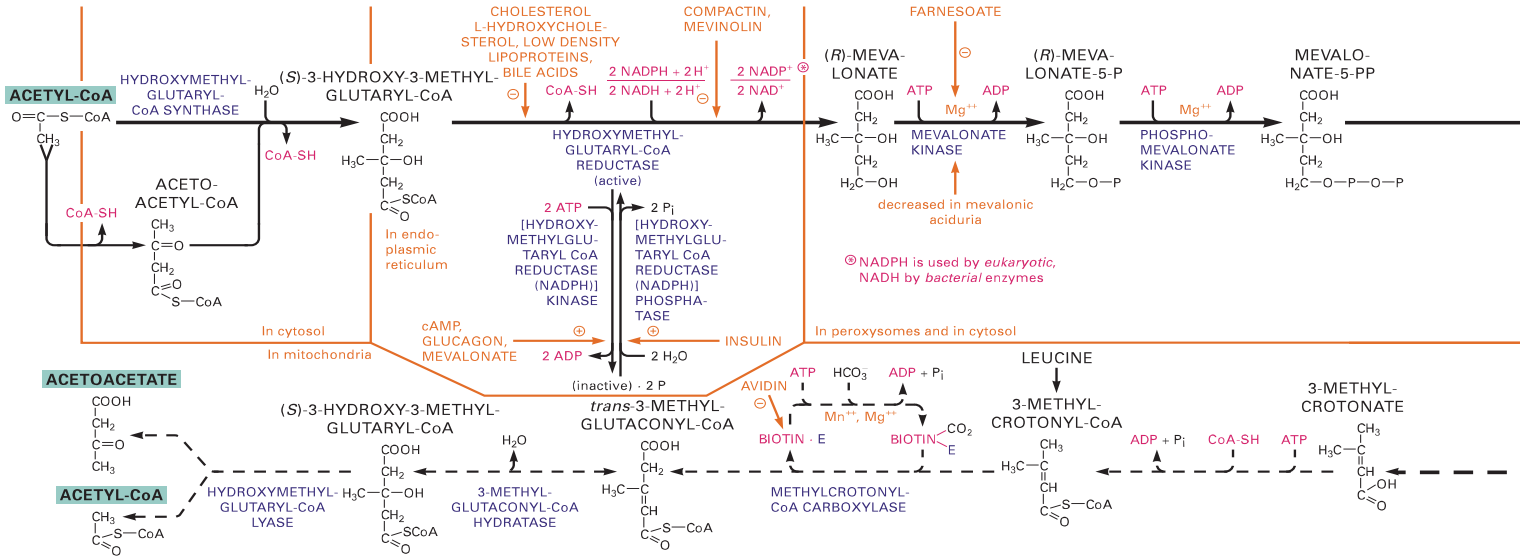
Gieselmann, V. Biochim. Biophys. Acta 1995;1270:103–136.

Merill, A.H, Sweeley, C.C., in Vace, D.E., Vance, J. (Eds.) *Biochemistry of Lipids, Lipoproteins and Membranes. New Comprehensive Biochemistry* Vol. 31. Elsevier, 1996, p. 309–339.

Svennerholm, L. Acta Chem. Scand. 1963;17:239–250.

Table 3.4.4-4. Examples of Glycoglycerolipids

Primary Lipophilic Residue (Formulae see Fig. 3.4.2-1)	Sugars or Derivatives	Ligands of Sugar	Occurrence
Neutral glycoglycerolipids			
Diacyl- or monoacylglycerol	$(–Glc_{,}–Gal)_1$		plants (esp. in chloroplasts), animal nerves, few bacteria
	$(–Glc, –Gal)_{2\ or\ higher}$		bacteria (also with -Man), plants (esp. in chloroplasts), animal nerves, brain, saliva
Alkylacylglycerol	$–Gal, –(Glc)_n$		mammalian spermatozoa and brain (-Gal), saliva $(-Glc_n)$
	$(–Gal, –Glc)_n$		bacteria
Glycerol diethers (with phytyl or other residues)	$(–Man, –Glc)_{1-24}$		archaea (also tetraethers with 2 glycerol residues opposing each other → monolayer membrane in some thermophilic species)
Acidic glycoglycerolipids			
Diacyl- or monoacylglycerol	glucuronic or galacturonic acids, sialic acid		some bacteria
	–Gal, –Glc–Gal	$-HSO_3^-$	animal brain
	–Gal, –Glc	$-HSO_3^-$	archaea, some bacteria
	$(–Glc, –Gal)_n$, also acylated	$-(glycerol\text{-}P)_n$	bacteria: lipoteichoic acids (immunogenic!)
Alkylacylglycerol	–Glc	$-HSO_3^-$	human gastric mucus
	–Gal	$-HSO_3^-$	mammalian spermatozoa (seminolipid)
Glycerol diethers (with phytyl or other residues)	$(–Gal, –Man, –Glc)_n$	$-HSO_3^-$	archaea
Phosphatidic acid = $(Acyl)_2$ glycerol-P	–inositol–GlcN–$(Man)_3$	-P-ethanolamine-protein	protozoa, higher animals (membrane anchor of proteins)

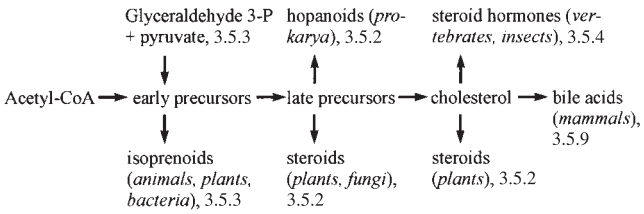


3.5 Steroids and Isoprenoids

Röbbe Wünschiers

3.5.1 Cholesterol

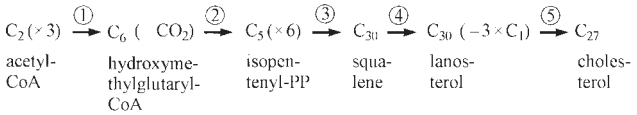
Cholesterol is a compound of essential importance for animals. It is an integral member of cellular membranes and influences their fluidity. Furthermore, a wealth of biologically active compounds is derived from cholesterol or intermediates of its biosynthesis in various species:



Cholesterol is not common in bacteria and plants. However, these organisms have functionally related structures (e.g., hopanoids, 3.5.2).

3.5.1.1 Biosynthesis (Fig. 3.5.1-1)

The biosynthesis can be subdivided as follows:



① Condensation of acetyl-CoA and acetoacetyl-CoA (formed from 2 acetyl-CoA) to hydroxymethylglutaryl-CoA (HMG-CoA). In animals, it occurs in the cytosol of liver (predominantly), small intestine, adrenals and gonads. The required acetyl-CoA is supplied from mitochondria by the citrate-pyruvate shuttle (Fig. 3.4.1-1). The reverse HMG-CoA lyase reaction is carried out by a mitochondrial enzyme and leads to ketone bodies (3.4.1.7).

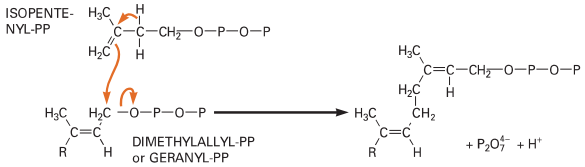


Figure 3.5.1-2. Nucleophilic Substitution Reaction of Isopentenyl-PP and Dimethylallyl-PP to Geranyl-PP

② Reduction to mevalonate, catalyzed by HMG reductase, an enzyme embedded into the membrane of the smooth endoplasmic reticulum (ER, but also present in peroxisomes). It is the key reaction for cholesterol biosynthesis and is therefore strictly regulated. The following reactions, pyrophosphorylation and decarboxylation to isopentenyl pyrophosphate (IPP) take place in peroxysomes (as major site) and in the cytosol. However, IPP (the origin of isoprenoids), is obtained in a different way (3.5.3).

③ Isomerization to an equilibrium with dimethylallyl-pyrophosphate (Fig. 3.5.1-2). For the condensation to squalene, the activated compound dimethylallyl-PP acts as a primer and isopentenyl-PP is added via a 'head to tail'-nucleophilic substitution. Subsequently, the geranyl-PP formed becomes a primer for an analogous condensation yielding farnesyl-PP. Squalene is synthesized by a 'head to head' condensation of 2 farnesyl-PP molecules. This, as well as the following reactions of cholesterol biosynthesis, takes place in the smooth endoplasmic reticulum.

④ In eukarya, squalene is converted to (S-)squalene-2,3-epoxide by an oxygen and NADPH dependent reaction. This compound undergoes a cyclization to lanosterol (in all animals except insects and in fungi) or cycloartenol (in plants, 3.5.2) catalyzed by lanosterol or cycloartenol synthases, respectively (Fig. 3.5.1-3).

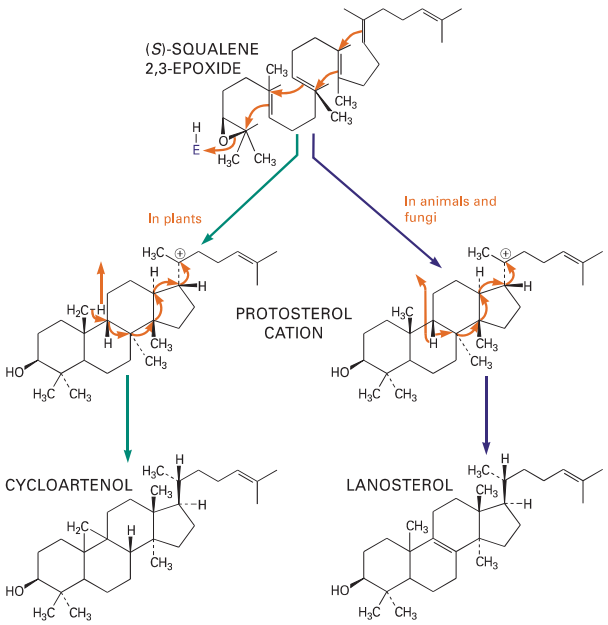
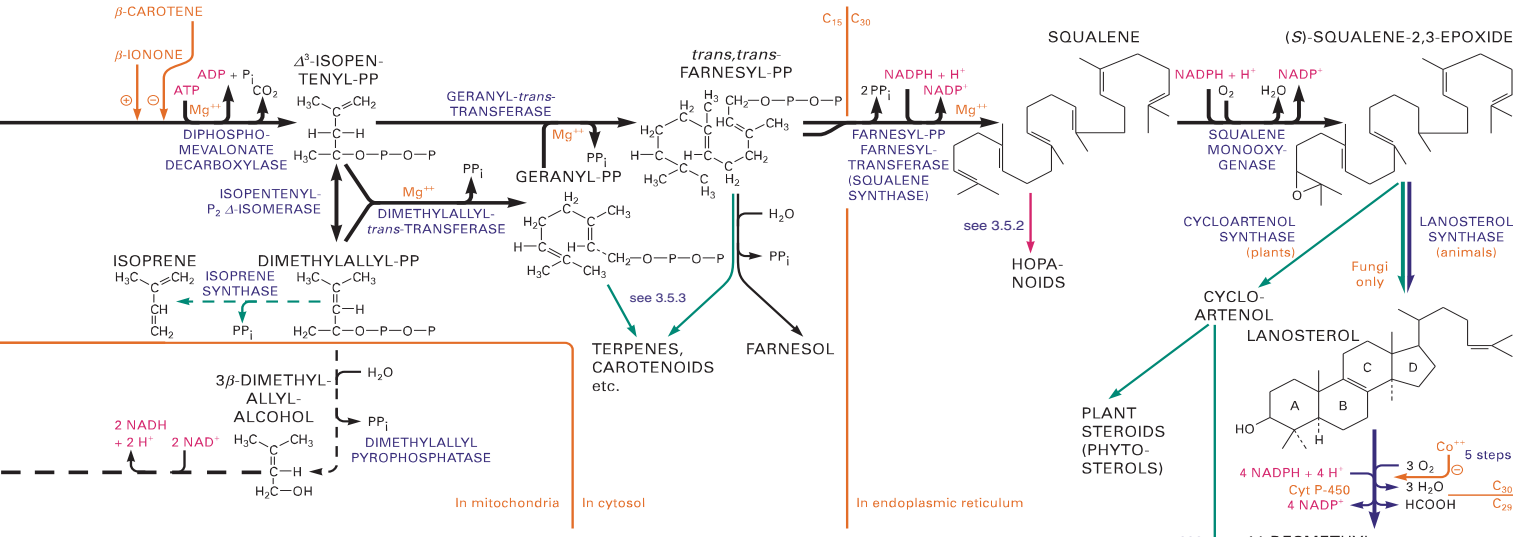


Figure 3.5.1-3. Cyclization of Squalene to Cycloartenol or Lanosterol



Cyclization starts with an electrophilic attack by the enzyme at C-3, followed by electron shifts. The resulting C-20$^+$ cation causes a backward rearrangement by H and CH_3 shifts. In animals, the final step is H elimination at C-9, yielding lanosterol. In plants, nucleophilic interaction of the enzyme with C-9 causes hydrogen migration from C-9 to C-8 and finally elimination of H from C-19 with formation of a cyclopropane ring, resulting in cycloartenol. Prokarya cyclize squalene to hopanoids (3.5.2).

⑤ Conversion of lanosterol to cholesterol: This is a sequence of 19 reactions, which include removal of three methyl groups (one at C-14 as formate, two at C-4 as CO_2) to zymosterol, followed by migration of a double bond and reduction of another. Depending on the enzymatic equipment of the particular tissue, this can take place by either of two different methods (via lathosterol or desmosterol).

An important conversion of cholesterol is esterification with long chain fatty acids (e.g., palmitic acid) in the ER, assisted by sterol carrier protein (SCP_2). This blocks the only polar group of cholesterol, which is essential for its presence in membranes. Ester splitting by an esterase takes place in a strictly regulated way.

While cholesterol degradation in mammals leads to bile acids (3.5.9), bacteria may either reduce cholesterol to coprostanol (in the lower parts of the mammalian intestine) or oxidize it to Δ^4-cholesten-3-one.

3.5.1.2 Turnover of Cholesterol (Fig. 3.5.1-4)

The daily demand of cholesterol in humans is 1 … 1.5 g. Food intake supplies 50 % or less of this. The intestinal resorption rate is only

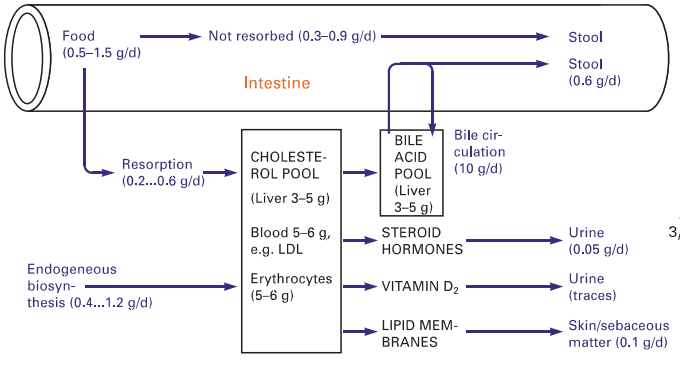


Figure 3.5.1-4. Flow Sheet of Cholesterol in Humans

Figure 3.5.1-5. Cholesterol Ring Structure (*all-trans* = chair representation)

Figure 3.5.1-1. Cholesterol Biosynthesis

about 40 %; it can be greatly reduced by administration of phytosterols (e.g., β-sitosterol, stigmasterol, campesterol).

The endogeneous cholesterol synthesis yields about 0.4 … 1.2 g/day. Cholesterol is degraded to bile acids (3.5.9) in the liver and excreted in the order of about 0.6 g/day. Small quantities are converted to steroid hormones and vitamin D_2. The flow of compounds in humans is shown in the graph.

The intracellular 'cholesterol pool' (Table 3.5.1-1) consists of cholesterol esters and free cholesterol. Considerable quantities of esters besides free cholesterol are also present in blood lipoproteins (6.2), which perform the lipid transport. Mammalian plasma membranes contain 40 … 50 % lipids, $^1/_2$ … $^1/_3$ of which is free cholesterol. In intracellular membranes (mitochondria, ER, nucleus), the cholesterol share is lower ($^1/_7$ … $^1/_{10}$). Cholesterol esters are practically absent from membranes. However, some precursors of cholesterol, such as lathosterol, 7-dehydrocholesterol and desmosterol can be found there.

Table 3.5.1-1. Cholesterol Distribution (Rat, Weight 341 g). (Jones, A., Glomset, J. in Danielsson *et al.*, p. 96)

	Free Cholesterol	Cholesterol Esters
Total	439 mg	60 mg
Plasma	0.4 %	8.7 %
Intestine	4.0 %	1.4 %
Liver	6.2 %	24.7 %
Skin	12.1 %	25.5 %
Muscle	16.9 %	2.5 %
Brain + nerves	21.2 %	0.2 %

3.5.1.3 Function of Cholesterol in Membranes

Unesterified cholesterol orients itself perpendicularly to the surface with the –OH group directed towards the aqueous phase. The planar, *all-trans* sterol structure (Fig. 3.5.1-5) and the polarity promote its association with phospholipids and lipoproteins. Cholesterol lowers the membrane fluidity by restricting the movement of fatty acyl chains (above the transition temperature). Coated pits, which assemble LDL receptors (6.2.4), are cholesterol-poor membrane areas. Also other membrane functions are affected by variations in the cholesterol content.

3.5.1.4 Regulation of Cholesterol Synthesis (Fig. 3.5.1-7)

The synthesis rate has to respond to the food intake and to the needed quantity of cholesterol, its metabolites and precursors. Thus it must be strictly regulated. Cholesterol synthesis exhibits a pronounced circadian rhythm with a maximum at midnight.

- Short term regulation (indicated by solid orange arrows) adapts itself within seconds to variations in supply or demand. This is primarily achieved by phosphorylation (inactivation) and dephosphorylation (activation) of hydroxymethylglutaryl-CoA (HMG-CoA) reductase, the rate-determining initial enzyme of cholesterol biosynthesis (Fig. 3.5.1-6). While insulin and thyroid hormones promote dephosphorylation, glucagon and cAMP enhance the (inactivating) protein kinase.
- The regulating protein kinase, in turn, is activated by cAMP dependent phosphorylation and inactivated by dephosphorylation, thus constituting a cascade mechanism (Fig. 3.5.1-6). The same protein kinase also regulates the activity of fatty acid synthesis by phosphorylating and inactivating the initial enzyme, acetyl-CoA carboxylase (3.4.1.2). Both pathways are coordinated this way.
- Furthermore, competitive feedback inhibition of several enzymes of cholesterol biosynthesis takes place. E.g., doubling of the physiological concentration of farnesyl-PP to 5 µmol/l inhibits mevalonate kinase more than 50 %. Still higher concentrations of farnesyl-PP (25 µmol/l) also inhibit its own conversion to squalene and favor its degradation (Fig. 3.5.1-7). Potent inhibitors of HMG-reductase are the fungal metabolites compactin and mevinolin.
- Long-term regulation (lasting hours, indicated by dashed orange arrows) occurs via changes in the transcription or translation rate or in the half-life time of mRNA or the enzymes.

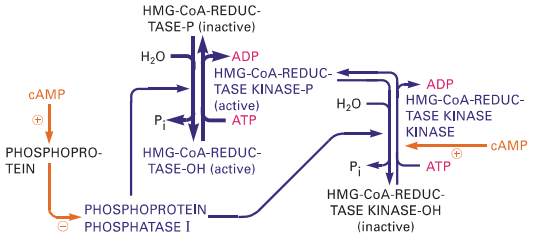


Figure 3.5.1-6. Phosphorylation and Dephosphorylation of HMG-CoA Reductase

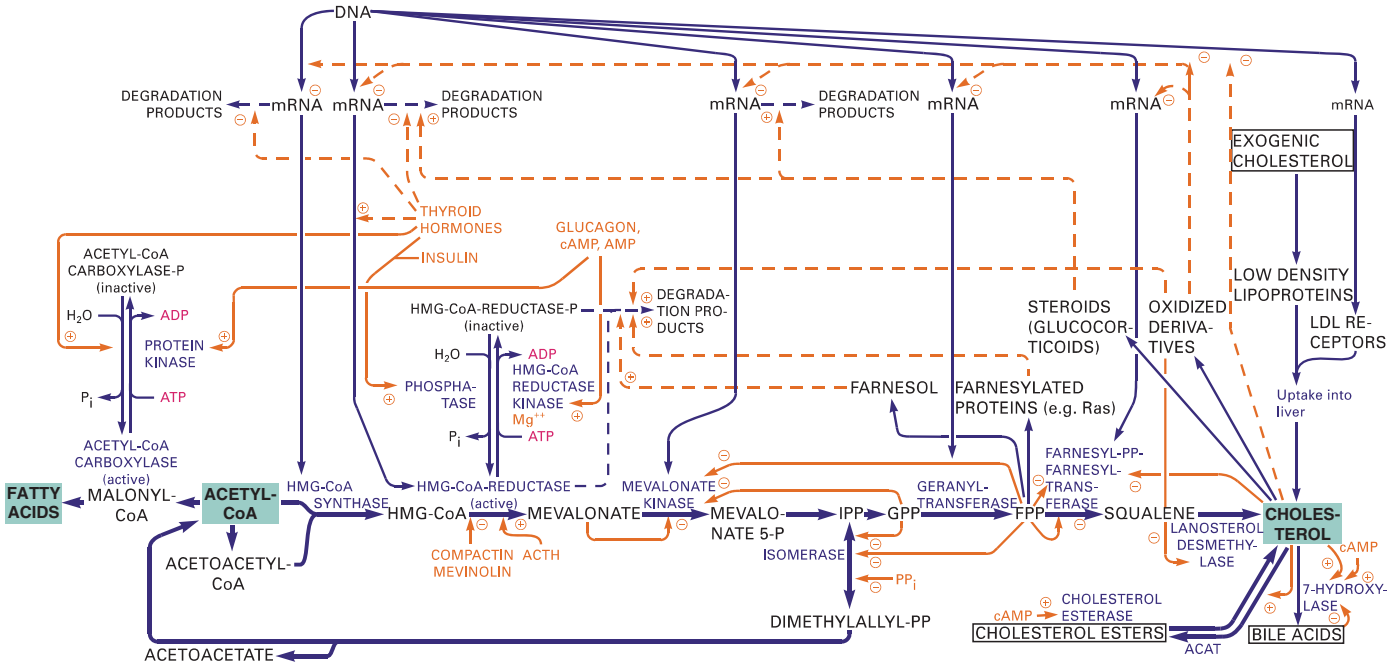


Figure 3.5.1-7. Reactions Involved in Cholesterol Homeostasis

Thyroid hormones increase the biosynthesis rate for HMG-CoA reductase (3.5.1.1). On the other hand, oxidized metabolites of cholesterol (e.g., 25-hydroxycholesterol) diminish the transcription rate of this gene, of the genes for HMG synthase, geranyl transferase, farnesyl-PP, farnesyl transferase and LDL receptor.

HMG reductase is regulated at additional levels: Glucocorticoids (being metabolites of cholesterol) and other compounds accelerate the mRNA breakdown. Thus, in rat liver the transcription rate of the HMG-CoA reductase gene can be modified 30-fold and the half-life time of the enzyme up to 35-fold, accounting well for the more than 450-fold change observed. In addition, HMG-CoA reductase is regulated by sterol-accelerated ubiquitination and degradation. However, no total blockade of the enzyme occurs, since sufficient intermediates have to be supplied for the synthesis of non-sterols (3.5.2; 3.5.3.). Similar regulation is exerted with HMG-CoA synthase and mevalonate kinase.

3.5.1.5 Cholesterol Homeostasis (Fig. 3.5.1-7)

Besides controlling the cholesterol synthesis, regulation mechanisms have also to encompass all other cholesterol converting reactions.

- In addition to its effects on cholesterol synthesis, 25-hydroxycholesterol decreases the transcription rate of the LDL receptor gene (6.2.4), and thus coordinates cholesterol synthesis with its uptake into the liver. Involved in this regulation is a common *cis*-acting DNA sequence SRE-1 (sterol regulatory element 1) located in the 5′ flanking region of the genes for HMG synthase and reductase, geranyl transferase and the LDL receptor. In case of low cholesterol concentrations, the transcription factor SREBP-1 (sterol regulatory element binding protein, 68 kDa) is liberated by proteolysis, binds to SRE-1 and activates transcription.
- After LDL has entered the liver cells, the cholesterol esters are hydrolyzed by the lysosomal cholesterol esterase. Excess cholesterol formed this way activates the acyl-CoA cholesterol acyl transferase (ACAT) (6.2.4) of the endoplasmic reticulum and achieves re-esterification in this way. The esters are deposited as lipid droplets. Cholesterol can be mobilized again by the neutral cholesterol esterase, which (at least in some organs) appears to be regulated by cAMP dependent phosphorylation and dephosphorylation.
- Cholesterol regulates its own degradation by stimulating the transcription of cholesterol 7α-hydroxylase, the initial enzyme of bile acid production (3.5.9). Additionally, this enzyme is controlled by the bile acids returning to the liver after reabsorption. The 7α-hydroxylase appears to be regulated by phosphorylation (activation) and dephosphorylation (deactivation), the activating kinase being a cAMP dependent enzyme.

Disturbances in cholesterol metabolism cause various diseases. In familial hypercholesterolemia the uptake of cholesterol into the cells via LDL receptor is diminished (6.2.4). Therefore, in spite of high plasma cholesterol concentration, the (intracellularly regulated) cholesterol synthesis still proceeds at high speed.

Literature:

Botham, K.M. Biochem. Soc. Transactions 1992;20:454–459.

Brown, M.S., Goldstein, J.L. Angew. Chem. 1986;98:579–599.

Danielsson, H., Sjövall, J. (Eds.) *Sterols and Bile Acids. New Comprehensive Biochemistry* Vol.12. Elsevier, 1985.

DeBose-Boyd, R. A. Cell Res 2008;18:609-621

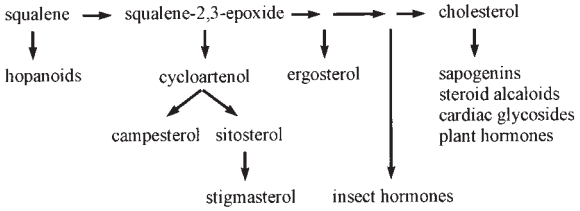
Meigs, T.E. *et al.* J. Biol. Chem. 1996;271:7916–7922.

Molowa, D.T., Cimis, G.M. Biochem. J. 1989;260:731–736.

Vance, D.E., Vance, J. (Eds.) Biochemistry of Lipids, Lipoproteins and Membranes. New Comprehensive Biochemistry Vol. 31. Elsevier, 1996.

3.5.2 Hopanoids, Steroids of Plants and Insects

Late precursors of cholesterol biosynthesis, as well as cholesterol itself, can give rise to a large number of important compounds for bacteria, plants, fungi and insects. (For early precursors see 3.5.3.)



3.5.2.1 Hopanoids (Fig. 3.5.2-1, left)

These pentacyclic compounds occur in all fossil sediments and petroleum deposits in huge quantities. They are widely distributed in prokarya as a membrane component and were found also in some eukarya (e.g., in the ciliate *Tetrahymena* and in some higher plants). Apparently, this evolutionarily very old group of compounds developed at a time, when the earth atmosphere still lacked oxygen. Whenever the epoxidation of squalene (leading to cholesterol or to cycloartenol) is impossible, squalene cyclizes in water directly to hopanoids (e.g., diploptene, diplopterol, tetrahymenol). Their side chains are then modified with ATP as energy source. Frequently, polyhydroxylated C_8-side chains are found, which are attached via ester or amide bonds to sugars, amino acids and nucleotides.

3.5.2.2 Phyto- and Mycosterols (Fig. 3.5.2-1)

While in most animals cholesterol is the major steroid product, other sterols (substituted at C-24) are prevalent in higher plants and fungi. Well known sterols in plants are campestrol, sitosterol and stigmasterol and in fungi ergosterol. They act as membrane components. The majority of phytosterols are conjugated as esters and glycosides. Other plant steroids (derived from cholesterol) are sapogenins (e.g., diosgenin), steroid alcaloids, cardiac glycosides and brassinosteroids (plant hormones) (Fig. 3.13-9).

Biosynthesis: In all higher plants and algae, 2,3-epoxysqualene is not cyclized to lanosterol, but to cycloartenol (Fig. 3.5.1-3). The 3-ring annelated to ring B is opened in later metabolic steps. Oxidative removal of the methyl groups at C-4 and C-14 proceeds analogously to cholesterol biosynthesis. Conversely, side chain methylations (at C-24, sometimes at C-23 or C-25) take place and are frequently species specific. A major number of plant sterols are synthesized this way.

On the other hand, cycloartenol may yield cholesterol, which can be converted via pregnenolone and progesterone into plant cardenolides (digitoxigenin, digoxigenin, g-strophantidin, etc.). The naturally occurring glycoside derivatives are named cardiac glycosides or commonly ‘digitalis’ [e.g., digoxin from *Digitalis* (foxglove), oubain from *Strophantus,* see Fig. 3.13-9]. Toad poisons (e.g., bufotoxin) with digitalis-like properties originate in cholesterol, too. Recently it has been discovered that compounds of this kind (ouabain, digoxin, bufadienolides) are synthesized in mammalian tissues in an analogous way.

Properties: β-Sitosterol decreases the intestinal resorption of cholesterol and is used for treatment of hypercholesterolemia. Ergosterol can be converted by illumination with ultraviolet light into ergocalciferol (vitamin D_2), an opening of ring B takes place. It has the same vitamin effects as cholecalciferol (vitamin D_3, 3.7.11). Sapogenins and steroid alcaloids are toxic and have hemolytic effects. Brassinosteroids cause cell elongation and cell division in plants. Cardiac glycosides, as well as the corresponding endogeneous hormones (see above) increase the contractility of the cardiac muscle (positive inotropic effect) by converting ATP more effectively into mechanical energy. Apparently the Na^+/K^+-ATPase (6.1.4) is inhibited, which causes intracellularly an increase of Na^+ and a

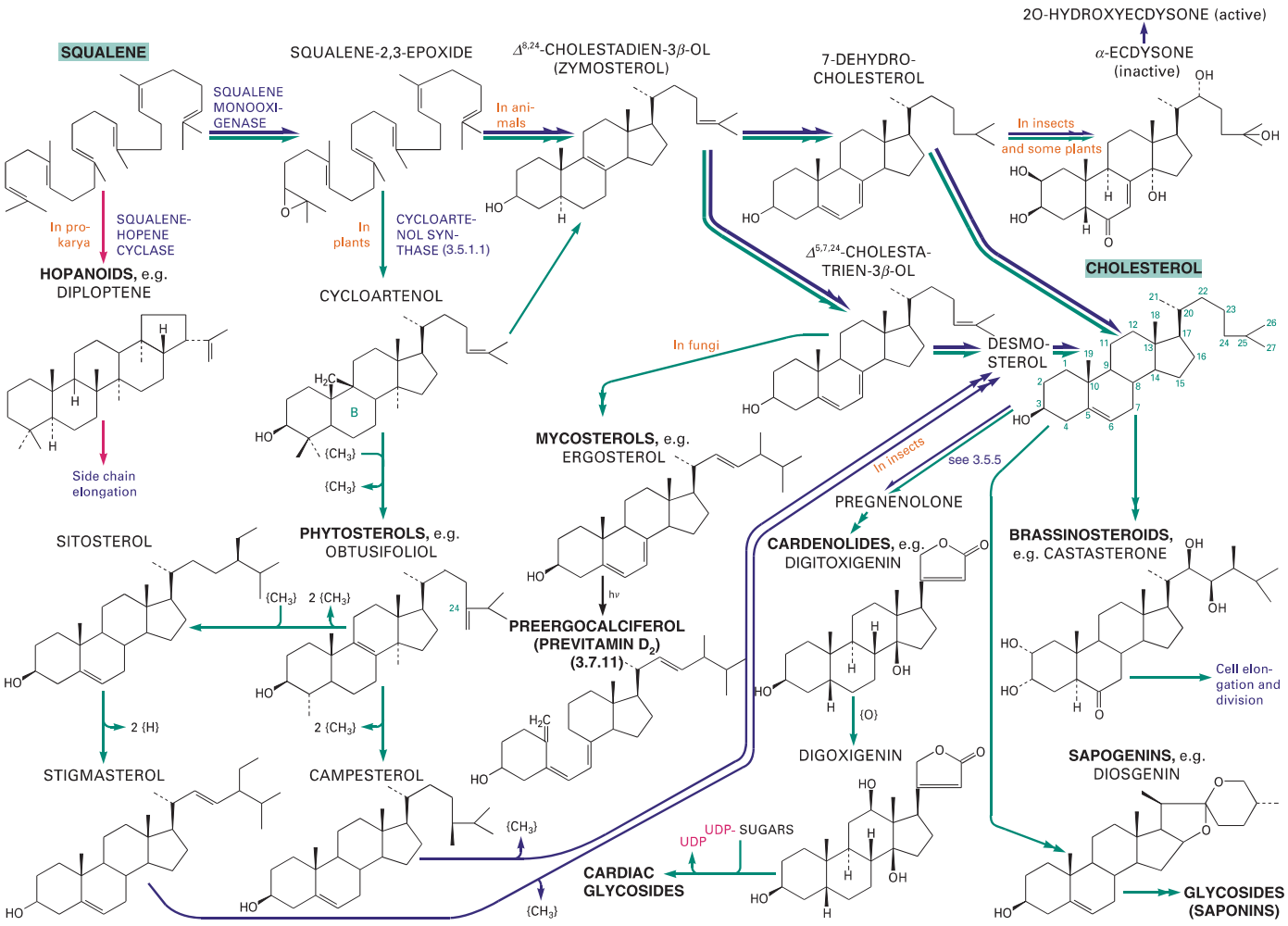


Figure 3.5.2-1. Metabolism of Hopanoids, Plant Sterols and Insect Hormones

decrease of K^+. The elevated Na^+ concentration, in turn, stimulates the Ca^{++}/Na^+ antiport mechanism and thus increases intracellular Ca^{++} (Table 6.1-4). This drug action mimics the effect of neuronal signals (7.4.5).

3.5.2.3 Ecdysone (Fig. 3.5.2-1, right)

Insects are unable to synthesize cholesterol *de novo*. They satisfy their demand either by food intake of cholesterol or by conversion of phytosterols into it in unique reaction sequences. The critical step is the demethylation at C-24. Highly specific inhibitors of this step are effective insecticides, which do not disturb other organisms.

From 7-dehydrocholesterol, hydroxylations lead to α-ecdysone and further to 20-hydroxyecdysone, which are highly effective hormones for molting and pupating. They also control vital functions in crustaceans, nematodes and mollusks and are likely the phylogenetically oldest steroid hormones. Phytoecdysteroids are a family of about 200 plant steroids structurally related to the invertebrate steroid hormone 20-hydroxyecdysone.

Literature:

Dinan, L. Phytochemistry 2001;57:325–339.

Goodwin, T.W. in Danielsson, H., Sjövall, J. (Eds.) *Sterols and Bile Acids. New Comprehensive Biochemistry* Vol. 12. Elsevier, 1985, p. 175–198.

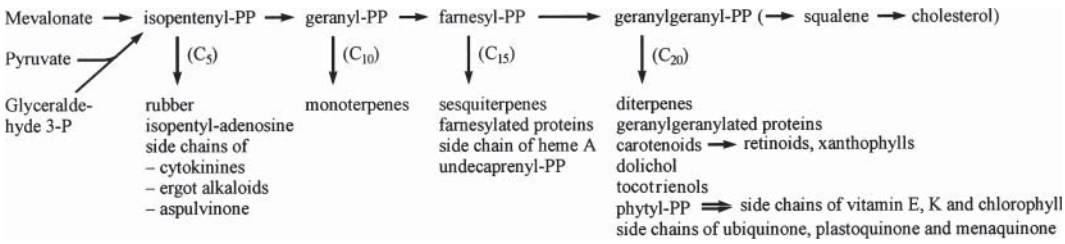
Schoner, W., Scheiner-Bobis, G. Nephrology Dialysis Transplantation 2008;23:2723–2729.

Sahm, H. *et al.* Adv. Microb. Physiology 1993;35:246–273.

Zhou, W. *et al.* Molecules 2004;9:185–203.

3.5.3 Isoprenoids

Some early intermediates of the cholesterol biosynthesis are the origin of linear and cyclic polymers in animals, plants and bacteria:
The common structural unit is isoprene (CH_2=C(CH_3)–CH=CH_2). In its activated form, isopentenyl-PP (CH_2=C(CH_3)–CH_2CH_2–PP) is used by animals and plants for synthesis of linear and cyclic polymeres by prenyltransferases.



Two distinct pathways exist for the biosynthesis of the universal isoprenoid precursors, isopentenyl diphosphate (IPP) and dimethylallyldiphosphate (DMAPP).

- In animals, fungi, archaea, and some bacteria, IPP and DMAPP are formed via the mevalonate (MVA) pathway (Fig. 3.5.1-1). In plants, the MVA pathway results in cytosolic IPP, which is then isomerized to DMAPP by the IPP isomerase. Mitochondrial isoprenoids (e.g., ubiquinone) are also synthesized from the MVA-pathway derived prenyl diphosphates that are imported from the cytosol.
- In plastides of plants, some apicomplexa protozoa (e.g., *Plasmodium*) and many eubacteria, IPP and DMAPP are formed by the independent 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway (1-deoxy-D-xylulose-5-phosphate, non-mevalonate pathway, or Rohmer pathway, Fig. 3.5.3-1). Since the MEP pathway is not present in humans, it is a target for the development of new herbicides and antimicrobial drugs.

3.5.3.1 Terpenes (Fig. 3.5.3-2)

Terpenes are a very large group of organic plant compounds, subdivided mainly into monoterpenes (C_{10}), sesquiterpenes (C_{15}) and diterpenes (C_{20}). Details are presented in 3.12.2.

The biosynthesis of terpenes generally takes place by intramolecular prenylation (analogous to Fig. 3.5.1-2, but within a single molecule), starting from geranyl-PP, farnesyl-PP or geranylgeranyl-PP, respectively. Consecutive rearrangements, introduction of hydroxy or carboxy groups etc. lead to an enormous number of compounds. Many of them are aromatic essences, some are of therapeutic importance (e.g., as antimalarial drugs).

3.5.3.2 *All-trans* Metabolites (Fig. 3.5.3-2)

Carotenoids (C_{40}) are formed from two geranylgeranyl molecules. The enzyme complex phytoene synthase in higher plants, yeasts, algae and fungi produces *cis*-phytoene (which is converted to the *trans* configuration in later steps), while most bacteria obtain *trans*-phytoene. Membrane-associated phytoene desaturase oxidizes and cyclizises phytoene via several steps to *all-trans* α- or β-carotene. By stepwise oxidation, plants and bacteria convert them to xanthophylls, which act as light absorbing pigments in photosynthesis (3.12.1). They are also able to eliminate free radicals (comparable to vitamin E in humans, 3.7.12) and act as photo-protective agents via the xanthophyll-cycle (Fig. 3.5.3-2). During light-stress, the di-epoxide compound violaxanthin is transformed via antheraxanthin to epoxide-free zeaxanthin, catalyzed by violaxanthin de-epoxidase. This compound is able to dissipate excess energy. In the dark, the reactions are reversed by zeaxanthin epoxidase.

Catalyzed by specific dioxygenases, β-carotene is split symmetrically to *all-trans* retinal. This compound is essential for the visual process in vertebrates (7.4.6) as well as for photosynthesis in halophilic archaea (3,12,1). In cells, retinal is in equilibrium with *all-trans* retinol (vitamin A, 3.7.1).

Uptake of retinoids from food and storage takes place as retinol-fatty acid esters. Intercellular transport and distribution proceeds via specific retinol-protein complexes in blood (RBP, CRBP, CRBP II, IRBP). On the other hand, small quantities of retinal are oxidized to retinoic acid. This compound is transported as a highly specific complex with cellular retinoic acid binding protein (CRABP). Retinoic acid shows hormone-like activity in cell development and cell division by acting on specific nuclear receptors (7.7). It is applied as a cancerostatic drug in promyelocytic leukemia. β-Carotenes are of importance for stimulation of immunological defense and prevention of arteriosclerosis. Catabolism of all retinoids proceeds by conjugation with glucuronic acids and excretion in the urine.

3.5.3.3 *Poly-cis* Metabolites (Fig. 3.5.3-2)

Dolichol-P (C_{70} … C_{120}) is an essential sugar carrier in animals, plants and fungi for formation of glycoproteins in the endoplasmic reticulum (4.4.1). Furthermore, dolichyl esters serve as a transport and storage

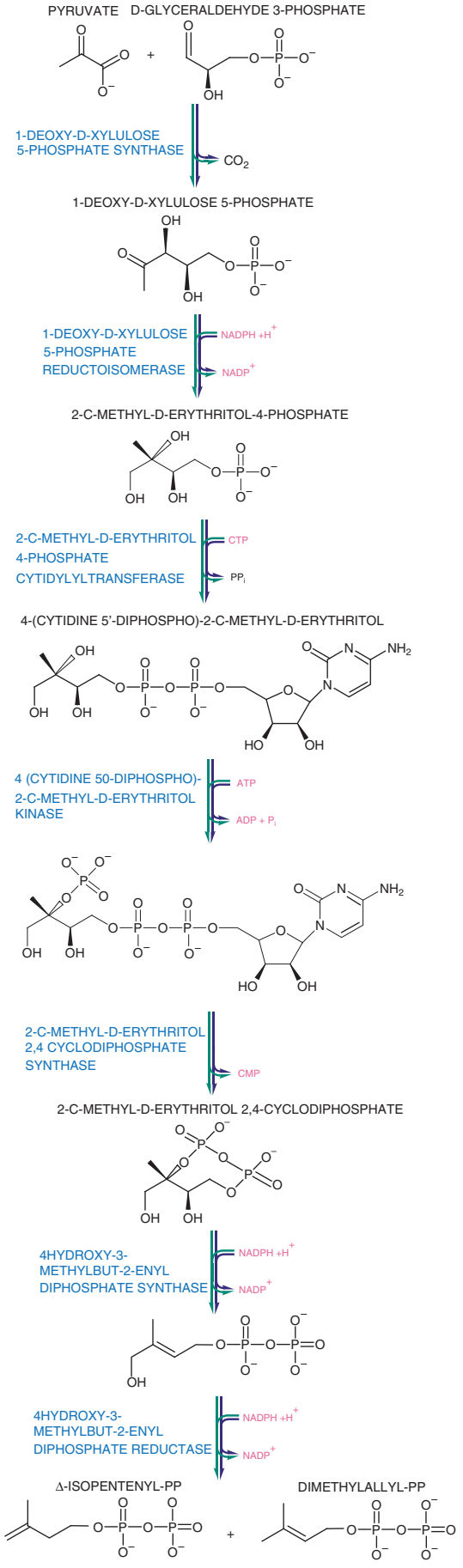


Figure 3.5.3-1. MEP pathway (Phillips *et al.*)

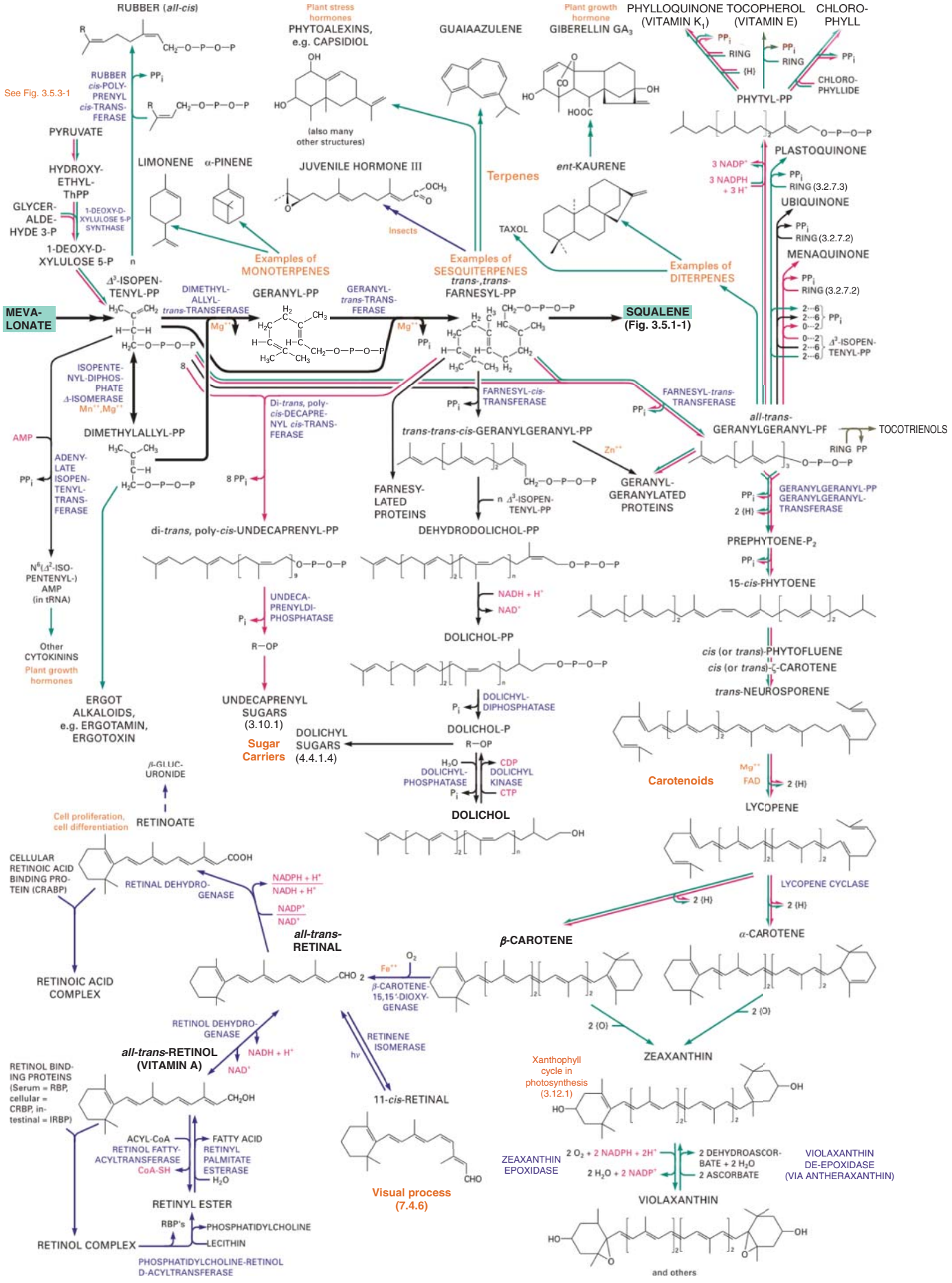


Figure 3.5.3-2. Metabolism of Terpenes, Carotenoids, Retinoids

form of fatty acids. *Cis*-geranylgeranyl-PP enters the pathway to dolichol via a highly affine polyprenyl-*cis*-transferase.

Bacteria use undecaprenyl-P (C_{55}) as carrier for glycopeptides in formation of cell wall murein (Fig. 3,10.1-1) and as a sugar carrier for synthesis of other membrane components, like lipopolysaccharide or teichoic acids. Both compounds carry 2 … 3 double bonds in *trans*-configuration at their non-hydroxyl end.

Isopentenyl-PP is the origin of natural rubber (e.g., from *Hevea brasiliensis,* 3.13.2.7), which consists of very long *all-cis* isoprenoid chains (ca. C_{3500} … C_{25000}).

3.5.3.4 Isoprenoid Side Chains (Fig. 3.5.3-2)

Bacterial pigments and cytokines (plant hormones) contain isoprenoid side chains. *All-trans*-geranylgeranyl-PP forms the isoprenoid moiety of tocotrienols (vitamin E group).

Phytyl-PP (C_{20}, originating from *all-trans*-geranylgeranyl-PP in plants and some bacteria) is a side chain in chlorophyll (5.4), α-tocopherol (vitamin E, 3.7.12) and phylloquinone (vitamin K_1). Longer isoprenoid side chains are present in menaquinone (vitamin K_2, C_{20} … C_{30}, 3.7.13), ubiquinone (C_{30} … C_{50}, 3.2.7.2) and plastoquinone (C_{45}, important in plant respiration).

Many proteins are attached to cellular membranes by isoprenoid anchors (via a thioether bond to mostly farnesyl, sometimes geranylgeranyl chains), e.g., Ras (7.5), Gγ proteins (7.4) and heme A (3.3.2.4). This prenylation determines the location and the function of the proteins. The enzymatic reaction is shown in Figure 3.5.3-3. Farnesylation takes place if X = serine, methionine or glutamine. Geranylgeranyl is attached if X = leucine or phenylalanine. Proteins of the Rab group, which contain –C–C– or –C–X–C– sequences, are geranylgeranylated similarly. The enzyme is assisted by the Rep protein.

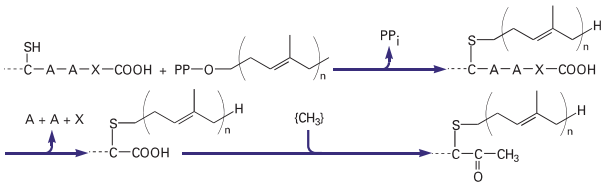


Figure 3.5.3-3. Isoprenoid Anchors of Proteins
(C = cysteine, A = aliphatic amino acid)

Literature:

Gershenzon, J., Croteau, R.B. in Moore, T.S. (Ed.) *Lipid Metabolism in Plants.* CRC Press, 1993, p. 339–388.

Hemming, F.W. in Wiegandt, H. *Glycolipids. New Comprehensive Biochemistry* Vol. 10. Elsevier, 1985, p. 261–305.

Packer, L. (Ed.) *Retinoids. Part A. Methods in Enzymology* Vol 189. Academic Press, 1990.

Packer, L. (Ed.) *Carotenoids. Part B. Methods in Enzymology* Vol 214. Academic Press, 1993.

Phillips, M.A. *et al.* Trends in Plant Science 2008;13:619–623.

Rekittke, I. *et al*. J. Am. Chem. Soc. 2008;130:17206–17207.

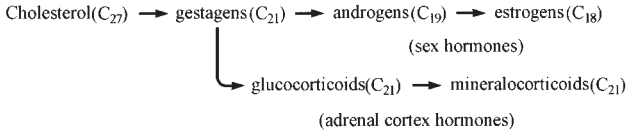
Rodriguez-Concepcion, M. and Boronat, A. Plant Physiol. 2002;130:1079–1089.

Young, AJ. R. *et al*. Pure & Appl.Chem 1997;69:2126–2130.

Zhang, F.L., Casey, P.J. Ann. Rev. Biochem. 1996;65:241–269.

3.5.4 Steroid Hormones

Cholesterol is the origin of a multitude of steroids, which are of high importance in metabolic processes. While the compounds in plants and insects are dealt with in 3.5.2 and 3.5.3, the following Section deals with vertebrate hormones.



3.5.4.1 Biosynthesis

The common synthesis pathway starts with the partial removal of the cholesterol side chain, yielding the gestagen progesterone. The further reactions depend on the enzyme equipment of the particular cells (Fig. 3.5.4-1). Removal of the remaining side chain leads to androgens, a further aromatization of ring A yields estrogens. On the other hand, different hydroxylations of progesterone produce gluco- or mineralocorticoids.

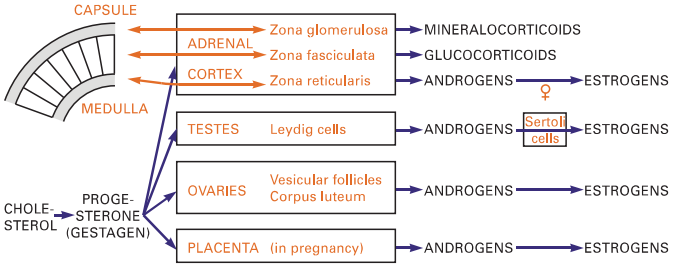


Figure 3.5.4-1. Sites of Steroid Hormone Biosynthesis

Reaction types: The major conversions are hydroxylations, reductions/isomerizations (usually combined), C-C bond splitting and aromatization (estrogen formation). Frequently, the enzymes do not show absolute specificity, so that several alternative reaction sequences can lead to the same end product.

Hydroxylations are mostly catalyzed by NADPH dependent mixed function steroid monooxygenases, which are associated with cytochrome P-450 (Fig. 3.5.4-2). They transfer one oxygen from O_2 to the steroid and the other one to hydrogen. They also play a major role in detoxification reactions.

$$R–H + O_2 + NADPH + H^+ \rightarrow R–OH + NADP^+ + H_2O$$

Other hydroxylases depend on different hydrogen donors, like NADH, $FADH_2$ or ascorbic acid.

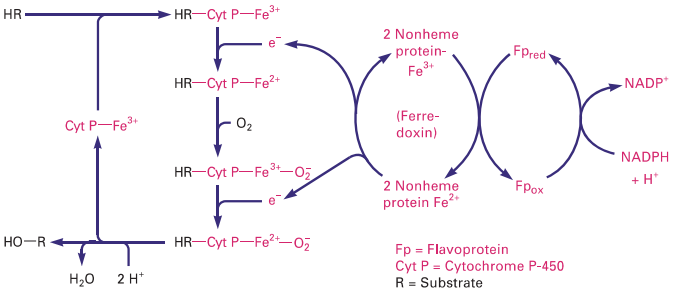


Figure 3.5.4-2. Mechanism of Cytochrome P-450 Dependent Hydroxylations

3.5.4.2 Biological Activation and Regulation of Steroid Hormones

Activation of most steroid hormones in mammals is a three-step cascade starting at the hypothalamus (for a general discussion, see 7.1.5). An external stimulus causes the hypothalamus to secrete releasing hormones (activating liberins and inhibiting statins), which cause a specific activation of the anterior pituitary lobe (adenohypophysis). In turn, trophic hormones (tropins) are secreted, which are transported by the bloodstream to peripheral tissues (adrenals, testis, ovary) and cause there the secretion of the steroid hormones.

The releasing and the trophic hormones (Fig. 3.5.4-3) are peptides and act on cell surface receptors, which transmit the signals intracellularly via various pathways (e.g., G-proteins and phosphorylation cascades, 7.4 and 7.5) to their site of action, e.g., gene expression in the nucleus (resulting in enzyme synthesis) or elevated NADPH production (which increases hydroxylation reactions).

Contrary to this 'second messenger' system, the steroid hormones produced by this cascade act directly. Since they are very lipophilic, they are able to diffuse through the membrane of their target cells. In the nucleus, they are bound to specific receptors with very high affinity (K = 10^{-8} … 10^{-10} mol/l). The hormone-receptor complexes attach

themselves to special DNA sequences (hormone responsive elements) and enhance transcription (7.7). Since in different organs different genes are activated, the same hormone may control the synthesis of various proteins.

The secreted hormones exert a feedback action towards the hypothalamus and the pituitary gland. Therefore, the reaction cascade is activated only for short times. Furthermore, the steroid hormones show only short half-life times (10 … 90 min.). Both effects prevent prolonged, high concentrations of steroid hormones and allow the system to respond quickly to changing situations.

As an exception, aldosterone is regulated by the renin-angiotensin system, which is dependent on blood volume and plasma electrolyte concentration. (7.1.8)

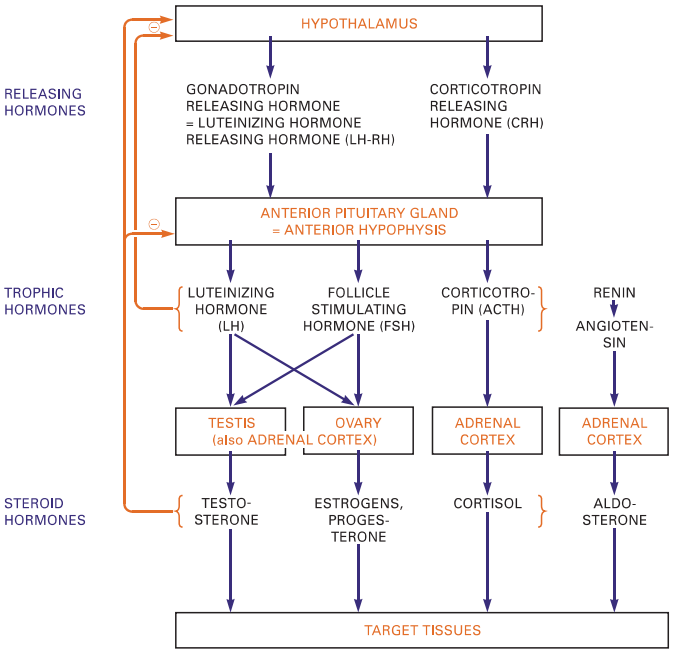


Figure 3.5.4.3. Regulation of the Synthesis and Secretion of Steroid Hormones (see also Fig. 7.1-4)

3.5.4.3 Transport of Steroid Hormones

For transportation in blood, the lipophilic steroid hormones are bound to plasma protein carriers (albumin, orsomucoid, transcortin, sex hormone binding globulin and α_1-fetoprotein). These complexes are physiologically inactive. At the surface of the target cell, the complexes release the hormone for diffusion into the cell.

3.5.4.4 Degradation of Steroids

Only bacteria are able to open the rings of the steroid skeleton. In animals, deactivation and modification for excretion occur in liver and to a smaller degree in kidneys and lung. Major reactions, which increase the solubility in the aqueous phase are:

- reduction of the 3-oxo to the 3-hydroxy group (α or β) and hydrogenation of the double bond in ring A (the aromatic ring in estrogens is not modified),
- conjugation at the C-3 hydroxyl with sulfate (as phosphoadenylyl phosphosulfate) or glucuronide,
- introduction of more hydroxy groups by cytochrome P-450 dependent enzymes,
- oxidative degradation of the side chain in C_{21} steroids.

Literature:

Cooke, B.A., King, R.J.B., van der Molen, H.J.(ed.) *Hormones and their Actions Part I and II New Comprehensive Biochemistry* Elsevier, 1988.

Edwards, D.P. Journal of Mammary Gland Biology and Neoplasia 2000;5:307–324.

Norman, A.W. and Litwack, G. in *Hormones* (Second Edition). San Diego: Academic Press, 1997, p. 49–86.

Norman, A.W. in Benjamin, C. (ed.) *Encyclopedia of Food Sciences and Nutrition.* Oxford: Academic Press, 2003 p. 3166–3174.

3.5.5 Gestagen

3.5.5.1 Biosynthesis of Progesterone (Figs. 3.5.5-1 and 3.5.5-2)

In mammals, the binding of follicle-stimulating hormone (FSH), luteinizing hormone (LH) or corticotropin (ACTH) to their receptors at the adrenal cortex, ovary or testes membranes starts a G-protein mediated reaction sequence (7.4). This leads to a transport of cholesterol from cytoplasmatic inclusion droplets into the mitochondria, assisted by sterol carrier protein SCP_2. The conversion of cholesterol to pregnenolone is effected by the trifunctional complex cholesterol monooxygenase, which is located at the inner membrane of mitochondria. The presumed reaction steps are 22-hydroxylation, 20-hydroxylation and side chain lyase reaction. The hydroxylations correspond to the monooxygenase scheme (3.5.4.1), using the $P450_{scc}$ variant and adrenodoxin as the Fe-S-protein.

Pregnenolone then passes through the mitochondrial membrane and is converted by a bifunctional enzyme complex to the gestagen hormone, progesterone (Fig. 3.5.5-1). This reaction takes place at the smooth endoplasmatic reticulum. Pregnenolone and progesterone are also the origin of androgens (3.5.6), estrogens (3.5.7), gluco- and mineralocorticoids (3.5.8); pregnenolone also of cardiac glycosides in plants and mammals (3.5.2.2).

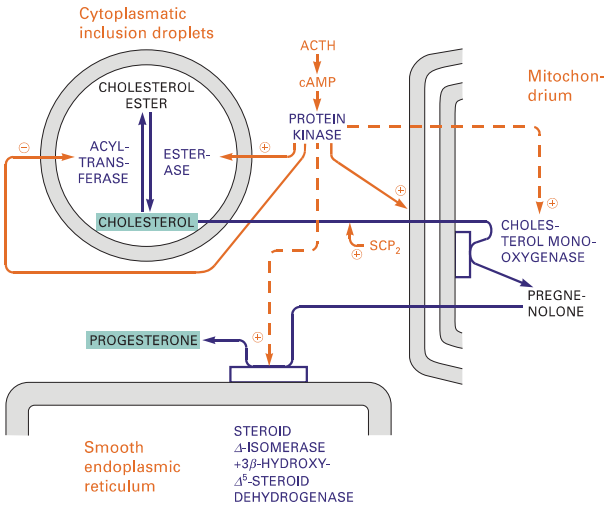


Figure 3.5.5-1. Conversion of Cholesterol to Progesterone

In bacteria and plants, progesterone biosynthesis proceeds analogously. However, the conversion of pregnenolone to progesterone is performed by 2 separate enzymes.

3.5.5.2 Gestagen Function, Transport and Degradation

In female mammals, progesterone plays an essential role as the only active gestagen. It is produced mostly in the corpus luteum of the ovaries during the second half of the cycle and in great quantities in the placenta during pregnancy. It binds tightly to receptor proteins in the nucleus and regulates the transcription to mRNA (4.2.2.3). Its physiological effects are

- during the second half of the female cycle:
  - preparation of the uterus for implantation of the fertilized ovum (modification of the endometrium),
  - general effects on the CNS (e.g., elevated body temperature).

- during pregnancy:
  - preservation of the mucous coat of the uterus,
  - prevention of further ovulations,
  - formation of lactating alveoli in the breasts (synergistically with estrogen).

The estrogen-progesterone interaction is biphasic: High estrogen concentrations (e.g., during pregnancy) stimulate the formation of progesterone receptors, thus increasing progesterone effects. On the other hand, they also stimulate 17β-hydroxysteroid dehydrogenase, which converts estradiol into the much less active estrone. This decreases the progesterone receptor formation again.

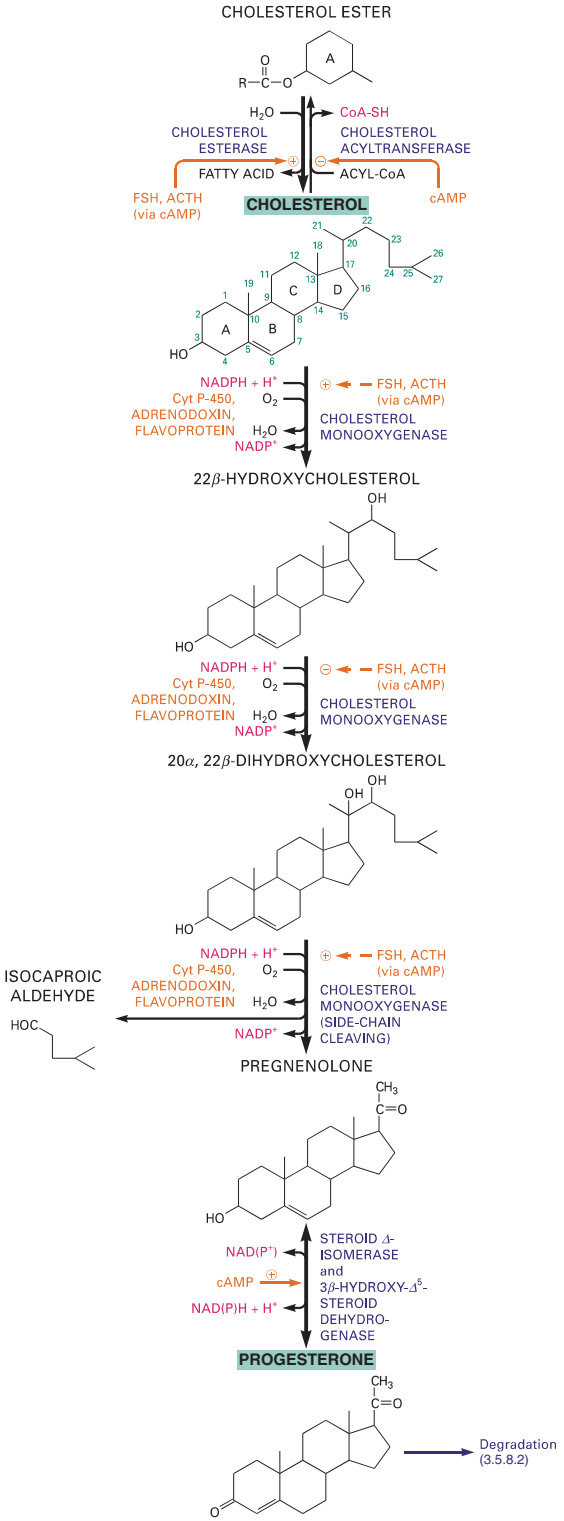


Figure 3.5.5-2. Biosynthesis of Progesterone

In the adrenal cortex of both sexes and in testis, progesterone is only an intermediate in biosynthesis of sex hormones and of gluco- and mineralocorticoids (adrenals only).

Progesterone is transported in serum bound to albumin, transcortin (corticosteroid-binding globumin) or orsomucoid (acidic α_1 glycoprotein). Degradation (see Fig. 3.5.8-2) involves reduction of the keto functions and the double bond and excretion as glucuronide.

Since ovulation is suppressed by an external supply of progesterone and/or estrogens, this principle is used for contraception. For oral administration, synthetic compounds are used, e.g., hydroxyprogesterone or ethinylnortestosterone derivatives (gestagens) and ethinylestradiol derivatives (estrogens). In most cases, a combination of both groups is used.

Literature:

Cooke B.A., King R.J.B., van der Molen H.J. (Eds.) Hormones and their Actions Part I and II. New Comprehensive Biochemistry. Elsevier, 1988.

Edwards, D.P. Journal of Mammary Gland Biology and Neoplasia 2000;5:307–324.

Goodman, H.M. in *Basic Medical Endocrinology* (Third Edition). San Diego: Academic Press, 2003, p. 391–456.

3.5.6 Androgens

Androgens are the determining factors for male development and behavior in vertebrates.

3.5.6.1 Biosynthesis (Fig. 3.5.6-1)

In male mammals, most of the androgens are produced in the Leydig cells of the testis. Additionally, androgens are synthesized in the ovaries and placenta of females and in the zona reticularis of the adrenal cortex. The respective enzymes are associated with the smooth endoplasmic reticulum of these cells. They are activated by luteinizing hormone (7.1.5), the concentration of which rises dramatically at puberty.

Either pregnenolone (Δ^5 pathway, e.g., in humans and rabbits) or progesterone (Δ^4 pathway, e.g., in rats and mice) is hydroxylated at the 17-position. The conversion between both pathways is catalyzed by a complex of two enzymes, it can take place at different levels of both pathways. Then the side chains are oxidatively removed, yielding 4-androstene-3,17-dione or dehydroepiandrosterone (DHA), respectively, which show weak androgen function. The splitting enzyme is closely linked with the 17-hydroxylase. Further redox reactions and isomerizations convert them into the important androgen, testosterone. 4-Androstene-3,17-dione can also be converted into adrenosterone, testosterone into 5α-dihydrotestosterone (5α-DHT), which shows even higher biological activity and is responsible for a series of specific androgen effects (see Table 3.5.6-1).

In testes and ovaries, most likely an additional pathway exists, which leads via 17α, 20α-dihydroxycholesterol directly to dehydroepiandrosterone. Up to 30 % of the androgens are produced this way in these organs.

In females, 4-androstene-3,17-dione and testosterone are formed in the theca interna cells of the ovaries as intermediates for estrogen production (3.5.7.1).

3.5.6.2 Transport and Degradation

Since all steroid hormones are highly lipophilic, they have to be modified to a more hydrophilic form for transport purposes. Dehydroepiandrosterone is transported in the circulation as (hormonally inactive) sulfate to its target organs, where it is desulfatized again. Another transport vehicle is the sex hormone binding globulin (SBG), which binds testosterone strongly and estradiol only slightly less.

The catabolism of androgens occurs in kidneys and liver. The compounds are hydrogenated at the Δ^4 double bond. The 3-oxo group is

then reduced to the 3α hydroxy function in humans and to 3β in most other vertebrates and thereafter (mostly) glucuronized. The products are excreted.

3.5.6.3 Biological Function of Androgens (Table 3.5.6-1)

Androgens bind to nuclear receptors; the complexes act mainly through control of gene expression (7.7). This leads to growth and development of male sex characteristics and causes strong anabolic effects. Some androgen derivatives are even effective in nano- or picomolar concentrations.

Table 3.5.6-1. Effects of Androgens in Male Vertebrates

Species	Stage	Organ	Effects	Mainly involved
All	embryo	External genitals	sex differentiation	testosterone, 5α-dihydro-testosterone
Most	embryo	liver, spleen	hemoglobin synthesis erythropoietin synthesis	5α-dihydrotestosterone 5α-dihydrotestosterone
Most	neonate	brain	sexual differentiation	5α-dihydrotestosterone
Most	neonate	liver	enzyme synthesis	androstenedione
All	puberty	testis	spermatogenesis	5α-dihydrotestosterone
Most	puberty	hair follicle	hair growth	5α-dihydrotestosterone
Most	puberty	vocal cords	thickening	5α-dihydrotestosterone
Most	puberty	muscle	growth, N-retention	testosterone
Human	puberty	sebaceous gl.	acne	5α-dihydrotestosterone
All	adulthood	brain	male libido	testostererone, estradiol
Birds	adulthood	brain	courtship display	
Deer	adulthood	antlers	growth	

3.5.6.4 Medical Aspects

Synthetic analogues of natural androgens are used as anabolic steroids for treatment of muscular dystrophy, psoriasis etc. Steroidal and non-steroidal anti-androgens block androgen receptors and are used, e.g., in tumor treatment.

Literature:

Cooke, B.A., King, R.J.B., van der Molen H.J. (Eds.) *Hormones and their Actions Part I and II. New Comprehensive Biochemistry.* New York: Elsevier, 1988.

Nazian, S. in Enna, S. J. and David, B. B. (Eds.) *xPharm: The Comprehensive Pharmacology Reference*. New York: Elsevier, 2007, p. 1–3.

Norman, A. W. and Litwack, G. in *Hormones* (Second Edition). San Diego: Academic Press, 1997, 341–360.

3.5.7 Estrogens

Estrogens control the development of the reproduction system and the reproductive functions in female vertebrates (7.1.5).

3.5.7.1 Biosynthesis (Fig. 3.5.7-1)

Estrone and estradiol are formed by oxidative removal of the C-19 methyl group from the androgens 4-androstene-3, 17-dione and testosterone and subsequent aromatization of ring A in a concerted aromatase reaction. The estrogen synthesis is controlled by luteinizing and the follicle stimulating hormones (7.1.5).

The aromatase system contains cytochrome P-450 and NADPH-cytochrome P-450 reductase (3.5.4.1). The enzyme is associated with

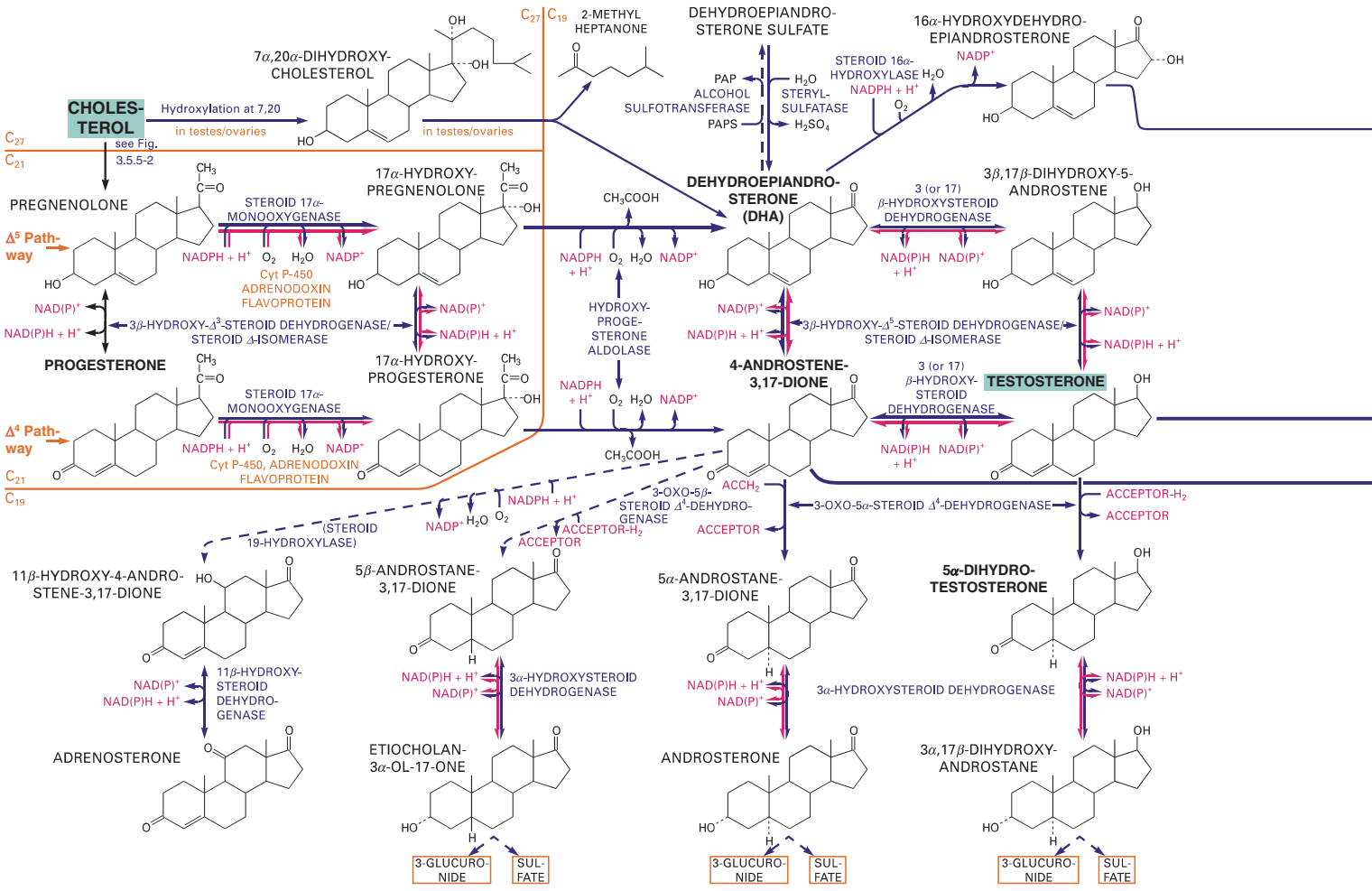


Figure 3.5.6-1. Metabolism of Androgens

the endoplasmic reticulum. Two successive hydroxylations convert the C-19 methyl group via the hydroxy to the oxo level. The mechanism of the consecutive formate elimination has not been completely resolved yet.

The aromatization reaction is unique in vertebrates. It takes place mostly in the granulosa cells of the ovaries, but also in the adrenal cortex (of less importance in pre-menopausal women) and in the Sertoli cells of the testis. 'Peripheral estrogen synthesis' occurs in adipose tissue and muscle (in females prevalent after menopause) and in liver and brain (the most important estrogen source in males, ca. 60%). In the placenta during pregnancy or in breast cysts, large quantities of estriol 3-sulfate are produced from dehydroepiandrosterone sulfate via initial 16α-hydroxylation, followed by aromatization. DHA originates both in the fetal adrenal cortex and in maternal circulation.

Estradiol and estrone can be interconverted by a redox reaction. Estriol is formed from estrone by 16-hydroxylation and 17-reduction reactions. 2-Hydroxylation of estrone and estradiol yields catechol estrogens, which are neuroendocrinologically active. This reaction can be performed by placental aromatase. Other steroids with estrogen function are found in some animals, e.g., equilin and equilenin in horses.

3.5.7.2 Transport and Degradation (Fig. 3.5.7-1)

Estrogens are transported in blood by the sex-hormone binding globulin. Degradation takes place mostly in the liver. Hydroxylation reactions of estrone and estradiol are performed by microsomal NADPH-dependent steroid hydroxylases. This takes place in different species at positions 6α, 6β, 7α, 11β, 14α, 15α, 16α, 16β and 18. Glucuronidization (prevalently) or sulfatation at the 3-position increases the solubility. Ca. 50 different metabolites are excreted in human urine. Some excretion also occurs via the intestine.

3.5.7.3 Biological Function of Estrogens (Table 3.5.7-1)

The control of the female reproductive functions includes a multiplicity of estrogen effects on gene expression. Estrogens combine with high affinity in the nucleus with specific estrogen receptors (65 … 70 kDa, 17.6). Estrogen receptors belong to a family of nuclear transcription factors like receptors for gonadal steroids, thyroid hormone, and retinoids. The complex effects an increase in the transcription rate (4.2.2) for proteins and peptides involved in growth control and DNA replication. This is in line with the estrogen dependency of some tumors. The estrogen activities decrease in the order estradiol > estriol > estrone. Generally, estradiol is active at concentrations, which are 100 to 1000 times lower than those required for hormone action of androgens.

Estrogens act anabolically on fat (causing reduced blood levels of cholesterol) and slightly anabolically on proteins. During menopause the estrogen level decreases and androgens from the adrenal cortex become more predominant.

Table 3.5.7-1. Effects of Estrogens in Female Vertebrates

Species	Stage	Organ	Effects
All	embryo	sex tract	development
Mammals	puberty	ovaries	estrus cycle
Mammals	puberty	endometrium of uterus	induction of cyclic proliferation
Mammals	puberty	epithelium of tubes and vagina	cyclic changes
Mammals	puberty	mammae	development
Human	puberty	hair, adipose tissue	adult distribution
Reptiles, amphibia, birds, fishes	adulthood	oviduct, liver	egg protein production

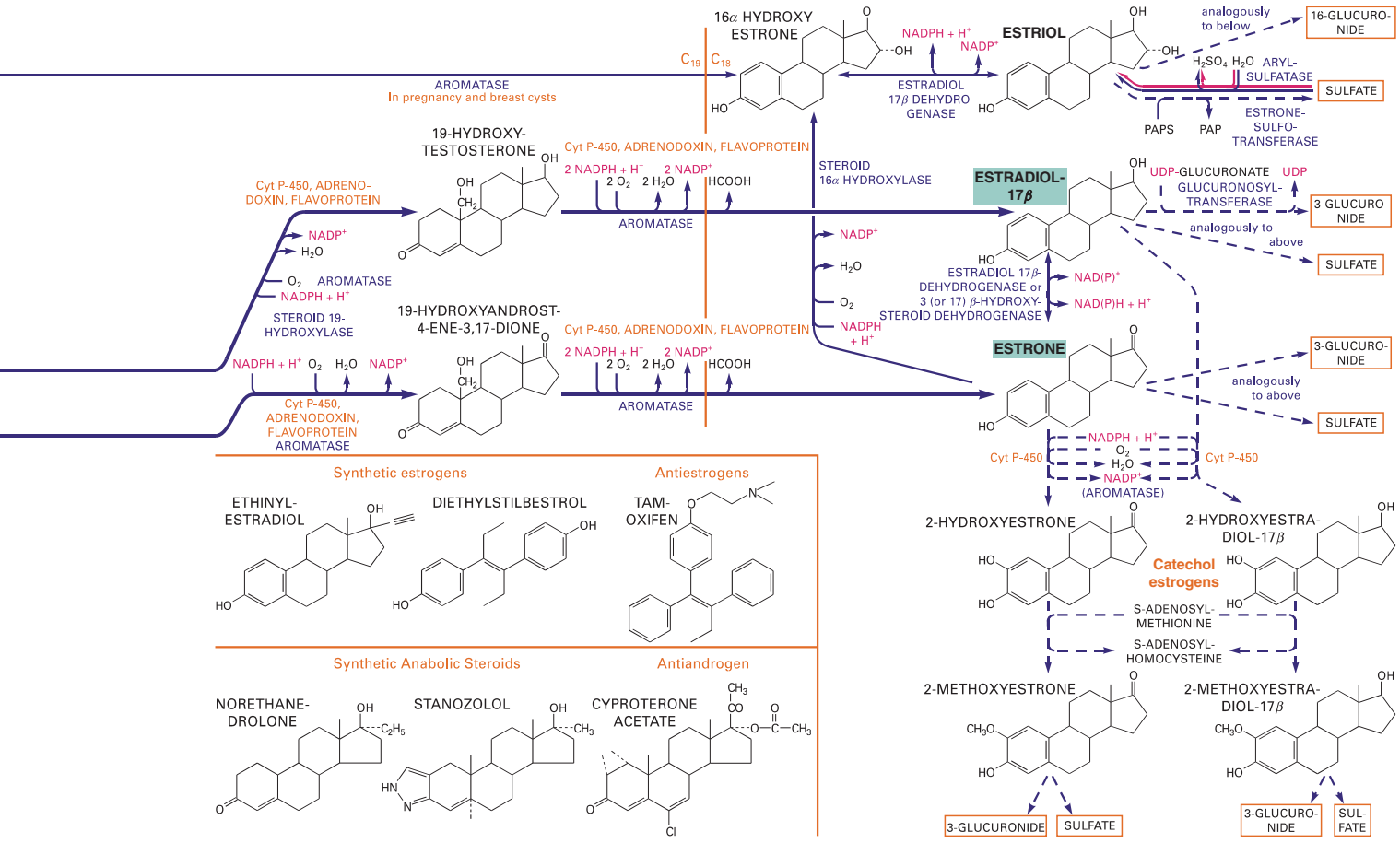


Figure 3.5.7-1. Metabolism of Estrogens

3.5.7.4 Medical Aspects

Estrogens are used for substitution therapy. Besides the natural hormones, ethinylestradiol and other estrogen derivatives, as well as stilbene derivatives can be applied (Natural stilbene derivatives occur in plants). Steroidal and non-steroidal antiestrogens (e.g., tamoxiphen) compete with estrogens at the receptors and thus act as inhibitors (species-dependent). These compounds, as well as aromatase inhibitors are used for tumor treatment.

Literature:

Cooke, B.A., King, R.J.B., van der Molen H.J. (Eds.): *Hormones and their Actions Part I and II. New Comprehensive Biochemistry.* Elsevier, 1988.

Hewitt, S. C. and Korach, K. S. Reviews in Endocrine & Metabolic Disorders 2002;3:193–200.

Hewitt, S. C. *et al.* Science 2005;307:1572–1573.

3.5.8 Corticosteroids

Corticoids are another class of hormones, which are produced by the adrenal cortex of vertebrates. They include the glucocorticoids, which are vital regulators of glucose metabolism and the mineralocorticoids, which control the mineral metabolism.

Table 3.5.8-1. Important Corticoids

Glucocorticoid Effects	Relative potency	Mineralocorticoid Effects	Relative potency
cortisol	1	aldosterone	1
cortisone	0.8	11-deoxycorticosterone	0.03
		cortisol	0.0003

3.5.8.1 Biosynthesis (Fig. 3.5.8-2)

As with all steroid hormones, glucocorticoids and mineralocorticoids are synthesized starting from cholesterol.

- Formation of glucocorticoids in the zona fascilata: From pregnenolone, the 17α-hydroxysteroids 11-deoxycortisol, cortisol and cortisone are synthesized via hydroxylations at 17α, 21 and 11β. This occurs preferably in humans and rabbits. In rats, however, these compounds are formed from progesterone.
- Formation of mineralocorticoids in the zona glomerulosa: Progesterone is the precursor of the 17-deoxycorticoids cortexone, corticosterone, 11-dehydrocorticosterone and aldosterone via hydroxylations at positions 21, 11β and 18. Oleyl-corticosterone is the active form of this hormone.

Although in both pathways the hydroxylations take place at the same positions 11β and 21, the enzymes are apparently different. Enzymatic defects lead to different diseases. In some species, a strict sequence of reactions does not exist due to only moderate specificity of the converting enzymes.

A striking phenomen is the repeated change between the compartments where the reactions take place (Fig. 3.5.8-1). Cholesterol from the cytosol is converted to pregnenolone in the mitochondria, isomerized and further hydroxylated in the smooth endoplasmic reticulum, while final hydroxylations and dehydrogenase steps take place in the mitochondria again. Secretion into the lymph requires another passage through the cytosol. This separation of the synthesis sites likely allows fine tuning of the regulation.

3.5.8.2 Transport and Degradation

In blood, cortisol is bound to transcortin (corticosteroid-binding globulin). The degradation of corticosteroids is generally initiated by hydration to tetrahydro derivatives (mostly 3α, 5β, but also 3α, 5α). Then glucuronidation and hydration to hexahydro derivatives usually follows. These are excreted in urine. The degradation of progesterone follows the same steps.

3.5.8.3 Biological Function (Table 3.5.8-1)

The function of steroid hormones can be classified in two major types of action. For the genomic, receptor-mediated effect, the steroid hormones pass through the cellular membrane and bind directly to nuclear receptor proteins, which induce DNA transcription (7.7). Non-genomic effects occur by interaction of the steroids with specific receptors potentially located in the cell membrane or with membrane lipids or non-specific proteins (Fig. 3.5.8-3).

Glucocorticoids are released from the adrenals in a diurnal rhythm: the blood level increases during the night and drops during the day.

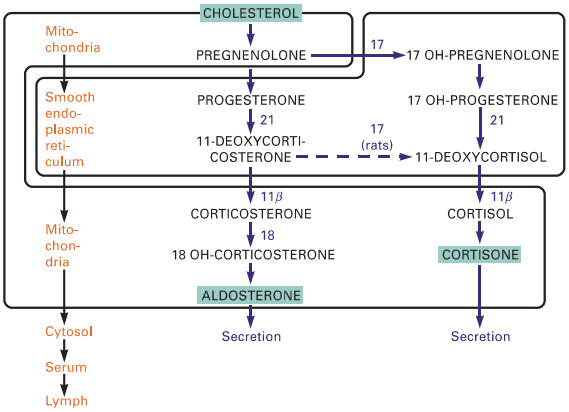


Figure 3.5.8-1. Sites Involved in Corticosteroid Synthesis
(Numbers indicate positions of hydroxylation)

Table 3.5.8-2. Diseases Caused by Defects in Corticoid Metabolism

Disease	Origin	Cause	Metabolic Effects	Clinical Picture
Cushing's syndrome	hypothalamus / pituitary gland (frequently tumors)	CRH or ACTH overproduced	cortisol ↑, ACTH ↑ (cannot be inhibited by cortisol)	steroid diabetes, muscular atrophy, bone destruction, decreased formation of antibodies
	adrenals, frequently tumors	cortisol overproduced	cortisol ↑ (no elevated ACTH)	same
Addison's disease	adrenals	chronic insufficiency of adrenal cortex	cortisol ↓ (in spite of elevated ACTH), Na^+ ↓, K^+ ↑	loss of water, myasthenia, hypoglycemia, blood pressure ↓, atrophy of genitals
Waterhouse-Friederichsen syndrome	adrenals	acute insufficiency of adrenal cortex	cortisol ↓↓	hemorrhagic necrosis due to meningococci-caused sepsis
Congenital adrenal hyperplasia	adrenal cortex	21-hydroxylase defective	plasma cortisol ↓, androgens ↑	androgenital syndrome, virilization, salt wasting syndrome
Conn syndrome (aldosteronism)	adrenal cortex, frequently tumors	aldosterone overproduced	aldosterone ↑, Na^+ ↑, K^+ ↓	elevated potassium excretion, alkalosis, edema
Secondary aldosteronism	kidney blood supply	arteriosclerosis, hypertonia, alkalosis	renin ↑, causing aldosterone ↑	same
Hypoaldosteronism	very rare	lack of 18-hydoxylase	Na^+ ↓, K^+ ↑	salt loss

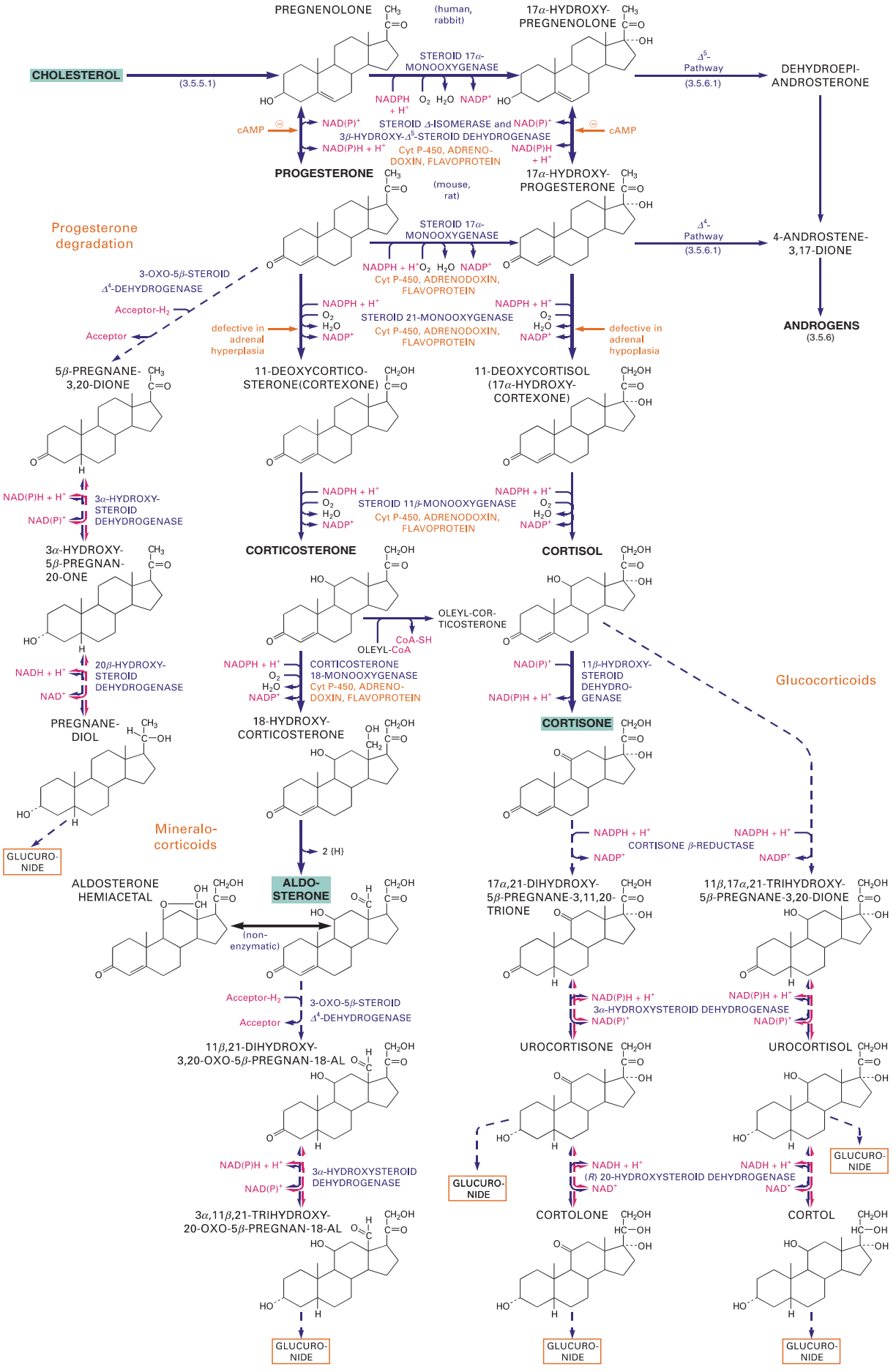


Figure 3.5.8-2. Metabolism of Glucocorticoids and Mineralocorticoids

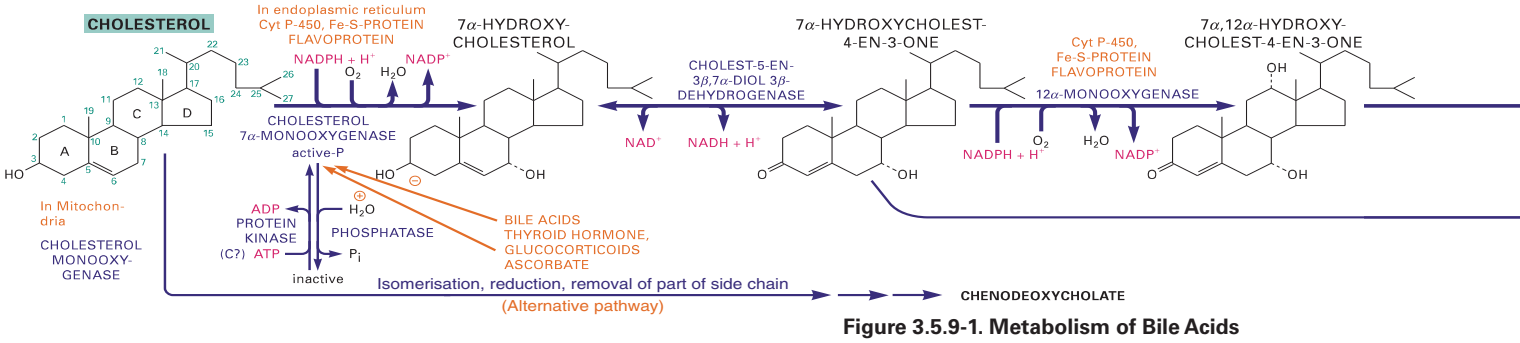


Figure 3.5.9-1. Metabolism of Bile Acids

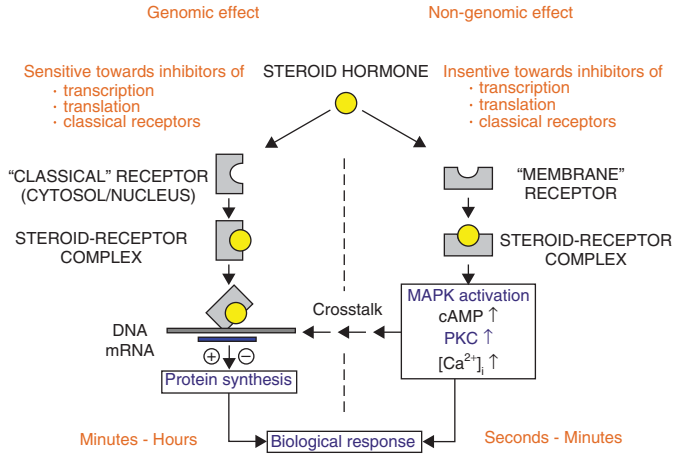


Figure 3.5.8-3. The Genomic and Non-Genomic way of steroid action (Zöllner *et al.*, 2008)

In stress situations, additional release takes place. Hormone actions proceed in almost all mammalian cell types:

- In extrahepatic tissues (muscle, skin, adipose tissue, lymphocytes etc.), catabolic enzymes are induced. Glucose resorption and glycolysis decrease, while the blood glucose level, proteolysis and lipolysis (regulated via epinephrine) increase.
- In liver, due to anabolic effects, protein biosynthesis, gluconeogenesis and glycogen synthesis rise.
- In the lymphatic system, an immunosuppressive effect occurs via inhibition of the response of T-lymphocytes (8.2.6) and the activation of macrophages, as well as a decrease in the number of circulating lymphocytes, eosinophiles, monocytes and macrophages.
- A strong suppressive effect on inflammation is exerted by high levels of (especially) cortisol and cortisone. This is due to:
  - biosynthesis of the glycoprotein lipocortin, which inhibits phospholipase A_2 and therefore decreases the liberation of arachidonic acid and its conversion into prostaglandins (7.4.8),
  - feedback inhibition of the hypothalamus, which suppresses the fever reaction to pyrogens,
  - downregulation of NO synthase II (7.8.2),
  - inhibition of leukocyte immigration into the inflamed tissue (8.2.6).

Mineralocorticoids: Aldosterone promotes in the proximal and distal kidney tubuli

- reabsorption of Na^+ and Cl,
- retention of water,
- secretion of K^+, H^+ and NH_4^+ ions.

This is effected by increased expression of the Na^+/K^+ channel, the Na^+channel and enzymes of the citrate cycle. Therefore, administration of aldosterone leads to a considerably decreased cytosolic K^+ level and to a tendency for edema formation. Similar effects are also shown by other corticoids, such as 11-deoxycorticosterone, cortisone, cortisol and corticosterone.

3.5.8.4 Medical Aspects

A number of serious diseases are caused by disturbances in the corticoid metabolism. Frequently, defective hydroxylases cause a shift in the corticoid hormone pattern (Table 3.5.8-2).

Glucocorticoids are used for substitution therapy of adrenal insufficiency. Synthetic glucocorticoid analogs with two double bonds in ring A (e.g., dexamethasone) do not show mineralocorticoidal side effects and are preferred. Artificial antiglucocorticoids are used for treatment of Cushing's syndrome and adrenocorticoidal carcinoma. Due to multiple side effects, only in very serious cases high doses of cortisol and cortisone are used as antiphlogistic drugs.

For treatment of essential hypertonia or of edema, antimineralocorticoids (e.g., spironolactone) are applied.

Literature:

Cooke, B.A., King, R.J.B., van der Molen, H.J. (Eds.) *Hormones and their Actions Part I and II New Comprehensive Biochemistry.* Elsevier, 1988.

Morris, D.J. *et al.* Steroids 2009;74:1–6.

Rogerson, F. M. and Fuller, P. J. Steroids 2000;65:61–73.

Zöllner, S. *et al*. Steroids 2008;73:966–972.

3.5.9 Bile Acids

In vertebrates, the major part of cholesterol is finally converted to bile acid conjugates in the liver and excreted via bile into the small intestine. The conjugates are bile acid amides with taurine or glycine or half-esters with sulfuric acid. Due to their low pK_a values (taurocholic acid ca. 2, glycocholic acid ca. 4), they dissociate completely at physiological pH values. Therefore they are good anionic detergents, which form mostly cylindric micelles. In the intestine, they act as emulsifiers on lipid food components (e.g., triglycerides, cholesterol and its esters). This enables their resorption by the intestinal mucosa (Fig. 6.2-2).

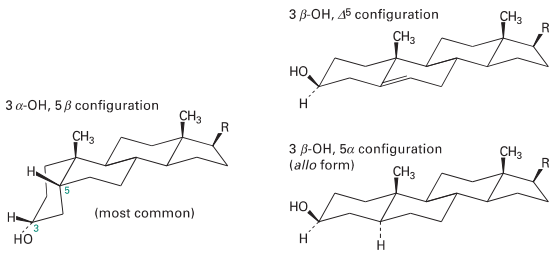


Figure 3.5.9-2. Configuration of Bile Acids

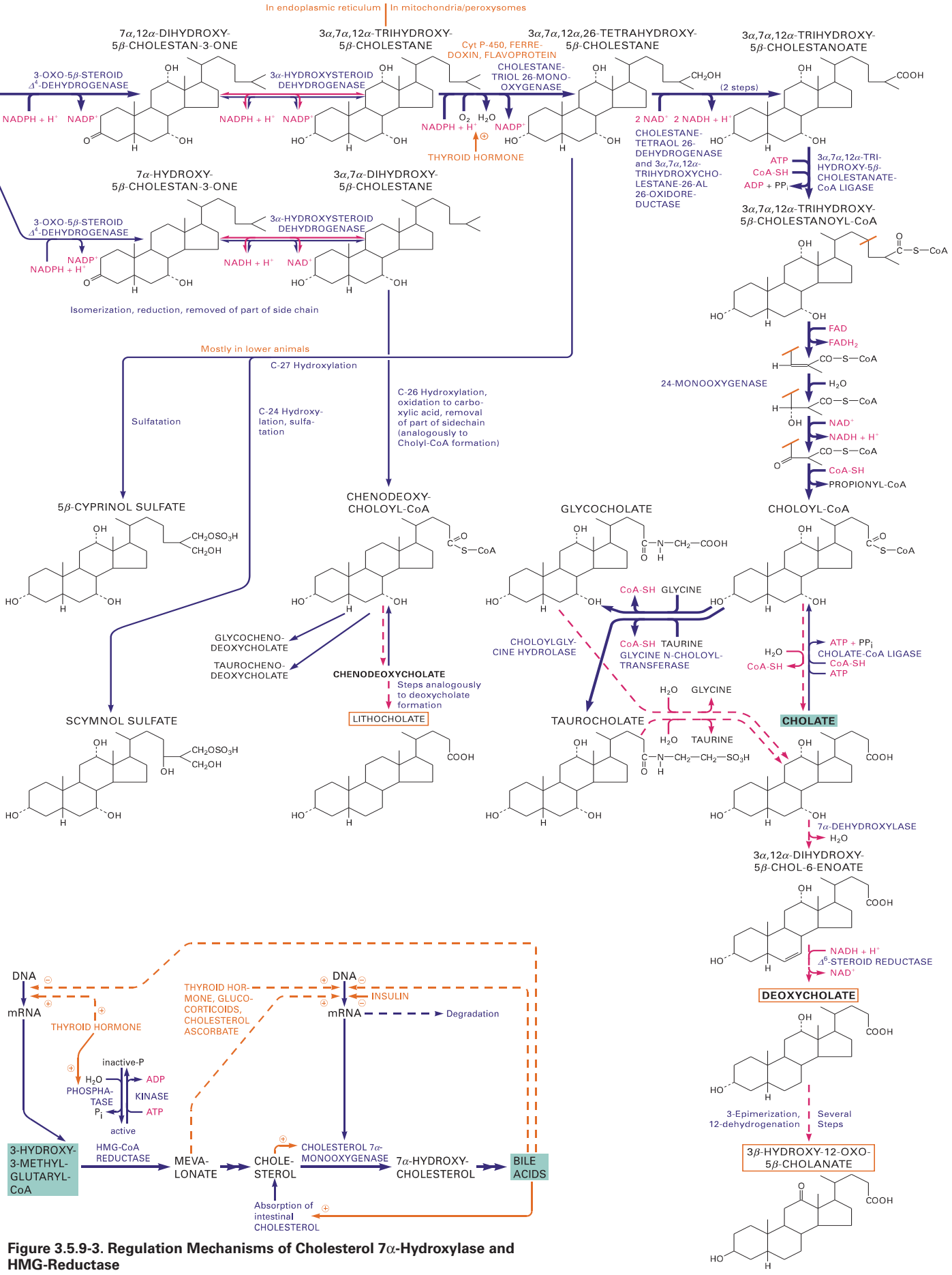


Figure 3.5.9-3. Regulation Mechanisms of Cholesterol 7α-Hydroxylase and HMG-Reductase

Table 3.5.9-1. Occurrence of Bile Acid and Bile Alcohol Conjugates (Examples)

Family	Trivial Name	Hydroxyl Groups at	Conjugated with	Species
Bile alcohols	5β-cyprinol	3α, 7α, 12α, 26	sulfuric acid	fishes
	scymnol	3α, 12α, 24, 26	sulfuric acid	fishes
	5α- or β-bufol	3α, 7α, 12α, 25, 26	sulfuric acid	toads, lung fishes
C_{27} Bile acids	3α, 7α, 12α, 22-tetrahydroxy-5β-cholestan-26-oic acid	3α, 7α, 12α,22	taurine	turtles
C_{24} Bile acids	cholic acid	3α, 7α, 12a	taurine, glycine	snakes, birds, mammals
	chenodeoxycholic acid	3α, 7α	taurine, glycine	birds, mammals
	ursodeoxycholic acid	3α, 7β	taurine	some mammals
	deoxycholic acid	3α, 12α	taurine, glycine	some mammals, intestinal bacteria
	lithocholic acid	3α	taurine, glycine	some mammals, intestinal bacteria
	α / β-muricholic acid	3α, 6β, 7α/ β		rat, mouse
	hyocholic acid	3α, 6α, 7α		pigs
	3α-hydroxy-7-oxo-cholanic acid	3α (+ 7-oxo)		some birds and mammals

3.5.9.1 Occurence (Table 3.5.9-1)

Bile alcohol derivatives occur only in fishes and amphibia and have frequently four or more hydroxyl groups, one of which is esterified with sulfuric acid. C_{27}-bile acids are found in amphibia and reptiles (except snakes) as taurine conjugates. The largest group are the C_{24} bile acids, which are derivatives of 5β-cholanic acid (5α = *allo* derivatives are less frequent). Taurine conjugates are widely distributed, while glycine conjugates occur only in mammals. In human bile, the ratio of cholic acid / chenodeoxycholic acid / deoxycholic acid / lithocholic acid conjugates is typically 1 / 1 / 0.4 / traces.

3.5.9.2 Biosynthesis (Fig. 3.5.9-1)

Bile acids are formed by classical and alternative biosynthetic pathways.

The classical (neutral) pathway operates entirely in the liver and comprises a cascade of fourteen steps catalyzed by enzymes located in the cytoplasm, microsomes, mitochondria and peroxisomes In the endoplasmic reticulum of liver in rats and likely in humans, cholic acid formation starts with hydroxylation at the 7-position of cholesterol, using the P-450 system (3.5.4.1). This is the rate-limiting step. Then the 3β-hydroxy-Δ^5 steroid is converted to the 3α-hydroxy-5β configuration. Small quantites of the 5α (*allo*) configuration are also formed, which constitute about 1 … 5 % of mammalian bile acids (Fig. 3.5.9-2). After hydroxylation at C-12 and at C-26 in mitochondria (in humans), oxidation to the carboxylic acid level takes place.

The removal of a part of the side chain possibly occurs in peroxisomes. A hydroxyl group is introduced at C-24; then in a reaction resembling the fatty acid degradation, propionyl-CoA and choloyl-CoA are formed. Propionyl-CoA enters a reaction sequence, which leads via methylmalonyl-CoA and succinyl-CoA to the citrate cycle (Fig. 3.2.5-2). Choloyl-CoA reacts with taurine or glycine to the respective conjugates. If the 12-hydroxylation does not take place, the reactions lead to chenodeoxycholyl- CoA and its conjugates.

In the alternative (acidic) pathway (present in all tissues) side-chain oxidation precedes steroid ring modification. Various oxysterols produced by several cell types can be converted into bile acids. The production of these oxysterols is catalyzed by several sterol hydroxylases (e.g., 24-hydroxylase, 25-hydroxylase, and 27-hydroxylase). Oxidation to the carboxyl function takes place early in the sequence, while the isomerization of the C-3 hydroxyl occurs later.

The contribution of the classic and alternative pathways to overall bile acid synthesis is unclear. Whereas the classic pathway appears to be the main pathway, since it is strongly regulated under physiological conditions, the alternative pathway may become the major pathway in patients with liver diseases.

All bile acids, which are formed in the organism directly from cholesterol, are named primary bile acids (in humans mostly cholate, chenodeoxycholate and the corresponding *allo*-bile acids). Some intestinal bacteria are able to metabolize them into so-called secondary bile acids (in humans primarily deoxycholate, lithocholate and the corresponding *allo*-bile acids).

Bile acids are subject to enterohepatic circulation by secretion via the gall bladder into the small intestine (ca. 20 … 30 g/day in humans) and reabsorption in the ileum (Fig. 6.2-2). In this way, about 90 % return via the portal blood to the liver. The rest is excreted in the feces.

3.5.9.3 Regulation of Biosynthesis (Fig. 3.5.9-3)

The bile acid formation is regulated at the 7α-hydroxylase (= monooxygenase) step. This is coordinated with the regulation of the hydroxymethylglutaryl-CoA reductase (the key enzyme of cholesterol synthesis, 3.5.1.4). Both enzymes are closely associated with each other in the endoplasmic reticulum. High cholesterol concentrations stimulate the hydroxylase directly, while mevalonate, thyroid hormone, glucocorticoids and vitamin C increase its synthesis. The mRNA concentration and the enzyme activity follow a diurnal rhythm, which is in context with relatively short half-life times of mRNA and hydroxylase protein (a few hours).

On the other hand, high bile acid concentrations repress the synthesis of both 7α-hydroxylase and HMG reductase. A possible regulation of the 7α-hydroxylase by phosphorylation (activation) and dephosphorylation (deactivation) is under discussion.

Thyroid hormone also promotes the 26-hydroxylation (experiments in rats). Since a 26-OH group prevents consecutive hydroxylation at the 12-position, the ratio of chendeoxycholate / cholate rises. Bile acid and cholesterol homeostasis is also maintained by specific transporters expressed in the liver and the intestine, controlling the enterohepatic circulation of bile acids.

3.5.9.4 Medical Aspects

A shift in the concentration ratio of cholesterol, bile acids and lecithin in the bile causes the formation of gallstones. Splitting of conjugates by bacterial infection of the jejunum leads to premature bile acid reabsorption and lack of conjugates in the small intestine. This causes a decrease of lipid resorption. If the intestinal pH value is lowered to such an extent, that no dissociation of bile acid conjugates takes place, similar deficits in lipid absorption occur (Zollinger-Ellison's syndrome). In Zellweger's syndrome, no peroxisomes are formed. Since no fatty acid oxidation is possible, 3α, 7α, 12α-trihydroxy-5β-cholestanoate accumulates. Hydrophobic bile-acids can induce cell apoptosis and thus contribute to liver injury during cholestasis.

The enzymes, nuclear receptors, and transcription factors involved in bile acid biosynthesis are potential pharmaceutical targets to control diseases of disturbed cholesterol homeostasis (e.g., hypercholesterolemia and atherosclerosis).

Literature:

Agellon, L. B. in Vance, D.E. and Vance, J.E. (eds.) *Biochemistry of Lipids, Lipoproteins and Membranes* (Fifth Edition). San Diego: Elsevier, p. 423–440.

Alrefai, W. and Gill, R. Pharmaceutical Research 2007;24:1803–1823.

Chiang, J.Y.L. Journal of Hepatology 2004;40:539–551.

Hofmann, A. and Hagey, L. Cellular and Molecular Life Sciences (CMLS) 2008;65:2461–2483.

Norlin, M. and Wikvall, K. Curr. Mol Med. 2007;7(2):199–218.

Rust, C. *et al.* J. Biol. Chem. 2009;284:2908–2916.

3.6 Nucleotides and Nucleosides

Röbbe Wünschiers

Nucleosides are composed of a purine or pyrimidine base linked to ribose or 2′deoxyribose (ribo- and deoxyribonucleosides, respectively). Nucleotides are usually phosphorylated at the 5′position of the pentose moiety. In a wider sense, structures with other N-containing ring systems are also named nucleotides, e.g., nicotinamide (3.7.9) or flavin derivatives (3.7.3). Nucleoside di- and triphosphates carry additional phosphate residues connected by pyrophosphate-type bonds. In cyclic nucleotides, the single phosphate forms a diester structure with two hydroxyl groups (mostly 3′, 5′) of the pentose. Dinucleotides are two nucleotide residues linked by a pyrophosphate bond, e.g., NAD^+ (3.7.9) or FAD (3.7.3). Figure 3.6.1-1 shows the structures.

Nucleotides fulfill many important tasks:

- ATP is the universal transmitter of energy in biological systems (3.6.1.3).
- Nucleotides are involved in the regulation of metabolic processes. Examples are cAMP (7.4.2) and cGMP (7.8). Also the ratio of the ATP/ADP/AMP levels exerts a control function.

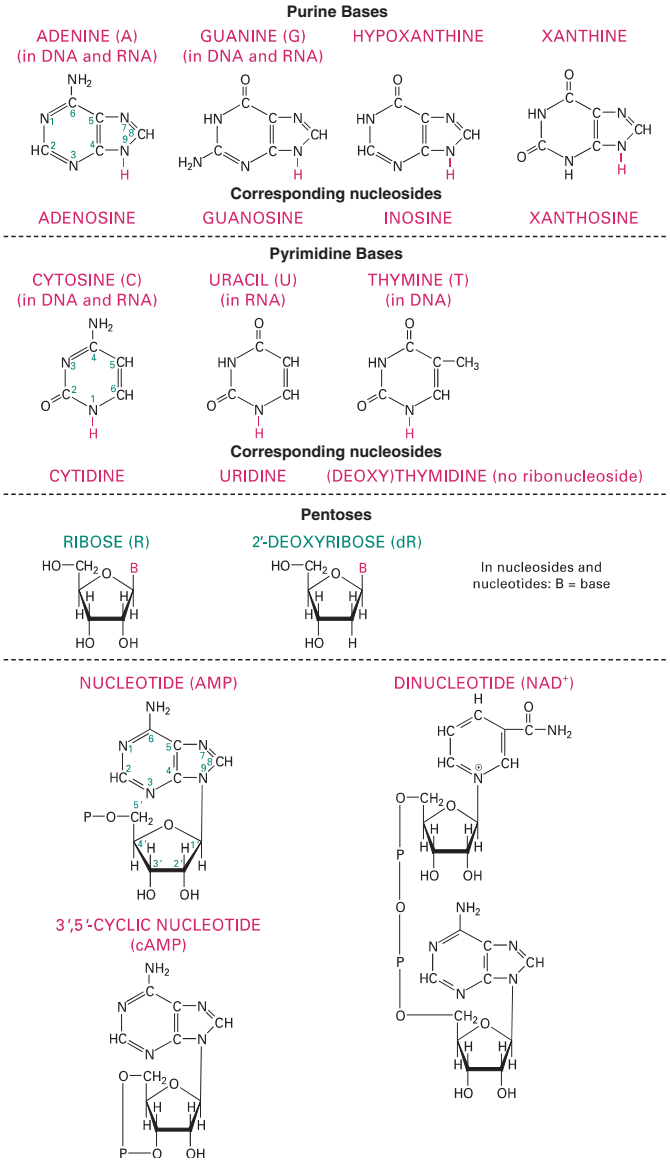


Figure 3.6.1-1. Structure of (Nucleo-) Bases, Nucleosides and Nucleotides

- Nucleotide derivatives represent the activated forms of many compounds during biosynthetic reactions (e.g., UDP-glucose, 3.1.2.2, CDP-choline, 3.4.3.2, phosphoadenylylsulfate, 3.2.5.6).
- Nucleotides are components of the coenzymes NAD^+ and $NADP^+$ (3.7.9), FAD and FMN (3.7.3), CoA (3.7.7) and serve as direct precursors for the synthesis of B-class vitamins like riboflavin, thiamine and folates.
- Ribonucleotides and deoxyribonucleotides are the building blocks of nucleic acids (3.8, 3.9).

The biosyntheses of the bases are multistep processes. However, there are effective ‘salvage pathways’ permitting the recycling of the bases.

3.6.1 Purine Nucleotides and Nucleosides

3.6.1.1 Biosynthesis of Inosine 5′-Phosphate (IMP, Fig. 3.6.1-2)

The ribose component of purine ribonucleotides is derived from ribose 5-P, an intermediate of the pentose phosphate cycle (3.1.6.1). The atoms of the base moiety are contributed by many compounds (Fig. 3.6.1-3). They are added stepwise to the preformed ribose. Striking interrelationships exist with the pathway for histidine synthesis (3.2.8.1).

Ribose 5-P, synthesized in the pentose phosphate pathway, is activated by an unusual transfer of a pyrophosphate group, resulting in 5-phosphoribosyl-1-pyrophosphate (PRPP). Then an amino group from glutamine displaces the pyrophosphate group and yields 5-phosphoribosylamine. Hydrolysis of the released pyrophosphate pushes the reaction. This step also involves a conversion of the ribose from the α to the β anomeric form (1.2.1). The enzymes catalyzing the initial and the second reactions, ribose-P pyrophosphokinase and amido-P ribosyltransferase are regulated.

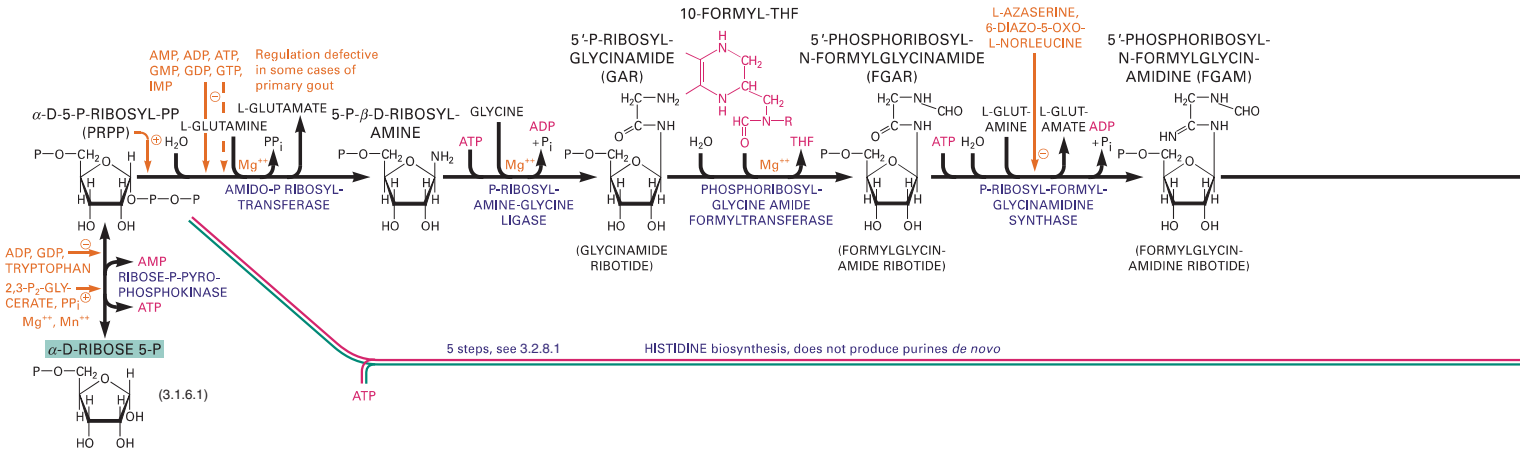
The first of these, which also takes part in the histidine (3.2.8.1), tryptophan (3.2.7.1) and pyrimidine (3.6.2.1) synthesis, is feedback inhibited by ADP and GDP as well as by tryptophan. The second one is the committed (unambiguous) step for purine synthesis and is likewise feedback inhibited by purine nucleotides (Fig. 3.6.1-5). Two independent, but synergistic inhibitory sites exist, one of them binding AMP (preferred), ADP and ATP, the other binding GMP (preferred), GDP and GTP. Also, a feed forward activation by PRPP takes place at this step.

In the following reaction, which is energized by ATP hydrolysis via an acylphosphate intermediate, glycine is added. Then N^{10}-formyltetrahydrofolate (3.7.6.2) contributes a formyl group. In a glutamine and ATP dependent reaction, the oxo group at C-2 is replaced by an imino group. ATP energizes the consecutive closure and aromatization of the imidazole ring, resulting in 1(5′-phosphoribosyl)-5-aminoimidazole (AIR). An unusual carboxylation reaction, which does not use biotin, adds CO_2 and yields the carboxy derivative. This compound forms an amide bond with aspartate, followed by elimination of the aspartate carbon skeleton as fumarate. Thus, aspartate contributes only its amino group to the formation of 1(5′-phosphoribosyl)-5-amino-4-imidazolecarboxamide (AICAR, compare the aspartate cycle, 3.2.9.1). Another formylation by N^{10}-formyltetrahydrofolate contributes the last member of the C_6 ring. By dehydration, this ring is closed and inosine 5′ phosphate is obtained.

In *E. coli*, several of the genes of purine biosynthesis are combined in monocistronic operons. The P-ribosylaminoimidazolecarboxamide formyltransferase and the IMP cyclohydrolase functions are combined on a single peptide chain.

In vertebrates, even more multifunctional enzymes exist:

- The phosphoribosylglycinamide formyltransferase (GART) gene encodes a trifunctional protein carrying out 3 steps of purine biosynthesis: P-ribosylamine-glycine ligase (GARS)/ P-ribosylglycineamide formyltransferase (also abbreviated as GART) / P-ribosylformylglycineamidine cycloligase (AIRS) and a smaller protein containing the GARS domain of GART as a functional protein. The GART gene is located on human chromosome 21 and may be involved in the phenotype of Down syndrome
- The chain encoded by the AIRC gene contains the P-ribosylaminoimidazole carboxylase and the P-ribosylaminoimidazole-succinocarboxamide synthase functions.



- The chain encoded by the IMPS gene contains the P-ribosyl-aminoimidazole-carboxamide formyltransferase and the IMP cyclohydrolase functions.

It has been found that, at least in *E. coli*, the 'P-ribosylaminoimidazole carboxylase' consists of two enzyme entities. The 5′-phosphoribosyl-5-carboxyaminoimidazole synthase (NCAIR synthase) performs an ATP-assisted carboxylation of the amino group at C-5, while the NCAIR mutase (or isomerase) shifts it to the C-4 position.

AICAR is released during histidine biosynthesis in plants and bacteria (3.2.8.1). Since, however, its formation requires the ribose residue and part of the purine ring of ATP, this pathway does not lead to a net formation of purines. Rather, the steps from AICAR to purines serve only the reconstitution of ATP consumed before.

For the intermediates of the purine biosynthesis pathway, the terms of the IUB 'Enzyme Nomenclature' are used above. Since other expressions frequently occur in the literature they are given in parentheses in Fig. 3.6.1-2.

3.6.1.2 Interconversions of Purine Ribonucleotides (Fig. 3.6.1-4)

Inosine 5′-monophosphate (IMP) does not accumulate in the cells. It is rapidly converted into adenosine and guanosine mononucleotides.

Adenosine 5′-monophosphate (AMP): Its synthesis involves an amino group transfer from L-aspartate in a two step pathway, which is energized by GTP hydrolysis. The cleavage of the intermediate adenylosuccinate liberates fumarate and yields AMP. This cleavage is performed by the same enzyme, which catalyzes an analogous reaction during IMP biosynthesis (3.6.1.1). Adenosine 5′-diphosphate (ADP) is obtained by a phosphate transfer from adenosine 5′-triphosphate (ATP), catalyzed by adenylate kinase (myokinase).

Besides being involved in the formation of AMP, the aspartate → fumarate interconversion also has the effect of increasing the amount of citrate cycle intermediates. This is of importance in muscle, where only a low level of other anaplerotic enzymes (3.1.3.4) is found. Fumarate formation from aspartate is enhanced by the AMP deaminase (adenylate deaminase) reaction, which converts the newly formed AMP back into IMP for another amination round (purine nucleotide cycle). Inherited myoadenylate deaminase insufficiency leads to cramps after muscle exercise.

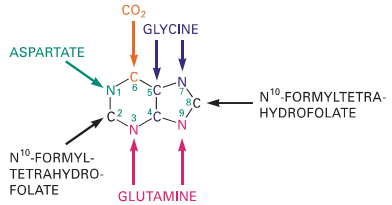


Figure 3.6.1-3. Origin of the Purine Ring Atoms

Guanosine 5′-monophosphate (GMP): An initial dehydrogenation reaction yields xanthosine 5′-phosphate, which then receives an amino group from glutamine. This reaction differs from transamination reactions (Fig. 3.2.2-3) in the absence of pyridoxal phosphate and the participation of ATP, the hydrolysis of which drives the reaction to completion. The use of glutamine as source of the amino group enables animals to avoid a high NH_4^+ level in blood. The bacterial enzyme also accepts NH_3 instead of glutamine.

The phosphate groups for the formation of guanosine 5′-diphosphate (GDP) and guanosine 5′-triphosphate (GTP) are supplied by ATP. While guanylate kinase is a highly specific enzyme, the diphosphate kinase is quite unspecific and accepts many ribo- and deoxyribo- di- and triphosphates.

Regulation of the purine nucleotide biosynthesis (Fig. 3.6.1-5): In animals, the first individual reactions leading to the formation of AMP and GMP are catalyzed by adenylosuccinate synthase and IMP dehydrogenase. These enzymes are feedback inhibited (competitively to IMP) by the end products of the pathway. On the other hand, GTP contributes to the formation of AMP and likewise ATP to the formation of GMP. By this reciprocal relationship, both individual pathways are coordinated, while the overall supply of the common precursor IMP is regulated by the combined action of adenosine and guanosine nucleotides. In *E. coli*, the purine repressor purR regulates the expression of all genes for purine and pyrimidine biosynthesis. Hypoxanthine and guanine serve as corepressors.

Salvage reactions: During degradation of RNA besides mononucleotides (3.8.3), nucleosides and free bases are also obtained (for reactions, see Fig. 3.6.1-4). This formation of non-phosphorylated compounds is essential during digestion for resorption purposes, but also occurs during normal cellular metabolism. It is a prerequisite for the membrane passage, which proceeds in most cases through a common permease protein via facilitated diffusion (the nucleoside is phosphorylated after membrane passage and thus removed from the diffusion equilibrium, compare 3.10.3).

The purine bases can be reconverted into their respective mononucleotides by reaction with 5-P-ribosyl-PP (PRPP, 3.6.1.1, 'salvage pathway'). This is catalyzed by adenine P-ribosyl transferase and by hypoxanthine P-ribosyl transferase (HGPRT, which reacts with guanine and with hypoxanthine, the base moiety of IMP). For the phosphorylation of nucleosides, there are several kinases with different specificity.

3.6.1.3 ATP and Conservation of Energy

Chemical reactions proceed spontaneously only if there is a decrease in free energy ($\Delta G < 0$, see 1.5.1). Therefore, necessary endergonic reactions of living beings (e.g., biosynthesis, movements, signal transfer) have to be 'pulled' by coupling to exergonic reactions, which usually employ 'high energy' compounds. The most versatile of them is adenosine triphosphate (ATP).

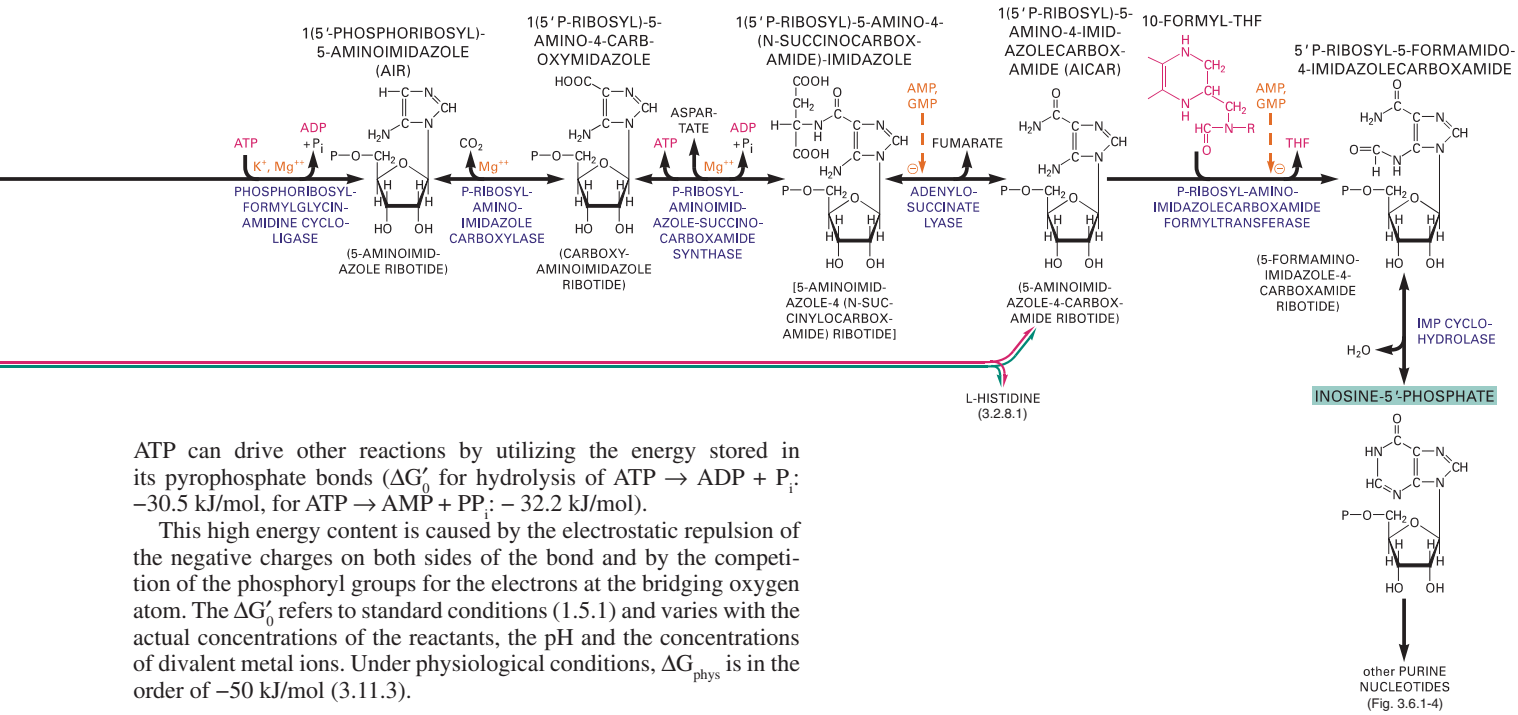


Figure 3.6.1-2. Synthesis of Purine Ribonucleotides

ATP can drive other reactions by utilizing the energy stored in its pyrophosphate bonds ($\Delta G'_0$ for hydrolysis of ATP → ADP + P_i: −30.5 kJ/mol, for ATP → AMP + PP_i: − 32.2 kJ/mol).

This high energy content is caused by the electrostatic repulsion of the negative charges on both sides of the bond and by the competition of the phosphoryl groups for the electrons at the bridging oxygen atom. The $\Delta G'_0$ refers to standard conditions (1.5.1) and varies with the actual concentrations of the reactants, the pH and the concentrations of divalent metal ions. Under physiological conditions, ΔG_{phys} is in the order of −50 kJ/mol (3.11.3).

ATP is obtained

- by substrate level phosphorylation. This involves the transfer of phosphate from another energy rich bond to ADP, which can proceed aerobically as well as anaerobically. Examples are the reactions of pyruvate kinase (3.1.1.1) and bacterial fermentations (3.10.5).
- by phosphorylation of ADP via the H^+ -transporting ATP synthase (3.11.4.5), which is driven by a proton gradient. In many cases, this gradient is established by aerobic oxidation of substrates (3.11.4), but also anaerobic mechanisms occur (3.10.6).

The metabolic role of ATP consists mainly of the transfer of phosphate or of the adenylyl group to other compounds. This can effect an increase in reactivity. Example: fatty acid activation (3.4.1.4). The phosphorylation of proteins can have signalling purposes, e.g., in receptor-tyrosine kinase cascades (7.5.3). Hydrolysis of protein-bound ATP can effect changes in the protein configuration. Examples: muscle contraction, 7.4.5; ion transport by Na^+/K^+ exchanging ATPase, 6.1.4.

In plants, adenine nucleotides, as the major nucleotides of energy metabolism, have significant pools in at least three cellular compartments; e.g. in wheat leaves approximately 45 % of adenine nucleotides is found in the plastid, 46 % in the cytosol, and 9 % in the mitochondria.

3.6.1.4 Ribonucleotide Reduction to Deoxyribonucleotides (Fig. 3.6.1-6)

The conversion of the purine diphosphoribonucleotides, as well as of pyrimidine diphosphoribonucleotides (NDP) into the respective diphospho-2′deoxyribonucleotides (dNDP) is catalyzed by ribonucleoside-diphosphate reductase according to the general formula

$$NDP + protein(SH)_2 = dNDP + protein(SS) + H_2O$$

where the dithiol $(-SH)_2$ is converted to a disulfide (-S-S-). After phosphorylation, the dNTPs are used for incorporation into DNA. Control mechanisms acting on the reductase and the DNA polymerase keep the dNTP pools small. By means of additional regulation systems (Fig. 3.6.1-7), the optimum concentration ratios of all four deoxyribonucleotides have to be kept at a certain level, since this is a prerequisite for a high fidelity of DNA replication.

Three main classes of ribonucleoside-diphosphate reductases exist depending on different metal cofactors for the catalytic activity:

- class I enzymes comprise a di-iron-oxygen cluster,
- class II enzymes contain a cobalt containing cobalamin cofactor (vitamin B_{12}),
- class III enzymes contain a 4Fe-4S iron-sulfur cluster coupled to S-adenosylmethionine (SAM).

The *E. coli* ribonucleoside-diphosphate reductase (a class I enzyme) is composed of 2 dimers, α_2 (2 * 86 kDa) and β_2 (2 * 43 kDa). Each α subunit contains several -SH groups and 2 allosteric regulatory binding sites, while each β subunit contains a stable tyrosyl radical Y• with an unpaired electron, which is generated by interaction with an unusual iron-oxygen center Fe(III)-O^{2-}-Fe(III).

The reductase reaction takes place by a radical mechanism at the interface between the α and β subunits (Fig. 3.6.1-6). The free radical of the enzyme abstracts (probably indirectly) the hydrogen from the 3′ position of the ribose, while a protonation from an enzyme -SH group leads to the removal of H_2O, leaving a cation radical at the 2′ position. This intermediate is reduced by a sulfhydryl group of the enzyme. Then hydrogen is returned to the 3′ position. This completes the formation of the deoxyribonucleotide. The sulfhydryl groups of the enzyme, which were oxidized during this process are reduced again by thioredoxin (12 kDa, contains 2 -SH groups in close proximity), which, in turn, is reduced by NADPH. (Thioredoxin is also involved in enzyme regulatory processes, e.g., 3.12.2). In a number of bacteria, thioredoxin is replaced by glutaredoxin, which thereafter is reduced by glutathione (3.2.5.7).

The enzyme activity is allosterically regulated at two levels (Fig. 3.6.1-7, p. 129): Binding of dATP (acting as a representative of deoxynucleotides) at the general regulatory site inhibits the overall activity. This is counteracted by ATP. This interplay responds to the needs of, e.g., DNA synthesis during cell division. On the other hand, binding of various deoxynucleotides regulatory site serves the coordination of their production rates.

The expression of this enzyme in eukarya varies during the cell cycle. It increases greatly when cells leave the G_0 and enter the S-phase (Fig. 3.9.1-2). As a consequence, the dNTP pools increase likewise.

Class II enzymes are found in bacteria (and in some of their phages) that can live under both aerobic and anaerobic conditions, and use a 5′deoxycobalamin cofactor (3.7.5.2) for radical generation.

Class III enzymes exist in strict, or facultative, anaerobic bacteria and some bacteriophages. The bacteriophage T4 enzyme comprises two subunits with a larger α2 subunit containing the active and allosteric sites. The smaller β2 subunit ('activase') carries a 4Fe-4S iron-sulfur cluster that reacts with SAM.

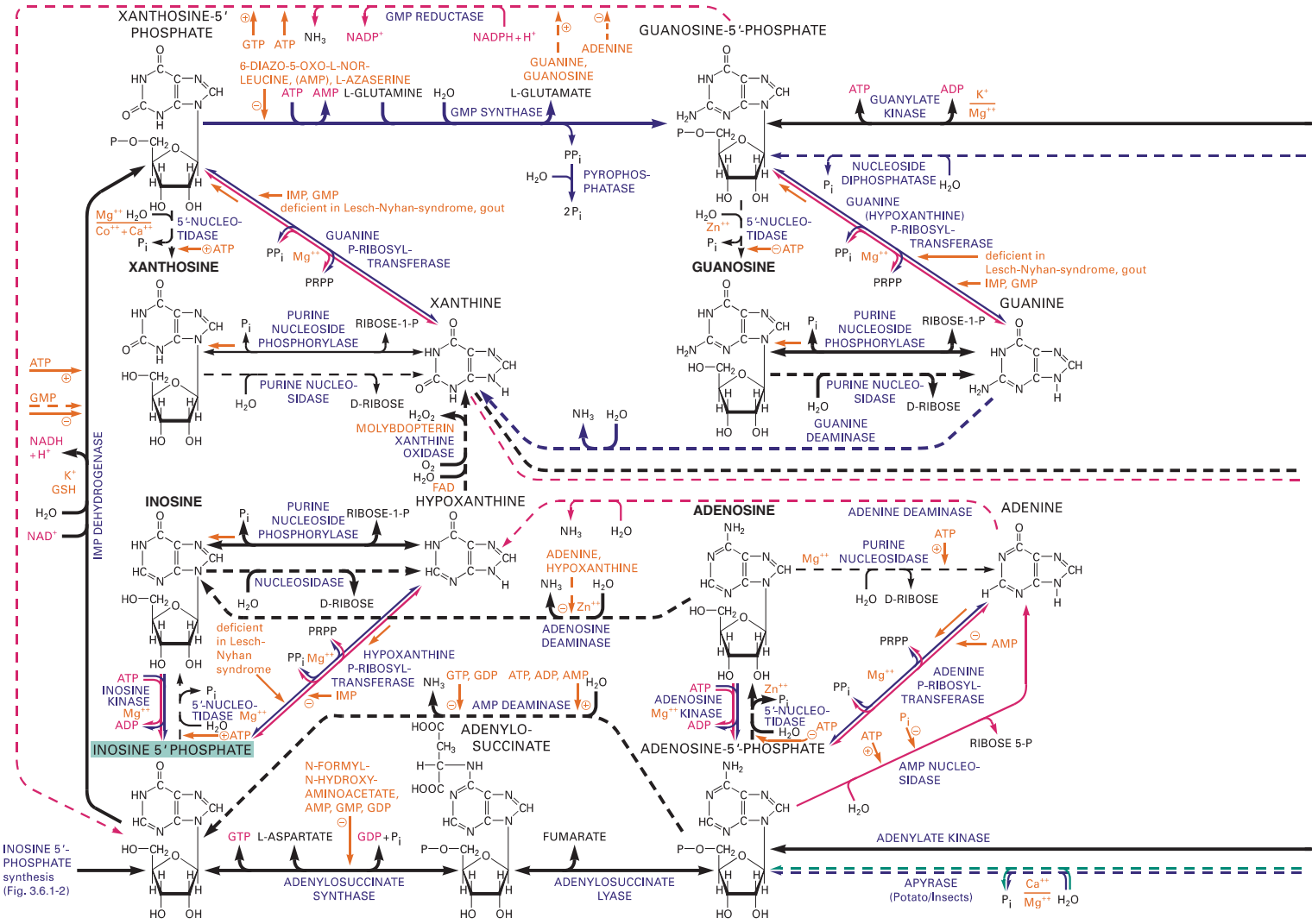


Figure 3.6.1-4. Interconversions and Degradation of Purine Ribonucleotides

3.6.1.5 Interconversions and Degradation of Purine Deoxyribonucleotides (Fig. 3.6.1-8)

Phosphorylation and degradation follow a similar pattern to the ribonucleotide analogues. In some cases, even the same enzymes accept ribo- and deoxyribo-substrates. There are also a number of base exchange reactions both at the nucleoside and mononucleotide level.

3.6.1.6 Catabolism of Bases (Figs. 3.6.1-4 and 3.6.1-8)

Purine nucleosides and nucleotides to be degraded are at first cleaved, releasing the free bases (adenosine and AMP are previously deaminated to inosine and IMP, respectively). The bases hypoxanthine and xanthine are oxidized to urate by xanthine oxidase. This is the final product of purine degradation in a number of animals. In blood, urate is transported in protein-bound form. Primates excrete it in the urine.

Xanthine oxidase is a homodimeric enzyme (2 *130 kDa), which contains molybdenum, FAD and 2 FeS clusters. In mammals, it is primarily present in the liver and in the small intestinal mucosa. The probable reaction mechanism is shown in Fig. 3.6.1-9 (p. 129). It starts by a nucleophilic attack of an enzyme residue on the C-8 position of xanthine and elimination of H as a hydride ion. This reduces the enzyme-bound Mo(VI) to Mo(IV). Water displaces the enzyme in the substrate complex and releases urate. Mo(IV) is oxidized by O_2. The H_2O_2 resulting from this reaction is decomposed by catalase.

Urate is a very effective antioxidant. It destroys reactive oxygen species ($HO^{\bullet}$, $O_2^{\bullet-}$, $^1O_2$, see 3.2.5.8) and reduces oxidized heme iron. In these respects, it is about as effective as ascorbate.

In terrestrial reptiles, birds and many insects (uricotelic animals), urate is also involved in the excretion of amino acid nitrogen, which ends up in this compound via purine intermediates. Due to its low solubility, it can be excreted as a paste of crystals and saves water this way. Non-primate mammals and other animals (except the uricotelic animals) degrade urate further by action of the urate oxidase (uricase) and excrete allantoate, urea or NH_3 (Fig. 3.6.1-4).

In higher plants, purine nucleotides are oxidatively degraded via urate and allantoin to CO_2 and NH_3, which is then reassimilated by the glutamine oxoglutarate aminotransferase.

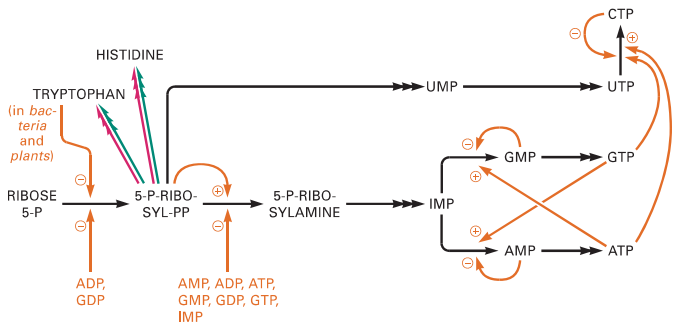


Figure 3.6.1-5. Regulation of Purine Nucleotide Biosynthesis

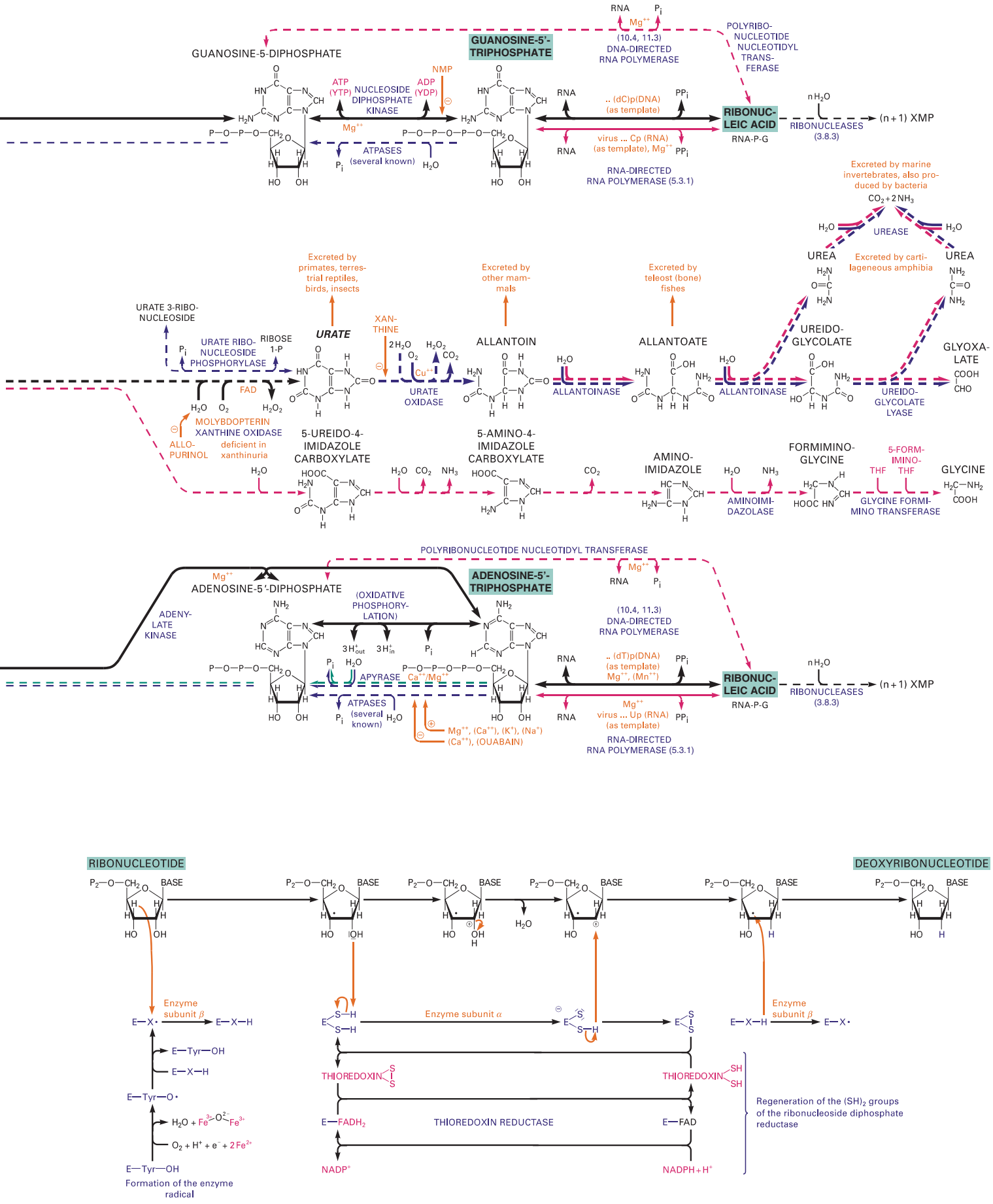


Figure 3.6.1-6. Reaction Mechanism of the Ribonucleoside-Diphosphate Reductase from *E. coli*

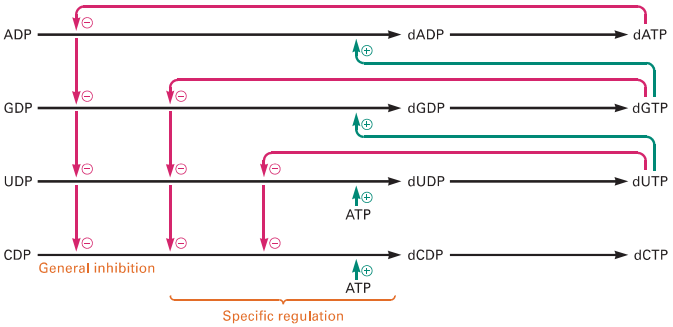


Figure 3.6.1-7. Regulation of Ribonucleoside-Diphosphate Reductase from *E. coli* (red: inhibition, green: activation)

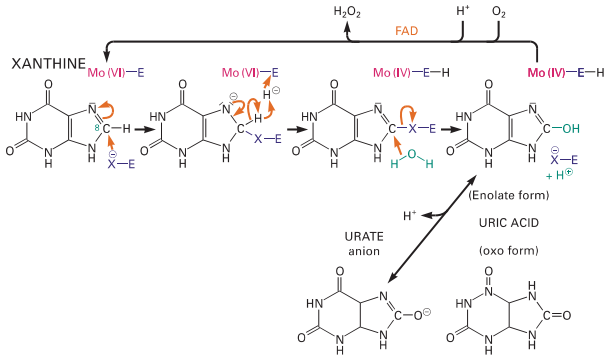


Figure 3.6.1-9. Reaction Mechanism of Xanthine Oxidase

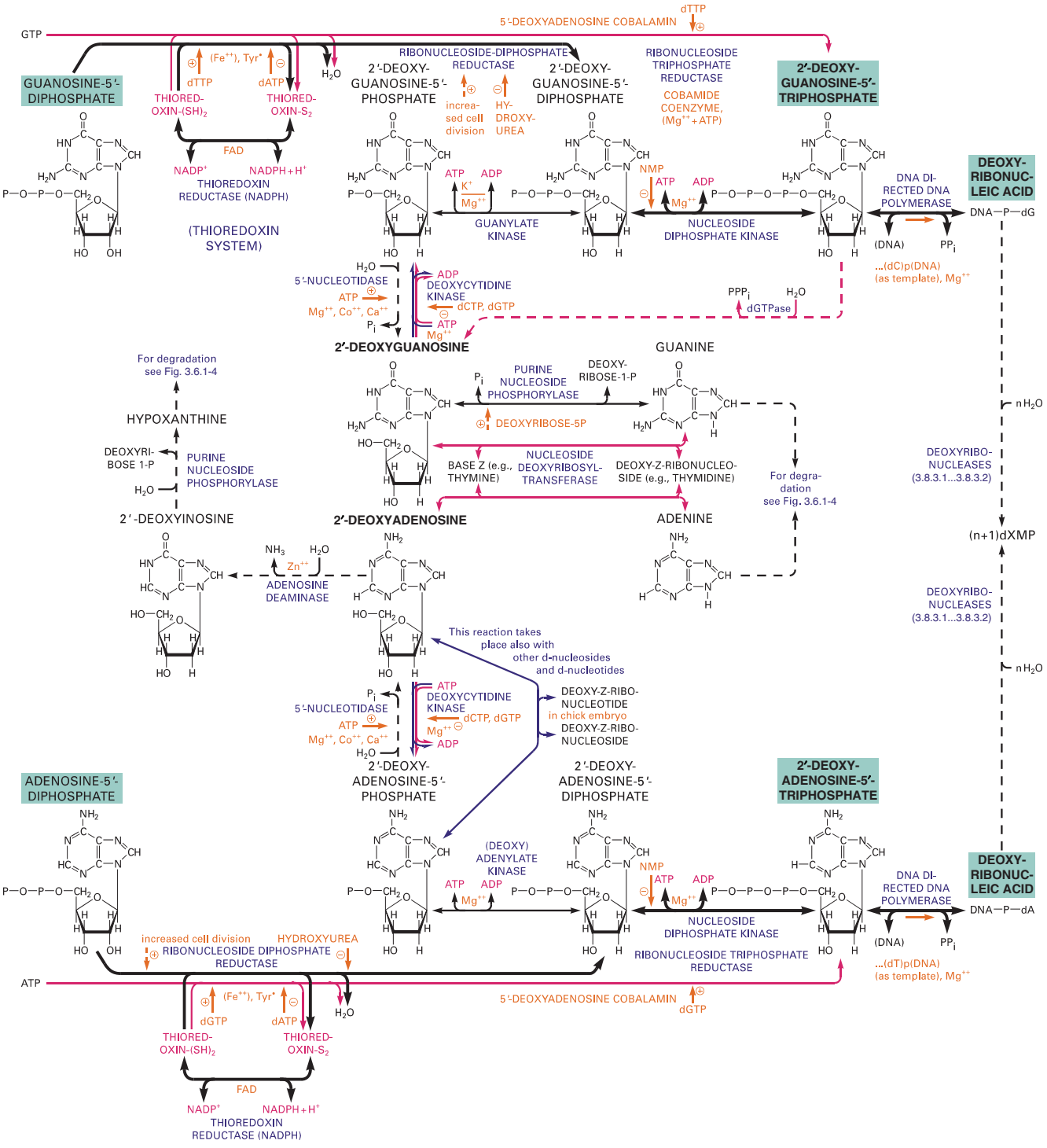


Figure 3.6.1-8. Interconversions and Degradation of Purine Deoxyribonucleotides

3.6.1.7 Medical Aspects

A severe congenital deficiency of hypoxanthine P-ribosyl transferase (3.6.1.2, Lesch-Nyhan syndrome) leads to an accumulation of PRPP, since it is insufficiently consumed by the salvage pathway reactions. Since this compound activates the purine biosynthesis (3.6.1.1) in a feed-forward way, excessive quantities of purines and consecutively their degradation product urate are formed. This causes gout symptoms (see below) and (possibly in combination with additional reactions) mental retardation, aggressive behavior etc.

Health problems arise from the low solubility of urate. While uric acid is almost insoluble, urate (the prevalent form at physiological pH, pK = 5.4) is only a little soluble. In blood, most of it is bound to proteins. Hyperuricemia (beyond the normal concentration of ca. 0.4 mmol/l) leads to deposition of sodium urate crystals in joints, tendon sheaths, renal medulla, etc. (gout).

- Hereditary primary hyperuricemia is caused by
  - disturbed tubular secretion of urate (mostly) or by
  - elevated production of this compound (by the Lesch-Nyhan syndrome or by defective feedback inhibition of the amido-P ribosyltransferase, etc.).
- Secondary hyperuricemia is caused, e.g., by elevated degradation of nucleic acids (leukemia, psoriasis).
- To some extent, meat-rich food may also contribute to the formation of gout.

One way of treatment is the inhibition of xanthine oxidase by allopurinol (an analog of hypoxanthine with interchanged C-8 and N-7). The more soluble substrates of xanthine oxidase pile up and are excreted.

A hereditary disease caused by defective adenosine deaminase is the severe combined immunodeficiency disease (SCID). Deoxyadenosine (and likewise adenosine, 3.6.1.6) accumulates and is phosphorylated, yielding dATP (up to 50 times the normal concentration), which is an allosteric inhibitor of ribonucleotide reductase (3.6.1.4). The inhibition of this enzyme leads to a lack of deoxynucleotides and to impaired DNA synthesis. This affects mostly the lymphocyte generation and causes severe malfunctions of B and T cells (8.1.3). This disease is lethal unless treated by administration of adenosine deaminase, which is protected against degradation by coupling to polyethylene glycol.

An imbalanced accumulation of dATP or dGTP occurs in some immune diseases. This leads according to the mutual regulation scheme of ribonucleoside-diphosphate reductase (Fig. 3.6.1-8) to a depletion of the other dNTPs, resulting in impaired DNA synthesis in S phase cells and insufficient DNA repair in resting lymphocytes, which results in cell death.

3.6.2 Pyrimidine Nucleotides and Nucleosides

3.6.2.1 Biosynthesis of Uridine 5′-Phosphate (UMP, Fig. 3.6.2-1)

While in purine biosynthesis the components of the purine ring are assembled step by step at a preformed ribose moiety (3.6.1.1), during pyrimidine formation the base ring is formed first and ribose is added afterwards.

Aspartate and carbamoyl phosphate are the source of the ring atoms. In eukarya, the latter compound is also a member of the urea cycle (3.2.9.1), but both pathways are separated by compartmentization. While the carbamoyl-P formation and the consecutive reactions for pyrimidine synthesis take place in the cytosol, the respective reaction of the urea cycle is a mitochondrial one. Also, carbamoyl synthase II involved in pyrimidine biosynthesis has different properties to the urea cycle enzyme: it uses glutamine as a nitrogen source (instead of ammonia) and it is not allosterically activated by N-acetyl glutamate. In bacteria, it is the only carbamoyl synthase present.

The next reaction, which is catalyzed by aspartate carbamoyl transferase (aspartate-carbamyl transferase, ATCase) represents the committed step of pyrimidine biosynthesis and is the subject of multiple regulation mechanisms.

The allosteric *E. coli* enzyme is composed of 2 homotrimeric catalytic (c_3) and 3 homodimeric regulatory (r_2) subunits, which result in the structure c_6r_6 with threefold rotational symmetry. The reaction takes place at the interface between two c-chains. Substrate binding effects a 10 degree turn and an increase of the distance between the catalytic trimers. Also the r-chains turn. This corresponds to the transition from the T- to the more active R-state (2.5.2). Studies of the catalytic mechanism proved that this transition is in agreement with the symmetry model of allosteric interaction (all subunits change simultaneously from the T- to the R-state).

The *E. coli* aspartate carbamoyl transferase is feedback-inhibited by the pathway end product CTP. The previous enzyme carbamoyl-P synthase is also subject to feedback inhibition, however by UMP.

The next step is the ring closure by release of water, followed by the oxidation to orotate. In eukarya, the dihydroorotate oxidase (EC number 1.3.3.1, also called 'dehydrogenase') is located on the outside of the mitochondrial membrane and donates the reducing equivalents to the quinone pool (3.11.4). In bacteria, the enzyme is likewise membrane-bound and interacts with quinones. (Bacterial NADH or NADPH dependent orotate reductases also exist. The reactions catalyzed by these enzymes proceed mainly in the opposite direction.) The reaction with 5′-phosphoribosyl pyrophosphate (PRPP, compare 3.6.1.1) leads to the addition of the ribose moiety. Hydrolysis of the released PP_i drives this reaction to completion. Decarboxylation of orotidine 5′-P is the last step in the biosynthesis of uridine 5′-P.

Although pyrimidine biosynthesis is well conserved, distinct differences appear between bacteria, yeast, and higher eukarya. In bacteria, each enzymatic step is catalyzed by an individual protein encoded by a different gene. In plants, the first three steps are carried out by three separate proteins, whereas in mammals the first three enzyme functions of this pathway are combined on a single peptide chain (Carbamoylphosphate synthase/aspartate transcarbamoylase/dihydroorotase, CAD protein, 240 kDa). In mammals and plants, the penultimate and the last enzymatic function, orotate P-ribosyltransferase and the orotidine-5′-P decarboxylase, are located on distinct domains of a bi-functional protein called UMP synthase (UMPSase). The catalysis by multifunctional proteins enables a simple coordination of the synthesis and facilitates the substrate transfer from one enzyme function to the next one (analogous to, e.g., 3.4.1.1 and 3.6.1.1).

3.6.2.2 Interconversions of Pyrimidine Ribonucleotides (Fig. 3.6.2-2)

Pyrimidine mononucleotides are converted by phosphorylation reactions to di- and consecutively to triphosphates. Uridine 5′-triphosphate (UTP) is aminated in an ATP-energized reaction by glutamine (in mammals) or by NH_4^+ (in *E. coli*), yielding cytidine 5′-triphosphate (CTP, Fig. 3.6.2-3). The amination reaction proceeds analogously to the conversion of xanthosine 5 phosphate (3.6.1.2), except for the release of P_i instead of PP_i.

A salvage reaction (of uracil) similar to the purine metabolism (3.6.1.2) has been found in bacteria.

3.6.2.3 Ribonucleotide Reduction and Interconversions of Pyrimidine Deoxyribonucleotides (Fig. 3.6.2-2)

The same enzymes that reduce purine ribonucleotides to deoxyribonucleotides also act on pyrimidine nucleotides. These conversions are subject to mutual control by the other nucleotides (3.6.1.4).

In addition to reduction of UDP, deoxyuridine phosphates can also be formed via deoxycytidine phosphates. In *E. coli*, only about 25 % of the dU and consecutively dT compounds (see below) are obtained by UDP reduction, while 75 % are generated via dCDP intermediates. There are two major pathways for this conversion. In bacteria, they proceed via the steps dCDP → dCTP → dUTP (deamination) → dUMP (dephosphorylation to dUMP by deoxyuridine pyrophosphorylase). The deamination is feedback inhibited by dTTP and dUTP. In eukarya (and in bacteriophage T2-infected *E. coli*), the reaction sequence is dCDP → dCMP → dUMP (deamination). The deaminase is inhibited by dTTP and activated by dCTP. There is also a conversion at the nucleoside level dCMP → dC → dU (deamination) → dUMP. The activities of the deaminase enzymes control the ratio dC/dT compounds available for DNA synthesis.

Deoxyuridine nucleotides do not occur in DNA. Their role is taken over by deoxythymidine nucleotides. Deoxythymidine

(Continuation on p.133)

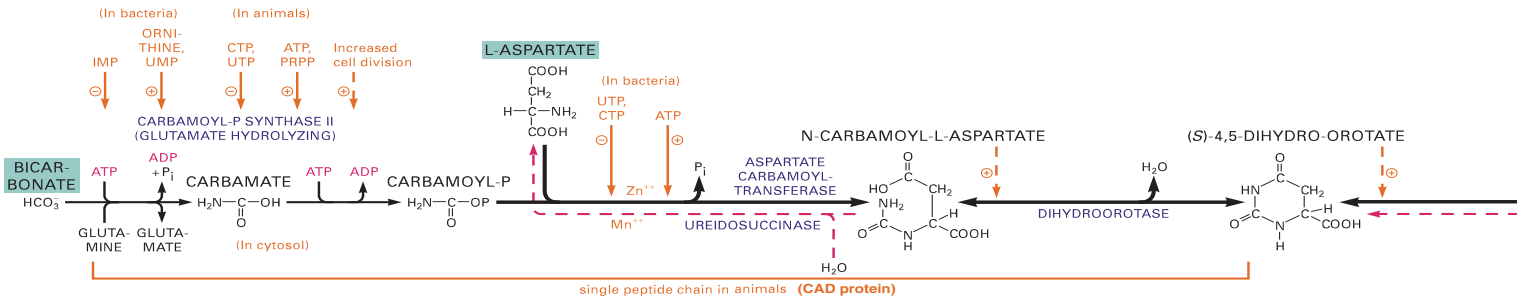


Figure 3.6.2-1. Synthesis of Pyrimidine Ribonucleotides

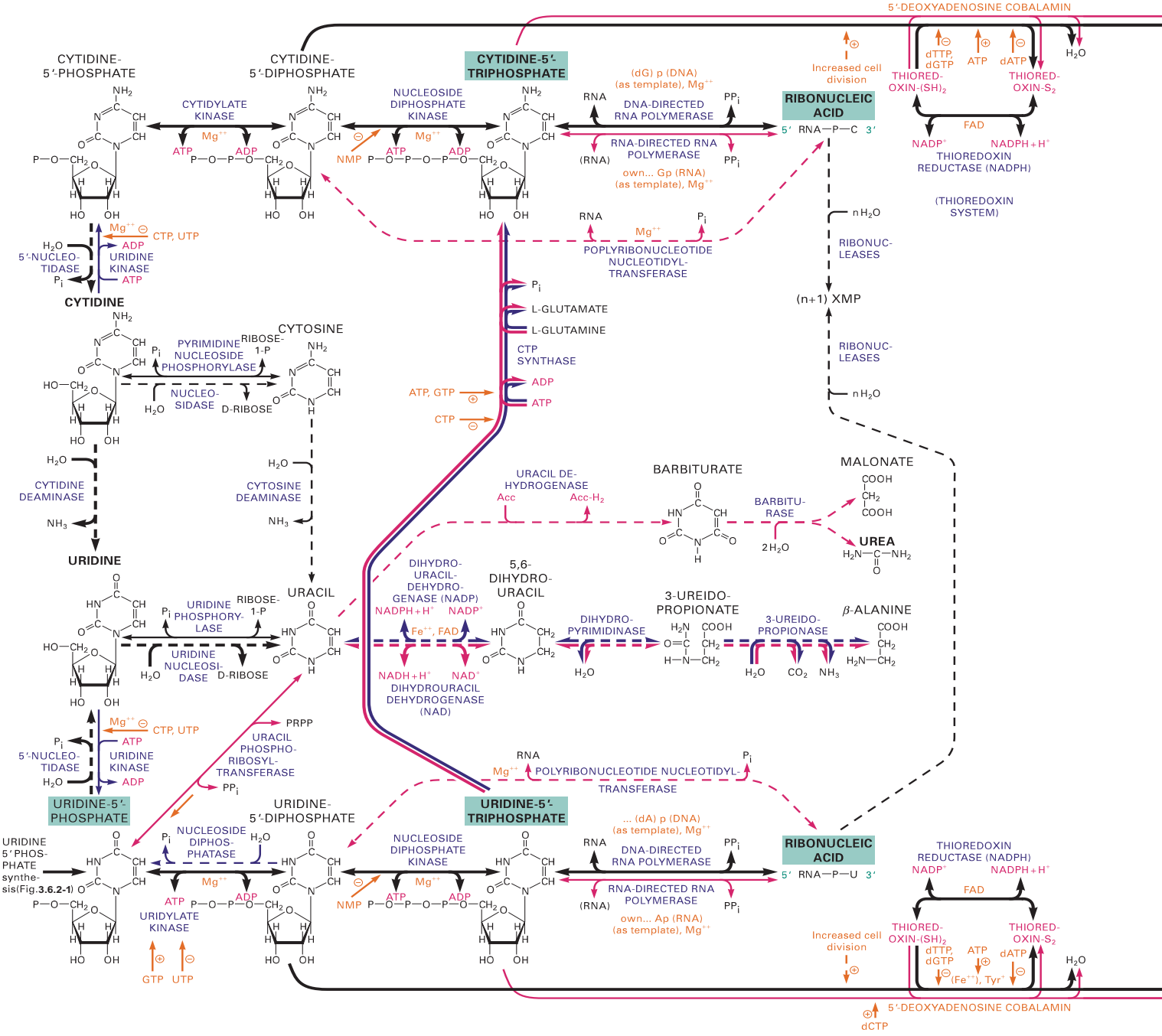
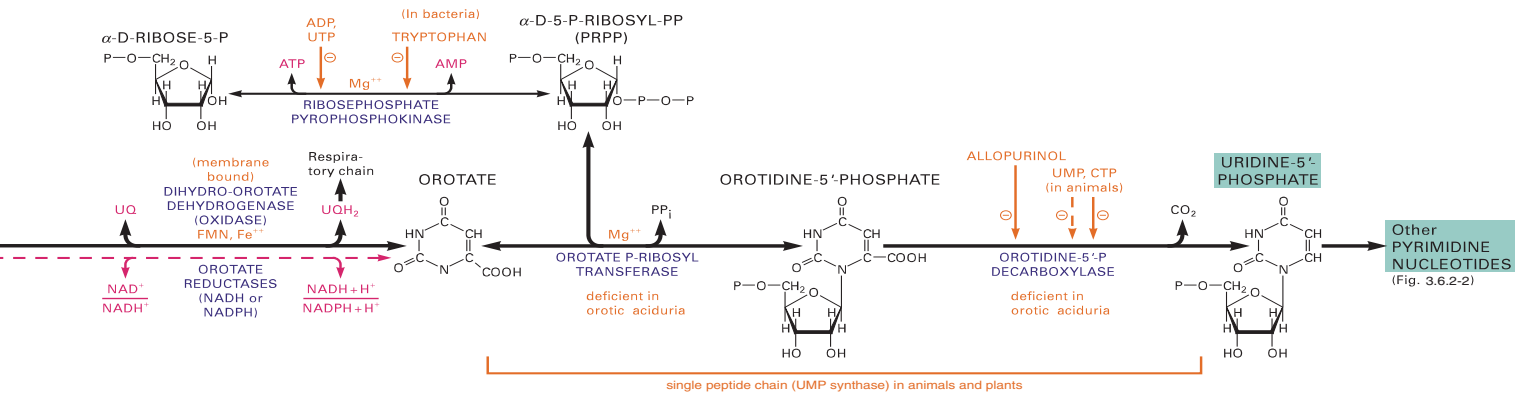
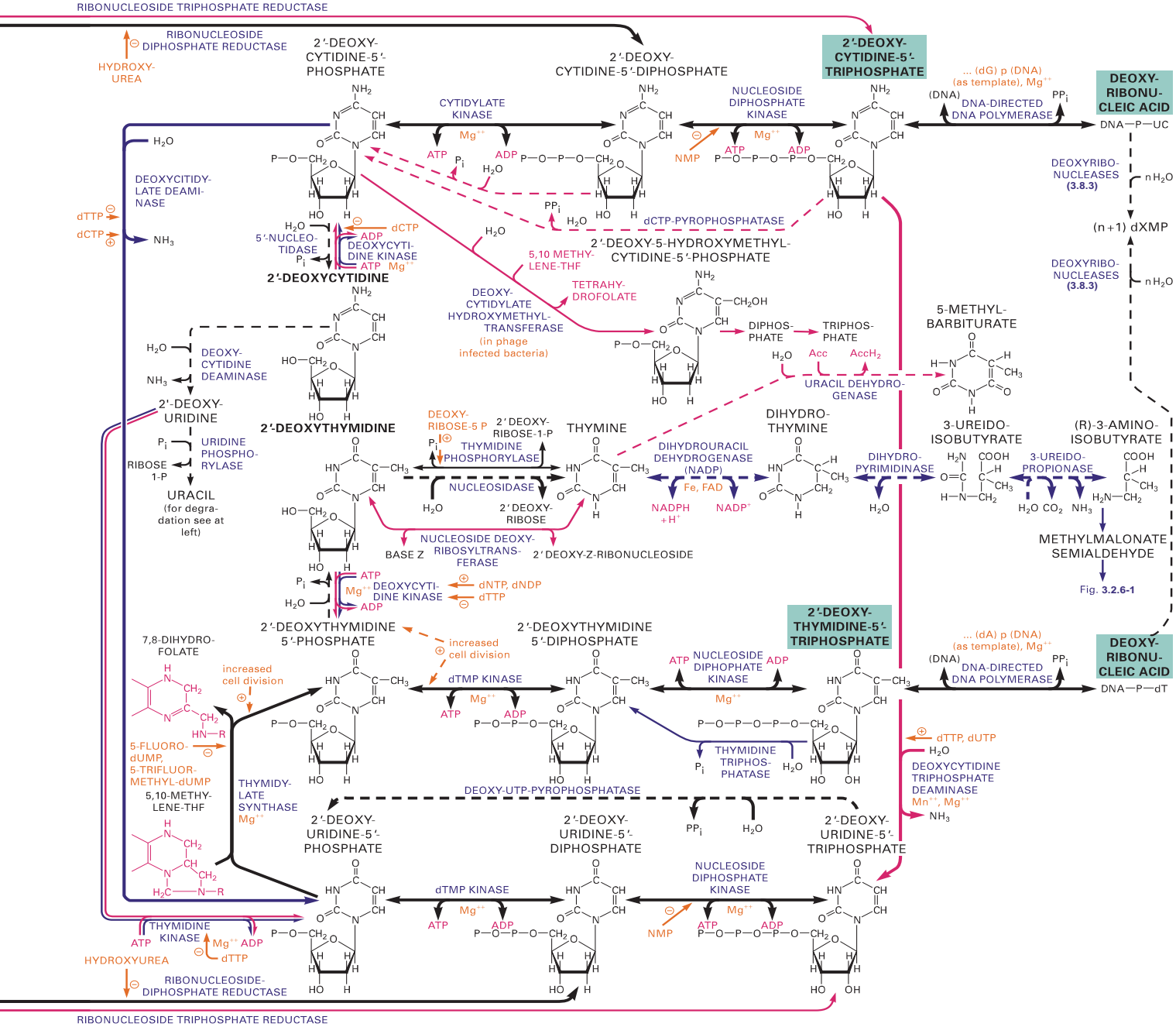


Figure 3.6.2-2. Interconversions and Degradation of Pyrimidine Nucleotides and Bases


α-D-RIBOSE-5-P
ADP, UTP
ATP
Mg++
(In bacteria)
TRYPTOPHAN
AMP
RIBOSEPHOSPHATE PYROPHOSPHOKINASE
α-D-5-P-RIBOSYL-PP (PRPP)
(membrane bound)
DIHYDRO-OROTATE DEHYDROGENASE (OXIDASE)
FMN, Fe++
UQ
UQH2
Respira-tory chain
OROTATE
OROTATE REDUCTASES (NADH or NADPH)
NAD+ NADH
NADH + H+ NADPH + H+
PPi
OROTATE P-RIBOSYL TRANSFERASE
deficient in orotic aciduria
OROTIDINE-5'-PHOSPHATE
ALLOPURINOL
UMP, CTP (in animals)
OROTIDINE-5'-P DECARBOXYLASE
deficient in orotic aciduria
CO2
URIDINE-5'-PHOSPHATE
Other PYRIMIDINE NUCLEOTIDES
(Fig. 3.6.2-2)
single peptide chain (UMP synthase) in animals and plants


RIBONUCLEOSIDE TRIPHOSPHATE REDUCTASE
RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
HYDROXY-UREA
2'-DEOXY-CYTIDINE-5'-PHOSPHATE
2'-DEOXY-CYTIDINE-5'-DIPHOSPHATE
2'-DEOXY-CYTIDINE-5'-TRIPHOSPHATE
CYTIDYLATE KINASE
NUCLEOSIDE DIPHOSPHATE KINASE
NMP
(DNA) DNA-DIRECTED DNA POLYMERASE
... (dG) p (DNA) (as template), Mg++
PPi
DEOXY-RIBONU-CLEIC ACID
DNA—P—UC
DEOXYRIBO-NUCLEASES (3.8.3)
n H2O
(n+1) dXMP
DEOXYCITIDY-LATE DEAMI-NASE
dTTP
dCTP
NH3
5'-NUCLEO-TIDASE
dCTP-PYROPHOSPHATASE
2'-DEOXY-5-HYDROXYMETHYL-CYTIDINE-5'-PHOSPHATE
2'-DEOXYCYTIDINE
5,10 METHY-LENE-THF
TETRAHY-DROFOLATE
DEOXY-CYTIDYLATE HYDROXYMETHYL-TRANSFERASE (in phage infected bacteria)
DIPHOS-PHATE
TRIPHOS-PHATE
5-METHYL-BARBITURATE
Acc
AccH2
URACIL DEHYDRO-GENASE
DEOXY-CYTIDINE DEAMINASE
2'-DEOXY-URIDINE
2'-DEOXYTHYMIDINE
DEOXY-RIBOSE-5 P
2'DEOXY-RIBOSE-1-P
THYMINE
THYMIDINE PHOSPHORYLASE
DIHYDRO-THYMINE
DIHYDROURACIL DEHYDROGENASE (NADP)
Fe, FAD
NADPH + H+
NADP+
DIHYDRO-PYRIMIDINASE
3-UREIDO-ISOBUTYRATE
3-UREIDO-PROPIONASE
(R)-3-AMINO-ISOBUTYRATE
H2O CO2 NH3
METHYLMALONATE SEMIALDEHYDE
Fig. 3.2.6-1
URIDINE PHOSPHO-RYLASE
RIBOSE 1-P
URACIL
(for degra-dation see at left)
NUCLEOSIDASE
2'DEOXY-RIBOSE
NUCLEOSIDE DEOXY-RIBOSYLTRANS-FERASE
BASE Z
2'DEOXY-Z-RIBONUCLEOSIDE
dNTP, dNDP
dTTP
2'-DEOXYTHYMIDINE 5'-PHOSPHATE
2'-DEOXYTHYMIDINE 5'-DIPHOSPHATE
2'-DEOXY-THYMIDINE-5'-TRIPHOSPHATE
increased cell division
dTMP KINASE
NUCLEOSIDE DIPHOPHATE KINASE
... (dA) p (DNA) (as template), Mg++
DEOXY-RIBONU-CLEIC ACID
DNA—P—dT
7,8-DIHYDRO-FOLATE
5-FLUORO-dUMP, 5-TRIFLUOR-METHYL-dUMP
5,10-METHY-LENE-THF
THYMIDY-LATE SYNTHASE
THYMIDINE TRIPHOS-PHATASE
DEOXY-UTP-PYROPHOSPHATASE
dTTP, dUTP
DEOXYCYTIDINE TRIPHOSPHATE DEAMINASE
Mn++, Mg++
2'-DEOXY-URIDINE-5'-PHOSPHATE
2'-DEOXY-URIDINE-5'-DIPHOSPHATE
2'-DEOXY-URIDINE-5'-TRIPHOSPHATE
THYMIDINE KINASE
HYDROXYUREA
RIBONUCLEOSIDE TRIPHOSPHATE REDUCTASE

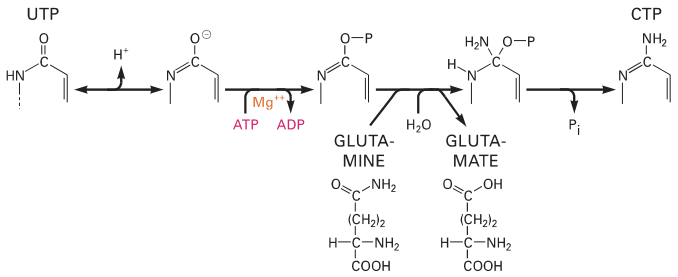


Figure 3.6.2-3. ATP-Dependent Replacement of an Oxo Group by an Amino Group in Nucleotides

5′-monophosphate (dTMP) is obtained from deoxyuridine 5′-monophosphate by a methylation reaction, which is catalyzed by thymidylate synthase and uses N^5, N^{10}-methylene tetrahydrofolate (THF) as a methyl donor. The folate moiety is simultaneously oxidized to the dihydrofolate (DHF) level (9.6.1) and has to be reduced again by the NADPH dependent dihydrofolate reductase:

$$dUMP + N^5,N^{10}\text{-methylene THF} = dTMP + DHF$$

$$DHF + NADPH = THF + NADP^+ + H^+$$

The activity of thymidylate synthase depends on the cell cycle and increases strongly when cells leave the G_0 and enter the S-phase (Fig. 3.9.1-2), similarly to ribonucleoside-diphosphate reductase (3.6.1.3). In higher plants, thymidylate synthase and dihydrofolate reductase are located on a bi-functional protein.

Afterwards dTMP is converted into deoxythymidine 5′-triphosphate (dTTP) by two phosphorylation steps. This compound is used for DNA synthesis. In order to prevent an accidental incorporation into DNA, the level of dUTP is kept low by the action of deoxyuridine pyrophosphorylase.

Apparently the presence of thymine bases (instead of uracil bases) in DNA is a protective measure. U, which is accidentally present in the template strand pairs with A and effects its incorporation into the daughter strand during replication. Thus, if C is converted into U by deamination as a result of damage, replication would result in an U-A pair instead of a C-G pair, which by the next round would yield an A-T pair. This change of the codon would lead to grave consequences, when the genes are being transcribed and translated thereafter (2.7.3).

In *E. coli*, which is infected by bacteriophages T2, T4 or T6, enzymes for hydroxymethylation of dCMP and its consecutive further phosphorylation are induced. The nucleotide triphosphate is introduced into DNA and pairs with G (Fig. 3.6.2-2).

3.6.2.4 Catabolism of Bases (Fig. 3.6.2-2)

Similarly to the purine analogues, the base residues of pyrimidine nucleotides and nucleosides have to be released for degradation. Cytosine compounds are converted into the respective uridine compounds by deamination. The general degradation pathway for uracil and thymine starts by reduction of the 5–6 double bond and consecutive hydrolytic ring opening. Deamination and decarboxylation lead to β-alanine and to 3-aminoisobutyrate, respectively.

An alternative pathway in bacteria begins with an oxidation reaction at C-6 to an oxo group, yielding barbiturate and to 5-methylbarbiturate, respectively. Then, a hydrolytic ring opening follows.

In some plants, toxic secondary metabolites are derived from pyrimidines for use as defense compounds.

3.6.2.5 Medical Aspects

Since the only task of thymidylate synthase is the provision of deoxythymidine nucleotides for DNA synthesis, its inhibition efficiently blocks DNA replication and cell division (‘thymineless death’).

This inhibition is most effective, when a high division rate requires a continuous supply of deoxythymidine nucleotides. An effective inhibitor is 5-fluorouracil, where the methyl group of thymine is replaced by a fluorine atom. After application, it is metabolized into fluordeoxyuridine 5monophosphate. During the thymidylate synthase reaction, a -SH group of the enzyme and methylene-THF form a covalent intermediary complex, which cannot dissolve in presence of the fluoro nucleotide and thus prevents further turnover. Another way to block the enzymatic reaction is the inhibition of dihydrofolate reductase, which reconstitutes the reaction partner tetrahydrofolate. This can be achieved by dihydrofolate analogues, e.g., aminopterine or methotrexate, which are strong competitive inhibitors of the mammalian enzyme ($K_i < 10^9$ mol/l, see 1.5.4). These inhibitors are used as anticancer drugs, but act likewise upon other fast growing cells (e.g., hair follicles, stem cells of the bone marrow). They are therefore quite toxic to humans.

Other dihydrofolate reductase inhibitors (e.g., trimethoprim) affect the bacterial enzyme about 10^5 times more strongly than the human enzyme and can be used as antibacterial drugs.

Inactivity of UMP synthase results in orotic aciduria, an autosomal recessive disease that is characterized by large amounts of orotic acid in the urine.

Literature:

Carter, N. *et al*. Adv. Exp. Med. Biol. 2008;625:141–154.

Jones, M.E. Ann. Rev. of Biochem. 1980;49:253–279.

Kafer, C. *et al.* Front. Biosci. 2004;9:1611–25.

Knox, A.J. *et al.* Gene 2009;429:23–30.

Kolberg, M. *et al.* Biochim. Biophys.Acta (BBA) / Proteins & Proteomics 2004;1699:1–34.

Nyhan, W.L. Molecular Genetics and Metabolism 2005;86:25–33.

Reichard, P. Ann. Rev. of Biochem. 1988;57:349–374.

Zrenner, R. *et al*. Ann. Rev. of Plant Biology 2006;57:805–836.

3.7 Cofactors and Vitamins

Ida Schomburg

Cofactors are chemical compounds that are required for many enzyme-catalyzed reactions and are tightly or loosely bound to the protein. When tightly bound or attached via a covalent bond to an enzyme they are called prosthetic groups. The heme group in hemoglobin is a prosthetic group (see 3.3.2.1). Inorganic prosthetic groups are usually transition metal ions such as cobalt, copper, iron, manganese or molybdenum. In nutrition they are essential trace elements. Loosely bound cofactors, termed coenzymes are organic substances that directly participate as substrates in an enzyme reaction.

Vitamins are compounds that are required in the diet of higher animals in small quantities, since they cannot be synthesized. A number of them are precursors of cofactors for enzymatic reactions (vitamin B group), while others are involved in the visual process and regulation of transcription (vitamin A), redox reactions (vitamins C and E), bone formation (vitamin D), blood coagulation (vitamin K), etc. The term cofactor is also used for other compounds such as hormones which act as coactivators or corepressors of metabolic processes.

The vitamins and cofactors belong to various chemical classes. A common way of classification refers to the solubility:

- Water-soluble vitamins: B_1, B_2, B_6, B_{12}, folate, pantothenate, biotin (= ‘vitamin B complex’), C. They are poorly stored. Intake in excess leads to secretion in the urine.
- Fat-soluble vitamins: A, D, E, K. They can be stored. Intake in excess can lead to hypervitaminoses (especially with A and E).

Some of their biosyntheses and actions are dealt with in different chapters. This chapter gives a survey of their properties and provides cross-references to the other chapters.

3.7.1 Retinol (Vitamin A)

Vitamin A is the common name for the alcohol retinol, the aldehyde retinal and retinoic acid, which are isoprenoids with lipid character (3.5.3). Several *cis-trans* isomers exist.

3.7.1.1 Biosynthesis and Interconversions

The biological source of the vitamin A group is the provitamin β-carotene, the biosynthesis of which takes place in plants and is shown in Figure 3.5.3-2. Animals are able to cleave carotene into retinal by a dioxygenase reaction (already in the intestine), thus they can use it in place of the vitamin.

Vitamin A is resorbed together with fat and stored as retinyl palmitate in lipid droplets in hepatic stellate (fat-storing) cells. For transport to the various organs, the ester is hydrolyzed and retinol is released into the bloodstream, where it is bound to specific retinol binding proteins. In the organs, interconversions of retinol, retinal and retinoic acid proceed via $NAD(P)^+$-dependent reactions.

In human nutrition, carotene (provitamin A) is mostly obtained from plant sources, while retinol is supplied by animal sources (e.g., fish liver oil). A daily intake of 900 µg retinol or 6 mg β-carotene (= 1000 retinol equivalents = 3333 IU) is required for adults. The body converts the provitamin carotene into vitamin A as it is needed, therefore high levels of carotene are not toxic. However, an excess of vitamin A in retinoid form is known as hypervitaminosis A and can be harmful or fatal. This occurs, when large doses of synthetic vitamin A (above 3000 µg/day) or great quantities of liver from animals adapted to polar environments are consumed.

Avitaminoses are common in undernourished populations. They cause night blindness, disturbances in bone development and in differentiation of epithelial tissue, inhibition of growth and of the reproductive function.

3.7.1.2 Biochemical Function

Retinoic acid is a regulator of gene transcription. After passing the cellular membrane, it binds to specific receptors: *all-trans* retinoic acid to the RAR receptor, 9-*cis* retinoic acid to the RXR receptor (7.7). The heterodimeric retinoic acid-RAR-RXR complexes bind in the nucleus to DNA hormone response elements (consensus sequence AGGTCA, repeated).

Among the genes regulated in this way are the genes for retinol binding proteins (3.7.1.1), PEP-carboxykinase (3.1.3.4) and apolipoprotein A1 (6.2.1). Retinoic acid is involved in the control of embryo- and morphogenesis (even small doses of vitamin A during pregnancy are teratogenic!), growth, differentiation and fertility.

The light-induced interconversion of 11-*cis*- and *all-trans*- retinal forms the basis of the visual process in animals (7.4.6). Halobacteria use the light-induced interconversion of *all-trans* and 13-*cis* retinal for proton pumping (3.12.1).

Possibly, retinyl phosphate is involved in the biosynthesis of glycosaminoglycans (2.9) in epithelial cells by playing a similar role as dolichyl phosphate (4.4). In this way, it indirectly sustains the integrity of cellular and mitochondrial membranes, keeps up a normal epithelium of the skin and mucosa and prevents disturbances of growth.

Literature:

Friedrich, W. *Vitamins*. De Gruyter, 1988, p. 63–140.

Roos, T.C., Jugert, F.K., Merck, H.F., Bickers, D.R. Phanrmacol Rev. 1998;50:315–333.

3.7.2 Thiamin (Vitamin B_1)

Vitamin B_1 (aneurin) is unphosphorylated thiamin. By pyrophosphorylation, it is converted into the coenzyme thiamin pyrophosphate (ThPP, thiamine diphosphate), which plays an essential role in oxidative decarboxylation and group transfer reactions.

3.7.2.1 Biosynthesis (Fig. 3.7.2-1)

The pathways of biosynthesis differ among pro- and eukarya (yeast, plants). Thiamin biosynthesis in *Escherichia coli* is composed of the separate formation of the pyrimidine and thiazole moieties. They are subsequently coupled to form thiamin phosphate.

Apparently in *Escherichia coli* and in *Salmonella typhimurium*, 1-(5′-phosphoribosyl)-5-aminoimidazole (AIR, see purine synthesis, 3.6.1.1) is a precursor of the pyrimidine residue, while in yeast it is formed in another way. In *E. coli*, possible precursors of the thiazole moiety are pyruvate and D-glyceraldehyde (compare Fig. 3.5.3-2). They condense to 1-deoxy-D-xylulose, which contributes the C-atoms 4′, 4, 5, 6, 7. Tyrosine is the origin of C-2′ and N-3′, while the sulfur may be provided by cysteine. The diphosphate ester of the pyrimidine moiety enters the coupling reaction yielding thiamin monophosphate, followed by a dephosphorylation to the vitamin. The conversion from the vitamin to the coenzyme thiamin pyrophosphate (ThPP) is a 1-step pyrophosphorylation.

Although much progress has been made in elucidating the thiamin biosynthetic pathway, not all details of the reactions are known so far. For plants the majority of enzymes in the pathway have not yet been characterized.

Rich sources of the vitamin are germinating grain and yeast. The recommended daily intake for humans is 1 … 1.5 mg. Avitaminosis leads to beriberi, which causes neurological disorders and affects the heart function.

3.7.2.2 Biochemical Function (Figs. 3.1.3-4 and 3.7.2-2)

The essential function of the coenzyme is the transfer of activated (= ThPP bound) aldehyde groups. ThPP is the prosthetic group of pyruvate dehydrogenase (3.1.3.1) and pyruvate decarboxylase (3.1.3.6), 2-oxoglutarate dehydrogenase (3.1.8.1) and transketolase (3.1.6.1, 3.12.2); thus it is involved in the course of glycolysis and of the citrate, pentose-P and Calvin cycles.

The C-2′ atom has a very acidic character due to its proximity with N^+ and S. The loss of the proton results in a carbanion, which easily adds to carbonyl groups (Fig. 3.1.3-4). In the decarboxylase and dehydrogenase reactions mentioned above, this effects the decarboxylation of the 2-oxo acid and the formation of a resonance-stabilized ThPP-bound carbanion, which represents an activated aldehyde. This residue is either released (decarboxylase enzyme) or transferred to lipoamide with concomitant oxidation (dehydrogenase enzymes). In the transketolase reaction, the reaction partner is a keto sugar, which is cleaved. An aldehyde is released, while the ThPP-bound activated aldehyde is transferred to an incoming compound (Fig. 3.7.2-2).

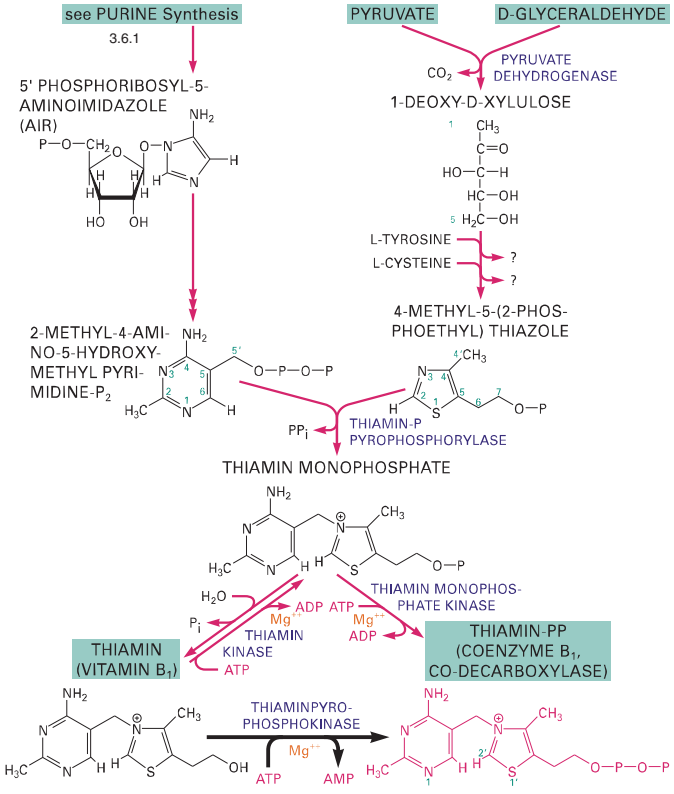


Figure 3.7.2-1. Biosynthesis of Thiamin *in E. coli*

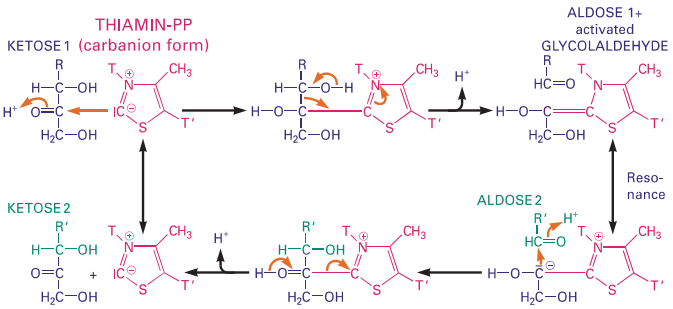


Figure 3.7.2-2 Reaction Mechanism of Transketolase (T = pyrimidine residue, T′ = C_2-PP residue)

Literature:

Friedrich, W. *Vitamins*. De Gruyter, 1988, p. 339–401.

White R.L., Spenser, J.D. in Neidhardt, F.C. (Ed.) *Escherichia coli and Salmonella.* 2nd Ed. ASM Press, 1996, p. 680–686.

3.7.3 Riboflavin (Vitamin B_2), FMN and FAD

Besides the nicotinamide nucleotides NAD^+ and $NADP^+$ (3.7.9), the flavin nucleotides flavin-adenine dinucleotide (FAD) and flavin mononucleotide (FMN) are the other important group of hydrogen carriers. They take part in more than 200 redox reactions. These coenzymes are derived from the vitamin riboflavin by phosphorylation (FMN) and by a consecutive adenylylation (FAD).

3.7.3.1 Biosynthesis and Interconversions (Fig. 3.7.3-1)

The vitamin riboflavin is biosynthesized from GTP and ribulose 5-P in plants, yeasts and many microorganisms.

The pathway starts by the opening of the imidazole ring of GTP and removal of a pyrophosphate residue. Deamination, reduction and removal of the last remaining phosphate yields 5-amino-6-ribitylamino-2,4-pyrimidinedione. The reaction of this compound with 3,4-dihydroxy-2-butanone-4-P (originating from ribulose 5-P) results in the bicyclic compound 6,7-dimethyl-8-ribityllumazine. This compound undergoes an unusual dismutation by transfer of a 4-carbon unit, which leads to the tricyclic compound riboflavin and back to pyrimidinedione. Phosphorylation of this vitamin finally results in flavin mononuceotide (FMN). A consecutive adenylylation yields flavin-adenine dinucleotide (FAD). The phosphorylation of riboflavin is strictly controlled in animals, e.g., by thyroid hormones and by aldosterone.

FAD and FMN are usually tightly bound to enzymes (flavoproteins). In some cases even a covalent bond exists (in most cases to the 8α-methyl group, e.g., in succinate dehydrogenase). Thus, reduced flavin coenzymes have to be reoxidized 'on the spot' and cannot diffuse to another enzyme. Protein binding greatly diminishes the light sensitivity of the free flavin coenzymes.

The vitamin is present in many vegetables and in meat, less in grain products. During digestion in animals, the various flavoproteins from food are degraded and riboflavin is resorbed. The daily requirement of a human adult is about 1.4 … 2 mg. Avitaminoses cause disturbances of growth and skin diseases. They are quite rare.

The major degradation and excretion product in humans is riboflavin. Other degradation steps (especially by bacteria) are hydroxylations of methyl residues and shortening of the ribityl chain.

3.7.3.2 Biochemical Function

The redox function of FMN and FAD is exerted by the tricyclic isoalloxazine moiety. For full reduction, it accepts two electrons and two protons (usually in the 1,5-positions, Fig. 3.7.3-1) and preserves electroneutrality. Contrary to $NAD(P)^+$, the flavin compounds can also take part in one-electron reactions by the formation of a semiquinone radical. Thus they can act as 'redox switches' between one- and two-electron mechanisms. The redox potential is more positive than the

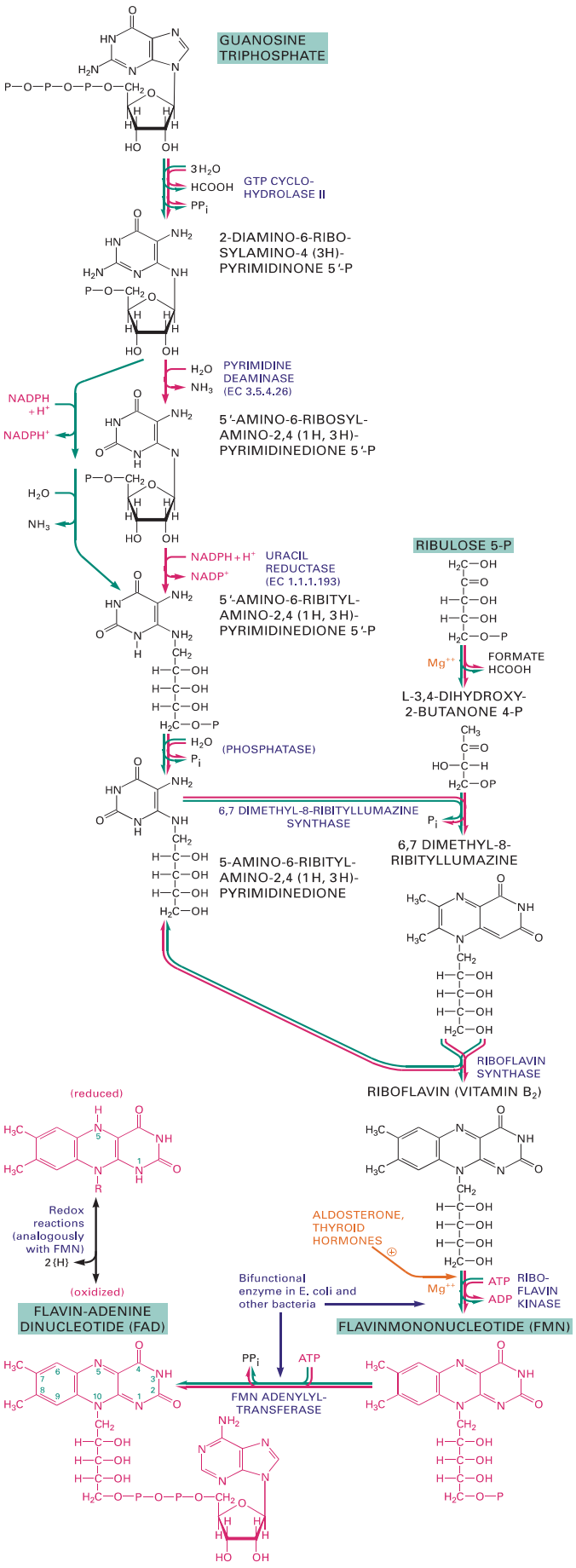


Figure 3.7.3-1. Biosynthesis of Riboflavin, FMN and FAD

potential of NAD(P)$^+$ (Table 3.11-2) and is further modified by the protein binding of the coenzyme. This enables flavin coenzymes to perform other oxidation reactions than NAD(P)$^+$.

Frequently they react as intermediates in the transfer of redox equivalents between NAD(P)H and various acceptors, e.g., in the reactions catalyzed by:

- dihydrolipoyl dehydrogenase (3.1.3.2) or GSSG reductase (3.2.5.7) – transfer to or from -S-S- groups

or in the transfer of redox equivalents between substrates and the respiratory chain (3.11.4), e.g., in the reactions catalyzed by:

- NADH dehydrogenase (Fig. 3.11-1) – transfer to FeS centers in complex I,
- glycerol-3-P dehydrogenase as part of the glycerol 3-P dehydrogenase shuttle (Fig. 3.11-2) – transfer to the ubiquinone (UQ) pool,
- acyl-CoA dehydrogenases during fatty acid oxidation (3.4.1.5) or succinate dehydrogenase (3.1.8.1) – transfer to the UQ pool, formation of double bonds.

In reduced form a number of flavoproteins are able to react with molecular oxygen (for the reason see 3.2.5.8), e.g., in

- oxidation reactions with formation of H_2O_2 (2 e$^-$ reactions):
  - oxidative deaminations, e.g., catalyzed by monoamine oxidase (3.2.7.4)
  - oxidation of aldehydes, e.g., catalyzed by glucose oxidase (3.1.5.1)
  - oxidation of alcohols to yield aldehydes or ketones, e.g., catalyzed by yeast alcohol oxidase
- monooxygenase reactions (4 e$^-$ reactions)
  - 'external' type: X-H + $FADH_2$ + O_2 = X-OH + FAD + H_2O as in the kynurenine 3-monooxygenase reaction (Fig. 3.2.7.4); FAD is then again reduced by NADPH
  - 'internal' type: X-CHOH-COO^- + O_2 = X-COO^- + H_2O + CO_2 as in the lactate 2-monooxygenase reaction. The flavin acts as an intermediate carrier of redox equivalents.

The different oxidation states show characteristic spectra with the following λ_{max}: FAD 370 and 450 nm (yellow), semiquinone radical 370, 450 and 590 nm (red), $FADH_2$ 370 nm (pale yellow).

A related compound is factor F_{420} (5-deazaflavin, 3.10.6.2), which occurs in methanogenic bacteria.

Literature:

Bacher, A. *et al.* in Neidhardt, F.C. (Ed.). *Escherichia coli and Salmonella.* 2nd Ed. ASM Press, 1996, p. 657–664.

Friedrich, W. *Vitamins*. DeGruyter (1988) 403–471.

3.7.4 Pyridoxine (Vitamin B_6)

Pyridoxine (pyridoxol), pyridoxal and pyridoxamine compose the vitamin B_6 group. Pyridoxal and pyridoxamine phosphates act as essential coenzymes in a large number of amino acid conversions, e.g., transaminase, decarboxylase and dehydratase reactions.

3.7.4.1 Biosynthesis and Interconversions (Fig. 3.7.4-2)

A number of bacteria, yeast and plants are producers of pyridoxine. Other bacteria and animals have to rely on external supply. In *E. coli* the condensation product of pyruvate and D-glyceraldehyde, 1-deoxy-D-xylulose (compare 3.7.2 and Fig. 3.5.3-2), is the origin of the C-atoms 2′, 2, 3, 4 and 4′. Erythrose 4-phosphate is converted into 4-hydroxy-L-threonine, apparently in a reaction analogous to serine biosynthesis from 3-P-D-glycerate (Fig. 3.2.4-1). This compound is the precursor of the C-atoms 5′, 5, 6 and N-1 of pyridoxine. The synthesis is feedback inhibited by pyridoxine. The dehydrogenation of unphosphorylated pyridoxine to pyridoxal is a bacterial reaction.

Vitamin B_6 biosynthesis in plants does not appear to depend on 1-deoxy-D-xylulose. Instead, the vitamin is synthesized from the pentose phosphate pathway intermediates, D-ribulose 5-phosphate and glyceraldehyde 3-phosphate (3.1.6). The nitrogen is provided by the glutaminase subunit of the pyridoxal 5′-phosphate synthase complex.

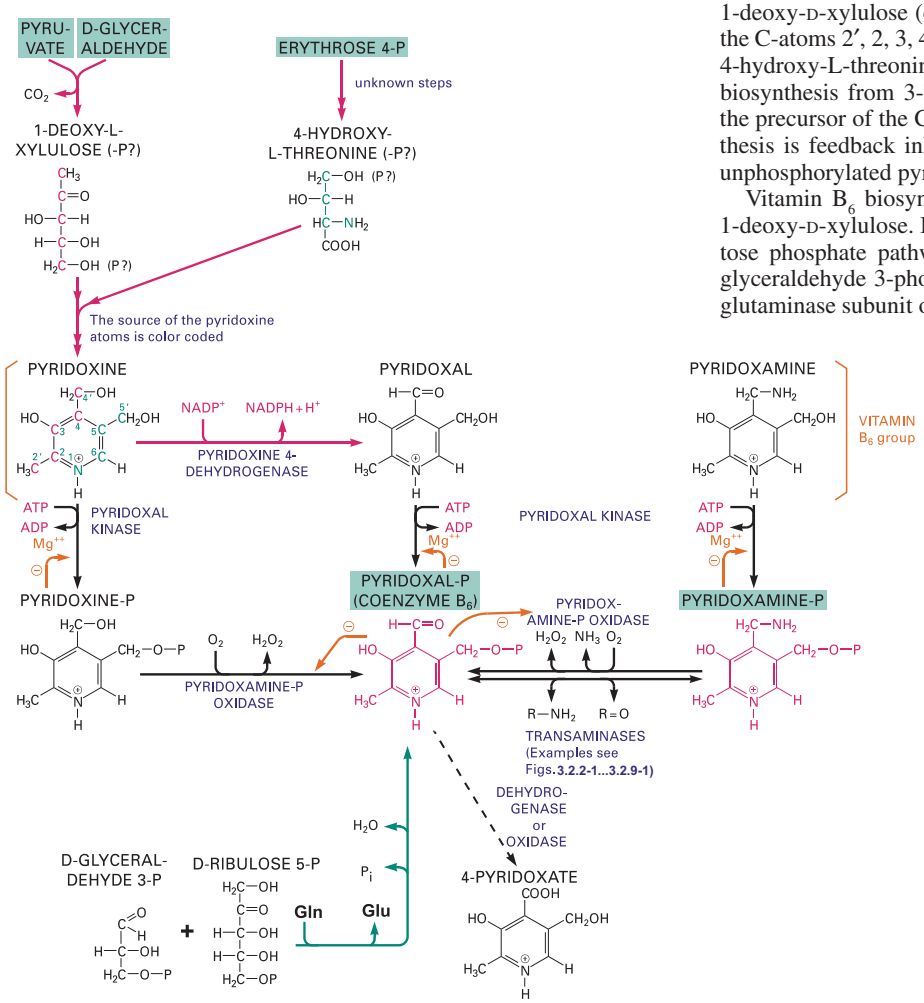


Figure 3.7.4-1. Precursors (in *E. coli*) and Metabolism of Pyridoxine

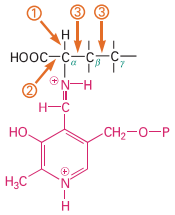


Figure 3.7.4-2. The Reactive Intermediate (Schiff Base) of Pyridoxal Catalyzed Reactions

In animals, the vitamins pyridoxine, pyridoxal and pyridoxamine are phosphorylated after resorption. The kinase is subject to product inhibition. Then an oxidase reaction converts pyridoxine phosphate and pyridoxamine phosphate into pyridoxal phosphate. This oxidase is also product-inhibited.

The recommended daily intake for adult humans is ca. 2 mg. Deficiencies lead to neuronal disorders (neuritis), dermatitis and impaired amino acid metabolism (xanthurenate excretion, 3.2.7.3).

Degradation of the phosphorylated compounds begins with a phosphatase reaction. Further degradation of the vitamin proceeds by oxidation to the carboxylic acid 4-pyridoxate, which is biologically inactive.

3.7.4.2 Biochemical Function

Pyridoxal phosphate is the essential cofactor for reactions, which act on the C-2 (C_α) atom of amino acids and involve cleavage on any of its bonds. These bonds are labilized by the formation of a Schiff base between the coenzyme and the amino acid (Fig. 3.7.4-2).

① C_α–H:
- – transaminations (for mechanism see Fig. 3.2.2-3),
- – racemizations (e.g., alanine racemase, Fig. 3.10.1-1)

② C_α–COOH:
- – decarboxylation, e.g., glycine dehydrogenase (decarboxylating) (for mechanism see Fig. 3.2.4-3), histidine decarboxylase (Fig. 3.2.8-1)

③ C_α–R or C_β–C_γ:
- – β-eliminations, e.g., glycine hydroxymethyltransferase (Fig. 3.2.4-1), L-serine ammonia-lyase (for mechanism see Fig. 3.2.4-2)
- – γ eliminations, e.g., cystathionine γ-lyase (Fig. 3.2.5-2)
- – β-replacement (e.g., cystathionine β-synthase, Fig. 3.2.5-2; the serine hydroxyl is replaced by the homocysteine residue)

Literature:

Friedrich, W. *Vitamins*. De Gruyter, 1988, p. 543–618.

Hill, R.E., Spenser, J.D. in Neidhardt, F.C. (Ed.) *Escherichia coli and Salmonella.* 2nd Ed. ASM Press, 1996, p. 695–703.

Tambasco-Studart, M. *et al*. Proc. Natl. Acad. Sci. USA 2005;102: 13687–13692.

Fitzpatrick, T.B. *et al*. Biochem. J. 2007;407:1–13

3.7.5 Cobalamin (Coenzyme B_{12}, Vitamin B_{12})

Cobalamin is one of the largest non-polymeric biological molecules. The central cobalt atom is coordinated with six ligands. Four of them are the pyrrols of the corrin ring, one is the unusual nucleotide dimethylbenzimidazole and the sixth can be hydroxyl or cyano (vitamin B_{12}), methyl (methylcobalamin) and 5′-deoxyadenosyl (deoxyadenosylcobalamin, coenzyme B_{12}). The cobalt ion can occur as Co^+, Co^{++} and Co^{+++}. The change in the oxidation states is essential for its coenzyme function.

3.7.5.1 Uptake of the Coenzyme and Reduction of the Vitamin

Coenzyme B_{12} is biosynthesized exclusively by microorganisms in a multistep procedure. The first steps are identical to the pathways to heme and chlorophyll (3.3.1). There are several variants of the biosynthetic reactions. The pathway in *Pseudomonas denitrificans* is described in 3.3.5 and shown in Figure 3.3-8.

The needs of mammals (in humans < 10 µg/day) are partially met by the B_{12} production of intestinal bacteria, partially from animal food sources. A specific transporter protein (intrinsic factor, 50 kDa, secreted from gastric cells) is required for the resorption of vitamin B_{12} from the intestinal lumen. Insufficient production of this factor causes severe B_{12} deficiency (pernicious anemia). For the transport in blood, B_{12} is bound to different plasma globulins, called transcobalamins. B_{12} is taken up in the most oxidized form (hydroxycobalamin, Co^{3+}). After two reduction steps, the deoxyadenosyl form (coenzyme B_{12}) is restored (Fig. 3.7.5-1).

3.7.5.2 Biochemical Function

In bacteria many different reactions are dependent on the coenzyme B_{12}, whereas in mammals there are only two of them (see below).

The cobalamin-dependent reactions can be classified according to the way the Co-C bond is cleaved:

- Homolytic cleavage, employing deoxyadenosyl cobalamin:
  - – Intramolecular rearrangements: The cleavage of the metalorganic bond ($Co^{3+} \rightarrow Co^{2+}$, Fig. 3.7.5-2) results in the formation

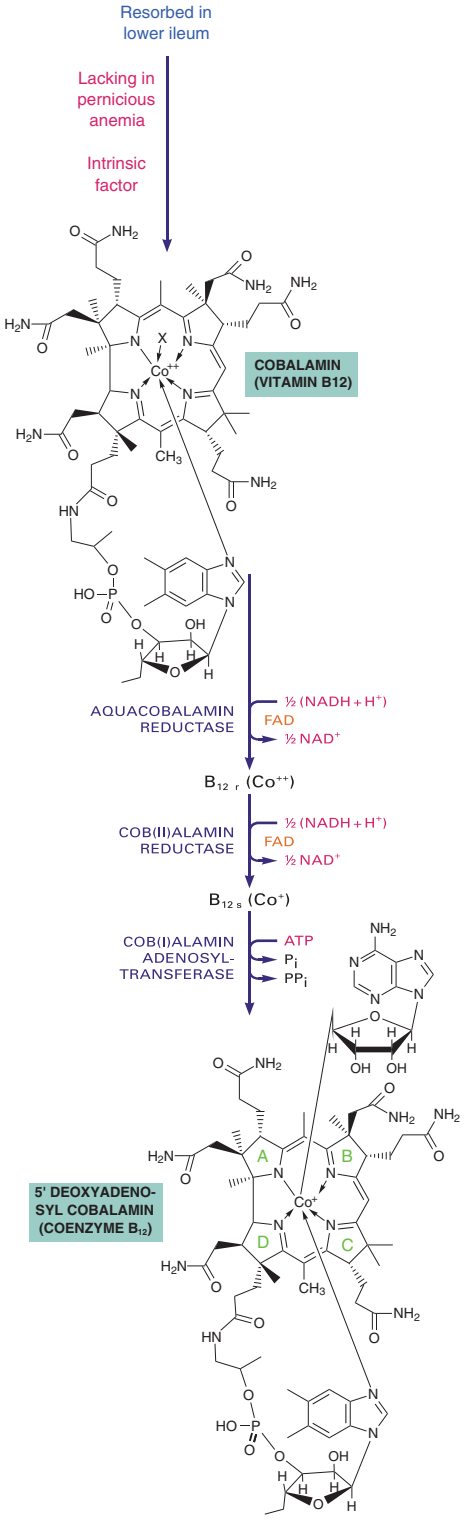


Figure 3.7.5-1. Conversion of Vitmin B_{12} into Coenzyme B_{12}

of a free radical at the deoxyadenosyl group. This radical is used to exchange the positions of hydrogen and various substituents at neighboring atoms.

Examples: methylmalonyl-CoA mutase in mammals (Fig. 3.2.5-2, exchanging H and a - CO-S-CoA group), propionate and glutamate fermentation by bacteria (Figs. 3.10.5-2 and 3.10.5-4).

– Reduction of ribonucleotides: Some bacterial ribonucleoside-triphosphate reductases (3.6.1-4) contain coenzyme B_{12} instead of the Fe(III)-O^{2-}-Fe(III) center to generate a free radical.

- Heterolytic cleavage, employing methyl cobalamin: In some methyl transferases, methylcobalamin acts as the methyl group donor (Co^{3+}-$CH_3 \rightarrow Co^{+} + CH_3$, Fig. 3.7.5-3).

Examples are homocysteine methyltransferases in mammals and bacteria (Fig. 3.2.5-2). B_{12} is then remethylated by 5-methyl-tetrahydrofolate. An occasional oxidation of the cofactor from the intermediate Co^{+} to the Co^{++} state requires a regeneration step, involving reduction by NADPH via flavo- or ferredoxin and methylation by 5-adenosylmethionine. Related reactions are methyl transfers in archaeal methanogenesis and bacterial acetogenesis (Figs. 3.10.6-2 and 3.10.6-3)

3.7.5.3 Siroheme and Coenzyme F_{430}

Siroheme (Fig. 3.3-1) is derived from precorrin 2 by introduction of an iron atom. It is the coenzyme of bacterial sulfite and nitrite reductases (3.10.6). Coenzyme F_{430} (Fig. 3.3-1) is a nickel-containing derivative of precorrin 3A. It acts as cofactor in the methyl-coenzyme M reductase reaction of methanogenesis (3.10.6.2) and functions similar to coenzyme B_{12} in methyl transfer reactions.

Literature:

Banerjee, R. Chemistry and Biology 1997;4:175–186.

Blanche, F. *et al.* Angew. Chem. Int. Ed. 1995;34:1001–1029.

Drennan, C.L. *et al.* Science 1994;266:1669–1674.

Kräutler, B. Biochem. Soc. Transactions 2005;33:806–810

Ludwig, M.L., Matthews, R.G. Ann. Rev. of Biochem. 1997;66:269–313.

Stubbe, J. Science 1994;266:1663–1664.

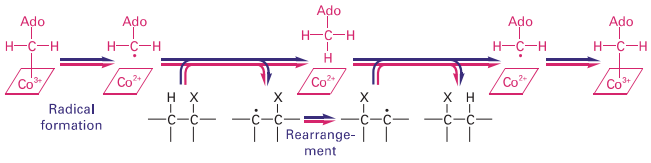
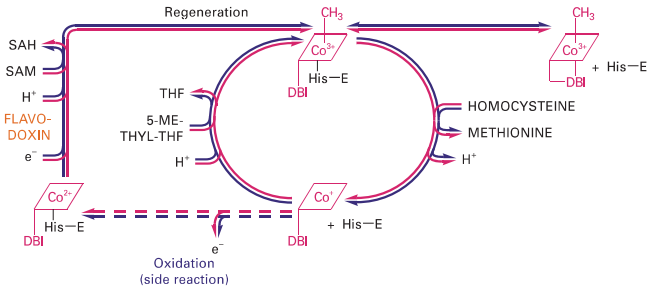


Figure 3.7.5-2. Mechanism of B_{12} Dependent Mutase Reactions Employing Deoxyadenosyl Cobalamin



Abbreviations: Ado = adenosine, DBI = dibenzimidazole, SAM = S-adenosylmethionine, SAH = S-adenosylhomocysteine, THF = tetrahydrofolate, E = homocysteine methyltransferase, Co = corrin ring

Figure 3.7.5-3. Reaction Mechanism of Homocysteine Methyltransferase Employing Methylcobalamin

3.7.6 Folate and Pterines

One vitamin and several cofactors contain a pteridine ring system:

- Folate/tetrahydrofolate/tetrahydrofolylpolyglutamate (THF)
- Biopterin/tetrahydrobiopterin (THB)
- Molybdenum (and tungsten) cofactors (MoCo)
- Methanopterin/tetrahydromethanopterin (THMPT).

The synthesis of folate compounds is restricted to plants and microorganisms, but the product is required as a vitamin (vitamin B_c) by animals. Biopterin synthesis seems to be only performed by animals, whereas the molybdenum cofactor can apparently be synthesized by all living organisms. Methanopterin is formed exclusively by anaerobic archaea (methanogens).

3.7.6.1 Tetrahydrofolate/Folylpolyglutamate (Fig. 3.7.6-1)

Biosynthesis: The synthesis of folate starts with a two step conversion of GTP to dihydroneopterin-P_3 by GTP cyclohydrolase I (GTPCH I). C-8 of GTP is released as formate. The enzyme, in contrast to GTPCH II (riboflavin synthesis, 3.7.3), incorporates C atoms of ribose (1′, 2′) into the pteridine ring system. Dihydroneopterin-P_3 is also an intermediate of the pathways leading to biopterin (3.7.6.3) and to methanopterin (3.7.6.5).

The specific part of the tetrahydrofolate biosynthesis proceeds via a (so far unspecified) removal of all phosphates, side chain shortening and pyrophosphorylation to 2-amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridine-P_2. The fusion of this intermediate with 4-aminobenzoate to 7,8-dihydropteroate is the target of the antibacterial action of sulfonamides. After the ATP dependent addition of glutamate, the resulting dihydrofolate is reduced to tetrahydrofolate. In eukarya, the synthesis is performed in mitochondria. Some steps are catalyzed by species specific multifunctional enzymes.

Polyglutamylation of reduced folate: Tetrahydrofolylpolyglutamate synthase catalyzes the sequential addition of usually 2 … 8 glutamate residues via γ-peptide linkages. Polyglutamylation increases or even enables the cofactor function of folates. It also inhibits the diffusion of folates through membranes and thus enables their accumulation and subcellular localization. For resorption of folates from the diet, these bonds must be hydrolyzed by pteroyl-poly-γ-glutamate hydrolase. The dietary requirement for humans is 0.4 mg/day.

Biochemical functions of folylpolyglutamate: Polyglutamylated tetrahydrofolate (THF) is the central cofactor of the one carbon (C_1) metabolism (see below).

3.7.6.2 General Reactions of the C_1 Metabolism (Fig. 3.7.6-1)

C_1 metabolism is a key regulatory factor of catabolic and anabolic pathways by providing a controlled flux through the C_1 pool. It encompasses the transfer of C_1 compounds and their reduction or oxidation. CO_2 fixation and biotin dependent reactions (3.7.8.2) are not usually discussed in this context.

THF carries C_1 units at the oxidation levels of formate (as 5- or 10-formyl-THF, 5,10-methenyl- or 5-formimino-THF), formaldehyde (as 5,10-methylene-THF) and methanol (as 5-methyl-THF). The origins of the C_1 units entering at the various levels are

- 10-formyl-THF: formate
- 5-formimino-THF: degradation products of histidine (3.2.8.2) and xanthine (3.6.1.6)
- 5,10-methylene-THF (mainly in mitochondria): serine (3.2.4.1, main supplier of C_1 units), betaine degradation products

C_1 Units are provided by

- 10-formyl-THF: for purine (3.6.1.1) and N-formylmethionyl-tRNA formation (in bacteria, archaea, mitochondria and chloroplasts, 4.1)
- 5,10-methylene-THF: for biosyntheses of thymidine 5′-P (3.6.2.3), coenzyme A (3.7.7.1) and serine (3.2.4.1, mainly in the cytosol)
- 5-methyl-THF: for methionine biosynthesis (3.2.5.4). Methionine in the form of S-adenosylmethionine serves as the donor of all energy driven methylation reactions.

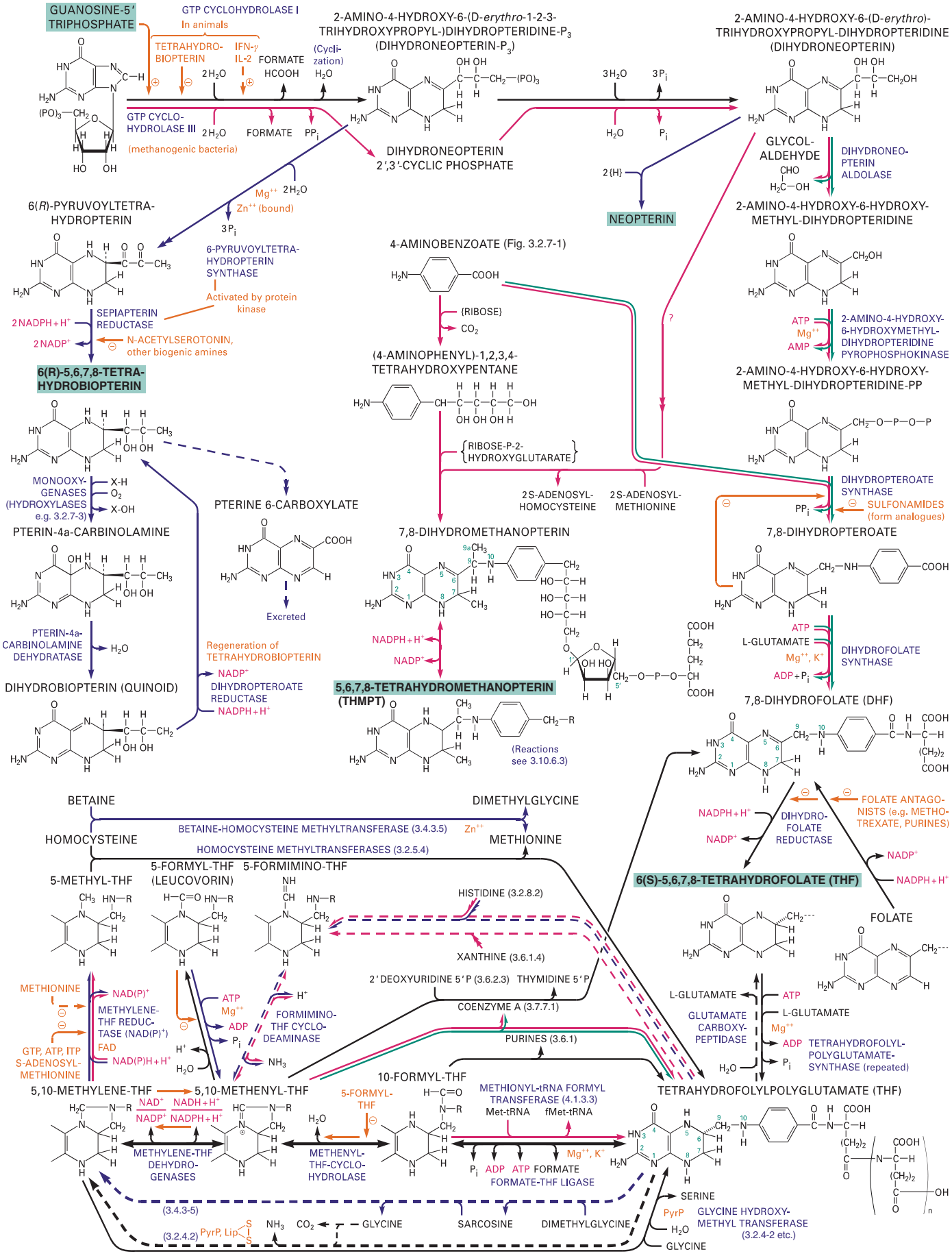


Figure 3.7.6-1. Biosynthesis of Pterines

The reversible interconversions of C_1 compounds take place by formate-THF ligase, methenyl-THF cyclohydrolase and methylene-THF dehydrogenase (NADPH, Fig. 3.7.6-1). In animals they are combined into a trifunctional enzyme. In mitochondria, the dehydrogenase activity is NAD-dependent and controls the metabolic flux from serine/glycine to formate (for protein synthesis and for cytoplasmatic needs). Glycine hydroxymethyltransferase (3.2.4.2) and 5-formyl-THF cycloligase interconvert 5,10-methylene-THF and 5-formyl-THF. Irreversible reactions are catalyzed by 10-formyl-THF dehydrogenase, which releases THF and by methylene-THF reductase (NADPH), which leads to 5-methyl-THF. In the case of cobalamin deficiency, the 5-methyl-THF-homocysteine hydroxymethyltransferase (3.2.5.4) is inoperative and a deleterious accumulation of 5-methyl-THF takes place.

Most regulation mechanisms are poorly understood. 5-Formyl-THF inhibits most enzymes of the C_1-THF interconversions and other C_1-THF dependent reactions. S-Adenosylmethionine inhibits the methylene-THF reductase.

3.7.6.3 Tetrahydrobiopterin (Fig. 3.7.6-1)

Biosynthesis: Tetrahydrobiopterin (THB) biosynthesis proceeds analogously to folate biosynthesis (3.7.6.1) up to the dihydroneopterin-P_3 level. The Zn-dependent 6-pyruvoyltetrahydropterin synthase catalyzes an internal redox transfer and eliminates the triphosphate residue. Two consecutive reduction steps are performed by the NADPH dependent sepiapterin reductase.

The rate limiting step of THB synthesis is the initial GTP cyclohydrolase I reaction. Feedback inhibition by THB is mediated by the GTP cyclohydrolase I feedback regulatory protein (GFRP, formerly p35). The *de novo* synthesis of the cyclohydrolase is suppressed by glucocorticoids and cell-specifically stimulated by, e.g., interferon γ, interleukin-1β or other cytokines.

Regeneration of THB is an essential part of the phenylalanine hydroxylating system. During the reactions catalyzed by aromatic amino acid hydroxylases, molecular oxygen is transferred to the corresponding amino acid and THB is oxidized to THB-4a-carbinolamine. Two enzymes are involved in its subsequent dehydratation and reduction to THB: pterin-4a-carbinolamine dehydratase (PCD) and dihydropteridine reductase (DHPR). Enzymic recycling of THB is essential for phenylalanine metabolism: (1) to ensure a continuous supply of reduced cofactor, and (2) to prevent accumulation of harmful metabolites produced by rearrangements of THB-4a-carbinolamine.

Biochemical functions of tetrahydrobiopterin: THB is the cofactor of the aromatic amino acid monooxygenases (phenylalanine, tyrosine and tryptophan monooxygenases, 3.2.7.3), of nitric oxide synthase (7.7.2) and of glyceryl-ether hydroxylase (e.g., PAF degradation, 3.4.3.3).

Since THB plays a role in the biosynthesis of serotonin (3.2.7.3) and catecholamines (3.2.7.4), defects in THB synthesis or regeneration may lead to, e.g., neurological diseases.

Additionally THB seems to be involved in other cofactor functions, e.g., in cytokine dependent proliferation of erythroid and T cells and in interactions with the IL-2 receptor.

Neopterin: This pteridine, together with dihydroneopterin, is secreted from monocytes/macrophages after stimulation of THB synthesis by interferon-γ.

Other pteridines: There is strong evidence that in algae, fungi and plants, a pteridine is one of the cofactors of the blue light receptor.

3.7.6.4 Molybdenum/Tungsten Cofactors (Fig. 3.7.6-2)

Biosynthesis: Molybdopterin (MPT) synthesis likewise seems to start from GTP. Ring opening and further steps (without release of formate) lead to the labile cyclic pyranopterin monophosphate (cPMP), which is sulfurylated and rearranged by MPT synthase to MPT, forming the Mo binding dithiolene structure. Incorporation of Mo (delivered by the Mo uptake and transport system) into MPT yields the very unstable molybdenum cofactor (MoCo). Enzymes of the xanthine oxidase family require a final maturation step by addition of a terminal sulfido ligand to the molybdenum center, as catalyzed by the enzyme molybdenum cofactor sulfurase. The last step of the synthesis is the incorporation into the target enzyme, which effects

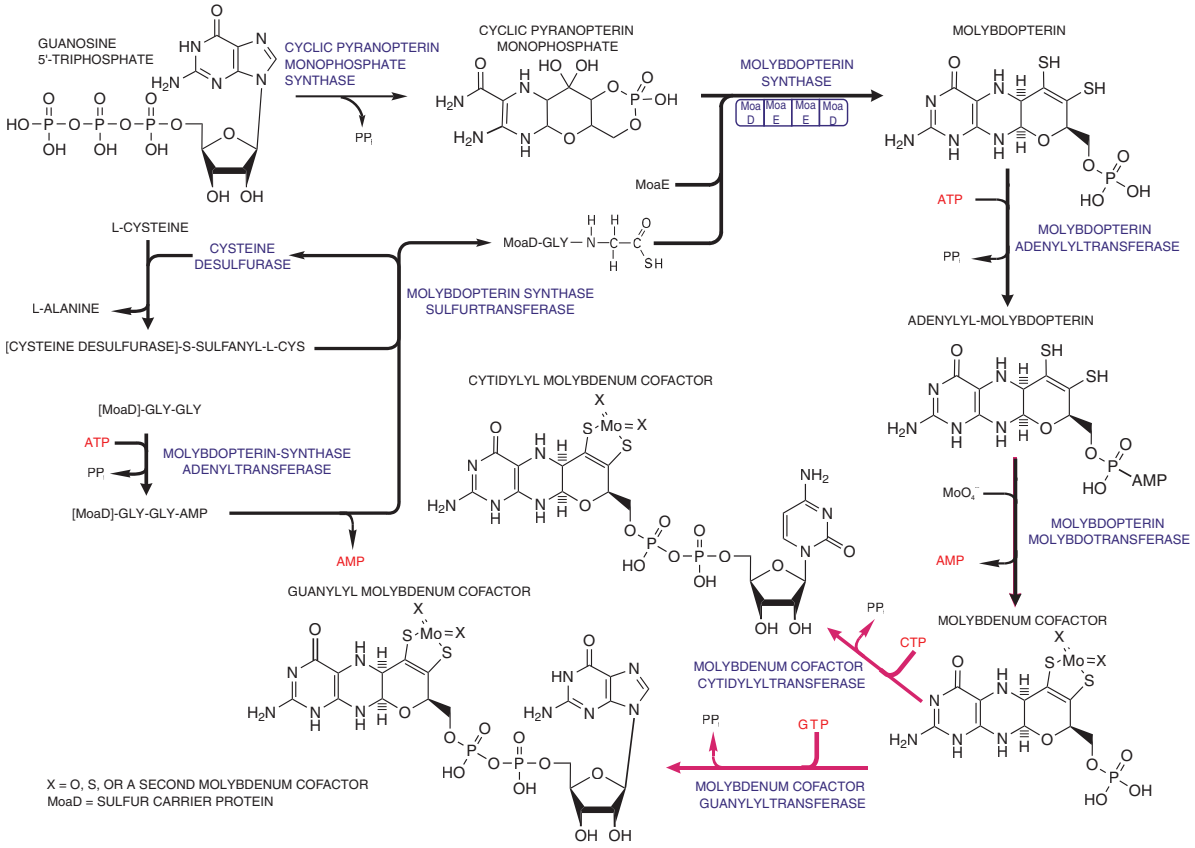


Figure 3.7.6-2. Biosynthesis of the Molybdenum Cofactor

the stabilization of the cofactor. In bacteria, MoCo is usually converted into a dinucleotide before incorporation into the enzyme takes place.

Biochemical function: All molybdoenzymes (except nitrogenase, 3.2.1) contain the organometallic molybdenum cofactor. Mammalian molybdoenzymes are sulfite oxidase, xanthine oxidase (3.6.1.6), aldehyde oxidase and amidoxime reductase. Molybdopterin is degraded to urothione. Defects in molybdopterin biosynthesis in humans are very rare. They are lethal in early childhood. In addition to the four molybdoenzymes occuring in mammals, plants possess assimilatory nitrate reductases which contain a molybdenum cofactor.

In plants and bacteria, a broad range of reactions are catalyzed by molybdoenzymes. Three enzyme families were defined by functional criteria:

- DMSO reductase family in bacteria, e.g., DMSO reductase, nitrate reductase (dissimilatory, 3.10.6), formate dehydrogenases (one of them contains tungsten)
- Xanthine oxidase family, e.g., xanthine oxidase/dehydrogenase (3.6.1.6, contains additionally FAD), aldehyde oxidase.
- Sulfite oxidase family in algae and higher plants, e.g., sulfite oxidase/dehydrogenase (3.10.6), nitrate reductase (assimilatory)

Related are tungsten enzymes, which form the majority of the aldehyde ferredoxin oxidoreductase family, e.g., aldehyde ferredoxin oxidoreductase.

3.7.6.5 Methanopterin (Fig. 3.7.6-1)

Biosynthesis: The biosynthesis of tetrahydromethanopterin (THMPT) resembles that of tetrahydrofolate and apparently branches off at the dihydroneopterin-P_3 level. Then S-adenosylmethionine dependent methylation at C-7 takes place. The sources of the C-9 and C-9a atoms have not been conclusively proven. The side chain of methanopterin is formed by the condensation of 4-aminobenzoate (3.2.7.1) with ribose. The mechanism resembles the indole 3-glycerol-P synthase reaction (3.2.7.1). Ribosyl 5-P and 2-hydroxyglutarate (obtained by reduction of 2-oxoglutarate, Fig. 3.1.8-2) are added afterwards.

Biochemical function: The C_1 derivatives of tetrahydromethanopterin are analogous to those of tetrahydrofolate. They are involved in methane synthesis, as described in 3.10.6.2.

Literature:

Auerbach, G., Nar, H. Biol. Chem. 1997;378:185–192.

DiMarco, A.A. *et al.* Ann. Rev. Biochem. 1990;59:361–370.

Hille, R., Nishino, T., Bittner, F. Coord. Chem. Rev. 2011; 255:1179–1205.

Leimkuhler, R.R., Wuebbens, M.M., Rajagopalan, K.V. Coord. Chem. Rev. 2011;255:1129–1144.

Mendel, R.R. Planta 1997;203:399–405.

Mendel, R.R., Schwarz, G.: Coord. Chem. Rev. 2011;255:1145–1158.

Neidhardt, F.C. *et al.* (Eds.) *Escherichia coli and Salmonella*. ASM Press, 1996, p. 665–679.

Rébeillé, F. *et al.* EMBO J. 1997;16:947–957.

Thöny, B., Auerbach, G., Blau, N. Biochem. J. 2000;347:1–16

3.7.7 Pantothenate, Coenzyme A and Acyl Carrier Protein (ACP)

Panthothenate (vitamin B_5) is an essential component of the important acyl group carriers, coenzyme A (CoA or CoA-SH, 3.1.3.3) and acyl carrier protein (ACP), which are present in all cells and take part in more than 100 reactions. Among them are fatty acid synthesis (3.4.1.1) and degradation (3.4.1.5), pyruvate oxidation (3.1.3.1) and glyceride synthesis (3.4.2.1).

3.7.7.1 Biosynthesis and Interconversions (Fig. 3.7.7-1)

The entire biosynthetic pathway has been elucidated in e.g., *Escherichia coli* and *Arabidopsis thaliana.* In humans all genes necessary for biosynthesis of coenzyme A from pantothenate have been identified.

2-Oxoisovalerate, the transamination product of l-valine, is methylated by 5,10-methylene-THF to yield 2-dehydropantoate, which is then reduced to (R)-pantoate. An ATP-dependent condensation with β-alanine (which is produced from aspartate by 1-decarboxylation, 3.2.2.4) leads to (R)-pantothenate. Since intestinal *E. coli* secretes considerable amounts of it, this becomes an important source of the vitamin to the mammalian host. Also, pantothenol can be oxidized to pantothenate. It is present in food, but is also used as a drug.

The synthesis of coenzyme A (CoA) from the vitamin pantothenate can be performed by all living beings and begins with its phosphorylation. This is the committed step (= first ambiguous step), which is regulated via allosteric inhibition by the end product CoA. Then CTP-dependent condensation with l-cysteine and decarboxylation takes place. Adenylylation and phosphorylation complete the synthesis of the coenzyme. In mammals, these two steps are carried out by a bifunctional enzyme, while in plants and bacteria separate entities exist. The active holo-acyl carrier protein (ACP, 3.4.1.1) is obtained by transfer of the 4′-phosphopantetheine residue from CoA to a serine residue of apo-ACP.

Dietary requirement for humans: Since pantothenate is present in practically all biological material and is, in addition to food, supplied by intestinal bacteria (see above), avitaminoses are practically unknown. The daily requirement for adults is estimated to be in the order of 6 … 10 mg.

Degradation: CoA is cleaved by a phosphodiesterase, yielding 4′-phosphopantetheine and adenosine 3′,5′-bisphosphate. Likewise, the cleavage of holo-ACP yields 4′-phosphopantetheine (in addition to apo-ACP).

3.7.7.2 Biochemical Function

The thioester bonds between acids and CoA or pantetheine are energy-rich bonds ($\Delta G'_0$ for hydrolysis = −31.5 kJ/mol, 3.1.3.3). Their formation corresponds to an activation of the acid for synthetic reactions, such as esterification, amidation (e.g., 3.1.7), anhydride formation (e.g., Fig. 3.10.5-4), C-C bond formation (e.g., 3.1.3.3) etc. Best known is acetyl-CoA, the ‘activated acetic acid’ (3.1.3.3).

$$\text{R-COOH} + \text{ATP} + \text{CoA-SH} = \text{Acyl-CoA} + \text{AMP} + \text{PP}_i$$

$$\text{Acyl-CoA} + \text{X} = \text{Acyl-X} + \text{CoA-SH}.$$

During fatty acid synthesis (3.4.1.1), the pantetheine residue of ACP acts as a ‘movable arm’ to move the bound acid from one catalytic center to the next. A similar task is performed by pantetheine ‘arms’ in non-ribosomal polypeptide synthesis (e.g., Fig. 3.10.8-1).

Literature:

Friedrich, W. *Vitamins*. De Gruyter, 1988, p. 809–835.

Jackowski, S. in Neidhardt, F.C. (Ed.). *Escherichia coli and Salmonella.* 2nd Ed. ASM Press, 1996, p. 687–694.

Kupke, T., Hernández-Acosta, P., Culiáñez-Macià, F.A. J. Biol. Chem. 2003;278:38229–38237.

3.7.8 Biotin

Biotin is the essential cofactor for carboxylation reactions. It is covalently enzyme-bound by an amide bond to the ε-amino group of a lysine residue.

3.7.8.1 Biosynthesis and Interconversions (Fig. 3.7.8-1)

The biosynthesis is performed by bacteria (including bacteria in the intestines of animals), yeasts and higher plants in an identical way.

It starts from pimeloyl-CoA. The origin of this compound is not completely known. In *E. coli* it possibly results from the condensation of three molecules malonyl-CoA and the release of two molecules CO_2. In *Bacillus subtilis* it it synthesized from ACP-bound fatty acids involving a C-C bond cleavage by a cytochrome P_{450} containing enzyme. The following condensation with L-alanine is pyridoxal-PP dependent and involves the loss of CO_2 and CoA-SH, similar to the δ-aminolevulinate synthesis (3.3.1). The product, 8-amino-7-oxopelargonate (KAPA) undergoes a transaminase reaction, yielding 7,8-diaminopelargonate. In this reaction, S-adenosylmethionine (3.2.5.4) acts as an amino group

donor (instead of its more common role in transmethylation reactions). An ATP-dependent introduction of CO_2 closes the imidazolidone ring. Then a stereospecific introduction of S (originating from L-cysteine) and closure of the thiophane ring takes place. The result is (+)-biotin (one of eight possible stereoisomers).

The binding of biotin to the respective enzymes requires an ATP-dependent activation, catalyzed by special ligases (biotin-[enzyme] ligases). In bacteria, the ligases are allosterically regulated by several effectors. If all of the apoenzyme has been saturated, additional activated biotin is transferred to a repressor and, by acting as corepressor, decreases the expression of the biosynthetic enzymes.

In animal food, protein-bound biotin is released by biotinidase and thereafter resorbed. Biotinylation of the carboxylase enzymes takes place mainly in the liver.

A supply of 100 … 200 µg/day is required for human adults. Avitaminoses are rare. However, genetic defects of all biotin-dependent enzymes have been observed, as well as of biotinidase, which causes insufficient biotin resorption. Biotin-binding proteins are present in whites and yolks of bird's eggs. These compounds form extremely tight complexes with biotin (notably avidin: $K = 10^{-15}$ mol/l). If native avidin is present in food, the avidin-biotin complex cannot be resorbed.

Degradation by microorganisms: In *Pseudomonads* (and probably in other microorganisms), β-oxidation shortens the side chain and removes the C atoms of the thiophane ring. The rest of the imidazolidone ring is released as urea. Mammals can only remove the side chain and form the sulfoxide compound. Most biotin is, however, excreted in unchanged form.

3.7.8.2 Biochemical Function (Fig. 3.7.8-2)

The essential function of biotin is the transfer of carboxyl groups. Enzyme functions involved in this process are biotin carboxylase (BC), which loads HCO_3^- on biotin, transcarboxylase (TC), which transfers CO_2 to an acceptor, biotin carrier protein (C), which provides an 'anchor' for the biotin side chain and biotin decarboxylase (BD). They can be present as a single multifunctional enzyme or as separate entitites. Various combinations of them carry out the following reaction types:

- Carboxylations (by BC, TC and C functions):

  The enzyme-catalyzed activation of bicarbonate leads via a carboxyphosphate intermediate to the carboxylation of biotin at the N-1 position. The side chain of biotin acts as a flexible arm, which enables the movement of the carboxylated biotin residue from the biotin carboxylase to the transcarboxylase reaction centers. There, CO_2 is transferred to an acceptor molecule (Fig. 3.4.1-2). Examples are pyruvate carboxylase (3.1.3.4), propionyl-CoA carboxylase (3.2.5.4) and methylcrotonyl-CoA carboxylase (3.2.6.2). Most important is acetyl-CoA carboxylase (3.4.1.1):

  $$HCO_3^- + ATP + biotin = ADP + P_i + biotin\text{-}CO_2$$
  $$Biotin\text{-}CO_2 + acetyl\text{-}CoA = biotin + malonyl\text{-}CoA$$

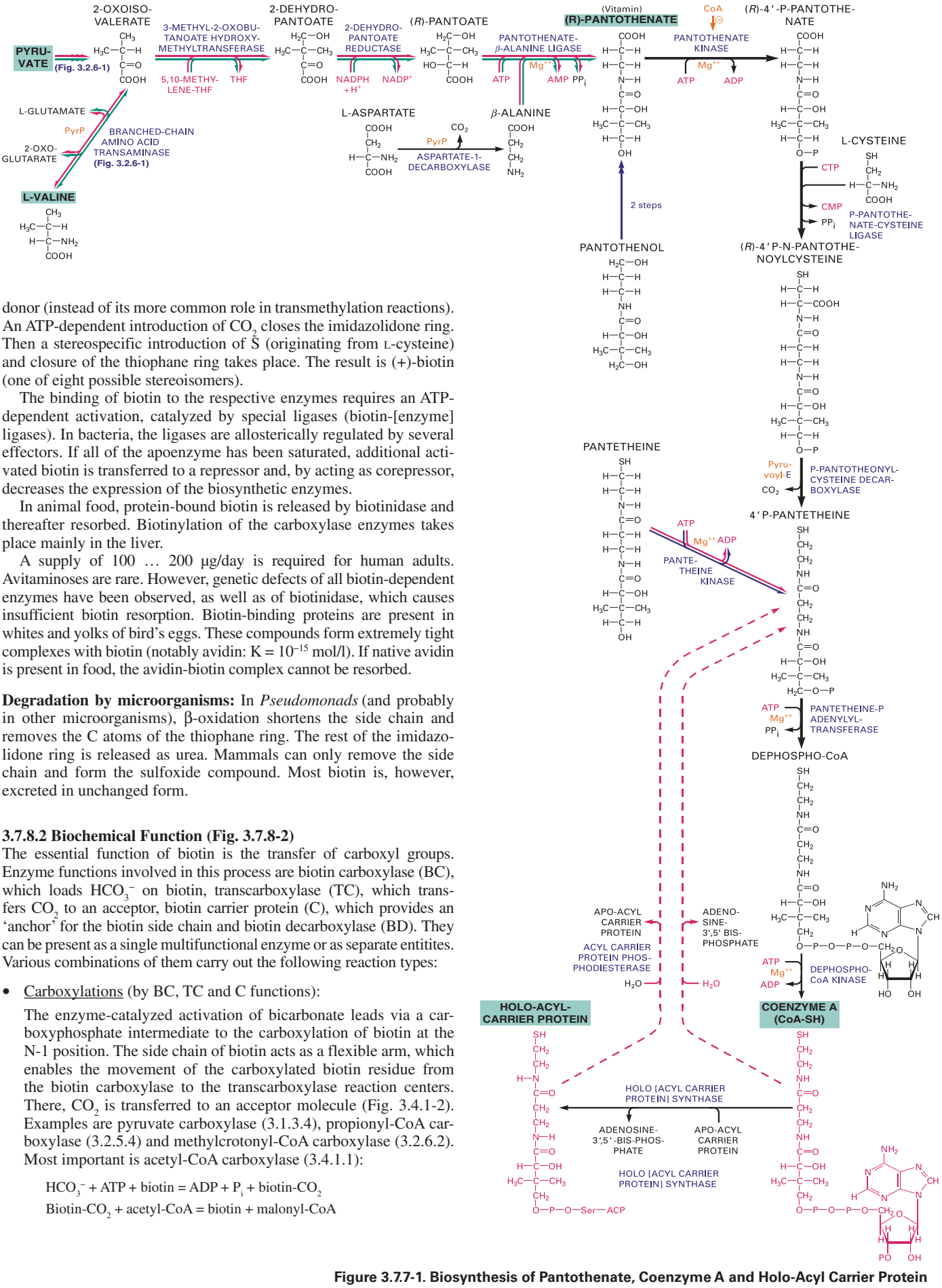


Figure 3.7.7-1. Biosynthesis of Pantothenate, Coenzyme A and Holo-Acyl Carrier Protein

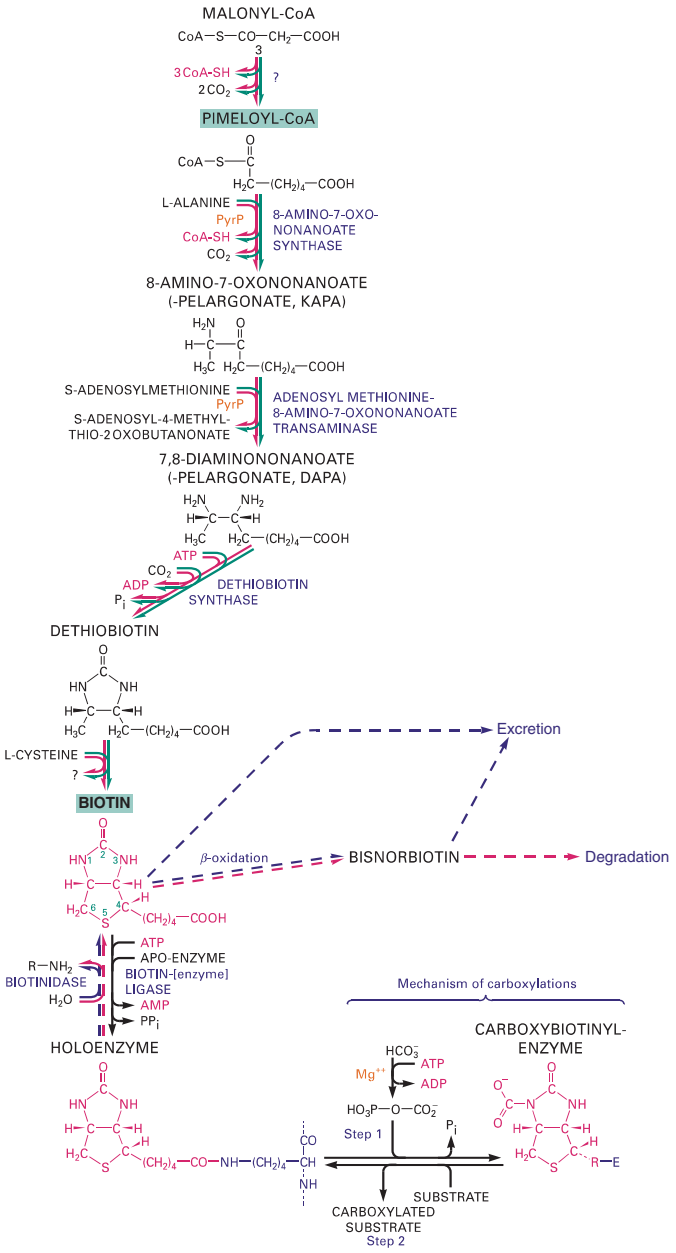


Figure 3.7.8-1. Biosynthesis and Degradation of Biotin

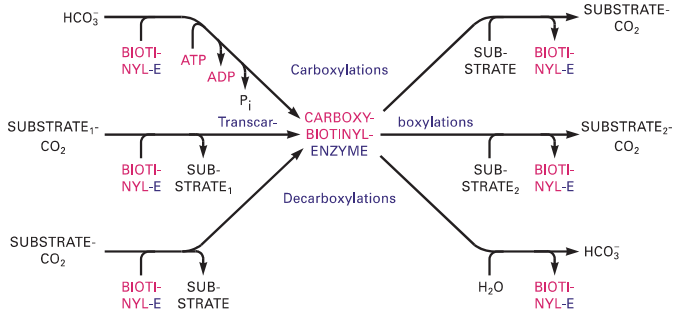


Figure 3.7.8-2. Reaction Types of Biotin

- Transcarboxylations (by 2 TC and 1 C functions):

  One carboxyl transferase function transfers CO_2 from a carboxyl donor to biotin, while the other moves it on to the carboxyl acceptor compound. No activating reaction is needed. An example is methylmalonyl-CoA carboxyltransferase (Fig. 3.1.5-2):

  Methylmalonyl-CoA + biotin = propionyl-CoA + biotin-CO_2
  Biotin-CO_2 + pyruvate = biotin + oxaloacetate

- Decarboxylations (by TC and BD functions)

  Biotin accepts CO_2 from a substrate, the decarboxylase function removes it. Examples are oxaloacetate decarboxylase or methylmalonyl-CoA decarboxylase (In bacteria, both enzymes energize sodium pumps in this way):

  Methylmalonyl-CoA + biotin = propionyl-CoA + biotin-CO_2
  Biotin-CO_2 + H_2O = biotin + H_2 CO_3.

Literature:

Friedrich, W. *Vitamins*. De Gruyter (1988), p. 752–805.

DeMoll, in Neidhardt, F.C. (Ed.) *Escherichia coli and Salmonella.* 2nd Ed. ASM Press, 1996, 704–709.

Stok J.E, De Voss J. Arch Biochem Biophys. 2000;384:351–360.

3.7.9 Nicotinate, NAD+, and NADP+

Nicotinate (niacin) and nicotinamide are precursors of the coenzymes nicotinamide-adenine dinucleotide (NAD±) and nicotinamide-adenine dinucleotide phosphate (NADP±), jointly known as nicotinamide (or pyridine) nucleotides. By interconversion with their reduced forms, NADH and NADPH, they participate in several hundred redox reactions. These coenzymes therefore occupy a central role in metabolic processes of all living beings.

3.7.9.1 Biosynthesis and Degradation of NAD+ and NADP+ (Fig. 3.7.9-1)

Nicotinate is not a vitamin in a strict sense, since it can be synthesized in animals from tryptophan (which, however, is an essential amino acid).

An additional supply of nicotinate is only required when the supply of tryptophan is insufficient (e.g., by food containing mainly maize or sorghum → pellagra) or in order to satisfy peak demands. 18 mg/day are considered sufficient for human adults.

The eukaryotic catabolism of tryptophan yields 2-amino-3-carboxymuconate semialdehyde as an intermediate (Fig. 3.2.7-4). This compound cyclizes nonenzymatically to quinolinate. In *E. coli*, aspartate oxidation produces iminoaspartate, whose reaction with dihydroxyacetone-P (glycerone-P) and a consecutive ring closure also results in quinolinate. The FAD-containing oxidase is feedback inhibited by NAD+. Quinolinate synthase is O_2-sensitive. Decarboxylation and conversion to nicotinate mononucleotide are performed by a ribosyltransferase.

Adenylate transfer from ATP yields deamido-NAD+, which is then converted into the amide, NAD+. Bacteria use NH_3 for this reaction, while eukarya obtain the amino group from glutamine or use NH_3. An additional phosphorylation at a ribose moiety results in NADP+. Also a reconversion to NAD+ by a phosphatase reaction takes place.

Members of the family Pasteurellaceae do not use either the NAD *de novo* biosynthesis pathway or the NAD salvage pathway and rely on the presence of NAD or related compounds in the growth medium. They have been classified into two distinct classes based on their specific requirement: members of the so called "V-factor-dependent" class, which includes the pathogen *Haemophilus influenzae*, require supplements in which the pyridine nucleotide source must possess an intact pyridine-ribose bond and the pyridine-carbonyl group must be amidated. Thus, they require either NAD, nicotinamide mononucleotide (NMN) or nicotinamide riboside (NR). On the other hand, members of the "V-factor-independent" class possess the enzyme nicotinamide phosphoribosyl transferase, which converts nicotinamide (NAm) to NMN, and are able to to synthesize NAD from NAm as well.

Salvage reactions: In aerobic bacteria, the half-life time of NAD+ is only about 90 minutes. The valuable nicotinate moiety has to be recovered by pyridine nucleotide cycles. In most cases, the primary step of NAD+ degradation is the cleavage by pyrophosphatase. Another NAD+ degradation reaction is the direct removal of the nicotinamide residue. In both cases, consecutive steps lead to nicotinate, which reacts with phosphoribosyl pyrophosphate to yield nicotinate mononucleotide. Its reconversion to NAD+ proceeds as described above.

Degradation: In bacteria, nicotinate is hydroxylated and then decarboxylated. Oxidative ring opening and further reactions lead to fumarate, which is a component of the citrate cycle (3.1.8.1).

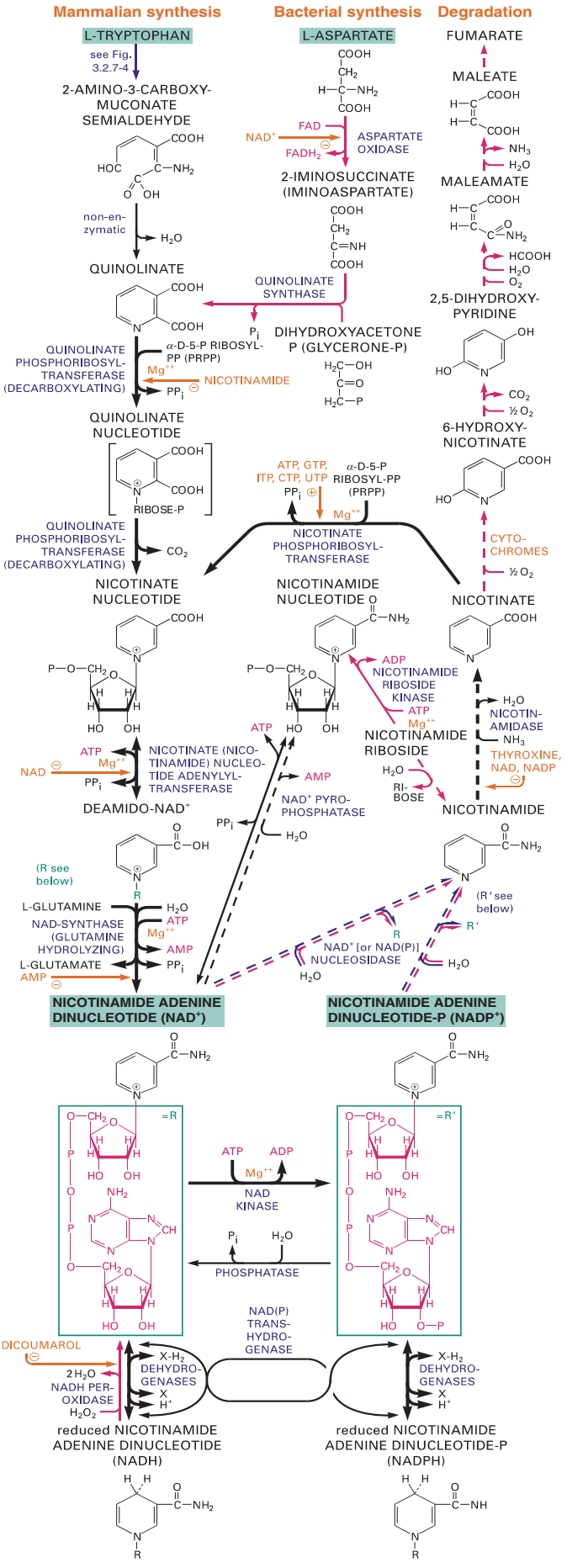


Figure 3.7.9-1. Biosynthesis of Nicotinamide, NAD+ and NADP+

3.7.9.2 Mechanism of the Redox Reactions, Stereospecificity

The reduction of the nicotinamide group proceeds by the uptake of 1 proton H^+ and 2 electrons e^- (formally a hydride ion, H^-). Another hydrogen of the substrate is released as a proton.

The positive charge on the pyridine nitrogen and the aromatic ring character is lost. This is accompanied by a spectral change (Fig. 3.7.9-2), which can be used for analytical work. The general reaction equation can be written as

$$NAD^+ \text{ (or } NADP^+) + XH_2 = NADH \text{ (or NADPH)} + X + H^+$$

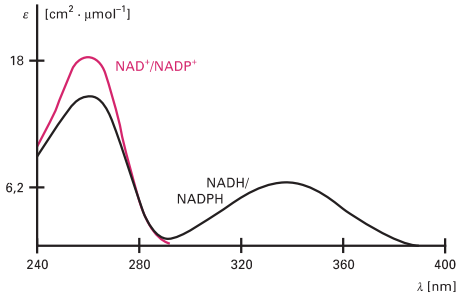


Figure 3.7.9-2. Absorption Spectra of NAD+/NADP+ and NADH/NADPH

This hydrogen transfer is stereospecific: the added H is located either above the plane of the nicotinamide ring (A side or pro-R side according to the Cahn-Ingold-Prelog 'RS'-system) or below (B side or pro-S side, Fig. 3.7.9-3). E.g., alcohol, malate and isocitrate dehydrogenases are 'A' enzymes, 2-oxoglutarate, glucose-6-phosphate and glutamate dehydrogenases are 'B' enzymes. Likewise, the removal of hydrogen from the substrate proceeds stereospecifically.

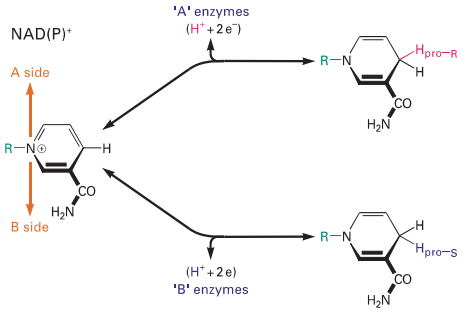


Figure 3.7.9-3. Stereospecifity of Dehydrogenase Reactions

3.7.9.3 Biochemical Function of the Nicotinamide Coenzymes

Although the redox potentials of the NAD^+/NADH and the $NADP^+$/NADPH systems are almost identical (E'_0 = −320 and −324 mV), in living organisms a ratio of the oxidized/reduced dinucleotides of about 200 … 1000 with NAD and 0.01 with NADP is maintained.

- Via a membrane-bound transhydrogenase inside of mitochondria and bacteria, hydrogen exchange between internal NAD^+ and $NADP^+$ takes place (Fig. 3.11-1). The reaction is driven by the proton gradient across the membrane and is essential for maintaining the different redox states of both coenzymes:

$$NADH + NADP^+ + H^+_{outside} = NAD^+ + NADPH + H^+_{inside}$$

- In other compartments, substrate concentrations cause the differences of the redox situation. The coenzyme specificity of most NAD^+ or $NADP^+$ employing enzymes prevents an equilibration of the redox states.
- Due to the prevalence of the reduced form, NADPH is well suited for reductive biosyntheses, e.g., of fatty acids and steroids. In erythrocytes, NADPH keeps up the glutathione (GSH) concentration, which is needed for the removal of membrane-damaging H_2O_2 (3.2.5.7). On the other hand, NADPH oxidation in activated neutrophil granulocytes and macrophages leads to

aggressive superoxide radicals, which play a role in cellular defense (3.2.5.8).

- NAD^+ usually accepts hydrogen during oxidation of metabolites and delivers them to the cytochrome system for terminal oxidation (3.11.4.4) or transfers them to other substrates during fermentations (3.10.5). NAD^+ supplies the ribose moiety of α-ribazole during coenzyme B_{12} biosynthesis (3.7.5). It is also a substrate for the modification of proteins by ADP-ribosylation (e.g., 7.4.1). A bacterial DNA ligase uses this reaction to energize the resealing of nicks in DNA repair (Fig.3.7.9-4).

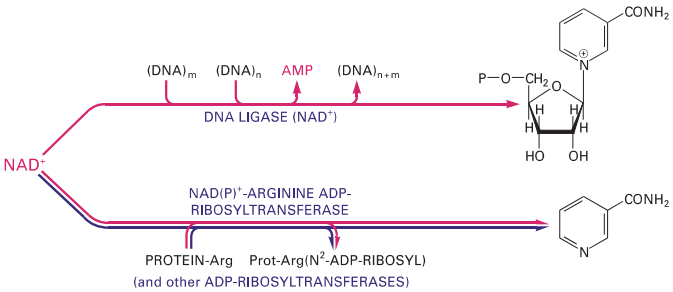


Figure 3.7.9-4. Ligase and ADP-Ribosylation Reactions Employing NAD^+

Literature:

Creighton, D.J., Murphy, N.S.R.K in Sigman, D.S., Boyer, P.D. (Eds.) *The Enzymes*. 3rd Ed., Vol. 19. Academic Press, 1990, p. 323–421.

Dolphin, D. *Pyridine Nucleotide Coenzymes: Chemical, Biochemical and Medical Aspects*. Wiley, 1987.

Friedrich, W. *Vitamins*. De Gruyter, 1988, p. 475–542.

Penfound, D., Foster, J.W. in: Neidhardt, F.C. (Ed.) *Escherichia coli and Salmonella*. 2nd Ed. ASM Press, 1996, p. 721–730.

3.7.10 Ascorbate (Vitamin C)

Vitamin C takes part in many reactions involving oxygen and frequently exerts a protective function.

3.7.10.1 Biosynthesis and Metabolism (Fig. 3.7.10-1)

Ascorbate can be biosynthesized by higher plants, algae, yeast and most animals. An adult goat produces ~13 g of vitamin C per day. This ability is lacking in some mammals (including humans and guinea pigs) and also in insects, invertebrates and most fishes. In those mammals, which require external ascorbate supply, the biosynthetic enzyme L-gulonolactone oxidase is absent.

The biosynthesis in animals and most plants proceeds from D-glucose via UDP-glucose, UDP-glucuronate, D-glucuronate, L-gulonate, L-gulonolactone and 2-dehydro-L-gulonolactone. Some algae proceed via the D-galacturonate pathway with essentially analogous steps. Plants possess alternate routes for the synthesis of ascorbate via L-gulonate or the L-galactose pathway starting from GDP-L-galactose. In some bacteria (*Acetobacter sp.,Gluconobacter sp., Ketogulonicigenium sp.*) the synthesis proceeds from D-sorbitol via L-sorbosone to L-ascorbate.

After resorption of ascorbate from food, it is transported in blood as an albumin complex to the various organs. In some of them, uptake is an active process, promoted by insulin. The highest concentrations are found in the adrenal and pituitary glands, eye, liver, pancreas, thymus and brain. In humans, scurvy is prevented by an intake of 10 mg/day. Stress, smoking and pregnancy increase the demand. A daily allowance of ca. 45-90 mg is recommended, but this is under discussion. It is claimed that elevated doses of ascorbate strengthen the immune system, prevent easy tiring and bleeding and even decrease the risk of cancer. However, after reaching a pool size of 1500 mg by a daily intake of 50 mg, degradation and excretion of ascorbate increases greatly.

Degradation and excretion: In primates, part of the body ascorbate is excreted unchanged in urine. Other excretion products are oxalate and dioxogulonate (formed from dehydroascorbate by a nonenzymatic, irreversible reaction). Also, complete oxidation to CO_2 takes place (the preferred method of catabolism in rodents).

3.7.10.2 Biochemical Function (Fig. 3.7.10-2)

Ascorbate participates in many important redox reactions. Many of them involve oxygen. The E_0' for the redox couple dehydroascorbate/ascorbate is + 58 mV ($2e^-$ reaction). However, in most cases, a one-electron transfer takes place. This leads to the monodehydroascorbate radical, which is then reconverted to ascorbate either directly or via dehydroascorbate.

Important examples for ascorbate involvement are:

- Dioxygenases, e.g., proline and lysine dioxygenases (3.2.3, 3.2.5.2). If the substrates are not available, but the cofactor oxoglutarate is present, the enzyme-Fe^{2+} is oxidized to Fe^{3+} and requires stoichiometric amounts of ascorbate to be reduced again. Under *in vivo* conditions, smaller amounts suffice. A Fe-ascorbate-substrate complex as an intermediate of the enzymatic reaction has been discussed in the literature. Also other enzymes, such as 4-hydroxyphenylpyruvate dioxygenase (3.2.7.3) and homogentisate 1,2-dioxygenase (3.2.7.3) require ascorbate, although their involvement is less clear. Ascorbate also promotes the hydroxylation of xenobiotics.
- A number of monooxygenases, e.g., Cu^{2+} containing dopamine β-monooxygenase (3.2.7.3). Stoichiometric amounts of ascorbate are required for reduction of the metal as part of the catalytic cycle.
- Reduction of heme proteins, which previously have been oxidized by H_2O_2 to the ferryl level (e.g., myoglobin).
- Direct reaction of ascorbate with activated oxygen species (3.2.5.8), mostly for detoxification and protection purposes. This way, the eye is protected against light-activated oxygen or chloroplasts against dangerous side reactions of the photosynthesis apparatus (generation of radicals or singlet oxygen). Non-enzymatic reactions are:
  - Removal of the superoxide radical

    $O_2^{-\bullet}$ + ascorbate + H^+ = H_2O_2 + monodehydroascorbate$^\bullet$
  - Removal of hydroxyl and peroxyl radicals

    $HO^\bullet$ + ascorbate = H_2O + monodehydroascorbate$^\bullet$
    $ROO^\bullet$ + ascorbate = ROOH + monodehydroascorbate$^\bullet$
  - Quenching of singlet oxygen $^1O_2$.

In plants, H_2O_2 is removed by the L-ascorbate peroxidase reaction (Fig. 3.7.10-2). Ascorbate also reacts with molecular oxygen, either nonenzymatically (especially in the presence of heavy metal ions) or catalyzed by ascorbate oxidase:

O_2 + 4 ascorbate = 2 H_2O + 4 monodehydroascorbate$^\bullet$

The monodehydroascorbate formed in these reactions is reconverted to ascorbate by NADH, by cytochrome b_5 (fatty acid desaturation system) or by ferredoxin (in plants, from photosystem I, 3.12.1). Alternatively, it disproportionates into ascorbate and dehydroascorbate. The latter is reduced to ascorbate by glutathione in a reaction catalyzed by dehydroascorbate reductase. The oxidized glutathione, in turn, is reduced again by NADPH.

Generally, ascorbate has a cytoprotective function. In certain conditions, however, ascorbate can generate reactive oxygen species, e.g., superoxide and hydroxyl radicals and thus cause toxic effects. This occurs mostly in presence of heavy metal ions.

Scurvy is caused by insufficient hydroxylation of proline (and to some extent, lysine) in procollagen, taking place in the endoplasmic reticulum, which leads to a reduced number of hydrogen bridges in collagen (2.3.1). This decreases the stability and the water-binding capacity of the collagen triple helix. Apparently, side reactions cause an oxidation of Fe^{2+} in the hydroxylating enzymes, which cannot be reversed due to the lack of ascorbate. Scurvy is characterized by weakness, bleeding, swelling of gums and brittle bones. Untreated scurvy is invariably fatal. Moeller-Barlow disease is childhood scurvy.

Literature:

Friedrich, W. *Vitamins*. De Gruyter, 1988, p. 931–1001.

Rice-Evans, C.A., Burdon, R.H. (Eds.) *Free Radical Damage and its Control. New Comprehensive Biochemistry* Vol. 28. Elsevier, 1994.

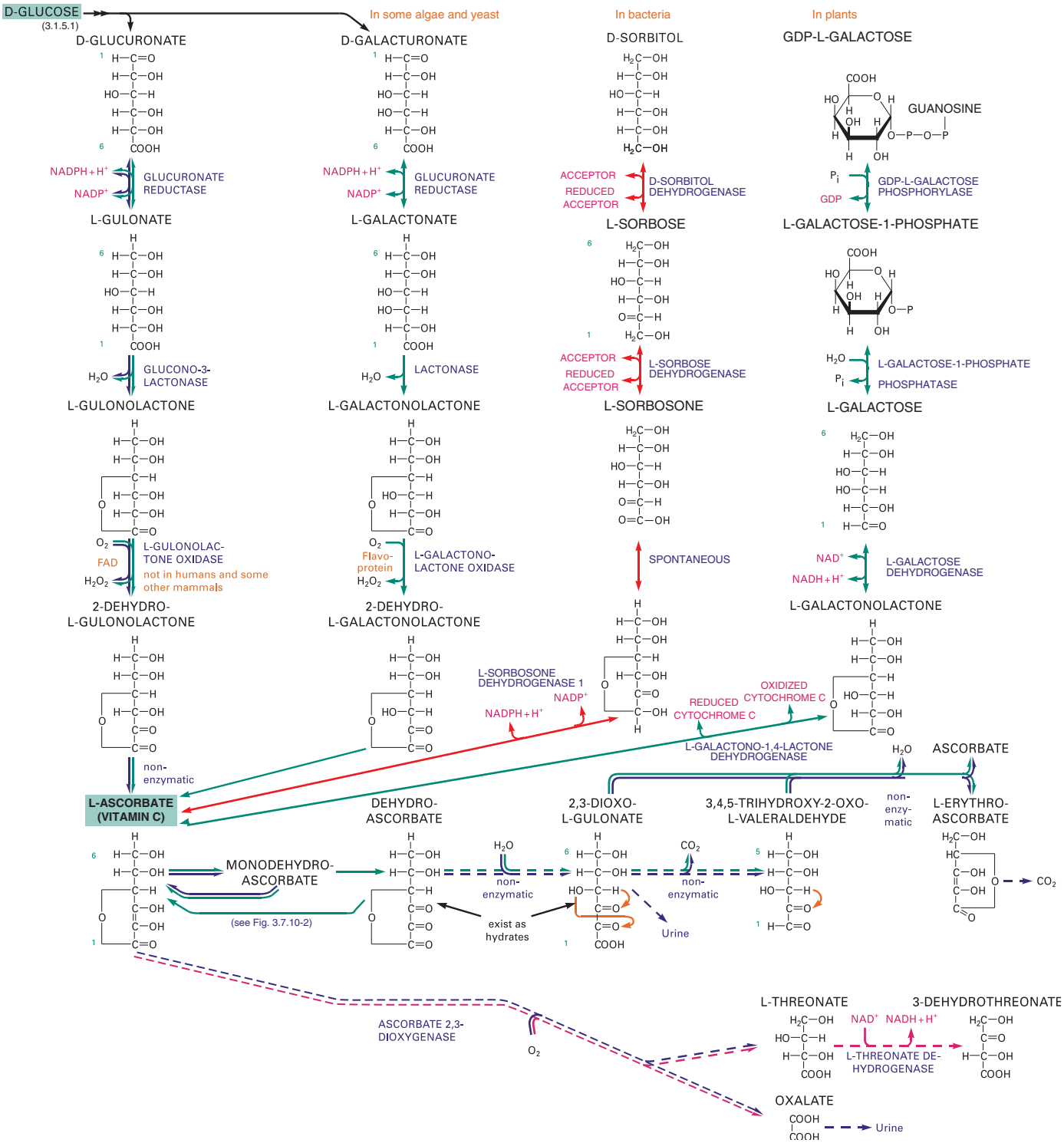


Figure 3.7.10-1. Biosynthesis and Degradation of Ascorbate

3.7.11 Calciferol (Vitamin D)

In higher animals, hydroxylation products of vitamin D (calciferol) play a central role in calcium metabolism.

Actually, the term 'vitamin D' encompasses a group of related compounds, but of different origin. They are not vitamins in the strictest sense, because one of them (D_3) can be formed in the skin by ultraviolet irradiation from a steroid (7-dehydrocholesterol, provitamin D), which is synthesized in the body.

3.7.11.1 Biosynthesis and Interconversions

The various D vitamins are formed from $\Delta^{5,7}$ steroids [which are intermediates in the biosynthesis of zoo- (3.5.1), phyto- and mycosteroids (3.5.2)] by light-induced opening of the B ring. They differ only in their side chains. The first two of the following are most important:

- Vitamin D_2 (ergocalciferol), obtained from ergosterol (3.5.2.2, present in plants and fungi)
- Vitamin D_3 (cholecalciferol), obtained from 7-dehydrocholesterol (3.5.1.1, present in higher animals)
- Vitamin D_4, obtained from 22-dihydroergosterol (present in plants and fungi)
- Vitamin D_5, obtained from 7-dehydrositosterol (present in plants).

Endogeneous formation: In a first reaction step, UV illumination causes an isomerization with the simultaneous opening of ring B of

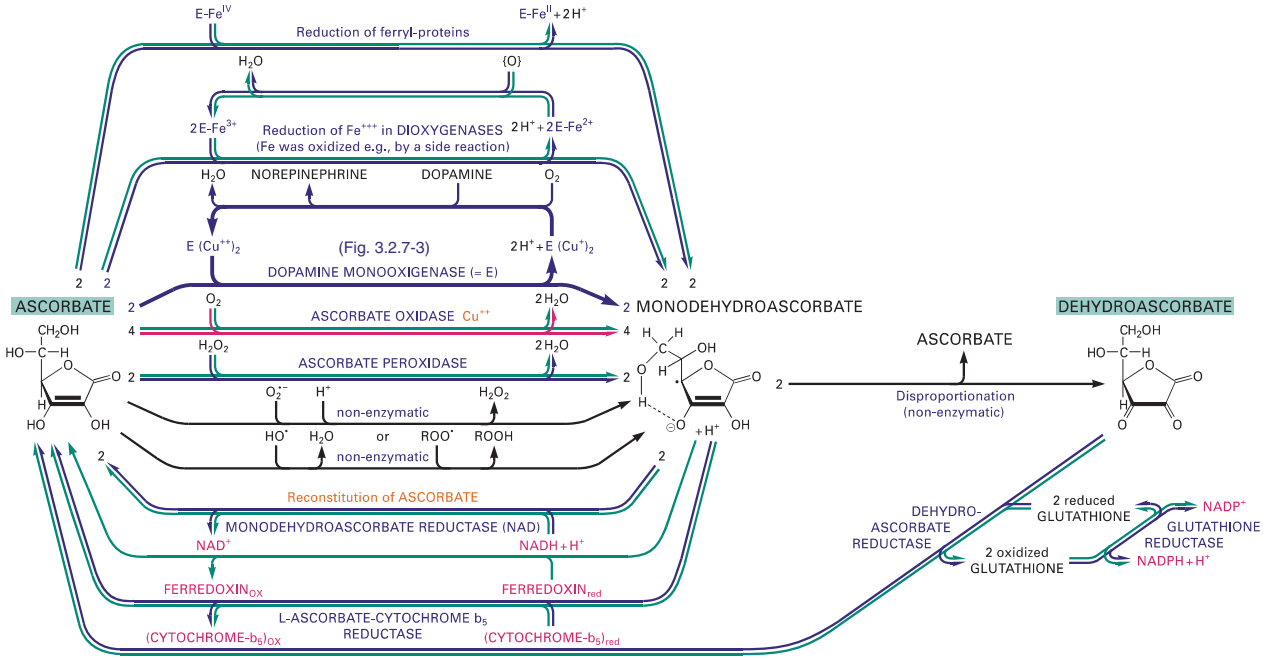


Figure 3.7.10-2. Important Reactions of Ascorbate

the sterol structure, yielding previtamin D. By another, temperature-dependent isomerization, vitamin D is obtained. This compound then enters the bloodstream and is bound to vitamin D binding protein (DBP). The optimum wavelength for the photoreaction is 295 nm. Irradiation at longer wavelengths leads to lumisterol, at shorter wavelengths to tachysterol, while too intense irradiation results in the formation of suprasterol I and II.

Supply by food: Although illumination can provide sufficient vitamin D, additional intake is recommended (for human adults 5 µg/day, for children 10 µg/day). After (bile acid assisted) resorption, dietary vitamin D is transported to the liver by chylomicrons or bound to DBP.

In the liver, hydroxylation at C-25 takes place through a cytochrome P-450 dependent enzyme (Fig. 3.5.4-2). The compound is then transferred to the kidney, where the biologically active 1α,25-dihydroxycalciferol (calcitriol, DHCC) is formed by another hydroxylation at C-1. An alternative hydroxylation leads to 24(R),25-dihydroxycalciferol. The hydroxylations also occur in the placenta and in monocytes/macrophages.

1α,25-dihydroxycalciferol is of great importance for the calcium homeostasis, therefore its formation is strictly regulated. Parathyroid hormone (parathormone, PTH, 7.1.7) stimulates the 1-hydroxylase and inhibits the 25-hydroxylase. High plasma levels of Ca^{++} repress the PTH synthesis and thus lower calcitriol formation. Calcitriol represses the synthesis of the 1-hydroxylase and enhances the synthesis of the 25-hydroxylase.

Degradation takes place by additional hydroxylations and/or oxidation of hydroxyl to carboxylic groups. These more polar compounds show little biological activity and are excreted via feces or urine.

3.7.11.2 Biochemical Function (Fig. 3.7.11-1)

In mammals, 'vitamin D' (actually 1α,25-dihydroxycalciferol) regulates the calcium and phosphate metabolism (7.1.7).

It activates the (re)absorption of Ca^{2+} and phosphate in the intestine and kidney and thus contributes to their deposition in bones, concomitant with PTH and calcitonin. On the other hand, it mobilizes Ca^{2+} from bone for short time regulation of the blood Ca^{++} level.

Apparently this is effected by binding of the vitamin to vitamin D receptors (7.7). The complexes induce the transcription of genes for calcium binding proteins in intestine, pancreas and bones (CaBP, calbindin, osteocalcin etc.). These proteins, which contain

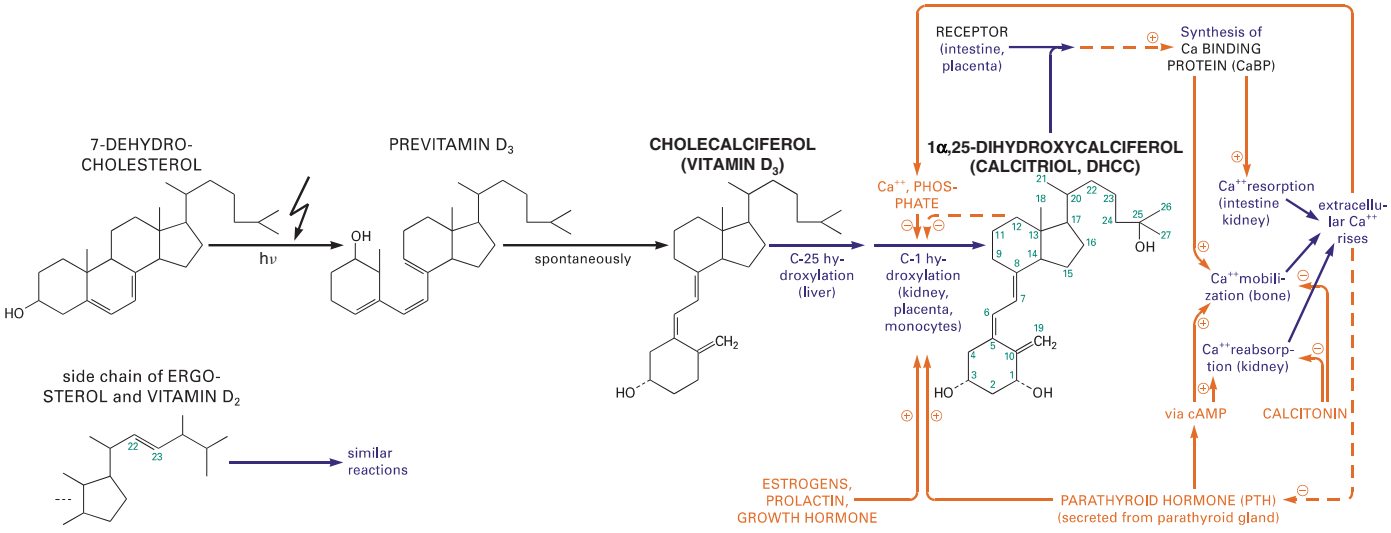


Figure 3.7.11-1. Formation and Effects of Vitamin D_3 and 1α,25-Dihydroxycalciferol

γ-carboxyglutamate, are involved in calcium uptake and bone mineralization. This way, the vitamin acts in a hormone-like fashion. It is also involved in the differentiation of macrophages into osteoclasts. To some extent, 24,25-dihydroxycalciferol is also involved in these reactions.

Medical aspects: Diseases associated with abnormal levels of vitamin D are

- Rickets (diminished mineralization of the skeleton, mostly during childhood). This is caused by reduced resorption of Ca^{2+} due to lack of vitamin D, sometimes also by deficient formation of calcitriol from calciferol.
- Osteomalacia (softening of bones) and osteoporosis due to impaired calcitriol formation. This is caused by kidney failure or by a reduced level of estrogens, which stimulate the hydroxylation step.
- Autoimmune diseases. Lack of vitamin D is a risk factor for autoimmune diseases such as lupus erythematosus, diabetes type I or multiple sclerosis. Vitamin D binds to nuclear vitamin D receptors (VDR) which are present in most immune cell types. These exhibit potent immunomodualry functions.
- D-hypervitaminosis. Demineralization of bones occurs by overdosing of vitamin D.

Literature:

Cancela, L. *et al.* In Cooke, B.A., King, R.J.B, van der Molen, H.J. (Eds.) *Hormones and their actions*. Part I. *New Comprehensive Biochemistry*. Elsevier, 1988, Vol. 18A, p. 269–289.

Darwish, H., DeLuca, H.F. Crit. Rev. Eukar. Gene Expr. 1993;3:89–116.

Friedrich, W. *Vitamins*. De Gruyter, 1988, p. 143–216.

3.7.12 Tocopherol (Vitamin E)

The tocopherol group encompasses eight compounds, which consist of a 8-chromanol ring and an isoprenoid side chain. The end product of biosynthesis is α-tocopherol (5,7,8-trimethyltocopherol), which also shows the highest bioactivity. Tocopherols protect lipids in animals and plants from oxidative damage. The various tocopherols differ in the number and position of the methyl groups on the ring and of the double bonds in the side chain. All of them are very lipophilic.

The compounds are biosynthesized from homogentisate only in plants, mainly in chloroplasts of leaves (Fig. 3.2.7-3). They are mainly localized in the cellular membranes of animals and plants (e.g., of mitochondria, erythrocytes and chloroplasts).

The recommended intake for human adults is 8 … 12 mg α-tocopherol/day. It is higher when the food contains many polyunsaturated fatty acids. Avitaminoses in humans are rare, except in premature infants or during disturbances of lipid resorption. However, degeneration of various organs and infertility due to imbalanced food have been observed in animal breeding.

The biochemical function of tocopherols is mainly the protection of membrane lipids from peroxidation (in particular polyunsaturated fatty acids). Tocopherols are internal ethers of hydroquinones. They act as radical quenchers by transition to the tocopheroxy state (semiquinone, Fig. 3.7.12-1). This is an unreactive radical with a half-life time of several hours, which interrupts chain reactions (3.2.5.8). Its formation is reversible. The reduction to tocopherol is most likely performed by ascorbate, which is oxidized in this way to monodehydroascorbate at the lipid-water interface (3.7.10.2). There is a synergism between both vitamins. Also β-carotene was proposed as a reductant. The further oxidation of the semiquinone to the quinone state involves a ring opening and is irreversible.

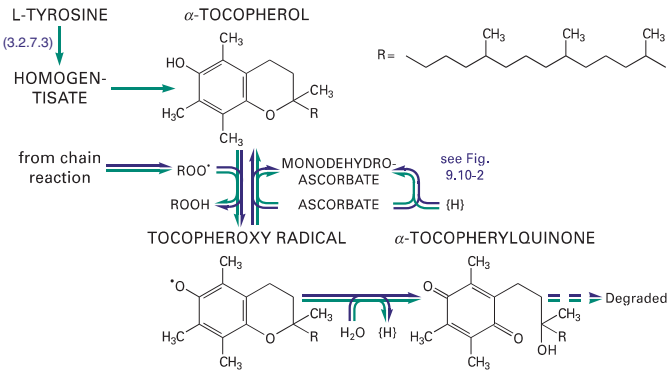


Figure 3.7.12-1. Oxidation States of Tocopherol and its Function as a Radical Quencher

Literature:

Rice-Evans, C.A., Burdon, R.H. (Eds.) *Free Radical Damage and its Control. New Comprehensive Biochemistry* Vol 28. Elsevier, 1994.

Friedrich, W. *Vitamins*. De Gruyter, 1988,143–216.

3.7.13 Phylloquinone and Menaquinone (Vitamin K)

Phylloquinone is a member of the photosystems in photosynthesizing plants (3.12.1). Menaquinone plays a role in anaerobic respiration of bacteria (3.10.6). In animals, both compounds (vitamins $K_1$and K_2, respectively) act as cofactors for the γ-carboxylation of glutamate. They are monomethyl naphthoquinones with an isoprenoid side chain.

Phylloquinone and menaquinone differ only in the length and the number of double bonds in the side chain.

For the biosynthesis of menaquinone and phylloquinone see 3.2.7.2 and Figure 3.2.7-1.

The recommended daily intake for human adults is 70 … 140 µg/day. It is usually met by food and by its production by intestinal bacteria. Degradation before excretion is usually limited to oxidative shortening of the side chain.

The biochemical function of these compounds in plants and bacteria is described in chapters 3.11 and 3.12. In animals, they are cofactors for the formation of γ-carboxyglutamate in proteins (Gla, Fig. 3.7.13-1) by a carboxylase activity of the rough endoplasmic reticulum. The proteins are secreted thereafter. This posttranslational modification is essential for Ca^{2+} binding and for attachment of these proteins to membrane phospholipids via Ca^{2+} bridges (e.g., during blood coagulation, Chapter 9).

Vitamin K hydroquinone abstracts a proton from glutamate, which is carboxylated afterwards. Simultaneously, a K-quinone 2,3-epoxide is formed. The energy gained from the oxygenation enables the generation of a very strong base in order to remove a proton from glutamate. The glutamyl carbanion then undergoes carboxylation to yield γ-carboxyglutamate. The vitamin K epoxide is reduced by a dithiol-linked vitamin K epoxide reductase to the quinone and further on to the quinol hydroquinone level. The reductase step can be inhibited by coumarin derivatives, e.g., dicoumarol and warfarin, resulting in an anticoagulant action.

γ-carboxyglutamate-containing proteins are coagulation factors II, VII, IX, X, proteins C and S (9.3.1, synthesized in the liver), osteocalcin (regulates calcification of bones), ovicalcin (in eggs) and several others.

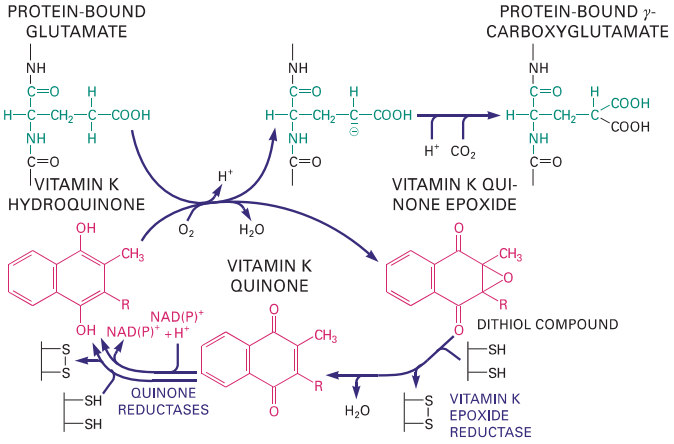


Figure 3.7.13-1. Vitamin K Dependent Carboxylation Reactions

Literature:

Dowd, P., Hershline, R., Ham, S. W., and Naganathan, S. Science 1995;269:1684–1691.

Friedrich, W. *Vitamins*. De Gruyter, 1988, p. 143–216.

Suttie, J.W. Ann. Rev. of Biochem. 1985;54:459–477.

3.7.14 Other Compounds

3.7.14.1 Lipoate as a Cofactor

Protein-bound lipoate is a cofactor of several dehydrogenase complexes, e.g., pyruvate dehydrogenase (3.1.3.1), 2-oxoglutarate dehydrogenase (3.1.8.1) and the glycine cleavage system (3.2.4.2).

The biosynthesis takes place by introduction of sulfur from cysteine into the hydrocarbon chain of octanoic acid (an intermediate of fatty acid biosynthesis, 3.4.1.1) at the positions C-6 and C-8. The reaction proceeds via a radical mechanism. Lipoate synthase is an iron-sulfur protein that uses S-adenosyl-L-methionine. Binding of ATP-activated lipoate to the ε-amino group of lysine in the respective enzymes proceeds in a way analogous to biotin (3.7.8). While most living beings can synthesize lipoate, it is required by some bacteria as a growth factor.

Biochemical function: The redox potential of lipoate/dihydrolipoate (E_0') is -290 mV. The lipoate moiety accepts ligands at the oxo level and oxidizes them to the carboxylic function with concomitant formation of a thioester. Then, acting as a mobile arm (length 1.4 nm), it moves the acyl group to the next enzyme function (compare pantothenate, 3.7.7). After release of the ligand, lipoate disulfide is reconstituted (Fig. 3.7.14-1).

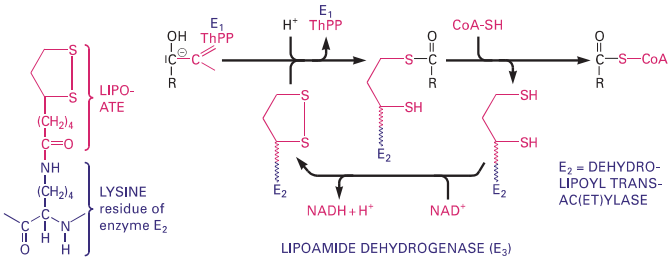


Figure 3.7.14-1. Structure and Biochemical Function of Lipoate
(Examples: Pyruvate and 2-Oxoglutarate Dehydrogenases, compare Figure 3.1.3-4)

Literature:

Berg, A., de Kok, A. Biol. Chem. 1997;378:617–634.

3.7.14.2 Essential Fatty Acids ('Vitamin F')

Mammals are unable to introduce double bonds into fatty acids beyond the Δ^9 position and have to obtain polyunsaturated fatty acids by food intake. Most important are linoleic and linolenic acids. For details see 3.4.1.1.

3.7.14.3 Essential Amino Acids

Mammals cannot synthesize nine of the amino acids and are dependent on their supply in food. For details see 3.2.1.

3.8 Nucleic Acid Metabolism in Bacteria

Susanne Peifer and Elmar Heinzle

3.8.1 Bacterial DNA Replication

DNA, the carrier of genetic information, has to be replicated before cellular division takes place (except meiotic division in higher eukarya, 3.9.1). The semi-conservative mechanism uses the parental DNA duplex strands as templates for replication; each parental strand forms a new duplex with a newly synthesized strand.

The mechanism of bacterial DNA replication described here applies to many bacterial chromosomes, episomes and plasmids. Unlike eukaryotic chromosomes (3.9.1.2), each of these circular, covalently closed, double-stranded molecules is replicated as a single unit (replicon). The same holds true for viral genomes of the circular duplex type.

Replication of bacterial chromosomes starts from a single origin. Separation of both DNA strands yields two replication forks, which proceed bidirectionally around the molecule. Both strands are used as templates for synthesis of new DNA strands (Θ-mechanism, Fig. 3.8.1-1). In some plasmids and bacteriophages, replication is unidirectional: one of the circular DNA strands is opened and the other one is used as a template for replication, sometimes resulting in a series of consecutive strand copies (σ-mechanism, rolling circle mechanism). The newly synthesized strand then directs the synthesis of the complementary strand.

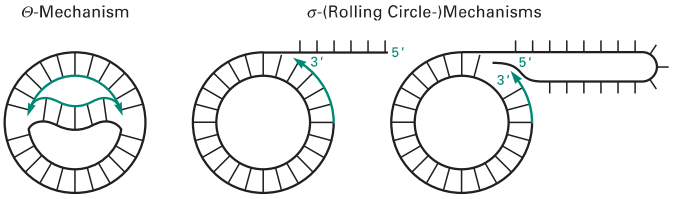


Figure 3.8.1-1. Mechanisms of Bacterial Replication

Table 3.8.1-1. Bacterial and Viral Replicons

Replicon	Size (base pairs)	Mechanism
Calothrix chromosome	$1.6 * 10^7$	Θ
Sorangium cellulosum chomosome	$1.3 * 10^7$	Θ
Escherichia coli chromosome	$4.7 * 10^6$	Θ
Mycoplasma chromosome	$5.8 * 10^5$	Θ
Bacteriophage T4 (linear chromosome)	$1.7 * 10^5$	Θ
E. coli F-Episome	$1 * 10^5$	σ
Bacteriophage λ	$4.8 * 10^4$	σ
ColE1 plasmid	$6.6 * 10^3$	σ
Bacteriophage ΦX174	$5.4 * 10^3$	σ

Replication proceeds via a polymerization reaction in 5′→3′ direction (Fig. 3.8.1-2):

$$(dNMP)_n + dNTP \rightarrow (dNMP)_{n+1} + PP_i \qquad \Delta G_0' = +2 \text{ kJ/mol},$$

that is driven by the subsequent hydrolysis of the liberated PP_i:

$$PP_i + H_2O \rightarrow 2\ P_i \qquad \Delta G_0' = -30 \text{ kJ/mol}.$$

The correct nucleotide (dATP, dCTP, dGTP, dTTP) is selected by its ability to form a Watson-Crick base pair (2.6.1) with the corresponding nucleotide at the template strand. The reaction is a nucleophilic attack of the terminal 3′ hydroxyl in the growing DNA chain on the α-phosphate of the incoming dNTP.

The replication rate in bacteria is about 500 … 1000 nucleotides/sec (vs. 30 … 50 nucleotides/sec in eukarya). The *Escherichia coli* chromosome is copied in about 40 minutes. Cells with doubling times < 40 minutes (e.g., *E. coli* 20 minutes) start another initiation round before the previous replication is completed. Thus, they are partially polyploid (having more than one DNA copy per cell).

3.8.1.1 Cell Cycle and Replication

Chromosomal DNA replication is linked to the cell cycle by a complex regulation system. Initiation, being the critical step, is tightly controlled. During the replication reaction, the *oriC* region (3.8.1.2) of the newly formed strand remains unmethylated and the duplex DNA binds to the cell envelope. Thus, premature DnaA binding is prevented (which would otherwise start another replication round). Only later in replication, the newly synthesized strand becomes also methylated. Furthermore, the concentration of DnaA protein plays an important role for the initiation of DNA replication by initiation of the DNA separation.

3.8.1.2 Initiation of Replication (Fig. 3.8.1-6, Table 3.8.1-2)

The chromosomal replication in bacteria takes place from a single replication origin (*oriC*, length 260 bp, Fig. 3.8.1-3) and is mediated by a single initiator protein, DnaA. This protein interacts specifically with

five nonpalindromic sequences (DnaA boxes, length 9 bp) at *oriC*. Four of these sites are almost identical and contain the consensus sequence TTAT(C/A)CA(C/A)A. The left side of *oriC* is an AT-rich region, consisting of three similar sequences (length 13 bp, each starting with GATC), which represent the melting site. FIS protein has a negative effect on the reaction, HU and IHF proteins, a high ATP concentration (>2 mM), high temperature (38°C) and transcriptional activation enhance unwinding.

The *oriC* region also contains no less than eleven 5′-GATC-3′ boxes, which are target sequences for Dam methylase. Their methylation is involved in replication initiation.

For initiation, 20 … 30 molecules of DnaA bind to the DnaA boxes with simultaneous ATP hydrolysis (pre-initiation complex). The DNA is bent to a loop around the DnaA core. This is supported by binding of IHF protein (one binding site), and/or FIS protein (four binding sites) and by unspecific binding of HU protein. It is assisted by simultaneous transcription of genes in the *oriC* region. As a result, strand separation (melting) of the double helix occurs at the adjacent melting site, forming an open complex.

Hexameric helicase (DnaB protein) is complexed by accessory DnaC protein (helicase loader) with simultaneous ATP hydrolysis and is loaded onto the single-stranded region of the DNA (prepriming complex). The rest of this region is covered by tetrameric single-stranded binding protein, SSB. Then, the primase enzyme (DnaG) is loaded by DnaB to form the primosome (at some origins, this is assisted by Pri proteins) and synthesizes two leading strand primers.

The DNA directed DNA Polymerase III (Pol III) holoenzyme also loaded by DnaB is assembled from 10 subunits in a defined order (Fig. 3.8.1-4). Assisted by the γ subunit complex, the dimeric β processivity factor (β subunit) forms a sliding clamp around the DNA. Pol III* is finally loaded onto this ring, which clamps to the template strand. The addition of DNA gyrase completes the formation of the replisome.

3.8.1.3 Elongation and Termination (Fig. 3.8.1-6, Table 3.8.1-2)

Gyrase (ATP-dependent topoisomerase type II, breaking and religating both strands) travels along the strands ahead of the helicase and introduces negative supercoils into the DNA (underwound state of the

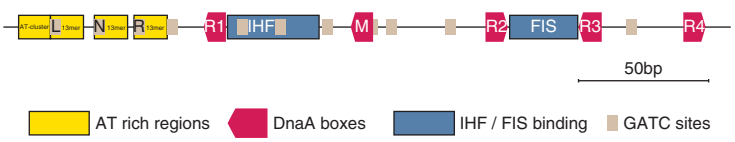


Figure 3.8.1-3. Structure of *oriC* in *Escherichia coli*

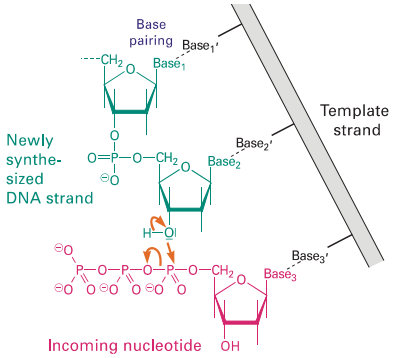


Figure 3.8.1-2. Reaction Mechanism of the DNA Replication

Table 3.8.1-2. Proteins Involved in Bacterial DNA Replication (*Escherichia coli*)

Type	Protein	No. subunits	Mol. mass (kDa)	Function
Initiator protein	DnaA	1	52	binds cooperatively to *oriC*; promotes double helix opening and DnaB loading
Integration host factor	IHF (heterodimeric)	2 (α, β)	11, 11	binds sequence-specifically to double stranded DNA, effects DNA bending
	FIS (homodimeric)	2	11 (* 2)	binds sequence-specifically to double stranded DNA, effects DNA bending
Histone-like protein	HU	1	19	binds sequence-unspecifically to double stranded DNA, effects DNA bending
Helicase	DnaB	6	50 (* 6)	unwinds parental strands; activates primase on single stranded DNA
Accessory protein	DnaC	6	29 (* 6)	complexes DnaB; delivers DnaB to DNA
Primase	DnaG	1	60	DNA-dependent RNA-polymerase; synthesizes RNA primers
Single strand binding protein	SSB	4	19 (* 4)	binds sequence-unspecifically, cooperatively to single stranded DNA, stimulates DNA-Pol III; facilitates DnaB loading
RNA-Polymerase	$\alpha_2\beta\beta + \sigma$	4 (5)	40(* 2), 155,160 + 70 (σ)	transcriptional activation of initiation; double helix destabilization
Type II topoisomerase	Gyrase	2	97 (*2), 90 (* 2)	ATP dependent topoisomerase type II, introduces negative supercoils into parental double helix ahead of the replication fork, removes positive supercoils. Inhibited by nalidixic acid and novobiocin.
DNA Polymerase III holoenzyme (core and τ act as dimers in replication)[1]	subunit α } core	1	130	catalyzes the main reaction of DNA replication: long strand processivity.
	subunit ε } core	1	27.5	3′→5′-exonuclease activity provides proofreading of newly synthesized DNA.
	subunit θ } core	1	8.6	unknown, binds ε
	subunit τ	1	71	core dimerization. Gene *dnaX* encodes also subunit γ by translational frameshift
	γ-complex: γ (dimeric), δ, δ′, χ, ψ	4	47.5(* 2), 39, 37, 17, 15	loads β-subunit to a primed template (‘matchmaker’, ATP-dependent), recognizes RNA primers at the lagging strand.
	subunit β	2	40.6 (* 2)	processivity factor, ‘sliding ring’, clamps Pol III to primed template, dissociates easily from rest of Pol III holoenzyme (→Pol III*).
DNA Polymerase I[2]		1	103	Trifunctional enzyme: removal of RNA primers from Okazaki fragments by its 5′→3′ exonuclease function, consecutive gap filling by polymerase action and proofreading by its 3′→5′ exonuclease activity. The enzyme also plays a major role in DNA repair (3.8.2). Protease splitting yields the ‘Klenow fragment’ (67 kDa, amino acid residues 324 … 928), which contains the polymerase and the 3′→5′ exonuclease function. It is widely used in research.
Ribonuclease	RNase H			removes RNA primer from Okazaki fragments
DNA Ligase			74	NAD dependent joining of Okazaki fragments
Terminator protein	Tus		36	inhibits helicase, replication fork arrest
Type II topo- isomerase	topoisomerase IV	2	81, 67	decatenation/catenation, relaxation of supercoiled DNA
SMC Protein (structural maintenance of chromosomes)	MukB (homodimeric)	2	177 (* 2)	chromosome partitioning

[1] Terminology of DNA polymerase III: (subunits α + ε + θ) = core; [$(core)_2$ + (subunit $\tau)_2$] = Pol III′; (Pol III′ + γ-complex) = Pol III*; (Pol III* + subunit β) = Pol III holoenzyme (HE).
[2] In *E. coli*, also DNA polymerase II has been found. This enzyme is likely to be involved in SOS DNA repair (3.8.2.6).

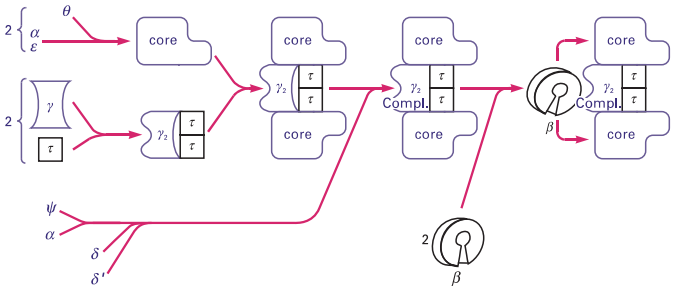


Figure 3.8.1-4. Assembly of the DNA Polymerase Holoenzyme

double helix). This assists the helicase (DnaB) in melting the parental helix by separating the strands. The energy is derived from simultaneous ATP hydrolysis. Reannealing of the strands is prevented by the SSB protein cover. Then DNA primase (DnaG), activated by helicase, synthesizes RNA primers (length 10 … 12 nucleotides) as starting sequences for the action of Pol III. The γ-complex of Pol III* (see footnote to Table 3.8.1-2) recognizes the RNA primers and loads a sliding ring of dimeric β subunits onto the primer-template junction (ATP dependent). Thereafter Pol III* moves to this β ring and starts DNA replication (Pol III can, however, carry out the primase reaction itself). Wrongly inserted nucleotides are removed by the 3′→5′ exonuclease function of Pol III.

Pol III (like all DNA polymerases) polymerizes dNTPs only in 5′→3′-direction, therefore only one template strand can be copied continuously this way (leading strand). The other template strand has a reverse orientation and can be replicated only discontinuously in 5′→3′-direction by ligating 1.5 … 2 kb long Okazaki fragments (lagging strand, Fig. 3.8.1-5). Thus multiple priming reactions are required.

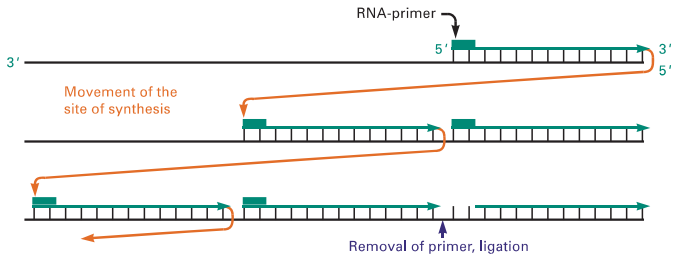


Figure 3.8.1-5. Synthesis of Okazaki Fragments (Lagging Strand, schematically)

A β-ring becomes attached to each new primer as described above. Upon completion of the previous Okazaki fragment, Pol III* moves to this β ring and synthesizes the next fragment until it reaches the primer of the previous one. Then it dissociates from DNA, leaving the β-ring behind and starts a new synthesis cycle. (The β-ring is later removed by Pol III, too).

Due to this mechanism, *E. coli* is able to perform synthesis of both strands with only 10 … 20 molecules Pol III* per cell. The coordination of the discontinuous lagging strand synthesis with the continuous leading strand synthesis is effected by the dimeric structure of Pol III (one enzyme for each strand) and by primase activation due to its association with helicase.

The RNA primer is then degraded by DNA directed DNA polymerase I (Pol I) with 5′-3′-exonuclease-function and/or ribonuclease H (RNase H). In both cases, the degradation proceeds in the 5′→3′ direction. Pol I then fills the gap with deoxynucleotides. Finally, the DNA chains are joined by DNA ligase.

Termination: Replication terminates in a (non-essential) 0.5 kb termination region opposite to *oriC* at the chromosome. At 10 termination (Ter) sites, 10 Tus proteins are bound. Since the Ter sites are non-palindromic (they do not read identically in both directions), asymmetric complexes are formed, which block only one of the approaching replication forks by inhibiting strand separation by helicase (DnaB). This enables termination of replication at a definite site in spite of possible speed differences by both forks.

The parental strands are then unlinked by topoisomerases (decatenation, 3.9.1.4) and the daughter chromosomes are separated by binding of MukB and Par proteins (partition).

3.8.1.4 Fidelity of Replication

The overall error rate is only about 1 per 10^9 … 10^{10} nucleotides. This high replication accuracy is due to

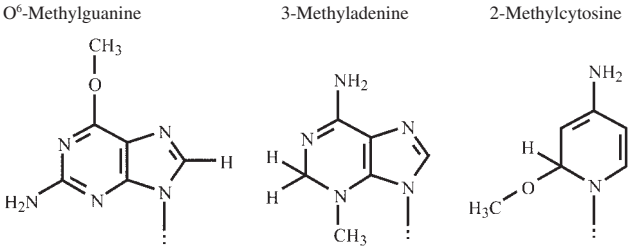
- Twofold dNTP base selection mechanism by DNA Pol III at the binding and the catalytic step (error rate about 1 per 10^4 nucleotides)
- 3′→5′-exonuclease activities of Pol III and Pol I removing misincorporated nucleotides (200 … 1000 fold fidelity improvement)
- postreplicative mismatch repair (see 3.8.2).

3.8.2 Bacterial DNA Repair

3.8.2.1 DNA Damage

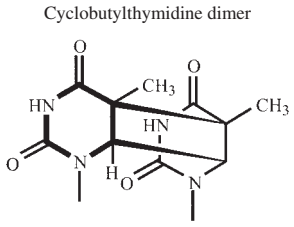
The huge DNA molecule can suffer damage in many ways by exogeneous or endogeneous agents. In many cases the bases of DNA are affected. This may cause mutations or interfere with replication and transcription. The most important damages are:

- Alkylation: Alkylating compounds can modify nucleotides nonenzymatically. They can be of endogeneous (e.g., S-adenosylmethionine) or exogeneous origin (e.g., N-methyl-N-nitrosourea). Products are, e.g., O^6-methylguanine (pairs with thymine!), 3-methyladenine or 2-methylcytosine.

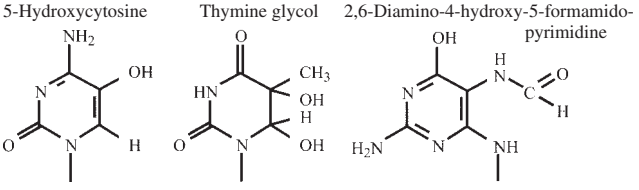


The physiological enzymatic methylation of DNA strands, however, yields N^6-methyladenine, e.g., within the sequence 5′ GATC 3′ (in bacteria, 3.8.1) or 5-methylcytosine (in eukarya, 3.9.1), which are not subject to repair mechanisms.

- Pyrimidine dimerization: UV irradiation (λ = 200 … 300 nm) promotes covalent linking of pyrimidines (mostly thymine) via a cyclobutane ring. This distorts the DNA structure.



- Spontaneous reactions: The most frequent spontaneous damage to DNA is hydrolytic loss of purine bases, leaving apurinic sites. Another hydrolytic reaction is the deamination of cytosine to uracil (U would be read as T in the next semiconservative replication!). Deaminations of adenine to hypoxanthine and guanine to xanthine occur more rarely. Also demethylation takes place (thymine to uracil).
- Oxidative damage: Radicals, especially of oxygen (caused e.g., by ionizing radiation) can induce unusual modifications of bases:



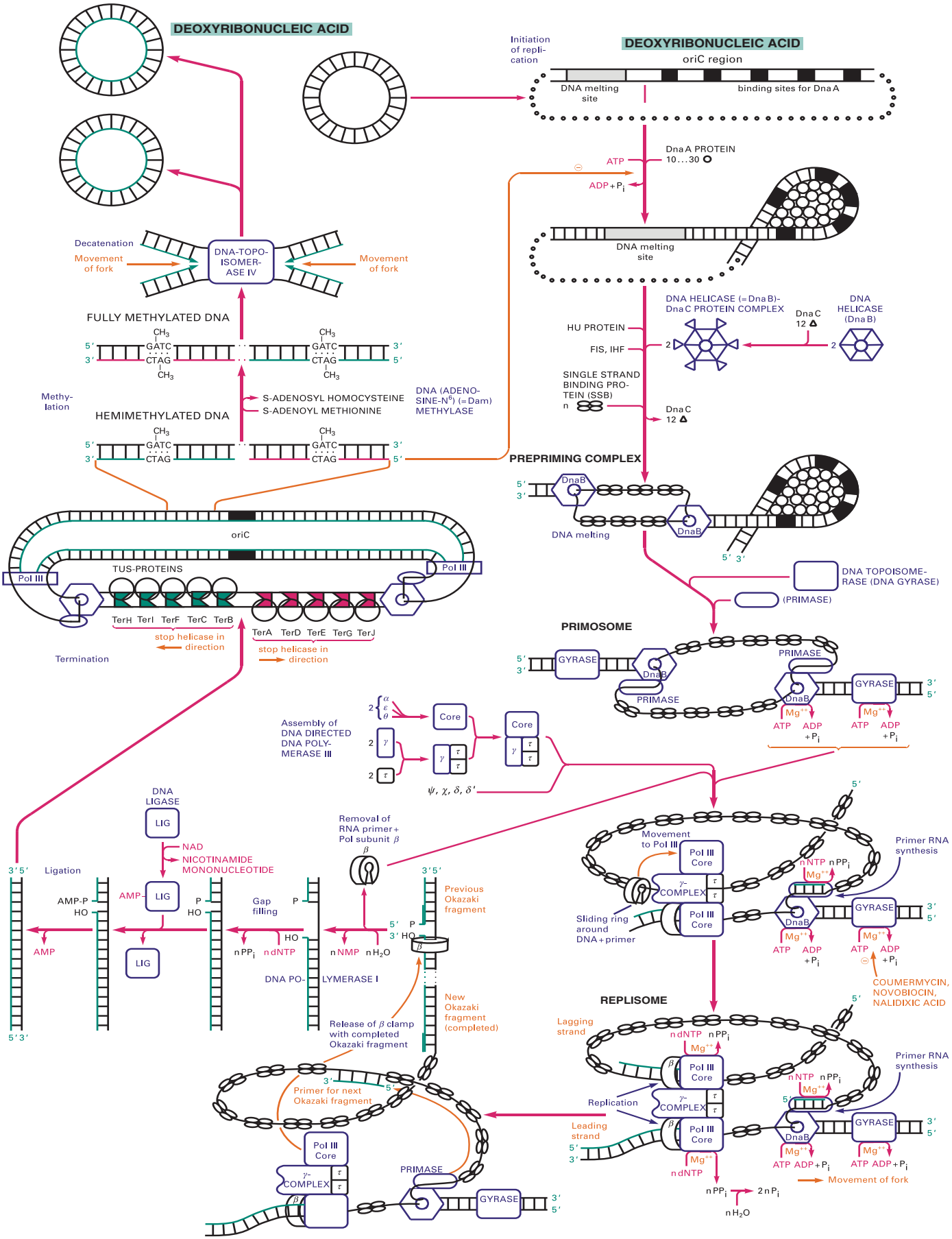


Figure 3.8.1-6. Bacterial DNA Replication
The newly synthesized DNA is drawn in green

- Bulky adducts: Compounds which form adducts with bases (e.g., *cis*-platin•1,2-GpG, benzo[a]pyrene•guanine or psoralene•thymine) cause major distortions of the DNA helix.
- Double strand breaks: Free radicals induced by ionizing radiation or other agents can break both strands of a DNA molecule. This is highly hazardous to the cell, since without repair no replication of the DNA molecule can take place.
- Replication errors: In spite of proofreading by DNA polymerase (3.8.1.4), occasional errors take place during DNA replication and have to be eliminated by repair systems. Otherwise the fidelity would drop 100…1000 fold.
- In heavily damaged cells, the SOS-response (3.8.2.6) tolerates a high level of replication errors in order to enable replication at all.

DNA repair systems: DNA repair fulfills the vital task of maintaining the integrity of DNA structure and sequence. In bacteria, several systems for DNA repair exist, which can be redundant with respect to the lesion as well as to the implicated proteins. Generally, repair proceeds by one of the following strategies:

- direct reversal of damage
- excision of damage and resynthesis according to the information of the complementary strand (including repair of base pair mismatches)
- repair by transfer of sequences from a homologous strand (recombination)

The repair systems make use of the redundant information in the DNA duplex.

3.8.2.2 Direct Reversal of Damage (Fig. 3.8.2-1)

Reversal of alkylation: Methylation of nucleotides is repaired by DNA repair methyltransferases (e.g., O^6-methylguanine-DNA-protein-cysteine S-methyltransferase). These proteins transfer the methyl group from DNA to a cysteine residue in their active site. Since this methyl group cannot be removed from the protein, the reaction leads to inactivation of the enzyme ('suicidal mechanism'). The inactivated enzyme, however, enhances transcription of new enzyme.

Photoreactivation: Cyclobutyl pyrimidine dimers are repaired in a photoreactivation reaction by deoxyribodipyrimidine photo-lyase, which catalyzes the cleavage of the cyclobutane ring in a light dependent reaction.

The enzyme contains two chromophores [FAD and in different species either 5,10-methenyltetrahydrofolate, λ_{max} ca. 380 nm (*E. coli*, yeast) or 8-hydroxy-5-deazariboflavin, λ_{max} ca. 440 nm] which absorb light and initiate dimer splitting by electron transfer to the pyrimidine dimer.

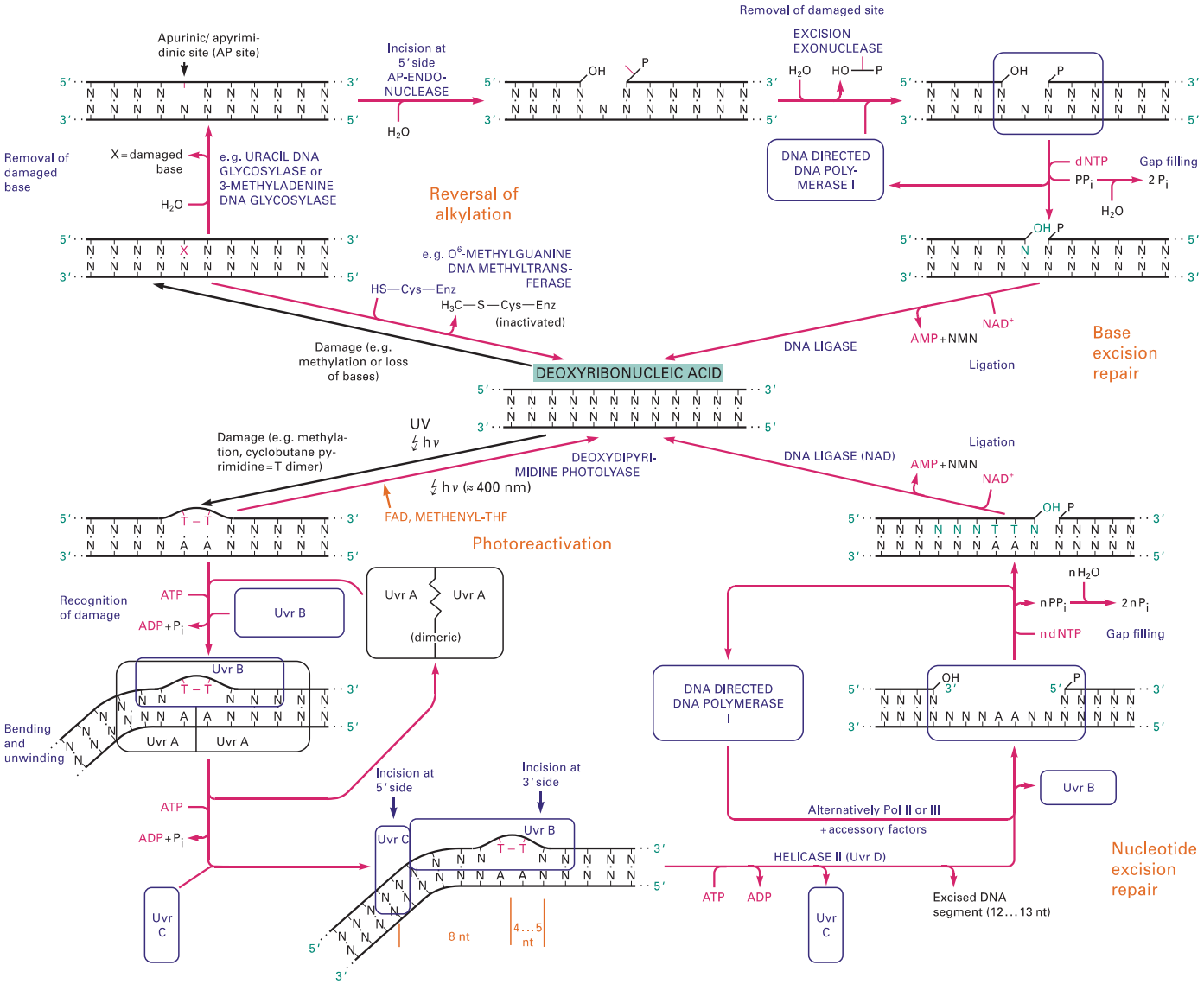


Figure 3.8.2-1. Direct Reversal and Excision DNA Repair Systems in Bacteria (red = damage, green = repair)

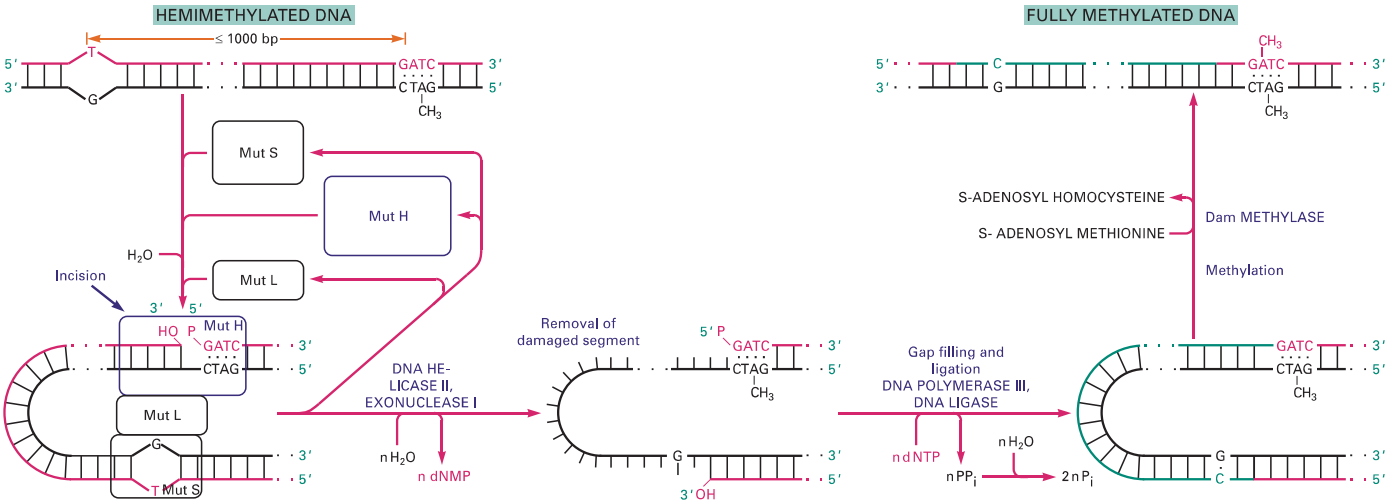


Figure 3.8.2-2. Long patch mismatch repair (Methyl Directed Pathway, *E. coli*)
(red = newly synthesized strand with mismatch, green = repaired segment)

3.8.2.3 Excision Repair Systems

Base excision repair (Fig. 3.8.2-1): Damage to single bases by oxidation, deamination, methylation or demethylation (dT→dU) or base-base mispairs are recognized by specific DNA glycosidases, which cleave the covalent sugar-base bond, leaving apurinic or apyrimidinic sites (AP sites). Examples are DNA- (uracil-, 3-methyladenine-, hypoxanthine- and formamidopyrimidine = 8-hydroxyguanine-) glycosidases. AP sites may also have risen from spontaneous base losses. AP endonuclease cleaves the DNA backbone 5′ of the AP site. Excision exonucleases (e.g., DNA deoxyribophosphodiesterase in *E. coli*) remove the damaged site. The gap is filled in and sealed by DNA polymerase (e.g., Pol I in *E. coli*) and DNA ligase. The removal of the damaged site (and some additional nucleotides) may also be performed by the 5′→3′ exonuclease activity of DNA polymerase I.

Alternatively to 'pure' DNA glycosidases, those with an associated DNA lyase activity carry out additional reactions. Since at the AP site the hemiacetal deoxyribose ring is at equilibrium with an open-chain aldehyde form, the latter can enter a β-elimination reaction leading to cleavage of the DNA chain 3′ of the AP site. The biological importance of this pathway is uncertain.

Nucleotide excision repair (NER, Fig. 3.8.2-1): Bulky adducts to DNA, as well as damages which cause minor distortions (e.g., pyrimidine dimers) can be removed by the ABC system, which responds to these distortions.

The *E. coli* enzyme activity 'ABC excinuclease' is carried out by the three proteins UvrA, UvrB and UvrC. Dimeric UvrA loads UvrB onto a damaged site. After that UvrC binds to UvrB, resulting in the UvrBC-DNA incision complex. In this complex, two incisions are made by UvrC: the first at the fourth or fifth phosphodiester bond on the 3′ side of the damage and the second at the eighth phosphodiester bond on the 5′ side. UvrD (helicase II) removes the damaged strand. The resulting gap is filled by Pol I and the nick is closed by DNA ligase.

The transcription-repair coupling factor (TRCF) connects NER with transcription. TRCF releases stalled RNA polymerase and the truncated RNA transcript from the damage site and recruits UvrA for consecutive repair as outlined before.

3.8.2.4 Mismatch repair

Very short patch repair (*E. coli*): This system corrects G•T mismatches caused by deamination of 5-methylcytosine to thymine in C(meC)(A/T)GG sequences of fully methylated DNA. The incision 5′ of the mispaired T is catalyzed by Vsr protein, an endonuclease. Thereafter, DNA polymerase I removes the mispaired T and less than 10 more nucleotides by its 5′→3′ exonuclease activity and fills the gap. The nick is then ligated. Mut S and Mut L proteins (see below) are also needed, possibly for damage recognition.

Long patch repair (Fig. 3.8.2-2): Small insertion or deletion mispairs as well as base-base-mismatches are repaired by the methyl-directed mismatch repair (MMR) of *E. coli* (a related system is the Hex-dependent pathway of *Streptococcus pneumoniae*). This system takes advantage of the fact that for a short time after replication the newly synthesized strand is not yet methylated by Dam-methylase,

Table 3.8.2-1. DNA Repair in Bacteria

System	Damage Specificity	Implicated Repair Proteins (*E. coli*)	Remarks
Reversal of alkylation	O^6-methylguanine, O^4-methylthymidine	O^6-methylguanine DNA repair methyltransferase I (Ada, repairs also O^6-methylthymidine and methylphosphotriesters), O^6-methylguanine DNA repair methyltransferase II (Ogt)	"suicidal" mechanism (only one methyltransfer catalyzed by one protein molecule)
Photoreactivation	pyrimidine dimers	deoxyribodipyridine photo-lyase, carrying as chromophores methenyltetrahydrofolate + FAD	light dependent (λ = 300 … 500 nm)
Base excision repair	uracil, thymine glycols, hypoxanthine, 8-oxoguanine, 3-methyladenine	DNA glycosidases, apurinic/apyridinic endonuclease, excision exonuclease, DNA polymerase I, DNA ligase	length of repair tract ≤ 10 nucleotides
Nucleotide excision repair (Uvr)	bulky adducts, e.g., cisplatin•GpG, pyrimidine dimers	UvrA, UvrB, UvrC, UvrD (Helicase II) proteins, DNA polymerase I, DNA ligase	length of repair tract 12–13 nucleotides
Very short patch mismatch repair	G•T mismatch (corrected to G•C) in 5′-CT(A/T)GG	Vsr, MutS, MutL proteins, DNA polymerase I	length of repair tract ≤ 10 nucleotides
Long patch mismatch repair (methyl-directed)	base-base-mismatch, small insertion/deletion mispairs	MutS, MutL, MutH proteins, DNA helicase II, single strand binding protein, exonuclease I (or exonuclease VII/exonuclease RecJ), DNA polymerase III holoenzyme, DNA ligase	length of repair tract ≈1000 nucleotides
Recombination repair	double strand breaks, distorting DNA sites	RecBCD, RecA, RuvC proteins	error-prone
SOS response	DNA polymerase stalled at lesion, DNA single strands remain	RecA, LexA, UmuC, UmuD proteins, DNA polymerase III	error-prone repair as part of the SOS regulon, lesion persists

while the parental strand already carries methyl groups at adenine-N^6 in GATC (= Dam) sequences. This differentiates between both DNA strands and insures that sequence information for repair is retrieved from the unmutated, parental strand.

According to the current model, the repair is initiated by the binding of MutS (95 kDa, oligomeric) to the lesion. After the binding of MutL to MutS and MutH to a nearby d(GATC) sequence (up to 1000 bp away in either direction), an incision is made by MutH (25 kDa) on an unmethylated DNA strand. Assisted by MutL (95 kDa), the proteins form a complex, which incises the non-methylated DNA strand 5′ from the GATC site. Subsequently, exonucleases degrade the faulty strand from here to a point beyond the site of the lesion. If the incision is located at the 3′ side of the mismatch, repair requires the 3′ to 5′ exonucleolytic activity of exonuclease I (ExoI); if the incision is located on the 5′ side, repair requires the 5′ to 3′ exonucleolytic activity of either exonuclease VII (ExoVII) or RecJ. The final steps are performed by SSB, DNA polymerase III and DNA ligase (filling the single-strand gap left by the excision).

3.8.2.5 Double-Strand Repair and Recombination

A very difficult situation arises if double strand breaks occur. Similarly, if during replication a damaged, distorting template sequence is reached, the replication by DNA polymerase III is interrupted and reinitiated at some point beyond this site, leaving a gap in the daughter strand. In both cases the strand defects cannot be cured by excision repair, since no correct template sequence is available in the duplex DNA. However, by using the correct sequence from a homologous DNA to fill the gap, an intact DNA sequence is obtained, which can be used as a template for damage correction in the other strand (recombination or postreplication repair, Fig. 3.8.2-3). This is possible because fast growing bacterial cells usually start another round of DNA replication before the previous one is completed and cell division has taken place (3.8.1), so that two or more copies of the same gene are available in one cell.

In case of double strand breaks, both free 5′ ends are degraded by exonuclease V (RecBCD protein) until a ‘chi-site’ (sequence 5′GCTGGTGG3′, spaced at about 10 kb in E. coli) is reached, where RecBCD switches to a helicase activity. One single-stranded DNA 3′ end is helped by the nuclease function of RecA protein (38 kDa) to invade a homologous sequence (Fig. 3.8.2-3). Expansion of the invaded stretch and gap filling lead to a double Holliday junction (crossover). Its subsequent resolution by RuvC protein brings about exchanges of the two homologous strands. Several variations of this reaction exist.

Similar mechanisms (Fig. 3.8.2-4) are used to fill the gap in the daughter strand opposite a defective site during replication. The defect can be repaired afterwards by the mechanisms described above. However, since mismatches are tolerated in the exchange reaction, recombinational repair is more error-prone than the other repair systems.

3.8.2.6 SOS Response (Damage Tolerance Mechanism, Fig. 3.8.2-5)

Agents which cause intense damage to DNA induce a complex emergency repair system in *E. coli* and similarly in many other bacteria. The cells stop dividing and start a special DNA repair mechanism, sacrificing a high level of fidelity.

If the proceeding replication fork meets damaged DNA sequences, usually in- or post-replication repair mechanisms (as described before) are initiated. In case of frequent damages, however, DNA polymerase III action stops. Thus, single stranded DNA sequences result. RecA protein (see 3.8.2-5) binds to them and is converted into an activated form, which aids in autoproteolysis of the LexA repressor protein. Under physiological conditions, this repressor is bound to SOS boxes (consensus sequence $CTGN_{10}CAG$) at more than 20 genes and represses their expression considerably. Removal of this block leads to increased expression of (among others) UvrA, UvrB and UvrC (see above at NER), LexA and RecA proteins as well as of helicase II, DNA polymerase III, UmuC and UmuD proteins. The latter two enable DNA polymerase III to resume replication and to proceed through sites of DNA damage despite missing or faulty information from the complementary strand in a way not known in detail yet. Although this results in a highly error-prone DNA replication, it enables some form of replication.

Literature:

Friedberg, E.C., Walker, G.C., Siede, W. *DNA Repair and Mutagenesis*. ASM Press, 1995.

Harris, R. S., Ross, K. J., Lombardo, M. J., Rosenberg, S. M. J. Bacteriol. 1998;180:989–993.

Messer, W. FEMS Microbiol Rev. 2002;26:355–74.

Modrich, P., Lahue, R. Annu. Rev. of Biochem. 1996;65:101–133.

Sancar, A. Annu. Rev. of Biochem. 1996;65:43–81.

(various authors) Trends in Biochem. Sci. 1995;20:381–439.

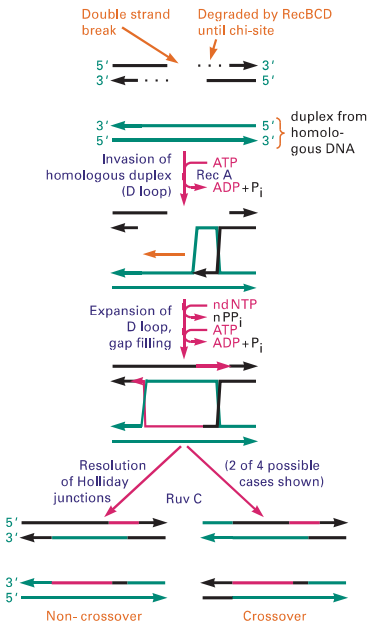


Figure 3.8.2-3. Recombination Repair of Double Strand Breaks

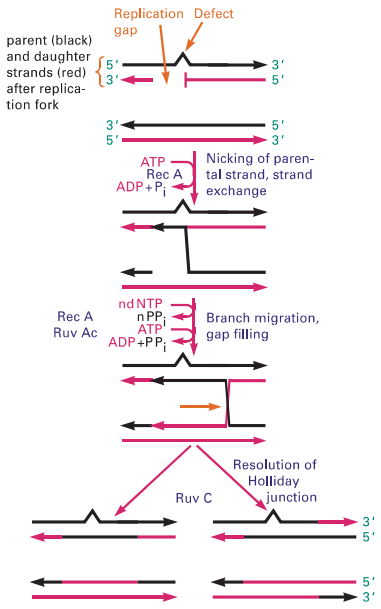


Figure 3.8.2-4. Recombination Repair of Gaps in Daughter Strands

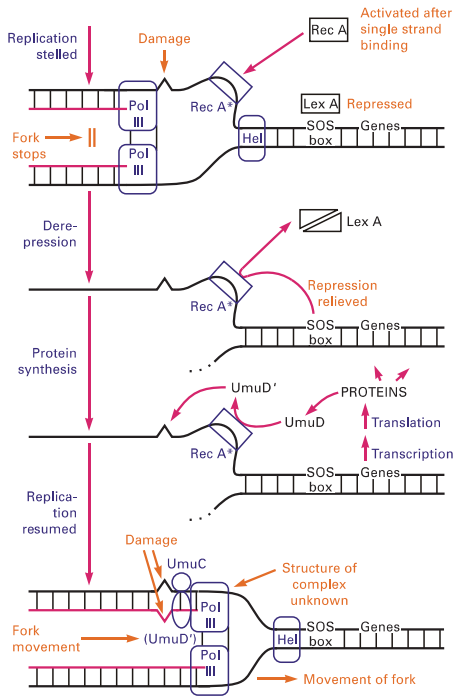


Figure 3.8.2-5. SOS Repair System

3.8.3 Degradation of Nucleic Acids

Nucleases catalyze the cleavage of phosphodiester bonds in DNA (DNases) or in RNA (RNases). DNases play a decisive role during DNA synthesis, repair and recombination (3.8.1, 3.8.2, 3.9.2), for relief of obstructive superhelical tension as well as in apoptosis (4.3.6, 7.6) and as a defense measure in restriction of foreign DNA (4.1.4.2). RNases are, e.g., of importance in RNA processing (4.1.1.1, 4.2.1.3) and also regulate the transcription by degradation of mRNA (4.2). Nucleases either remove terminal nucleotides (exonucleases) or act inside of the nucleic acid molecule (endonucleases).

3.8.3.1 Exodeoxyribonucleases (Exo-DNases, Table 3.8.3-1)

Exo-DNases are characterized by their cleavage direction, their preference for a single stranded or a double stranded substrate and by producing mono- or (more rarely) oligonucleotides. In a number of cases, they perform several functions (see below). The nucleases may dissociate after each catalytic event (non-processive or distributive action) or they may remain bound to the polymer until many reaction cycles are completed (processive or non-distributive action).

The reaction scheme of *E. coli* exonuclease III is schematically shown in Figure 3.8.3-1. The enzyme is multifunctional: it acts also as an endonuclease specific for apurinic DNA sites (3.8.2.3) and as a 3′ phosphatase. Bacterial DNA polymerase I exerts a 3′ → 5′ exo-DNase activity for proofreading (similarly to eukaryotic DNA Pol δ and ε) and additionally a 5′ → 3′ exo-DNase/RNase function for DNA repair and for removal of Okazaki RNA primers (3.8.2.3; 3.8.1.3).

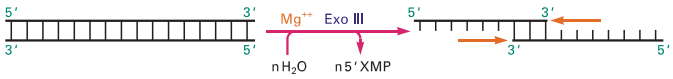


Figure 3.8.3-1. Reaction Scheme of *E. coli* Exonuclease III

3.8.3.2 Endodeoxyribonucleases (Endo-DNases, Table 3.8.3-2)

The endonucleases often show a strong preference for either single stranded or duplex DNA. Endonucleases that function in repair of lesions identify the damaged DNA site and incise (nick) the DNA at one side of the lesion as a first step towards excision. A second principal characteristic of these enzymes is the recognition of base sequences. A striking example is the cytosine-specific cleavage by T4 endonuclease IV. Pancreatic DNase, *E. coli* endonuclease I or spleen DNase produce oligonucleotide digests with characteristic sequence patterns at the 3′ and 5′ termini.

Restriction endonucleases occur in a variety of microorganisms. More than 5000 of them have been identified so far (http://rebase.neb.com/). They recognize sequences of 4 … 8 nucleotides in a DNA duplex with extraordinary accuracy and cleave both strands.

The organisms always produce a companion DNA methyltransferase, which recognizes the same sequence in endogeneous DNA and modifies it immediately after replication by methylation of A or C residues. This protects the organism's own DNA from degradation by the restriction enzyme. Thus, the restriction endonuclease and the cognate methyltransferase form a restriction-modification (R-M) system. Since it cleaves (restricts) infecting DNAs (e.g., viruses) and thus prevents them from parasitizing the cell, the R-M system is also called the 'immune system of the microbes'.

At least four different kinds of R-M systems exist:

- Type I enzymes, e.g., EcoK1, carry methylase and nuclease activity on the same protein and require Mg^{++}, ATP and S-adenosylmethionine for cleavage. They cleave randomly and remote (> 400 bp) from the recognition sequence.
- Type II enzymes, e.g., EcoR1, recognize mostly palindromic nucleic acid sequences and cleave within or near these sequences. (Palindromic sequences repeat each other at the other duplex strand in inversed order, resulting in twofold rotational symmetry. For examples see Fig. 3.8.3-2.) The enzymes require Mg^{++} for activity. Their homodimeric structure corresponds to the palindromic substrate. Many of them generate 'sticky end' ('cohesive end') duplex fragments with 5′ protruding termini (e.g., EcoRI) or with 3′ protruding tails (e.g., HhaI). Other enzymes cleave at the center of the recognition sequence and produce 'blunt end' fragments (e.g., HaeIII). The cognate methylase is a separate enzyme. The cleavage specificity is schematically shown in Figure 3.8.3-2. These enzymes are indispensible tools for molecular cloning techniques and for DNA sequence analysis. They are named by the 3-letter abbreviation of the source organism.
- Type IIS enzymes recognize asymmetric sequences of 4 … 7 bp length. They cleave at a defined distance of up to 20 bp to one side of their recognition sequence.
- Type III enzymes, e.g., StylT1, have similar characteristics as Type I enzymes, but cleave at specific sites only a short distance (24 … 26 bp) away from the recognition sequence.

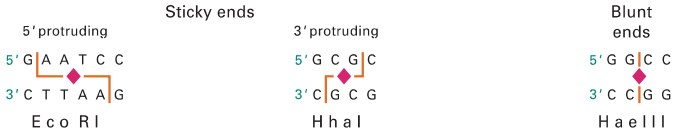


Figure 3.8.3-2. DNA Cleavage by Type II Restriction Endonucleases
(♦ = twofold symmetry axis)

Table 3.8.3-1. Examples for Different Types of Exonucleases

Specificity for	3′ → 5′ direction	Cleavage in 5′ → 3′ direction	either direction
Single stranded DNA	*E. coli* exonuclease I, mammalian DNase III	exo-DNase (phage SP_3 encoded)	*E. coli* exonuclease VII
Double stranded DNA	*E. coli* exonuclease III	exo-DNase (phage λ encoded), mammalian DNase IV	*E. coli* exonuclease V (ATP-dependent)
DNA or RNA	Venom exonuclease	Spleen exonuclease	
RNA	Exo-RNase II, RNaseQ, RNase BN	Yeast RNase	

Table 3.8.3-2. Examples for Different Types of Endonucleases

Specificity for	3′ P endproducts	Cleavage to 5′ P endproducts	other end products
Single stranded DNA	*Aspergillus* DNase K_1	DNase IV (phage-T_4-encoded), *yeast* DNase	Crossover junction endo-RNase (acts only on Holliday junctions)
Double stranded DNA	pancreatic DNase II	Type I, II and III restriction enzymes (10.7.2), pancreatic DNase I	
DNA or RNA	Micrococcal nuclease, spleen endonuclease	*Aspergillus* nuclease S_1, Mung bean nuclease, potato nuclease	
RNA-DNA duplex		Exo-RNase H (acts on RNA : DNA hybrids, degradation of Okazaki fragments)	
RNA	RNase T_2 = RNase II, pancreatic RNase Polynucleotide phosphorylase	RNase III and RNase P (processing of tRNA and rRNA precursors)	*Bacillus subtilis* RNase (yields 2′,3′-cyclic phosphates)

3.8.3.3 Ribonucleases (RNases, Tables 3.8.3-1 and 3.8.3-2)

Similarly to DNases, RNases differ by exo- and by endo-activity, preferences for termini (exo-enzymes) or in some cases for specific sequences (endo-enzymes). Some nucleases even cleave both DNA and RNA. A number of RNase type reactions are not catalyzed by proteins, but rather by RNA sequences (ribozymes). E.g., the RNA component of RNase P catalyzes the processing of untranslated prokaryotic RNA (4.1.1.2, as well as some reactions in eukaryota); the protein component has only assistant function. In a number of cases, eukaryotic group I or II introns are removed from the rRNA by its own action (self-splicing, e.g., in *Tetrahymena*, 4.2). It is speculated that RNA catalysis and self-replication preceded enzyme-protein catalysis during evolution.

Literature

Lin, S.M., Roberts, R.J. (Ed.) *Nucleases.* CSH Monograph Series, 1985.

Kornberg, A., Baker, T.A. *DNA Replication*. 2nd Ed. Freeman, 1992.

Roberts, R.J., Macelis, D. Nucl. Acid Res. 1996;24:223–235.

3.9 Nucleic Acid Metabolism in Eukarya

Helmut Burtscher

3.9.1 Eukaryotic DNA Replication

With the exception of the second meiotic division of germ cells, DNA replication is a necessary prerequisite of cellular division. DNA must be replicated with high fidelity to ensure correct information transfer to daughter cells. In eukarya several replicative DNA polymerases, some of them with intrinsic exonuclease activity, are involved in this extremely accurate process.

Although in eukarya it is actually chromatin that replicates (the complex of DNA, histones and other proteins, attached to the nuclear scaffold, 2.6.4), the following section concentrates on the replication of DNA, which proceeds in a series of integrated protein-protein and protein-DNA interactions and enzymatic reactions.

3.9.1.1 Cell Cycle and DNA Replication

DNA replication must be precisely linked to the cell cycle (summarized in Fig. 3.9.1-1). It depends on the passage through a special point in the G_1-phase, at which cells commit to another round of replication and cell division (see below). DNA is replicated exclusively in the S-phase. For the next round of replication, a passage through the M phase must normally have taken place.

The mechanisms underlying these controls are still being elucidated. More details of the regulatory aspects are given in Section 4.3.

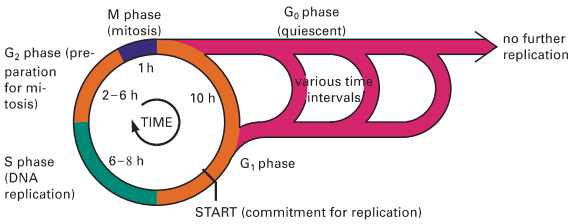


Figure 3.9.1-1. Phases of the Mammalian Cell Cycle

- The chromatin must be in a replication competent state: A pre-replication complex of DNA and 'licensing factors' (3.9.1.2) must have been formed, which dissolves during the later phases of the cell cycle and thus prevents entering an untimely round of replication.
- Also, external signals (e.g., nutrient supply in unicellular organisms or hormonal messages in multicellular organisms) are required to initiate DNA replication and cellular division. In yeast, coordination of cell growth and DNA replication occurs at a unique point in the G_1 phase of the cell cycle, which is called START (Fig. 3.9.1-2). This control step irreversibly commits the cell to another round of cell division, as most external cell cycle regulators have no effect on DNA replication beyond this point. Here, binding of cyclins (cell cycle regulating proteins, 4.3) to the cyclin dependent kinase Cdc 28 causes activation of transcription factors leading to expression of genes necessary for DNA replication. Similarly, cell cycles in higher eukarya depend on passage through the restriction or cell cycle commitment point (R) in the G_1 phase (4.3.3), where most regulatory processes leading to cell proliferation take place. After the R point, cells no longer need growth factors for completion of the cell cycle.

In most organisms active genes are replicated early in the S-phase, while inactive genes, telomeres, telomere-linked regions and heterochromatin are replicated later on. The influence of various factors like histones, nucleosomes, the nuclear matrix and matrix associated enzymes (e.g., DNA topoisomerase II, histone acetyltransferases, histone deacetylases, poly-ADP-ribosyltransferase etc.) as well as the unfolding mechanism of the densely packed chromatin are very complex and will not be discussed here.

3.9.1.2 Initiation of Replication (Fig. 3.9.1-2, Tables 3.9.1-1 and 3.9.1-2)

Initiation of replication takes place at origins or initiation zones. These DNA regions are recognized by specific proteins or protein complexes which catalyze localized unwinding and load replication proteins onto the exposed single strands. The order of assembly of replication factors is highly conserved across all species.

Unlike bacterial chromosomes, eukaryotic chromosomes contain multiple origins of DNA replication. On a typical mammalian chromosome these are spaced 50 … 100 kb apart. Not all replication origins are activated simultaneously. Some origins are used early, some late in S-phase and not all seem to be utilized in every cell or during every cell cycle. Early replicating origins seem to be associated with transcriptionally active regions whereas late replicating origins are located near repressed regions. However, it looks as if it is not chromatin accessibility or pre-replication complex formation but a later step that determines origin firing.

Re-initiation from the same origin during a single cell cycle must be blocked by all means. This is ensured by redundant mechanisms, many if not all of which are regulated by cyclin-dependent kinases.

Table 3.9.1-1. Replication of Genomes

Species	Size (bp, haploid)	Number of Chromosomes	Number of Replication Origins (total)	Duration of Genome Replication
Yeast	$1.4 * 10^7$	16	250…400	30 min
Human	$3 * 10^9$	2 * 22 + 2	20,000…50,000	8 h

In yeast (*S. cerevisiae*), replication origins are called autonomously replicating sequences (ARS). A stretch of 100 to 150 nucleotides is required for maximal function, arranged in domains A and B. Domain A contains an essential ARS consensus sequence, 5′-(A/T)TTTA(T/C)(A/G)TTT(A/T)-3′. Origins in yeast resemble promoters (3.9.2). Replication enhancer sequences can be located at distances up to 1 kb on either side of the start position and in both directions.

A complex of six essential proteins in yeast (origin recognition complex, ORC, homologues in other eukarya) binds in ATP-dependent manner to ARS and remains there throughout the cell cycle. The activating proteins Cdc6, Cdt1 and a complex of Mcm proteins ('licensing factors') become available during mitosis (when the nuclear membrane dissolves) and are part of the pre-initiation complex at the end of M-phase, which is essential for initiation of replication. Cdt1 recruits Cdc45 into the complex. Thereafter, Cdc6 and Cdt1 leave the complex and are either degraded or sequestered for most of the cell cycle (The Mcm complex moves along with the replication fork). Complex formation and dissolving is controlled by the cyclin/cyclin dependent kinase complexes Clb1, 2, 3, 4-Cdc28 and Clb5, 6-Cdc28, respectively (4.3.2).

Identity and structure of replication origins in higher eukarya are probably similar but more complex. Initiation sites can be distributed over broad regions (initiation zones) of several kb length possibly due to more complex regulation requirements. The chromosomal structure seems to be a major influence. Today, methods are available for isolation of active origins, e.g., from CHO and other cells.

3.9.1.3 DNA Polymerases (Pol, Table 3.9.1-3)

Eukarya have an amazing number of DNA polymerases, e.g., at least 14 DNA-dependent DNA polymerases are encoded in the human genome. These polymerases are grouped in families based on sequence homologies in their catalytic domains; they can also be grouped according to function: Replicative polymerases which copy the genome with high fidelity and trans-lesion or bypass polymerases which are non-processive and copy DNA with low catalytic efficiency in an error-prone but nevertheless essential manner.

In yeast three DNA polymerases (α, β, ε) are essential for DNA replication. Their sequences are highly conserved (human and yeast DNA polymerase α show 31 %, bovine and yeast DNA polymerase δ 44 % homology).

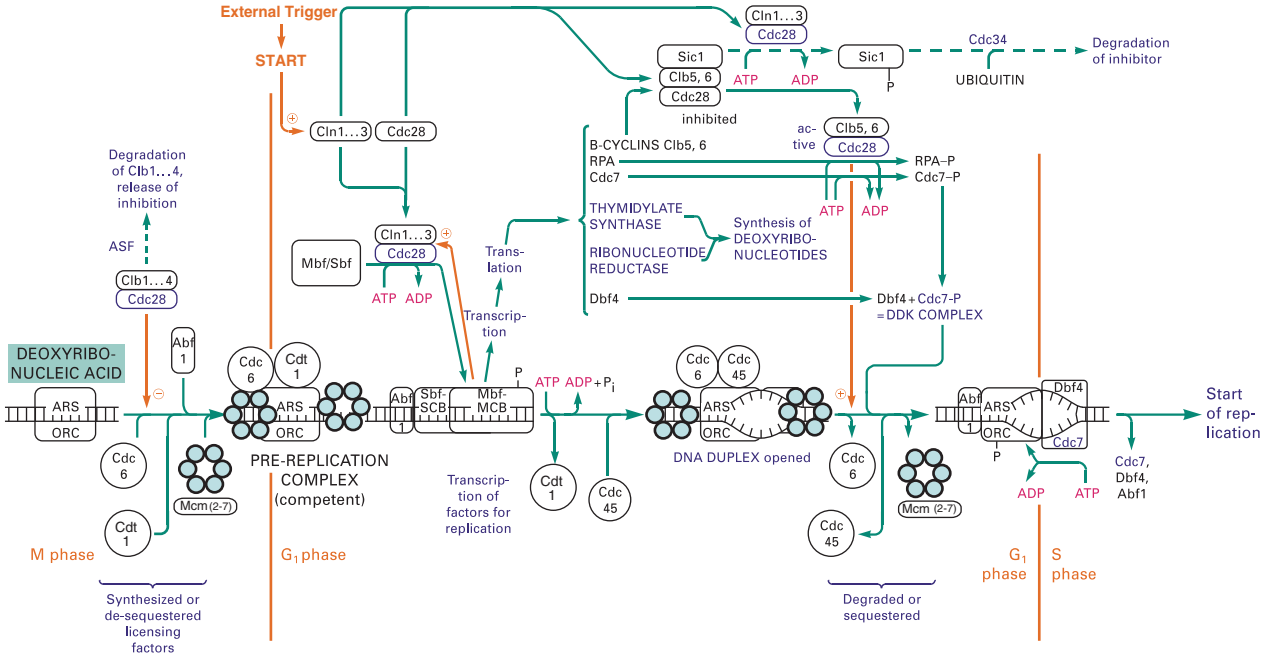


Figure 3.9.1-2. Model of Replication Initiation in Yeast
(Spatial Arrangements of the Factors are unknown)

Table 3.9.1-2. Essentials for Initiation of DNA Replication in *Yeast (Saccharomyces cerevisiae)*

Type	DNA Sequence/Protein Factor	Mol. mass kDa	Function
Origin DNA element	ARS		Autonomously replicating sequence, core (A site) + 3 adjacent B sites
DNA element	Abf binding sequence		3′ or 5′ from core ARS, consensus sequence (A/G)TC(A/G)(T/C)NNNNNACG
DNA elements	MCB/SCB		Cell cycle boxes, located in the promoter of genes transcribed in late G_1 phase, bound by transcription factor Mbf and Sbf, respectively
Proteins:			
Origin recognition complex	ORC (subunits: Orc 1…6)	104, 71, 72, 61, 55, 50	Remains bound to ARS (core + 3 sequence) throughout the cell cycle
Activating proteins (in nucleus only in M + G_1 phases)	Cdc6 Cdc45 Cdt1	58 74 68	Activates ORC/ARS, required for binding of Mcm proteins and recruitment of Cdc45 Activates ORC/ARS, required for recruitment of Mcm-complex Licensing factor, associates with Cdc6 to enable association of mcm proteins with chromatin
Mcm proteins (in nucleus only in G_1 and S phases)	Mcm2 Mcm3	99 108	Activates ORC/ARS, possibly ‘licensing factor’ Binds 3′ to ARS, activates ORC/ARS, opens duplex (helicase?). Analogy in humans: P1 = DNA Pol-primase accessory factor, targets Pol α?
also:	Mcm4 (Cdc54), Mcm5 (Cdc46), Mcm6, Mcm7 (Cdc47)		
Cyclins	G1 type: Cln1, 2, 3 B type: Clb5, 6 Clb1, 2, 3, 4	62, 62, 65 50, 44 55, 56, 49, 54	G_1 phase cyclins, associate at START with kinase Cdc28, activate Mbf/Scf Late G_1 and S phase cyclins, associate with kinase Cdc28, inhibited by Sic1 G_2 and M-phase cyclins
Protein kinases	Cdc28 (cyclin-dependent, present throughout cell cycle)	34	a) Activation at START after association with G_1 cyclins Cln1, 2, 3, budding b) Association with B cyclins Clb5, 6: phosphorylates Cdc7, Swi6, Cdc10, RF-A Generally: causes DNA synthesis, spindle formation
	Cdc7	58	Expression depends on Mbf1 bound to MCB sequence, becomes activated by Dbf 4, phosphorylates Orc or other proteins in late G_1 phase
Activating protein	Dbf4	80	Targets Cdc 7 protein kinase to DNA
Inhibitor	Sic1 (p40)	32	Inhibits Cdc28–Clb5, 6 complex, degraded after phosphorylation by Cdc28/Cln
Ubiquitin-protein ligase	Cdc34	34	Ligates Sic1-P with Ub for degradation, thus activates Cdc 28/Clb5,6-complex
Transcription factors	Abf1	81	ARS binding factor, enhancer, required for origin function
	Mbf (subunits: Mbp1, Swi6) and Sbf	94, 91	Bind after phosphorylation to MCB and SCB DNA sequences, respectively, start transcription of ribonucleotide reductase, thymidylate synthase, Pol α-primase, Clb5, 6, Cdc6, 7, Cdc46, Dbf4, Orc6 etc.
Polymerase / primase and replication factor RF-A		See 4.2	

Table 3.9.1-3. Replicative Eukaryotic DNA Polymerases

Polymerase	Source	Number of subunits	Mol. mass kDa	Function
α (I)	yeast	4	167, 79, 62,48 (primase)	Initiation of both strands. Subunit p68 is periodically phosphorylated in a cell-cycle dependent manner. Tightly associated primase activity. Starts Okazaki fragment synthesis. Synthesizes an RNA primer of
	human	4	180, 68, 58, 48	8 … 10 nt length, extends it with an initiator DNA of less than 50 nt. No exonuclease activity.
β (IV)	yeast	1	68	Involved in base excision repair (3.9.2.3). Essential in mammals (as shown in mice), this enzyme is mutated
	human	1	39	in up to 30% of human cancers; not essential for viability in yeast.
γ (m)	yeast	1	144	Replication of mitochondrial DNA, similar enzyme in chloroplasts, human enzyme consists of a homodimer
	human	3	140, 55 (2)	of accessory subunits and a catalytic subunit. The enzyme is encoded in nuclear DNA.
δ (III)	yeast	3	125, 55, 40	Synthesizes leading strand and completes Okazaki fragment synthesis. No primase activity. Requires PCNA
	human	4	125, 66, 50, 12	(Table 3.9.1-5) for binding to DNA. Long strand processivity. Contains 3′ → 5′ exonuclease for proofreading (by removing mononucleotides). Also involved in DNA-repair and recombination. Essential.
ε (II)	yeast	5	255, 80, 34, 30, 29 258, 59, 17, 12	Similar to Pol δ. Long strand processivity even without presence of PCNA. 3′ → 5′ exonuclease activity, degrades to 6 … 7 nt long oligonucleotides. Participates in DNA replication (leading strand) and repair.
	human	4		Essential.

3.9.1.4 Replication Forks (Fig. 3.9.1-6)

The general aspects of DNA replication are discussed in Chapter 3.8.1. Eukaryotic replication forks move with a speed of about 30 … 50 nt/sec. (in vitro, on 'naked' DNA, about 500 nt/sec. are observed). The large number of replication origins shortens the replication time of mammalian chromosomes from more than 1000 hours (calculated for a single origin) to ca. 8 hours.

Although it is still generally assumed that the replication machinery travels along the DNA, there are also indications for a relatively immobile replication apparatus, where the DNA fork is being threaded through. (Analogous possibilities also exist for transcription complexes, 4.2.1.2).

After the initiation phase (which is presented in Section 4.3.2 for yeast), helicases unwind the DNA double strand and DNA topoisomerase I relieves the positive superhelical tension (Fig. 3.9.1-3, one strand is nicked and released after passing past the other strand). Both reactions are energized by ATP hydrolysis. DNA directed DNA polymerase α/primase (Pol α /primase) attaches itself to the single strands, the still open portion of which is covered by RPA (= RFA). The primase synthesizes a short stretch of RNA complementary to the template DNA, which is extended as initiator DNA (ca. 34 nt) by Pol α. Then a polymerase switch takes place: RF-C displaces Pol α by binding to the 3′ end of the initiator DNA. It then recruits PCNA to form a trimeric ring around the DNA strand, which enables DNA directed DNA polymerase δ (Pol δ) to join and to continue DNA replication. PolΔ is able to proofread the newly synthesized DNA strands.

Analogous to bacteria (3.8.1.3), replication of one template strand can proceed continuously in 5′→3′ direction (leading strand as described above). Since the other template strand has a reverse orientation, its replication can be performed only discontinuously (lagging strand) by ligating short Okazaki fragments (about 100 … 200 nucleotides in eukarya), which are also synthesized in 5′→3′ direction. Leading and lagging strand polymerases presumably act in concert, resulting in the periodic formation of unreplicated looped lagging strand DNA (trombone model). Although more effort seems to be needed for lagging strand processing, both strands are synthesized at identical rates.

The Okazaki fragments are started by RNA primers and initiator DNA. The Pol α /primase performing these reactions is consecutively replaced via a polymerase switch by Pol δ, which continues replication until the previous Okazaki fragment is reached. Removal of the RNA primers requires the combined action of the 5′→3′ exonuclease MF1 and RNase H. The resulting gap is filled by a DNA polymerase and closed by DNA ligase I.

There is still some uncertainty about the detailed mechanism of DNA replication. In particular, the exact spatial arrangement of the DNA replication system is still unclear. Several models have been proposed which differ in various aspects. Therefore, Figure 3.9.1-6 will probably not be exact in all details.

The completed DNA strands wind around histones and are further compacted, using scaffolding proteins (2.6.4). The relaxing of negative supercoil tensions, which occur during these reactions, is achieved by DNA topoisomerase II. Another function of this enzyme is the removal of 'knots' of the DNA strands (decatenation). The enzyme breaks both DNA strands and guides them through another DNA duplex.

Eukaryotic topoisomerase type IB (Fig. 3.9.1-3) acts preferentially on double stranded DNA. A tyrosine hydroxyl group of the enzyme temporarily forms a covalent bond with the 3′ phosphate group of one of the DNA strands. After the free 5′ end has turned in order to relax the twisting stress, release of the nick takes place. No supply of additional

Table 3.9.1-4. Eukaryotic DNA Polymerases With Special Functions

Polymerase	Source	Number of subunits	Mol. mass kDa	Function
η	yeast	1	72	Bypass of UV-induced lesions, accumulates at replication forks after DNA damage; defects cause xeroderma
	human	1	78	pigmentosum variant disease
ι	yeast			Bypass of normal replication-blocking DNA lesions
	human	1	80	
κ	yeast			Bypass of normal replication-blocking bulky DNA lesions
	human	1	99	
λ	yeast	?	?	Base-excision repair, non-homologous end joining, has DNA polymerase and terminal transferase activities
	human	1	63	
μ	yeast	?	?	Non-homologous end joining, sister-chromatid exchange
	human	1	55	
θ	yeast	?	?	DNA repair, e.g., interstrand DNA crosslinks (?), ATPase activity associated
	human	1	290	
ζ	yeast	3	173, 112, 29	Bypass of normal replication-blocking DNA lesions
	human	2	353, 24	
Revψ	yeast	-	-	(in yeast part of POLZ)
	human	1	138	Incorporation of dC opposite abasic sites
ν	yeast	?	?	unclear
	human	1	100	

Table 3.9.1-5. Other Important Proteins Involved in DNA Replication

Enzyme	Source	No. subunits	Mol. mass kDa	Function
DNA Helicases	eukarya	1	46…110	Associate with various DNA (and RNA) polymerases and catalyze DNA unwinding by disruption of hydrogen bonds between duplex DNA strands (ATP dependent). Action proceeds either in 5′→ 3′ or in 3′ → 5′ direction, yielding stretches of single stranded DNA for replication, repair (3.9.2), recombination, transcription and conjugation. More than 20 eukaryotic helicases have been characterized, their role and mechanism of action still remains to be elucidated in most cases.
DNA Topoisomerase I	eukarya	1	95…165	Removes positive superhelical tension by inducing a transient single strand nick into the DNA template ahead of the replication fork. Inhibited by camptothecin.
RPA (RFA)	yeast mammals	3 3	69, 36, 13 70, 34, 11	Unspecific single strand and sequence-specific double-strand DNA binding protein. The 36 kDa subunit is periodically phosphorylated by a cyclin-dependent kinase during the cell cycle. The 69 kDa subunit contains single strand DNA binding and origin-unwinding activity.
RFC	yeast mammals	4 3	120, 40, 37, 36 140, 41, 37	Essential accessory factor to DNA polymerases δ and ε. It binds to the primer terminus (ATP dependent) and recruits PCNA as a ring around DNA. The subunits are highly homologous among each other and between eukarya. Possibly analogous to the complex of *E. coli* DNA Pol III (3.8.1.2).
Proliferating cell nuclear antigen (PCNA)	eukarya	3 (homotrimeric)	26 (*3)	Essential accessory factor to DNA Pol δ. Binds possibly to DNA analogous to the β-subunit of *E. coli* DNA Pol III (3.8.1.2). DNA is thought to be threaded through a ring (sliding clamp) of PCNA, preventing DNA polymerase to dissociate from the DNA during synthesis. Participates in nucleotide excision repair of UV-damaged DNA (3.9.2.2).
Ribonuclease H				Degradation of RNA primers in Okazaki fragments.
Exonuclease MF1	yeast human	1 1	118 43	Degradation of RNA primers in Okazaki fragments in 5′→ 3′ direction. Cooperates with RNase H.
DNA ligase	yeast mammals	1 (Cdc9) 1 (Ligase I)	ca. 85 102	Catalyzes the formation of phosphodiester bonds at single-strand breaks in double-stranded DNA. Of the 3 mammalian DNA ligases, ligase I is used in replication.
Telomerase	yeast human	many many (>9)	103 (cat. subunit) 127 (cat. subunit)	RNA dependent DNA polymerase, contains essential RNA. Extends 3 DNA strand ends by addition of telomeres.
DNA Topoisomerase II	eukarya	2 (homodimeric)	100…180 (*2)	Required for segregation of daughter DNA molecules after replication. Relaxes supercoils; acts via cutting and re-ligating of double strands. It is also important as a nuclear scaffolding protein.
Histones	eukarya	2* (2a, 2b, 3, 4)	14, 14, 15, 11	Basic chromosomal proteins, essential components of nucleosomes. Histone H1 binds to DNA between nucleosomes (2.6.4)

energy is required. (Contrary to this, in bacterial topoisomerase type IA the 5′ phosphate group is temporarily bound by tyrosine.)

The dimeric topoisomerase type II (Fig. 3.9.1-4) breaks both strands of duplex DNA 4 bp apart. A covalent bond between 5′ phosphate groups of the DNA and a pair of tyrosyl groups of the enzyme is formed. Then another unbroken DNA duplex passes through the gap. The reaction requires ATP in order to achieve the changes in the protein configuration, which allow the passage. [A similar mechanism is used by bacterial gyrase (3.8.1.3), where a DNA duplex of about 140 bp is wrapped around the enzyme, crossing itself. Duplex strand passage through the gap produces negative supercoils.]

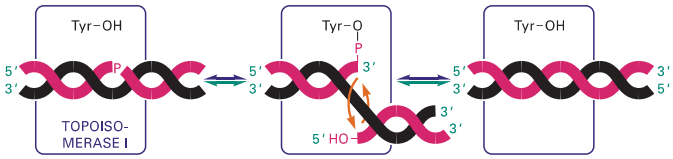


Figure 3.9.1-3. Mechanism of DNA Topoisomerases IB (3′, Rotation Model)

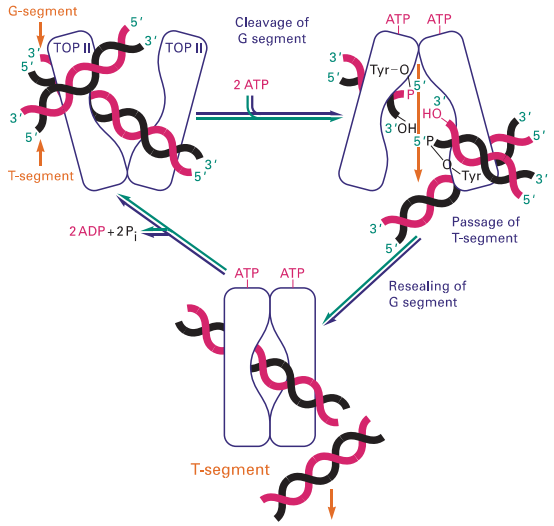


Figure 3.9.1-4. Mechanism of DNA Topoisomerases II

3.9.1.5. Telomeres (Fig. 3.9.1-5)

In replicating the lagging strand of linear DNA a problem arises, when the 3′ end of the template strand is reached. The last RNA primer cannot be replaced by DNA since there is no further template for synthesizing another primer/Okazaki fragment. Therefore, with every round of replication, shortening of DNA would occur. In order to avoid gradual loss of essential genes, linear DNA carries telomere extensions of up to 2000 short, tandemly repeated DNA sequences at its 3′ end (species specific, 5 to 8 bases long, in humans TTAGGG).

In vitro, the G-rich sequences frequently fold back onto themselves by non-Watson-Crick G-G base pairing, forming hairpin loops

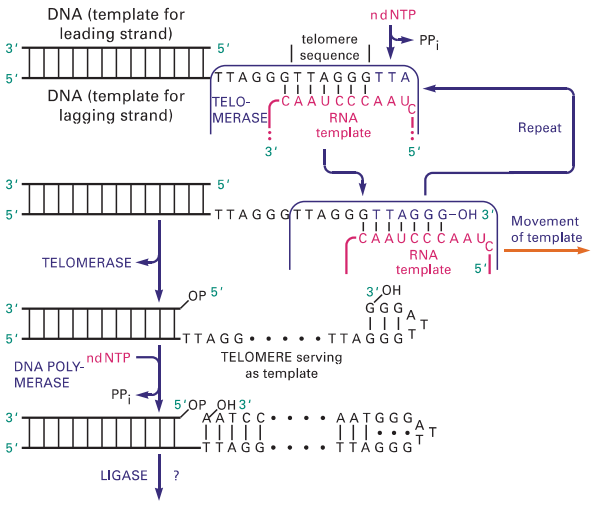


Figure 3.9.1-5. Formation of Telomeres

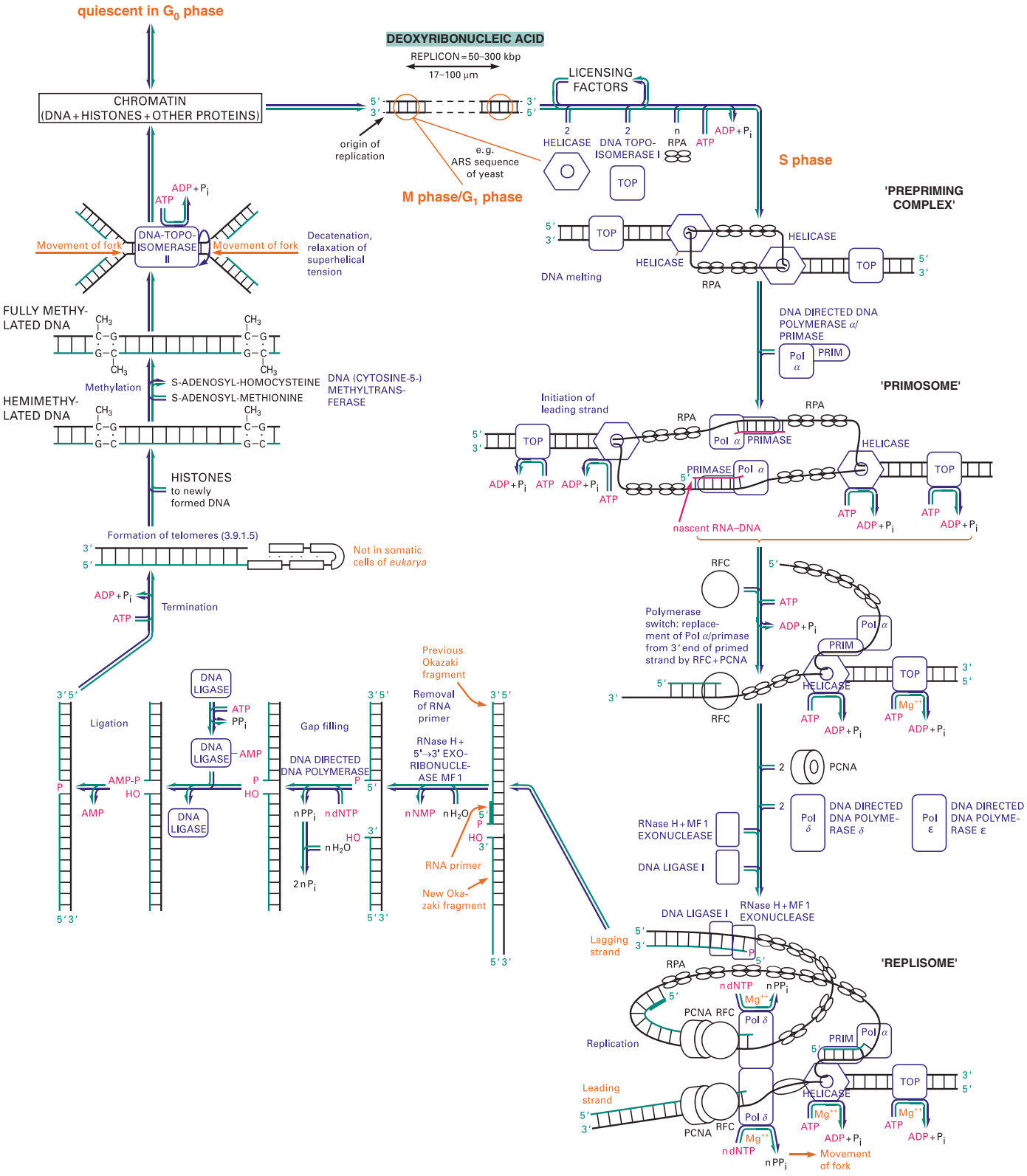


Figure 3.9.1-6. Eukaryotic DNA Replication (Not all details covered)

(t-loop) and covering the strand ends. This way, telomeres might keep chromosomes from fusing end-to-end, thus preventing chromosome breakage and loss during cell division. Chromosomes lacking telomeres are not stably maintained.

The formation of these extensions is catalyzed late in S phase by telomerase. The enzyme contains an essential RNA component serving as template for the synthesis of the telomere repeats. By synthesizing DNA from RNA templates, it resembles reverse transcriptase (5.4.1, except for using an internal template). The telomeres are bound by various telomere binding proteins (TBPs, like POT1, TRF1, TRF2) which play an important role in telomere length control.

While lower eukarya like yeast and also the germ line cells of higher eukarya contain telomerase, most somatic cells of higher eukarya do not express this enzyme (fetal cells, blood stem cells and

many mouse cells seem to be an exception). Therefore telomeres shorten with every round of DNA replication in somatic cells. This might limit their life-span and be part of a molecular clock measuring the age of cells. Aberrant telomerase activity has been implicated in the loss of normal cellular function and seems to be an inherent property of malignant cells making telomerase a promising target for tumor therapy.

3.9.1.6 Fidelity of Replication

Accurate replication is essential for maintaining genome stability. Base selectivity alone would lead to more than 1 error for each 100 nucleotides polymerized. Additional mechanisms, like proofreading and post-replicative mismatch repair (3.8.2.4) reduce the overall error rate to about 1 per 10^{10} nucleotides. This rate, as well as the relative contributions of the mechanisms for enhancing fidelity are similar to bacteria (3.8.1.4). The most frequent error is the deletion of a single nucleotide. Among the replicative DNA polymerases pol ε is the most accurate eukaryotic DNA polymerase, while pol α is the least accurate one (it lacks proofreading ability). In some hereditary forms of human cancer, impaired mismatch repair leads to several thousand-fold higher error rates in DNA replication.

Literature:

Aparicio, O.M. *et al:* Cell 1997;91:59–69.

Bell, S.P., Dutta, A. Annu. Rev. Biochem. 2002;71:333–374.

Bloom, J., Cross, F. R. Nat. Rev. Mol. Cell Biol. 2007;8:149–160.

Diller, J.D., Raghuraman, M.K. Trends in Biochem. Sci. 1994;19:320–325.

Loeb, L.A., Monnat, R.J. Nat. Rev. Genetics 2008;9:594–604.

Roca, J. Trends in Biochem. Sci. 1995;20:156–160.

Sugino, A. Trends in Biochem. Sci. 1995;20:319–323.

Stillman, B. Cell 1994;78:725–728.

Toyn, J.H., *et al.* Trends in Biochem. Sci. 1995;20:70–73.

Wang, T.A., Li, J.J. Current Opin. in Cell. Biol. 1995;7:414–420.

Wang, J.C. Ann. Rev. of Biochem. 1996;65:635–692.

Wang, J.C. Nature Rev.Mol.Cell Biol. 2002;3:430–440.

3.9.2 Eukaryotic DNA Repair

3.9.2.1 DNA Damage and Principles of Repair

DNA is a highly vulnerable molecule due to its enormous size. It can be damaged by many exogenous agents (e.g., environmental chemicals, UV or ionizing radiation). Also, spontaneous DNA damage arises from hydrolytic removal of bases, deamination of cytosine to uracil and (more rarely) adenine to hypoxanthine and guanine to xanthine, non-enzymatic DNA methylation by S-adenosyl methionine or damage by free oxygen radicals (reactive oxygen species, ROS) (3.2.5.8, 3.8.2.1) as well as replication errors. In humans, these types of spontaneous damage occur at a daily rate in the order of 50,000 per cell.

The ability to minimize the number of inheritable mutations caused by damage to DNA is essential for all organisms in order to maintain genome integrity. Besides mutagenicity, DNA damage can also interfere with transcription and replication. Therefore, an intricate network of various systems for recognition and repair of DNA damage has evolved.

DNA repair in eukarya appears to follow the same basic reactions as in bacteria (3.8.2), but is more complicated due to the larger genome and its more complex structure. Similarities in DNA repair mechanisms extend from yeast to humans. Some of the components can even be exchanged *in vivo*.

Repair is often associated with DNA replication and transcription when DNA is supposedly more accessible and repair is more essential than in a quiescent state. Transcriptionally active regions are repaired faster than 'silent' regions (transcription-coupled repair). Genes transcribed by RNA polymerases I or III are not subject to this specific type of DNA repair. Eukaryotic cells have checking mechanisms that can arrest cell cycle progress at defined "checkpoints" when damaged DNA is detected (4.3.6).

Figure 3.9.2-1 gives an overview of the basic features of different repair pathways. Covalent modifications (ubiquitinylation, SUMOylation, etc.) which also play an important role in many of these processes have been omitted.

3.9.2.2 Direct Reversal of Damage (Fig. 3.9.2-1)

Reversal of alkylation: Homologues to the specialized alkyltransferases of *E. coli* for direct repair of alkylation damage (3.8.2.2) have also been found in eukarya. O^6-methylguanine-DNA-protein-cysteine S-methyltransferase (O^6-MGT I) is a 'suicide' enzyme that removes methyl groups from O^6-methylguanine to a cysteine-SH group of the enzyme at the expense of its own inactivation.

Photoreactivation: A deoxyribodipyrimidine photolyase (similar to bacteria, 3.8.2.2) has been found in *S. cerevisiae*, some other fungi and in algae, but so far not in mammalian cells. A similar enzyme in animals repairs the 6-4 photoproduct (another pyrimidine dimer).

3.9.2.3 Excision Repair Systems (Fig. 3.9.2-1)

Excision repair relies on the redundant information in the DNA duplex to remove a damaged base or nucleotide and to replace it with the correct base by using the complementary strand as a template.

Base excision repair (BER): With this mechanism, damages caused by ionizing radiation, oxidizing and methylating agents (including products of the endogenous metabolism) and by loss of bases are repaired. It has only a limited substrate range since damaged bases must be recognized by specific DNA glycosylases.

In base excision repair single altered bases are targeted. Each of the DNA-glycosylases (e.g., uracil DNA-glycosylase or 3-methyladenine DNA-glycosylase; the EC list uses 'glycosidase') recognizes specifically a particular type of modified bases and removes them by hydrolyzing the N-glycosidic bond between the base and the deoxyribose-P backbone of the DNA. The resulting 'apurinic' or 'apyrimidinic' site (AP site) is recognized by specialized AP-endonucleases which cut the DNA backbone 5′ (upstream) to this location. A deoxyribophosphodiesterase (dRPase, an exonuclease) removes the base-free sugar phosphate. DNA polymerases (β, possibly also α or ε) fill the gap with the correct nucleotide. Then the strand is closed by action of a DNA ligase. In mammals, this is DNA ligase III.

Nucleotide excision repair (NER): This mechanism is coupled to transcription with selective preference to the transcribed strand. Since NER responds to distortions in DNA structure and not to individually recognized lesions, it is an almost universal repair system which removes a wide spectrum of structurally unrelated damages.

Examples are most UV-induced photoproducts (e.g., cyclobutane pyrimidine dimers, 3.8.2.1), bulky chemical adducts (e.g., benzo[a] pyrene•guanine adducts caused by smoking, cis-platin•guanine adducts), methylated bases (e.g., O^6-methylguanine) and intrastrand as well as interstrand crosslinks. However, sometimes it also acts (inefficiently) on mismatches with no strand preference, which can lead to fixation of mutations.

Although there are no sequence homologies to the *E. coli* UvrABC endonuclease system (3.8.2.3), the DNA strand in eukarya is cut similarly on both sides of the lesion by a specialized excision nuclease complex. Eukaryotic nucleases hydrolyze the 3rd to the 8th phosphodiester bond 3′, followed by the 20th to 25th phosphodiester bond 5′ of the lesion, removing a patch of 22 to 32 (in humans 27 to 29) nucleotides. The incision step alone requires many different proteins (10 in yeast, 16 in humans). A selection is shown in Figure 3.9.2-1 and Table 3.9.2-1. RFA temporarily covers the separated strands. The gap is then filled by DNA polymerase δ or ε and the continuity of the phosphodiester backbone is restored by DNA ligase.

Some of the repair factors in humans are also part of the transcription machinery (4.2), e.g., TFIIH, TFIIJ. Coupling of repair to transcription is effected by CSA/ERCC8 and CSA/ERCC6 proteins, which temporarily replace RNA polymerase stalled by the NER system.

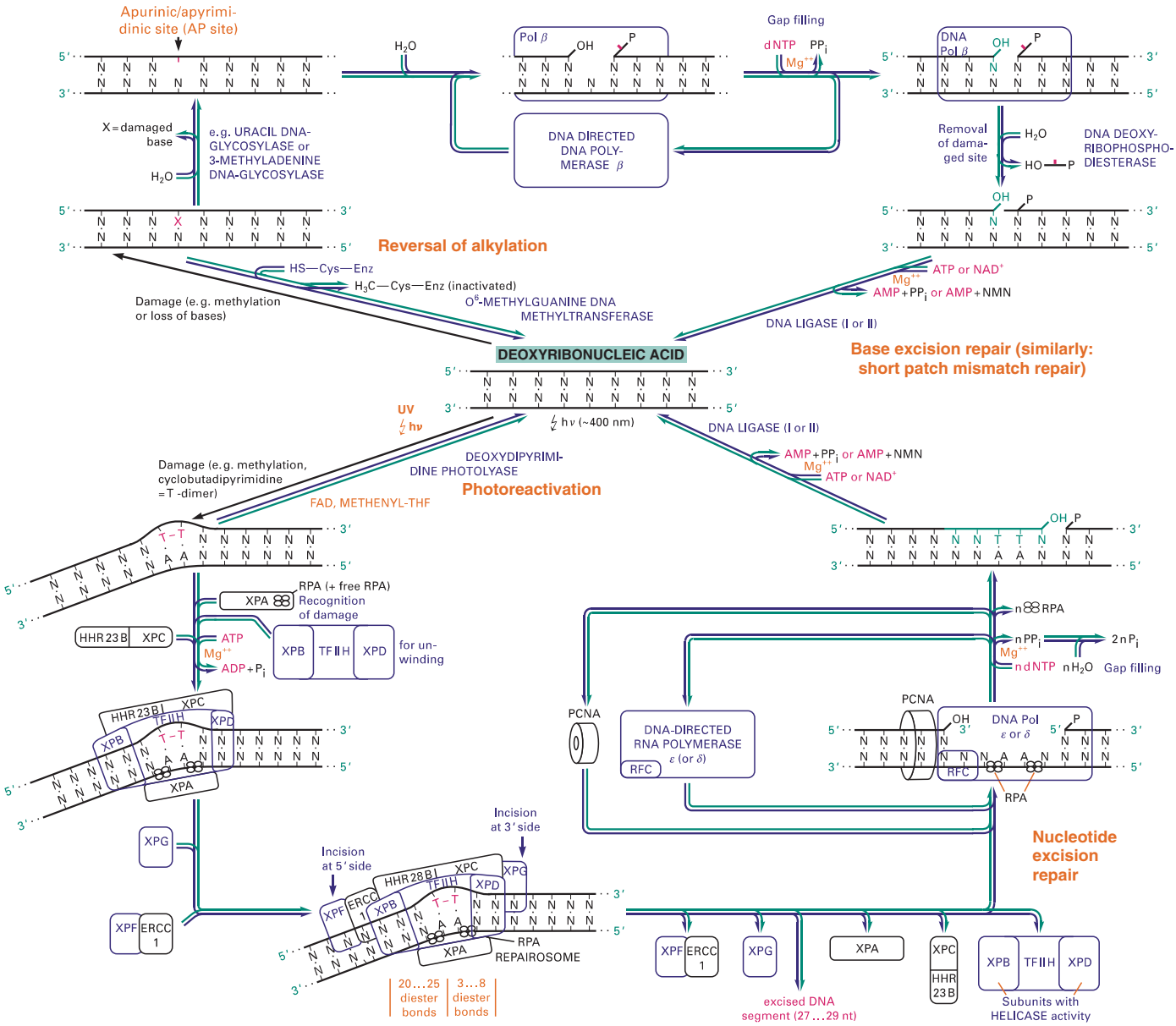


Figure 3.9.2-1. Systems of Eukaryotic DNA Repair. The terminology of the human system is used.

Most of the activities first described in *E. coli* and in *S. cerevisiae* have also been found in mammalian cells.

Table 3.9.2-1. Protein Factors in Eukaryotic Nucleotide Excision Repair (Selection)

Gene product	Mol. mass kDa (human)	Function
Damage excision:		
XPA	31	recognition of damage (zinc finger)
RPA (RFA)	(see 3.9.1.4)	same, binds to single-stranded DNA
XPC/RAD23B	102, 43	stabilization of preincision complex
TFIIH	(see 3.9.1.4)	formation of preincision complex, transcription-repair coupling
XPB (ERCC3)	89	3′→5′ helicase, TFIIH subunit
XPD (ERCC2)	87	5′→3′ helicase, TFIIH subunit
XPG (ERCC5)	133	DNA incision, 3′ side of lesion
XPF (ERCC4)/ ERCC1	104, 33 (1:1 complex)	DNA incision, 5′ side of lesion
Gap filling:		
RFC	(see 3.9.1.4)	loads PCNA onto DNA (ATP-dependent)
PCNA, DNA polymerase β, δ or ε, DNA ligase – see 3.9.1.3		

3.9.2.4 Mismatch Repair

Base pairs in a double-stranded DNA helix not conforming to the Watson-Crick rules (A-T, G-C, 2.6.1) are called mismatches. Such incorrect base pairs arise for several reasons:

- occasional errors in proofreading mechanisms during DNA-replication
- formation of incorrect heteroduplexes between DNA-strands during recombination
- deamination of 5-methylcytosine to thymidine, causing a G-T mismatch.

Thus, this type of repair system is preferably connected to replication and recombination events. It follows the same principles as the excision repair systems mentioned above. The main problem is the distinction of the correct from the incorrect base, since both molecules are normal constituents of DNA. As in *E. coli*, mismatches in mammalian cells can be repaired by two different mechanisms depending on the sequence context:

- A short patch system with preference for G-T mismatches originating from deamination of (5-methyl-)C at CpG sites. It resembles base excision repair (3.9.2.3) and likewise shows restricted specificity. After thymine removal, 10 or less nucleotides (down to a single

one in mammals) are excised. The resulting gap is closed by DNA polymerase β.

- Likewise, the mismatch-specific thymine-DNA-glycosylase of humans is also able to remove dU generated at CpG sites by deamination. There are also hints for a mismatch repair mechanism in mammals which removes A (mispaired to G) by incisions at both sides.
- a long patch mechanism using MutS and MutL homologues similar to the bacterial methyl-directed pathway (3.8.4, recognition of the correct strand at methylated moieties after replication, or of the defective strand at breaks). This system detects mismatches in many different contexts. Fairly large pieces of DNA (up to >1000 bases) are excised, this may occur in either direction. The gap filling is similar to nucleotide excision repair, possibly by DNA polymerase δ or ε.

3.9.2.5 Double-strand Repair and Recombination

One of the severest DNA lesions occurring in a cell is a double-strand break, because it completely disrupts the integrity of the DNA molecule. This is the most frequent effect of ionizing radiation. Double strand ends of broken chromosomes frequently cause DNA recombination and mutations. However, double-strand breaks also occur during natural recombination, e.g., in immunoglobulin gene maturation or during meiosis.

Repair can be achieved by homologous recombination (crossover between damaged DNA and correct 'sister' DNA). Several of the enzymes required for this type of repair in simpler organisms are conserved in eukarya. E.g., in yeast RAD51 is a functional homologue to RecA of *E. coli*. The RAD52 enzyme is mainly involved in repair of double strand breaks. The repair mechanisms in *Saccharomyces cerevisiae* appear to be similar to bacteria (3.8.2.5). Generally, homologous recombination does not occur often in higher eukarya.

Although the genes involved in recombination are conserved, in higher eukarya sometimes 'illegitimate' events occur. These are reactions which rely on very short stretches of homology or no homology at all between the recombining molecules ('non-homologous end joining').

Other enzyme activities: Eukaryotic cells contain a DNA-dependent protein kinase (DNA-PK, 465 kDa). The two Ku subunits of this enzyme (80 and 70 kDa) show ATP dependent helicase activity and bind to DNA ends, hairpins, nicks and gaps irrespective of the DNA sequence and thus likely recognize damage. Cellular targets of DNA-PK include many DNA binding regulatory proteins (e.g., c-Jun, c-Fos, c-Myc, SRF), TFIID, RNA polymerase II (CTD domain), DNA topoisomerases I and II etc. (3.9.1.4). It also autophosphorylates and deactivates itself. Its exact role in the repair of double strand breaks is still under investigation. The enzyme is also involved in immunological processes [V(D)J recombination, 8.1.4]. In SCID cells (severe combined immunodeficiency 3.6.1.7) it appears to be defective.

3.9.2.6 Translesion Synthesis (TLS):

In recent years a large number of new eukaryotic DNA dependent DNA polymerases have been discovered. They have no proofreading function and their error rates vary between 1 in 10 (Pol ζ, specialized to bypass thymine dimers) and 1 in 10^4 (measured on undamaged DNA). These enzymes share almost no primary sequence homology with replicative DNA polymerases, but do have some similar structural features and a common overall architecture. They are highly conserved from yeast to humans (Table 3.9.1-4).

Their main function is to temporarily bypass lesions to allow completion of DNA replication with relaxed emphasis on protecting the accuracy of genetic information. Given that these enzymes are essential and mutations in their genes are causing severe disease phenotypes, the cost (increased mutation rate)/benefit (completion of replication and avoidance of even more severe damage) ratio must be positive for translesion synthesis. Without this kind of DNA damage tolerance, cells could face the risk of replication fork collapse, chromosomal aberrations and cell death.

Their exact mode of action is still a matter of debate, with one model proposing a polymerase switch directly at the fork while another is more in line with the classical post-replication repair model, assuming gap-filling after the fork has processed. Interaction between TLS polymerases and modified (ubiquitinylated) PCNA is important for TLS regulation.

3.9.2.7 DNA Repair and Human Diseases

Defects in DNA repair cause severe diseases, e.g., most human cancers result from DNA damage. Human DNA repair has mostly been studied in genetic diseases with impaired repair function. Examples are xeroderma pigmentosum (high sensitivity against UV radiation), Fanconi anemia (cross-link sensitive), Cockayne's syndrome and trichothiodystrophy, all of which are caused by defects in nucleotide excision repair. In ataxia telangiectasia the repair of double-strand breaks is defective; the patients are extremely sensitive to ionizing radiation. Damage in genes for long patch mismatch repair mechanisms (hMSH2 and hPMS1 on chromosome 2, hMLH1 on chromosome 3p21 or hPMS2 on chromosome 7) are associated with hereditary nonpolyposis colorectal cancer (HNPCC; Lynch syndrome II, Muir-Torre syndrome), which is characterized by instabilities of simple repeated sequences. This is one of the most common genetic diseases of humans (1 in 200 individuals).

Literature:

Bergink, S., Jentsch, S. Nature 2009;458:461–467.

Branzei, D., Foiani, M. Nat. Rev. Mol. Cell Biol. 2008;9:297–308.

Friedberg, E.C., *et al. DNA Repair and Mutagenesis*. 2nd Ed. ASM Press, 2005.

Hanawalt, P.C., *et al.* Current Biol. 1994;4:518–521.

Lehmann, A.R., Fuchs, R.P. DNA repair 2006;5:1495–1498.

Misteli, T., Soutoglou, E. Nature Rev. Mol. Cell Biol. 2009;10:243–254.

Sancar, A. Ann. Rev. of Biochem. 1996;65:43–81.

Umar, A., Kunkel, T.A. Eur. J. Biochem. 1996;238:297–307.

Waters, L.S., et al. Microbiol. Mol. Biol. Rev. 2009;73:134–154.

Wood, R.D. Ann. Rev. of Biochem. 1996;65:135–167.

3.10 Special Bacterial Metabolism and Biosynthesis of Antimicrobials

Julia Garbe, Annika Steen and Max Schobert

3.10.1 Bacterial Envelope

The bacterial envelope consists of the cytoplasmic or inner membrane, the murein (= peptidoglycan) and, in the case of Gram-negative bacteria, the outer membrane. Structural differences of the envelope, which can be assayed by Gram staining are important for bacterial taxonomy.

The envelope is important for a number of tasks:

- Mechanical and chemical protection
- Osmotic barrier
- Adhesion to surfaces, extracellular polymeric substances or other cells
- Detection and selective uptake of substrates
- Exchange of genetic material (conjugation)
- Membrane vesicle formation in Gram-negative bacteria (trafficking of biological material, toxins and interaction with other cells).

Murein synthesis (Fig. 3.10.1-1, lower part): The peptidoglycan consists of linear polysaccharide strands of alternating N-acetyl-glucosamine □ and N-acetyl-muramic acid ⊗ (MurNAc) units linked by β (1→4) glycosidic bonds. The strands of about 30 units are cross-linked by short peptides attached to the lactyl group of MurNAc (Fig. 2.9-3) forming a covalently closed sacculus. This sac-like structure is extremely rigid and helps to maintain cell structure (e.g., rods, cocci, spirilloids, filaments). Moreover, it preserves the integrity of the cytoplasmic membrane from rupture in medium with low osmolarity.

After the assembly of the disaccharide-peptide monomer unit, the polymerization of the peptidoglycan takes place at the outside surface of the cytoplasmic membrane. The units of the chains are synthesized at undecaprenyl membrane anchors in the cytoplasm and are transferred through the membrane to the growing chain. The glycan strands run around the cell axis, the peptide crosslinks parallel to the axis (Fig. 3.10.1-2). During cell wall synthesis, new material is inserted

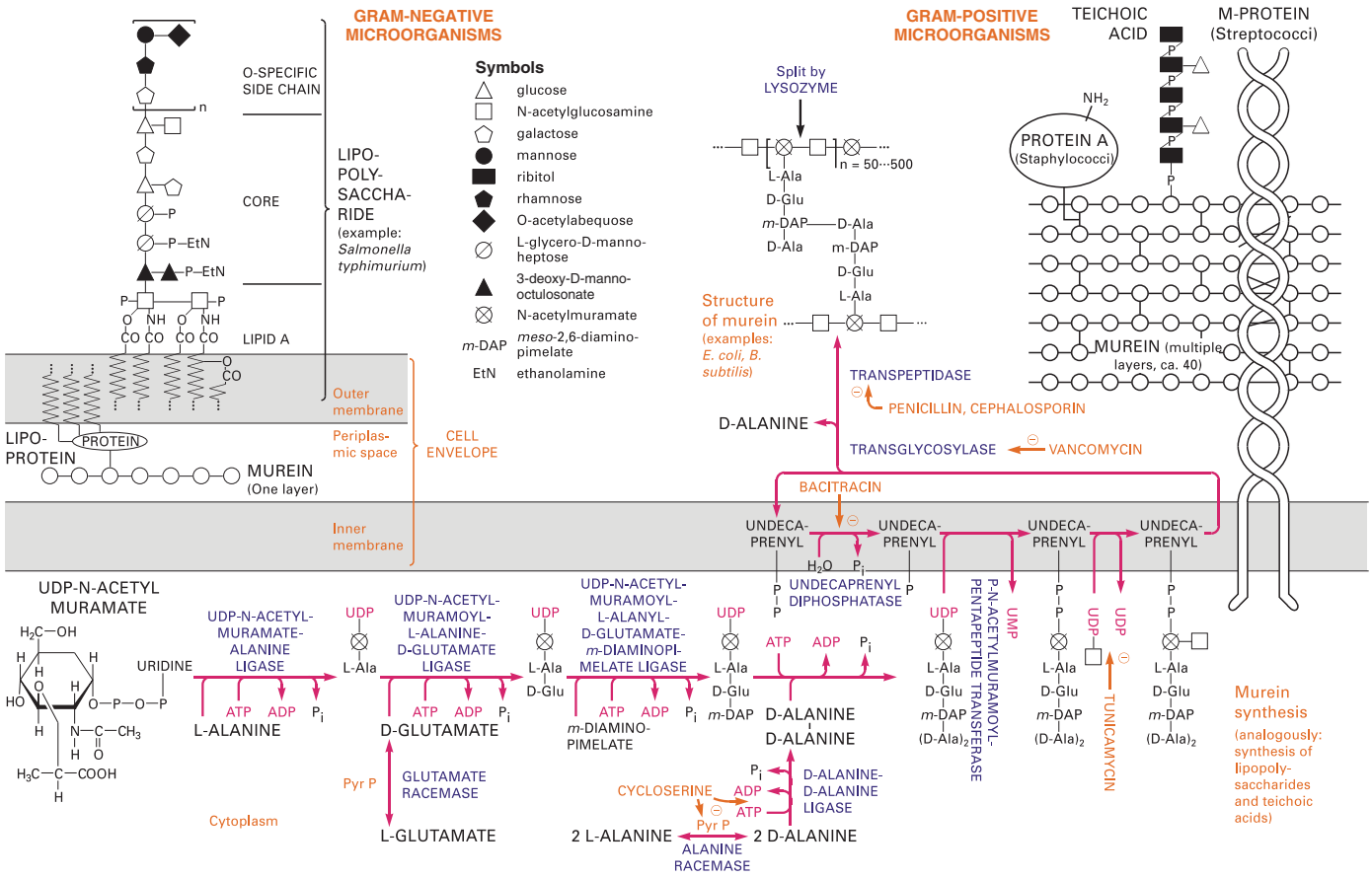


Figure 3.10.1-1. Bacterial Cell Envelope

around the equator of the cell. Details of the peptidoglycan structure vary among bacteria. Particles of approximately 2 nm can pass the peptidoglycan layer.

Gram-negative bacteria have been shown to reutilize all components of their cell wall, which is important in the competition for food during cell growth. In *E. coli,* at least nine enzymes, one permease and one periplasmatic protein are involved in this recycling process.

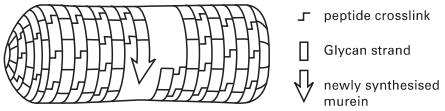


Figure 3.10.1-2. Growth of the Murein Layer

Gram-negative bacteria (Fig. 3.10.1-1, left): The inner membrane is surrounded by one giant molecule of murein. It is covered by the outer membrane, which differs chemically from all other biological membranes, because the phospholipids in the outer leaflet are replaced by the unique lipopolysaccaride (LPS). The exact chemistry of the LPS, especially of the polysaccharides, varies among bacteria. It is composed of the Lipid A, the core polysaccharides and the O-antigen.

- Lipid A is a glycolipid. Its fatty acids are linked via amide linkages to glucosamine. This part of the LPS is also known as the endotoxin. When released by lysed cells, endotoxin overstimulates host defenses and can induce a lethal endotoxic shock.
- The core polysaccharide contains 2-oxo-3-deoxy-octanoic acid, which is linked to the lipid A and other heptoses and hexoses.
- The O-specific side chain or O-antigen consists of mostly hexoses and is the most variable part of the whole LPS. This highly polar saccharide surface renders Gram-negative bacteria resistant to detergents, e.g., bile salts. The O-antigen units are assembled at a lipid carrier molecule, identified as undecaprenyl phosphate (3.5.3) and are transferred to the LPS core. Porins span the outer membrane and allow passive diffusion of low molecular weight substances (3.10.3).

Gram-positive bacteria (Fig. 3.10.1-1, right): Gram-positive bacteria possess a multilayered peptidoglycan pervaded with various proteins and polymers. This murein wall grows from the innermost to the peripheral layer.

An outer membrane is absent in the bacterial envelope. Accessory anionic polymers such as wall teichoic acids, lipoteichoic acids or teichuronic acids exert various functions such as binding of divalent cations, support folding of extracellular metalloproteins and act as diffusion barriers. Wall teichoic acids are quite diverse in structure and abundance, depending on bacterial strain or nutrient availability. Their overall structure consists of a chain of phosphodiester-linked glycerol or ribitol residues attached to a terminal linkage unit, allowing a covalent attachment to the murein layer. Lipoteichoic acids differ from wall teichoic acids in the chirality of their chain compounds and are anchored to the membrane by glycolipids. Like lipopolysaccharides of Gram-negative bacteria, teichoic acids are potent immunogens. The main chain of teichuronic acid consists of sugar monomers directly linked by glycolytic bonds.

Serotyping, diagnosis: Due to antigenic surface components, some bacterial species, especially pathogens, can be classified into different immunological groups. For example, the Gram-positive *Streptococci* can be differentiated according to the polysaccharides attached to murein (Lancefield groups A, B, C, F, G), teichoic acids (D) and lipoteichoic acids (N). Differences in the sugar composition of the O-specific side chains allow the identification of several thousand serotypes of the Gram-negative *Salmonellae*.

Mycobacteria: The mycobacterial cell wall has some unique characteristics. It consists of the cell membrane surrounded by the inner and outer layer. The inner layer is composed of peptidoglycan, arabinogalactan and mycolic acids, which form a covalently linked insoluble complex, characteristic of mycobacteria. The outer layer consists of both proteins and lipid-linked polysaccharides. Due to this unusual and robust cell wall, mycobacteria are impermeable to a number of compounds including drugs or acids.

Literature:

Schäffer, C., Messner, P. Microbiology 2005;151:643–651.

Delcour, J. *et al*. Antoine van Leeuvenhoek 1999;76:159–184.

King. J.D., *et al*. Innate Immunity 2009;15:261–312.

Hett, E.C., Rubin, E.J. Microbiol. Mol. Biol. Reviews 2008;72: 126–156.

Park , J.T., Uehara, T. Microbiol. Mol. Biol. Reviews 2008;72: 211–227.

3.10.2 Bacterial Protein Export across the Cytoplasmic Membrane

Extracellular, periplasmic and outer membrane proteins are translocated across the bacterial cytoplasmic membrane either via the general secretion (Sec-) pathway or the twin arginine translocation (Tat-) pathway. The Sec-system is involved in the transport of unfolded proteins and in the insertion of membrane proteins. Unfolded proteins acquire their final structure in the periplasm supported by special chaperones. In contrast, the transport of folded or cofactor-containing proteins, such as compounds of respiratory or photosynthetic transport chains, is mediated by the Tat-system. The TC-system (see 6.1.1.2) lists the Sec family under 3.A.5, while the Tat-system has the number TC 2.A.64.

Sec-dependent transport (Fig. 3.10.2-1, Table 3.10.2-1): Secretory proteins are tagged N-terminally for translocation either via the secA/B or SRP-dependent pathway. The signal sequence consists of a positively charged n-region followed by a hydrophobic h-region and a

Table 3.10.2-1. Factors Involved in Protein Export via Sec-pathway of *Escherichia coli*

Component	Factor	Mol. Mass (kDa)	Function
Export chaperone	SecB	18	Binds and stabilizes unfolded preprotein
Export ATPase	SecA	102	Peripheral membrane protein; associated with SecY chaperone, binds signal peptide of preprotein, directs it to translocase; initiates translocation.
Translocase	SecY SecE SecG	48 14 15	Trimeric integral membrane complex, forming the translocation channel (homologous to eukaryotic Sec61complex)
	SecD SecF	67 39	Integral membrane proteins, facing the periplasm. Regulate separation of SecA
Signal peptidase	LepB	36	Membrane bound protease facing the periplasm, cleaves signal peptide
Signal recognition particle (SRP)	P48 = Ffh RNA	48 4.5S	Chaperone, binds signal peptide of nascent protein; GTPase; activates GTPase of FtsY (homologous to eukaryotic SRP 54)
SRP receptor	FtsY		GTPase; interacts with SRP and activates SRP GTPase; homologous to eukaryotic SR

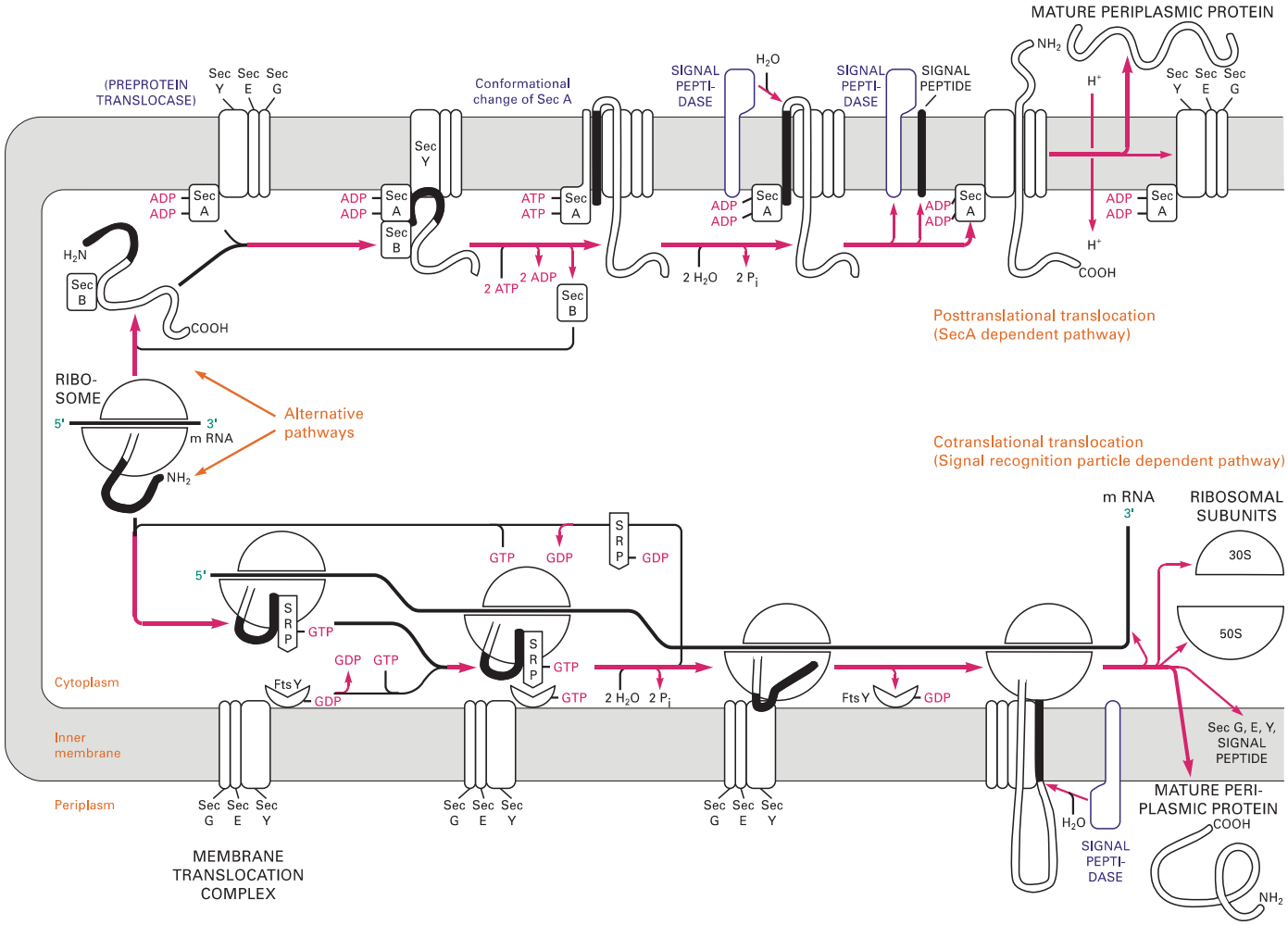


Figure 3.10.2-1. Pathways of Bacterial Protein Export via the Sec-Pathway through the Cytoplasmic Membrane

short polar c-region with a type I signal peptidase cleavage site in case of cleavable signal sequences.

- Posttranslational targeting, SecA/B dependent pathway (Fig. 3.10.2-1 upper part) is mediated by the chaperone SecB, which guides the unfolded protein to the membrane-anchored Sec-translocase. The Sec-translocase consists of the proteins SecE, SecG and SecY, which form a trimeric transmembrane channel. Protein export is initiated by the signal peptide recognizing ATPase SecA and driven by ATP and the proton motive force. As soon as the C-terminus of the signal peptide reaches the periplasmic surface of the membrane, it is cut off by the signal peptidase.
- Cotranslational targeting or SRP-dependent pathway (3.10.2-1 lower part) occurs when a signal recognition particle (SRP) binds to the secretory protein during its emergence from the ribosome. SRP is recognized by the FtsY receptor and transported via the Sec-translocase. In some cases, protein export is also initiated by SecA but it seems not to be essential. This pathway is used by membrane proteins and some preproteins.

Tat-dependent transport (Table 3.10.2-2): The N-terminal signal peptides of folded proteins recognized by the Tat-system share the same overall organization of Sec signal sequences with two important differences: Tat signal peptides contain two characteristic, conserved arginine residues and a significantly longer n-region. The Tat-translocase is composed of membrane-integrated proteins, forming either TatAC or TatABC complexes, which then function as a transmembrane channel. The energy for protein export is provided by the proton motive force and is more 'costly' than the Sec-dependent transport.

Table 3.10.2-2. Factors involved in protein export by the Tat-system of *Escherichia coli*

Compound	Protein	Mol. mass (kDa)	Function
Translocase	TatA	9 kDa	Forms protein conducting channel together with TatC, single N-terminal transmembrane α-helix
	TatB	18kDa	Slight homology to TatA, not present in all bacteria, exact function unknown, forms complex with TatC
Translocase	TatC	30 kDa	Six transmembrane helices with both termini at the cytoplasmic side, recognizes twin arginine motif of the tat signal peptide

Literature:

Sargent, F. *et al*. FEMS Microbiol. Lett. 2006;254:198–207.

Driessen, A.J., Nouwen, N. Ann. Rev. Biochem. 2008;77:643–667.

3.10.3 Protein Transport across the Outer Membrane of Gram-Negative Bacteria (Fig. 3.10.3-1)

Due to an additional outer membrane, Gram-negative bacteria need special secretion systems, which mediate protein transport across the outer membrane. Among the proteins translocated across the outer membrane are exoenzymes and virulence factors. Transport across the outer membrane can be divided into the Sec dependent (type II and V) and the Sec-independent (type I, III, IV, VI) pathways (Fig. 3.10.3-1). An exception is the type IV dependent system, which is found also in Gram-positives. In the TC-system (6.1.1.2), the types I-IV are listed as TC 3.A.1, TC 3.A.5, TC 3.A.6 and TC 3.A.7.

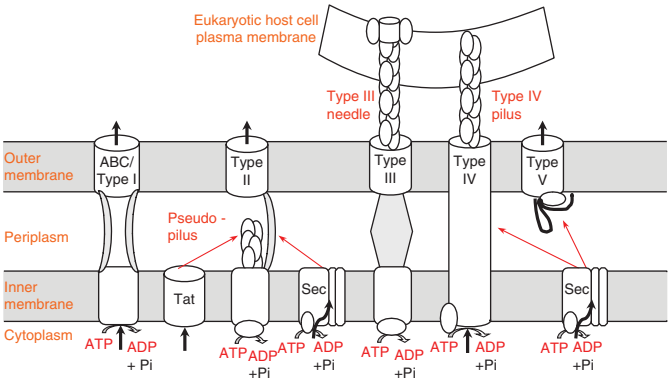


Figure 3.10.3-1. Protein Secretion Pathways

Sec/Tat-dependent secretion: The type II secretion system (T2SS) transports a wide variety of extracellular proteins including toxins and is closely related to type IV pili in *Pseudomonas aeruginosa*. Transport across the outer membrane occurs via a multiprotein complex with 12–16 accessory proteins. A NTPase forms a hexameric ring at the cytoplasmic side of the inner membrane and promotes the helical assembly of the pseudopili. These pseudopili are incorporated into a pseudopilus structure that pushes folded exoproteins through the secretion channel. The proteins are transported via both Sec- or Tat-dependent pathways into the periplasm.

The type V secretion system (T5SS) comprises autotransporter and two-partner secretion systems. T5SS is one of the most widely distributed secretion systems in Gram-negative bacteria. Typically, virulence factors use this autotransporter system (e.g., protease Ig A1 of *Neisseria gonorrhoeae*). The secreted protein consists of three domains. The first domain is the Sec-dependent signal sequence at the N-terminus for translocation across the inner membrane, the second domain comprises the exported protein ('passenger domain') and the third domain contains a C-terminal ß-barrel which forms a pore in the outer membrane through which the passenger domain passes the cell surface. In two-partner secretion systems a translocator mediates secretion of an effector protein through the outer membrane.

Sec-independent secretion: The type I secretion system (T1SS) is similar to an ABC transporter (3.10.4.2) and is involved in secretion of toxins, proteases and lipases (e.g., α-hemolysin of *Escherichia coli*). It consists of three proteins, which span the inner and outer membrane, two cytoplasmic membrane proteins, the ATP-binding cassette (ABC), the membrane fusion protein (MFP) and an outer membrane protein (OMP). The cargo proteins contain a C-terminal signal sequence and are transported directly across both membranes without a periplasmic intermediate.

Type III secretion system (T3SS) is a large complex consisting of about 20 different proteins. It spans both bacterial membranes and forms a needle-like structure allowing the injection of virulence factors across the host membrane into the eukaryotic host cytoplasm. The transport is ATP-dependent. Some components of T3SS are closely related to the flagellar basal body. Protein substrates for T3SS are presumed to carry a signal recognition sequence at the N-terminus.

Type IV secretion system (T4SS) is ancestrally related to the bacterial conjugation machinery and is characterized by a remarkable functional versatility. T4SS are multicomponent cell-envelope-spanning complexes. They consist of a secretion channel and often a pilus, a surface filament or protein, and translocate proteins or ssDNA-protein (relaxase) complexes. A well-investigated system is the translocation of DNA-protein complexes of *Agrobacterium tumefaciens*. T4SS recognize C-terminal and internal motifs of secretion substrates.

Type VI secretion system (T6SS) is a novel kind of complex multicomponent secretion machine first characterized in *Vibrio cholerae* and *Pseudomonas aeruginosa.* This system is often involved in interaction with eukaryotic cells but its precise role and mode of action are still unknown.

Literature:

Cascales, E. EMBO Rep 2008;9:735–741.

Christie, P.J., *et al*. Ann. Rev. Microbiol. 2005;59:451–485.

Cornelis, G.R. Nature Reviews Microbiol. 2006;4:811–825.

Delepelaire, P. Biochim. Biophys Acta 2004;1694:149–161.

Filloux, A., *et al.* Microbiology 2008;154:1570–1583.

Hauser, A.R. Nature Reviews Microbiol. 2009;7:654–665.

Mazar, J., Cotter P. A. Trends Microbiol. 2007;15:508–515.

Voulhoux, R., *et al.* EMBO J. 2001;20:6735–6741.

3.10.4 Bacterial Transport Systems

The bacterial membrane is a permeability barrier to most hydrophilic nutrients, except for gases or small molecules like O_2, H_2O, CO_2, NH_3, ethanol or hydrophobic molecules as aliphatic- or aromatic compounds. Therefore, bacteria have developed various mechanisms to transport nutrients into the cell.

Transport in Gram-positive bacteria involves only the inner membrane, because murein is no barrier for nutrients, except for polymers like starch and cellulose, which must be degraded extracellularly. In Gram-negative bacteria, nutrients must also pass the outer membrane. Hydrophilic molecules smaller than 600 Da cross the outer membrane through protein channels (porins) via passive transport.

3.10.4.1 Passive Transport

Passive transport across bacterial membranes occurs by either simple or facilitated diffusion, in which the diffusion process takes place through protein channels. These processes are energy independent. Both simple and facilitated diffusion use the concentration gradient to move a compound from a compartment of higher concentration to a compartment of lower concentration. Therefore, in the case of Gram-negative bacteria, for uptake of a nutrient its concentration in the periplasm must be lower than in the extracellular space. Since in most natural environments microorganisms live under substrate limitation, this requirement can only be fulfilled by removal of these compounds from the periplasm by active (= energy-dependent) transport across the inner membrane. This concentration gradient can also be achieved, if - for instance - a nutrient is rapidly metabolized or chemically modified upon entry into the bacterial cell. A well-characterized example for facilitated diffusion is the glycerol transporter GlpF of *E. coli* (Fig. 3.10.4-1 right, TC 1.A.8.1.1).

3.10.4.2 Active Transport

Transport of molecules across the membrane against a concentration gradient requires energy. Three main types of transport systems are involved in active translocation of molecules across the bacterial membrane. These include the secondary transport system, phosphotransferase system and ATP-binding-cassette (ABC) transporters.

Primary transport systems are energy conserving and couple translocation of solutes to chemical or photochemical reactions as, e.g., proton export driven by respiration or photosynthesis. These systems convert light or chemical energy into electrochemical gradients, which are used as driving force by secondary transport systems. An example for primary transport systems is proton translocation by respiration processes, which generates the proton motive force (PMF). The proton motive force is composed of a chemical pH gradient (ΔpH; c $[H^+_{outside}] > c\ [H^+_{inside}]$ and the electrical membrane potential (ΔΨ [mV]; $\Psi_{outside}$ +, Ψ_{inside} -):

$$\begin{aligned} PMF\ [mV] &= (-2.303 * R * T/F) * \Delta pH + \Delta\Psi\ [mV] \\ &= -59 * \Delta pH + \Delta\Psi \quad (at\ 25\ ^oC). \end{aligned}$$

Secondary transport systems (Fig. 3.10.4-1, left): Secondary transport is driven by electrochemical gradients generated by primary transport systems. A number of permeases are not coupled to the proton gradient (PMF), but to a secondary Na^+ gradient. This is especially important for bacteria living in alkaline environments, which cannot establish a H^+ gradient across the cytoplasmic membrane, therefore they exclusively use a Na^+ gradient.

Secondary active transporters use the energy of this disequilibrium to transfer nutrients into the cytoplasm or to secrete metabolic end products. A large family of transporters present in all organisms is the 'Major facilitator superfamily'. Transport occurs by various mechanisms:

- Symport: The uptake of nutrients is coupled to a component with a favorable electrical (ΔΨ) or chemical (e.g., ΔpH) gradient, such as lactose (in) / H^+ (in); lactate (in) / H^+ (in).
- Antiport: Contrary to symport, uptake is coupled to extrusion of a second component, or vice versa, e.g., Na^+ (out) / H^+ (in).
- Uniport: Only the membrane potential ΔΨ is used to transport cations into the cell or anions out of the cell, e.g., lysine$^+$ (in).

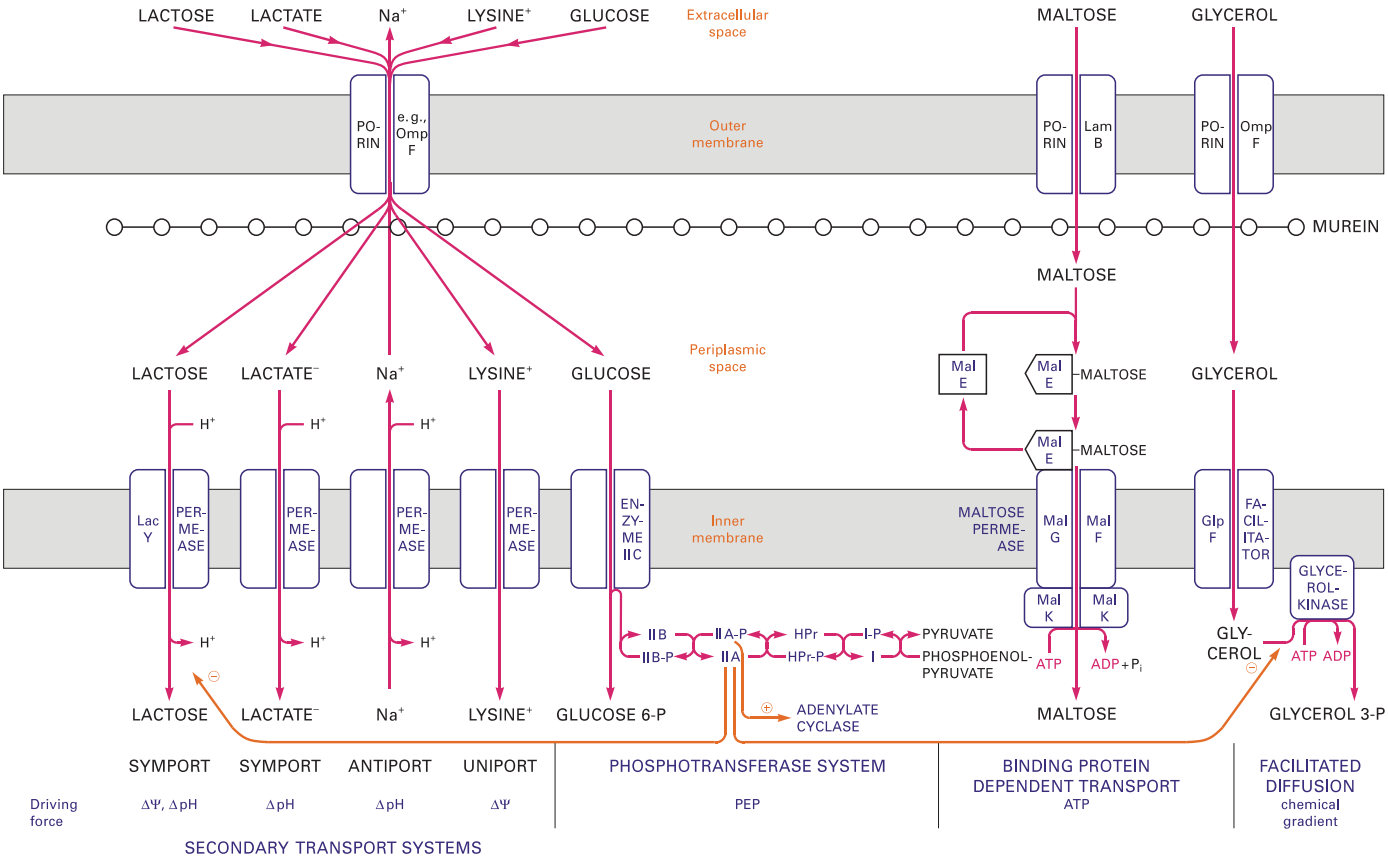


Figure 3.10.4-1. Examples for Bacterial Transport Systems

Phosphotransferase system (PTS) (Fig. 3.10.4-1, center): Substrates that enter the cells by this system are phosphorylated, yielding membrane impermeable derivatives in the cytoplasm (e.g. glucose into glucose 6-P). Compared to other uptake mechanisms, this system is energy conserving, because phosphorylation is also required for further metabolization of the nutrients.

As energy source, phosphoenolpyruvate is used. In Gram-negative bacteria, phosphate transfer is mediated by a sequence of enzymes: EI, histidine protein Hpr, EIIA, EIIB and the transmembrane protein EIIC (TC-system: TC 8.A.7, TC 8.A.8). Often, EIIA, EIIB and EIIC are fused to a single polypeptide.

In *E. coli*, EIIA of the glucose system regulates other uptake systems: if glucose transport is high, $EIIA_{Glc}$ is dephosphorylated (see Fig. 3.10.5-1). This form allosterically inhibits other uptake systems, like glycerol kinase or lactose permease. If glucose transport is low (glucose shortage), $EIIA_{Glc}$ is phosphorylated and activates adenylate cyclase, thus more cAMP is produced. The catabolite repressor protein (CRP) binds cAMP. This enables the transcription of genes, which encode enzymes involved in uptake or metabolization of alternative nutrients.

ATP-binding-cassette (ABC) transport systems (Fig. 3.10.4-1, right): These transporters are integral membrane proteins and form another large family of transporters, which carry out ATP-dependent import of substrates. In the periplasm of Gram-negative bacteria, the substrate is bound by a binding protein with high affinity (K_m = 0.1 to 1 µmol/l) and high specificity, which transfers the substrate to the membrane permease. ATP hydrolysis induces a conformational change of the permease, which allows the substrate to enter the cytoplasm (Example: maltose bound in the periplasm to MalE, see Fig. 15.4-1, TC-system 3.A.1.1.1). Gram-positive bacteria lack an outer membrane and a periplasm, therefore the binding protein is anchored in the inner membrane.

Iron transport: Iron is an essential nutrient of bacteria. However, in aerobic environments iron is present as oxidized ferric ion (Fe^{3+}), which is insoluble and unavailable for transport. For ferric ion uptake, bacteria synthesize and secrete siderophores. These molecules chelate ferric ion in the environment and the siderophore-iron complex is imported into the cells by ABC transport systems. Siderophores have a variety of chemical structures and require distinct uptake systems.

3.10.4.3 Efflux Systems (Fig. 3.10.4-2)

Efflux systems play an important role in bacterial drug resistance, since they pump out drugs and other harmful substances from the cytoplasm. These systems can be classified into single- and multi-component pumps. Single-component pumps transport their substrates across the cytoplasmic membrane. They contain members of the Major faciliator and ABC superfamilies (TC-system 8.A.1 and 3.A.1., respectively), which catalyze the efflux by proton motive force or ATP hydrolysis, respectively. Multi-component pumps are found in Gram-negative bacteria and are composed in some cases of a cytoplasmic pump of the major faciliator superfamily together with a membrane fusion protein and an outer membrane protein.

Another important group is the 'resistance-nodulation-cell-division' (RND) superfamily, which is also composed of a cytoplasmic-, membrane fusion and an outer membrane protein. These pumps transport their substrates across the inner and outer membrane and are driven by proton motif force (Fig. 3.10.4-2, TC-system 2.A.6 and 2.A.7). Two well known systems for single and multiple-component pumps are the H^+/tetracycline antiporter and the AcrBA-TolC ABC transporter in *E. coli*.

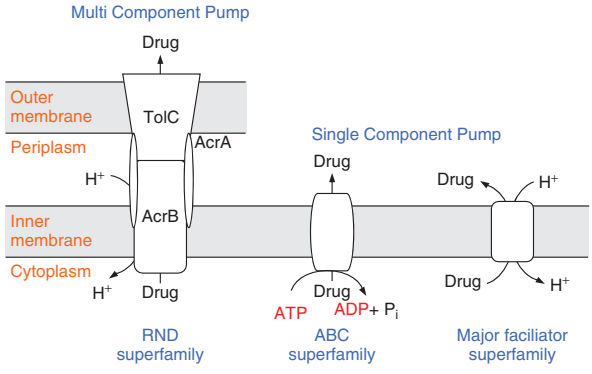


Figure 3.10.4-2. Bacterial Drug Efflux Pumps

Literature:

Kumar, A., Schweizer, H.P. Adv. Drug Deliv. Rev. 2005;57: 1486–1513.

Lengeler J.W., Drews, G., Schlegel, H.G. *Biology of the Prokaryotes,* San Deigo: Academic Press,. 1999.

Postma, P.W., *et al. Phosphoenolpyruvate: Carbohydrate Phosphotransferase Systems.* In: Neidhardt, F.C. (Ed). *Escherichia coli and Salmonella, cellular and molecular biology.* Washington D.C.: ASM Press, 1996.

Tchieu, J.H., *et al*. J. Mol. Microbiol. Biotechnol. 2001;3:329–346.

3.10.5 Bacterial Fermentations

Fermentations are anaerobic energy conserving-processes with substrate-level phosphorylation but no electron transport coupled phosphorylation. Most fermentations involve redox reactions of a single organic substrate (S), which serves both as electron donor and acceptor, since no external electron acceptor is involved (Fig. 3.10.5-1). Depending on the fermentative pathway, one or more organic products (P, reduced and oxidized) and ATP are generated (Fig. 3.10.5-1a and b). Moreover, simple fermentations involve only the conversion of the substrate to product, with the concomitant production of ATP not involving a redox reaction (Fig. 3.10.5-1c), or the use of two substrates, one of which is oxidized and the other reduced, while one or both can participate in the generation of ATP (Fig. 3.10.5-1d).

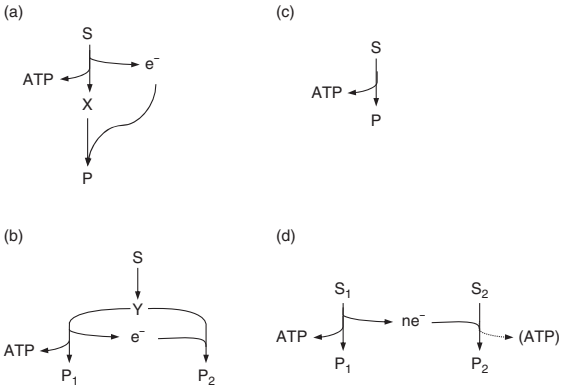


Figure 3.10.5-1. General Principle of Fermentation

The end products of fermentation processes are usually excreted and many fermentations have economic importance, for example in food production. Typical reduced fermentation products from glucose are ethanol, H_2 and propionate. Typical oxidized products are CO_2 and succinate (Fig. 3.10.5-2).

Fermentation allows ATP generation only by substrate-level phosphorylation, which results in a low ATP yield and, therefore, requires a high substrate throughput. For example, in homolactate fermentation, 2 mol ATP are obtained from 1 mol glucose, while aerobic respiration theoretically results in 33 mol ATP and about 29 mol ATP under laboratory conditions. Some specialized fermentations, such as diacetyl fermentation by *Leuconostoc spec*., have been found to involve ion transport mechanisms. These generate a membrane potential, which can be used for ATP synthesis by membrane bound ATP synthase (3.11.4.5).

3.10.5.1 Fermentation of Sugars (Figs. 3.10.5-2 and 3.10.5-3, Table 3.10.5-1)

Fermentation of sugars may start with the breakdown of glucose, which is initiated by reactions of the Embden-Meyerhof-pathway of glycolysis (3.1.1.1) or, in some bacteria, the Entner-Doudoroff pathway (3.1.5.3). A key intermediate in both pathways is pyruvate, since

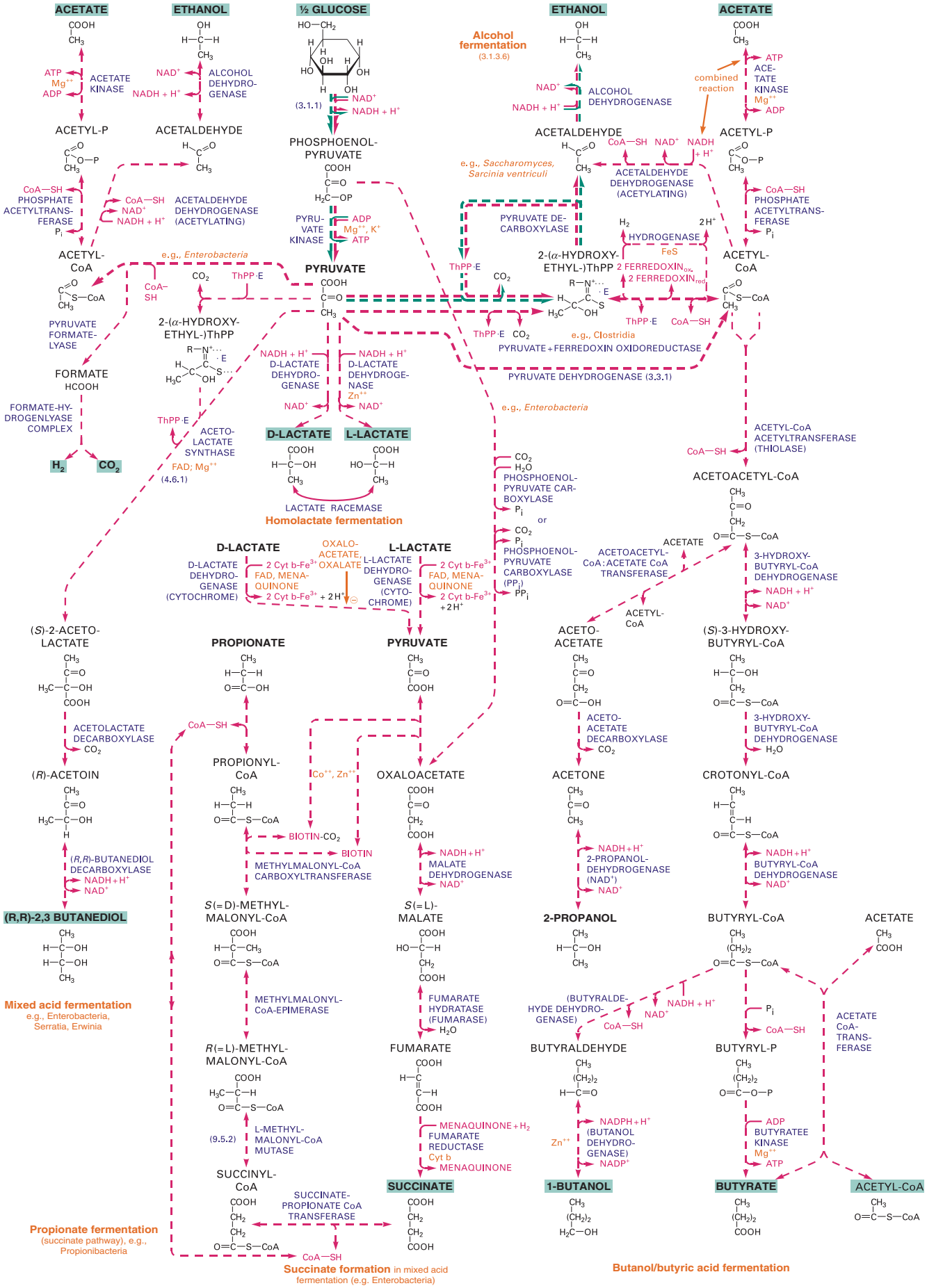


Figure 3.10.5-2. Glycolytic Fermentations and Lactate Converting Fermentations

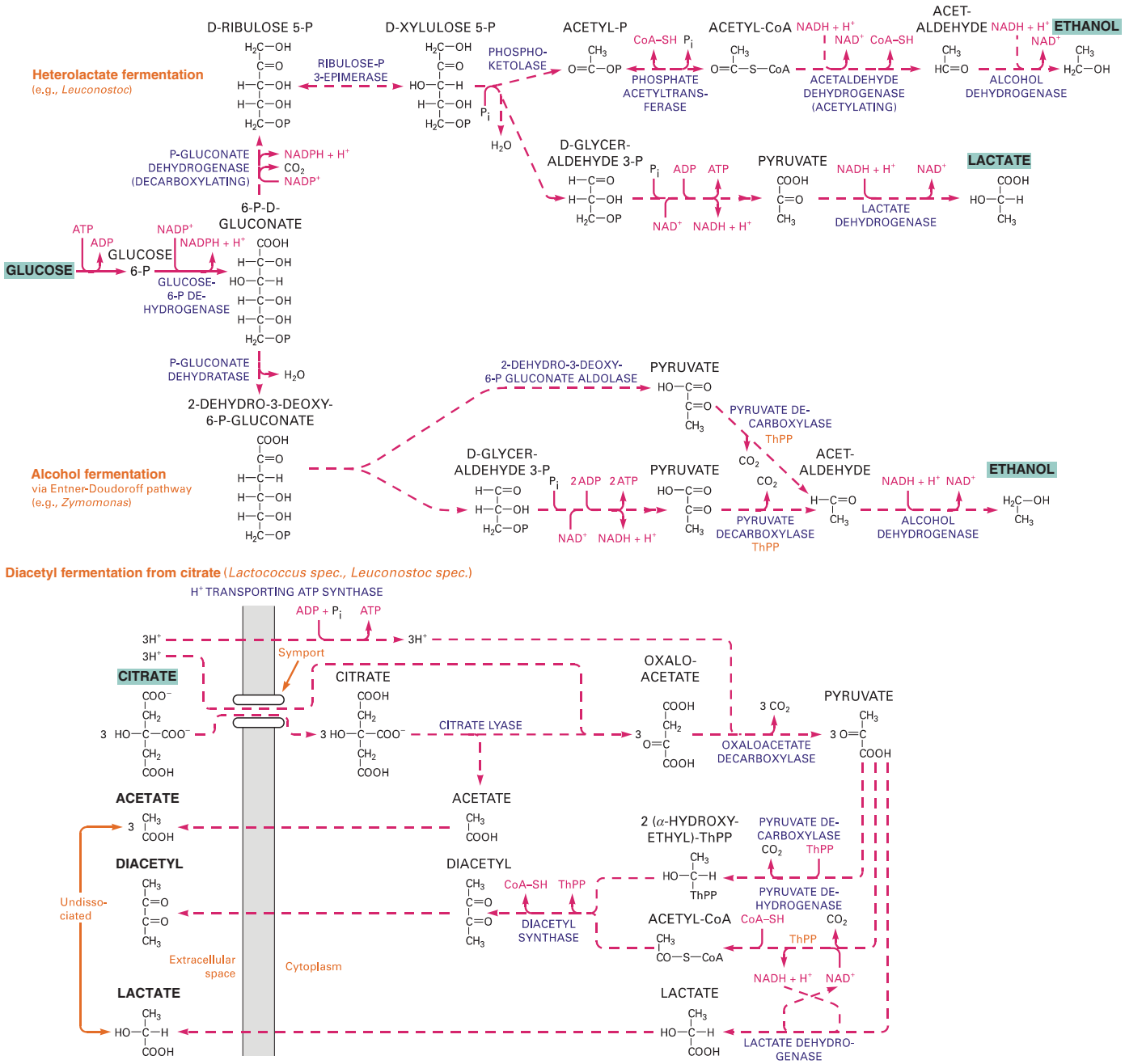


Figure 3.10.5-3. Non-Glycolytic Carbohydrate Fermentations

it is the branch point for several metabolic pathways. In some cases, for example, mixed acid fermentations of *Enterobacteria*, two or more phases of fermentation can be observed. In the first phase, mainly organic acids are produced and excreted, leading to acidification of the medium. When a critical pH is reached, additional enzyme systems leading to pathways with neutral end products are synthesized. Examples for these are the formate:hydrogen lyase complex of *E. coli* or acetolactate synthase of *Enterobacter aerogenes*.

It has to be emphasized that not all branches of the fermentation pathways are present in the organisms listed in Table 3.10.5-1. For example, *E. coli* does not produce acetoin or butanediol and various species of *Clostridia* will show different patterns of organic acid production.

3.10.5.2 Fermentations of Non-Sugar Carbon Sources and Especially of Nitrogenous Compounds (Fig. 3.10.5-4, Table 3.10.5-2)

These fermentations involve highly specialized pathways. With fermentation of amino acids, the critical step of amino group removal by anaerobic microorganisms is performed in many different ways, such as reductive cleavage of the C-N bond, C-C rearrangement, α,β-elimination of ammonia, shift of -NH_2 from α- to β-position, conversion to oxo-acids and consecutive reduction. The Stickland reaction, an example of amino acid fermentation, involves the fermentation of two different amino acids simultaneously. During this reaction, the first amino acid is oxidized while ATP is generated. The accumulating reduction equivalents are regenerated through a reductive deamination of the second amino acid. An example for this type of reaction is the oxidation of alanine and concurrent reduction of two molecules of glycine in *Clostridium sporogenes*.

Literature:

Madigan, M.T., Martinko, Brock, J.M. *Biology of Microorganisms*. Prentice Hall, 2005.

Lengeler J.W., Drews, G., Schlegel, H.G. *Biology of the Prokaryotes,* San Diego: Academic Press, 1999.

Konings, W.N., Lolkema, J.S., Poolman, B. Arch. Microbiol. 1995;164:235–242.

Gottschalk, G.: *Bacterial Metabolism.* 2nd Ed. Springer (1986).

Table 3.10.5-1. Fermentation of Sugars (Examples)
The stoichiometries given in both tables refer only to the energy metabolism. Due to anabolic reactions the actual yield of end products is lower.

Type		Gross Reaction, ATP yield		Pyruvate Converting Enzyme	Organisms (Examples)
Glycolytic Fermentations and Lactate Converting Fermentations					
Alcohol fermentation	#	glucose → 2 ethanol + 2 CO_2 (+2 ATP)		pyruvate decarboxylase	*Sarcina ventriculi, Saccharomyces cerevisiae*
Homolactate fermentation	#	glucose → 2 lactate (+2 ATP)		lactate dehydrogenase (L or D)	*Lactobacillus lactis, Lactococcus lactis, Enterococcus faecalis*
Butanol / butyric acid fermentation	#	glucose → butanoate, acetate, H_2, CO_2 (+ 2 ATP) or 2 glucose → butanol + acetone + 4 H_2 + 5 CO_2 (+4 ATP)	*	pyruvate: ferredoxin oxidoreductase	*Clostridium spec.*
Mixed acid fermentation	#	glucose → formate, ethanol, acetate, lactate, fumarate, CO_2, H_2	**	pyruvate formate-lyase (formate C-acyltransferase)	*Escherichia spec., Salmonella spec.*
Mixed acid fermentation	#	glucose → ethanol, acetate, lactate, CO_2, H_2, acetoin, 2,3-butanediol	**	pyruvate formate-lyase	*Enterobacter spec., Erwinia spec.*
Homoacetate fermentation	#	glucose → 3 acetate (+4 ATP)		pyruvate: ferredoxin oxidoreductase	*Clostridium thermoaceticum, Acetobacterium woodii*
Propionate ferment.: succinate pathway	#	3 lactate → 2 propionate + acetate + CO_2 (+3 ATP)	***	lactate dehydrogenase (cytochrome)	*Propionibacterium spec.*
acrylate pathway	#	3 lactate → 2 propionate + acetate + CO_2 (+1 ATP) via different intermediates		pyruvate: ferredoxin oxidoreductase	*Clostridium propionicum, Peptostreptococcus elsdenii*
Non-glycolytic Fermentations					
Alcohol fermentation via Entner-Doudoroff pathway	#	glucose → 2 ethanol + 2 CO_2 (+1 ATP)		pyruvate decarboxylase	*Zymomonas spec.*
Heterolactate fermentation	#	glucose → lactate + ethanol + CO_2 (+1 ATP)		lactate dehydrogenase	*Leuconostoc mesenteroides*
		3 fructose → lactate + acetate + 2 mannitol + CO_2 (+2ATP)		lactate dehydrogenase	*Lactobacillus brevis*
Diacetyl fermentation	#	3 citrate → 3 acetate + diacetyl + lactate		pyruvate decarboxylase, pyruvate dehydrogenase, lactate dehydrogenase	*Lactococcus spec., Leuconostoc spec.*

Table 3.10.5-2. Fermentation of Nitrogenous Compounds (Examples)

Compound Fermented		Gross Reaction	Remarks	Organisms (Examples)
Fermentation of Amino Acids				
Alanine fermentation		3 alanine → 2 propionate + acetate + CO_2 + 3 NH_3	Via acrylate pathway	*Clostridium propionicum*
Glycine fermentation		4 glycine → 3 acetate + 4 NH_3 + 2 CO_2		*Peptostreptococcus micros*
Glutamate fermentation	#	5 glutamate → 6 acetate + 2 butyrate + 5 CO_2 + H_2 + 5 NH_3	Via mesaconate, citramalate	*Clostridium tetanomorphum*
Glutamate fermentation		2 glutamate → butyrate + 3 acetate + 2 NH_3	Via 2-hydroxyglutarate, glutaconyl-CoA	*Acidaminococcus fermentans*
Cofermentation of pairs of amino acids	#	e.g., alanine + 2 glycine → 3 acetate + CO_2 + 3 NH_3	'Stickland reaction'	*Clostridium botulinum, C. sticklandii*
Arginine fermentation		arginine → ornithine + CO_2 + 2 NH_3	Non-redox fermentation	*Pseudomonas aeruginosa*
Fermentation of Heterocyclic Compounds				
Fermentation of urate		uric acid → acetate + 4 CO_2 + 4 NH_3	All heterocyclic compounds are degraded via xanthine	*Clostridium purinolyticum, C. acidiurici*

[#] Reaction sequences shown in Figures 3.10.5-2 to 3.10.5-4.
[*] The end products vary, depending on the fermentation conditions. Instead of acetone, 2-propanol can be formed.
[**] The ratio of the end products depends predominantly on the pH of the medium.
[***] Acetyl-CoA is formed in variable amounts by pyruvate decarboxylase, it is hydrolyzed to acetate.

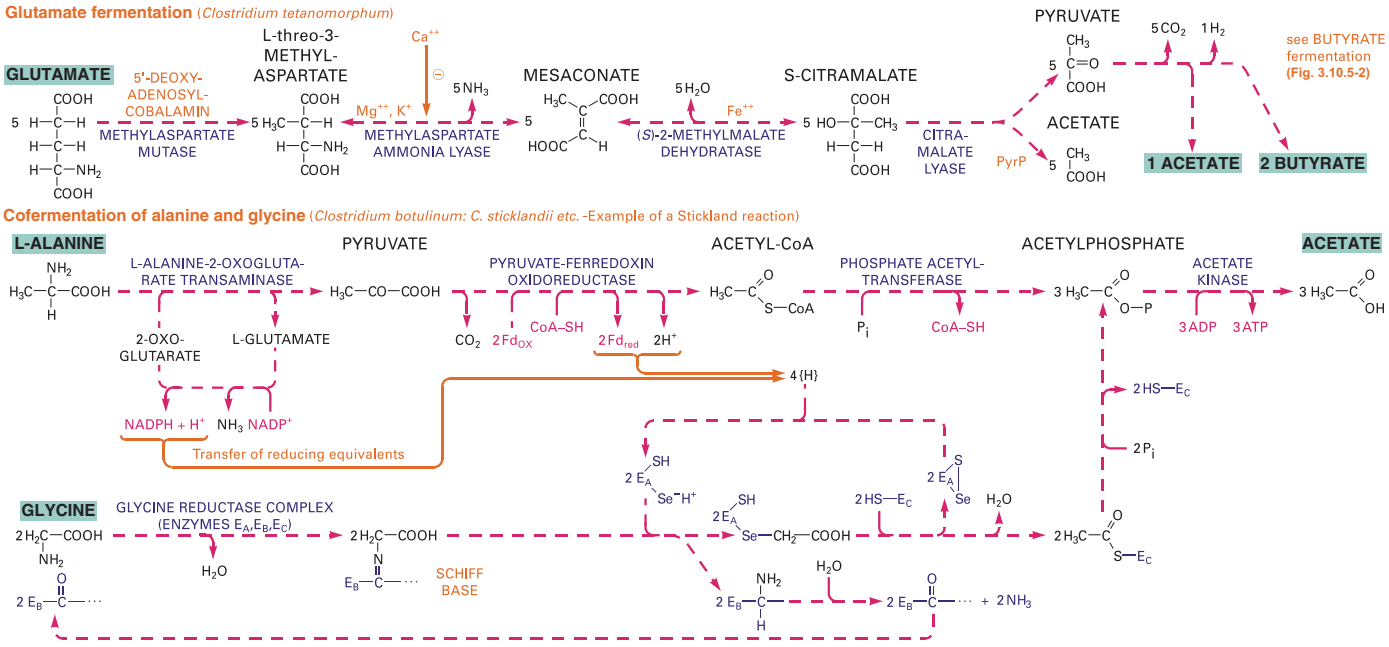


Figure 3.10.5-4. Fermentations of Nitrogeneous Compounds (For Other Bacterial Degradations of Amino Acids, see Chapter 3.2)

3.10.6 Anaerobic Respiration

Aerobic and anaerobic respirations are energy conserving processes which involve an electron transport chain. A substrate is oxidized and electrons are transferred to an electron transport chain, which transfers electrons to external terminal acceptors and in parallel uses energy to pump protons (or in some cases sodium ions) out of the cell. In this way, electron flow causes proton translocation out of the cells and generates an electrochemical gradient (proton motive force), which drives the H^+-dependent ATP synthase. While aerobic respiration uses oxygen as terminal acceptor, typical electron acceptors of anaerobic respiration are oxidized inorganic compounds as CO_2, sulfate, nitrate or Fe^{3+}. But also organic electron acceptors, such as fumarate are used. In most cases, the energy source (electron donor) are reduced organic compounds (carbohydrates, fatty acids, other organic acids), although several inorganic compounds (H_2, NH_4^+, CO) can also function as electron donors (anaerobic lithotrophy, 3.10.7).

3.10.6.1 Redox Reactions and Electron Transport (Fig. 3.10.6-1)

Electrons from the oxidation of the energy source (electron donor reactions) are fed into membrane-bound electron transport chains via various dehydrogenases (e.g., NADH dehydrogenase). Transfer of the electrons to the acceptor substrate is catalyzed by terminal reductases that take the role of the terminal oxidases in aerobic respiration. The system of *Wolinella succinogenes* is shown in Figure 3.11-1. Several components of the electron transport chain are involved in translocation of protons, as the cytochrome bc_1 complex. Note that Figure 3.10.6-1 compiles various anaerobic respirations systems which are not found together in a single organism.

Energy yield: The theoretical energy yield depends on the difference of the redox potentials of the electron donor and the electron acceptor (Table 3.10.6-1). Since almost all electron acceptors have redox potentials lower than that of O_2/H_2O, anaerobic electron transport is less energy efficient than aerobic respiration.

Table 3.10.6-1. Redox Potentials

	Redox Couple	E_0 [mV]
Typical electron donors	$2H^+/H_2$	−410
	pyruvate/lactate	−197
	fumarate/succinate	+33
Typical electron acceptors	CO_2 / CH_4	−244
	HSO_3^- / HS^-	−110
	APS / HSO_3^- + AMP	−60
	fumarate/succinate	+33
	NO_2^-/NH_3	+340
	NO_3^-/NO_2^-	+430
	NO_3^- / N_2	+740
	N_2O / N_2	+1355
For comparison: aerobic	½ O_2 / H_2O	+816

In anaerobic respiration, the most important electron acceptor conversions are:

- Reduction of nitrate and nitrite, denitrification (3.10.6.2).
- CO_2 reduction to methane by methanogenic archaea (3.10.6.3).
- Homoacetogenic bacteria reduce CO_2 to acetate (carbonate respiration, 3.10.6.4).
- Reduction of ferric ions and manganese (II) ions (3.10.6.5).
- Sulfate and sulfur respiration is carried out by a variety of unrelated, strictly anaerobic bacteria and archea. Both acceptors are reduced to H_2S (sulfidogenesis). Sulfate needs to be activated before reduction; adenosine-5′-phosphosulfate (APS) is the actual electron acceptor. Sulfite reductases contain siroheme, an iron-tetrahydroporphyrin (3.7.5.3). Sulfidogenic bacteria often use H_2 as energy source.
- Other reactions include the reduction of organic electron acceptors glycine, dimethylsulfoxide (DMSO), trimethylamine-N-oxide (TMAO) or fumarate and inorganic substances as ClO_4^-.

Regulation: Facultative anaerobes (*e.g. E. coli*) can switch from oxygen to the use of alternative electron acceptors depending on their availability. Since the energy yield depends on the redox potential difference, these organisms utilize electron acceptors in the order O_2 > NO_3^- > fumarate. The synthesis of these alternative electron transport systems is transcriptionally regulated. In general, the energetically

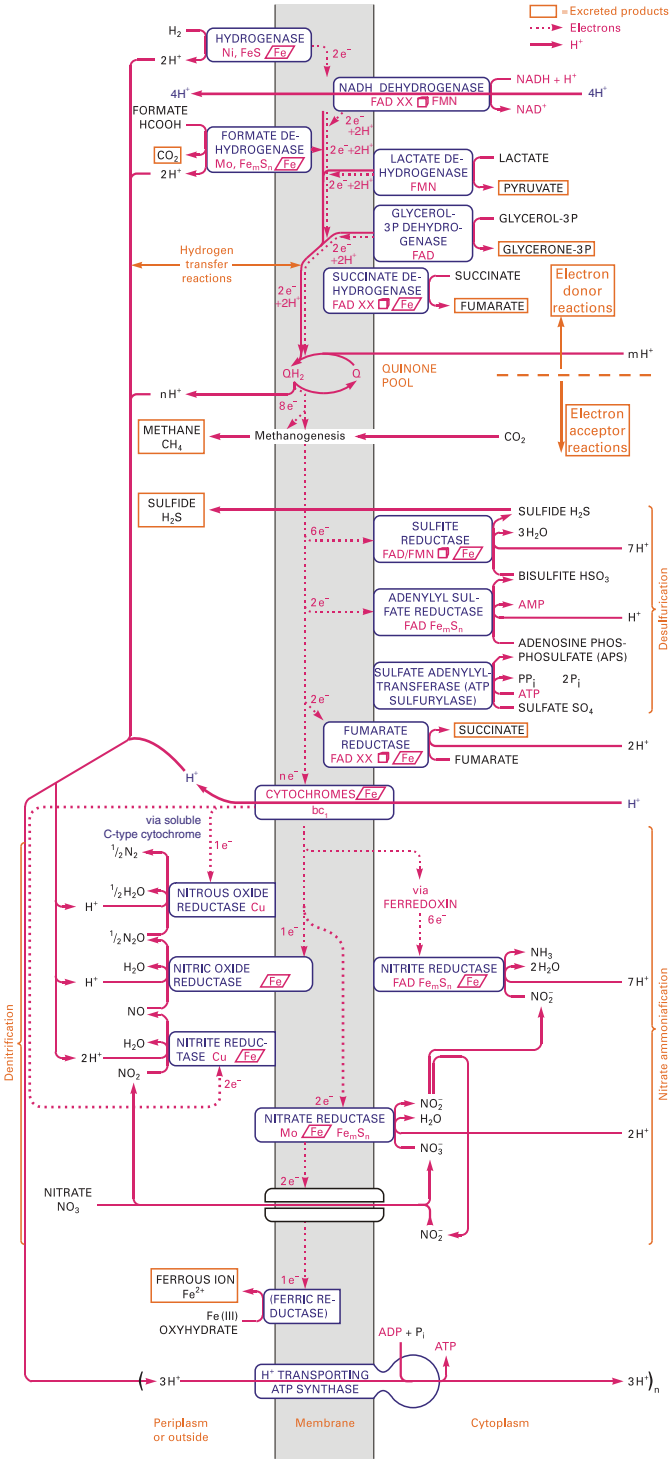


Figure 3.10.6-1. Systems of Anaerobic Respiration (in various organisms)

more favorable electron acceptor represses the expression of the genes necessary for growth with less favorable ones.

3.10.6.2 Denitrification, Ammonification (Fig. 3.10.6-1)

Inorganic nitrogen compounds are reduced by diverse facultative anaerobic bacteria. In denitrifying bacteria, nitrate is reduced to N_2 in four steps via the following intermediates: $NO_3^- \rightarrow NO_2^- \rightarrow NO \rightarrow N_2O \rightarrow N_2$. Each of the four reductases accepts electrons of an electron transport chain, either from quinol or via c-type cytochromes. Beside nitrate, nitrite or nitrous oxide (N_2O) are also used as electron acceptors. Denitrification is the only microbial process generating N_2. It is important during nitrate removal in sewage plants and a detrimental process in agriculture, where nitrate fertilizer is lost.

Nitrate ammonification is found in strict and facultative anaerobes: $NO_3^- \rightarrow NO_2^- \rightarrow NH_4^+$. This reaction is coupled to energy conservation and is not inhibited by NH_4^+. (Dissimilatory nitrate reduction in contrast to assimilatory nitrate reduction, which is inhibited by NH_4^+).

3.10.6.3 Methanogenesis (Fig. 3.10.6-2)

Methanogenesis is the reduction of CO_2 or small organic molecules (e.g., methanol, methylamine, formate or acetate) to methane (CH_4) utilizing H_2 as electron donor. It is only found in strictly anaerobic methanogenic archaea (methanogens). These reactions employ unique coenzymes which are found in methanogens and a limited number of bacteria as *Methylobacterium extorquens* or *Streptomyces*. It is estimated that methanogens produce 1 Giga ton (10^9 t) of methane per year and, therefore, play an important role in the global carbon cycle.

Methanogenic archea can be further divided into two groups, those with or without cytochromes. While energy metabolism of methanogens without cytochromes is still not fully understood, methanogens with cytochromes use the following pathway: CO_2 (oxidation state +4) reduction to methane (CH_4, oxidation state –4). Throughout the pathway, the carbon atom remains covalently bound to special one-carbon carriers.

In the first step, CO_2 is reduced to a formyl group (+2) bound to methanofuran, the electron donor is a still unidentified ferredoxin. The formyl group is transferred to methanopterin (5,6,7,8-tetrahydromethonopterin (THMPT, Fig. 3.10.6-2). The product is dehydrated to yield the methenyl group, which is then reduced to the methylene level (0). The 5-deazaflavin derivative coenzyme F_{420} (3.7.3.2) serves as electron carrier from the hydrogenase, utilizing H_2 as electron donor. Further reduction (again involving coenzyme F_{420} and a hydrogenase) yields methyl-H_4-MPT (–2) from which the methyl group is transferred to coenzyme M (2-mercaptoethanesulfonic acid). This reaction is catalyzed by the membrane bound methyl-H_4MPT–coenzyme M methyltransferase, which contains a corrinoid (cobamide), with the concomitant translocation of two Na^+ ions out of the cytoplasm. The resulting methyl-CoM is reduced to methane; here, the nickel porphinoid coenzyme F_{430} (3.7.5.3) serves as cofactor. The actual reductant is 7-mercaptoheptanoylthreonine phosphate (HS-HTP, coenzyme B) which is oxidized and forms a heterodisulfide with coenzyme M. The reduced forms of CoM and HTP are regenerated by reduction of the heterodisulfide via disulfide reductase. This heterodisulfide reductase receives reducing equivalents via an electron transport chain. This chain consists of cytochrome b, methanophenazine (an electron carrier used by archaea instead of quinone) and a hydrogenase which reduces methanophenazine by oxidation of H_2. Moreover, protons are translocated simultaneously into the periplasm, generating a proton motive force (free energy sequestered by an electrochemical gradient, analogously to the electromotive force). Beside this proton gradient, a sodium gradient is generated.

Methanogenesis is an unusual anaerobic respiration, since the electron acceptor, CO_2 is not reduced directly by a membrane bound terminal oxidase, but in a series of reactions distinct from the electron transport events. In some orders of methanogens, redox enzymes contain selenocysteine, while others use pyrrolysine-containing methyltransferases (Sec and Pyl, respectively, 4.1.3.4, Fig. 1.3-2).

Energy yield and conservation: Since E_0 (CO_2/CH_4) is quite negative (–244 mV), a strong reductant is needed, e.g., hydrogen [E_0 ($2H^+/H_2$): –410 mV].

$$CO_2 + 4\,H_2 = CH_4 + 2\,H_2O \qquad \Delta G'_0 = -131\ \text{kJ/Mol.}$$

In methanogenic habitats, however, H_2 concentration is very low (1-10 Pa) and $\Delta G'_0$ decreases to –40 to –17 kJ per mol CO_2. Thus, the decrease in free energy allows only the generation of < 1 mol ATP per mol CH_4 produced.

3.10.6.4 Acetogenesis (Fig. 3.10.6-3)

Many acetogenic bacteria are Gram-positive and belong to the genus *Clostridium* or *Acetobacterium* (not to be mixed up with the Gram-negative *Acetobacter*).

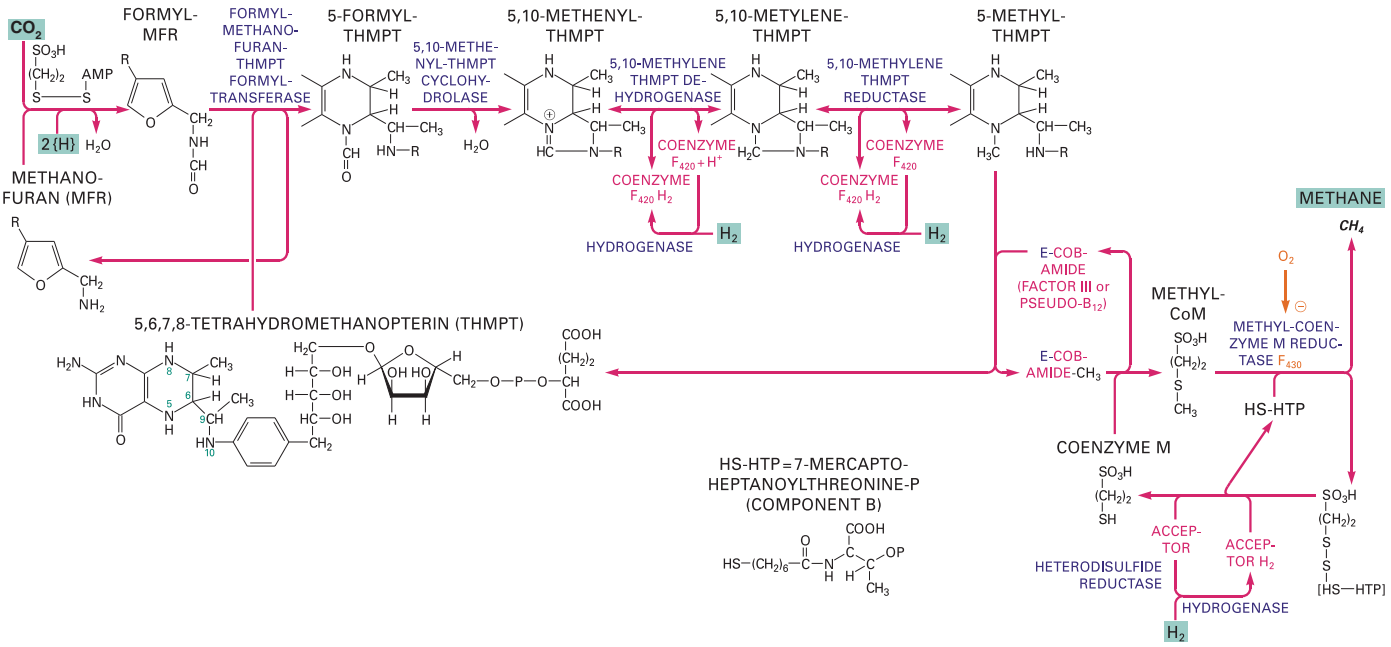


Figure 3.10.6-2. Methanogenesis (e.g., in *Methanothermobacter thermoautotrophicus*)

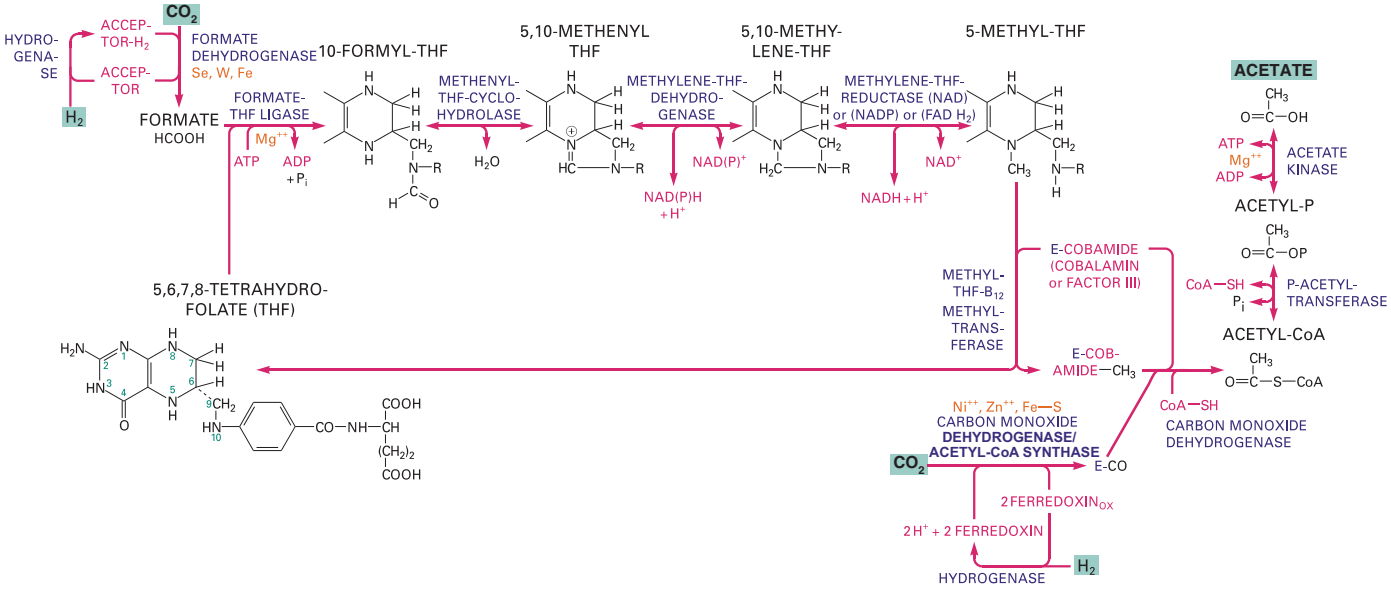


Figure 3.10.6-3. Acetate formation (e.g., in *Moorella thermoacetica*)

Acetogenesis shares some conceptual similarities to methanogenesis, however, different cofactors and enzymes are used. During acetogenesis, CO_2 is reduced with H_2 as an electron donor, leading to the formation of acetate. It can take place during anaerobic bacterial growth on CO_2 as sole carbon source (anaerobic lithotrophy) as well as part of homoacetate fermentation (Table 3.10.5-1). E.g., *Moorella thermoacetica* (formerly *Clostridium thermoaceticum*) uses CO_2 and H_2 formed by the pyruvate:ferredoxin oxidoreductase and hydrogenase reactions for acetate synthesis. Some of the enzymes involved are different in various organisms. E.g., formate dehydrogenases or methylene-tetrahydrofolate reductases may use ferredoxin, cytochromes *b* or NADH as electron donors.

The key enzyme in acetogenesis is the carbon monoxide dehydrogenase/acetyl-CoA synthase which reduces CO_2 to CO and catalyzes the addition of a carboxyl group to the CH_3 moiety. It contains Ni, Zn and Fe-S centers.

Acetogenesis requires one ATP (formate-THF ligase) and generates one ATP via substrate level phosphorylation (acetate kinase). Energy is generated via anaerobic respiration involving hydrogenases and an electron transport chain generating a chemiosmotic gradient (sodium or proton motive force, depending on organism).

3.10.6.5 Reduction of Ferric Ions and Manganese (IV) Ions (Fig. 3.10.6-1)

Metals, such as ferric iron (Fe^{3+}) are used as electron acceptors in anaerobic respirations by bacteria from the genus *Shewanella* and *Geobacter*. *Shewanella oneidensis* has been shown to reduce in addition manganese (IV) ion (Mn^{4+}). Remarkably, bacteria like *S. oneidensis* are capable of transferring electrons via multi-heme c-type cytochromes across the outer membrane to solid metal oxide substrates. This extracellular respiration raises interest in respect of microbial fuel cell development.

Literature:

Drake, H.A. *Acetogenensis*. Chapman & Hall, 1994.

Fredrickson, J.K., *et al*. Nature Rev. Microbiol. 2008;6:592–603.

Gralnick, J.A., Newman, D.K., Mol. Microbiol. 2007;65:1–11.

Ragsdale, S.W., Pierce, E. Biochim Biophys Acta 2008;1784: 1873–1898.

Rother, M., Krzycki, J.A. Archaea vol 2010, Article ID 453642 (2010) doi:10.1155/2010/453642.

Schäfer, G., *et al*. Microbiol. Mol. Biol. Rev. 1999;63:570–620.

Thauer , R.K., *et al*. Nature Reviews Microbiology 2008;6:579–591.

Zumft., W.G. Microbiol Mol Biol Rev. 1997;61:533–616.

3.10.7 Chemolithotrophy

Chemolithotrophic organisms (bacteria and archaea) derive energy from the oxidation of reduced inorganic compounds, e.g., sulfide, hydrogen, ammonia or Fe^{2+}. Often these compounds are produced by chemoorganotrophs, which use organic molecules such as hexoses. Chemolithotrophic microorganisms often obtain carbon from CO_2.

3.10.7.1 Redox Reactions and Electron Transport (Fig. 3.10.7-1)

Electrons are released during the oxidation of the donor substrate and are fed into a membrane-associated respiratory chain. This occurs at the redox level corresponding to the redox potential of the substrate/product couple (Table 3.10.7-1, columns 1…3, Fig. 3.10.7-2). Finally, the electrons are transferred to a terminal electron acceptor such as oxygen, nitrate or sulfate, which is then reduced.

Example for the calculation of the theoretical energy yield (Definitions see 1.5.2): Oxidation of nitrite to nitrate; the electrons are taken up by oxygen (2-electron reaction):

$E'_0\,[NO_3^-/NO_2^-] = +430$ mV $E'_0\,[O_2/H_2O] = +816$ mV

$\Delta G'_0 = n * F * \Delta E'_0$

$\Delta G'_0 = 2 * 0{,}0965$ [kJ * mV^{-1} * mol^{-1}] * (430 – 816) [mV]

$\Delta G'_0 = -75$ [kJ * mol^{-1}] (see Table 3.10.7-1).

3.10.7.2 Energy Conservation and Reductive Power

Energy conservation takes place via chemiosmotic coupling (oxidative phosphorylation, 3.11). At various reaction steps (oxidation of nutrients) external excess H^+ is generated. Additionally, protons are translocated out of the cell, e.g., by the terminal oxidase. This builds up an electrochemical gradient, which drives the H^+-dependent ATP synthase.

In many cases, the redox potential of the substrate couple is more positive than that of NAD^+/NADH (–340 mV). Therefore, direct reduction of NAD^+ (or $NADP^+$) in order to supply reducing equivalents for the anabolic reactions is not possible. Rather, reversed electron pumping (to a more negative redox potential) is necessary in order to produce NADH (Fig. 3.10.7-3). This reaction is also driven by the electrochemical gradient, i.e. protons flow back into the cell and thus partially reverse the H^+ outflow by the oxidation reactions.

$$n\,H^+_{out} + NAD^+ + 2\,e^- = (n-1)\,H^+_{in} + NADH.$$

Most chemolithotrophic bacteria derive the carbon for biosynthesis from CO_2 (autotrophy). Fixation of CO_2 takes place via the ribulose bisphosphate pathway (see 'Calvin cycle', 3.12.2). The key enzyme

of this pathway is ribulosebisphosphate carboxylase (rubisco).This enzyme is found in facultative anaerobic purple bacteria, chloroplasts, cyanobacteria and in most of the chemolithotrophic bacteria.

Microorganisms which obtain energy by oxidation of an inorganic compound but require an organic carbon source are called mixotroph.

3.10.7.3 Nitrification

Ammonia and nitrite serve both as electron donors. The oxidation of these compounds to nitrite and nitrate, respectively, is a key reaction in the global nitrogen cycle. The oxidation is carried out in specialized microorganisms (e.g., *Nitrosomonas*: ammonia oxdiation; *Nitrospira*: nitrite oxidation) and uses oxygen as terminal electron acceptor.

Ammonia monooxygenase catalyzes the first step in ammonia oxidation and generates hydroxylamine. This reaction is driven by a high positive redox potential and requires electrons, which are supplied by the second reaction, the consecutive hydroxylamine oxidoreductase reaction (Table 3.10.7-1, Fig. 3.10.7-4).

Nitrifying bacteria oxidize nitrite to nitrate by the membrane-bound nitrite/nitrate-oxidoreductase.

3.10.7.4 Anammox

Anammox (anoxic ammonia oxidation) is the anaerobic oxidation of ammonia (electron donor) using nitrite as electron acceptor. In contrast to nitrification, which is performed by strict aerobic bacteria, anammox is observed in unusual obligate anaerobic bacteria such as *Brocadia anamoxidans*. These bacteria contain a membrane-enclosed compartment, the annamoxosome where the oxidation of ammonia takes place, because of the highly toxic intermediate hydrazine (Fig. 3.10.7-5).

$$NH_4^+ + NO_2^- = N_2 + 2\,H_2O. \qquad \Delta G_0' = -357\ \text{kJ}.$$

3.10.7.5 Sulfur, Hydrogen Sulfide and Thiosulfate Oxidation

Oxidation of reduced sulfur compounds may either proceed directly to sulfate involving sulfide- and sulfite dehydrogenases and sulfur dioxygenases (Fig. 3.10.7-1). Another pathway oxidizes sulfite (HSO_3^-) with simultaneous adenylylation yielding adenosine 5′ phosphosulfate as an intermediate (APS, Fig. 3.10.7-6). APS is then phosphorylytically cleaved to sulfate and ADP, which disproportionates to ATP and AMP. This way, half of the original AMP is converted to ATP. In addition, electrons are fed into an electron transport chain (Fig. 3.10.7-6).

3.10.7.6 Iron Oxidation

Reduced metal ions such as Fe^{++} also serve as electron donors for chemolithotrophs under aerobic conditions (compare anaerobic iron

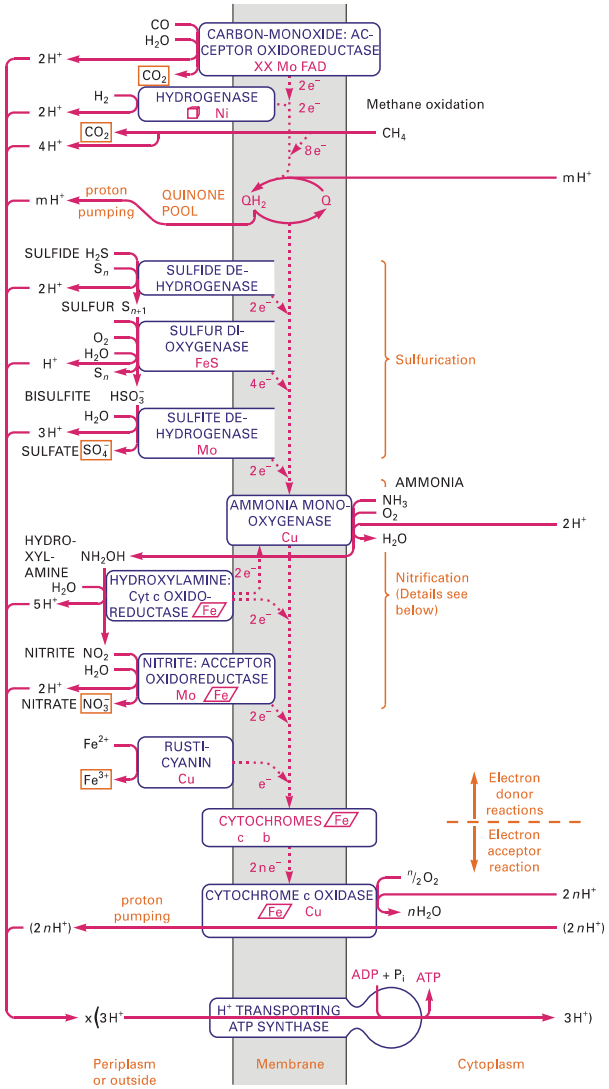


Figure 3.10.7-1. General Chemolithotrophic Reactions (in various organisms)

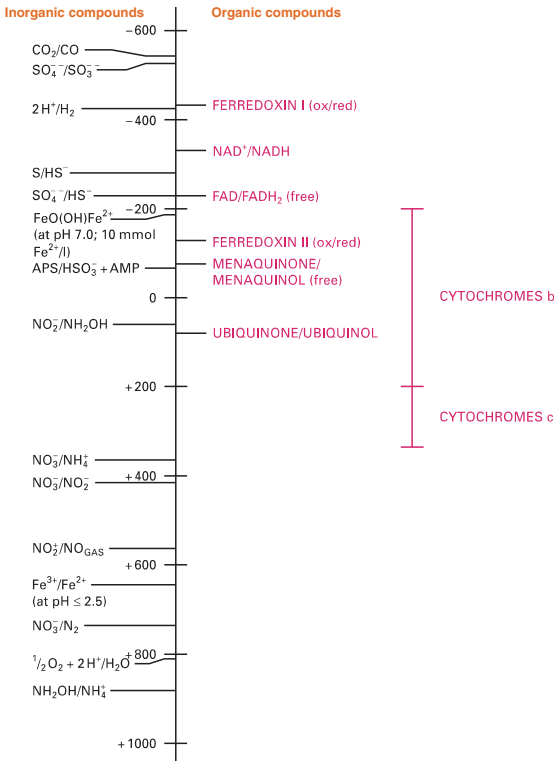


Figure 3.10.7-2. Redox Potentials E_0 [mV, 25 °C, pH = 7.0] of Substrate Couples

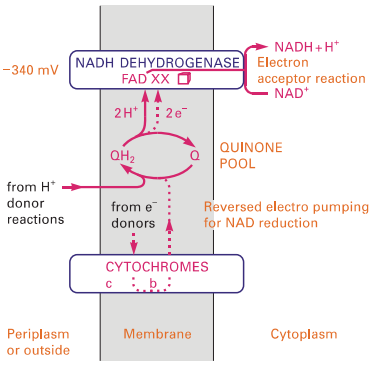


Figure 3.10.7-3. Reversed Electron Flow Driven by the Electrochemical Gradient

Table 3.10.7-1. Redox Reactions in Chemolithotrophic Bacteria

Reductant	Redox Couple[1]	E'_0 [mV]	Reactions	G_0 [kJ]*mol–1][2]	Enzymes; Cofactors	Organisms
Carbon monoxide	CO_2/CO O_2 (0.21 atm)/H_2O	–540 +816	$CO + H_2O \rightarrow CO_2 + 2\,H^+ + 2\,e^-$ $½\,O_2 + 2\,e^- + 2H^+ \rightarrow H_2O$		carbon monoxide dehydrogenase	'Carboxidobacteria', e.g. *Pseudomonas*
			$\Sigma = CO + ½\,O_2 \rightarrow CO_2$	–261		
Hydrogen	$2\,H^+/H_2$	–410	$H_2 \rightarrow 2\,H^+ + 2\,e^-$ $½\,O_2 + 2\,e^- + 2\,H^+ \rightarrow H_2O$		hydrogenase, [NiFe] or [Fe]	'Knallgas' bacteria
			$\Sigma = H_2 + ½\,O_2 \rightarrow H_2O$	–237		
Sulfide	S_0/HS^-	–260	$HS^- \rightarrow S_0 + H^+ + 2\,e^-$ $½\,O_2 + 2\,e^- + 2\,H^+ \rightarrow H_2O$		'sulfide oxidase'	*Thiobacillus; Beggiatoa; Wolinella succinogenes* (uses fumarate as el.acc.)
			$\Sigma = HS^- + ½\,O_2 + H^+ \rightarrow S_0 + H_2O$	–207		
	HSO_3^-/HS^-	–110	$HS^- + 3\,H_2O \rightarrow HSO_3^- + 6\,H^+ + 6\,e^-$ $½\,O_2 + 2\,e^- + 2\,H^+ \rightarrow H_2O$		sulfite reductase; siroheme, FeS	*Thiobacillus; Sulfolobus*
			$\Sigma = HS^- + 1½\,O_2 \rightarrow HSO_3^-$	–536		
Sulfur	HSO_3^-/S_0	–45	$S_n + O_2 + H_2O \rightarrow S_{n-1} + HSO_3^- + H^+$	–332	'sulfur dioxygenase'; FeS	*Thiobacillus*
Sulfite	SO_4^{2-}/HSO_3^-	–520	$HSO_3^- + H_2O \rightarrow SO_4^{2-} + 3\,H^+ + 2\,e^-$ $½\,O_2 + 2\,e^- + 2H^+ \rightarrow H_2O$		sulfite: cytochrome c oxidoreductase; heme-Fe, Mo^{++}	*Thiobacillus(T. denitrificans* uses NO_3^- as electron acceptor anaerobically)
			$\Sigma = HSO_3^- + ½\,O_2 \rightarrow SO_4^{2-} + H^+$	–258		
	APS/HSO_3^-	–60	a) $HSO_3^- + AMP \rightarrow APS + 2\,e^-$ b) $APS + P_i \rightarrow SO_4^{2-} + ADP + H^+$ c) $ADP \rightarrow ½\,AMP + ½\,ATP$ $½\,O_2 + 2\,e^- + 2\,H^+ \rightarrow H_2O$		a) adenosine phosphosulfate reductase; FAD, FeS, b) sulfate adenylyl transferase; Mg^{++}, c) adenylate kinase, Mg^{++}	*Thiobacillus* *Thiobacillus*
			$\Sigma = HSO_3^- + ½\,O_2 + P_i + ½\,AMP + H^+ \rightarrow SO_4^{2-} + ½\,ATP + H_2O$	~ –227		
Ammonia	NH_2OH/NH_3 NO_2^-/NH_2OH	+900 +60	a) $NH_3 + O_2 + 2\,H^+ + 2\,e^- \rightarrow NH_2OH + H_2O$ b) $NH_2OH + H_2O \rightarrow NO_2^- + 5\,H^+ + 4\,e^-$		a) NH_3 monooxygenase; Cu, b) hydroxylamine: cytochrome c_{554} oxidoreductase, heme-Fe, Mo^{++}	nitrosofying bacteria, e.g. *Nitrosomonas*
		+340	$½\,O_2 + 2\,e^- + 2\,H^+ \rightarrow H_2O$			
			$\Sigma = NH_3 + 1½\,O_2 \rightarrow NO_2^- + H_2O + H^+$			
	NO_2^-/NH_3			–276		
Nitrite	NO_3^-/NO_2^-	+430	$NO_2^- + H_2O \rightarrow NO_3^- + 2\,H^+ + 2\,e^-$ $½\,O_2 + 2\,e^- + 2\,H^+ \rightarrow H_2O$		nitrite : cytochrome c_{550} oxidoreductase; heme-Fe, Mo^{++}	nitrifying bacteria: e.g., *Nitrobacter*
			$\Sigma = NO_2^- + ½\,O_2 \rightarrow NO_3^-$	–75		
Fe^{++} at low pH (ca.2.0)	Fe^{+++}/Fe^{++} O_2 (0.21 atm)/H_2O	+770 +1100	$Fe^{++} \rightarrow Fe^{+++}$ $¼\,O_2 + e^- + H^+ \rightarrow ½\,H_2O$		rusticyanin; Cu	*Thiobacillusferrooxidans, Sulfolobus*
			$\Sigma = Fe^{++} + ¼\,O_2 + H^+ \rightarrow Fe^{3+} + ½\,H_2O$	~ –32		
Fe^{++} at ca. neutral pH	$Fe(OH)_3$ (sat.)/ Fe^{2} (10 µmol/l)	– 150	$Fe^{++} + ¼\,O_2 + 2½\,H_2O \rightarrow Fe(OH)_3 + 2\,H^+$	?	?	*Gallionella*

[1] These reactions proceed from reduced to oxidized state.
[2] Calculated from E_0 or from free energies of formation per mole of substrate oxidized.

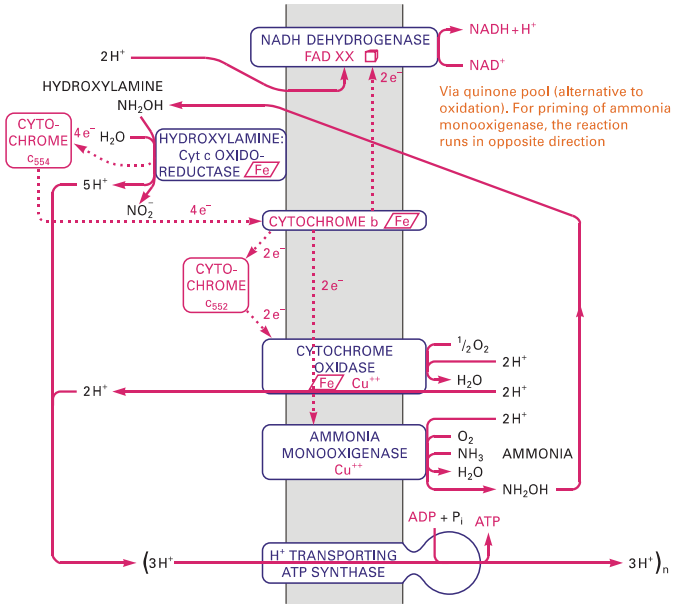


Figure 3.10.7-4. Details of Electron Flow and Energy Conservation During Oxidation of Ammonia in *Nitrosomonas*

reduction, 3.10.6.4.). However, under aerobic conditions Fe^{++} occurs often as poorly soluble metal salts. At low pH values solubility of these ions increases, moreover, Fe^{++} is chemically stable at low pH in the presence of oxygen. The free energy of the reaction increases at pH 2 compared to neutral pH. This may be the reason why many Fe^{++} oxidizing chemolithotrophs prefer highly acidic habitats. These microorganisms are even able to generate an acidic environment by oxidizing the inorganic sulfur in metal sulfides such as pyrite (FeS_2) to SO_4^- in addition to iron oxidation.

Literature:

Jetten, M.S. *et al.* Biochem Soc Trans. 2005;33(1):119–123.

Fuchs, G. *Allgemeine Mikrobiologie.* Thieme, 2006.

Lengeler J.W., Drews, G., Schlegel, H.G. *Biology of the Prokaryotes,* 1999.

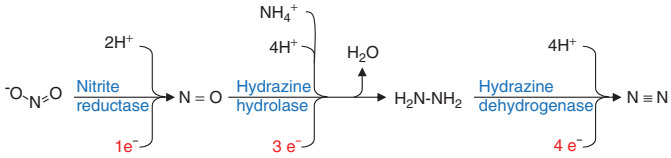


Figure 3.10.7-5. Anammox Reaction

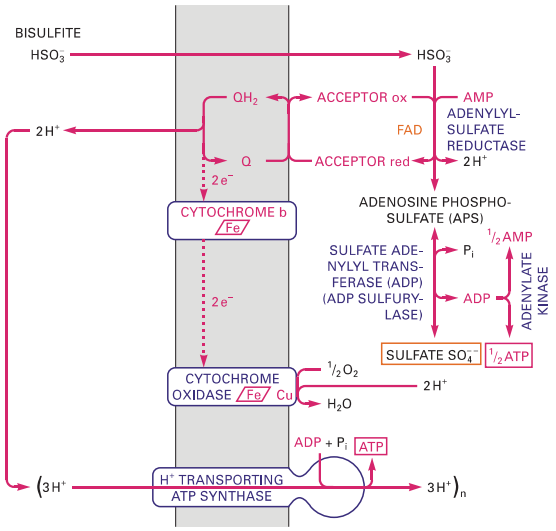


Figure 3.10.7-6. Electron Flow and Substrate Level Phosphorylation During Sulfite Oxidation by the 'APS Reductase Pathway' in *Thiobacillus*

3.10.8 Quinoenzymes, Alkane and Methane Oxidation

3.10.8.1 Alkane Oxidation (Fig. 3.10.8-1)

The capacity for utilization of aliphatic hydrocarbons is quite widespread among bacteria and fungi. n-Alkanes with chain lengths of C_{10} … C_{18} are among the most frequently occurring substrates, they are rapidly oxidized. Since these substances are insoluble in water, uptake and enzymatic attack are difficult. Many bacteria synthesize specialized lipids (e.g., trehalo- and rhamnolipids) as constituents of their cell walls, in which these hydrocarbons 'dissolve' and in which they are transported to the membrane, where they are oxidized at the cytoplasmic side. The fatty acids formed this way enter the usual degradation pathway via β-oxidation (3.4.1.5), supplying acetyl-CoA as source of cell constituents as well as NADH for reduction reactions.

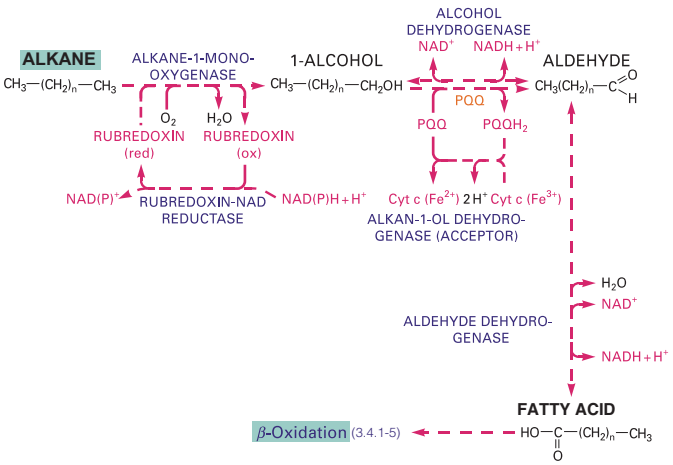


Figure 3.10.8-1. Alkane Oxidation

3.10.8.2 Quinoenzymes (Fig. 3.10.8-2)

The redox carrier pyrrolo-quinoline quinone (PQQ) is the coenzyme of methanol dehydrogenase in methylotrophs and other bacterial alcohol- and glucose dehydrogenases, e.g. the alcohol-dehydrogenase in acetic acid bacteria. The biosynthesis of PQQ is still not completely understood, it starts with tyrosine and glutamate. In contrast to some earlier erroneous findings, it is now accepted that PQQ is not covalently bound to proteins and is produced by Gram-negative bacteria.

Meanwhile, three additional quinoid cofactors have been found, which are all covalently linked to the respective enzymes: Tryptophantryptophylquinone (TTQ) in bacterial methylamine oxidase (formed by posttranscriptional condensation of two tryptophan molecules), trihydroxyphenylalanine quinone (TPQ, topa quinone) in mammalian serum amine oxidase, pea and bacterial amine oxidases (originating from tyrosine), lysinetyrosylquinone (LTQ) in mammalian lysyl oxidase

3.10.8.3 Oxidation of C_1 Compounds (Fig. 3.10.8-3)

Methylotrophs are able to grow on C_1 compounds (compounds lacking a C–C bond, e.g., methanol, methylamine, dimethylamine). Methylotrophs able to grow on methane are called methanotrophs, which are obligate C_1 utilizers (obligate methylotrophs). A number of interesting enzymes are involved in these reactions, e.g., methane monooxidase and PQQ-dependent methanol dehydrogenase. A critical step in oxidation of C_1 compounds is the oxidation of formaldehyde to formate. Here two unusual pathways exist in the same organism (e.g., *Methylobacterium extorquens*): the tetrahydrofolate (THF, 3.7.6.1) and the tetrahydromethanopterin (THMPT, 3.7.6.5) pathway. Interestingly, the cofactor tetrahydromethanopterin was initially identified in methanogenic archaea. Apparently, archaeal genes encoding the corresponding enzymes were transferred to the beta-proteobacterium *M. extorquens* via lateral gene transfer. At least in *M. extorquens*, an NAD-dependent formaldehyde dehydrogenases seem to play no significant role in formaldehyde oxidation.

3.10.8.4 Assimilation of Carbon in Methylotrophs

During growth of methylotrophs on C_1 compounds carbon is assimilated as formaldehyde or CO_2 via several pathways: formaldehyde (as methylene tetrahydrofolate) and CO_2 are assimilated via the serine

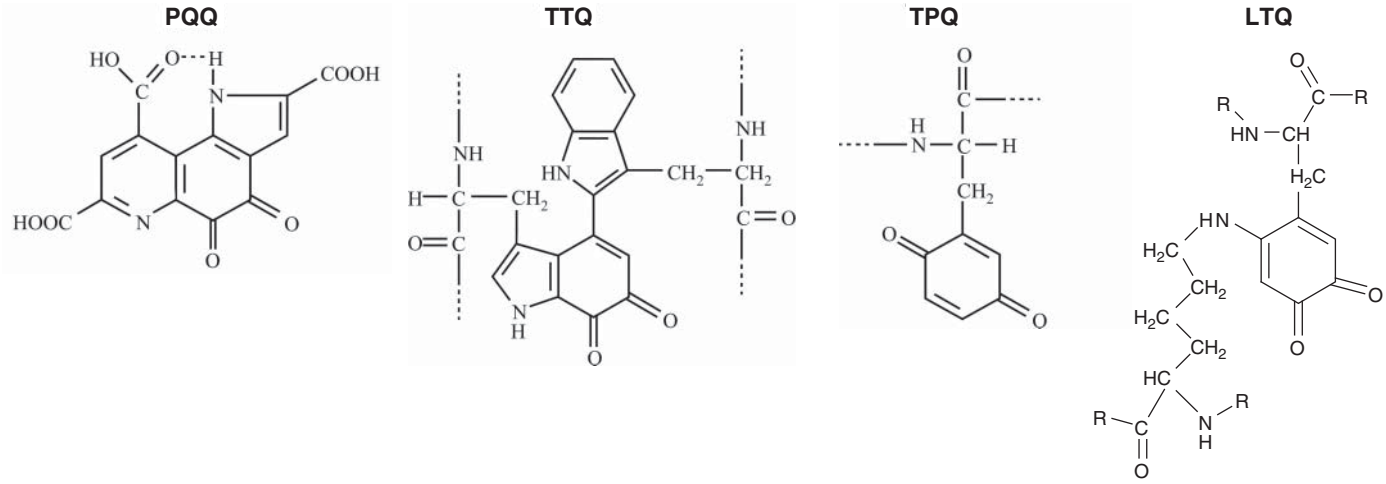


Figure 3.10.8-2. Coenzymes of Quinoenzymes

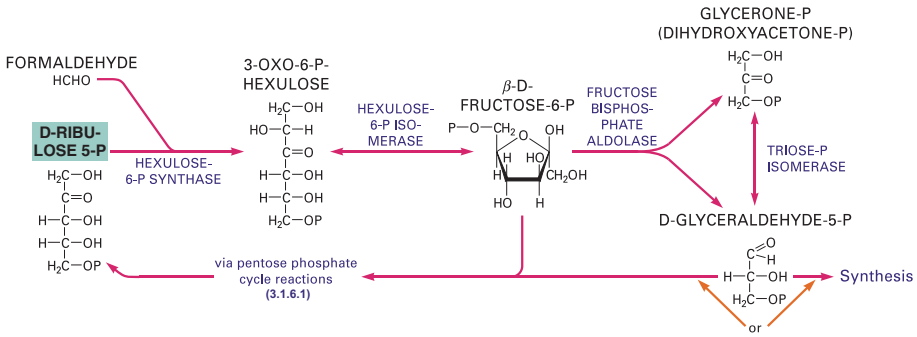


Figure 3.10.8-5. Ribulose Monophosphate pathway

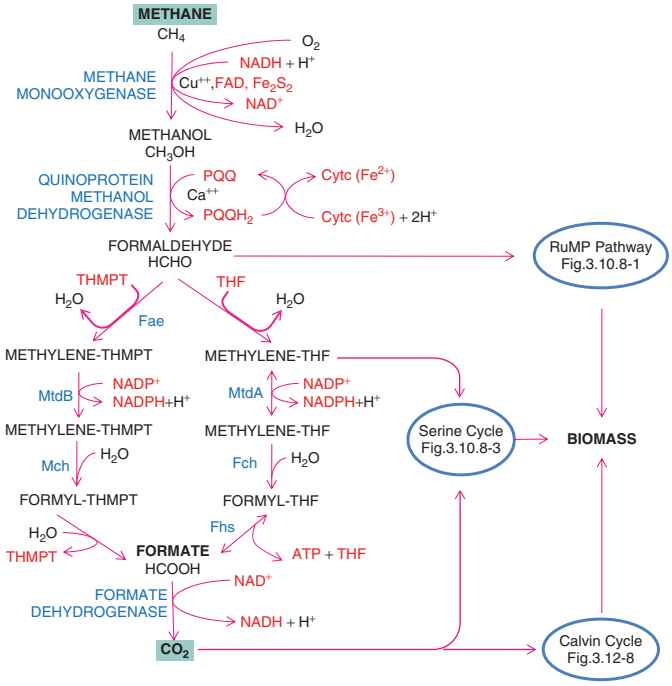


Figure 3.10.8-3. Methane Oxidation

Abbreviated enzyme names: Fae: THMPT-dependent formaldehyde activating enzyme; MtdB: NAD(P)-dependent methylene-THMPT dehydrogenase; Mch: methenyl-THMPT cyclohydrolase; Ftr: formyl-transferase+Ftr-hydrolase complex;.MtdA: NADP-dependent methylene-THMPT dehydrogenase; Fch: methenyl-THF cyclohydrolase; Fhs: formyl-THF synthetase.

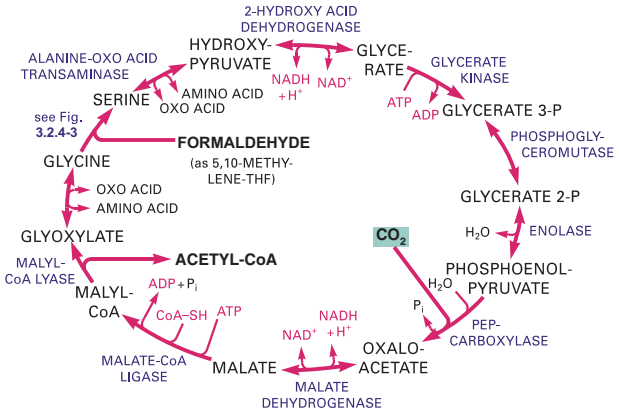


Figure 3.10.8-4. Serine Pathway

pathway (Fig. 3.10.8-4, compare 3.1.3.4 and 3.2.4.2), while the ribulose monophosphate pathway assimilates formaldehyde (Fig. 3.10.8-5). Assimilation of CO_2 via the Calvin cycle is also observed in some microorganisms, but seems to play a minor role.

Literature:

Hakemian A.S., Rosenzweig A.C. Annu. Rev. Biochem. 2007;76:223–241.

Chistoserdova, L. *et al.* Science 1998;281:99.

Vorholt, J. A. Arch. Microbiol. 2002;178:239–249.

Matsushita K. *et al.* Appl. Microbiol. Biotechnol. 2002;58:13–22.

Anthony, C. Arch. Biochem. Biophys. 2004;428:2–9.

3.10.9 Antibiotics

Antibiotics are natural or synthetic compounds which prevent growth of microorganisms, either by inhibiting growth (bacteriostatic) or killing (bacteriocidal) of the microbial cell. Antibiotics belong to a variety of chemical classes (see Table 3.10.9-1) and usually have low molecular masses (below a few kDa).

Due to the dramatic increase in antibiotic resistance caused by various factors, the World Health Organization (WHO) defined a number of 'critically important' antibiotics, among them fluoroquinolones, macrolides, and third- and fourth-generation cephalosporins. Several important antibiotics produced by fungi and bacteria are discussed in detail below.

3.10.9.1 Penicillin and Cephalosporin (Fig. 3.10.9-1)

Both antibiotics are derivatives of tripeptides, which are converted to typical ring structures: penicillins show a five-membered thiazolidine ring, cephalosporins a six-membered dihydrothiazine ring, which are fused to a four-membered β-lactam ring. Because of the ring substituents, penicillins are hydrophobic, whereas cephalosporins are hydrophilic. Their activity spectrum is determined mostly by the structure of the acylamino side chain.

Biosynthesis: The biosynthesis is an example of a non-ribosomal peptide synthesis by a multienzyme complex. As far as is known, all naturally occurring penicillins and cephalosporins are synthesized from the same three amino acids, L-2-aminoadipic acid, L-cysteine and L-valine. They are condensed to the tripeptide δ-(L-2-aminoadipyl)-L-cysteinyl-D-valine (ACV). These reactions are catalyzed by the single enzyme δ-(L-2-aminoadipyl)-L-cysteinyl-D-valine synthetase. The bicyclic isopenicillin N (IPN) is then formed from the linear ACV tripeptide. IPN is the branch point of penicillin and cephalosporin biosynthesis.

In the penicillin biosynthesis (only in fungi), the hydrophilic L-2-aminoadipic acid side chain of isopenicillin N is exchanged for a hydrophobic acyl group, e.g., phenoxyacetyl in penicillin V or phenylacetyl in penicillin G.

The biosynthesis of cephalosporins has been reported in bacteria (cephamycin C) and fungi (cephalosporin C). The L-2-aminoadipic acid side chain of IPN is epimerized to the D-enantiomer to yield penicillin N. Then ring expansion to deacetoxycephalosporin C (DAOC) takes place. Oxidation of the methyl group at C-3 yields deacetylcephalosporin C. In *A. chrysogenum*, this hydroxy group is then acetylated to cephalosporin C, while in bacteria different groups are attached. E.g., in *S. clavuligerus*, several additional steps lead to the end product cephamycin C (7-methoxycephalosporin).

For the synthesis of semisynthetic penicillins, biosynthetic penicillins (e.g., penicillin G) are cleaved *in vitro* either enzymatically by penicillin amidase or chemically. The resulting 6-aminopenicillanic acid is chemically modified by addition of different side chains leading to a variety of different penicillins.

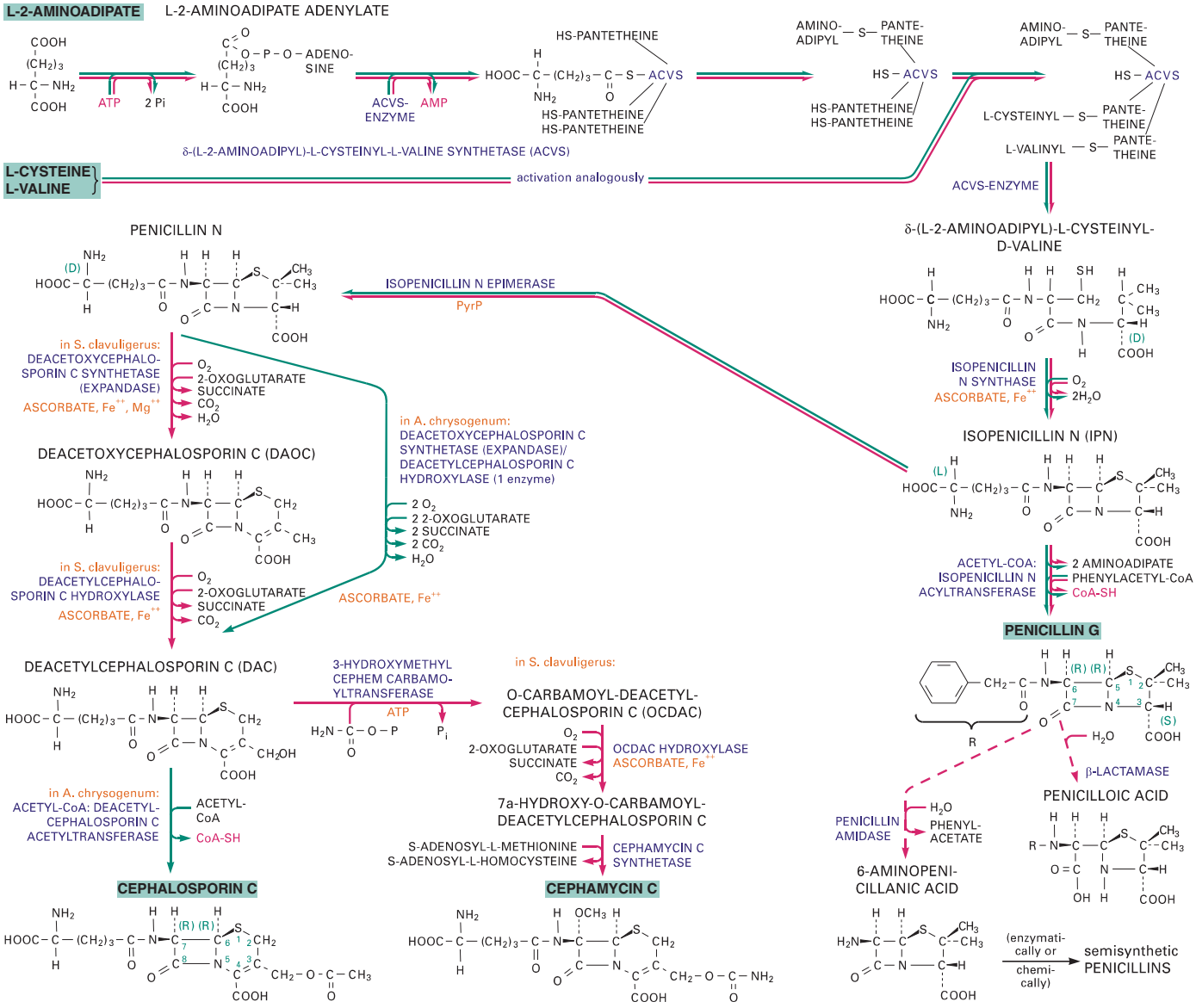


Figure 3.10.9-1. Biosynthesis of Penicillins and Cephalosporins

Resistances: Resistance against β-lactam antibiotics is mainly due to hydrolytic cleavage of the β-lactam ring by β-lactamases or removal of the side chain by acylhydrolases. These enzymes are secreted by resistant microorganisms into the periplasmic space. In addition, a decrease in the penetrability of the external cellular membrane of Gram-negative bacteria or modifications of the target sites (penicillin binding proteins in Gram-positive and -negative bacteria) can cause resistance.

3.10.9.2 Streptomycin (Fig. 3.10.9-2)

Streptomycin belongs to the aminoglycoside antibiotic group and is produced by *Streptomyces griseus*. Aminoglycosides are broad spectrum antibiotics, effective against both Gram-negative and Gram-positive bacteria. Streptomycin is a water soluble molecule exhibiting a basic pH.

Biosynthesis: A set of 25–30 genes is expected to be involved in Streptomycin biosynthesis. It is synthesized from the precursors dTDP-dihydrostreptose, NDP-N-methyl-L-glucosamine and streptidine-6-P, which are formed by independent biosynthetic pathways. All of the precursor molecules are derived from D-glucose.

For formation of dTDP-dihydrostreptose, glucose is converted to the dTDP derivative. The hydroxy group at C-4 is oxidized to the oxo function, followed by reduction of the hydroxy group at C-6. Then, epimerization at C-3 and C-5 and finally a ring narrowing to a five-membered ring takes place. The biosynthesis of N-methyl-L-glucosamine requires the transformation of a D-sugar to its L-enantiomer. Some steps have not yet been completely resolved. The pathway also needs activation of the sugar as a nucleoside diphosphate derivative (compare 3.1.2.2). It is unclear whether this occurs with UTP or CTP. Streptidine 6-P is synthesized via *myo*-inositol (for reaction see 3.1.5.4) which is then oxidized at C-1 and transaminated to give inos-1-amine. After phosphorylation, the compound is transamidinated by arginine. The same procedure is repeated at the C-3 position. For the final steps of streptomycin biosynthesis, streptidine-6-P and dTDP-streptose most likely react first to an intermediate named pseudo-disaccharide-6-P. Addition of N-methyl-L-glucosamine yields dihydrostreptomycin-6-P. This is the last soluble intermediate which can be found within streptomycin producing cells. Oxidation to streptomycin-6-P takes place at the cytoplasmic membrane. NAD^+ serves as an electron acceptor. The phosphorylated product is still inactive; dephosphorylation to the active antibiotic is achieved by an extracellular alkaline phosphatase.

Resistances: Resistance of bacteria against streptomycin can be due to decreased penetration into the cells, a mutated target site or to bacterial modification of the streptomycin molecule. So far, phosphorylation at the C-6 bound hydroxyl group of the streptidine moiety or at the C-3 bound hydroxy group of the N-methylglucosamine moiety has been described, as well as a mannosylation at the C-4 bound hydroxyl of the N-methylglucosamine residue.

Table 3.10.9-1. Classification of Antibiotics according to their chemical structure

Chemical Class	Subgroups (Selection)	Antibiotic (Example)	Producing Organisms	Site of Antibiotic Effect
1 Carbohydrate derivatives	aminoglycosides	streptomycin	bacteria, e.g., *Streptomyces griseus*	Bacterial protein biosynthesis, inhibits the function of the 30 S ribosomal subunit by causing translational errors
2 Macrocyclic lactone compounds	macrolides	erythromycin	bacteria, e.g., *Saccharopolyspora erythrea*	Bacterial protein biosynthesis, inhibits the function of the 50 S ribosomal subunit
3 Quinones and related compounds	tetracycline	tetracycline	bacteria, e.g., *Streptomyces aureofaciens*	Bacterial protein biosynthesis, inhibits the function of the 30 S ribosomal subunit
4 Amino acid and peptide analogs	β-lactam antibiotics	penicillin	fungi, e.g., *Penicillium chrysogenum, Emericella nidulans*	Biosynthesis of the bacterial cell wall
	cephalosporins	cephalosporin	bacteria, e.g., *Streptomyces clavuligerus;* fungi, e.g., *Cephalosporium acremonium*	Biosynthesis of the bacterial cell wall
	amino acid derivatives	chloramphenicol	bacteria, e.g., *Streptomyces venezuelae*	Bacterial protein biosynthesis, inhibits the function of the 50 S ribosomal subunit
5 N-containing heterocyclic compounds	nucleoside antibiotics	puromycin	bacteria, e.g., *Streptomyces alboniger*	Eukaryotic and bacterial protein biosynthesis
		nikkomycin	bacteria, e.g., *Streptomyces tendae*	Fungicide and insecticide, inhibits the biosynthesis of chitinous cell walls (chitin synthase)
6 O-containing heterocyclic compounds	polyether antibiotics	monensin	bacteria, e.g., *Streptomyces cinnamonensis*	Acts as ionophore at cytoplasmic membrane
7 Alicyclic compounds	steroid antibiotics	fusidic acid	fungi, e.g., *Fusidium coccineum*	Bacterial and eukaryotic protein biosynthesis
8 Aromatic compounds	condensed aromates	griseofulvin	fungi, e.g., *Penicillium griseofulvum*	Fungicide, effective against fungi with chitinous cell wall
9 Aliphatic antibiotics	polyene AB	fumagillin	fungi, e.g., *Aspergillus fumigatus*	Inhibition of the eukaryotic DNA synthesis
10 Chinolones	fluor-4-chinolone	ciprofloxacin	synthetic	Interference with DNA replication by inhibiting DNA gyrase or topisomerase IV
11 Oxazolidinones	cyclic lactone	2-oxazolidinone	synthetic	Blocks translation by inhibiting initiation complex formation

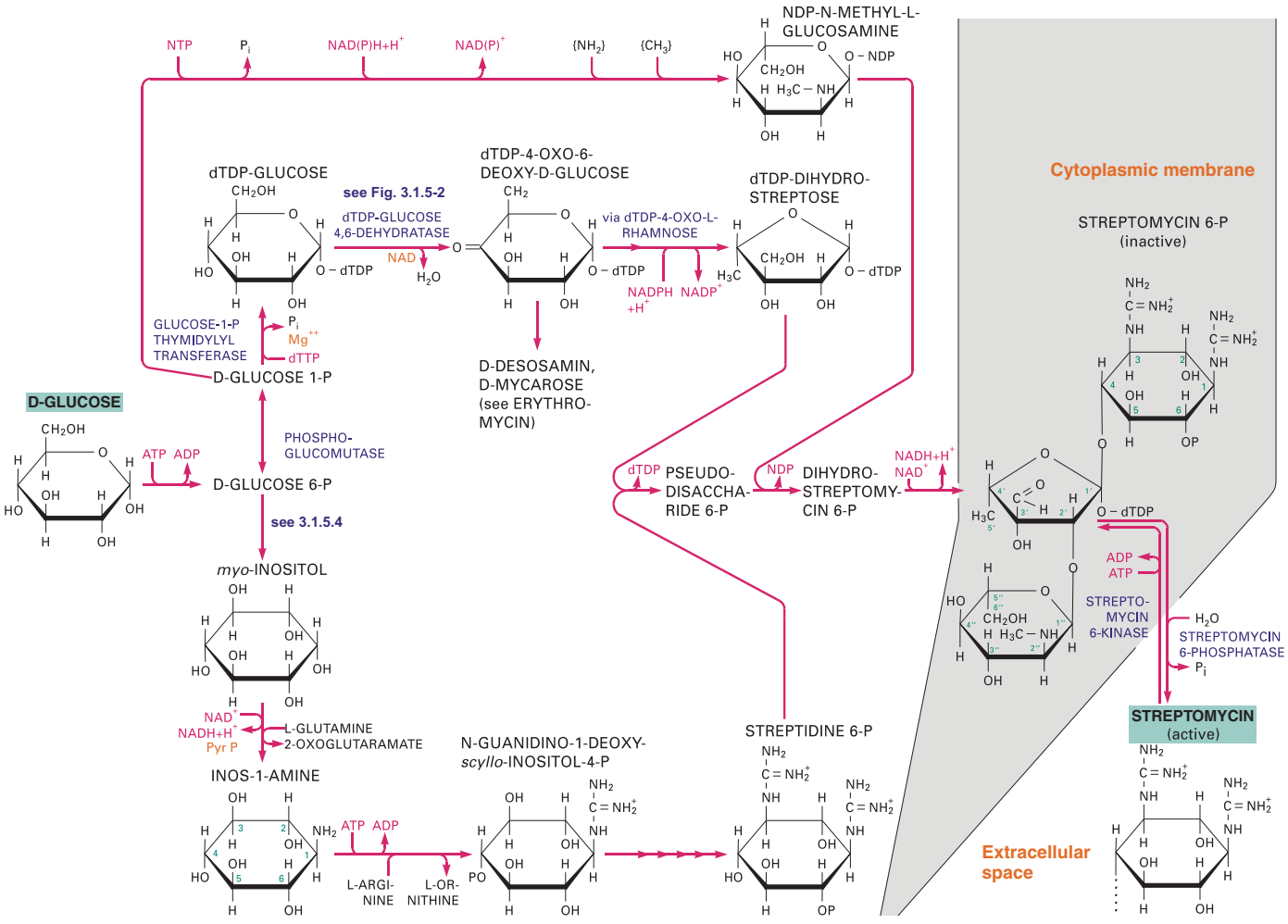


Figure 3.10.9-2. Biosynthesis of Streptomycin

3.10.9.3 Erythromycin (Fig. 3.10.9-3)

Erythromycin is a macrolide antibiotic consisting of a 14-membered lactone ring which is produced by the actinomycete *Saccharopolyspora erythraea.* Two sugar substituents, D-desosamine and L-cladinose, are essential for its biological activity.

Biosynthesis: The biosynthesis of macrolides proceeds via the polyketide pathway. To one molecule of propionyl-CoA, six units of 2-methylmalonyl-CoA are consecutively added. During each condensation step, one C-atom is eliminated via CO_2. Apart from the lack of reductive steps, these reactions resemble those of the fatty acid

synthesis (3.4.1.1). Upon ring closure from the instable intermediates, the first stable intermediate 6-deoxy-erythronolide B is formed. After oxidation at C-6, glycosylation with both dTDP-activated L-mycarose and D-desosamine yields erythromycin D. By methylation of the bound L-mycarose moiety to L-cladinose and oxidation at C-12, erythromycin A is formed. These two last steps may also proceed in reverse order. Erythromycin E results from oxidation of both the methyl group at C-2 of the macrolide ring and the C-1 of the L-cladinose ligand.

Resistances: Resistances of bacteria against erythromycin are mostly due to modification of the target site of the antibiotic (e.g., MLS-resistance by N^6-dimethylation of the base 2058 of 23 S rRNA).

3.10.9.4 Tetracycline (Fig. 3.10.9-4)

Tetracyclines inhibit the bacterial protein biosynthesis and are produced by different *Streptomyces* species, e.g., *Streptomyces aureofaciens*.

Biosynthesis: As described for the biosynthesis of erythromycin, tetracyclines are also synthesized via the polyketide pathway. Polyketides are synthesized by the polymerization of acetyl and propionyl subunits in a process similar to fatty acid biosynthesis (3.4.1.1).

Malonamoyl-CoA is formed from asparagine via L-oxosuccinamate (3.2.2.4). Enzyme-bound malonamoyl-CoA is condensed with malonyl-CoA (formed from acetate, 3.4.1). Consecutively, seven additional malonyl-CoA units are added, with the elimination of one C-atom as CO_2 in each step. The first intermediate is 6-methylpretetramide, which has been demonstrated by the use of knockout mutants. After hydroxylation and oxidation reactions, 4-dedimethylamino-4-oxo-anhydrotetracycline is a branch point of the biosynthetic pathway. Halogenation by chloride haloperoxidase leads via several following steps to chlorotetracycline. For formation of tetracyclines an amino group is introduced, followed by the addition of two methyl groups and subsequent hydroxylation and reduction reactions. Another branch of the pathway at the stage of 5α (11α)-dehydro-tetracycline leads to formation of oxytetracycline. Some steps of the biosynthetic pathway have not been clarified yet.

Resistances: Tetracycline enters bacteria by an active transport mechanism through the cytoplasmic membrane. The major mechanisms of resistance are decreased uptake and active transport out of the cells.

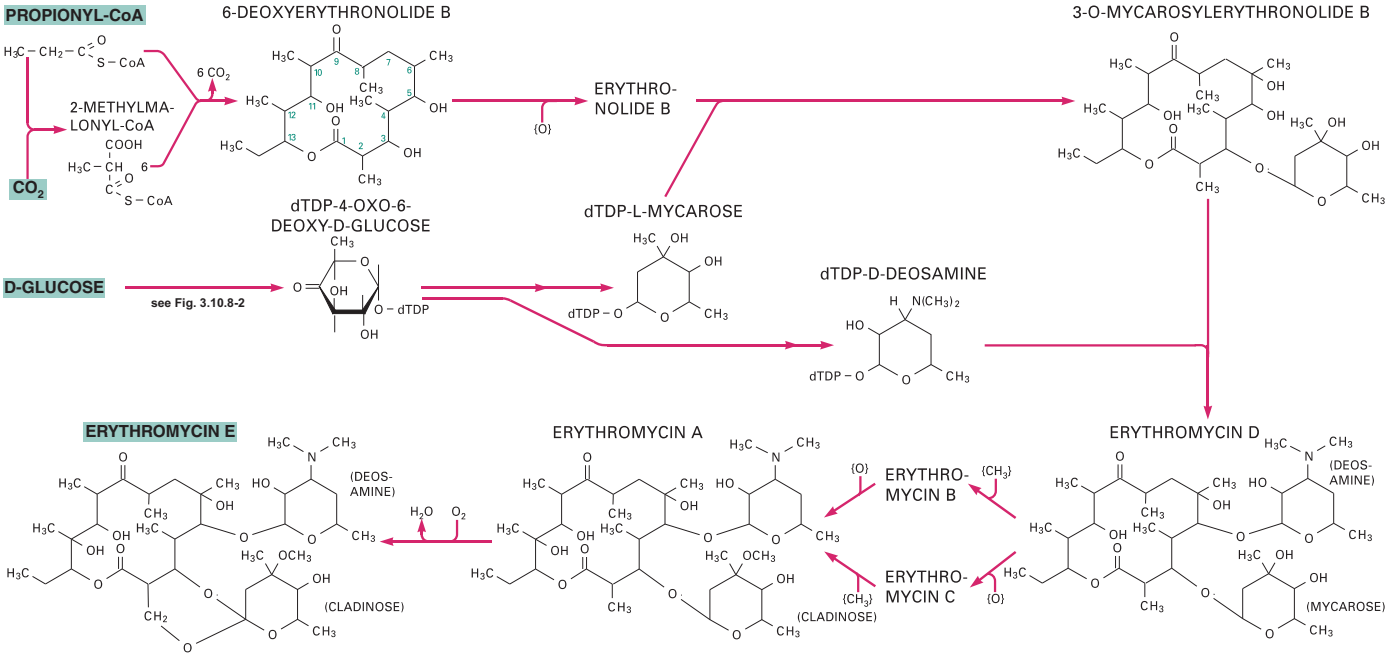


Figure 3.10.9-3. Biosynthesis of Erythromycin

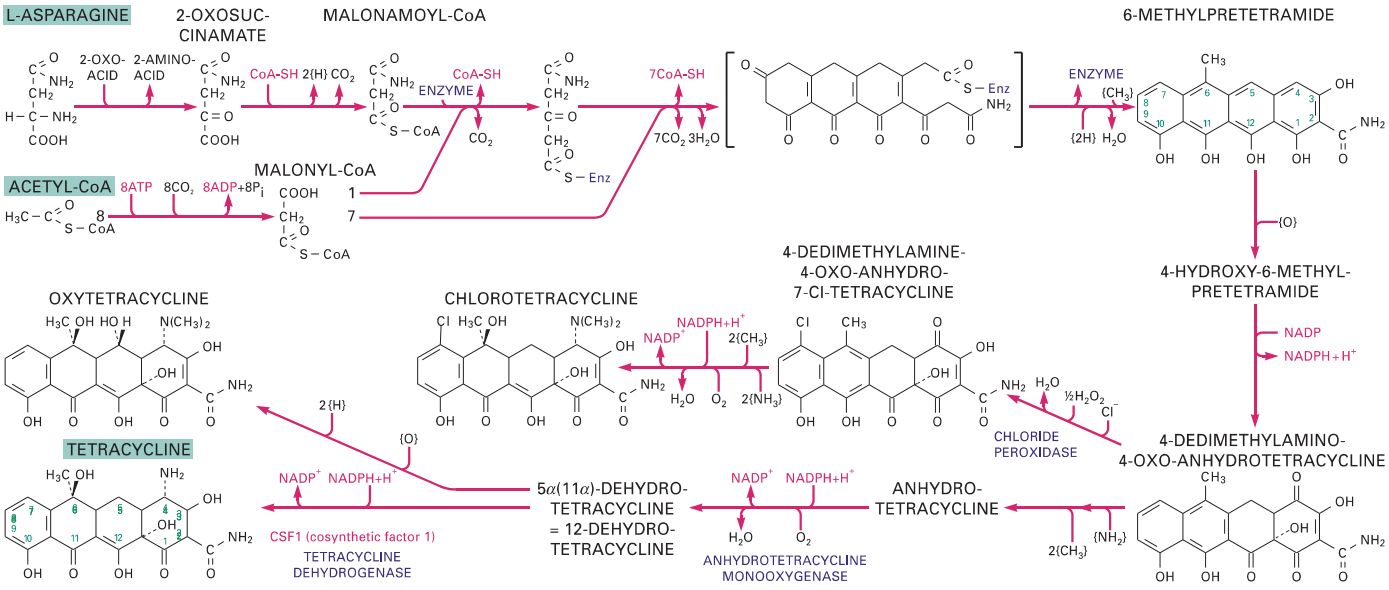


Figure 3.10.9-4. Biosynthesis of Tetracycline

Literature:

Aharonowitz, Y. *et al*. Bio/technology 1993;11:807–810.

Berdy, J. in: Verrall, M. S. (Ed.) *Discovery and Isolation of Microbial Products*. Ellis Horwood Publishers, 1985; p. 9–31.

Collignon, P. *et al*. Clin. Infect. Dis. 2009;49:132–141.

Brakhage A. A. Microbiol. Mol. Biol. Rev. 1998;62:547–585.

Kleinkauf, H., von Döhren, H. Eur. J. Biochem 1996;236:335–351.

Mascaretti, O.A. *Bacteria versus antibacterial agents: an integrated approach.* Washington D.C.: ASM Press, 2003.

Piepersberg, W. in: Vining, L., Stuttard, C. (Eds.) *Biochemistry and Genetics of Antibiotic Biosynthesis*. Butterworth, Heinemann, Stoneham, 1994; p. 71–104.

3.11 Electron Transfer Reactions and Oxidative Phosphorylation

Martina Jahn and Dieter Jahn

3.11.1 General Principles

ATP is the general energy currency of all living cells. It is employed in activation processes during the catabolic degradation of substrates, to energize transport, during signaling and for many biosynthetic anabolic reactions. In nature two different principles for ATP formation are known. Both utilize the oxidation of substrates for the formation of ATP from ADP and P_i. They are present in all kingdoms of life.

- Fermentation processes use the energy liberated from a substrate directly at the substrate level (3.10.5).
- Oxidative phosphorylation couples the oxidation of compounds with the establishment of an ion gradient across a membrane, which in turn provides the energy for ATP synthesis. This process is called chemiosmotic coupling. These systems usually encompass
  - Electron transporting, membrane-localized multiprotein complexes which combine electron transfer with H^+/Na^+ transfer across the membrane.
  - Membrane-associated primary dehydrogenases to supply high energetic electrons from substrates for transport and to pump H^+/Na^+ across the membrane.
  - Electron carrying cofactors, quinones, iron-sulfur clusters, heme, flavins and copper ions.
  - Terminal oxidases for the final transfer of the transported electron to electron acceptors like oxygen, nitrate or fumarate. They pump H^+/Na^+ across the membrane.
  - H^+/Na^+ -transporting ATP synthase, which couples ion translocation via the membrane barrier to ATP synthesis.

The H^+/Na^+ gradient for energy generation is formed at the cytoplasmatic membrane of prokarya and the inner membrane of the mitochondria of eukarya. The ion pumps are driven by the electron transport out of the prokaryotic cell or into the intermembrane space of the mitochondria.

3.11.2 Different Types of Electron Transport Chains

In the mitochondria of eukarya, mainly NADH is used as electron donor and primarily oxygen as an electron acceptor. Respiratory nitrate reduction has been observed for some fungi.

In prokarya, respiratory chains with oxygen as final electron acceptor may start with the oxidation of, e.g., Fe^{2+} to Fe^{3+} in some chemolithotrophs (*Thiobacillus ferrooxidans* or *Gallionella ferruginea*). Another example is the nitrification process, where electrons from ammonia are transferred to oxygen. Resulting nitrite is further oxidized to yield nitrate.

$$2\ NH_3 + 3\ O_2 \rightarrow 2\ NO_2^- + 2\ H^+ + 2\ H_2O \qquad (\Delta G'_0 = -272\ kJ / mol\ NH_3)$$

$$2\ NO_2^- + O_2 \rightarrow 2\ NO_3^- \qquad (\Delta G'_0 = -75\ kJ / mol\ NO_2)$$

Prokarya that are capable of an anaerobic life style employ many different organic and inorganic electron acceptors and donors. During methanogenesis under anaerobic conditions methanogenic archaea use formate, methanol, acetate, methylamines, carbon monoxide and H_2 as electron donors to reduce CO_2. Some methanogens can use ethanol, 2-propanol, butanol and various ketones instead. Three major groups of methanogens can be distinguished.

$$CO_2 + 4\ H_2 \rightarrow CH_4 + H_2O \qquad (\Delta G'_0 = -136\ kJ / mol\ CH_4)$$

$$4\ CH_3OH \rightarrow 3\ CH_4 + CO_2 + 2\ H_2O \qquad (\Delta G'_0 = -106\ kJ / mol\ CH_4)$$

$$CH_3COOH \rightarrow CH_4 + CO_2 \qquad (\Delta G'_0 = -37\ kJ / mol\ CH_4)$$

In contrast to mitochondria, the construction of electron transport systems is highly variable in bacteria and archaea, depending on the nature and energy difference of reductant and oxidant.

Examples of electron acceptors are CO_2, methanol or acetate in methanogenesis (3.10.6.2), nitrate in ammonification (3.10.6.1), nitrate in sulfur, H_2S or $S_2O_3^{2-}$ oxidation (3.10.6) and oxidized metals (3.10.6.4) including Fe(III), Mn(IV), U(VI), Se(VI), Cr(VI), Co(III), As(V), Tc(VII) and V(V). In these systems, various carbohydrates and amino acids, acetate, H_2, lactate, glycerol, formate are the electron donors.

3.11.3 The Energetic Basis of the Oxidative Phosphorylation

The amount of ATP formed during respiration depends on the free energy of the performed reactions, which can be calculated from the redox potential difference ΔE of the reactions. If the standard conditions for the reaction are used, the change of the standard free energy $\Delta G'_0$ is obtained by the equation

$$\Delta G'_0 = n * F * \Delta E'_0, \qquad [1.5\text{-}6b]$$

where n= number of the transferred electrons, F = Faraday constant (96.4 kJ/V* mol), E'_0 = redox potential at a concentration of 1 mol/l, 25°C and pH 7.

In mitochondria, the reaction

$$NADH + H^+ + ½\ O_2 \rightarrow NAD^+ + H_2O$$

yields a $\Delta\ E'_0$ of −1135 mV. With the standard conditions of pH = 7.0 with n = 2 electrons, equal concentrations of NAD^+ and NADH, atmospheric O_2 pressure at 37°C, a change in the free energy $\Delta G'_0$ = -219 kJ/mol can be calculated. For conditions different from the standard ones outlined, ΔE_0 in the above formula has to be modified according to the equation

$$\Delta G = \Delta G'_0 + R*T*2.303*\log \frac{[NAD^+]}{[NADH]*[½\ O_2]} \qquad [3.11\text{-}1]$$

(application of [1.5-2])

This is usually the case in a living cell since equal concentrations of NAD and NADH rarely occur under physiological conditions. Using these formulae one can calculate the energy recovery for every type of respiration since the redox potential differences are usually known (Table 3.11-1).

The phosphorylation potential ΔG_{ATP} is defined for the back reaction, i.e., for hydrolysis of ATP

$$ADP + P_i + H^+ \rightarrow ATP + H_2O$$

and depends on the concentration ratio of the reactants according to

$$\Delta G_{ATP} = \Delta G'_0 + R * T * 2.303 * \log \frac{[ADP] * [P_i]}{[ATP]} \qquad [3.11\text{-}2]$$

With $\Delta G'_0$ = −30.5 kJ/mol, ΔG_{ATP} varies between −46 to −67 kJ/mol from bacteria to mitochondria.

In the mitochondrial respiratory chain the stoichiometry of transported H^+ per oxidized NADH is 10. The ATP synthase requires 3 H^+ per ATP generated. In addition, one H^+ per generated ATP is consumed for the export of ATP from mitochondria. Thus, for cytosolic ATP the ratio is 4 H^+ per ATP and the oxidative phosphorylation ratio P/O is therefore 2.5 ATP/NADH ('P/O quotient').

3.11.4 Electron Transport System in Mitochondria and Bacteria (Fig. 3.11-1)

The respiratory chain of mitochondria consists of four multienzyme complexes, termed complexes I to IV (Table 3.11-2), which are all membrane associated or transmembrane proteins of the inner mitochondrial membrane. They contain as redox active cofactors flavines, iron sulfur clusters, hemes (3.3.2) and Cu centers (Table 3.11-3). They are functionally combined with ATP synthase (complex V).

The complexes in bacteria (e.g., *E.coli*) living in aerobic environments and oxidizing NADH are set up in a similar way, although differences exist (Fig. 3.11-1 and following sections). A complex analogous to the mitochondrial complex III is missing, the cyctochrome *bd* and *bo* complexes function as terminal oxidases.

3.11.4.1 Complex I: NADH:Ubiquinone Oxidoreductase

Complex I, also called NADH dehydrogenase, performs the two electron oxidation of NADH + H^+. This process is used in mitochondria to drive the transport of four protons per two electrons from the matrix side to the intermembrane space or in prokaryotes from the cytoplasm to the periplasm. The *E. coli* complex I is composed of 14 'core' subunits with a total mass of 550 kDa, while the human enzyme consists of 45 subunits with a combined mass of 980 kDa. In spite of the difference in size, both complexes contain equivalent redox components and show the same overall structure. Since the sequences in the 'core' subunits have been highly conserved during evolution, apparently bacteria and mitochondria use the same reaction mechanisms which are discussed here together, based on the recent structure determination of the *E. coli* and the *Thermus thermophilus* enzymes.

All NADH dehydrogenases are L-shaped and consist of two major parts. A hydrophobic arm is embedded into the membrane, while a hydrophilic arm protrudes into the matrix or the periplasm, respectively. The latter contains the primary electron acceptor FMN at the tip, followed by a sequence of one [2Fe-2S] cluster and six [4Fe-4S] clusters. Another [2Fe-2S] cluster and [4Fe-4S] cluster apparently are not in the main path of electron transfer. At the interface to the hydrophobic arm there is likely the binding site for an ubiquinone UQ. The membrane-embedded arm is composed of the subunits Nuo L, M, N, A, J and K and the easily dissociating Nuo H (*E. coli* nomenclature). The homologous subunits L, M and N each contain 14 conserved, slightly tilted transmembrane helices and a helix interrupted in the center by a peptide 'loop'. In addition, a long helix HL passes parallel to the membrane surface through the subunits L, M and N. It is in contact with the three 'interrupted' helices in L. M and N.

Based on this structure, the enzyme functions as follows: The two-electron carrier FMN receives two electrons from NADH. The series of iron-sulfur complexes mediate the one-electron transfer to the final cluster N2 and from there to the bound UQ, which thereafter collects another electron, takes up two protons and enters the 'Q cycle' in the membrane (see below). The charge transfer also causes major conformation changes in the subunits at the angle of the 'L'. This effects the translocation of a proton into the intermembrane space (or periplasm, respectively); possibly involving cluster N2 and its tandem cysteines. It also apparently moves the helix HL in piston-like fashion, which tilts the transmembrane helices in the subunits L, M and N, exposing charged residues and thus effecting translocation of 3 more protons.

3.11.4.2 Complex II: Succinate:Ubiquinone Oxidoreductase

Complex II (also called succinate dehydrogenase) is a member of the citrate cycle and the only enzyme of that cycle that is an integral part of the membrane. The function of this complex can be described as

ubiquinone (UQ) + 2 H^+ + 2 e^- → ubiquinol (UQH_2)

Table 3.11-1. Free Energies from Redox Reaction during Various Respiratory Processes

Reductive half reaction	$\Delta G'_0$ (kJ/2 electrons)
½ O_2 → H_2O	–219
2 NO_3^- → N_2	–206
Fe^{3+} → Fe^{2+}	–209
CH_3Cl → CH_4 + HCl	–135
MnO_2 → Mn^{2+}	–134
Se (VI) → Se (IV)	–129
Cr (VI) → Cr (IV)	–90
As(V) → As(0)	–46
SO_4^{2-} → HS^-	–20
CO_2 → CH_4	–14

Table 3.11-2. Complexes of the Respiratory Chain in Mammalian Mitochondria

Complex	Mol. Mass (kDa)	Sub-units	Active Groups	Inhibitors
I: NADH: ubiquinone oxidoreductase	907	43	FMN, 3 [2Fe-2S] and 3 [4Fe-4S] centers, bound ubiquinone	rotenone, piercidine
II: Succinate:ubiquinone oxidoreductase	130	5	FAD, 1 [2Fe-2S], 1 [4Fe-4S] and 1 [4Fe-4S] center, Cyt b_{558}	malonate
III: Ubiquinol:cytochrome-c reductase (Cyt bc_1 complex)	248	11	Cyt b_{566}, Cyt b_{562}, Cyt c_1, [2Fe-2S] center (Rieske), ubiquinone	antimycin, myxothiazol
IV: Ferrocytochrome-*c*: oxygen oxidoreductase (Cyt aa_3 complex)	210	13	Cyt a, 2 Cu_A, Cu_B-Cyt a_3 center	CN^-, SH^-, N_3^-, CO

Table 3.11-3. Standard Redox Potentials for the Electron Transfer System of Oxidative Phosphorylation

Redox Pairs	Standard Redox Potential (E'_0) in mV
NADH + H^+/NAD^+ + 2 H^+ + 2 e^-	–320
Complex I	
FMN + 2 H^+ + 2 e^-/$FMNH_2$	–300
Fe-S center (Fe^{3+}) + 1 e^- /Fe-S center (Fe^{2+})	–250 … –50
Succinate/Fumarate + 2 H^+ + 2 e^-	+30
Complex II	
FAD + 2 H^+ + 2 e^- /$FADH_2$	0
Fe-S centers (Fe^{3+}) + 1 e^- /Fe-S centers (Fe^{2+})	–260… 0
UQ + 2 H^+ + 2 e^- /UQH_2	+45
Complex III	
Cytochrome b_L (Fe^{3+}) + 1 e^- /Cytochrome b_L (Fe^{2+})	–10
Cytochrome b_H (Fe^{3+}) + 1 e^- /Cytochrome b_H (Fe^{2+})	+30
Rieske Fe-S center (Fe^{3+}) + 1 e^- /Fe-S center (Fe^{2+})	+280
Cytochrome c_1 (Fe^{3+}) + 1 e^- /Cytochrome c_1 (Fe^{2+})	+210
Cytochrome *c* (Fe^{3+}) + 1 e^- /Cytochrome *c* (Fe^{2+})	+230
Complex IV	
Cytochrome *a* (Fe^{3+}) + 1 e^-/Cytochrome *a* (Fe^{2+})	+10
CuA_{ox} + 1 e^- /CuA_{red}	+40
Cu_B-Cytochrome a_3 (Fe^{3+}) + 1 e^-/ Cu_B-Cytochrome *a* (Fe^{2+})	+390
0.5 O_2 + 2 H^+ + 2 e^-/H_2O	+820

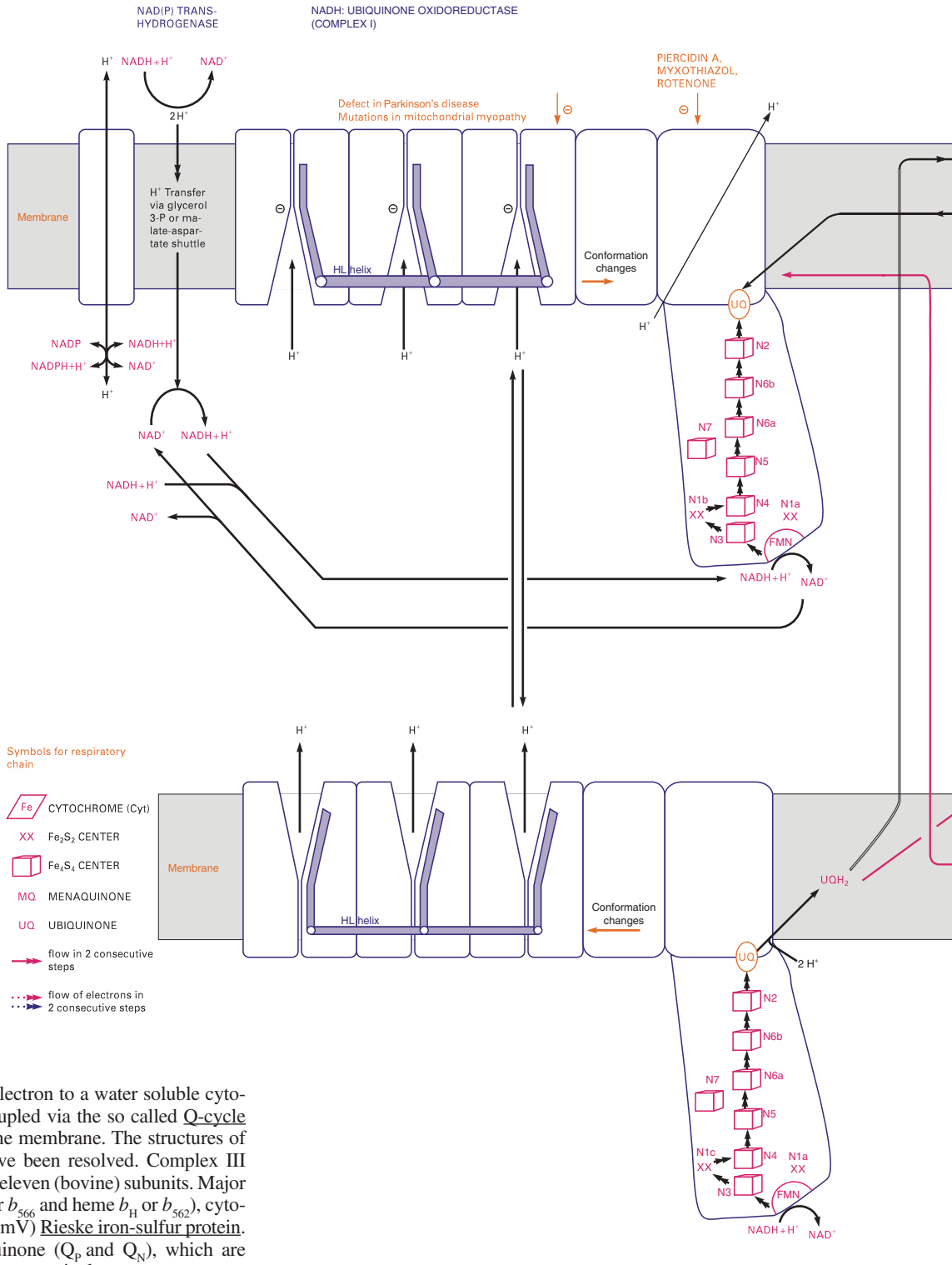
Table 3.11-4. Complexes of the Respiratory Chain in *Escherichia coli*

Complex	Sub-units	Active Groups	Inhibitors
NADH dehydrogenase (ubiquinone)	14	FMN, 3 [2Fe-2S] and 3 [4Fe-4S] centers	rotenone, piercidine
Succinate dehydrogenase	3	FAD, 3 [Fe-S] centers, Cyt b_{556}	
Cytochrome *bd* complex	2	Cyt b_{558}, Cyt b_{595}, Cyt *d*	
Cytochrome *bo* complex	3	Cyt b_{562}, Cyt b_{595} = Cyt *o*, Cu_B	CN^-, SH^-, N_3^-, CO

The enzyme is a trimer. Each monomeric unit consists of two subunits forming the catalytically active head and one or two subunits which form the membrane bound part of the enzyme. One subunit of the head carries the substrate binding site and a covalently bound FAD molecule. Here, the reaction succinate → fumarate + 2 H^+ + 2 e^- is catalyzed. The other subunit of the head possesses one of each of the [2Fe-2S], [4Fe-4S] and [3Fe-4S] clusters, which are responsible for the electron transport to ubiquinone. The membrane spanning subunit(s) carry a heme *b* molecule (often denoted as cytochrome b_{558}), which also may take part in the electron transfer. The path of two protons is under discussion.

In addition to complex I and II several other membrane-associated FAD-containing dehydrogenases may also insert electrons into the freely diffusing ubiquinone. These include oxidoreductases for choline, *sn*-glycerol-3-phosphate and the electron-transfer protein of the fatty acid β-oxidation system (3.4.1.5) which also accepts electrons from sarcosine, dimethylglycine and other substrates.

As a surprising feature, in *E. coli* a transfer of electrons from a biosynthetic process (heme biosynthesis) to quinone pools takes place. The FMN enzyme protoporphyrinogen IX oxidase (3.3.1.3, HemG) removes six electrons from its substrate and transfers them to ubiquinone and menaquinone.



3.11.4.3 Complex III: Ubiquinol: Cytotochrome-c Reductase

Complex III (also called cytochrome bc_1 complex), catalyzes the oxidation of ubiquinol (UQH_2) and transfers an electron to a water soluble cytochrome *c* molecule. This transfer is coupled via the so called Q-cycle with the transport of protons through the membrane. The structures of various cytochrome bc_1 complexes have been resolved. Complex III exists as a dimer with three (bacteria) to eleven (bovine) subunits. Major parts are two cytochromes *b* (heme b_L or b_{566} and heme b_H or b_{562}), cytochrome c_1 and the high potential (+ 280 mV) Rieske iron-sulfur protein. There are two binding sites for ubiquinone (Q_P and Q_N), which are closely located to heme b_L and heme b_H, respectively.

The reaction cycle (Q-cycle) starts with the binding of an ubiquinol molecule (UQH_2) to the first binding site *a* (Q_p). Oxidation of UQH_2 occurs in two steps. First, one electron is transferred to the iron-sulfur Rieske subunit and subsequently to cytochrome c_1. Then two protons are released into the intermembrane space of the mitochondrium. Thus, a semiquinone anion (UQ^-) is formed at the Q_p binding site. The second electron is then transferred to heme b_L (cytochrome b_1, L = low redox potential of −100 mV), converting UQ^- into UQ. Next the electron is moved from heme b_L on the intermembrane space side to heme b_H (H = high redox potential of + 50 mV) at the matrix side of the membrane. There, at a second ubiquinone binding site (Q_n), another molecule of ubiquinone accepts the electron with the generation of a semiquinone anion (UQ^-).

In the second half of the Q-cycle, analogous to the first half, a second molecule of UQH_2 is oxidized at the Q_p site, one electron being passed via the Rieske complex to cytochrome c_1 and the other transferred to heme b_L and then to heme b_H. From there, the b_H-electron is transferred to the 'waiting' semiquinone anion (UQ^-) at the Q_n site. With the addition of two protons from the mitochondrial matrix, the production of UQH_2 and its release from the Q_n site into the coenzyme Q pool, the Q-cycle is completed. Consequently, the Q-cycle at complex III takes two protons from the matrix side and releases four protons into the intermembrane space for each pair of electrons transferred.

The electrons from cytochrome c_1 of complex III are transferred to the water soluble cytochrome *c*. Like ubiquinone in the membrane, cytochrome *c* with its heme cofactor is a mobile electron carrier in the

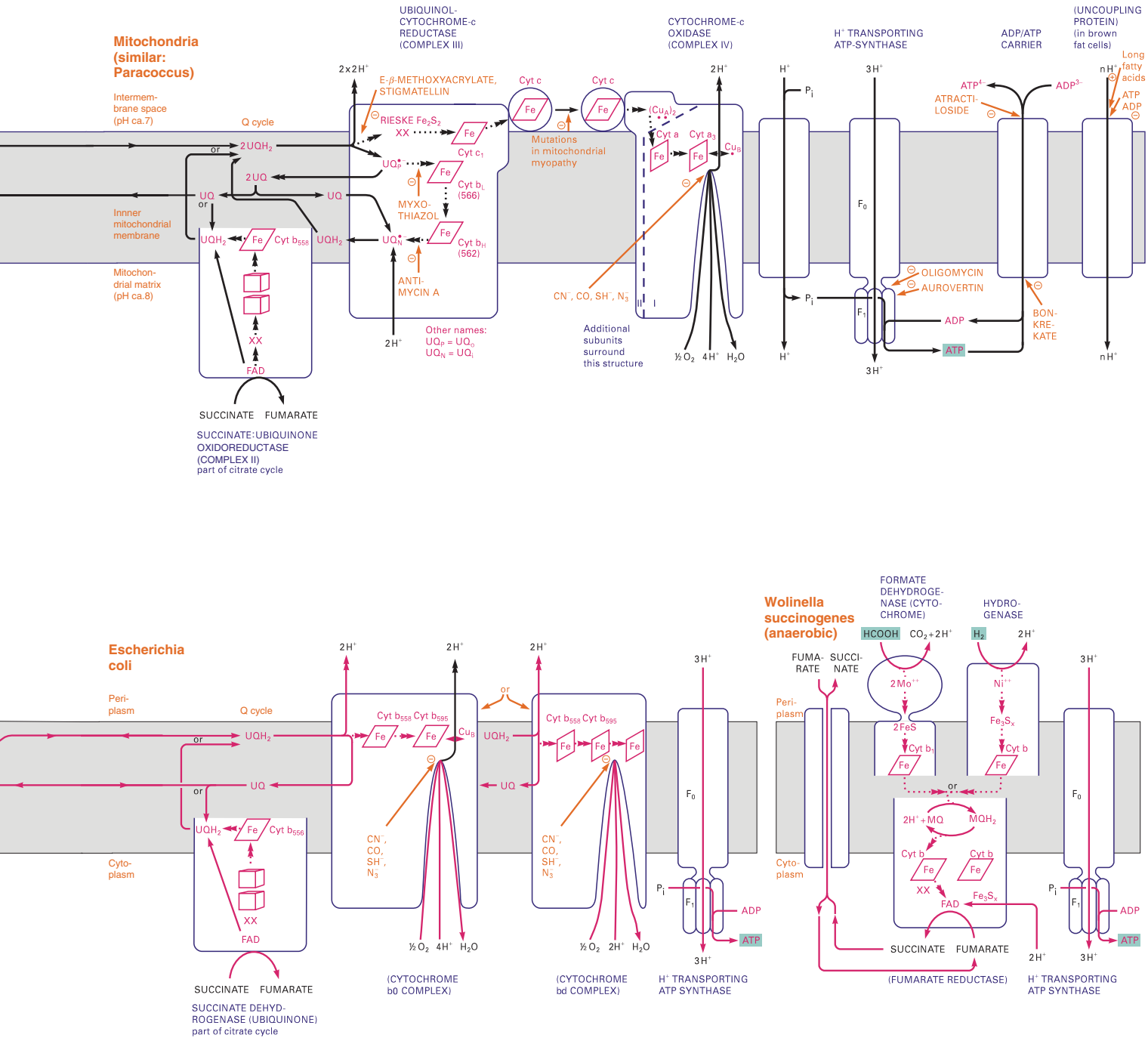


Figure 3.11-1. Respiratory Chain and Oxidative Phosphorylation in Mitochondria (Top) and Bacteria (Below)

intermembrane space of the mitochondrium. It associates to complex III to acquire its electrons and move from there to the cytochrome-*c* oxidase complex IV.

3.11.4.4 Complex IV: Ferrocytochrome-c: Oxygen Oxidoreductase

The mitochondrial complex IV (also named cytochrome-*c* oxidase) consists of 13 subunits. The simpler cytochrome *c* oxidase from the bacterium *Paracoccus denitrificans* has only three subunits, which are identical to the three large subunits I, II and III of the mammalian enzyme (not to be confused with the complexes I, II, and III described above). The structures of various eukaryotic and prokaryotic enzymes are known. All three major subunits have transmembrane helices. Subunit III seems to be of structural function only. Subunits I and II contain the redox active groups, which are two heme *a* (*a* and a_3) molecules, two different copper centers, and additional magnesium and zinc ions. The first copper center Cu_A contains two Cu^{2+} ions in a bi-metal center, while Cu_B carries one Cu^{2+} molecule coupled to cytochrome A_3 in the Cu_B-cytochrome center.

Complex IV accepts four electrons, one at a time from cytochrome *c*, and uses them to reduce O_2 to H_2O according to the overall reaction:

$$4e^- + 8H^+_{in} + O_2 \rightarrow 2H_2O + 4H^+_{out}.$$

The electrons enter the oxidase at the Cu_A center, move from there via cytochrome *a* to the cytochrome a_3- Cu_B center. There, in a deep cleft of the complex, an O_2 molecule is fixed between the iron of the heme and the copper. Four electrons and four protons are transferred to the oxygen in order to form water. The other four protons are transferred from the matrix to the intermembrane space in an exergonic reaction via different amino acid residues in the tunnel.

The supply of hydrogen to the respiratory chain in mitochondria follows an organized mode of action. NADH + H^+ are provided either

from substrates by intramitochondrial dehydrogenases or by extramitochondrial NADH + H^+, which delivers its hydrogen into the matrix by hydrogen shuttles across the inner mitochondrial membrane. Most important are the shuttles employing glycerol-3-phosphate or aspartate-glutamate-malate (Fig. 3.11-2).

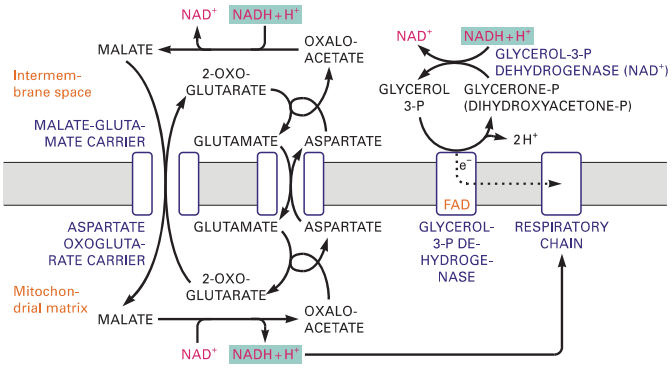


Figure 3.11-2. Carrier Systems for the Transport of Extramitochondrial Hydrogen to the Respiratory Chain

In the respiratory chain of *E. coli*, QH_2 is oxidized directly (without a Cyt bc_1 complex) either

- by the cytochrome bo complex, which (similarly to the complex IV of mitochondria) contains a Cu-heme center or
- by the cytochrome bd complex which contains heme proteins (Cyt b_{558} and b_{595}) and a chlorine-Fe protein (cytochrome d, Fig. 3.3-3).

At high oxygen concentrations, the Cyt bo complex is operative, while at lower pressures the Cyt bd complex is activated.

Obligate anaerobic bacteria lack an oxidase-type complex for obvious reasons. Many different combinations of redox reactions have been found. A survey is given in Section 3.10.6 'Anaerobic respiration'.

As an example, *Wolinella succinogenes* uses fumarate as terminal electron acceptor (Fig. 3.11-1). The dehydrogenases for H_2 (hydrogenase), for formate and other substrates reduce menaquinone (MQ, 3.2.7.2), which is reoxidized by fumarate reductase. The formate dehydrogenase of *W. succinogenes* contains 2 Mo and the hydrogenase Ni in the active centers. Because of the low redox potential of the system, menaquinone (with a 110 mV lower redox potential) is used as carrier instead of ubiquinone.

3.11.4.5 Complex V: ATP Synthase (Fig. 3.11-3)

The fifth mitochondrial complex synthesizes ATP, driven by a proton gradient. The enzyme F_1F_0 ATP synthase (F-type ATPase, TC-system 3.A.2.1) is a multiprotein complex consisting of a hydrophilic domain (F_1), which protrudes into the matrix space and a membrane bound part (F_0). The enzyme reaction is reversible, as has been shown with the reconstituted enzyme in artificial membranes. In biological systems, F_1F_0 ATPase (V-type ATPase, TC-system 3.A.2.2) performs the reverse reaction and pumps protons into vacuoles, coated vesicles, across the plasma membrane in kidneys etc. Its structure is closely related to the ATP synthase.

The F_1 and F_0 parts of the enzyme are both composed of several subunits (Table 3.11-5). The enzyme structure in bacteria, plants and animals has been highly conserved. In every case the F_1 complex contains five different types of subunits with the stoichiometry $\alpha_3\beta_3\gamma\delta\varepsilon$. The F_0 part differs between bacteria and eukarya. While in *E. coli* there are three different subunits with the stoichiometry ab_2c_{10}, the eukaryotic enzymes contain up to 16 different kinds of subunits.

Table 3.11-5. Complexes of the ATP Synthase

Complex	Subunits	Molecular mass (kDa)
Membrane bound part F_0	Mitoch.: a, b, c, d, e, f, g, F_6, A6L, OSCP *E. coli:* a, b_2, c_{10}	68, 56, 44, 20, 26, 13, 11, 12, 8, 22 30, 17, 8
Matrix part F_1	Mitoch./*E. coli:* α_3, β_3, γ, δ, ε	55, 52, 30, 15, 5.6

The structure and the subunits of F_1 have been resolved at atomic resolution. Three α and three β subunits form a $\alpha_3\beta_3$- complex that constitutes the so-called "knob". The α and β subunits alternate in a quasi-symmetric cyclic head, which contains the three active sites at the interface of the α and β subunits. This complex is associated with the F_0 part by a slender central stalk formed by the subunits γ and ε (*E. coli* nomenclature). The γ subunit, tightly associated with the F_0 part, projects into the F_1 complex in a slightly bent form. There is a second stalk made of subunit δ and 2 subunits b, which joins the $\alpha_3\beta_3$ complex with F_0. It is attached to the membrane via interaction with the membrane spanning subunits a and b.

The F_0 part of the *E. coli* enzyme consists of one a, two b and 10 c subunits. The c subunits form a ring around the central stalk. Apparently, the subunits a, b, α, β and δ constitute the 'stator' of a motor while subunits c, γ and ε form the 'rotor'. H^+ flow through the basal F_0 structure protonates the c subunits and causes a structural change, which forces their rotation against the a und b subunits. This three-step-rotation provides the necessary structural changes in the active site formed by the α and β subunits of the F_1 part (Fig. 3.11-4):

- One β subunit binds ADP + P_i while entering the 'loose' (L) conformation,
- the next subunit synthesizes ATP from ADP + P_i in the 'tight' (T) stage,
- the third subunit releases ATP, assuming the 'open' (O) conformation. This is the energy consuming step.

The final step in the oxidative phosphorylation system of mitochondria is the export of ATP against the import of ADP through the inner mitochondrial membrane by an ATP/ADP translocase, which is driven by the membrane potential Ψ. This reaction is closely coupled with the phosphate carrier, which uses up a proton per phosphate imported into the matrix (Fig. 3.11-1).

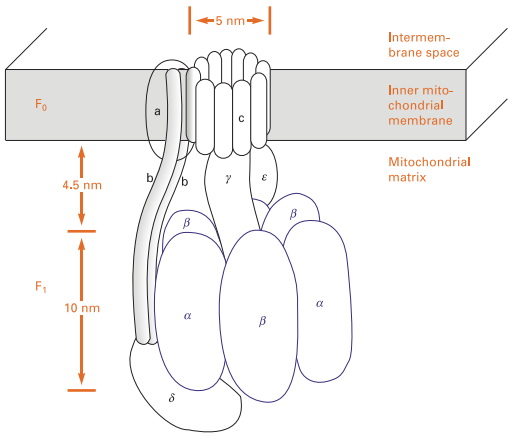


Figure 3.11-3. Structure of the ATP Synthase (*E. coli*)

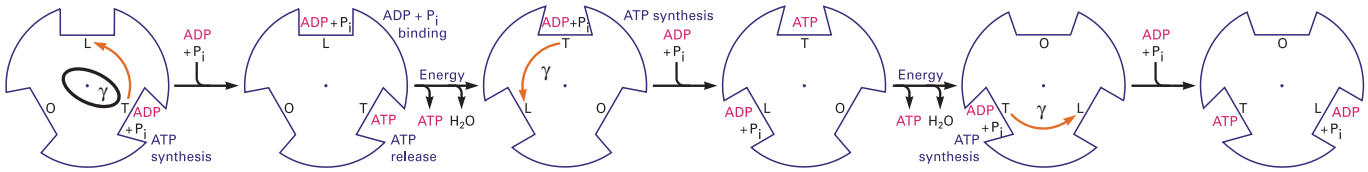


Figure 3.11-4. Rotational Conformation Changes of the ATP Synthase (as seen from the matrix side, schematized)

Heat generation in mitochondria of brown adipose tissue takes place when protons pumped out by the mitochondrial chain are recycled into the matrix by the uncoupling protein. It requires free long-chain fatty acids as activators of the controlling protein thermogenin and is inhibited by ATP and ADP (Fig. 3.11-1).

In some bacteria living in alkaline or Na^+-rich environments, NADH oxidation can be used to pump Na^+ instead of H^+. Other bacteria employ decarboxylation reactions to establish a Na^+ gradient. This gradient is used to energize ATP synthesis by a Na^+-transporting F-type ATP synthase (instead by the H^+ transporting enzyme) or by secondary transport processes (3.10.4.2).

Literature:

Abrahams, J.P. *et al*. Nature 1994;370:621–628.

Branden, G. *et al.* Biochim. Biophys. Acta. 2006;1757:1052–1063.

Brzezinski, P., Gennis, R.B. J. Bioenerg. Biomembr. 2008;40: 521–531.

Cole, J.A., Richardson, D.C. in Boeck, A. *et al.* (Eds.) *Escherichia coli* and *Salmonella* - Cellular and Molecular Biology – online book ***ecosal.*** Washington, D.C.: ASM Press, 2008**.**

Crofts, A.R. *et al.* Biochim. Biophys. Acta 2008;1777:1001–1019.

Curtiss, R.III *et al*. in Boeck, A. *et al.* (Eds.) *Escherichia coli* and *Salmonella* - Cellular and Molecular Biology – online book *ecosal,* Washington D.C.: ASM Press, 2008.

Efremov, R.G., Baradaran, R., Sazanov, L.A. Nature 2010; 465:441–445.

Horsefield, R. *et al.* Curr. Protein Pept. Sci. 2004;5:107–118.

Hosler, J.P. *et al.* Annu. Rev. Biochem. 2006;75:165–187.

Möbius, K. *et al.* Proc. Natl. Acad. Sci. USA 2010;107:10436–10441.

Ohnishi, T. Nature 2010;465:428–429.

Richter, O.M., Ludwig, B. Biochim. Biophys. Acta 2009;1787: 626–634.

Schneider, D. *et al.* Biochim. Biophys. Acta 2008;1777:735–739.

Vik, S.B. Photosynth. Res. 2006;87:231 / author reply 233.

von Ballmoos, C. *et al.* Annu. Rev. Biochem. 2009;78:649–672.

3.12 Photosynthesis

Dieter Oesterhelt and Josef Wachtveitl

Photosynthesis is the light-energized formation of organic compounds (initially carbohydrates) in specialized bacteria, algae and green plants. Thus, light is the original energy source for the organic material composing living organisms. Besides the most important reaction

$$6\ CO_2 + 6\ H_2O = 6\ O_2 + C_6H_{12}O_6 \qquad \Delta G'_0 = 2840\ kJ/mol$$

at least five additional pathways for the integration of CO_2 into organic compounds exist in nature.

With an energy input of ca. 10^{18} kJ/year, about 10^{11} t C/year are converted into biomass on the earth. 3.5–2.0 Billion years ago, early forms of life changed, in this way, the original no-O_2, high-CO_2 atmosphere into the present high-O_2, low-CO_2 atmosphere.

Photosynthesis is composed of two parts: in the light reaction (3.12.1), electrons are energized by light of definite wavelengths for transport of protons (enabling, e.g., ATP formation) or for supplying reducing compounds (NADPH or NADH), while in the dark reaction (3.12.2), the synthetic conversions take place. Many components of this system are related to the oxidative phosphorylation system (3.11). While in oxidative phosphorylation, chemical reactions are used to build up a proton gradient (Mitchell- theory), in photosynthesis light supplies the energy for proton translocation,

3.12.1 Light Reaction (Fig. 3.12-1)

Pathways of electron flow (compare 3.11.1): In chloroplasts of green plants (and similarly in cyanobacteria), the energy of the electrons can be used in 2 ways (Fig. 3.12-2):

- Cyclic electron flow: Protons are moved from the stroma to the thylakoid space (2.2.3) forming a proton gradient which allows the operation of the H^+ -transporting ATP synthase:

  $$ADP + P_i + 3\ H^+_{thyl.\ space} = ATP + 3\ H^+_{stroma}.$$

  The synthase structure and the reaction mechanism corresponds closely to the mitochondrial enzyme (3.11.5).

  When the proton gradient becomes too small, the chloroplast enzyme is inhibited in order to avoid ATP degradation by the reverse reaction

- Noncyclic electron flow: The electrons are used for reduction of NADP:

  $$NADP^+ + H^+ + 2\ e^- = NADPH \qquad \Delta G'_0 = 61.7\ kJ/mol,\ \Delta E'_0 = -320\ mV$$

  Electrons consumed in the latter reaction are replaced by the cleavage of water:

  $$H_2O = ½\ O_2 + 2\ H^+ + 2\ e^- \qquad \Delta G'_0 = 157\ kJ/mol;\ \Delta E'_0 = -815\ mV$$

  resulting in an overall reaction of:

  $$H_2O + NADP^+ = NADPH + ½\ O_2 + H^+ \qquad \Delta G'_0 = +219\ kJ/mol;\ \Delta E'_0 = -1135\ mV$$

  The photon energy amounts to

  $$E = h * \nu * N = \frac{h * c * N}{\lambda} = \frac{0.120}{\lambda}\ [\text{Joule / Einstein}]$$

  where 1 Einstein = 1 mol = N = 6.022 * 10^{23} photons, h = Planck's constant (6.626 *10^{-34} J/sec), c = speed of light (2.998 * 10^8 m/sec), λ = wavelength of the light (m).

  Therefore, light of the wavelength 700 nm (700 * 10^{-9} m) has an energy content of 171 kJ/Einstein. This is light of the longest wavelength (lowest energy), which in most cases can be used by plants for photosynthesis. Actually, in order to reduce one molecule of $NADP^+$ according to the above formula, four photons are required. About a third or less of their energy is used for this purpose and additionally 1/8 or less for proton movement into the thylakoid space, which enables ATP synthesis. These reactions take place in two different photosystems.

Photosynthetic purple bacteria (e.g., *Rhodobacter*) have also two mechanisms for electron flow (Fig. 3.12-3):

- The cyclic electron flow for proton translocation from cytoplasm to periplasm exists in a form different to green plants.
- Since the bacteria have only one photosystem and are unable to split water to supply electrons for NAD^+ reduction, they must obtain them from, e.g., H_2S or $H_2S_2O_3$ (reverse = noncyclic electron flow).

Halophilic archaea have only a light-driven proton transport system, which operates by a different mechanism (Fig. 3.12-7).

Structure of the photosystems in higher plants and cyanobacteria: Two different photosystems exist. In Figure 3.12-1 only the cofactors of reaction center monomers are shown in highly schematized form.

The reaction center of the water-splitting and oxygen-releasing photosystem II (PSII) is formed by the proteins D1 and D2, which carry together six chlorophylls (four of them forming a cluster, P_{D1}, P_{D2}, Chl_{D1}, Chl_{D2} as the 'primary donor' P680), two pheophytins (chlorophylls with two protons instead of Mg^{++}), two plastoquinones (phylloquinones in cyanobacteria) and two carotinoids. These are almost symmetrically arranged. A tyrosine moiety (Tyr_Z), which is involved in the electron transfer chain, is part of the D1 protein, another one (Tyr_D) belongs to the D2 protein. A Mn_4Ca-cluster is the essential component of the oxygen evolving complex (OEC). Two of these reaction centers are arranged as a dimer, which is surrounded by the intrinsic light-harvesting CP47 and CP43 proteins and by a large number of peripheral light-harvesting proteins, forming the light harvesting 'antenna' complex (LHC).

In green plants, the pigments of the LHC are chlorophyll (3.3.4), β-carotenes, and xanthophylls (3.5.3.2). In the 'xanthophyll cycle',

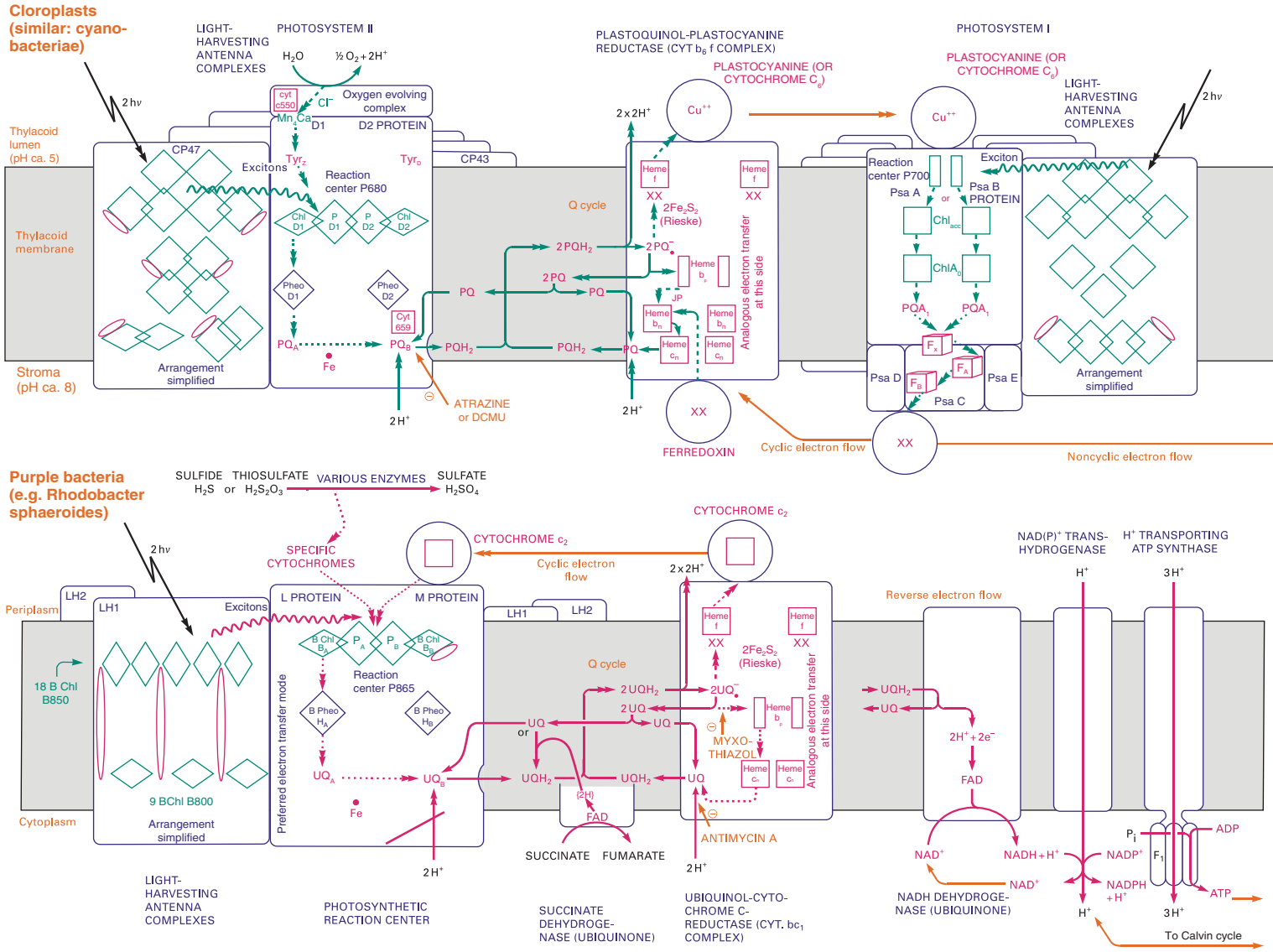


Figure 3.12-1. Photosynthetic Systems in Green Plants and Cyanobacteria (Top); in Purple Bacteria (Below)
The structures (in particular of the light-harvesting complexes) have been simplified.

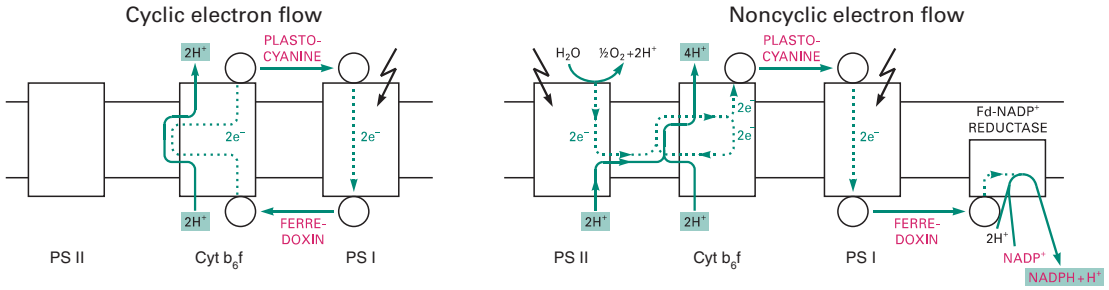


Figure 3.12-2. Electron Flow in Green Plants and in Cyanobacteria

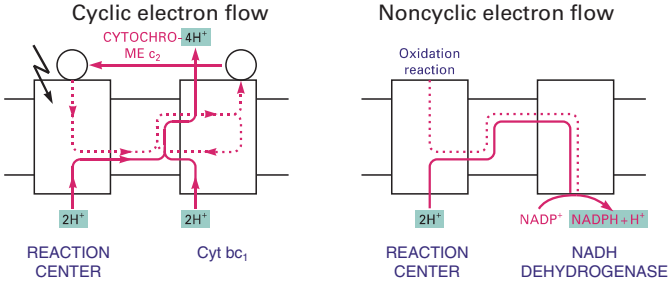
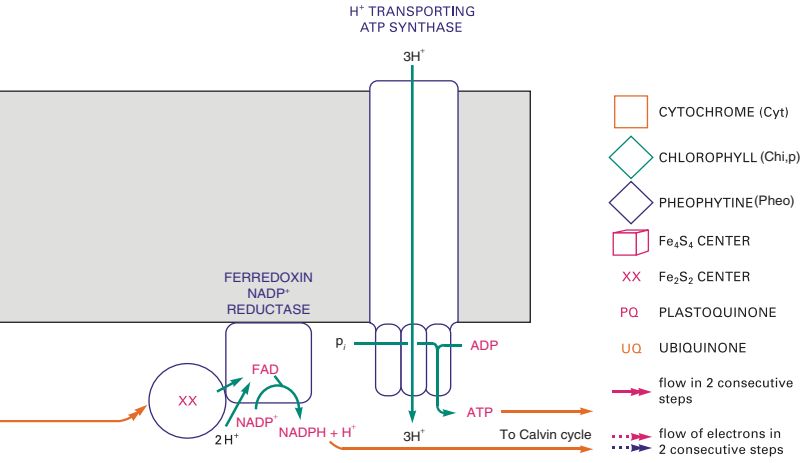


Figure 3.12-3. Electron Flow in Purple Bacteria

violaxanthin can be interconverted into zeaxanthin and back, depending on the light intensity (Fig. 3.5.3-2). While the former compound does not accept energy from excited states of chlorophylls, the latter is open for this energy transfer and dissipates the energy into heat via a short- living excited state. This results in a protective role by eliminating the dangerous triplet state of chlorophyll ($^3Chl^*$) at high light intensity, which could give rise to singlet oxygen (3.2.5.8). Cyanobacteria and red algae additionally use phycobilisomes as the major light- harvesting complexes. These are large rod-shaped, membrane-attached antenna complexes, which contain phycocyanobilin, phycoerythrobilin and other pigments (3.3.3). While chlorophylls a and b absorb in the blue and red regions, these pigments fill the 'green gap' (Fig. 3.12-4).



The structure of PS II has been resolved in high resolution in cyanobacteria, however, the photosystem in higher plants appears to be closely related.

Photosystem I (PSI) can be considered to be a light-driven plastocyanin-ferredoxin oxidoreductase. The main proteins PsaA and PsaB carry the components of the electron transfer chain in pseudo-symmetric C_2 fashion. They consist of a pair of chlorophylls a (eC-A1 and eC-B1, likely representing the primary donor P700), associated with two more pairs of chlorophylls a (eC-A2, eC-B2, eC-A3 and eC-B3, also named chl_{acc} and A_0 pairs), two phylloquinones (Q_K-A and Q_K-B, also named A_1 pair) and a central [4Fe-4S] cluster F_X. A pair of [4Fe-4S] clusters F_A and F_B is bound to the protein PsaC. Additionally, docking sites exist for ferredoxin (or flavodoxin) at the stromal surface and for plastocyanin (or cytochrome c_6 in cyanobacteria) at the luminal surface. The basic structures of the plant and the cyanobacterial PSI are closely related, however the plant system is monomeric and the bacterial one is mostly trimeric. As a core antenna in green plants, 79 chlorophylls are tightly coordinated by PsaA/PsaB for fast energy transfer, surrounded by more chlorophylls, β-carotenes and xanthophylls (3.5.3.2). Again, cyanobacteria have phycobilisomes attached to the PS core.

The composition of the antenna complexes is listed in Table 3.12-1, the absorption spectra are given in Figure 3.12-4.

Cytochrome b_6f Complex: This complex provides the electronic connection between PSII and PSI. In connection with the quinone pool, it provides proton translocation from the stromal to the thylacoid (luminal) side. In plants and cyanobacteria, it is a symmetrical dimer of a Rieske [2Fe-2S] protein, a cytochrome f (containing heme f), a cytochrome b_6 polypeptide (containing hemes b_p, b_n c_n), subunit IV and some minor proteins.

Structure of the photosystem in purple bacteria: There is only one photosystem, which resembles the photosytem II of plants and cyanobacteria and shows a twofold symmetry as well. The reaction center is a cluster of four bacteriochlorophylls (two of them closely associated = P 865). Two bacteriophaeophytins take the place of the phaeophytins and two ubiquinones take the place of the plastoquinones. Most purple bacteria have two antenna complexes containing bacteriochlorophylls and carotenoids. LH1 as core complex forms a tight ring around the reaction center, while several LH2 rings are arranged around this core. The cytochrome bc_1 complex is closely related to the complex III of the mitochondrial chain (3.11.1.3). It lacks the heme c_n present in the b_6 f complex.

Light absorption step: An absorbed light quantum excites an electron in one of the LHC molecules, which transfers its energy ('exciton') by resonance interaction via other LHC molecules quickly (ca.10^{-13} sec, > 90% efficiency) to the reaction center. In photosytem II of plants and cyanobacteria, or in the only reaction center of purple bacteria, it excites a pigment in the cluster of four closely associated chlorophylls (P680 → P680* or P865 → P865*, respectively).This pigment, in turn, donates an electron extremely quickly to a primary acceptor (pheophytine,

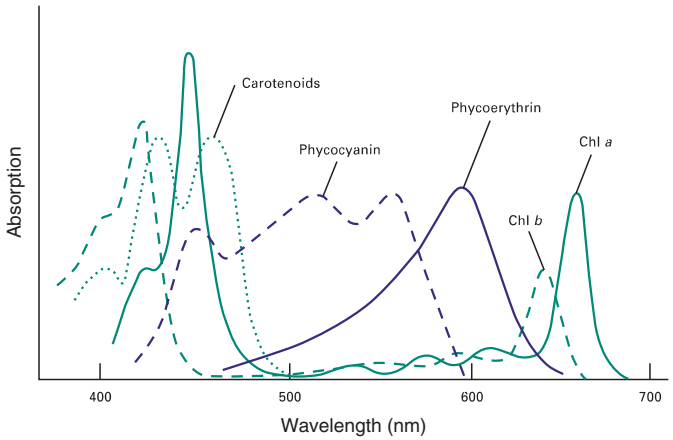


Figure 3.12-4. Absorption Spectra of Light Absorbing Chromophores (Line colors: green-plants, blue-cyanobacteria).

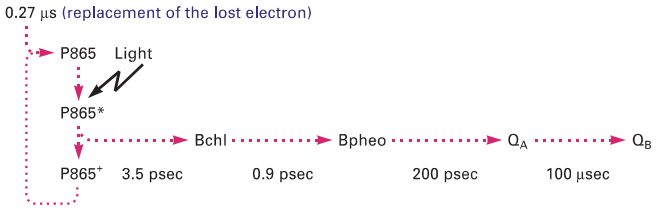


Figure 3.12-5. Time Course of Electron Transfer in Purple Bacteria

Table 3.12-1. Cofactors of the Light Harvesting Complexes (LHC)

Purple bacteria	Plants	Cyanobacteria (Blue Algae)
(1 reaction center)	(2 photosystems)	
32 bacteriochlorophylls, 16 carotenes	Ph. Sys. I: ca. 200 Chl., a > b 50 carotenoids	phycocyanobilin phycoerythrobilin
	Ph. Sys. II: ca. 250 Chl., a > b 110 carotenoids	phycoviolobilin phycouvobilin

3.3.4 or bacteriochlorphyll, 3.3.4), causing the reaction to become increasingly irreversible. Via the quinones PQ_A or UQ_A, the electron finally reaches phylloquinone B (PQ_B, 3.2.7.2) or ubiquinone B (UQ_B, 3.2.7.2) respectively, where two electrons and two protons (from the cytoplasm) accumulate, forming a hydroquinone (quinol, Fig. 3.12-5). In photosystem II only the D1- side is operative, while in photosystem I of plants and of cyanobacteria both branches may contribute to electron transfer.

Regeneration of the reaction center: In plants and cyanobacteria, $P680^+$ replaces the lost electron by abstraction of another electron from the Mn_4Ca-protein complex (oxygen evolving complex, OEC) via a tyrosine residue, Tyr_z. After four repetitions, OEC^{4+} reacts with water and is reduced again.

$$OEC^{4+} + 2\ H_2O = OEC^0 + 4\ H^+ + O_2$$

In purple bacteria, in the case of 'cyclic electron flow', the lost electron of $P865^+$ (the special pair) is returned from the cytochrome bc_1 complex via diffusing cytochrome c_2. No extra reducing power for other purposes becomes available in this way. In case of 'noncyclic electron flow' in these bacteria, an oxidation reaction (of H_2S, S, $H_2S_2O_3$, succinate etc.) takes place:

$H_2S = S_{solid} + 2\ H^+ + 2\ e^-$ (in periplasm)
or succinate = fumarate + $2\ H^+ + 2\ e^-$ (in cytoplasm).

The liberated electrons enter the reaction center via a bound cytochrome complex (e.g. in *Rhodopseudomonas viridis*, 0.27 µsec) or via soluble cytochrome c_2 (e.g., in *Rh. sphaeroides*, µsec to msec) and reduce the special pair again.

Cytochrome b_6f and bc_1 complexes: The hydroquinone (quinol) formed in the primary photosynthetic reaction transfers its hydrogen via the 'quinone pool' to the cytochrome complexes b_6f (in plants) or bc_1 (in bacteria), where protons are released to the thylakoid space or to the periplasm,

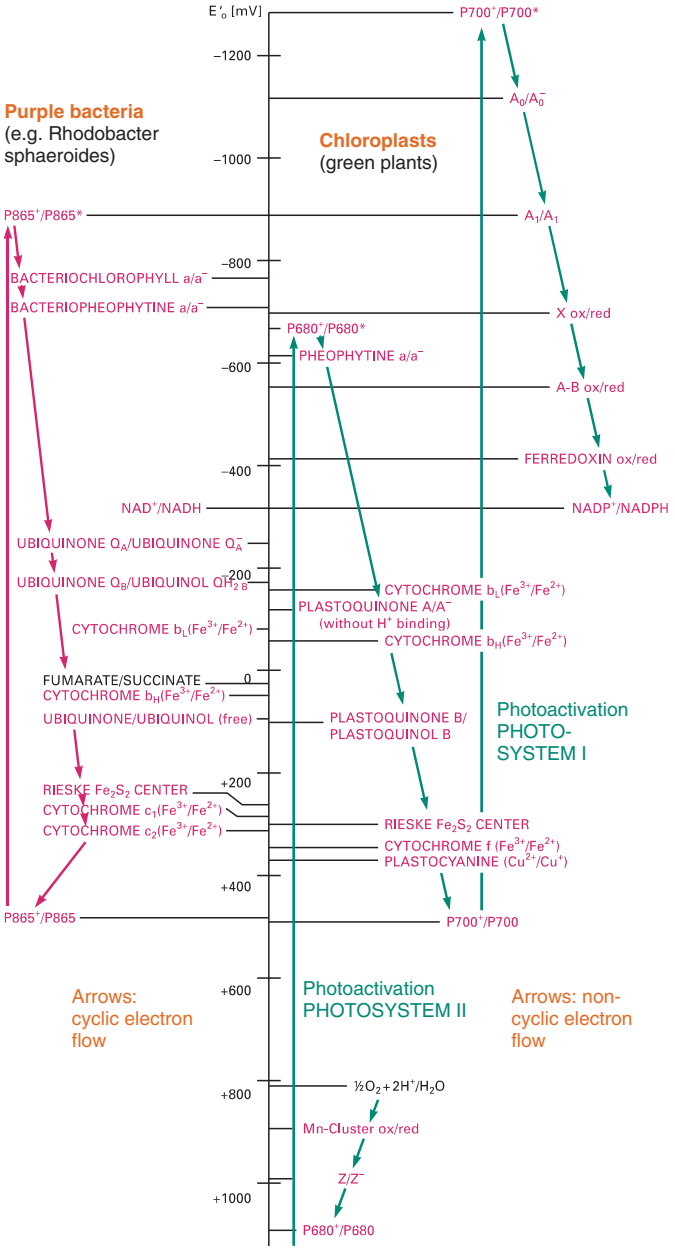


Figure 3.12-6. Standard Redox Potentials in Photosynthesis (Purple Bacteria/Plants and Cyanobacteria)
In vivo, the actual potentials can differ due to protein binding, variant concentration ratios, etc.

respectively. These complexes closely resemble the mitochondrial ubiquinol-cytochrome c reductase (complex III). Correspondingly, a 'Q cycle' operates for transfer of additional protons to the thylakoid space or to the periplasm, respectively. For details, see 3.11.4.3. The corresponding electrons are finally transferred to photosystem I (in plants via plastocyanine, in cyanobacteria via cytochrome c_6) or returned to the reaction center (in purple bacteria: cyclic electron flow via cytochrome c_2).

NAD^+ or $NADP^+$ reduction: In plants and cyanobacteria, illumination excites the primary donor P700 in photosystem I to release an electron to the primary acceptor chlorophyll A_0 (the role of the chlorophyll$_{acc}$ is unclear). Then it is transferred to phylloquinone A_1 and further on to the iron-sulfur cluster F_x. This electron transfer proceeds either through the cofactor sequence bound to the protein PsaA or to the ones bound to PsaB. From F_x, the electron reaches the iron-sulfur clusters F_A and F_B, which are bound to the peptide PsaC. These clusters release 2 electrons to the two [4Fe-4S] clusters in ferredoxin (or to the FMN in flavodoxin). These are then conferred either to the $NADP^+$ reductase (noncyclic electron flow), or alternatively back to the cytochrome b_6f-complex

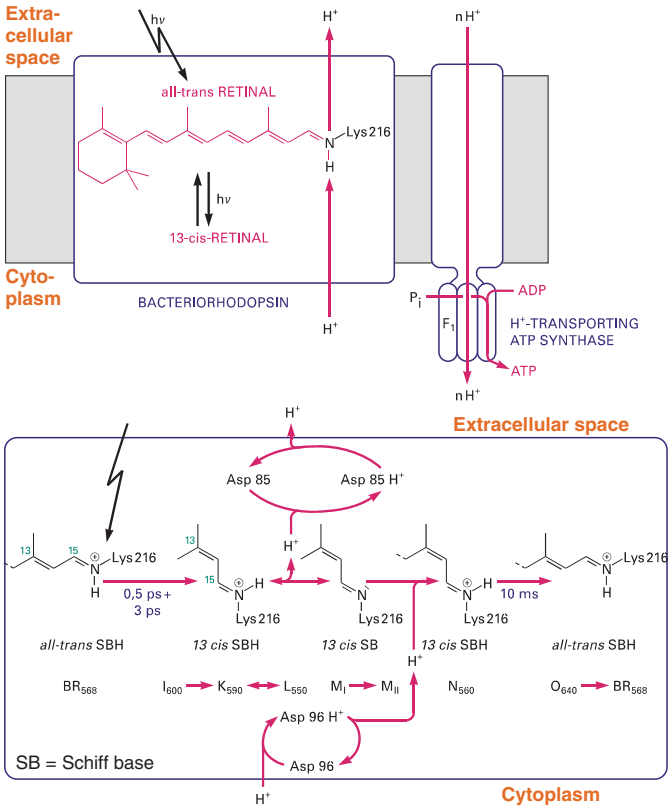


Figure 3.12-7. Photosynthesis and Reaction Mechanism in Halophilic Archeaea

for additional proton transfer (cyclic electron flow). This allows a fine adaptation to the requirements of the cell, since NADPH reduction equivalents or ATP energy can be supplied in variable ratios. The graph of the reduction potential of the steps passed through resembles a 'Z' (Fig. 3.12-6, for details of the redox potentials see 3.11.4).

As described above, purple bacteria cannot follow this mechanism. They have to obtain reducing power from the environment to be able to reduce $NADP^+$ (noncyclic = reverse electron flow, since the electrons have to flow 'uphill' of the redox potential).

Halophilic archaea (Fig. 3.12-7): The photosystem of these archaea is unrelated to photosynthesis in higher plants. It uses bacteriorhodopsin, a small retinal protein (26 kDa) with 7 transmembrane passes, which pumps protons upon absorption of photons through the membrane, quantum yield $\psi = 0.65$.

It is mediated by light- induced *trans-cis* isomerization of the retinyliden chromophore and involves the following steps:

- Isomerization of retinal from the *all-trans* to the 13 *cis*-configuration [BR_{568} to J state (0.5 psec) and on to K and L states]
- Transfer of a proton from the protonated Schiff base (SBH) to the carboxylate of Asp85 (L to M_I states), followed by its release to the extracellular medium
- Modification of chromophore/protein structure. This changes the accessibility from the extracellular side to accessibility from the cytoplasmatic side (M_I to M_{II} states)
- Transfer of a proton from Asp 96 to the Schiff base (M to N state, several msec)
- Thermal *cis-trans* reisomerization (N to O state, several msec)
- Restoration of the initial state (O to BR_{568} state).

Isomerization of retinal (11-*cis* ↔ *all-trans*) also plays a role in the visual process of vertebrates (7.4.6).

Literature:

Barros, T. *et al*. EMBO J. 2009;28:298–306.
Cramer, W.A., Zhang, H. Biochim. biophys. Acta 2006;1757:339–345.

Fromme, P. *Photosynthetic Protein Complexes. A Structural Approach.* Wiley-VCh Verlag, 2008 (Very detailed survey).

Guerkova-Kuras, M. *et al*. Proc.Nat. Acad.Sci. (USA) 2001;98:4437–4442.

Haupts, U. *et al.* Biochemistry 1997;36:2–7.

Holzwarth, A.R. *et al.* Proc. Natl. Acad. Sci. (USA) 2006;103: 6895–6900.

Jansson, S. Biochim. Biophys. Acta 1994;1184:1–19.

Loll, B. *et al.* Nature 2005;438:1040–1044.

Stroebel, D. *et al*. Nature 2003;426:413–418.

3.12.2 Dark Reactions

As described above, the light reactions provide both the energy carrier ATP and the reductant NADPH. For the consecutive synthesis of biological material (initially carbohydrates), CO_2 and water are also required.

Calvin cycle (Fig. 3.12-8): CO_2 fixation takes place in a cyclic process within the stroma by carboxylation of ribulose 1,5-diphosphate and concomitant cleavage into two 3-phosphoglycerate molecules. This is followed by phosphorylation and reduction reactions. Then an aldol condensation and a series of transfer reactions takes place, mostly using reactions closely related to the pentose phosphate cycle (3.1.6.1). As a result, the carboxylation of 6 C_5 molecules yields 1 C_6 molecule (glucose-P or fructose-P) and the reconstitution of the original 6 C_5 molecules:

$$6\,C_5 + 6\,CO_2 \rightarrow 6\,C_5 + 1\,C_6$$

according to the overall reaction of the Calvin cycle

$$6\,CO_2 + 12\,H_2O + 18\,ATP^{4-} + 12\,NADPH = C_6H_{12}O_6 + 18\,ADP^{3-} + 18\,P_i^{2-} + 12\,NADP^{+} + 6\,H^{+}.$$

The produced hexose is converted in chloroplasts into starch (3.1.2.2) or in the cytosol into sucrose (3.1.4.1).

The enzyme ribulose bisphosphate carboxylase/oxygenase (Rubisco) catalyzes the key reaction of the Calvin cycle

$$\text{Ribulose bisphosphate} + CO_2 + H_2O = 2\ \text{3-phospho-D-glycerate}$$
$$\Delta G'_0 = -35{,}1\ \text{kJ/mol}.$$

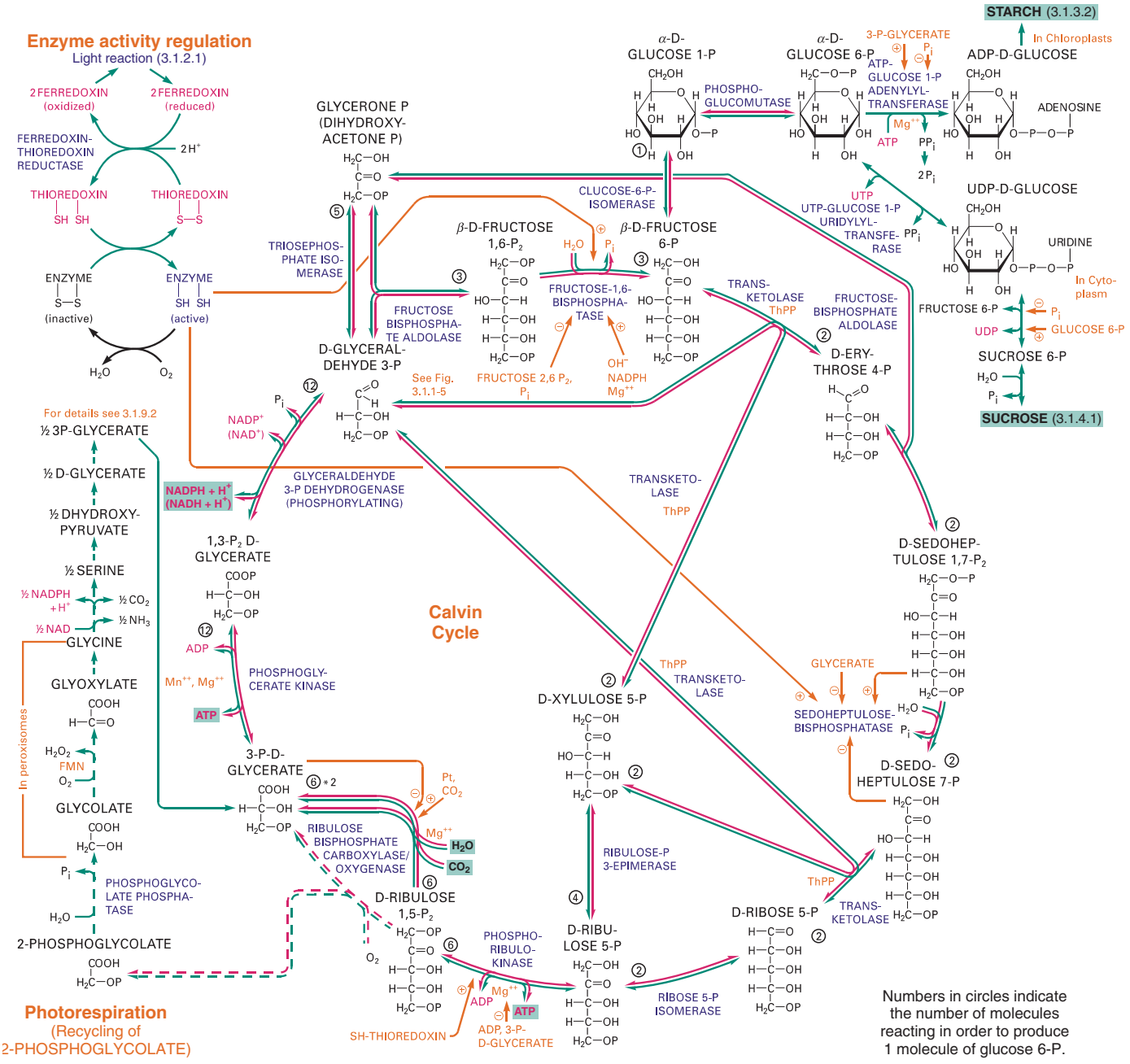


Figure 3.12-8. CO_2 Fixation by the Calvin Cycle and its Regulation

The enzyme is apparently the most abundant enzyme in the biosphere. It consists of eight large and eight small subunits (51 … 58 and 12 … 18 kDa). It has a low catalytic efficiency (kcat = 3.3 sec^{-1} per large subunit). Although the carboxylase reaction is usually preferred, it also performs an oxygenase side reaction (Fig. 3.12-9, see also 'photorespiration' below).

Regulation of the Calvin cycle: The cycle has to operate only if sufficient NADPH and ATP from the light reaction are available in order to prevent useless degradation reactions. This is performed by light-induced activation of rubisco, fructose bisphosphatase (FBPase) and sedoheptulose bisphosphatase (SBPase).

- The pH in the stroma increases during the light reaction (3.12.1), since protons are pumped out. It approaches the pH optimum of rubisco, FBPase and SBPase.
- Reduced ferredoxin, the reaction product of photosystem I, reduces thioredoxin, which in turn activates FBPase and SBPase by reduction of enzyme -SS- bridges (Fig. 3.12-5). Simultaneously, phosphofructokinase (3.1.1.2) is deactivated by this reduction and thus decreases the competing glycolysis reaction (3.1.1.1).
- Mg^{++}, which flows into the stroma during illumination, activates rubisco, FBPase and SBPase.
- NADPH, which is produced by the light reaction, activates FBPase and SBPase.

During dark, these reactions are switched off. The energy supply of photosynthesizing cells is then provided the same way as in non-photosynthesizing cells by glycolysis (3.1.1.1), pentose phosphate cycle (3.1.6.1) and oxidative phosphorylation (3.11).

Photorespiration and C_4 cycle: The rubisco side reaction with O_2 yields at first 3-phosphoglycerate and 2-phosphoglycolate, which later on is partially oxidized, resulting in CO_2 liberation (photorespiration, Figure 3.12-8, see also 3.1.9.2). This counteracts photosynthesis and requires additional energy input for recycling. The rate of this reaction increases relatively to the rate with CO_2 at higher temperatures and at low CO_2 concentration at the site of synthesis (e.g., on hot, bright days), and limits the growth rate of plants.

A number of plants (C_4 plants, mostly tropical ones) have developed a mechanism for increasing the CO_2 concentration in the fluid phase of chloroplasts from ca. 5 µmol/l to ca. 70 µmol/l (Fig. 3.12-10). So-called mesophyll cells surround the bundle-sheath cells, which contain the Calvin cycle enzymes. The mesophyll cells, which lack

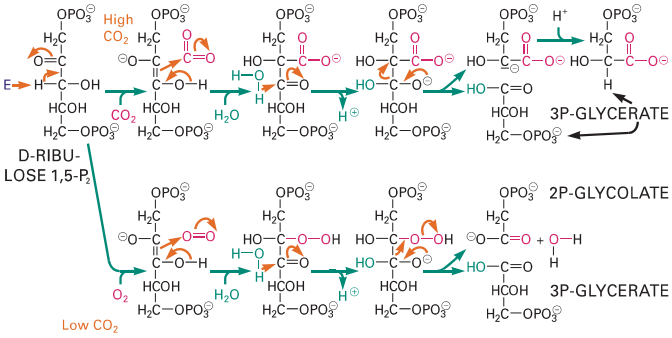


Figure 3.12-9. Carboxylase (Top) and Oxygenase (Below) Reaction Mechanisms of Rubisco

rubisco, perform a CO_2 fixation by the highly exergonic (and thus practically irreversible) reaction (3.1.3.4):

$$\text{Phosphoenolpyruvate} + HCO_3^- \rightarrow \text{oxaloacetate} + P_i$$

and transfer this bound CO_2 through a number of further reactions to the chloroplasts of bundle-sheath cells, where it is released to be used in the Calvin cycle. Several reaction types exist (Fig. 3.12-10). These reactions require five energy-rich P bonds/ CO_2 (instead of three in the Calvin cycle). Therefore, this mechanism is of advantage only in hot, sunny climates.

Literature:

Furbank, R.T., Taylor, W.C. The Plant Cell 1995;7:797–807.

Gutteridge, S., Gatenby, A. The Plant Cell 1995;7:809–819.

Heldt, H.W., Flügge, U.I. in Tobin, A.K. *Plant Organelles*. Cambridge University Press, 1992.

Heldt, H.W. *Plant Brochemistry and Molecular Biology.* Oxford University Press, (1998).

Portis, A.R. Ann. Rev. Plant Physiol. Plant Mol. Biol. 1992;43:415–437.

3.13 Plant Secondary Metabolism

Antje Chang

Plant metabolism can be divided into primary and secondary metabolism. The term primary metabolism encompasses all processes and compounds that are essential for the fundamental functions of life,

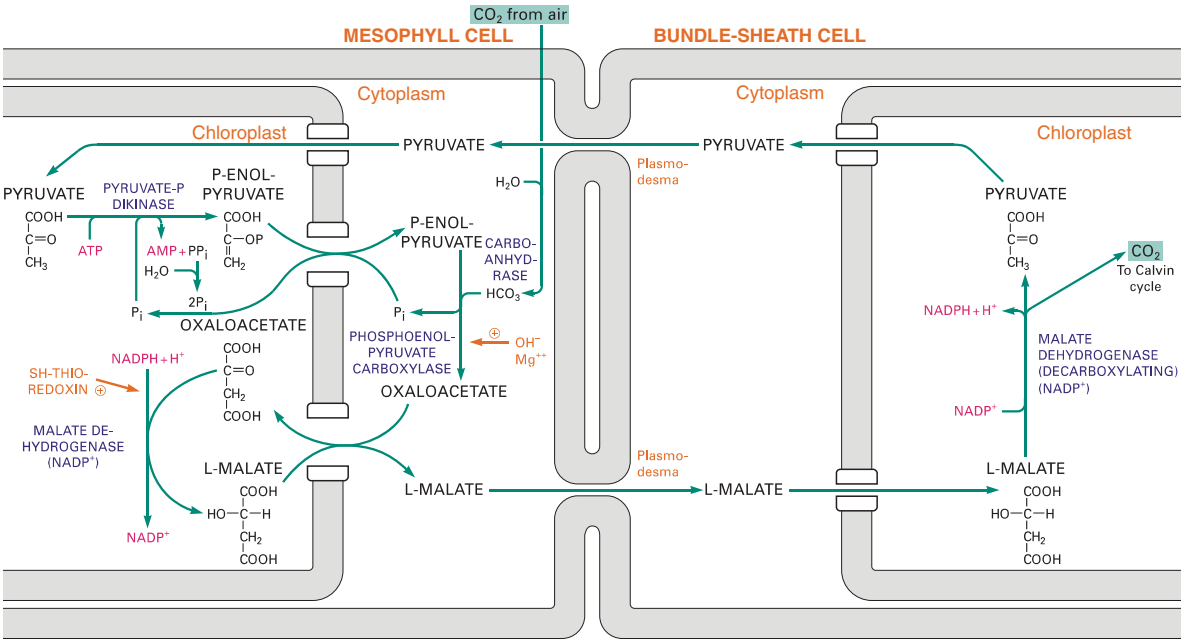


Figure 3.12-10. CO_2 Pumping by the C_4 Cycle (NADP$^+$-Malate Enzyme Type, e.g., in Maize and Sugar Cane)

like growth, development, and reproduction. In contrast, secondary metabolism which is characterized by its immense chemical diversity, is required for the survival of the individual in its respective environment. Therefore, these natural products, traditionally referred to as secondary metabolites have an ecological function for the organism in its interaction with its biotic and abiotic environment. Their role had been overlooked for a long time, but is widely accepted now.

The functions, which in general can be regarded as the plant's chemical interaction, are studied in the field of so-called chemical ecology, considering the following aspects:

- Chemical defense (constitutive or induced defense against pathogens and herbivores). Plants have developed different strategies for the defense against herbivores and pathogens:
  - The bioactive compounds are synthesized constitutively and accumulated in specialized cells (e.g., hair) or in subcellular compartments (e.g., vacuole), and are released by plant tissue destruction.
  - Non-toxic precursors (e.g., glycosylated precursor of toxic aglycons) are stored apart from the corresponding specific enzyme, e.g., a glycosidase. After destruction of the cell compartments the enzymatic reaction is initiated and the toxic aglycone is released.
  - The formation of defensive compounds, e.g., phytoalexins and proteinase inhibitors, may be induced by signal substances (elicitors) as a response to the attack by pathogens (e.g., by phytoalexins) and herbivores (e.g., by proteinase inhibitors).
- Attraction of pollinators and seed distributors (flower pigments, volatile compounds).
- Adaptation to the environment (e.g., UV protection).

Secondary metabolism is not only found in plants, but also in bacteria (e.g., antibiotics 3.10.9), fungi and marine sessile organisms. This chapter will focus on the plant secondary compounds, since 80 % of the secondary metabolites are produced by higher plants. Many of these reactions originate from pathways of the primary metabolism, therefore only the differing parts are described here and references are given for the common reactions. The biosynthetic origins of the secondary metabolites are also often used as base for their classification (Table 3.13-1).

Table 3.13-1. Major Groups of Plant Secondary Metabolites

Classes of secondary metabolites	derived from
Phenolic compounds: polyphenols, phenols, phenylpropane derivatives, flavonoids, stilbenes	shikimic acid, phenylalanine, polyketide
Terpenoids/isoprenoids: hemiterpenes, monoterpenes, sesquiterpenes, diterpenes, triterpenes, tetraterpenes, polyterpenes	C_5-unit ('activated isoprene')
Pseudo-alkaloids: terpenoid alkaloids, piperidine alkaloids	terpenes, polyketides, acetate
Alkaloids: Nicotiana alkaloids, pyrrolizidine alkaloids, tropane alkaloids, benzylisoquinoline alkaloids, indole alkaloids, purine alkaloids	amino acids

3.13.1 Phenolics

Unlike animals, plants, fungi, and bacteria are able to perform the *de novo* biosynthesis of aromatic metabolites. In higher plants most of the phenolics are formed by the shikimate pathway with aromatic amino acids as intermediates (3.2.7.1). Another major pathway leading to aromatic natural products is the polyketide pathway, which proceeds via linear coupling of acetate units. Flavonoids are an example of mixed biosynthesis of aromatic metabolites (3.13.1.3).

3.13.1.1 Biosynthesis

Shikimate pathway: The biosynthesis of the three aromatic amino acids L-phenyalanine, L-tyrosine, and L-tryptophan by the shikimate pathway is described in detail in 3.2.7.1 and Figure 3.2.7-1. The pathway is localized in plastids of plants and in the cytoplasm of bacteria and fungi. Originating from D-erythrose 4-phosphate and phosphoenolpyruvate,

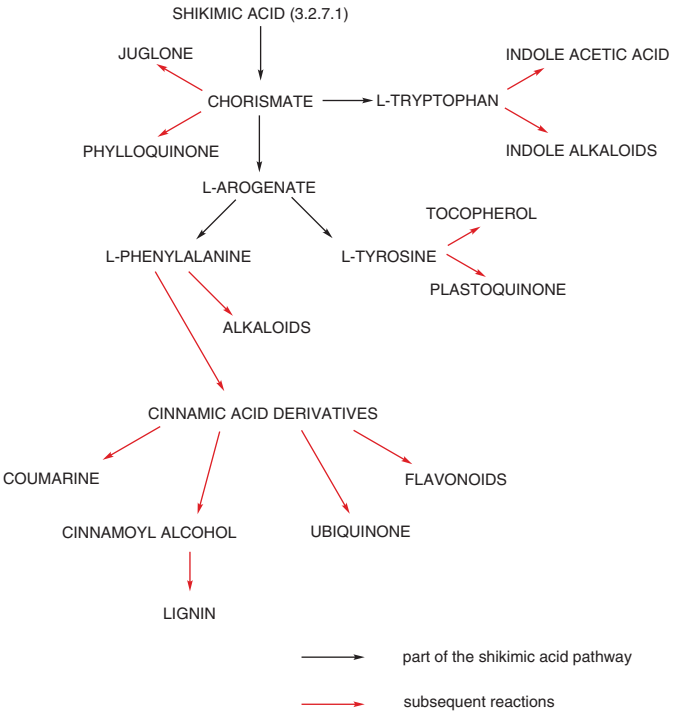


Figure 3.13-1. Products Produced by the Shikimate Pathway

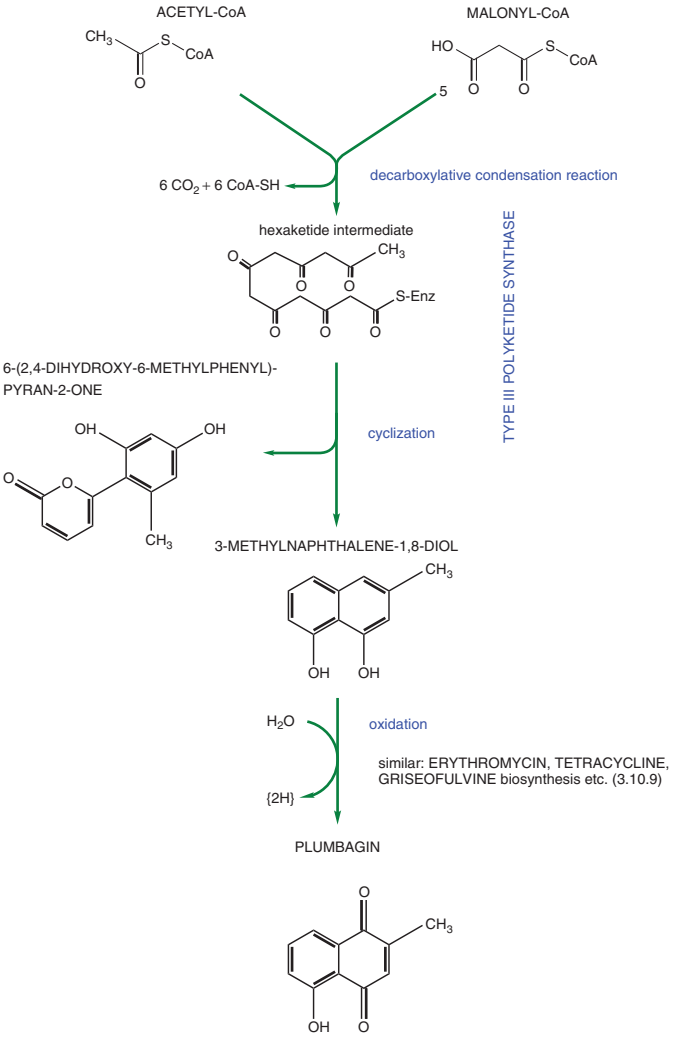


Figure 3.13-2. Polyketide Pathway (biosynthesis of plumbagin, putative reaction in *Plumbago indica*)

the pathway includes shikimate, chorismate and prephenate as intermediates. Contrary to the general pathway, part of the sequence is reversed in higher plants: prephenate is first transaminated to arogenate, the dehydratase/decarboxylase or dehydrogenase/decarboxylase reactions take place afterwards (arogenate pathway). These aromatic amino acids are precursors of numerous aromatic compounds in bacteria, fungi, and plants. A survey of these interrelationships is given in Figure 3.13-1.

Polyketide pathway (Fig. 3.13-2): Polyketides are natural products found mainly in bacteria and fungi, but also in plants and animals. They are synthesized by linear condensation reactions of acetate units, deriving from malonyl-CoA via decarboxylation. This is a process similar to fatty acid biosynthesis (3.4.1.1). The polyketide synthases are multi-enzyme complexes that produce a wide range of structural diverse secondary metabolites, also depending on the kind of starter molecule. In plants, the polyketide pathway is involved in mixed biosyntheses, like in the biosynthesis of flavonoids (3.13.1.3) and stilbenes (3.13.1.4), where a phenylpropane is the starter molecule. Several type III polyketide synthases are known in plants, such as chalcone synthase or stilbene synthase. Related reactions are found in the biosynthesis of, e.g., erythromycin (3.10.9.3), tetracycline (3.10.9.4) and other antibiotics.

3.13.1.2 Phenylpropane Derivatives (Fig. 3.13-3)

Phenylpropanes encompass a broad range of plant secondary metabolites. They are mainly synthesized from phenylalanine. Phenylalanine ammonia lyase (PAL) is a key enzyme between the primary and secondary metabolism, producing *trans*-cinnamate by release of ammonia. The activity of PAL is influenced by light and temperature and is regulated by feedback inhibition.

trans-Cinnamate is a central intermediate for a wide range of derivatives (Table 3.13-2, Fig. 3.13-4). They are synthesized mainly by hydroxylation and methylation reactions catalyzed by specific enzymes. Examples are phenylpropanoids, i.e., eugenol, anethol, and estragol, which are major constituents of essential oils. The corresponding alcohols (4-coumarol, sinapol, coniferol, ferulol) are formed by reduction of carboxylic groups and represent the monomeric components of lignin (monolignol).

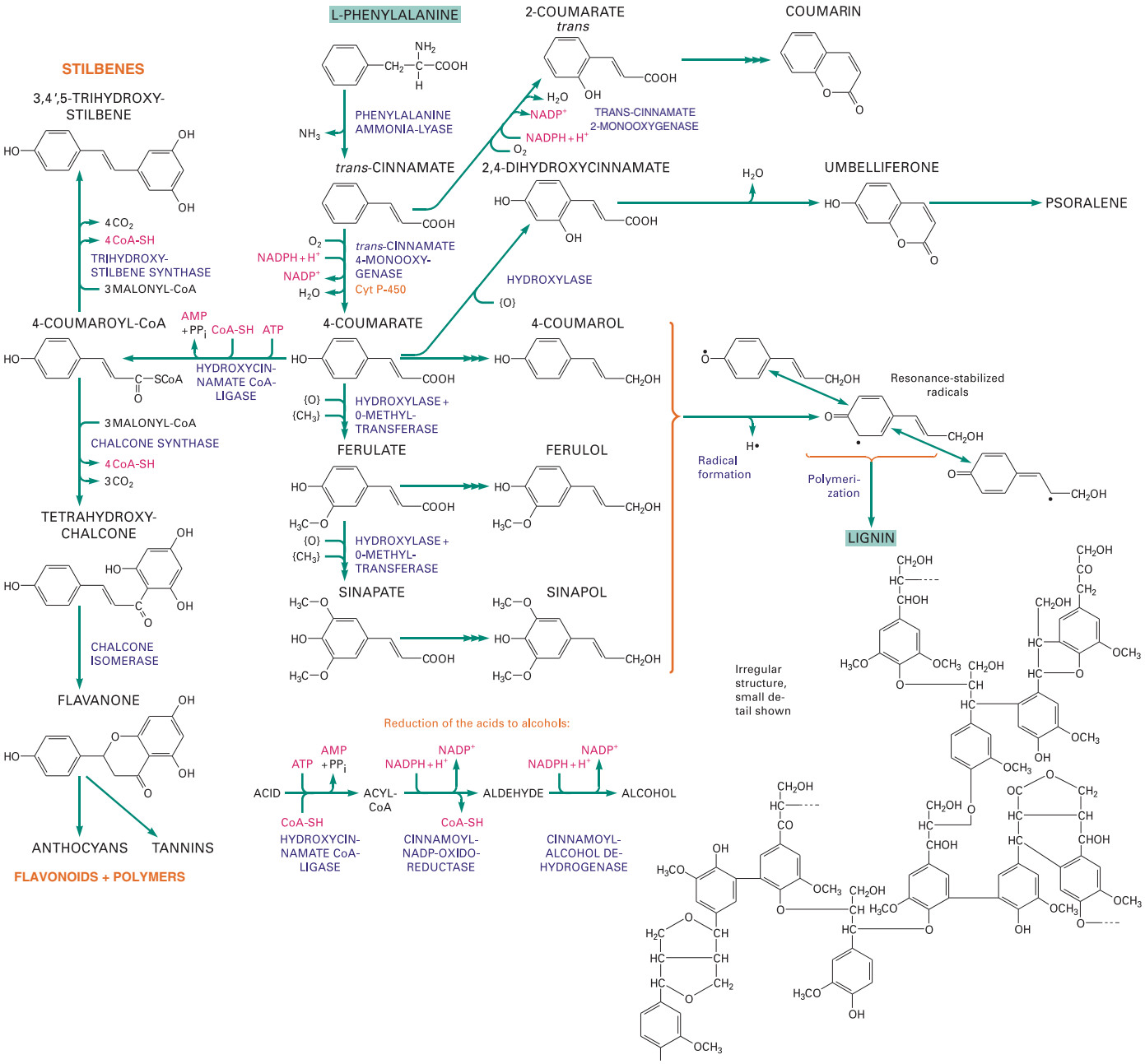


Figure 3.13-3. Phenylpropanoid Compounds in Plants

The polymerization reaction leading to lignin in the cell walls of the plants is catalyzed by lignin peroxidase (Fig. 3.13-3). The extracellular process is initiated by the formation of a radical, presumably by H_2O_2 (3.2.5.8) and progresses via chain reaction mechanisms. The result is a closely meshed, irregular network. Its overall composition depends on the ratio of the originating alcohols and the reaction conditions and varies among different species. Lignin is the second most frequent compound in the biosphere (after cellulose, the annual synthesis rate is ca. $2 * 10^{10}$ t). It brings about the pressure resistance of plant cell walls (3.1.6.3). Only a few organisms, mostly fungi, can degrade lignin. Suberin has a similar structure with alcoholic groups esterified by (mostly) long-chain fatty acids. It occurs in cork, the endodermal cells of roots and other parts of plants.

The pathway to coumarin starts with hydroxylation of *trans*-cinnamate, resulting in *trans*-2-coumarate (Fig. 3.13-3). The product accumulates in the vacuole of the mesophyll cells in the form of glucosylated *cis*- and *trans*-isomers. When the plants are wounded, a specific glucosidase in the cytoplasm catalyzes the hydrolysis of the *cis*-isomer, producing coumarin by lactonization.

Coumarin is a toxin found in many plants, e.g. in woodruff (*Galium odoratum*) or tonga bean (*Dipteryx odorata,* common name: cumaru). Coumarin derivatives have been used in the perfume industry. They are important in pharmacology due to their anticoagulant effect and likewise as rat poison, causing internal hemorrhage and death (e.g., Warfarin®).

Figure 3.13-4. *Trans*-Cinnamate Derivatives

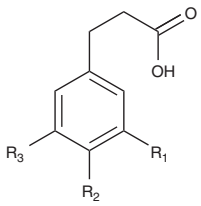


Table 3.13-2. Some *Trans*-Cinnamate Derivatives

	R_1	R_2	R_3
3-Coumarate	OH	H	H
4-Coumarate	H	OH	H
Caffeate	OH	OH	H
Ferulate	OCH_3	OH	H
Sinapate	OCH_3	OH	OCH_3

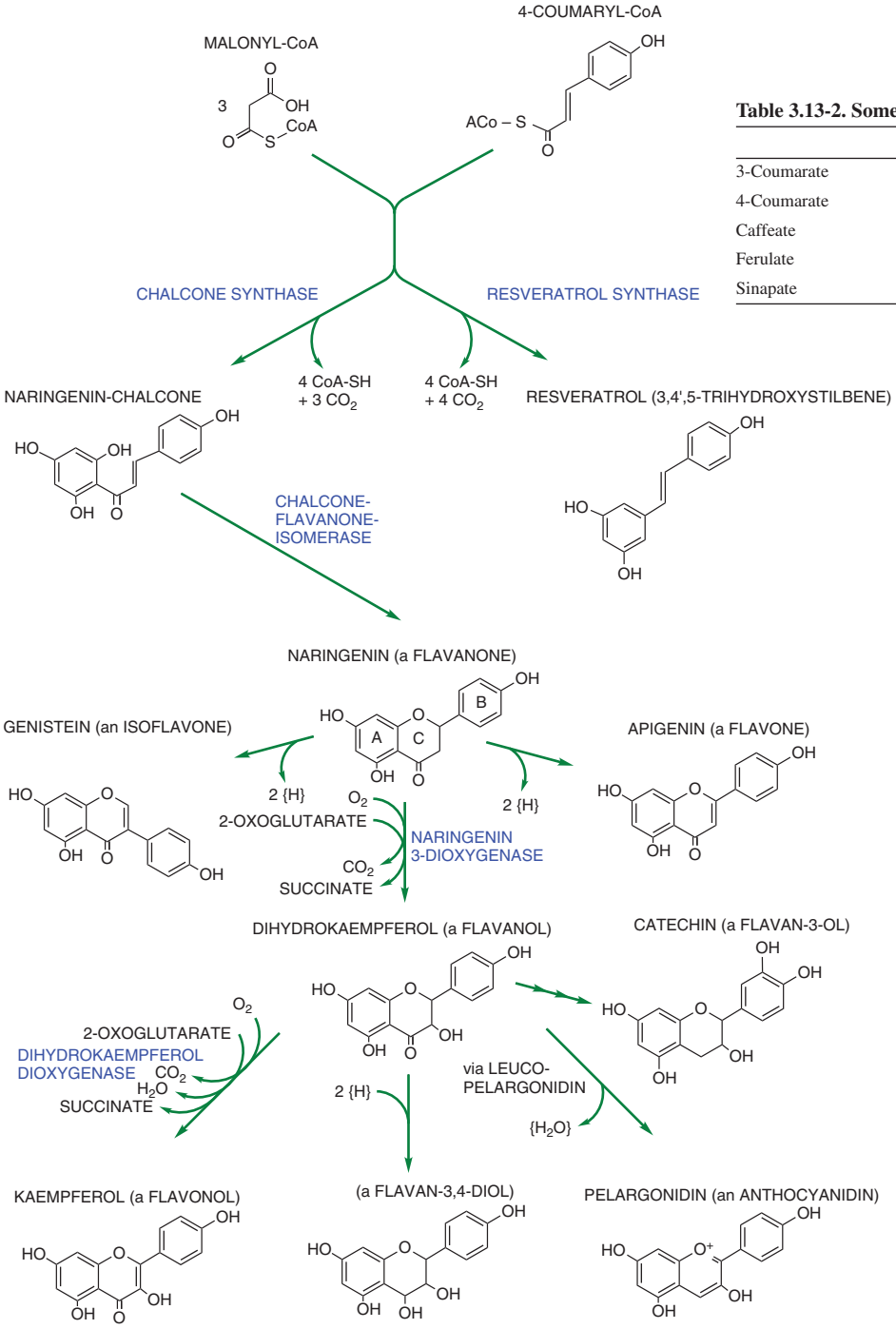


Figure 3.13-5. Biosynthesis of Flavonoids and Stilbenes

3.13.1.3 Flavonoids

The flavonoids are a large group of plant secondary metabolites. They display a great variety in structure and function and are widely distributed in the plant kingdom.

The biosynthesis (Fig. 3.13-5) combines the products of the shikimate pathway and of the polyketide pathway (3.13.1.1). 4-Coumaroyl-CoA ligase activates 4-coumarate to its CoA derivative. Thereafter, chalcone synthase catalyzes the addition of three malonyl-CoA units (originating from the polyketide pathway) and removal of 3 CO_2 to naringenine chalcone, forming the flavan backbone that is characteristic of all flavonoids. These compounds can be assigned to several subgroups depending on the substitution pattern, as listed in Table 3.13-3. Some flavonoid structures are shown in Figure 3.13-6.

Table 3.13-3. Subgroups of Flavonoids

Flavonoid subgroup	Examples	Source
Flavanone	hesperetin, naringenin, eriodictyol	grapefruit, orange
Flavone	luteolin, apigenin, tangeritin	pepper, celery
Flavonol	quercetin, rutin, kaempferol, myricetin	onion, endive
Flavanol	catechin, gallocatechin, epicatechin, theaflavin	red grape, apple, green tea
Flavanonol	taxiflorin, dihydrokaempferol	gingko
Isoflavone	genistein, daidzein, licoricidin	soybean
Anthocyanidin	cyanidin, delphinidin, malvidin, pelargonidin, peonidin, petunidin	cherry, blueberry, red grape

Flavonoids accumulate in cell vacuoles, mostly in their glycosylated form. Many color pigments in flowers and fruits serve as attractants of pollinators and animals for seed distribution. Anthocyanins, the glycosides of anthocyanidines, are water-soluble vacuolar pigments. Their colors depend on the substitution patterns of the B-ring, the pH-value in the vacuole, the binding of metal ions etc.

Flavonoids in the epidermis serve as UV-protection for the inner cell layers, e.g,. the mesophyll cells. These compounds play an important role in the interaction of rhizobia and plants. They act as plant signals activating the expression of nodulation genes, thus initiating the formation of N_2-fixing root nodules. Some flavonoid metabolites are produced by plants as phytoalexins (stress compounds) or antibiotics or exert antioxidant activity.

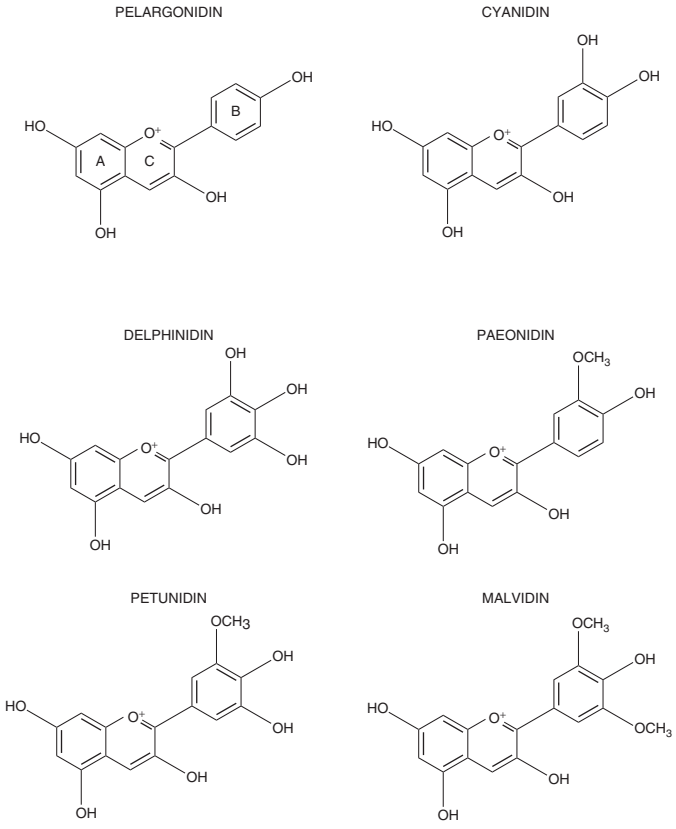


Figure 3.13-6. Some Flavonoids

3.13.1.4 Stilbenes

The biosynthesis of stilbenes (Fig. 3.13-5) is similar to that of the flavonoids. Three malonyl-CoA units (produced via the polyketide pathway) react with 4-coumaroyl-CoA (3.13.1.3). In this manner, 4 CO_2 are removed by decarboxylation and a diphenylethylene structure is formed. The resveratrol synthesis is shown as an example. This compound is a phytoalexin, which is produced by plants under the attack of bacteria and fungi. It has anti-cancer and anti-inflammatory activity.

3.13.1.5 Tannins (Fig. 3.13-7)

Tannins are plant polyphenols, widely occurring in gymnosperms and angiosperms. They can be classified chemically into two main groups, hydrolyzable (gallotaninns) and non-hydrolyzable (condensed) tannins, formed from flavonoid units (3.13.1.3). The gallotannins are glycosylated derivatives of gallic acid, which is derived from shikimic acid (3.2.7.1). The hydroxyl groups of a hexose (usually D-glucose) in the center of hydrolyzable tannins is esterified with numerous gallic acid molecules. The condensed tannins (proanthocyanidins) are oligo- or polymers of flavonoids units.

Tannins are mainly localized in the vacuoles or in specialized cells of the tree bark, wood, fruit, leaves, roots and plant galls for protection

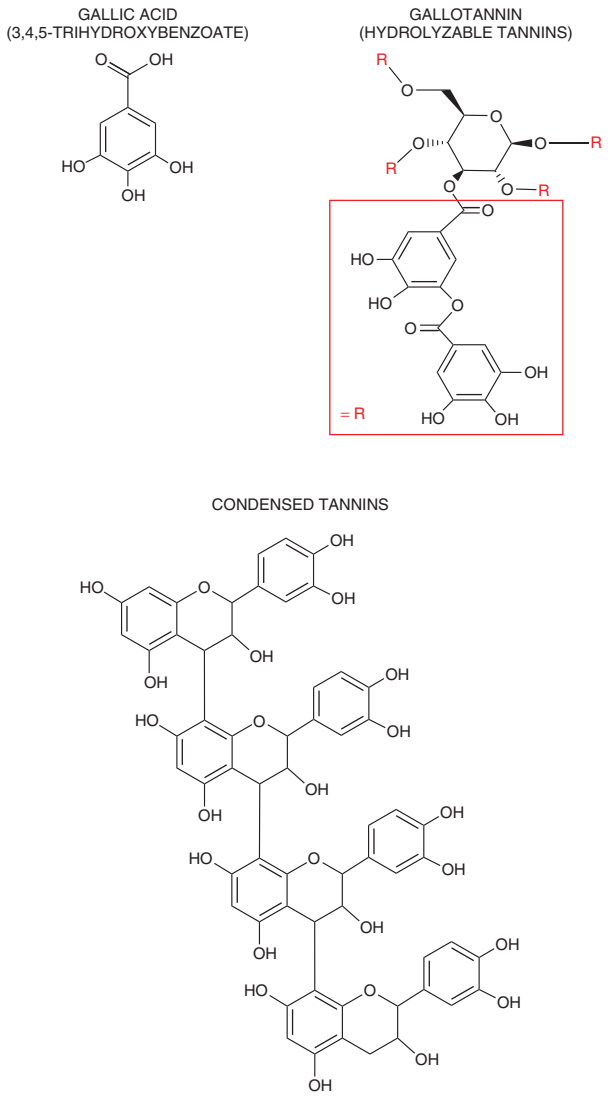


Figure 3.13-7. Gallic Acid and Tannins

against herbivores and pathogens. When the plant is wounded, the tannins are released and their phenolic groups bind to amino groups of plant proteins, converting the proteins into an indigestible form. This drastically reduces the food quality of the plant for herbivores. In addition, tannins have a bitter and astringent taste.

3.13.2 Terpenoids

The ubiquitously occurring terpenoids are the largest group of natural products, showing a wide structural diversity in carbon skeletons and functional groups, particularly within the plant kingdom. A part of these compounds is essential for plant development and hence is assigned to the primary metabolism, e.g., hormones, members of the electron transport system or pigments for light absorption. Most of the terpenoids, however, have an important function in the secondary metabolism, e.g., components of the essential oils, steroids, waxes, resins and natural rubber. A major number serve as defensive compounds against herbivores and pathogens, or as in the case of colors and scents, as attractants for pollinators. Due to the biological activity many of them have pharmacological significance.

3.13.2.1 Biosynthesis

All terpenoids are derived from the C_5-units 3-isopentenyl-PP (IPP) and dimethylallyl-PP (DMAPP), and are classified into hemiterpenes (C_5), monoterpenes (C_{10}), sesquiterpenes (C_{15}), diterpenes (C_{20}), triterpenes (C_{30}), tetraterpenes (C_{40}) and polyterpenes, according to the number of linked C_5-units. The biosynthesis of the precursors IPP and DMAPP of all terpenoids proceed via two different pathways, a mevalonate-dependent and a mevalonate-independent pathway:

- Mevalonate pathway**:** It is localized typically in the cytosol and is identical to the first part of cholesterol synthesis (Fig. 3.5.1-1) via the intermediate mevalonate up to IPP, which is catalyzed to DMAPP by isopentenyl-diphosphate isomerase.
- Rohmer pathway (non-mevalonate pathway or DOXP/MEP pathway, Fig. 3.5.3-1): It is localized in plastids and the precursors of it can be obtained from pyruvate and D-glyceraldehyde 3-P via producing the intermediates 1-deoxy-D-xylulose 5-P, 2-C-methyl-D-erythritol 4-P and further phosphorylated intermediates leading to IPP and DMAPP.

3.13.2.2 Hemiterpenes and Monoterpenes (Fig. 3.13-8)

The C_5-structure isoprene is a representative example of **hemiterpenes**. The synthesis takes place in the chloroplasts and is induced by light and high temperature. Isoprene is released by the cleavage of the diphosphate unit from dimethylallyl-PP (Fig. 3.5.1-1). It can contribute to the emission of organic aerosols together with other terpenoids in forest atmosphere, especially in coniferous forests.

Monoterpenes (C_{10}, Fig. 3.13-8) derive from geranyl-PP, which is formed by adding IPP to a DMAPP starter unit (head-to-tail addition, for mechanism see Fig. 3.5.1-2). In some cases, neryl-PP (i.e. the native substrate for the monoterpene synthase in tomato, *Solanum lycopersicum*) or linalyl-PP are the starting compounds. Most of them are volatile and typical scent and aroma compounds from higher plants. They are often found as components of the essential oils, together with the sesquiterpenes. Monoterpenes occur as acyclic, mono- or bicyclic molecules.

- Acyclic monoterpenes: Geraniol is the main constituent of rose oil and citronella oil. The essential oils of various *Citrus* species contain citronellol.
- Examples of monocyclic monoterpenes are menthol and limonene from *Mentha* species and thymol from thyme (*Thymus vulgaris*). All of them are synthesized by cyclization of geranyl-PP, a typical enzyme being limone synthase, synthesizing limonene. The resulting structures are further diversified by additional rearrangements and oxidations.
- Bicyclic monoterpenes are formed by two sequential cyclization reactions of geranyl-PP. Examples are pinene (in pine resin), camphene and thujene (a neurotoxic compound in absinth). Structures containing ketone, alcohol, and ether groups are, e.g., camphor, borneol and eucalyptol.

3.13.2.3 Sesquiterpenes (Fig. 3.13-8)

Sesquiterpenes (C_{15}) represent the largest group within the terpenoids. Several hundred sesquiterpene backbone structures have been identified and thousands of naturally occurring derivatives have been isolated. They are found in all tracheophyta, in mosses, fungi, brown and red algae, and in insects. The precursor is farnesyl-PP, synthesized by the transfer of two 3-isopentenyl-PP to a dimethylallyl-PP starter unit (head to tail addition, Fig. 3.5.1-1). Sesquiterpenes can be classified into acyclic, mono-, bi-, and tricyclic compounds.

- Acyclic sesquiterpenes are not common. Farnesol, an alcohol derivative of farnesyl diphosphate is present in essential oils of, e.g., rose flower, sandalwood and lemon grass.
- Monocyclic sesquiterpenes are based on several structural skeletons, e.g., bisabolane, germacrene, elemane and humulane. The cyclization reactions are catalyzed by specific cyclases. Bisabolol, an alcohol derivative has an anti-inflammatory effect and occurs in the essential oil of chamomile (*Matricaria chamomilla*) and in bergamot oil (*Citrus bergamia*).
- Bicyclic sesquiterpenes are, e.g., cadinenes and caryophyllenes. The latter are constitutents of many essential oils, e.g., clove (*Syzygium aromaticum*), hemp (*Cannabis sativa*), rosemary (*Rosmarinus officinalis*) and cinnamon (*Cinnamomum verum*). The phytoalexin capsidiol (Fig. 3.5.3-2) derives from the germacrene structure and can be found in pepper (*Capsicum anuum*) and tobacco (*Nicotiana tabacum*) in response to fungal infection. Azulenes (e.g., guaiazulene of chamomile (*Matricaria chamomilla*), Fig. 3.5.3-2) contain a condensed aromatic 5- and 7-ring system.

3.13.2.4 Diterpenes (Fig. 3.13-8)

Diterpenes (C_{20}) consist of four C_5-units and derive from geranyl-geranyl-PP. They occur in plants and fungi. Most of them are primary metabolites, such as the phytohormone gibberilic acid or phytol (Fig. 3.5.3-2), which is esterified to chlorophyll, both promoting growth and elongation during germination (Fig. 3.5.3-2). Paclitaxel (formerly named taxol), isolated from the bark from the pacific yew tree (*Taxus brevifolia*), has an anti-cancer effect and is used as a mitotic microtubule inhibitor in cancer therapy.

3.13.2.5 Triterpenes (Fig. 3.13-9)

Triterpenes (C_{30}) are derivatives of the acyclic squalene. This compound is synthesized through a head-to-head condensation of two farnesyl-PP molecules catalyzed by squalene synthase. In plants, squalene is converted to the tetracyclic cycloartenol by cycloartenol synthase. Cycloartenol is a precursor of plant steroids (phytosterols, Fig. 3.5.2-1), e.g., sitosterol, stigmasterol and campesterol (occurring in, e.g., soybean oil or rapeseed oil). Squalene can also be converted into α/β-amyrin by 2,3-oxidosqualene α- or β-amyrin cyclase. Amyrin is a precursor of pentacyclic triterpenes (see below). In animals, squalene is the precursor of cholesterol (3.5.1.1).

Cardiac glycosides occur only in glycosylated form in nature. Their aglycones can be classified as

- cardenolides (Fig. 3.5.2-1, exclusively synthesized in plants) and
- bufadienolides (formed in plants and in toads of the genus *Bufo*).

The characteristic features are additional 5-membered or 6-membered lactone rings, respectively. The glycosides of *Digitalis lanata* and *Digitalis purpurea* (digoxin and digitoxin) and other plants are important for pharmacological purposes, being the active components of drugs for treatment of heart insufficiency. Their effect is based on the inhibition of the Na^+/K^+ ATPase (TC 3.A.3.1.1, see sections 6.1.4 and 7.2.1), which is also responsible for their toxicity in higher doses.

Ecdysone (3.5.2.3 and Fig. 3.5.2-1) and 20-hydroxyecdysone are the major steroidal hormones of molting insects, which synthesize them from cholesterol or from plant sterols. Ecdysone analogues and some derivatives (e.g., abutasterone) were also isolated from the fern

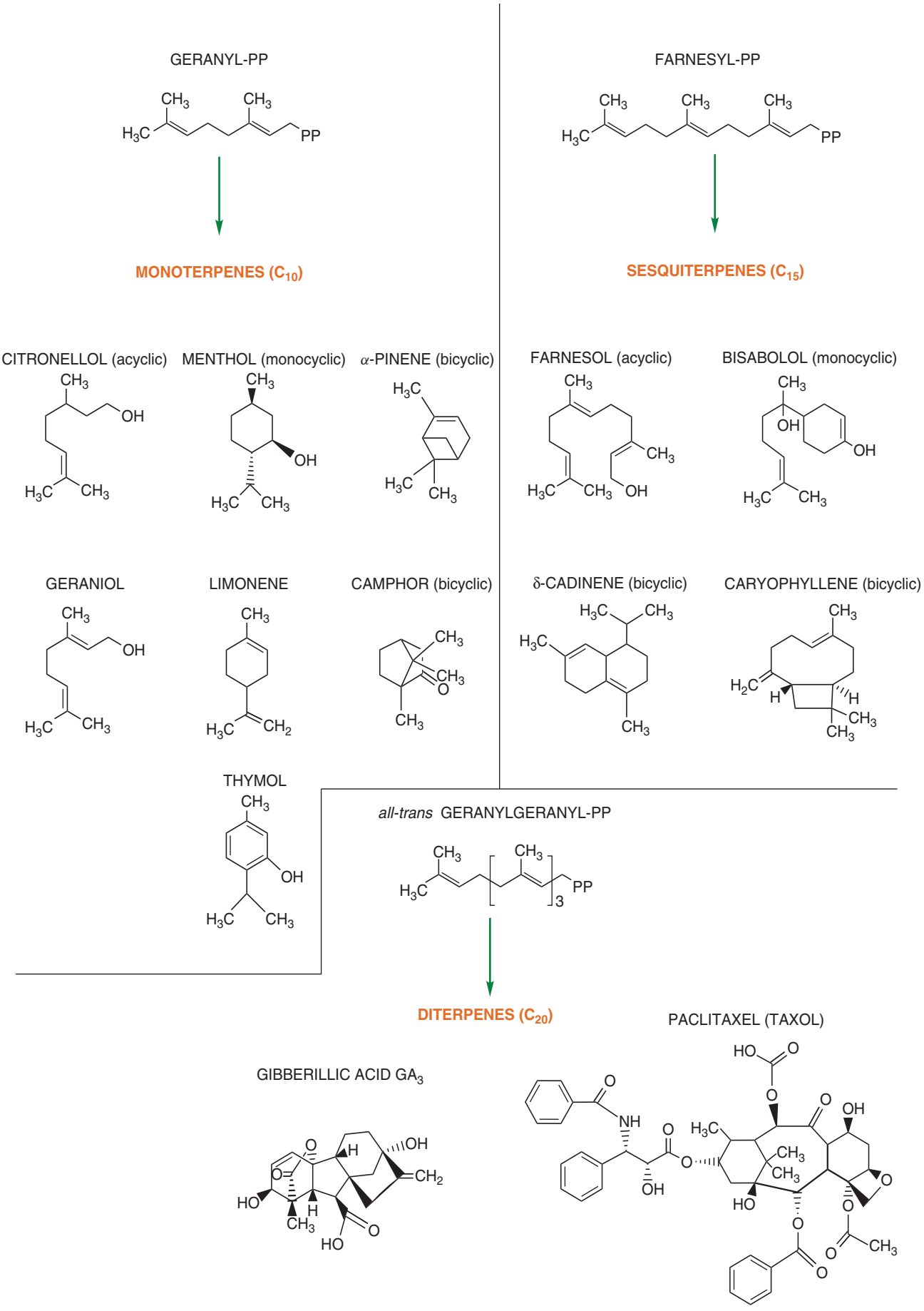


Figure 3.13-8. Examples for Monoterpenes, Sesquiterpenes and Diterpenes

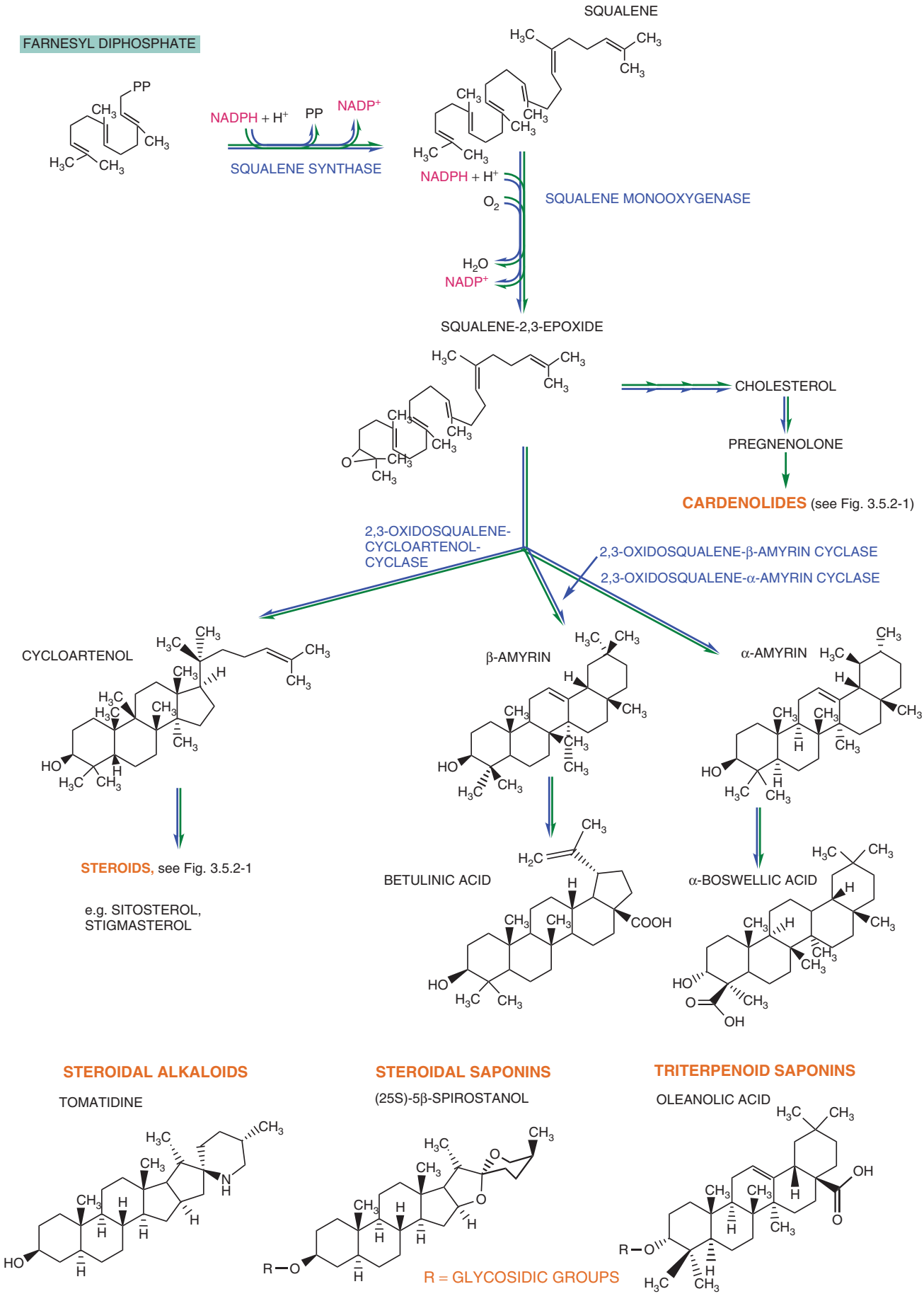


Figure 3.13-9. Triterpenes

Polypodium vulgare. These phytoecdysteroids act as insect feeding deterrents, disturb the precise synchronization of insect development and lead to the appearance of malformed animals.

Steroidal alkaloids are a group of nitrogen-containing steroids. Most of them are synthesized in higher plants. The insertion of nitrogen into the terpene structure results in an alkaline character and thus they share common properties with alkaloids. They belong to the so-called pseudoalkaloids (3.13.3.2). The nitrogen does not derive from amino acids but is inserted as NH_3 at a late stage of the biosynthesis. Like most of the alkaloids, the steroidal alkaloids are toxic to herbivores. The poisonous glycoalkaloids solasodine and solanine, isolated from potato (*Solanum tuberosum*) and tomatine from tomato (*Solanum lycopersicum*) are responsible for the toxicity of the green parts of the potato tuber and of unripe tomato fruits.

Non-glycosylated pentacyclic triterpenes are not a widespread group. Amyrin can be found in the latex of poinsettia ('Christmas Star', *Euphorbia pulcherrima*). Betulinic acid occurs in betulin (the pigment from the bark of white birch, *Betula pendula*). It shows anti-inflammatory and anti-tumor activities. Boswellic acid from frankincense (the resin of *Boswellia sacra*) is studied for anti-inflammatory applications.

Saponins (glycosylated pentacyclic triterpenes), are a group of plant secondary metabolites, which are localized in roots, rhizomes and seeds. Due to their amphiphilic character they can act as detergents, destroying cell membranes by surface action. Thus these agents protect against pathogens, fungi, and herbivores. Many of them are of pharmaceutical interest. According to the ring system of the aglycones, they are classified into steroidal saponins (with a spirostane structure, e.g., spirostanol) and triterpenoid saponins (e.g., oleanolic acid or solanine).

3.13.2.6 Tetraterpenes (Fig. 3.5.3-2)

A head-to-head condensation of two geranylgeranyl-PP molecules results in the C_{40} skeleton of phytoene, the precursor of all tetraterpenes. The carotinoids are the major group of the tetraterpenes, encompassing carotins and xanthophylls as their oxidation products. These compounds and their derivatives play important roles in the primary metabolism of plants (e.g., as pigments in the light-harvesting complexes of the photosynthesis system), as well as in other kingdoms of life. More details are given in chapters 3.5.3.2 … 4.

Under the aspect of secondary metabolism in plants, the lipophylic carotinoids are responsible for many colors in fruits and flowers varying from yellow (i.e., violaxanthine from pansy, *Viola tricolor*) to red (i.e., lycopene from tomato, *Solanum lycopersicum*). They accumulate in plastids.

3.13.2.7 Oligo- and Polyterpenes

Several plants are able to form polymers of the IPP and DMAPP derived isoprene, the polyterpenes. Natural rubber from the latex of the rubber tree (*Hevea brasiliensis*) consists of 500 to 5000 linearly bound C_5-units forming an *all-cis*-polyisoprene (Fig. 3.5.3-2). The biosynthesis is localized in the laticifers. Starting from geranyl diphosphate, isoprene units are added successively by rubber *cis*-polyprenyl*cis*transferase. Guttapercha is composed of isoprene units forming a *trans*-polyisoprene in latex. It is isolated from dried leaves of *Palaquium gutta*, a tropical tree, a native of Southeast Asia.

3.13.3 Nitrogen-Containing Secondary Metabolites

The large group of nitrogen-containing secondary compounds encompasses the glucosinolates, the cyanogenic glycosides and the alkaloids. Most compounds in this group derive from amino acids. Also non-proteinogenic amino acids belong to this group.

3.13.3.1 Cyanogenic Glycosides and Glucosinolates (Fig. 3.13-10)

The two classes of secondary metabolites share common properties. Their biosynthetic pathways are evolutionarily related. Both occur in a non-toxic glycosylated form. These water-soluble compounds are stored in the vacuoles of specialized cells. In case of plant damage, they release toxic compounds by the action of an enzyme localized in a different cell compartment.

Cyanogenic glycosides are widely distributed in the plant kingdom and can be encountered in gymnosperms and angiosperms, e.g., in seeds of bitter almonds (*Prunus dulcis*), in the tuberous root of cassava (*Manihot esculenta,* a tropical native of South America) and in sorghum (*Sorghum bicolor*). Dhurrin, a cyanogenic glycoside in sorghum is localized in the vacuole of the epidermis cells. When the plant tissue is destroyed, cytosolic β-glucosidase cleaves the glucoside bond, releasing 4-hydroxy-(S)-mandelonitrile, which, in turn, is split by cytosolic mandelonitrile lyase into 4-hydoxybenzaldehyde and the toxic hydrogen cyanide.

Glucosinolates are nitrogen- and sulfur-containing compounds, likewise derived from amino acids and α-D-glucose. They occur in almost all plants of the Brassicaceae and related families and have deterrent effect against herbivores, due to their bitter and sharp taste, which is characteristic for, e.g., horseradish, cauliflower, cabbage, mustard, and broccoli. Sinigrin, a glucosinate from horseradish derives from L-methionine and α-D-glucose and is accumulated in the vacuole. When the plant is damaged by a herbivore, cytosolic myrosinase (thioglucosidase) cleaves the glucoside and degrades sinigrin to allyl isothiocyanate, which is a effective deterrent.

Non-proteinogenic amino acids are metabolites produced by plants serving as an efficient defense against herbivores. They are toxic substances due to their structural similarity to proteinogenic amino acids. As an example, canavanine, isolated from several Fabaceae (leguminous plants) is structurally related to L-arginine (Fig. 3.13-11). After consuming these compounds, the herbivores mistakenly insert those amino acids into their own proteins, causing inactivation.

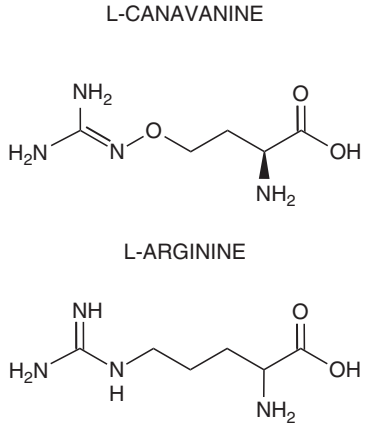


Figure 3.13-11. Canavanine and L-Arginine

3.13.3.2 Alkaloids

Alkaloids are a large class of naturally "alkali-like" secondary metabolites containing heterocylic nitrogen. They can be found in ca. 20 % of the vascular plants, mainly in angiosperms. Many of them are poisonous and function as a defense against herbivores and pathogens (e.g., colchicine). Thus, the highest concentration can often be detected in those tissues that are most important for the reproduction and survival of the plant or with the highest probability of attack, e.g., seeds, flowers, young and also growing peripheral tissues. The structures of ca. 12,000 plant alkaloids have been elucidated, and many of them show biological activity, mostly with toxic effects. They are used as stimulants (e.g., caffeine, nicotine) and drugs (e.g., morphine, quinine).

Their structural diversity is classified according to their occurrence in certain plant lineages (e.g., *Nicotiana* alkaloids from tobacco), their amino acid origin (Table 3.13-4) or to their structural skeleton:

- 'True' alkaloids contain a heterocyclic nitrogen atom, originating from an amino acid.
- Proto-alkaloids are also amino acid-derived, but their nitrogen is outside of the ring system (including peptides and polyamines).
- In pseudoalkaloids the nitrogen is bound in a heterocycle, but does not originate from amino acids. An example is coniin, a piperidine derivate. It is a neurotoxin, found in *Conium maculatum*, which paralyzes and disrupts the peripheral nervous system. Paclitaxel, another example is a diterpenoid-derived pseudoalkaloid (3.13.2.4)

Table 3.13-4. Major Types of Alkaloids and Their Amino Acid Precursor

Alkaloid class	Biosynthetic precursor	examples
Pyridine (*Nicotiana*) alkaloids	aspartate	nicotine
Quinolizidine (Lupin) alkaloids	lysine	lupinine
Purine alkaloids	glycine	caffeine
Pyrrolizidine alkaloids	ornithine	senecionine
Indole alkaloids	tryptophan	strychnine
Ergoline alkaloids	tryptophan	D-lysergic acid diethylamide (LSD)
Isoquinoline, benzylisoquinoline alkaloids	tyrosine	codein, morphine
Tropane alkaloids	ornithine	atropine

***Nicotiana*/tobacco alkaloids (Fig. 3.13-12):** Nicotine, nornicotine, anabasine, and anatabine can be found in the nightshade family (Solanaceae), mainly in the *Nicotiana* species, with nicotine being the major metabolite. Their biosynthesis is exclusively localized in the roots. The alkaloids are subsequently transported into the shoot via the xylem. Nicotine is strongly toxic to the nervous system and functions as a defense compound, especially as a natural insecticide. The backbone structure of nicotine is composed of two heterocyclic rings: the pyridine ring originates from nicotinic acid (Fig. 3.7.9-1), while the N-methylpyridine ring is synthesized from ornithine via putrescine. The heterocycles of anabasine and anatabine originate from nicotinic acid and lysine, respectively.

Quinolizidine alkaloids (Fig. 3.13-13): Quinolizidine alkaloids are also termed 'lupin alkaloids' due to their major occurrence in the genus *Lupinus* (Fabaceae). The basic structure, quinolizidine, is a bicyclic 6-membered ring system, sharing a nitrogen atom. The biosynthesis is located in the chloroplasts. It starts from lysine, which is transformed to cadaverine by lysine decarboxylase. Thereafter, one or more cadaverine molecules are integrated, yielding bi-, tri- or tetracyclic structures. The exact mechanism is not fully understood yet.

Lupinine, a bicyclic compound is isolated from yellow lupin (*Lupinus luteus*). Cytisine, a tricylic alkaloid from the golden chain (*Laburnum anagyroides*) is highly toxic and can be found in all parts of the plant. Sparteine (*Lupinus scoparius*) and lupanine from the white lupin (*Lupinus albus*) are typical compounds with a tetracylic structure.

Purine alkaloids (Fig 3.13-14): The precursor of purine alkaloids is xanthine, a degradation product of the purine pathway (Fig. 3.6.1-4). Caffeine, theobromine, and theophylline are methylated derivatives of xanthine and have stimulating effects. They occur in seeds of coffee (*Coffea arabica*) and cacao (*Theobroma cacao*) and in the leaves of the tea plant (*Camellia sinensis*).

Pyrrolizidine alkaloids: Pyrrolizidine alkaloids are a group of more than 400 structures. They are esters of the 'necine base' and one or more 'necic acids'. The 'necic acids' derive from branched-chain aliphatic amino acids isoleucine or valine (3.2.6). The 'necine base' 1-hydroxymethylpyrrolizidine is biosynthesized from putrescine and spermidine (3.2.9.3) via homospermidine.

Esterification results in a backbone structure (senecionine *N*-oxide, in the case of *Senecio* plants, Fig. 3.13-15), which is later structurally diversified by one- or two step transformations, e.g., dehydrogenations, position-specific hydroxylations, epoxidations, and O-acetylations.

Pyrrolizidine alkaloids occur in distantly related plant families of angiosperms, e.g., in the genera *Senecio* and *Eupatorium* (Asteraceae), *Heliotropium* and *Cynoglossum* (Boraginaceae) as well as in certain orchids such as *Phalaenopsis*. In *Senecio* species the biosynthesis of the alkaloids is localized in the roots and the synthesized polar *N*-oxides are subsequently transported into the vacuoles of the aerial parts via the phloem. If an animal is feeding on these plants, the *N*-oxides are spontaneously transformed into the pro-toxic free

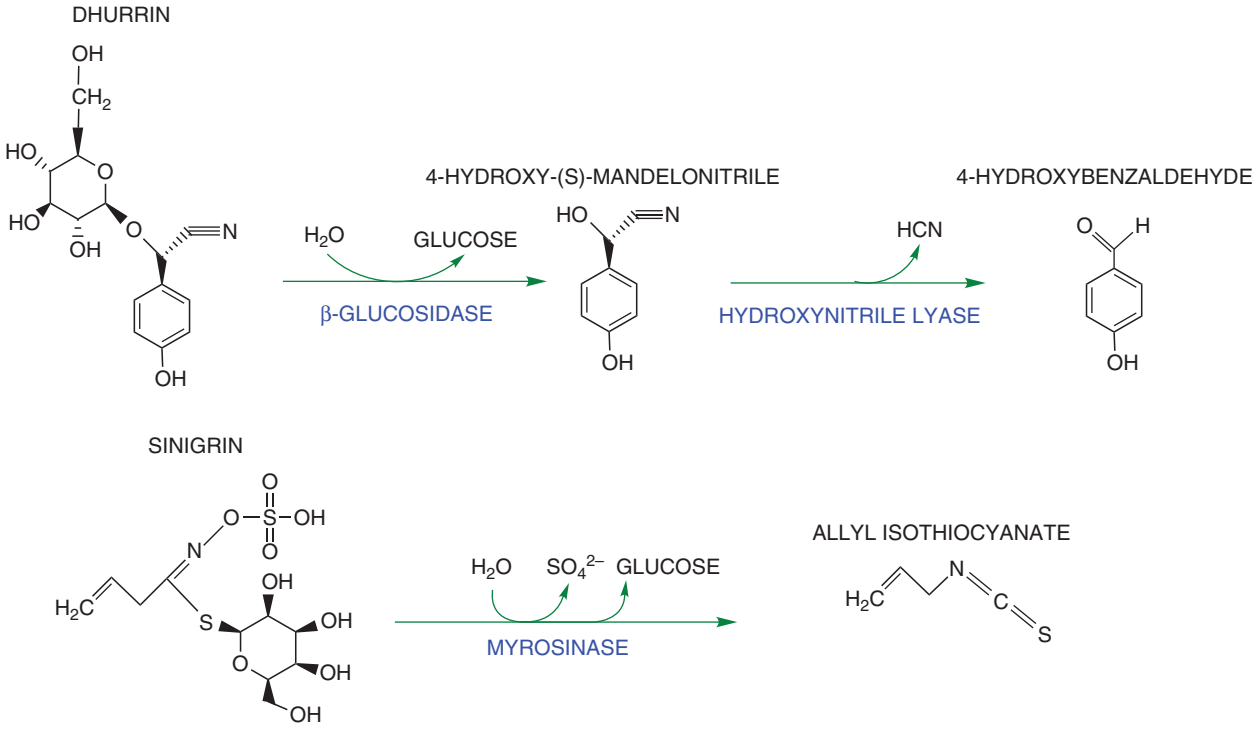


Figure 3.13-10 Examples of Cyanogenic Glycosides and Glucosinolates

base in the intestine. After resorption, the free base is bioactivated by cytochrome P450 enzymes into reactive pyrrolic intermediates (Fig. 3.13-16). These compounds react with proteins and nucleic acids and thus exert severe cell toxicity. They are strong feeding deterrents for livestock, wildlife, and insects.

Monoterpene indole alkaloids (Fig. 3.13-17): Monoterpene indole alkaloids encompass a group of more than 2500 compounds that were isolated mainly from the plant families Rubiaceae, Loganiaceae, and Apocynaceae. The alkaloids are synthesized from geranyl-PP (obtained via the Rohmer/non-mevalonate pathway, Fig. 3.5.3-1), which is converted into secologanin, a monoterpene (3.13.2.2). This compound undergoes an addition reaction with tryptamine (3.2.7.3) catalyzed by strictosidine synthase (Fig. 3.13-18). The resulting strictosidine is the central intermediate for all monoterpene indole alkaloids, e.g., yohimbine, catharanthine, strychnine, quinine and bisindole alkaloids. A number of these multi-step pathways have been described.

Many of these compounds show strong biological activity. An example is strychnine and its derivative brucin from the seeds of *Strychnos nux-vomica.* Both cause strong muscular convulsions, which could

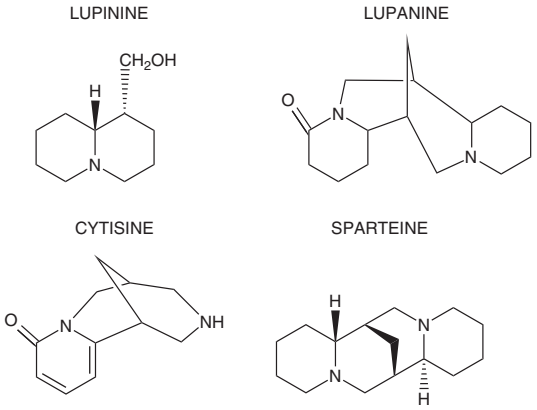


Figure 3.13-13. Examples of Quinolizidine Alkaloids

Figure 3.13-12. Biosynthesis of Nicotine

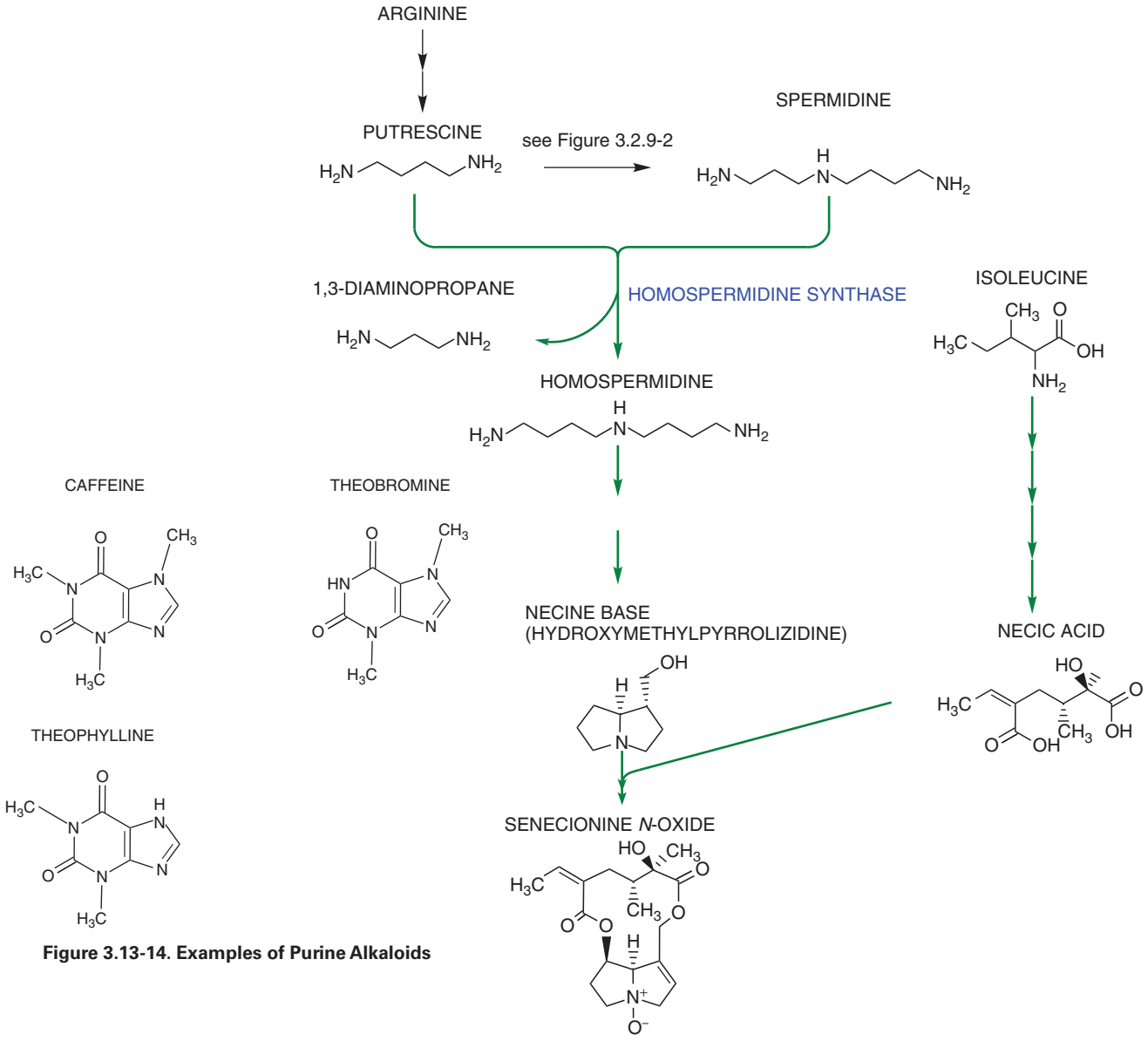


Figure 3.13-14. Examples of Purine Alkaloids

EXAMPLES OF PYRROLIZIDINE ALKALOIDS

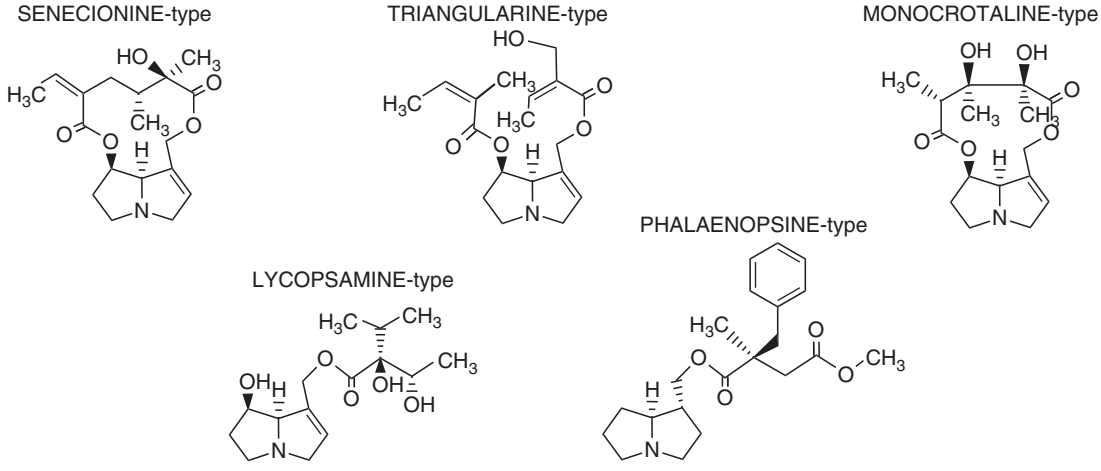


Figure 3.13-15. Synthesis and Structures of Pyrrolizidine Alkaloids

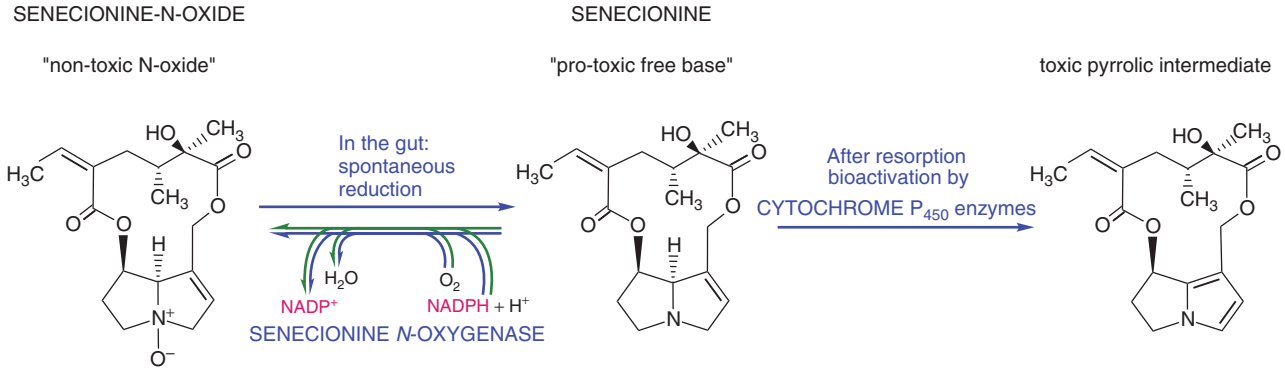


Figure 3.13-16. Bioactivation of Pyrrolizidine Alkaloids

lead to death by exhaustion. Extracts from other *Strychnos* species contain the bisindole alkaloids toxiferin and tubocurarin, which are the components of curare, an arrow poison from South America. These alkaloids inhibit the neuromuscular transmission resulting in paralysis of the peripheral nerves, causing respiratory paralysis and death.

Indole alkaloids are used as anti-cancer, anti-malarial and anti-arrhythmic agents. The pharmacological use of, e.g., vinblastine and vincristine as anti-cancer drugs is due to their inhibition of microtubule formation during mitosis, disruption of mitotic spindle assembly and arrest of tumor cells in the M phase of the cell cycle (4.3.5). Ajmaline is used in the treatment of cardiac arrhythmia. It is produced in *Rauwolfia serpentina* cell cultures involving many enzymatic steps.

Ergoline alkaloids (Fig. 3.13-19): Like the monoterpene indole alkaloids (see above), ergoline alkaloids are tryptophan-derived secondary metabolites. They can be divided into three compound classes:

- Lysergic acid amides (e.g., ergometrine).
- Lysergic acid peptide derivatives (e.g., ergotamine and ergotoxine). This group contains a complex cyclolactam-tripeptide structure generated from the three amino acids α-hydroxyalanine, proline, and phenylalanine.
- Clavine alkaloids, derivatives of 6,8-dimethylergolines. They are biologically inactive.

Lysergic acid derivatives show strong biological activity: Ergometrine causes rhythmical contractions of the uterus (German name

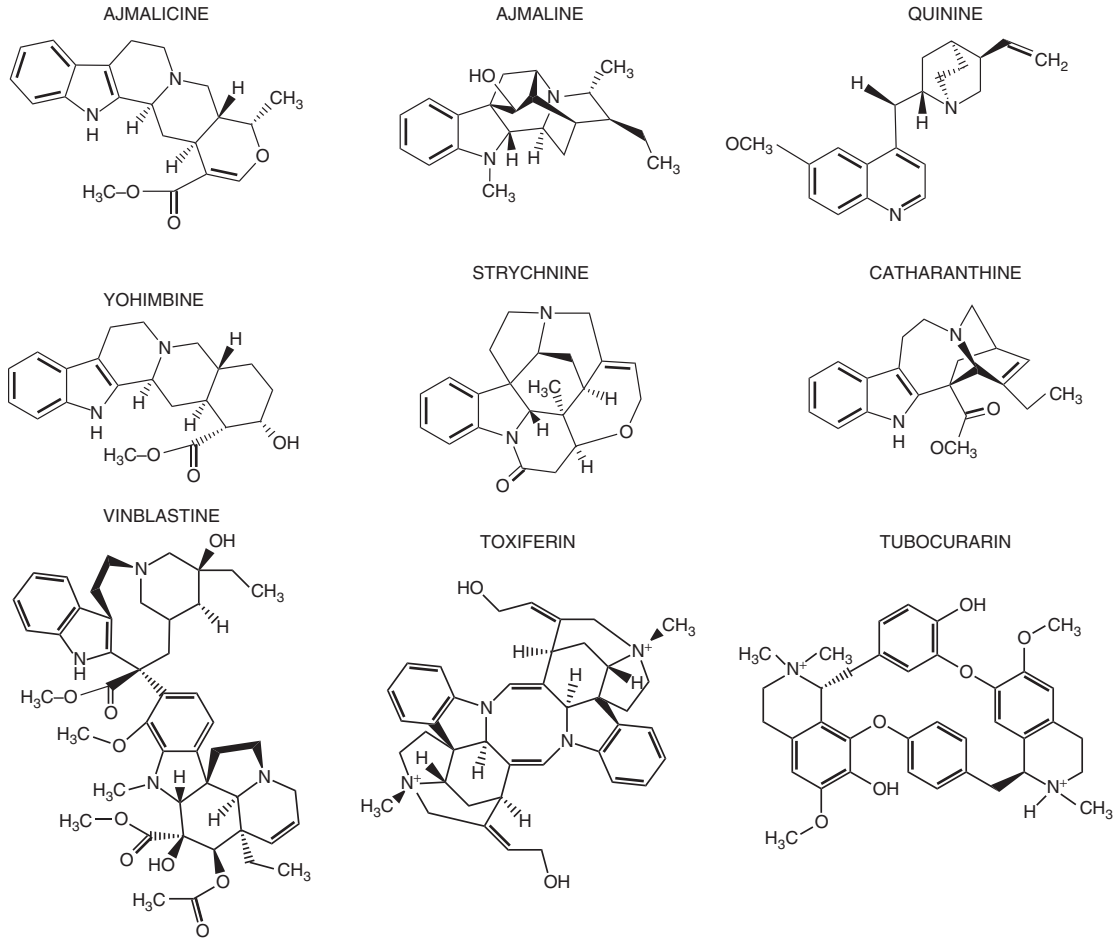


Figure 3.13-17. Indole Alkaloids

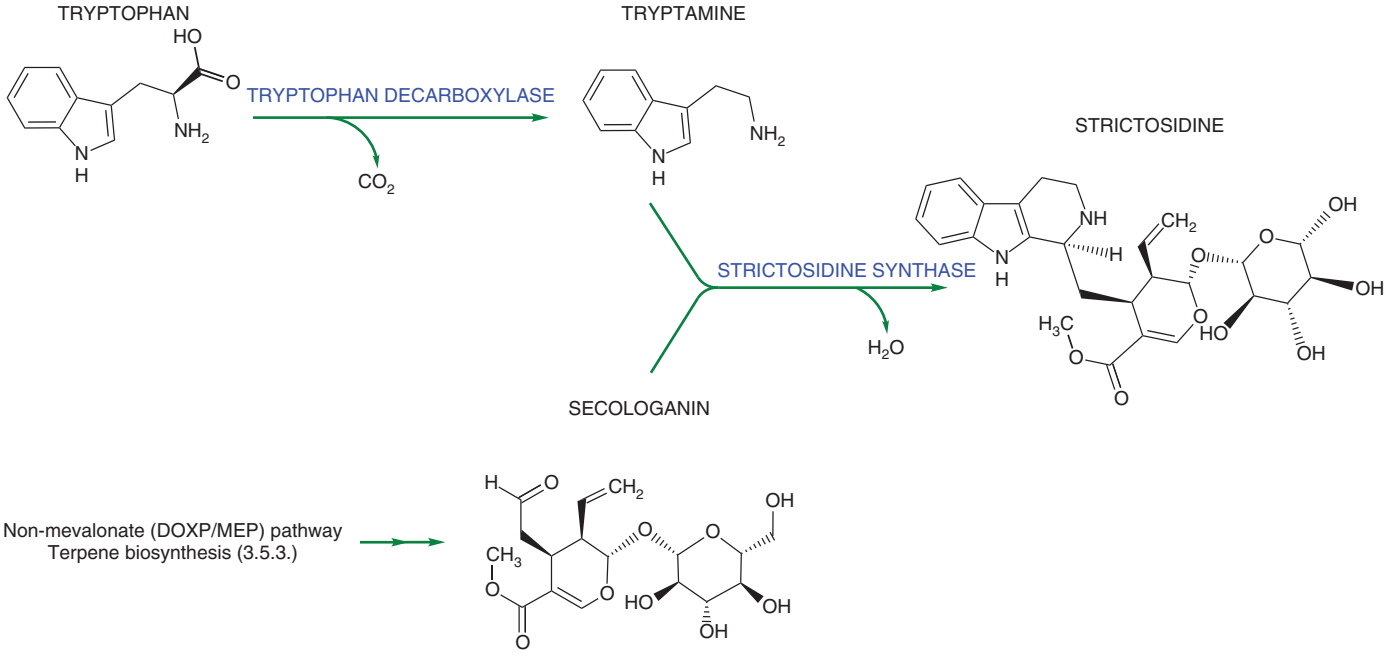


Figure 3.13-18. Biosynthesis of Strictosidine

'Mutterkorn' for the alkaloid group), ergotamine and ergotoxine have styptic effects. The peptide alkaloids also show sympatholytic effects and inhibit the action of epinephrine, norepinephrine and serotonin. The synthetic derivative lysergic acid diethylamide (LSD) produces hallucinogenic effects.

Ergotamine and ergotoxine alkaloids were first isolated from the fungus *Claviceps purpurea*. This fungus infects different genera of grains and grasses and Convolvulaceae, forming a violet-black dormant form (sclerotium), which is resistant against low temperature and drought. The sclerotium contains up to 1–2 % alkaloids. During the Middle Ages infections of the grain with the fungi frequently caused food poisoning (ergotism).

Benzylisoquinoline alkaloids (Fig. 3.13-20): These compounds occur mainly in the plant families Papaveraceae, Ranunculaceae, Berberidaceae, and Menispermaceae. Presently more than 2500 structures are elucidated. The most prominent natural products, which are mainly isolated from the latex, are codeine, morphine, and papaverine from opium poppy (*Papaver somniferum*), chelidonine from *Chelidonium majus* and berberine from *Berberis vulgaris*.

The compounds can be classified into the morphine-type, the benzylisoquinoline-type, the benzophenanthridine-type and the protoberberine-type alkaloids.

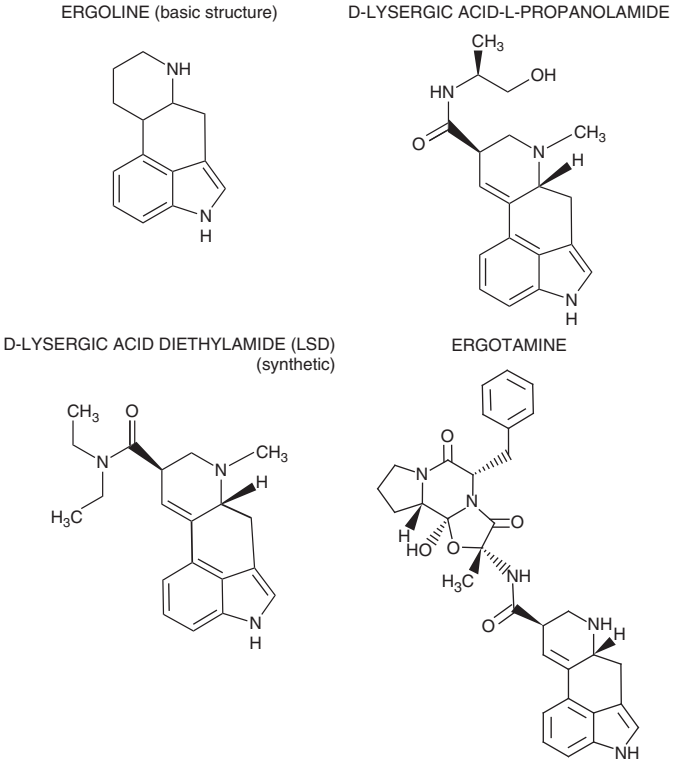


Figure 3.13-19. Ergoline Alkaloids

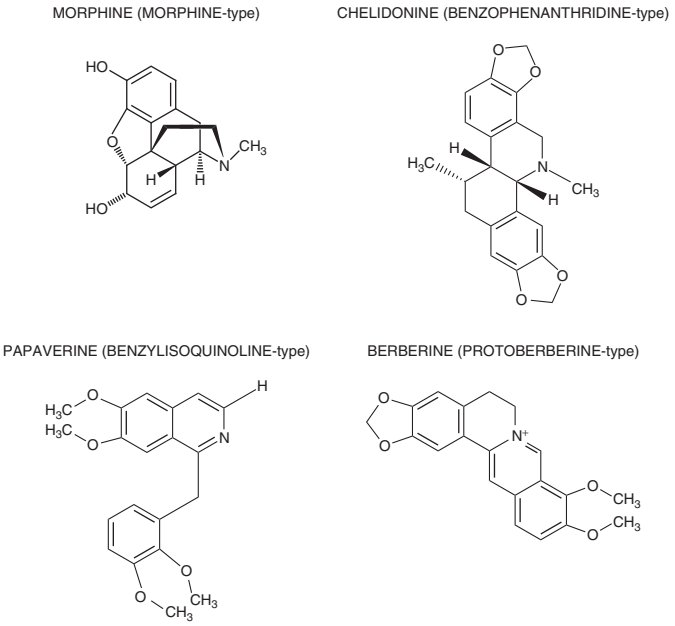


Figure 3.13-20. Benzylisoquinoline Alkaloids

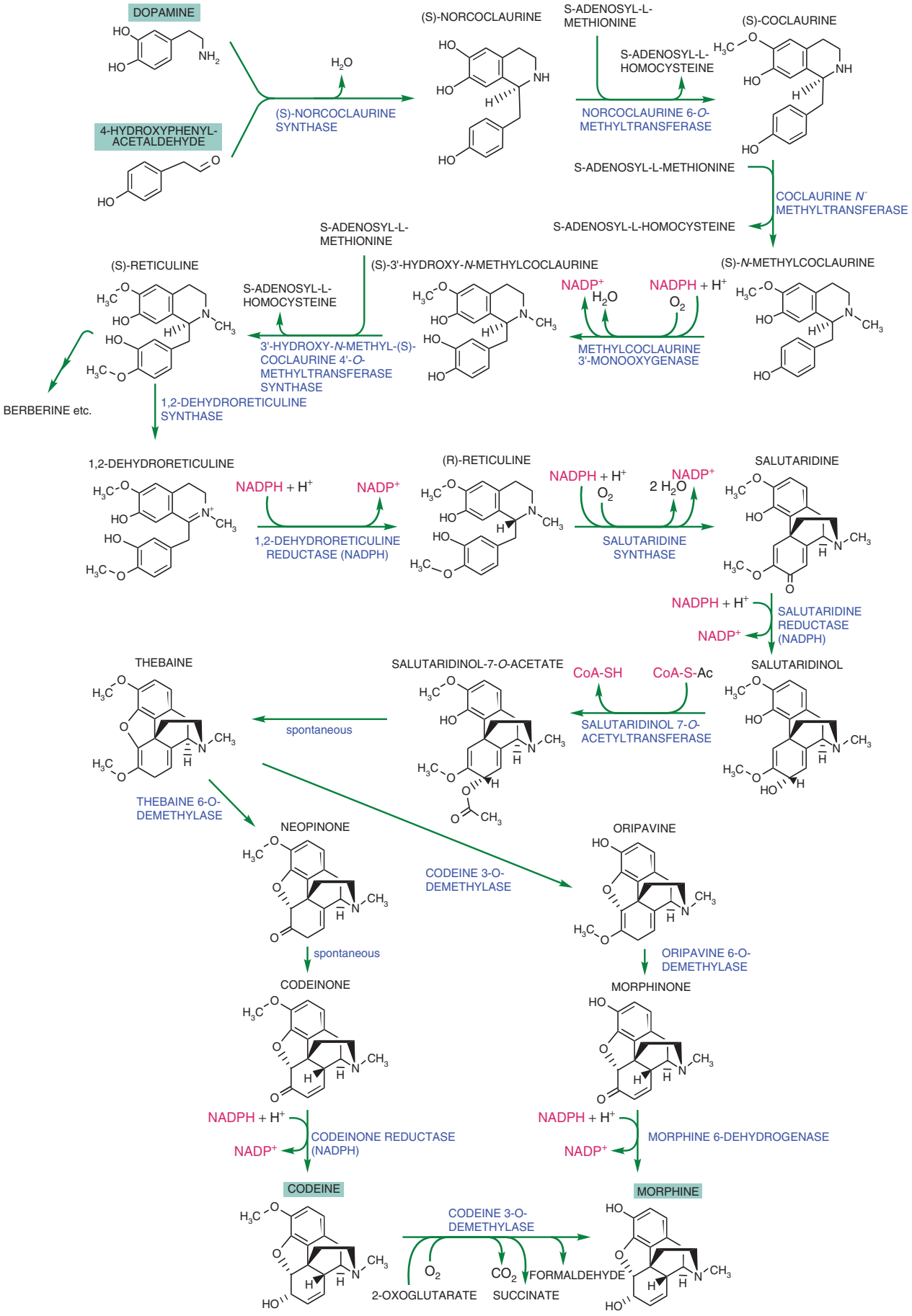


Figure 3.13-21. Morphine Biosynthesis

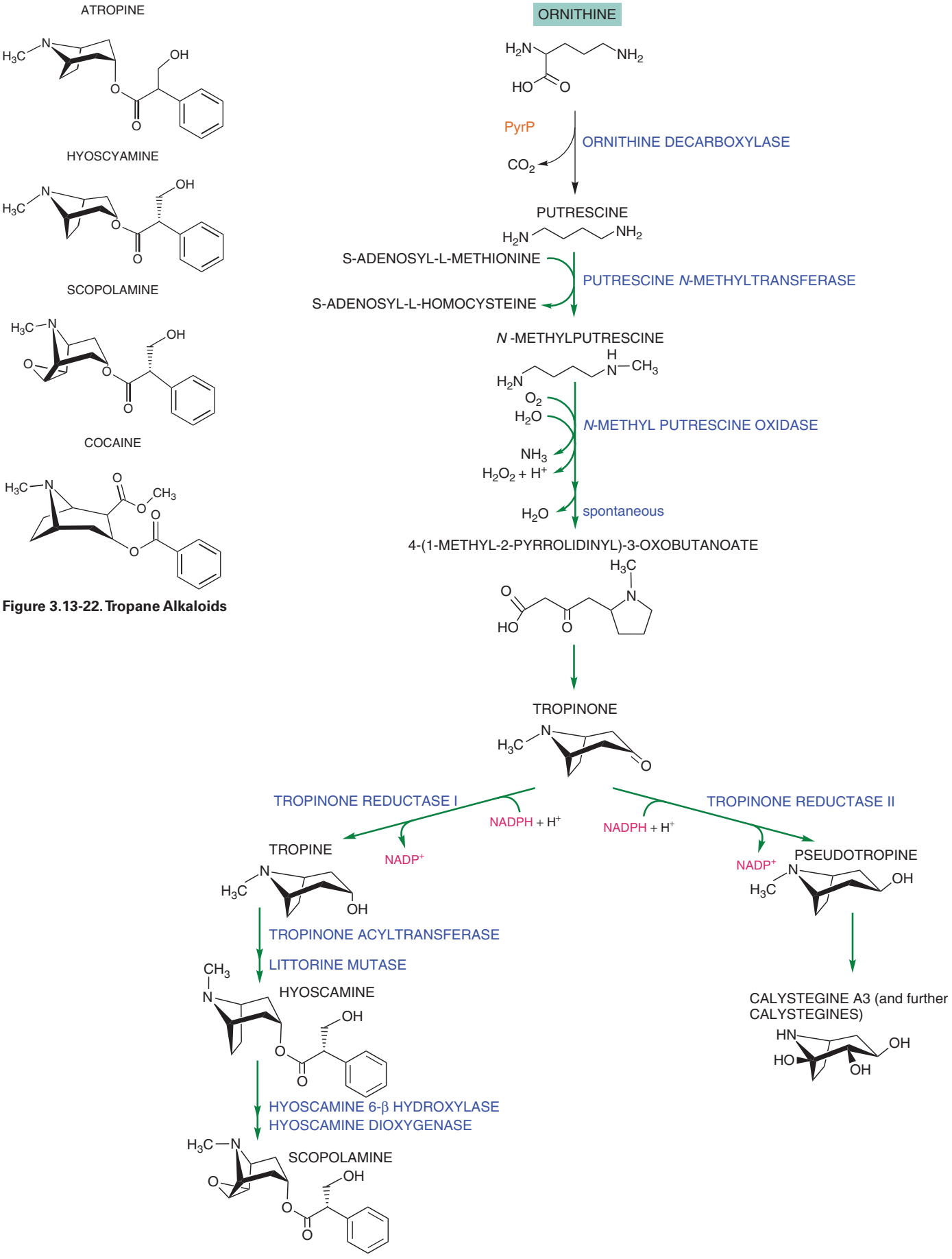


Figure 3.13-22. Tropane Alkaloids

Figure 3.13-23. Biosynthesis of Tropane Alkaloids

Morphine is the major alkaloid from opium poppy, one of the oldest medicinal plants and is a highly potent narcotic and analgesic opiate drug. It acts directly on the peripheral and central nervous system to decrease pain or to cause respiratory depression.

Morphine biosynthesis (Fig. 3.13-21): Almost all enzymes of this pathway have been described and the pathway is well understood. The first step in the biosynthesis is the condensation reaction of the tyrosine derivatives dopamine (3.2.7.3) and 4-hydroxyphenylacetaldehyde. The product norcoclaurine is O-methylated at position 6 yielding (S)-coclaurine. Then a N-methylation and a 3′-hydroxylation lead to (S)-3′-hydroxy-N-methylcoclaurine. The last step to (S)-reticuline is a 4′-O-methylation. This compound is also the branching point leading to various other benzylisoquinolines, e.g., berberine, palmatine and sanguinarine.

The specific morphine pathway starts with the two-step epimerization of (S)-reticuline. The subsequent synthesis to salutaridine and salutaridinol takes place through an intramolecular carbon-carbon phenol coupling. Afterwards salutaridinol is acetylated. Depending on the pH, the product salutaridinol 7-O-acetate can spontaneously cyclize to thebaine, a pentacyclic morphinan alkaloid, which is a precursor for synthetic morphine derivatives, e.g., diacetylmorphine (heroin). The final steps in the morphine biosynthesis consist of two demethylations yielding codeinone and reduction to codeine, which finally is demethylized to morphine.

Tropane alkaloids (Fig. 3.13-22): The occurrence of the tropane alkaloids is restricted to some genera of the Solanaceae and Erythroxylaceae. The compounds are found in, e.g., deadly nightshade (*Atropa belladonna*), and coca plant (*Erythroxylum coca*).

The initial part of the biosynthesis (Fig. 3.13-12) is shared with the formation of the *Nicotiana* alkaloids (see above). The backbone structure is tropane, a nitrogen-containing bicyclic ring system. It derives from ornithine/arginine via putrescine forming a *N*-methyl-Δ^1-pyrrolinium cation, which is further metabolized to tropinone. The subsequent reductions performed by tropinone reductase I and tropinone reductase II lead to tropine and pseudotropine, respectively. The former is catalyzed to L-hyoscyamine or DL-hyoscyamine (atropine) and scopolamine, whereas pseudotropine is the precursor for the biosynthesis of calystegines.

Atropine inhibits competitively the muscarinic actions of acetylcholine, e.g., it causes the relaxation of the circular eye muscle resulting in the dilation of the pupil. Besides hyoscyamine and scopolamine, cocaine shows high biological activity. Cocaine can be isolated from the leaves of the coca plant Erythroxylum coca. It is a powerful addictive stimulant of the nervous system.

Literature:

Ahimsa-Müller, M.A. J. Nat. Prod. 2007;70:1955–1960.

Bohlmann, J., Gershenzon, J. Proc. Natl Acad. Sci. USA 2009;106:10402–10403.

Facchini, P.J., De Luca,V. Plant J. 2008;54:763–784.

Flores-Sanchez, I.J., Verporrte, R. Plant Physiol. Biochem. 2009;47:167–175.

Hartmann, T., Ober, D. Top. Curr. Chem. 2000;209:207–244.

Hunter, W.N. J. Biol. Chem. 2007;282:21573–21577.

Munk, K. *Taschenlehrbuch Biologie Botanik.* Stuttgart: Thieme, 2008.

O'Connor, S.E., Maresh, J.J. Nat. Prod. Rep. 2006;23:532–547.

Richter, G. *Biochemie der Pflanzen.* Stuttgart: Thieme, 2008.

Stöckigt, J., Panjikar, S. Nat. Prod. Rep. 2007;24:1382–1400.

Träger, B. Phytochemistry 2006;67:327–337.

Weiler, E., Nover, L. *Allgemeine und molekulare Botanik.* Stuttgart: Thieme, 2008.

4 Protein Biosynthesis, Modifications and Degradation

4.1 Protein Synthesis in Bacteria

Martina Jahn and Dieter Jahn

4.1.1 Bacterial Transcription

Transcription is the selected transfer of genetic information stored in the DNA into single-stranded RNA. The base sequence of the produced RNA is identical with that of one strand of the DNA duplex, but with the exception of the use of uracil instead of thymine nucleotides (Fig. 4.1.1-1).

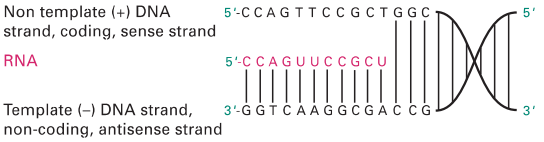


Figure 4.1.1-1. Principle of Transcription

In bacteria, this leads to the synthesis of three classes of RNAs:

- messenger RNA (mRNA), encoding proteins
- combined ribosomal/transfer RNA (rRNA/tRNA) transcripts from which tRNAs and rRNAs are obtained by cleavage (Fig. 4.1.1-3)
- additional RNA transcription units:
  - small regulatory and antisense RNAs (micro RNAs)
  - 6S RNA as a RNA polymerase regulator
  - 4.5S RNA (*E. coli*) as part of the signal recognition particle for protein export
  - Catalytic RNA, ribozymes (RNA part of RNAse P)

A usual transcriptional unit for a bacterial protein encoding gene (Fig. 4.1.2-2) consists of several typical regions including:

- a promoter region, which binds the RNA polymerase, positive and negative transcriptional regulatory proteins and RNAs
- an upstream non-coding region between the transcriptional and translational start site harboring the information for the ribosome binding site
- the actual coding region (which frequently contains several contiguously arranged structural genes), beginning at the translational start site (position +1). This region is also involved in the binding regulatory proteins
- a termination sequence.

4.1.1.1 Bacterial RNA Polymerase, Promoters and Initiation of Transcription (Fig. 4.1.1-2)

All classes of bacterial RNAs are synthesized by one type of DNA-directed RNA-polymerase. The core RNA polymerase consists of five subunits ($\alpha_2\beta\beta'\omega$) and requires an additional σ-(sigma) factor for activity (Tables 4.1.1-1 and 4.1.1-2). Unlike DNA polymerases, no primer is needed for initiation. The core RNA polymerase is DNA binding *per se*, however, without any DNA sequence specificity. The sigma factor confers specificity to the regions of transcriptional initiation, called promoter regions. Depending on the environmental conditions, the different sigma factors direct the enzyme to different promoters for the transcription of appropriate genes. During heat shock, genes for chaperones and proteases are induced by the heat-shock specific sigma factor σ^{32}. However, the majority of genes is under control of the vegetative sigma factor σ^{70}. Here additional regulatory mechanisms are required. They are mediated by additional transcription factors with specific binding sites in the promoter region. Additionally, small regulatory RNAs (micro RNAs) might participate in gene regulation. Some details about gene regulations are given below.

Table 4.1.1-1. Components of the Bacterial RNA Polymerase Holoenzyme

Subunit	Copies in holoenzyme	Gene	Mol. mass kDa	Function
α	2	*rpoA*	36.5	enzyme assembly, regulatory factor binding, C-terminal αCTD domain binds upstream of promoter elements
β	1	*rpoB*	150	catalytic subunit binds NTP, forms phosphodiester bonds
β′	1	*rpoC*	155	binds nonspecifically DNA, contains 2 Zn^{++}
σ	1	*rpoD*	70	recognizes the promoter and initiates synthesis
ω	1	*rpoZ*	10	Protective chaperon function, binds to β′ subunit

Table 4.1.1-2. The Sigma Subunits of *E. coli* RNA Polymerase

Sigma factor	Gene	Environmental stimuli	Molecules per cell, exponential growth
σ^{70}	*rpoD*	vegetative growth	700
σ^{54}	*rpoN*	reduced nitrogen assimilation	110
σ^{38}	*rpoS*	stationary phase	1
σ^{32}	*rpoH*	heat-shock	10
σ^{28}	*rpoF*	flagellin and chemotaxis genes	370
σ^{24}	*rpoE*	extreme/extracellular heat shock adaptation	10
σ^{19}	*fecJ*	iron depletion	1

RNA polymerase initiates transcription at the transcriptional start site upstream of the translational start codon. Therefore, the resulting transcript contains a 5′ untranslated region, which often contains a ribosome binding site. Subsequently, transcription proceeds through the coding region (elongation) and continues until a site for termination is reached. The regulation of bacterial transcription is discussed in detail below. The mechanism of the actual bond-forming reaction is analogous to that of DNA replication (Fig. 3.8.1-2) except that it involves ribonucleotides instead of deoxyribonucleotides.

Initiation: Bacterial σ^{70} dependent promoters usually consist of 2 modules with the following consensus sequences (the subscripts indicate the frequency of occurrence):

–35 sequence	–10 sequence (Pribnow-box)
(in highly effective promoters)	(general)
$T_{0.69}T_{0.79}G_{0.61}A_{0.56}C_{0.54}A_{0.54}$	$T_{0.77}A_{0.76}T_{0.60}A_{0.61}A_{0.56}T_{0.82}$

The RNA polymerase holoenzyme binds unspecifically to the DNA and slides along the DNA until it recognizes the DNA promoter region by its σ subunit. There it binds very tightly with a K_M of up to 10^{-14} mol/1. Two turns of the DNA from the –12 base pair to the +4 base pair are unwound and both DNA strands are separated to form the initiation complex (Fig. 4.1.1-2). The initial two ribonucleoside triphosphates are joined, the first being most commonly ATP, less frequently GTP. Then the σ subunit dissociates from the complex.

Elongation: RNA is synthesized at the template in the 5′→3′ direction, while the transcription bubble moves along the DNA at a rate of 50 to 100 nucleotides/sec. Usually 12 nucleotides of the newly formed RNA form a hybrid with the template DNA before both strands become separated. The movement of the transcription bubble

Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Second Edition. Edited by Gerhard Michal and Dietmar Schomburg.

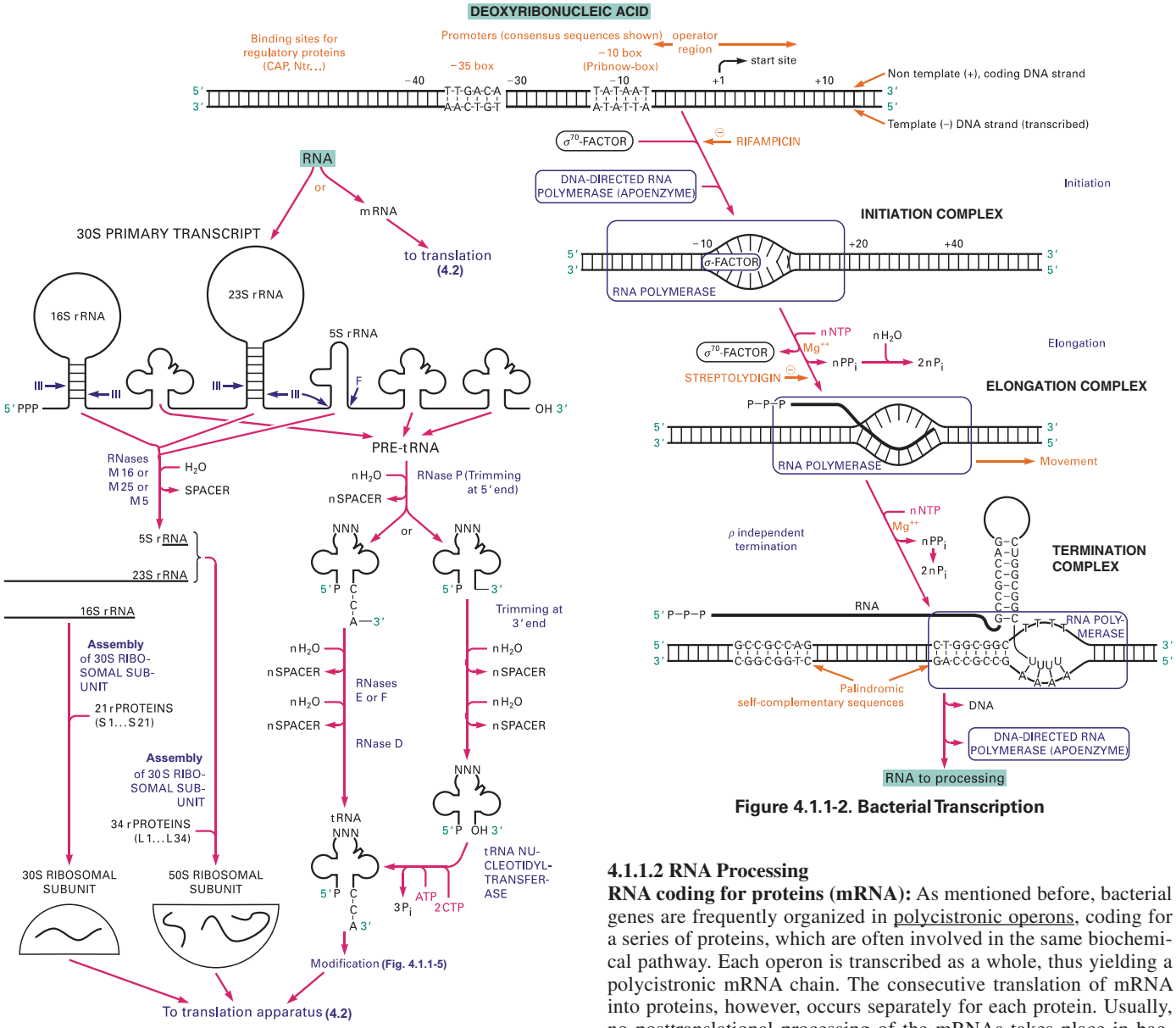


Figure 4.1.1-2. Bacterial Transcription

Figure 4.1.1-4. Processing of Untranslated RNAs

within the elongation complex is not uniform, since purified RNA polymerase has been observed to pause *in vitro* for seconds or even minutes. Duration and frequency of these pauses can be decreased by binding of antitermination factors to RNA polymerase (e.g., NusA and G together with phage lambda N protein). Transcription may also be resumed by action of GreA, which removes a short piece from the 3′ end of the synthesized RNA.

Termination: There are two classes of termination sites:

- ρ-independent DNA sites do not require any additional factor. They consist of a palindromic, G≡C-rich sequence followed downstream by an A=T-rich region. It is thought that the RNA oligo-U transcribed from the A-rich sequence destabilizes the DNA/RNA hybrid. Simultaneously, the RNA forms a G≡C-rich stem-loop structure, which retracts RNA from the transcription bubble and causes the structure to disassemble.
- No common DNA or RNA motifs were found for the much less abundant ρ-dependent termination sites, which require factor ρ for termination. The hexameric factor ρ presumably moves along nascent RNA in 5′→3′ direction until it catches up with the RNA polymerase stalled at a pausing site and releases the RNA from the enzyme.

4.1.1.2 RNA Processing

RNA coding for proteins (mRNA): As mentioned before, bacterial genes are frequently organized in polycistronic operons, coding for a series of proteins, which are often involved in the same biochemical pathway. Each operon is transcribed as a whole, thus yielding a polycistronic mRNA chain. The consecutive translation of mRNA into proteins, however, occurs separately for each protein. Usually, no posttranslational processing of the mRNAs takes place in bacteria. In a few cases, group II introns (4.2.1.3) have been found in bacteria. They have inherent ribozyme activity, which is involved in their splicing.

Untranslated RNAs: *E. coli* cells possess seven separated operons containing the genes for rRNAs and tRNAs. The 30S primary transcripts are about 5500 nucleotides long. They usually comprise one copy each of the 16S rRNA, 23S rRNA and 5S rRNA and several tRNA genes. The individual sequences are cut from the primary transcript while transcription is still going on (Fig. 4.1.1-3).

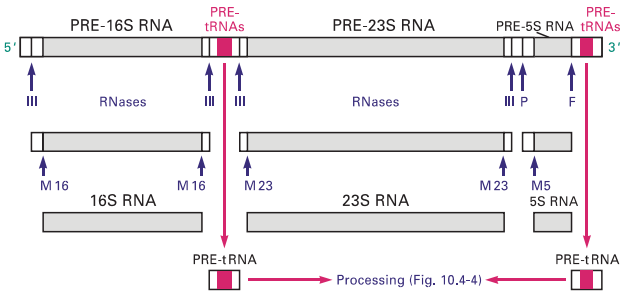


Figure 4.1.1-3. Cleavage of the 30S Primary Transcript

Ribonucleases (RNases) III, P, F cut at definite positions, producing pre-16S, 23S and 5S rRNAs, as well as pre-tRNAs (Fig. 4.1.1-4). Likely, stem-loops in the RNA act as recognition sequences for RNase III. Both the 5′ and the 3′ ends of the pre-rRNAs are trimmed by action of the RNases M16, D, M23 and M5 to their final lengths. Then methylation takes place, yielding N^6, N^6-dimethyl adenine and other methylated bases (Fig. 4.1.1-5).

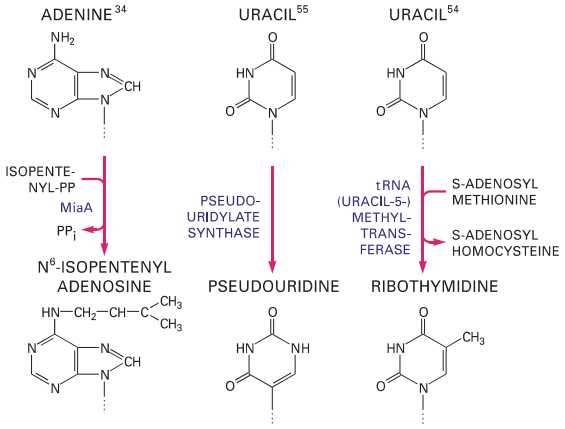


Figure 4.1.1-5. Examples for Modification of tRNA

The self assembly of the large and the small ribosome subunits from rRNAs and ribosomal proteins is a sequential process. Some proteins bind at distinct sites of the rRNAs, causing conformational changes and thus creating proper binding sites (scaffolds) for other ribosomal proteins.

The pre-tRNAs contain additional nucleotide sequences at both ends (Fig. 4.1.1-4), which have to be removed. The 19 nucleotide long 5′ extension is trimmed by RNase P. This enzyme contains a RNA component (M1 RNA) which actually catalyzes the cleavage. Thus, RNase P is a ribozyme. The 3′ extension of tRNA is removed by RNases E or F and D. Finally, the bases of the tRNAs are modified similarly to the eukaryotic tRNAs (4.2.1.8), but smaller in number.

Accuracy of Transcription: The error rate of transcription is about 1 per 10^4 nucleotides, thus considerably higher than in DNA replication (3.8.1.4). This rate is a compromise between speed and accuracy and is apparently tolerable for the following reasons:

- The products of transcription are not transferred to the progeny
- Transcription of an encoding gene takes place repeatedly
- The genetic code contains synonyms for amino acids (2.7)
- Many amino acid substitutions do not affect protein activity.

4.1.1.3 Inhibitors of Transcription

The antibiotic rifampicin inactivates only RNA polymerase of bacteria by blocking the initiation step, while streptolydigin blocks the elongation step. Actinomycin D binds to DNA and inhibits both bacterial and eukaryotic transcription.

4.1.2 Regulation of Bacterial Gene Expression

Bacteria continuously adapt to their environment for successful survival and growth. Consequently, they possess multiple receptors for environmental stimuli and connected regulatory circuits for the metabolic and gene expression control. Metabolic control is the quickest response of a bacterial cell to its environment. It acts on the existing components of the metabolism, the enzymes, their complexes and control compounds. Gene expression control ensures the formation of the proteins and RNAs needed under certain environmental conditions. It takes place at the level of transcription, translation, protein modification and degradation. It requires some time and is employed for middle-term and long-term adaptations.

4.1.2.1 Regulation of Transcription Initiation

Core RNA polymerase, albeit nonspecifically DNA binding, is not able to recognize the promoter region by itself. Binding of a sigma subunit to the core enzyme confers the ability of promoter recognition. Even at this early stage, different sigma subunits provide the cell with the possibility to respond to important environmental stimuli, including depletion of the carbon, nitrogen and iron sources or to heat stress. The seven sigma factors of *E. coli* are listed in Table 4.1.1-2. *Streptomyces avermitilis* has genes for 60 sigma factors.

An outdated concept stated that the σ^{70} recognizes the consensus promoters −35 (TTGACA) −10 (TATAAT) for the so called 'housekeeping', constitutively expressed genes in *E. coli*. However, transcriptional profiling using the DNA array technology revealed, that all genes are regulated by the growth rate, including the housekeeping genes. Moreover, many tightly controlled genes are also σ^{70}-dependent. Furthermore, transcriptional initiation by the sigma subunit binding RNA polymerase is controlled by many additional transcription factors. At certain promoters more than ten regulatory proteins and RNAs are involved in transcriptional control. This way various different, even oppositional stimuli are integrated into a useful gene regulatory response. In most cases gene regulation is not an on/off switch, but rather a continuous 'more or less' procedure.

RNA polymerase binding to promoters can be influenced by regulatory proteins in two different ways:

- Negative control: A repressor protein binds to or near the promoter and prevents RNA polymerase binding or activity.
- Positive control: An activator protein binds to or near the promoter and assists in RNA polymerase binding or transcription initiation.

The same regulator can act as a repressor of a certain gene and as an activator of another gene. The regulatory proteins show an extremely strong binding to their binding site (K_M typically 10^{-13} mol/l with high selectivity). Often DNA binding is mediated by a helix-turn-helix motif.

4.1.2.2 Examples for Gene Regulation in Bacteria

DNA binding regulatory proteins can respond to environmental stimuli (by stimulons, Table 4.1.2-1). In such a case binding of a signal molecule activates or inactivates the regulator. Thus, a repressor effects either repression or de-repression, an activator either activation or deactivation of gene transcription.

Many extracellular stimuli are mediated by so-called two-component regulatory systems. As an example, the redox regulating ArcAB system is shown (Fig. 4.1.2-1). A membrane bound receptor kinase ArcB senses the environmental signal (here a more negative redox state) and autophosphorylates itself at a histidine residue. This phosphate group gets transferred to an aspartate residue of the response regulator protein ArcA. This transfer induces a conformational change in ArcA, so that it can bind to specific sequences in the upstream regulator sequence (URS) of all the operons belonging to the same regulon and thus modulate their transcription.

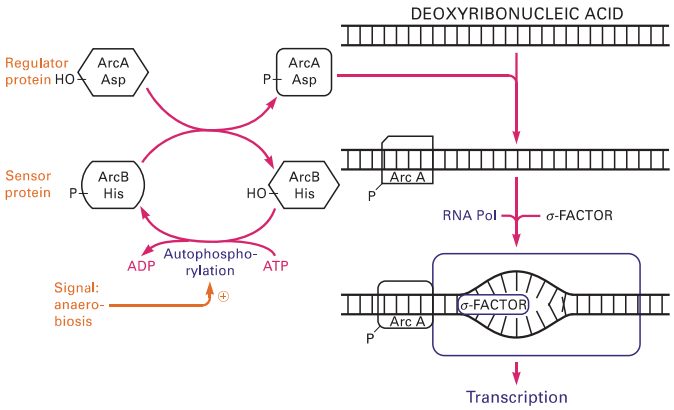


Figure 4.1.2-1. Regulation by the 2-Component System ArcAB

Table 4.1.2-1. Examples of Stimulons (*E. coli*)

Regulon	Responding to	Mechanism
Crp	carbon starvation	cAMP level (see below)
Heat shock	elevated temperature	sigma subunit (see above)
SOS	DNA damage	cleavage of LexA (3.8.2.6)
Ntr	nitrogen availability	NtrB/NtrC system (see below)
SoxRS, OxyR	oxidative stress	redox activation by superoxide (SoxR) or hydrogen peroxide (OxyR)
Fnr, ArcAB, NarXL	oxygen limitation, presence of alternative electron acceptors	oxygen and nitrate detection, redox measurements at the membrane

Control of the *E. coli lacZYA* operon (Fig. 4.1.2-2), encoding the enzymes of lactose metabolism: Two signals, the presence of lactose and the absence of glucose, are integrated at this promoter. In absence of lactose the tetrameric *Lac* repressor binds to several sites of the *lac* promoter, preventing RNA polymerase from initiating transcription. No enzymes for lactose utilization are formed. Once lactose becomes available, its intracellular transglycosylation product allolactose (Gal-1β-6-Glc) is bound by the *Lac* repressor which loses its DNA-binding capacity (Fig. 4.1.2-3). Transcription of the *lac* operon starts.

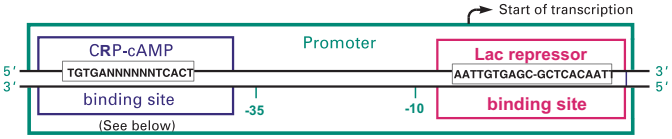


Figure 4.1.2-2. Structure of the *Lac*-Promoter/Operator Region in *E. coli*

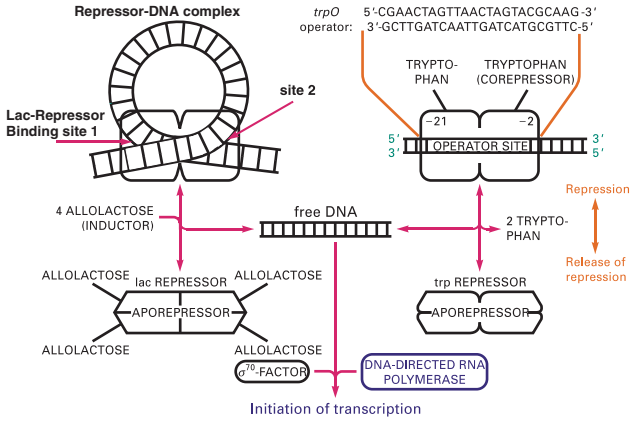


Figure 4.1.2-3. Repression Mechanisms

Additionally the expression of the *lac* operon is positively regulated by Crp (cAMP binding protein, Table 4.1.2-1) in response to the concentration of the preferred carbon source glucose. During glucose starvation (with the need for the utilization of alternative carbon sources including lactose), the cyclic AMP (cAMP, 7.4.2) levels are increased due to interaction of adenylate cylase with components of the phosphotransferase system (PTS, 3.10.4) for sugar uptake and phosphorylation. When the intracellular signal molecule cAMP is bound to Crp, the dimeric protein undergoes a conformational change, which allows it to bind to DNA (Fig. 4.1.2-4). This increases transcription up to 50-fold by facilitating the formation of a stable transcription initiation complex by RNA polymerase. Addition of the preferred nutrient glucose decreases the cAMP level again (catabolite repression) and leads to decreased gene translation.

Repression and attenuation of *trp* loci (coding for enzymes of tryptophan biosynthesis): The repression mechanism is inverse to that of the *Lac* repressor. The free *Trp* repressor has no affinity for DNA and leaves the transcription unaffected. If tryptophan is present, it reacts (as a co-repressor) with the repressor, which consecutively binds to

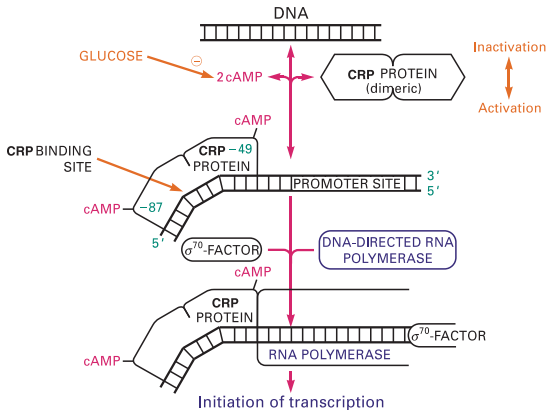


Figure 4.1.2-4. Activation by CAP Controlled Operons

the *trp* gene promoter regions and stops initiation of transcription. Analogous reactions take place at the *trp*P and *aro*H operators, which regulate the expression of other genes of tryptophan biosynthesis.

Additionally, tryptophan biosynthesis is regulated by attenuation (Fig. 4.1.2-5). The presence of the amino acid is sensed by performing a 'test translation'. The 5′ end of the transcribed *trp* mRNA encodes a 14 amino acid long leader peptide with two consecutive tryptophan codons, followed by the attenuator sequences 1, 2, 3, 4 which can form two alternative secondary structures (1•2 / 3•4 and 2•3), the first of which contains a termination hairpin (3•4). Since transcription and translation in prokarya are coupled in the cytoplasm, translation of the mRNA for this leader peptide provides a measurement for the tryptophan concentration in the cell. Depending on the position of the translating ribosome, the mRNA can form two different hairpin structures allowing or preventing further transcription of the *trp* operon.

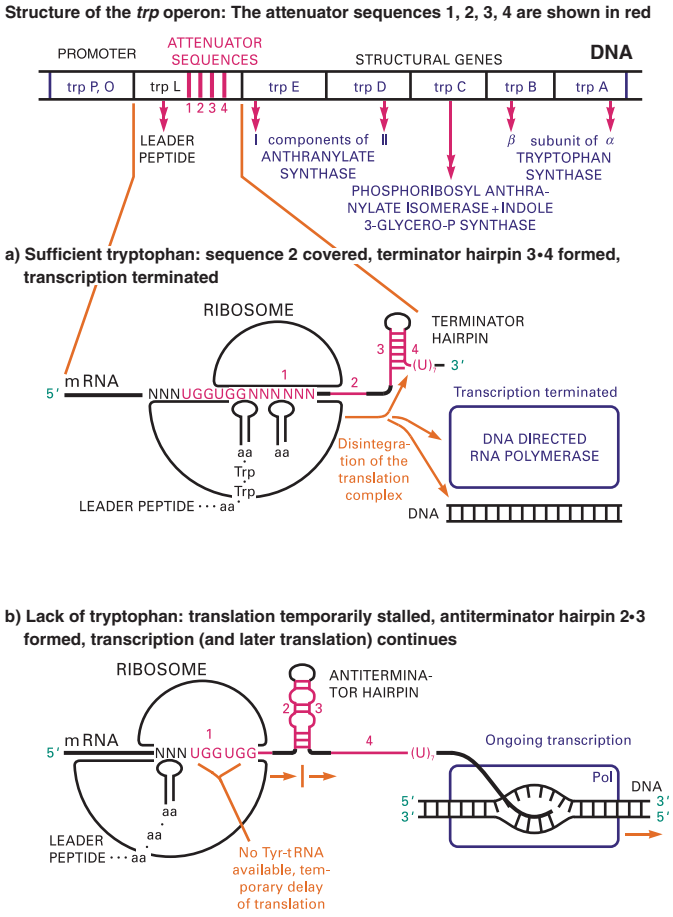


Figure 4.1.2-5. Attenuation of the *trp* Operon

If there is a sufficient supply of tryptophan, the translating ribosome, closely following the RNA polymerase, proceeds to the stop codon at the end of the leader peptide. As a result it covers the first part of the attenuator RNA sequence. This allows the formation of the terminator hairpin (3•4) and results in the termination of transcription. When the level of tryptophan is low, the ribosome, which translates the leader peptide, gets delayed at the two consecutive UGG tryptophan codons. This leaves the attenuator sequence temporarily uncovered and allows the formation of the alternative secondary structure 2•3, which does not contain the terminator hairpin. Therefore transcription can proceed through the structural genes; the enzymes for biosynthesis of this amino acid are produced.

Stringent response (Fig. 4.1.2-6): In various bacteria, starvation for amino acids, carbon sources, iron and fatty acids causes a drastic decrease in transcription of genes involved in RNA formation and protein biosynthesis. Amino acid starvation leads to the accumulation of non-aminoacylated tRNAs. Occupation of the ribosome A-site (4.1.2.3) with such a non-aminoacylated tRNA allows the binding of (p)ppGppp synthase I (RelA) to the large subunit of the ribosome and its activation. The enzyme catalyzes the transfer of the βγ-pyrophosphate of ATP to the 3´-hydroxyl group of GTP or GDP to form (p)ppGpp. Formed pppGpp is converted by 5´-phsophohydrolase into the alarmon (alarm hormone) ppGpp. The cellular ppGpp concentration is controlled by a degrading system consisting of ppGpp 3′ pyrophosphohydrolase (SpoT) and nucleoside-5´-bisphosphate kinase (Ndk). The influence of fatty acid metabolism on this system is mediated by direct interaction of fatty acid synthase with SpoT. The alarmon ppGpp controls multiple genes, often in combination with the transcriptional regulator DskA. Genes for amino acid and nucleotide biosynthesis are induced, while genes of protein biosynthesis, like rRNA and tRNA gene, are repressed.

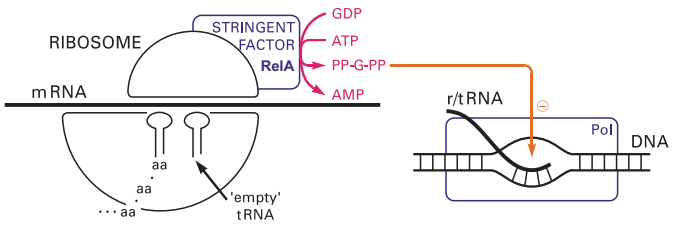


Figure 4.1.2-6. Stringent Response

Literature:

Braeken, K. *et al.* Trends Microbiol. 2006;14:45–54.

Gollnick, P. *et al.* Annu Rev Genet 2005;39:47–68.

Müller-Hill, B. *The lac Operon: a short history of a genetic paradigm.* deGruyter, 1996.

van Hijum, S.A. *et al.* Microbiol. Mol. Biol. Rev. 2009;73:481–509, Table of Contents.

Vassylyev, D.G. Curr. Opin. Struct. Biol. 2009;19:691–700.

4.1.3 Bacterial Protein Synthesis

During translation amino acids are linked by peptide bonds to form long unbranched peptide chains, which grow from their amino (N) terminus to their carboxy (C) terminus. The nucleotide sequence of a mRNA-transcript determines the amino acid sequence of the protein. Each nucleotide triplet of the mRNA codes for one amino acid (2.7.2). The translation of this nucleotide 'language' into amino acid 'language' is mediated by transfer RNAs (tRNAs): Each of these adapters is specifically charged with a specific amino acid and recognizes the corresponding specific codon on the mRNA via its complementary anticodon.

4.1.3.1 Transfer RNAs and Aminoacyl-tRNA Synthetases

Transfer RNAs (tRNAs) are composed of one RNA strand with an average length of 76 nucleotides. The 'cloverleaf' representation of a tRNA secondary structure (Fig. 4.1.3-1) shows base-pairing of complementary regions within the molecule. The 'leaves' are designated as acceptor (or amino acid) arm, D arm, anticodon arm, variable arm and TΨC arm. The highly ordered three-dimensional structure resembles a slender 'L' which is composed of the acceptor helix and the anticodon helix. The -CCA sequence at the 3′ end of the acceptor helix is charged with the appropriate amino acid. The other helix carries the 3-base anticodon, which base pairs with the codon on the mRNA. The nucleotide number in the extra arm differs between 3 and 23. Atypical interactions between bases stabilize the structure in addition to the Watson-Crick base pairing. tRNAs contain numerous base modifications, e.g., dihydrouridine (D) and pseudouridine (Ψ, cf. 4.1.1.2 and 4.2.1.8). Some modifications influence the recognition between anticodon and codon. tRNAs are classified according to the length of the extra arm: Class I tRNAs have very short ones (e.g., $tRNA^{Cys}$), class II tRNAs have long ones (e.g., $tRNA^{Ser}$).

Aminoacyl-tRNA synthetases (also named 'ligases') catalyze the aminoacylation of tRNAs with amino acids (aa). An ester bond is formed with the 3′ hydroxyl group of the terminal adenosine nucleotide of tRNA. The charging reaction (Fig. 4.1.3-2) is driven by hydrolysis of both high energy bonds of ATP:

$$\left.\begin{array}{r}\text{Amino acid} + \text{ATP} = \text{aminoacyl-AMP} + \text{PPi}\\ \text{Aminoacyl-AMP} + \text{tRNA} = \text{aminoacyl-tRNA} + \text{AMP}\\ \text{PP}_i + \text{H}_2\text{O} = 2\ \text{P}_i\ \text{(by inorganic pyrophosphatase)}\end{array}\right\}\ \begin{array}{l}\text{In total}\\ \Delta G'_0 = -29\ \text{kJ/mol}\end{array}$$

Each aminoacyl-tRNA synthetase is highly specific for a single amino acid. Usually one enzyme aminoacylates all tRNAs specific for the same amino acid (isoacceptor-tRNAs, see 'Genetic code', 2.7). Based on the structure of the tRNA binding region, aminoacyl-tRNA synthetases are divided into two classes. Class I enzymes ligate amino acids to the 2′-OH of the tRNA. Class II enzymes transfer the aminoacyl group directly to the 3′-OH. Aminoacyl-tRNA synthetases recognize the corresponding (cognate) tRNA by various identity elements, which can be the anticodon, tertiary structural features, certain bases in the acceptor helix, or specific modifications. If a wrong tRNA is bound by an aminoacyl-tRNA synthetase, the aminoacylation is prevented or the already misacylated tRNA is immediately deacylated.

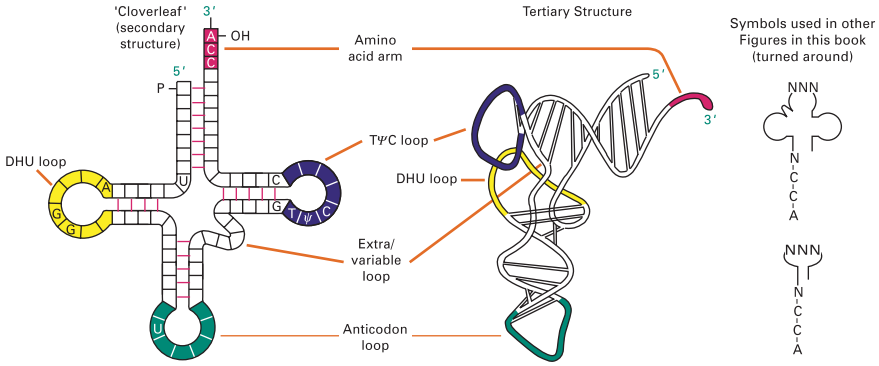


Figure 4.1.3-1. Structure of tRNAs

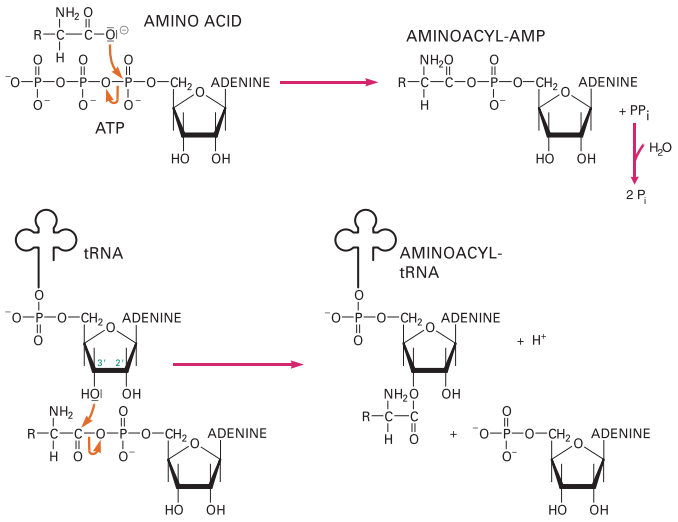


Figure 4.1.3-2. Aminoacylation of tRNAs

This proofreading is of high importance for translational fidelity. Only aminoacyl-tRNA synthetases for amino acids, which are easily discriminated (e.g., tyrosine) lack the proofreading ability.

In bacteria, the first amino acid of a growing peptide chain is formyl-methionine (fMet). A specialized initiator tRNA ($tRNA_I^{fMet}$) is charged with methionine by the normal methionine-tRNA synthetase. Then the methionine is formylated by 10-formyl-tetrahydrofolate in order to block the N terminus during peptide synthesis.

4.1.3.2 Bacterial Ribosomes and Translational Factors

Each growing *E. coli* cell contains 15,000 or more ribosomes which account for almost 25% of its dry weight. They are composed of 65% rRNA and 35% protein. Bacterial ribosomes (70S, 2700 kDa) are composed of a large and a small subunit, each containing RNA and proteins (Table 4.1.3-1). The ribosome self-assembles during synthesis of its components. The shape is shown in Figure 4.1.3-3. Its crystal structure was elucidated only a few years back. The high resolution structure thus confirmed what has been suspected for more than a decade – the ribosome is a ribozyme. The major catalytic steps, including the peptidyltransferase step, are performed by the RNA part of the macromolecular machine.

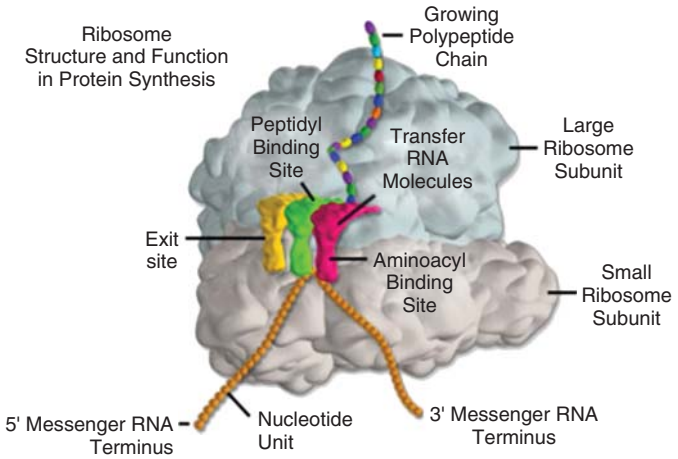


Figure 4.1.3-3. Structure of the Bacterial Ribosome

The mRNA binding site of the ribosome is located in a 'channel' of the small subunit, whereas the catalytic activity of the ribosome (the peptidyl transferase) is located in the large subunit. tRNAs are bound by both subunits at three different sites:

- Aminoacyl- (A-)site, binds the incoming aminoacyl-tRNA (except $fMet\text{-}tRNA_I^{fMet}$, 4.1.3.4)
- Peptidyl- (P-)site, binds at initiation $fMet\text{-}tRNA_I^{fMet}$ and, during elongation, the peptidyl-tRNA
- Exit- (E-)site, ejects the deacylated tRNA

Usually, 10–100 ribosomes in series (called a polysome) migrate along the same mRNA.

Table 4.1.3-1. Composition of *Bacterial* Ribosomes (*E. coli*, 2700 kDa)

	Small subunit (30S), 900 kDa	Large subunit (50S), 1800 kDa
RNA	16S rRNA, 1542 nt	23S rRNA, 2904 nt + 5S rRNA, 120 nt
Protein	21 polypeptides	33 different polypeptides, a total of 36

A number of protein factors are required for protein synthesis. They are involved in initiation complex formation, aminoacyl-tRNA binding, translocation and termination (Table 4.1.3-2).

Table 4.1.3-2. Translation Factors in *E. coli*

Factor	Mol. Mass (kDa)	Function
Initiation factor 1 (IF-1)	8	prevents premature binding of tRNA to the A site
Initiation factor 2 (IF-2)	97	directs initiator-tRNA to the ribosomal P-site, GTPase
Initiation factor 3 (IF-3)	21	prevents association of the ribosomal subunits, directs mRNA to 30S subunit
Elongation factor Tu (EF-Tu)	43	directs aa-tRNA to ribosomal A site, GTPase (very abundant *E. coli* protein)
Elongation factor Ts (EF-Ts)	28	promotes GDP-release from EF-Tu
Elongation factor G (EF-G)	77	catalyzes ribosomal translocation, GTPase
Release factor 1 (RF-1)	36	causes termination at UAA and UAG stop codons
Release factor 2 (RF-2)	38	causes termination at UAA and UGA stop codons
Release factor 3 (RF-3)	46	stimulates RF-1 and RF-2, GTPase
Ribosomal recycling factor (RRF)	21	promotes together with EF-G the release of peptidyl-tRNA and mRNA

4.1.3.3 Bacterial Translation (Polypeptide Synthesis, Fig. 4.1.3-4)

Initiation: Bacterial translation starts soon after the ribosome binding sequence and the translation initiation codon (see below) of the mRNA have been synthesized by RNA polymerase during transcription. The mRNA sequence at the ribosome binding site is complementary to the 3´end of the 16s rRNA. This leads to the attachment of the 30S subunit to the mRNA via base pairing. This process additionally requires IF3. After factor IF1 has bound and blocked the aminoacyl-(A) site, N-formyl-methionyl-tRNA is guided by the GTP-carrying form of IF2 to the start codon in the peptidyl (P)-site. IF3 is now leaving the complex and the 50S subunit can enter the complex. This promotes hydrolysis of IF1 and IF-2 bound GTP and leads to the subsequent release of the initiation factors. This reaction is the rate limiting step in translation.

Elongation: All further aminoacyl-tRNAs are delivered to the ribosomal aminoacyl-(A) site by the elongation factor Tu (EF-Tu) in a ternary complex consisting of aminoacyl-tRNA•EF-Tu•GTP. The only exception is the integration of selenocysteinyl-tRNA, described below.

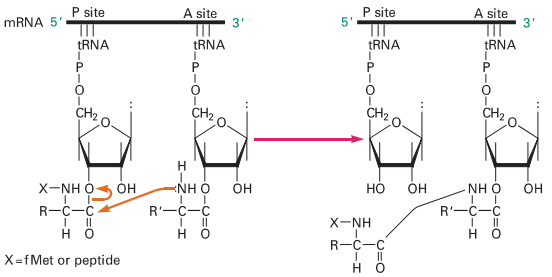


Figure 4.1.3-5. Principle of Ribosomal Protein Synthesis

After GTP hydrolysis, EF-Tu•GDP leaves the ribosome. For recycling, EF-Ts in complex with EF-Tu promotes the release of GDP from EF-Tu. The subsequent binding of GTP induces a large conformational change in EF-Tu, which then can bind aminoacyl-tRNA again. Peptidyltransferase is an enzyme activity of the 23S rRNA. It catalyzes the transfer of fMet or of the peptidyl chain in the following cycles from the tRNA at the P-site to the free amino group of the aminoacyl-tRNA at the A-site (Fig. 4.1.3-5). The reaction is an RNA-catalyzed nucleophilic displacement of the P site tRNA by the amino group of the aminoacyl-tRNA at the A site.

Next, the tRNAs and the mRNA move one codon site relative to the 30S ribosome. The peptidyl-tRNA is now at the P-site and the deacylated tRNA leaves the ribosome from the E-site. This step requires hydrolysis of GTP bound to EF-G. Then EF-G dissociates

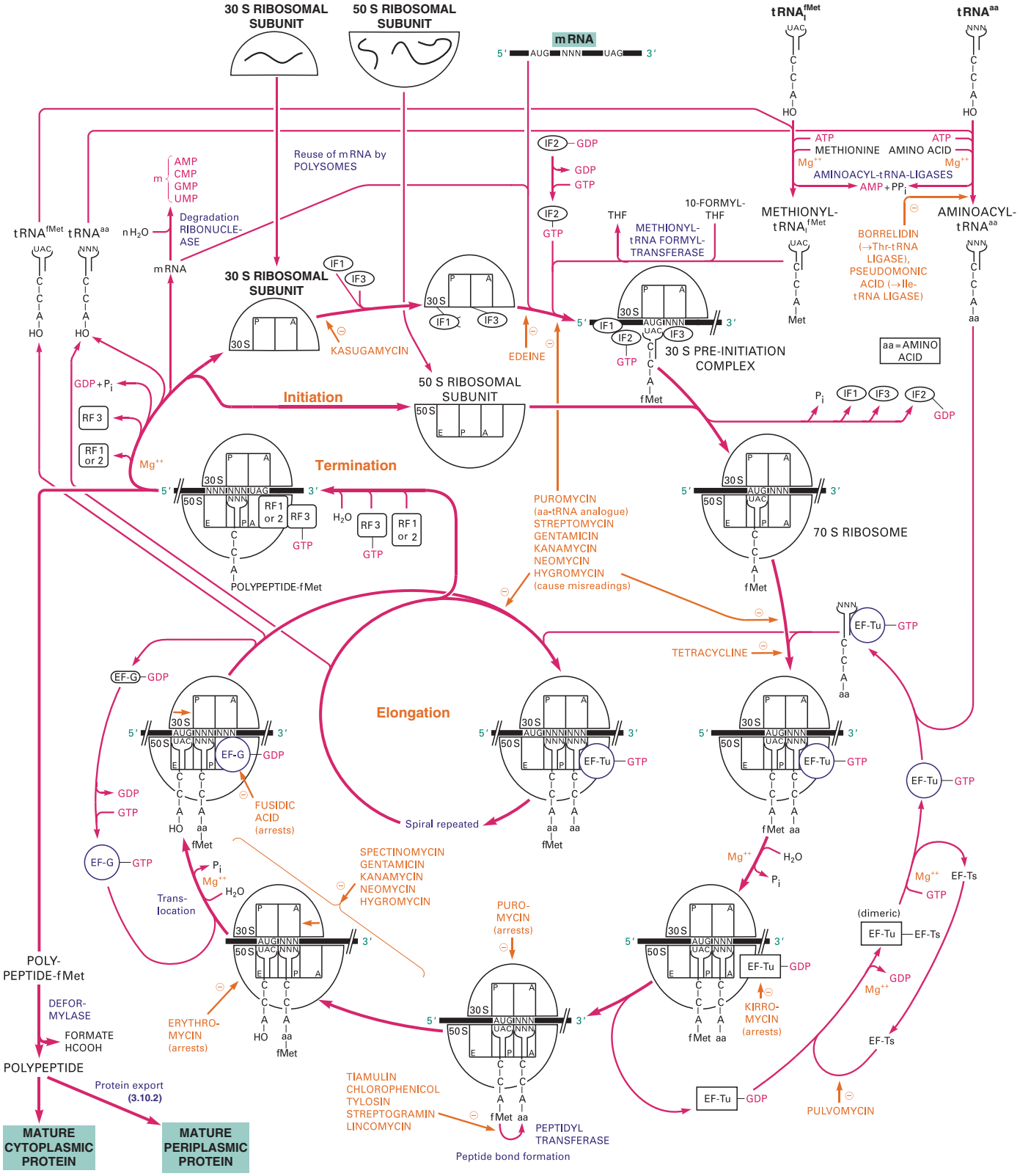


Figure 4.1.3-4. Bacterial Translation

from the complex. A new aminoacyl-tRNA can occupy the free A-site now, starting another elongation cycle. It is assumed, that these steps involve conformation changes and movements of the ribosomal subunits relative to each other. Since EF-G•GTP has a similar tertiary structure as the aminoacyl-tRNA•EF-Tu•GTP complex, it may interact with the same ribosome site during this process.

Termination: If a stop codon (UAA, UGA, UAG) arrives at the ribosomal A-site, protein synthesis terminates. Instead of another aminoacyl-tRNA, release factors RF-1 or RF-2 (depending on the codon) together with RF-3•GTP bind to the stop codon. After GTP hydrolysis, the polypeptide chain leaves the tRNA. Thereafter, EF-G and the ribosomal recycling factor (RRF) catalyze the liberation of mRNA and tRNA. Then the ribosomes separate into their subunits. The initiation factor IF-3 prevents their re-association until it is released from the pre-initiation complex.

The formyl group and, in ca. 50% of all proteins also the N-terminal methionine, are enzymatically removed. Folding mechanisms, which produce the correct tertiary structure of the protein, are dealt with in 4.5.1.

Altogether, formation of a single peptide bond ($\Delta G_0' = 21$ kJ/mol) requires hydrolysis of 4 energy-rich phosphate bonds ($\Delta G_0' = -124$ kJ/mol): 1 ATP → AMP for tRNA charging + 2 GTP → 2 GDP for delivering the aminoacyl-tRNA to the A-site and translocation. This makes protein synthesis the most energy consuming process of the living cell.

The overall error rate of translation is ca. 1/10 000. Consequently, the probability for correct synthesis of a protein with 300 amino acids is 0.97. Many inhibitors, most of them antibiotics, act at various steps of the protein synthesis (Fig. 4.1.3-4).

4.1.3.4 Selenocysteine

A number of bacteria, archaea and eukarya synthesize proteins which contain the amino acid selenocysteine. Selenocysteine (abbreviated Sec or U) is present in the catalytic center of some oxidoreductases and contributes to the catalytic mechanism by its high reactivity (Table 4.1.3-3). It is introduced into these proteins by an unusual decoding of mRNA. This process represents an extension of the genetic code. In bacteria, selenocysteine is synthesized in a pyridoxal-phosphate dependent reaction from serine bound to the special $tRNA^{Sec}$ and selenophosphate, which contains the only selenium-phosphorus bond known in biochemistry, via an aminoacryl intermediate (Fig. 4.1.3-6).

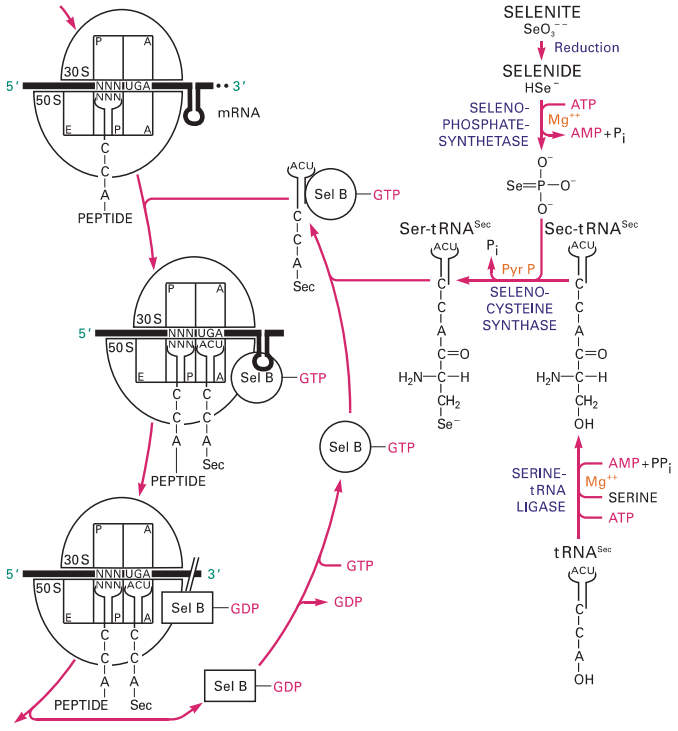


Figure 4.1.3-6. Biosynthesis of Selenocysteine (Sec)

During translation, selenocysteine is inserted into the polypeptide chain at an UGA codon, which in other contexts functions as a stop codon. In bacteria, a mRNA stem-loop structure adjacent to this codon determines UGA as selenocysteine codon. Instead of EF-Tu•GTP, a special elongation factor SelB•GTP, which recognizes and binds to the mRNA stem-loop, delivers selenocysteyl-$tRNA^{Sec}$ to the ribosome. In eukarya and archaea, the corresponding special mRNA structure is located in the 3′ untranslated region of the transcript.

Table 4.1.3-3. Proteins Containing Selenocysteine (Selection)

Bacteria	Archaea	Eukarya
Selenophosphate synthetase	Selenophosphate synthetase	Selenophosphate synthetase
Formate dehydrogenase	Formate dehydrogenase	Glutathione peroxidase
Hydrogenase	Hydrogenase	Thioredoxin reductase
Glycine reductase	Heterodisulfide dehydrogenase	5′-Tetraiodothyronine deiodinase
	Formyl-methanofurane dehydrogenase	Selenoproteins P and W

Literature:

Baron, C., Böck, A. *tRNA*. Washington D.C.: ASM Press, 1995, p. 15.

Berk, V., Cate, J.H. Curr. Opin. Struct. Biol. 2007;17:302–309.

Böck, A. *et al*. Trends Biochem. Sci. 1991;16:463–467.

Dunkle, J.A., Cate, J.H. Annu. Rev. Biophys. 39: 227–244.

Ibba, M., Soll, D. Annu. Rev. Biochem. 2000;69:617–650.

Johansson, M. *et al*. Curr. Opin. Microbiol. 2008;11:141–147.

Low, S.C., Berry, M.J. Trends Biochem. Sci. 1996;21:203–208.

Moore, P.B. J. Biol. 2009;8.

Stadtman, T.C. Annu. Rev. Biochem. 1996;65:83–100.

4.1.4 Degradation of Nucleic Acids

Nucleases catalyze the cleavage of phosphodiester bonds in DNA (DNases) or in RNA (RNases). DNases play a decisive role during DNA synthesis, repair and recombination (3.8.1, 3.8.2, 3.9.2), for relief of obstructive superhelical tension and as a defense measure in restriction of foreign DNA (4.1.4.2). RNases are of importance in RNA processing (4.1.1.1, 4.2.1.3) and also regulate the transcription by degradation of mRNA (4.2.5). Nucleases either remove terminal nucleotides (exonucleases) or act inside of the nucleic acid molecule (endonucleases).

4.1.4.1 Exodeoxyribonucleases (Exo-DNases, Table 4.1.4-1)

Exo-DNases are characterized by their cleavage direction, their preference for a single stranded or a double stranded substrate and by producing mono- or (more rarely) oligonucleotides. The nucleases may dissociate after each catalytic event (non-processive or distributive action) or they may remain bound to the polymer until several successive reaction cycles are completed (processive or non-distributive action).

The reaction scheme of *E. coli* exonuclease III is schematically shown in Figure 4.1.4-1. The enzyme is multifunctional. It acts as an endonuclease specific for apurinic DNA sites (3.8.2.3) and as a 3′ phosphatase. Bacterial DNA polymerase I exerts a 3′ → 5′ exo-DNase activity for proofreading (similarly to eukaryotic DNA Pol δ and ε) and additionally a 5′ → 3′ exo-DNase/RNase function for DNA repair and for removal of Okazaki RNA primers (3.8.2.3, 3.8.1.3).

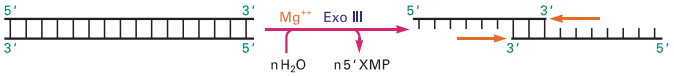


Figure 4.1.4-1. Reaction Scheme of *E. coli* Exonuclease III

4.1.4.2 Endodeoxyribonucleases (Endo-DNases, Table 4.1.4-2)

The endonucleases often show a strong preference for either single stranded or duplex DNA. Endonucleases that function in repair of lesions identify the damaged DNA site and incise (nick) the DNA at one side of the lesion as a first step towards excision. A second principal characteristic of these enzymes is the recognition of DNA sequences. A striking example is the cytosine-specific cleavage by T4 endonuclease IV. Pancreatic DNase, *E. coli* endonuclease I or spleen DNase produce oligonucleotide digests with characteristic sequence patterns at the 3′ and 5′ termini

Restriction endonucleases occur in a variety of microorganisms. More than 2000 of them have been identified so far. They recognize sequences of 4 ... 8 nucleotides in a DNA duplex with extraordinary accuracy and cleave both strands. The organisms always produce a companion DNA methyltransferase, which recognizes the same sequence in endogeneous DNA and modifies it immediately after replication by methylation of A or C residues. This protects the organism's own DNA from degradation by the restriction enzyme. Thus, the restriction endonuclease and the cognate methyltransferase form a restriction-modification (R-M) system. Since it cleaves (restricts) infecting DNAs (e.g., viruses) and thus prevents them from parasitizing the cell, the R-M system is also called the 'immune system of the microbes'.

At least four different kinds of R-M systems exist:

- Type I enzymes carry methylase and nuclease activity on the same protein and require Mg^{++}, ATP and S-adenosylmethionine for cleavage. They cleave randomly and remotely (> 400 bp) from the recognition sequence.
- Type II enzymes recognize mostly palindromic nucleic acid sequences and cleave within or near these sequences. (Palindromic sequences repeat each other at the other duplex strand in inversed order, resulting in twofold rotational symmetry. For examples see Fig. 4.1.4-2.) The enzymes require Mg^{++} for activity. Their homodimeric structure corresponds to the palindromic substrate. Many of them generate 'sticky end' ('cohesive end') duplex fragments with 5′ protruding termini (e.g., *Eco*RI) or with 3′ protruding tails (e.g., *Hha*I). Other enzymes cleave at the center of the recognition sequence and produce 'blunt end' fragments (e.g., *Hae*III). The cognate methylase is a separate enzyme. The cleavage specificity is schematically shown in Figure 4.1.4-2. These enzymes are indispensible tools for molecular cloning techniques and for DNA sequence analysis. They are named by the 3-letter abbreviation of the source organism.
- Type IIS enzymes recognize asymmetric sequences of 4–7 bp length. They cleave at a defined distance of up to 20 bp to one side of their recognition sequence.
- Type III enzymes have similar characteristics as Type I enzymes, but cleave at specific sites only a short distance (24–26 bp) away from the recognition sequence.

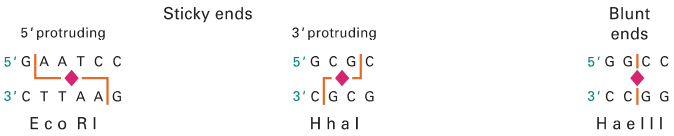


Figure 4.1.4-2. DNA Cleavage by Type II Restriction Endonucleases
(♦ = twofold symmetry axis)

4.1.4.3 Ribonucleases (RNases, Tables 4.1.4-1 and 4.1.4-2)

Similarly to DNases, RNases differ by exo- and by endo-activity, preferences for termini (exo-enzymes) or in some cases for specific sequences (endo-enzymes). Some nucleases even cleave both DNA and RNA.

A number of RNase type reactions are not catalyzed by proteins, but rather by RNA sequences (ribozymes). For example, the RNA component (M1 RNA) of RNase P catalyzes the processing of untranslated prokaryotic RNAs, like tRNAs. (4.1.1.2, as well as some reactions in eukarya); the protein component has only assistant function. In a number of cases, eukaryotic group I or II introns are removed from the rRNA by its own action (self-splicing, e.g., in Tetrahymena, 4.2.1.3). It is speculated that RNA catalysis and self-replication preceded enzyme-protein catalysis during evolution.

Literature:

Bloch, K.D. *Curr Protoc. Immunol.* Chapter 10, Unit 10.18. New York: Wiley, 2001.

Kornberg, A., Baker, T.A. *DNA Replication.* New York: Freeman, 1992.

Lin, S.M., Roberts, J.R. *Nucleases.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1985.

Roberts, R.J., Macelis, D. Nucleic Acids Res. 1996;24:223–235.

Table 4.1.4-1. Examples for Different Types of Exonucleases

Specificity for	3′ → 5′ direction	Cleavage in 5′ → 3′ direction	Either direction
Single stranded DNA	*E. coli* exonuclease I, mammalian DNase III	exo-DNase (phage SP_3 encoded)	*E. coli* exonuclease VII
Double stranded DNA	*E. coli* exonuclease III	exo-DNase (phage λ encoded), mammalian DNase IV	*E. coli* exonuclease V (ATP-dependent)
DNA or RNA	Venom exonuclease	Spleen exonuclease	
RNA	Exo-RNase II, RNaseQ, RNase BN	Yeast RNase	

Table 4.1.4-2. Examples for Different Types of Endonucleases

Specificity for	3′ P endproducts	Cleavage to 5′ P endproducts	other end products
Single stranded DNA	*Aspergillus* DNase K_1	DNase IV (phage-T_4-encoded), yeast DNase	Crossover junction endo-RNase (acts only on Holliday junctions)
Double stranded DNA	pancreatic DNase II	Type I, II and III restriction enzymes (10.7.2), pancreatic DNase I	
DNA or RNA	Micrococcal nuclease, spleen endonuclease	*Aspergillus* nuclease S_1, Mung bean nuclease, potato nuclease	
RNA-DNA duplex		Exo-RNase H (acts on RNA : DNA hybrids, degradation of Okazaki fragments)	
RNA	RNase T_2 = RNase II, pancreatic RNase Polynucleotide phosphorylase	RNase III and RNase P (processing of tRNA and rRNA precursors)	*Bacillus subtilis* RNase (yields 2′,3′-cyclic phosphates)

4.2 Protein Biosynthesis in Eukarya

Röbbe Wünschiers

4.2.1 Eukaryotic Transcription

Transcription is a process that transcribes genetic information from DNA into RNA by the polymerization of ribonucleotide precursors into an RNA molecule. (For a drawing of the principle see Fig. 4.1.1-1). In eukarya, this takes place in the nucleus, in mitochondria and chloroplasts. Transcription is not self-contained. There is a close connection between transcription and other nuclear processes like DNA replication and DNA repair. The initial opening of the chromatin structure is not dealt with here.

4.2.1.1 RNA Polymerases (Pol, Table 4.2.1-1)

Transcription is performed by DNA-directed RNA polymerases. Unlike DNA polymerases, RNA polymerases do not need a primer to start the reaction. While bacteria contain only one RNA polymerase (4.1.1.1), there are three different DNA-dependent RNA polymerases I, II, and III (Pol I, II, and III) in eukaryotic cells, which transcribe different genes. In higher plants, two additional polymerases, Pol IVa and IVb exist, which are involved in small RNA-mediated gene silencing.

Pol I, II, and III are multi-subunit complexes: two large polypeptides (ca. 200 and 140 kDa) are associated with about 12 smaller subunits, some of which are common to Pol I, II, and III. For example, Pol II subunits Rpb5, Rpb8, Rpb10, and Rpb12 are shared among all three eukaryotic DNA-dependent RNA polymerases. Additionally, there are different Pol in mitochondria and chloroplasts.

The basic components of the transcriptional machinery were well conserved in evolution. The order of the transcription complex assembly is basically the same for yeast, *Drosophila* and humans. The two large polymerase subunits are even homologous to the two largest subunits of the *E. coli* RNA polymerase (4.1.1.1).

Table 4.2.1-1. Eukaryotic Nuclear RNA Polymerases

Polymerase	Function
Pol I	synthesizes precursors of ribosomal RNA (28 S, 18 S, 5.8 S rRNA)
Pol II	transcribes all protein-coding genes (yielding mRNA); genes for snRNAs (U1, U2, U3, U4, U5), and genes for miRNAs and snoRNA
Pol III	transcribes genes for tRNA, 5S rRNA, snRNA (U6), in addition transcribes genes of RNase P RNA, RNase MRP RNA, 7SL RNA, 7SK RNA, Vault RNAs, Y RNAs
Pol IV and V (higher plants)	maintain small RNA-mediated gene silencing

4.2.1.2 mRNA Transcription by RNA Pol II (Fig. 4.2.1-1)

mRNA synthesis is a complicated and time-consuming process which can take up to 40 minutes for a large gene (e.g., fibronectin). Since all proteins with their central role in cellular functions and structure are synthesized via mRNA, this process must be strictly regulated at different levels: at DNA binding sites (enhancers, silencers), via proteins (basic and specific transcription factors), by modification (capping, polyadenylation) and processing (splicing). The regulation is discussed in detail in 4.2.2.

Transcription factors: These proteins play essential roles in initiation and elongation. Only the so-called basic (= general) transcription factors (required for the transcription of virtually all genes) are listed in Table 4.2.1-2 and shown in Figure 4.2.1-1. Their size in the Figure does not always reflect the actual proportions. They provide the low-level, 'basic' rate of transcription. (For the regulating, gene-specific transcription factors see Table 4.2.2-2.) Although the basic transcription factors and Pol II accurately initiate transcription *in vitro*, mediator (MED), a large complex of >20 subunits is required for transcription from most Pol II promoters *in vivo*.

Transcription factors often consist of a modular arrangement of distinct functional domains (for details, see 4.2.2.2):

- DNA-binding domains, e.g., zinc fingers or helix-turn-helix
- Transactivation domains, which mediate cooperative associations with other proteins

Initiation: In order to start transcription, a core promoter element at the DNA, e.g., the TATA box (4.2.2.1) must be recognized and bound by the TATA-binding protein (TBP) subunit of the basic transcription factor TFIID, introducing sharp kinks into the DNA. TFIIB and (in some cases) TFIIA stabilize this interaction, forming a pre-initiation complex. Frequently, these steps are controlled by activation or repression mechanisms (4.2.2).

In the case of promoters not containing the TATA-sequence, the initiator motif of DNA (Inr, which encompasses the start site) can mediate initiation. It is recognized by Inr-binding proteins (such as TFII-I to which TFIID associates) or by the subunits TAF_{II} 250 and TAF_{II} 150 of TFIID. This commonly occurs with 'housekeeping' genes (4.2.2.3).

Pol II together with transcription factors TFIIE, IIF and IIH is then recruited to form the initiation complex. TFIIB, assisted by TFIIF, acts as a bridge to Pol II. This is the rate-limiting step. At this stage the DNA strands start to become separated in an ATP requiring reaction forming an 'open initiation complex'. The MED complex facilitates the formation of the pre-initiation complex by direct interaction with Pol II and basic transcription factors.

Table 4.2.1-2. Basic Transcription Factors for RNA Polymerase II

Name	Species	Subunits	Mol. mass (kDa)	Functions / Properties
TFIIA	human yeast	3 2	37, 19,13 32, 14	Initiation, stabilizes TFIID association with promoter by increasing affinity of TATA-binding protein (TBP) subunit for the TATA box, counteracts repression of negative cofactors.
TFIIB	human yeast (TFe)	1 1	35 38	Required for binding of RNA Pol II to the initiation complex; functions in transcription start site selection and in stabilizing TFIID-promoter binding;. Target for steroid hormone receptors (7.7). Two domains, one interacts with TBP, the other with the small subunit of TFIIF.
TFIID	human yeast (TFd)	TBP + 12 TAF_{II} TBP + > 8 TAF_{II}	38 15 … 250 27 18 … 250	Central role in transcriptional activation. Initiates assembly of pre-initiation complex either by binding of TBP (TATA-binding protein) to the TATA-box of DNA or by binding of TAF to other promoter sequences. Individual TAF_{II} coactivators (TBP associated factors) are the specific target for many transcriptional activators and repressors which modulate TBP/TFIID binding to DNA (11.4.3). DNA topoisomerase I is associated with TFIID.
TFIIE	human yeast (TFa)	2 2	56, 34 66, 43	Necessary for recruitment of TFIIH, stimulates TFIIH dependent kinase, ATPase and helicase activity. Involved in DNA strand separation. Not necessary for transcription at all promoters. Probably an $\alpha_2\beta_2$ tetramer.
TFIIF	human yeast (TFg)	2 3	70, 30 105, 54, 30	Essential for initiation of transcription and transcript elongation (suppresses pausing). Activated via phosphorylation by TAF_{II} 250. TFIIF increases affinity of Pol II for TBP-TFIIB-promoter complex and is required for recruitment of TFIIE/TFIIH to the pre-initiation complex. Significant homologies to bacterial σ factors.
TFIIH	human yeast (TFb)	> 9 9	89 … 34 105 … 38	Shows helicase, DNA-dependent ATPase and kinase (= TFIIK) activity. Catalyzes phosphorylation of Pol II CTD, which plays a role in promoter clearance and possibly also in the elongation phase. Involved in DNA excision repair (3.9.2.3).
TFII I	human yeast	1 –	45 –	Similar to TFIIA, binds to DNA initiator motif (Inr), may help forming an alternative pre-initiation complex; interacts with regulatory proteins.
TFIIS = S II	human	1	38	Transcription elongation factor, facilitates passage of Pol II through pause sites by removing a sequence from the 3′ end of the nascent mRNA transcript. Highly conserved.
S III /elongin	human	3	110, 18, 15	Important for elongation (suppresses pausing) and termination. Activated by phosphorylation. Partial homology to bacterial protein.

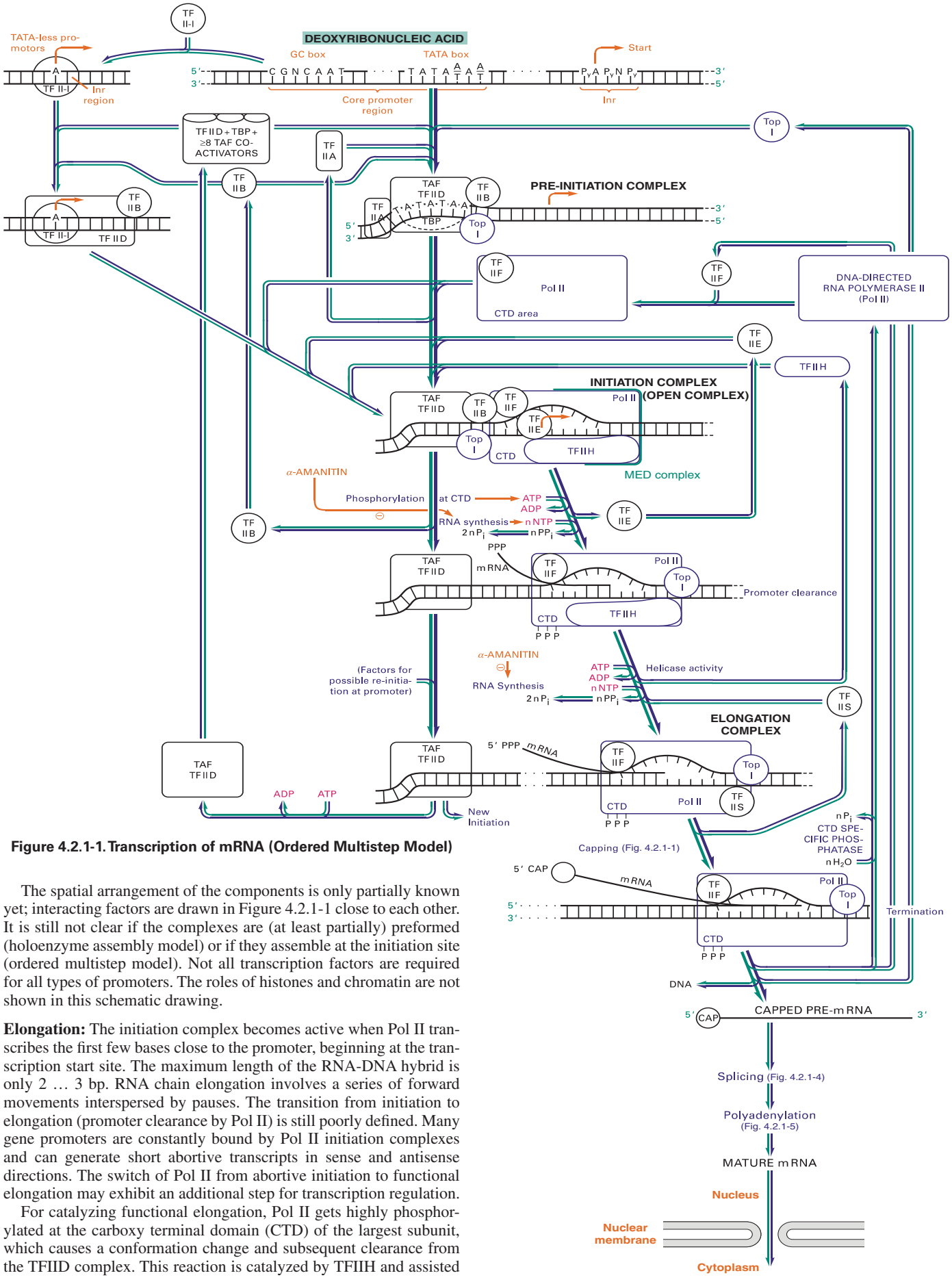


Figure 4.2.1-1. Transcription of mRNA (Ordered Multistep Model)

The spatial arrangement of the components is only partially known yet; interacting factors are drawn in Figure 4.2.1-1 close to each other. It is still not clear if the complexes are (at least partially) preformed (holoenzyme assembly model) or if they assemble at the initiation site (ordered multistep model). Not all transcription factors are required for all types of promoters. The roles of histones and chromatin are not shown in this schematic drawing.

Elongation: The initiation complex becomes active when Pol II transcribes the first few bases close to the promoter, beginning at the transcription start site. The maximum length of the RNA-DNA hybrid is only 2 … 3 bp. RNA chain elongation involves a series of forward movements interspersed by pauses. The transition from initiation to elongation (promoter clearance by Pol II) is still poorly defined. Many gene promoters are constantly bound by Pol II initiation complexes and can generate short abortive transcripts in sense and antisense directions. The switch of Pol II from abortive initiation to functional elongation may exhibit an additional step for transcription regulation.

For catalyzing functional elongation, Pol II gets highly phosphorylated at the carboxy terminal domain (CTD) of the largest subunit, which causes a conformation change and subsequent clearance from the TFIID complex. This reaction is catalyzed by TFIIH and assisted

by TFIIE. The CTD domain contains a tandemly repeated heptapeptide $(YSPTSPS)_n$ (one letter code for amino acids, n = 27 in yeast; 43 in *Drosophila* and 52 in mammals). The respective serine residues are reversibly phosphorylated during the transcription cycle. Serine-5 phosphorylation (Ser5P) along the CTD repeats is an indication of early transcription elongation, whereas serine-2 phosphorylation (Ser2P) takes place in later-stage elongation. Since the MED complex interacts with the unphosphorylated form of CTD, Pol II dissociates from the MED complex when elongation starts.

Upon clearing from the TFIID-complex, topoisomerase I presumably moves to the elongation complex and facilitates elongation. Also, TFIIH exerts its ATP dependent helicase activity, after TFIIE (which inhibits this activity) has left the complex. Highly supercoiled DNA does not require this activity, possibly due to its inherent energy content. Purified Pol II alone only transcribes 1.5 … 5 nt/sec. By action of elongation factors (TFIIS, TFIIF and elongin / SIII), transcription rates *in vivo* increase to 20 … 33 nt/sec.

Termination: The transcription machinery continues for 0.5 to 2 kb beyond (downstream) of the poly(A) signal (see Fig. 4.2.1-5) and then dissociates (termination). No generally conserved termination regions have been identified in DNA transcribed by Pol II so far. For several characterized genes of higher eukarya, transcription pause sites in close proximity to the polyA signal slow down elongating Pol II, enabling 5′-3′ exonucleases to bind to the 3′ cleavage product and to degrade the respective downstream transcript. A specific transcription termination sequence found in the human gastrin gene ($5'\text{-}T_9A_2T_5AT_4AT_4AT_5\text{-}3'$, inverted repeat) functions independently of its distance from the promoter but is strongly orientation dependent. Termination takes place immediately upstream of this sequence.

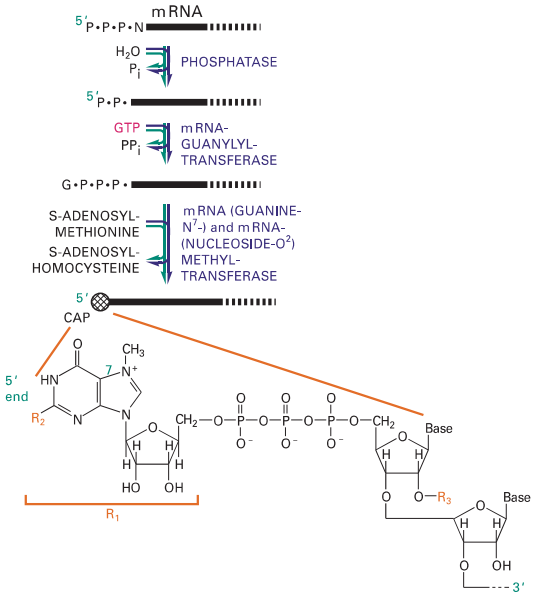


Figure 4.2.1-2. Capping of mRNA

Cap structure	R_1	R_2	R_3	Present in
7mG cap	see Figure 4.2.1-2	NH_2	CH_3	mRNA
m_3G cap	see Figure 4.2.1-2	$N(CH_3)_2$	CH_3	U 1, 2, 4, 5
γ-Monomethyl-P cap	CH_3	–	H	U 6

4.2.1.3 Processing of mRNA

During the course of transcription, mRNA is also processed by capping, splicing and polyadenylation. These are coupled reactions that influence each other. Capping and polyadenylation are important for the efficiency of translation, but not always absolutely required. Only spliced, 'mature' mRNA is transported to the cytoplasm where it gets translated into protein by ribosomes (4.2.3).

Capping (Fig. 4.2.1-2): Nascent mRNA's and snRNAs (4.2.1.4) are cotranscriptionally modified at their 5′ ends by addition of a 'cap', mostly 7-methylguanosine, attached via a 5′-5′ triphosphate bridge. This cap is important for pre-mRNA processing, mRNA export and translation. The removal of the cap is considered to be the first irreversible step in mRNA degradation (4.2.3). In addition to the nuclear capping pathway, a cytoplasmic form of capping enzyme exists generating capped ends from cleaved RNAs.

Splicing (Fig. 4.2.1-4): Splicing is an integral and essential step of gene expression in eukarya (and in a few cases also in archaea and bacteria). In this procedure, the non-translated portions (introns) are removed from the nuclear pre-mRNA and the remaining translated exons are joined with each other.

Split genes have been identified in all types of eukaryotic cells. There can be more than 50 exons in a single gene, some of them as short as 10 nucleotides. The number of introns in pre-mRNA varies from none to dozens (average ca. 8), their length from ca. 70 up to 200,000 nucleotides (average in vertebrates ca. 137). In lower eukarya, introns are short and often flanked by large exons. In higher eukarya, large introns separate usually short exons. This may lead to different mechanisms for splicing ('intron definition' in lower vs. 'exon definition' in higher eukarya). Almost all protein-coding genes in vertebrates have introns (notable exception: histones). The primary transcription unit is typically four to ten times larger than the final mRNA which has an average length of 1 to 2 kb (Fig. 4.2.1-3).

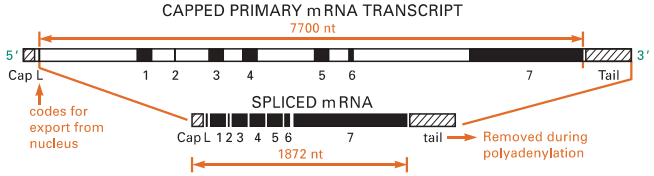


Figure 4.2.1-3. Effect of mRNA Splicing (Ovalbumin Gene)

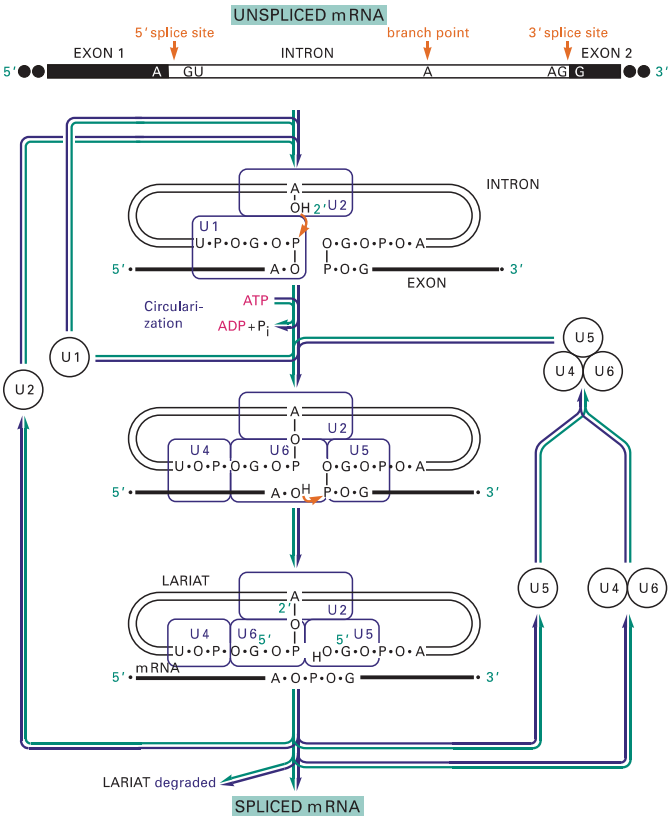


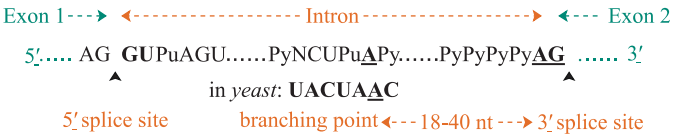
Figure 4.2.1-4. Splicing Mechanism for mRNA

The pre-mRNA processing steps, including splicing and polyadenylation, are mediated by RNA-binding proteins (RBPs) and trans-acting RNAs, jointly forming ribonucleoprotein complexes (RNP). Various classes of RBPs are associated with small non-coding RNAs within RNP complexes that actively participate in fundamental cellular functions, such as DNA replication and translational control. mRNPs (RBPs associated with mRNA) can control splicing by formation of the exon-junction complex (EJC) about 20 nucleotides upstream of exon–exon junctions.

Accuracy of splicing is crucial for cell function. Many human diseases are caused by mutations that interfere with RNA splicing. For example, approximately 25 % of the human globin gene mutations in thalassemia occur in sequences responsible for correct splicing. On the other hand, alternative splicing can effect genetic variability (e.g., formation of different immunoglobulins, 8.1.4 or MHCs, 8.1.7) and produce proteins with different functions (e.g., sexual dimorphism in *Drosophila*).

Splicing mostly occurs cotranscriptionally but can also proceed posttranscriptionally. It takes place within the spliceosome, which is a large RNA-protein complex (60S). The spliceosome is a dynamic structure, which is assembled stepwise on the pre-mRNA at each individual intron. It consists of small nuclear ribonucleoproteins, snRNPs (U1, U2, U4, U5, U6, see below, containing U-rich RNA of 57 … 217 nucleotides and about 80 % protein), several minor snRNPs and spliceosome associated proteins (SAP) which perform important auxiliary functions (not shown in Fig. 4.2.1-4).

Similarities in structure and reaction mechanism to self-splicing group II introns suggest a catalytic role of the spliceosomal RNAs. Highly conserved consensus sequences indicate exactly the exon-intron boundaries. The consensus sequences in higher eukarya show the following structure (invariant nucleotides are printed in bold, the branching nucleotide is underlined):



The first step in splicing is the recognition of these sequences at the 5′ splice site and at the branching point by complementary sequences of U1 and U2 snRNPs. The intron excision proceeds in two consecutive transesterification reactions. A nucleophilic attack by the 2-OH group of the branching point A on the phosphate at the 5′ splice site causes exon-intron separation and ATP-dependent lariat formation by the intron. After U1 has left the complex, U4, U5 and U6 bind. Thereafter, the newly formed 3-OH at the 5′ splice site exerts a nucleophilic attack on the phosphate at the 3′ splice site. This causes elimination of the intron and ligation of the two exons. Then the snRNPs dissociate from the mRNA. The lariat is later linearized by an RNA debranching enzyme and degraded.

Besides this mechanism (group III introns) other splicing mechanisms exist. They proceed with an external guanosine nucleotide instead of the branching point A (group I, rRNA) or with a ligation procedure after endonuclease splitting of the pre-mRNA (group IV, tRNA). Splicing of group II introns (e.g., mitochondria and chloroplast mRNA) resembles the group III mechanism, but without participation of snRNPs. Frequently, the group I or II procedures are performed by RNA activity only (self-splicing, e.g., rRNA in *Tetrahymena*). Exons of different RNA strands can also be combined by splicing mechanisms (trans-splicing, e.g., mRNA in *Trypanosoma*).

Polyadenylation (Fig. 4.2.1-5, Table 4.2.1-3): Posttranscriptional addition of poly(A) to the 3′ end of mRNA is an important biological process conserved from bacteria to humans. In eukarya, polyadenylation is essential for the stability of many mRNAs and also influences the efficiency of translation. In addition, polyadenylation can contribute to posttranscriptional control by targeting RNA for degradation by 3′ to 5′ exoribonucleases. Most histone-mRNAs and all sn-RNAs (except U6), however, are not polyadenylated.

The start of polyadenylation takes place ca. 10 to 30 nucleotides downstream from the polyadenylation signal (consensus sequence: AAUAAA). Also necessary are downstream elements, which consist of poorly defined G/U-rich sequences. Sometimes an U-rich region upstream of the polyadenylation signal acts as an enhancer of polyadenylation. The array of protein factors involved in polyadenylation has been called a ‘poly(A)osome’. The total number of factors involved is still unknown.

Polyadenylation is initiated by cleavage of the mRNA which has been transcribed considerably beyond the adenylation start site. Involved are TFIIS (see above) and CPSF (cleavage and polyadenylation specificity factor), which binds to the polyadenylation signal and stimulates polyA-polymerase (PAP). When the growing poly(A) tail has reached a length of 10 to 12 nucleotides, nuclear poly(A) binding protein II (PAB II) also binds. This strongly stimulates polyadenylation and enables poly(A)-polymerase to synthesize a stretch of $(A)_{200\ldots250}$ (in mammals) in a single processing event. Poly(A) tails in yeast have a length of ca. $(A)_{70\ldots90}$, bacterial (*E. coli*) tails of about $(A)_{15\ldots40}$. Later on, the poly(A) tail gets shortened in the cytoplasm. The cleavage-polyadenylation complex can also connect to the phosphorylated form of CTD allowing additional interaction with elongating Pol II.

The mature mRNA moves to the nuclear surface (probably by diffusion) and is exported through nuclear pores. This is an energy requiring, probably carrier-mediated process as is the export of ribosomal subunits and tRNAs.

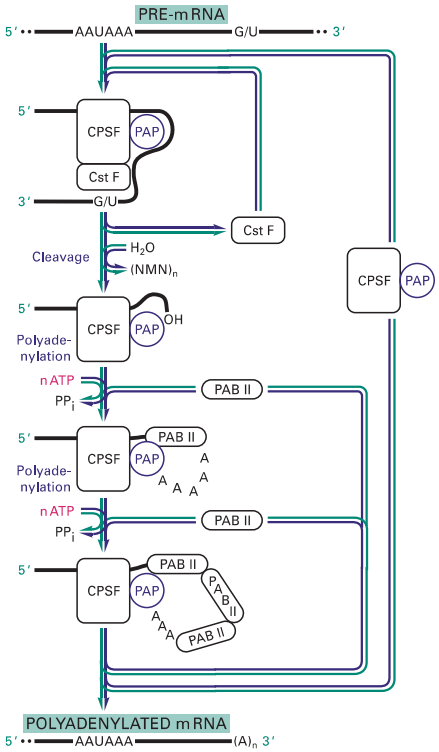


Figure 4.2.1-5. Polyadenylation of mRNA

Table 4.2.1-3. Some Factors Involved in Mammalian Polyadenylation

Name	Subunits	Mol. mass (kDa)	Function
Poly(A) polymerase (PAP)	1	82 (yeast 64)	involved in cleavage and polyadenylylation
cleavage / polyadenylation specificity factor (CPSF)	4	160, 100, 73, 30	binds specifically to the AAUAAA signal, stimulates poly(A) polymerase
Cleavage stimulating factor (CstF)	3	77, 64, 50	binds to U or G/U rich upstream elements, associates with cleavage factors CF1 and CF2
poly(A) binding protein II (PAB-II)	1	49	binds to the growing poly (A) tail, stimulates polyadenylation to $(A)_{200\ldots500}$

4.2.1.4 snRNA Transcription

snRNAs (small nuclear RNAs, e.g., U1…6) are constitutively expressed (10^5 … 10^6 copies per human cell). Some snRNA genes show a TATA element (4.2.2.1) in proximity to the proximal sequence element (PSE). In humans, the occurrence of both TATA box and PSE most often indicates Pol III-specific transcription whereas the absence of a TATA element results in transcription by Pol II. Initiation of transcription starts at a PSE (at position ca. −50) and is enhanced by a distal sequence element (DSE, at position ca. −200). After transcription by Pol II, the sRNAs U1, U2, U4 and U5 are transported to the cytosol, where they acquire a modified 2,2,7-trimethylguanosine cap structure (m_3G cap) at their 5′ termini (similar to mRNA, Fig. 4.2.1-2). Addition of proteins, leading to snRNPs (small nuclear ribonucleoparticles) also takes place there. The hypermethylated cap is essential for transport back to the nucleus. U6 is transcribed by Pol III, obtains a -monomethyl triphosphate cap and remains in the nucleus.

4.2.1.5 rRNA Transcription by RNA Pol I (Fig. 4.2.1-6)

Eukaryotic ribosomes (80S) consist of two subunits (60 and 40S), which contain ribosomal RNAs (rRNAs, large subunit: 28S, 5.8S and 5S rRNA; small subunit: 18S rRNA) and ribosomal proteins (large subunit ca. 45, small subunit ca. 33).

About 50% of RNA synthesis in a cell refers to transcription of rRNA genes. rRNAs are transcribed by RNA polymerase I (only 5S rRNA is transcribed by RNA polymerase III). Many transcription units for rRNAs are arranged in clusters of tandem head-to-tail repeats. They are localized in a structure called the nucleolar organizer region (NOR). There, a special structure (nucleolus) forms where most of the ribosome biogenesis takes place: transcription of rRNA genes, processing of the transcripts to mature rRNAs and assembly with proteins to form both ribosome subunits.

Within the nucleus of higher eukarya several nucleoli are present, in yeast only one. Nucleoli contain, in addition to ribosomal proteins, a large number of non-ribosomal proteins involved in ribosome biogenesis and maintenance of nucleolar structure. During mitosis, nucleoli disassemble in the prophase and reassemble during telophase. DNA topoisomerase I, RNA polymerase I and UBF remain associated with the NOR, but no transcription takes place during that interval.

The promoter for rRNA synthesis consists of at least 2 elements, a GC-rich upstream control element (UCE, 4.2.2.1) and a core region at the transcription start site. Both elements are recognized by the upstream binding factor (UBF, Table 4.2.1-4), which causes sharp bends to the DNA. Transcription factor IB (TFIB), in humans known as selectivity factor I (SLI), consisting of TBP and at least 3 TAF_I, binds to the core region followed by TFIC recruitment, leading to the pre-initiation complex. UBF attaches as a dimer to the UCE and the core region, cooperating with TFIB. UBF binds DNA through high-mobility-group boxes (HMG). By the subsequent joining of Pol I the initiation complex is formed. Elongation proceeds upon binding of DNA topoisomerase I.

Eukaryotic transcription units for ribosomal genes carry short DNA sequences at their 3′ ends which contain a SalI site ('Sal-Box') and cause termination of transcription. TTF-I protein binds to this region as a monomer and causes DNA bending. Terminator sites for RNA polymerase I function only in one orientation.

Table 4.2.1-4. Transcription Factors for RNA Polymerase I

Name	Species	Subunits	Mol. mass (kDa)	Functions / Properties
TFIB	human (SL1)	TBP + 3 TAF_I	38 95, 64, 53	Initiation, promoter selection, recruitment (together with UBF) of RNA polymerase I. Equivalent to TFIID.
TFIC		?	?	Role in formation of pre-initiation complex
UBF (upstream binding factor)	human, highly conserved	dimer of differently spliced gene products	97, 94	Required for the formation of stable initiation complexes by RNA polymerase I and TFIB, specific for large rRNA genes. Causes DNA bending.
TTF1	human	1	130	Termination factor, abundant

4.2.1.6 Processing of rRNA (Fig. 4.2.1-6)

After termination, the 80S RNP precursor contains a 5′ leader sequence and 18S, 5.8S and 28S rRNAs separated by spacer RNAs. It is processed by cleavage of the 5′ leader, splicing and nucleolytic degradation of the spacer RNA. As in mRNA, splicing is directed by small nuclear ribonucleoproteins (snRNPs). This leads to 20S (containing 18S rRNA) and 32S intermediates (containing 5.8 and 28S rRNAs) which are further processed to yield mature 28, 18 and 5.8S rRNAs.

Post-transcriptional modification of the nucleotides results in methylation of about 100 nucleotides per ribosome at the 2′ OH of ribose and isomerization of more than 100 uridine residues per ribosome to pseudo-uridine (see tRNA-modification, 4.2.1.8).

The rRNAs are complexed with ribosomal proteins in a self-organizing mode, forming both the large and small ribosomal subunits which are then separately transported to the cytoplasm.

4.2.1.7 tRNA Transcription by RNA Pol III (Fig. 4.2.1-7)

As in bacteria, there are multiple tRNA genes in eukarya (e.g., about 1300 in a human cell). Mature transfer RNAs (tRNAs) are mostly 75 to 80 nucleotides long. They are transcribed by RNA polymerase III.

tRNA gene promoters consist of 2 separated 10 bp elements (Boxes A and B) located downstream of the transcription start site. TFIIIC binds to box B, then box A orients TFIIIC towards the start site. TFIIIC causes correct positioning of TFIIIB (preinitiation complex), which then recruits Pol III. This DNA/TFIIIB/Pol III initiation complex is very stable and may pass through many rounds of tRNA transcription. Transcription and elongation start immediately after assembly of the initiation complex. The transcription of snRNA U6 (4.2.1.4) proceeds similarly.

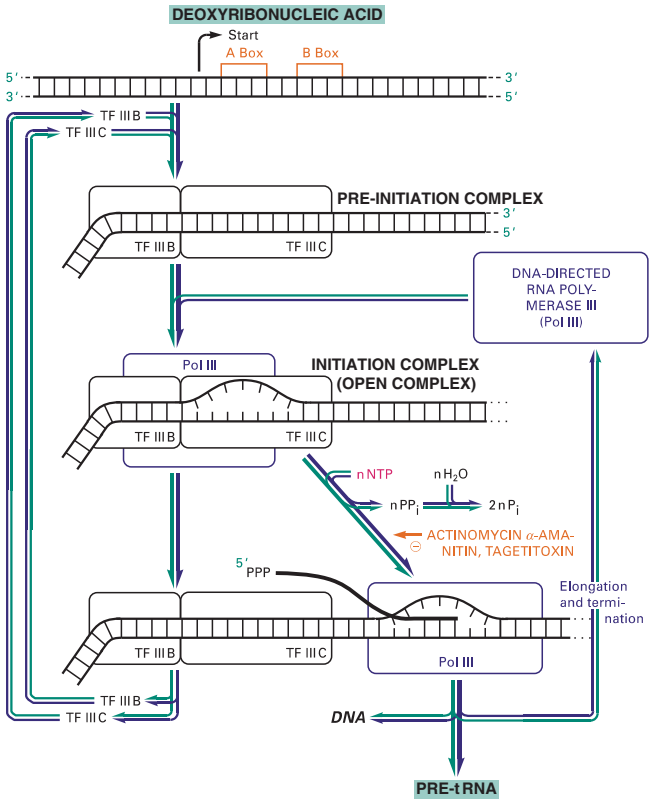


Figure 4.2.1-7. Transcription of tRNA

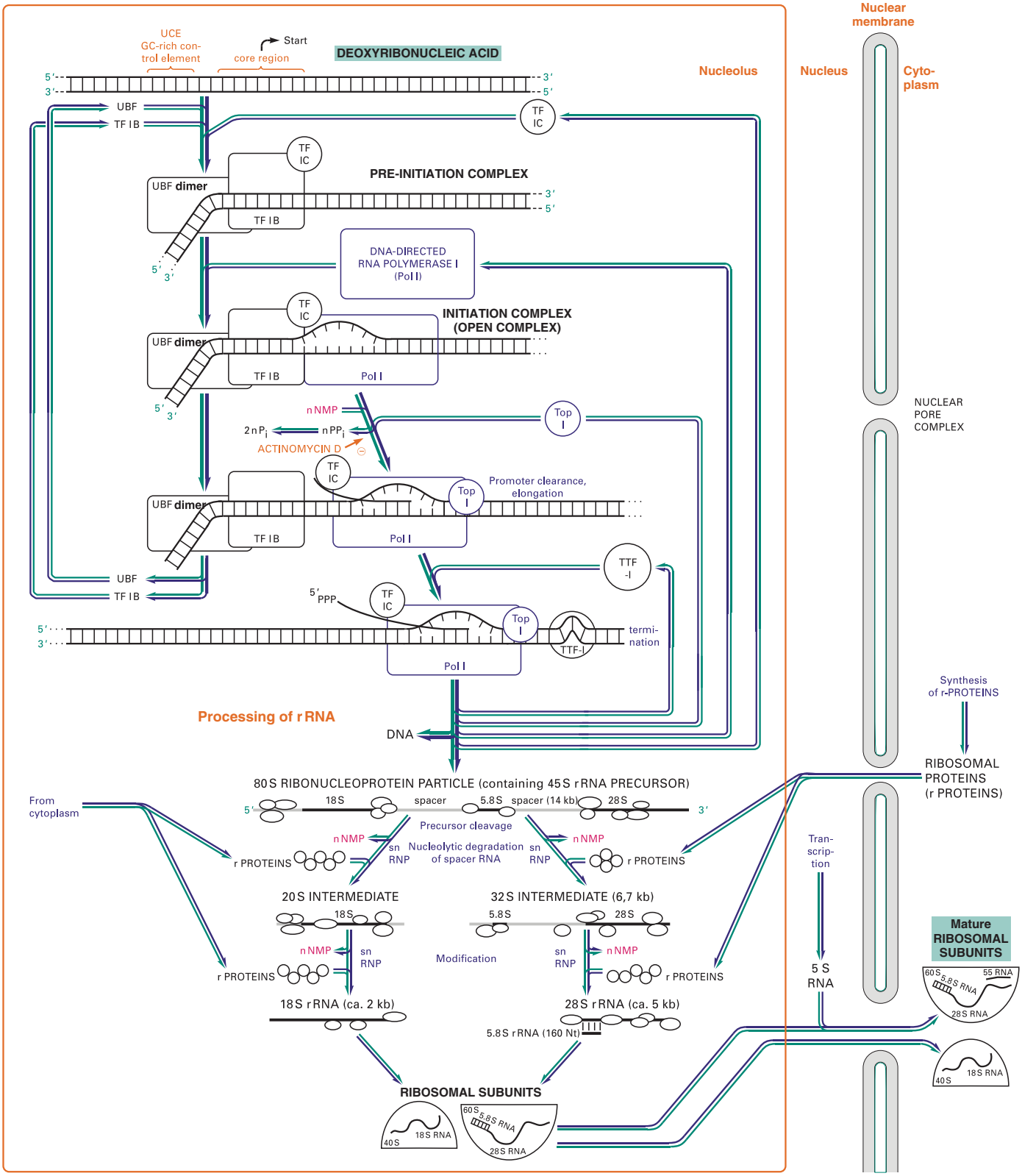
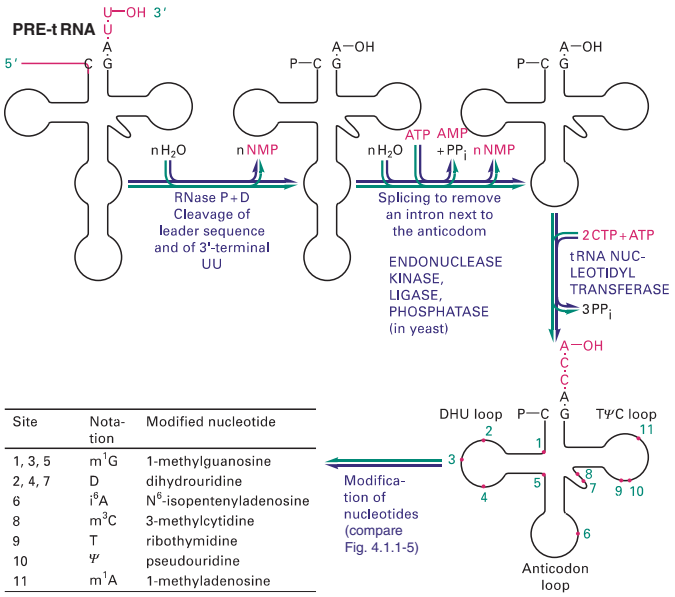


Figure 4.2.1-6. Transcription of rRNA

4.2.1.8 Modification / Processing of tRNAs (Fig. 4.2.1-8)

Pre-tRNAs are processed by cleavage of a 5′ leader sequence and by splicing to remove an intron close to the anticodon loop. Upon maturation the UU sequence at the 3′ end is replaced by CCA. The enzyme catalyzing synthesis and repair of the CCA residue is the ATP(CTP): tRNA nucleotidyl transferase ("CCA adding enzyme").

Eukaryotic tRNAs contain a large variety of modified nucleotides for fine tuning of activity, fidelity and stability, which are formed post-transcriptionally. The 2 hydroxyl groups of about 1% of all riboses are methylated. There are between 7 and 15 unusual bases per molecule, e.g., methylated or dimethylated A, U, C or G residues and pseudo-uridine (ψ).



Site	Nota-tion	Modified nucleotide
1, 3, 5	m^1G	1-methylguanosine
2, 4, 7	D	dihydrouridine
6	i^6A	N^6-isopentenyladenosine
8	m^3C	3-methylcytidine
9	T	ribothymidine
10	Ψ	pseudouridine
11	m^1A	1-methyladenosine

Figure 4.2.1-8. Modification of tRNA (Example: $tRNA^{Tyr}$ from yeast) The sites to be modified are shown in red

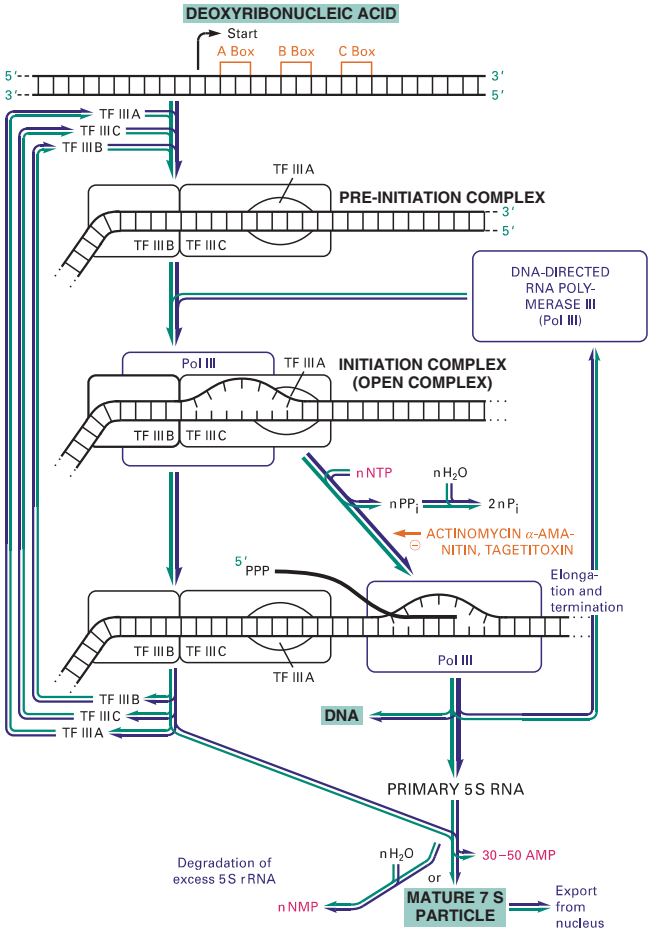


Figure 4.2.1-9. Transcription of 5S rRN

Mature tRNAs are then transported to the cytoplasm. The three-dimensional structure of eukaryotic tRNAs is similar to bacterial tRNAs (Fig. 4.1.3-1).

4.2.1.9 5S rRNA Transcription by RNA Pol III (Fig. 4.2.1-9, Table 4.2.1-5)

5S rRNA is a short (120 nt) molecule, which is highly conserved in sequence and structure (five stem loops). It is transcribed from a group of tandemly arranged genes outside of the nucleolus. The model plant *Arabidopsis thaliana* comprises ca. 1000 copies of 5S RNA genes per haploid genome. The transcription of 5S rRNA genes is similar to tRNA transcription, starting with binding of TFIIIA, followed by TFIIIB and TFIIIC and Pol III recruitment. After elongation and termination, the primary transcript undergoes only minor processing, e.g., removal of 10 … 50 nucleotides from the 3′ end. Surplus 5S RNA is degraded in the nucleus.

Table 4.2.1-5. Transcription Factors for RNA Polymerase III

Name	Species	Subunits	Mol. mass (kDa)	Functions / Properties
TFIIIA	human yeast	1 1	42 10	Assembly factor for positioning of TFIIIB; binds to an internal control region of 5S rRNA genes. Role in export of 5S rRNA.
TFIIIB	human yeast	$TBP+TAF_{III}$ $TBP+2\ TAF_{III}$	– 90, 67	Required for expression of all Pol III transcribed genes. Role in initiation, binds upstream of transcription start site. Bends DNA upon binding. Equivalent to TFIID.
TFIIIC	human yeast	6 6	230…55 145…55	Binds to two intragenic promoter elements (box A and box B) of tRNA genes. Causes positioning of TFIIIB; not always required for transcription.
PBP	human	1	90	Proximal element binding protein

4.2.1.10 Inhibitors of Transcription

Actinomycin D binds tightly and specifically to double-stranded DNA and stops transcription in general. α-Amanitin inhibits the translocation step in the elongation process. Some RNA polymerases, however (e.g., RNA polymerase II from *Aspergillus nidulans*) are resistant to α-amanitin. RNA polymerase I is markedly less sensitive to α-amanitin. This property can be utilized to distinguish between different RNA polymerases. Tagetitoxin (a phytotoxin of *Pseudomonas syringae* pv. *tagetis*) inhibits Pol II and III transcription at concentrations that leave Pol I unaffected.

Literature:

Caponigro, G., Parker, R. Microbiol. Reviews 1996;60:233–249.

Casamassimi, A., Napoli, C. Biochimie 2007;89:1439–1446.

Cramer, P. Adv. Protein Chemistry 2004;67:1–42.

Dieci, G. *et al*. Trends in Genetics 2007;23:614–622.

Erhard, K.F. *et al*. Science 2009;323:1201–1205.

Freiman, R.N. Biochim. Biophys. Acta - Gene Regulatory Mechanisms 2009;1789:161–166.

Glisovic, T. *et al*. FEBS Letters 2008;582:1977–1986.

Jawdekar, G.W., Henry, R.W. Biochim. Biophys. Acta - Gene Regulatory Mechanisms 2008;1779:295–305.

Juven-Gershon, T. *et al*. Current Opinion in Cell Biology 2008;20:253–259.

Kornberg, R.D. Proc. Natl. Academy Sci. USA 2007;104:12955–12961.

Krishnamurthy, S., Hampsey, M. Current Biology 2009;19:R153–R156.

Lange, H. *et al*. Trends in Plant Science 2009;14:497–504.

Latchman, D.S. *Eukaryotic transcription factors*. 5th Ed. Academic Press, 2007.

Moore, M.J., Proudfoot, N.J. Cell 2009;136:688–700.

Nikitina, T.V., Tishchenko, L.I. Molecular Biology 2005;39:161–172.

Proudfoot , N, O'Sullivan, J. Current Biology 2002;12:R855–R857.

Pugh, B.F. Current Opinion in Cell Biology 1996;8:303–311.

Ross , J. Trends in Genetics 1996;12:171–175.

Russell, J, Zomerdijk, J.C.B.M. Trends in Biochemical Sciences 2005;30:87–96.

Schoenberg, D.R., Maquat, L.E. Trends in Biochemical Sciences 2009;34:435–442.

Toor. N., Keating KS, Pyle AM. Current Opinion in Structural Biology 2009;19:260–266.

Vandromme, M. *et al*. Trends in Biochemical Sciences 1996;21:59–64.

4.2.2 Regulation of Eukaryotic Transcription

Eukaryotic transcription (4.2.1) is regulated by an interplay between specific DNA sequence elements (promoters, enhancers, silencers) and a diverse group of special proteins (transcription factors, TF), which recognize these DNA regions.

- Core promoter DNA elements are always required for accurate and efficient initiation of transcription. They are recognized by the basic transcription factors (Table 4.2.1-2); the complex provides the basic transcription rate.
- DNA response elements (enhancers, silencers, Fig. 4.2.2-1 and Table 4.2.2-1) increase or decrease the basic transcription rate of a given promoter. Specific transcription factors can bind to these '*cis*-acting' DNA elements, modulating the rate of transcription by interacting either directly or via co-activators with the general transcription apparatus.

A differentiation is to be made between *trans*-acting protein factors (transcription factors etc.), which can interact with every site of a genome containing their recognition sequence, and *cis*-acting DNA regions (enhancers, silencers), which only interact with their corresponding promoter.

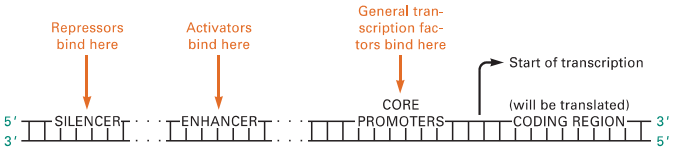


Figure 4.2.2-1. Regulating DNA Sequence Elements (Any Order is Possible)

In a number of cases it has been shown that in transcriptionally active genes the sequences at the start site and up to some distance upstream of it (DNase I hypersensitive sites) are free of nucleosomes (2.6.4), allowing free access of the protein factors to the promoters.

4.2.2.1 Structure of Core Promoter DNA Elements

These DNA regions determine the starting point and the basic initiation frequency of transcription. In eukarya, different core promoters exist for each of the RNA polymerases (4.2.1.1).

RNA polymerase II core promoters (Fig. 4.2.2-2): Polymerase II core promoters comprise two major types: focused and dispersed. Focused promoters (with either a single transcription start site or with multiple narrowly clustered start sites) appear to be more ancient and widespread than dispersed promoters. However in vertebrates, the majority of genes seem to comprise dispersed core promoters. These are typically found in CpG islands with multiple transcription start sites spread over a region of ca. 50 … 100 nucleotides. Pol II core promoter elements, such as TATA box, TFIIB recognition element (BRE), initiator (Inr) motif, downstream core promoter element (DPE) and motif ten element (MTE) are in most cases part of focused core promoters. So far, no universal core promoter elements have been found. The Inr motif (consensus sequence PyPyANTAPyPy) is a recognition site for TFIID and often comprises the start site in focused core promoters. The TATA box (consensus sequence: $T_{0.82}A_{0.97}T_{0.93}A_{0.85}A_{0.63}/T_{0.37}A_{0.88}A_{0.50}/T_{0.37}$, the

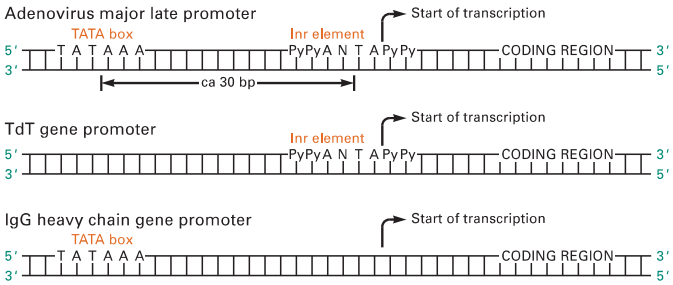


Figure 4.2.2-2. Examples of RNA Polymerase II Core Promoters

indices indicate the frequency of the nucleotide) is usually located at about –30 nucleotides (upstream of the transcription start site).

Human snRNAs, which are transcribed by polymerases II and III, use a TBP– TAF (TBP associated factors) complex (see 4.2.1.2) called small nuclear RNA activating protein complex (SNAPc). This complex binds specifically to the proximal sequence element at the DNA (PSE, position ca. –50), which is a core promoter element common to snRNA genes. In case of transcription by Pol II, the TATA-box is missing.

RNA polymerase I core promoters (Fig. 4.2.2-3): RNA polymerase I promoters (and the majority of RNA polymerase III promoters) lack TATA boxes. The promoter for rRNA synthesis consists of at least 2 critical elements, a GC rich upstream control element (UCE, location at –200 … –100 nucleotides from the start site) and a core region at the transcription start site (location at –50 … +20 nucleotides). Both elements are recognized by the upstream binding factor (UBF, Table 4.2.1-4). UBF also supports the recruitment of TATA-binding protein (TBP), a part of several transcription factor complexes including TFIB.

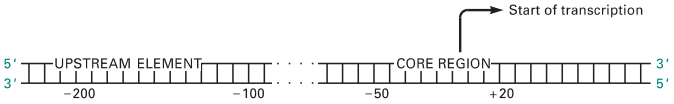


Figure 4.2.2-3. Example of RNA Polymerase I Core Promoters

RNA polymerase III core promoters (Fig. 4.2.2-4): Three main types of polymerase III promoters exist. They contain downstream elements within the genes [A and B boxes (which are each 10 bp long and bind transcription factor TF IIIC) and also I and C (internal control region) boxes]. Many polymerase III promoters (e.g., t-RNA and 5S rRNA promoters) have no TATA boxes.

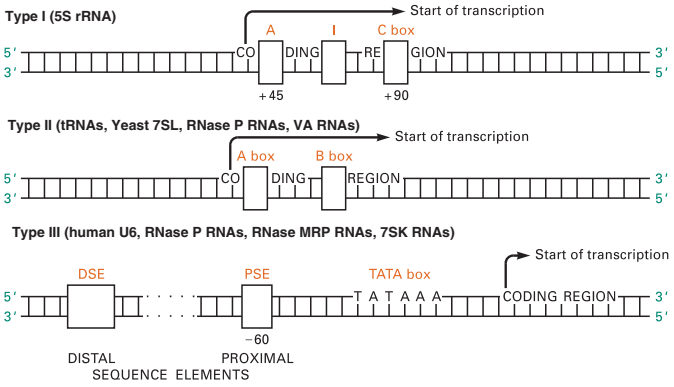


Figure 4.2.2-4. Examples of RNA Polymerase III Core Promoters

4.2.2.2 Structure of Specific Transcription Factors

The basal activity of promoters can be modulated by additional, specific transcription factors (Table 4.2.2-1). They contain these functional domains:

- a DNA binding domain that recognizes specific recognition or response sequences within its target promoters. The structures (zinc fingers, helix-turn-helix etc.) are frequently highly conserved between species. These domains exist also in many basic transcription factors (4.2.1.2, 4.2.1.6, 4.2.1.8).
- an activation domain of variable composition (acidic, glutamine-rich, proline-rich, serine/threonine-rich etc.) that is required for transcription stimulation. This activation domain makes either direct protein-protein contact with components of the basic transcription machinery or acts via co-activators.
- in some cases (e.g., nuclear receptors) a hormone binding domain.

A number of specific transcription factors react as homo- or heterodimers. Their components are connected by dimerization domains.

In plants, specific families of transcription factors usually have more members (e.g., Myb family), suggesting a higher frequency of adaptive responses to the environment as compared to other organisms.

Table 4.2.2-1. Examples of Domains in Transcription Factors

Functional Domain	Structure, Action	Occurs in Transcription Factors (Examples)
DNA binding Domains:		
Zinc finger	1 Zn is complexed by Cys_2-His_2 or by Cys_2-Cys_2 2 Zn are complexed by Cys_6	TFIIIA, Sp1 steroid receptors Gal 4 (yeast)
Helix-turn-helix	Pair of tilted α-helices, binds via hydrogen bonds, salt bridges, etc.	homeodomain proteins
Dimerization Domains:		
Leucine zipper	Dimerizes proteins with helices by hydrophobic attraction of several Leu on one side of the helices.	Fos, Jun, Myc, CREB = ATF-2, CREM
Basic helix-loop-helix	A basic region followed by 2 helices, connected by a loop.	Max, Myc, MyoD

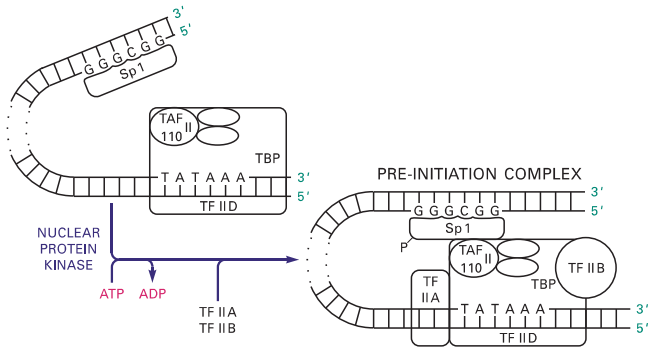


Figure 4.2.2-5. Contact of Upstream Elements with Core Promoters (Schematically)

4.2.2.3 Modulation of the Transcription Rate

The modulating specific transcription factors bind to DNA response elements (enhancers, repressors, silencers). These response elements can be located within the genes or upstream or downstream up to several 1000 bp away in either orientation. This often requires the bending of the DNA to enable contact between the specific and the basic transcription factors or coactivators (Fig. 4.2.2-5).

The specific transcription factors have to be activated by binding of ligands or by phosphorylation in order to bind to their target DNA sites. Some examples for this mechanism are shown in Figure 4.2.2-6. For steroid receptors, see 7.7.

The (more frequent) enhancement or the repression of gene expression proceeds by influencing the initiation (in most cases), promoter clearance and/or elongation steps (4.2.1). As an example, enhancers of RNA polymerase II often act on the TAF_{II} coactivator subunits of transcription factor TFIID (4.2.1.2) to stabilize preinitiation complexes or to assist in recruiting Pol II into preinitiation complexes. Regulation of gene expression can also be exerted by controlling the rate of the transcription factor synthesis or transcription factor access to the nucleus.

These mechanisms effect transcription of genes

- constitutively in all cell types ('housekeeping genes')
- or only in certain organs or cell types
- or on demand under certain conditions.

In eukarya, the expression ratio of 'turned on' and 'turned off' genes can be up to 10^9/1 (In bacteria, the ratio is about 10^3/1).

Cis-regulatory elements often contain multiple binding sites for transcription factors which assemble to multicomponent enhancer complexes called enhanceosomes. One of the best-studied enhanceosomes is that of the human interferon-β (IFN-β) gene. The IFN-β enhancer region contains several positive regulatory domains (PRDs), to which respective regulatory factors bind and coassemble (Fig. 4.2.2-7).

In addition to regulation at the transcriptional level, gene expression can also be modulated by the degradation rate of mRNA (4.2.5), by regulated splicing (e.g., 4.2.1.3), by the translocation rate to the cytosol or by modification of the translation rate (4.2.3). An important role within the cellular control of gene expression is seen in the function of RNA mediated silencing and activation pathways.

Table 4.2.2-2. Examples of Eukaryotic Enhancers and Repressor Elements

Response Element	Consensus Sequence (Palindromes in bold letters)[1]	Binds Specific Transcription Factors	Modulates Expression of Genes Coding for (Examples)
GC box	GGGCGG	stimulatory or specificity protein 1 (Sp1)	enhancer for housekeeping genes transcribed by Pol II and Pol III, regulated by Rb protein (Table 4.3-1).
cAMP response element (CRE)	**TGACGTCA** (similar: in TRE element)	CRE binding protein (CREB = ATF-2). Similar: CREM. Requires association with CBP.	somatostatin, c-Fos, PEPCK, VIP, PTH, tyrosine hydroxylase, fibronectin. Absent in some other genes responsive to cAMP, e.g., growth hormone, prolactin.
Serum response element (SRE)	**GTCCATATTAGGAC**	Serum response factor (SRF, assoc.with TCF, Elk-1, Sap-1)	c-Fos
Phorbol ester (TPA[2], AP-1) response element (TRE)	**TGACTCA** (also part of VDRE)	AP-1 (Jun•Fos)	collagenase, stromelysin, c-Myc, c-Sis, Pro-1
Vitamin D receptor response element (VDRE)	**aGGTGACTCACCt**	vitamin D receptor (7.7)	osteocalcin, osteonectin, calbindin, calreticulin, alkaline phosphatase
Glucocorticoid receptor response element (GRE)	**TGTTCT** (palindromic half site)	glucocorticoid receptor (7.7)	growth hormone, bone sialoprotein, chicken lysozyme
Estrogen receptor response element (ERE)	**aGGTCANNNTGACCt**	estrogen receptor (7.7)	estrogen responsive genes, e.g., osteocalcin, chicken ovalbumin
CCAAT box	gGCCAATct	CCAAT-binding transcription factor (CP1), CTF	α-globin and many other products
CACCC box / Rb control element (RBE)	gcCACCC	Rb protein (tumor suppressor, Table 4.3-1) and others	globin
Myc control site	**CACGTG**	Myc/Max, Mad/Max, Mxi/Max, Max/Max dimers	regulators of cell growth and differentiation
B sequence motif	**GGGANNPyTCC (GGGAAATTCC)**	nuclear factor kB (after release from IkB complex)	immunoglobulins, IL-2, IL-2 receptor, IFN-, GM-CSF, TNF- (in B cells, certain T cells and monocytes)
Octamer DNA bind. motif	ATGCAAAT	Oct-1, 2, 3, 4 (OTF-1, 2, 3, 4)	snRNA (by Oct-1), -globin (by Oct-2, in lymphoid cells)
E-Box	ccgaa**CACATGTG**cccgc	basic helix-loop-helix transcription factors	scleraxis (bone development, e.g.,in rats)

[1] Less conserved bases are printed with lower case letters.
[2] TPA (12-O-tetradecanoyl-phorbol-13-acetate) is a potent tumor promoter, potentiating the effect of a subcarcinogenic dose of an initiating carcinogen.

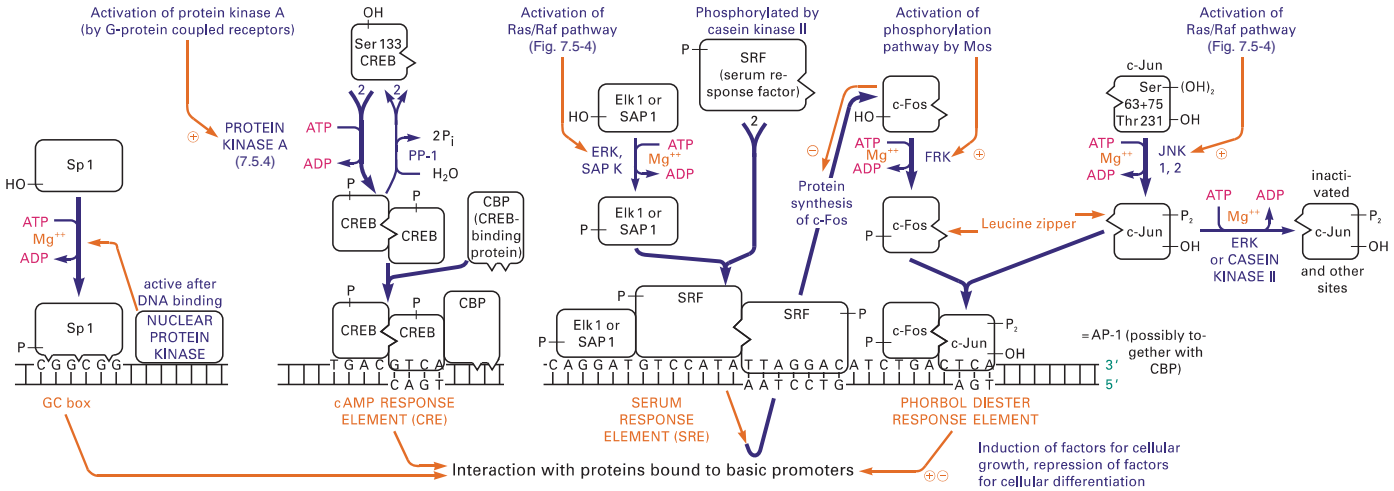


Figure 4.2.2-6. Activation and Binding of Regulatory Proteins to Modulating DNA Elements (Examples, Schematically)

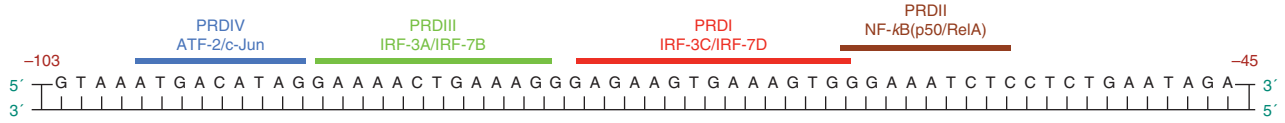


Figure 4.2.2-7. Positive Regulatory Domains (PRDs) and Respective Transcription Factors of the Human Interferon (IFN-β) Enhanceosome

These pathways may involve different small RNAs, e.g., short interfering RNAs (siRNAs), microRNAs (miRNAs), PIWI-interacting RNAs (piRNAs), and small activating RNAs (saRNA), respectively.

Literature:

Brodersen, P., Voinnet, O. Trends in Genetics 2006;22:268–280.

Cowell, I.G. Trends in Biochemical Sciences 1994;19:38–42.

Hawkin, P.G., Morris, K.V. Cell Cycle 2008;7:602–607.

Juven-Gershon, T. *et al*. Current Opinion in Cell Biology 2008;20:253–259.

Latchman, D.S. *Eukaryotic transcription factors*, 5th Ed. Academic Press, 2007.

Li , L-Clue. *RNA and the Regulation of Gene Expression: A Hidden Layer of Complexity*. Essex: Caister Academic Press. 2008; 189–199.

Moazed, D. Nature 2009;457:413–420.

Pah, H.L., Baeuerle, P.A. Current Opinion in Cell Biology 1996;8:340–347.

Panne, D. Current Opinion in Structural Biology 2008;18:236–242.

Russell, J., Zomerdijk, J.C.B.M. Trends in Biochemical Sciences 2005;30:87–96.

Shiu, S.H., Shih, M.C., Li, W.H. Plant Physiology 2005;139:18–26.

Sikorski, T.W., Buratowski, S. Current Opinion in Cell Biology 2009;21:344–351.

Treisman, R. Trends in Biochemical Sciences 1992;17:423–426.

Vandromme, M. *et al*. Trends in Biochemical Sciences 1996;21:59–64.

4.2.3 Eukaryotic Translation

Translation of the genetic code (i.e. the biosynthesis of proteins from amino acids as encoded in the RNA message) proceeds on ribosomes and uses charged tRNAs as activated forms of the amino acids. In eukarya, it takes place in the cytosol or at the membranes of the endoplasmic reticulum (4.4.1) or of the nucleus. The steps of eukaryotic protein synthesis resemble the bacterial ones (4.1), but are more complicated. On the other hand, the protein synthesis machinery in mitochondria and chloroplasts is simpler than in bacteria.

4.2.3.1 Components of the Translation System

Transfer RNAs: The general structure of eukaryotic tRNAs is analogous to bacterial tRNAs (Fig. 4.1.3-1). The tRNAs are charged at their 3′ ends (sequence -CCA) with their cognate amino acids by specific aminoacyl-tRNA ligases (also named 'synthetases'). This reaction requires ATP and takes place in two steps: Activation of the amino acid followed by charging of the tRNA (see 4.1.3.1 and Fig. 4.1.3-2). Contrary to bacteria, in many higher eukarya a number of ligases associate to a multienzyme particle or even fuse to a single polypeptide (GluPro-ligase).

Messenger RNAs: While bacterial mRNA is usually translated without modification, eukaryotic mRNAs are extensively processed before leaving the nucleus, resulting in a complicated structure (Fig. 4.2.3-1). They contain

- a methylated 5′ cap (for details, see Fig. 4.2.1-2)
- a 5′ untranslated leader sequence (5-UTR, usually < 100 bases), involved in regulation of transcription initiation and mRNA degradation
- the coding sequence, to be translated into proteins
- a 3′ untranslated sequence (3-UTR, length up to 1000 bases), involved in mRNA localization, initiation (or repression) of translation and mRNA degradation (by controlling the decapping rate, 4.2.5)
- the poly (A) tail, essential for the stability of mRNA (see Fig. 4.2.1-5)

Ribosomes: Eukaryotic ribosomes (80S or 4300 kDa) consist of two subunits of 40S (1400 kDa) and 60S (2900 kDa). They contain RNA and many different proteins (Table 4.2.3-2). Although their structure

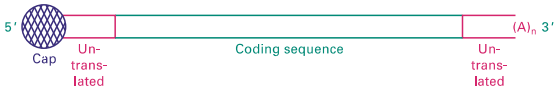


Figure 4.2.3-1. Structure of Eukaryotic mRNA

is more complex than that of bacterial ribosomes, the basic features are similar. This includes the general shape (Fig. 4.1.3-3) and the presence of the aminoacyl- (A), peptidyl- (P) and exit (E) interaction sites with tRNA and mRNA (4.1.3.1). The eukaryotic ribosome comprises additional ribosomal proteins and additional rRNA expansion segments, mainly located around the periphery of the subunits. The translation processes proceeds in four major phases: initiation, elongation, termination, and ribosome recycling.

Table 4.2.3-1. Composition of Eukaryotic Ribosomes (Rat liver)

	Small subunit (40S)	Large subunit (60S)
RNA	18S, 1874 nt, 700 kDa	28S, 4718nt + 5.8S, 160 nt + 5S, 120 nt; total 1820 kDa
Protein	33 polypeptides, 700 kDa	49 polypeptides, 1000 kDa

Translation factors (Table 4.2.3-2): In eukaryotic translation, a large number of protein factors are involved. E.g., there are at least 10 initiation factors (eIF) in eukarya as compared to 3 in bacteria. While bacteria regulate protein synthesis almost completely at the transcriptional level, eukarya also exert control during translation involving such factors.

4.2.3.2 Polypeptide Synthesis (Fig. 4.2.3-2)

Initiation: Eukaryotic protein synthesis starts with an initiator $tRNA_i^{Met}$ which is always charged with methionine (not formylmethionine as in bacteria) and is structurally different from the regular $tRNA^{Met}$. It is, however, charged by the normal ligase.

After completing a round of polypeptide synthesis, the ribosomes dissociate into the subunits. Reassociation is prevented by binding of initiation factors eIF-1A and eIF-3 to the 40S and eIF-6 to the 60S subunit. Then a 43S initiation complex forms from the 40S subunit, eIF-1A, eIF-2•GTP, eIF3 and Met-$tRNA^{Met}_i$. This complex localizes the capped 5′ end of the mRNA and migrates along the mRNA to the first AUG codon (scanning model). Met-$tRNA^{Met}_i$ and the start codon of mRNA are placed at the P-half site of the 40S ribosomal subunit.

Initiation is facilitated if AUG occurs in the context GCCGCC(A/G)CC AUGG. In less than 10% of all cases, initiation at non-AUG codons (CUG, ACG, GUG) or at downstream AUG codons takes place, possibly caused by secondary structural features of the mRNA. Eukaryotic mRNAs are mostly monocistronic (coding for only one gene).

eIF5 joins the 43S initiation complex and triggers hydrolysis of eIF-2 bound GTP, which causes eIF-2•GDP and the other initiation factors to leave. Then the 40S ribosomal subunit combines with the 60S subunit to form the 80S initiation complex. eIF2-GTP is regenerated from eIF2 by the action of eIF-2B (GEF = guanine nucleotide exchange factor). The eIF2 cycle is a control point in protein synthesis: phosphorylation of the subunit prevents regeneration of the GTP form.

Elongation: The steps of peptide bond formation and movement to the next codon proceed analogous to the mechanism in bacteria (4.1.3.3). The anticodon sequences of all further charged tRNAs recognize (by base pairing) complementary codon sequences in the mRNA (2.7). This takes place at the aminoacyl (A) site of the ribosome. tRNA also interacts with the ribosome itself.

Due to the flexibility in base pairing at the third position of a codon (wobble hypothesis, see Fig. 2.7-2) only ca. 35 tRNAs are needed to translate all 61 codons specifying amino acids. The wobble mechanism is mediated by post-transcriptional modifications at the tRNA wobble position 34 (particularly uridine 34).

The rRNA of the large ribosomal subunit exerts peptidyltransferase activity, which catalyzes the transfer of Met (or of the peptidyl chain in the following cycles) from the tRNA at the P-site to the N end of the aminoacyl residue attached to the tRNA at the A-site (see Fig. 4.2.3-2). Then the amino acid acceptor ends of both tRNAs move at the 60S ribosomal subunit from A- to P-half sites and from P- to E-half sites, respectively, while their anticodons and the mRNA remain at their sites at the 40S subunit. Thus, the tRNAs are located at hybrid sites: A/P-site and P/E-site. In a second step, both tRNAs and the mRNA move one codon site relative to the 40S ribosome. The peptidyl-tRNA now occupies the P site in both ribosomal subunits, while the deacylated tRNA leaves the ribosome from the E-site. A new charged

Table 4.2.3-2. Eukaryotic Translation Factors

Factor	Subunits	Mol. Mass (kDa)	Prokaryotic homologue	Function
eIF-1A	1	17		Binds to the 40 S ribosomal subunit, prevents ribosome reassociation in the absence of an initiation complex.
eIF-2	3 (α,β,γ)	52, 38, 35		Binds initiator tRNA (Met-$tRNA_i$) in GTP-dependent manner. Very important control reaction of translation. Its activity is shut off by phosphorylation of Ser 51 in the subunit, which is catalyzed, e.g., by PKR (double-stranded RNA dependent kinase) or HRC (heme controlled repressor).
eIF-2B	5	82, 67, 58, 39, 26		Mammalian guanyl nucleotide exchange factor (= GEF), regenerates GTP at eIF-2.
eIF-3	8 … 10	500 … 750	IF-3	Cooperates with eIF-1A in prevention of ribosome reassociation.
eIF-4A	1 (2 isoforms)	45		Subunit of eIF-4F. Assembles with eIF-4E and eIF-4G at the mRNA cap, melts secondary intermolecular structures which otherwise prevent binding of the 43S pre-initiation complex. Single-stranded RNA dependent ATPase and helicase.
eIF-4B	2 (homodimer)	69		Stimulates the RNA-dependent ATPase and helicase activity of eIF-4A.
eIF-4C	1	17		Binds after ribosome separation to the 40S subunits, facilitates consecutive steps.
eIF-4D	1	16		Appears to function late in initiation.
eIF-4E	1	24		Subunit of eIF-4F. Cap binding protein (CBP). Cooperates with eIF-4A for mRNA melting. One of the least abundant eIF's.
eIF-4F	3	100 … 150 (eIF-G),45 (eIF-4A), 24 (eIF-4E)		The functions of subunits 4A, 4E and 4G are described in the respective lines. Polio-, rhino- and hoof and mouth-disease viruses split the eIF-G subunit, causing inactivation of translation in the host and stimulation of viral mRNA translation.
eIF-4G	1	100 … 150		Other names: p420 and eIF-4. Cooperates with eIF-4A for mRNA melting.
eIF-5	1	49		Possibly effects release of eIF factors by triggering hydrolysis of eIF-2 bound GTP after the ribosome joins the AUG triplet.
eIF-6	1	25		Binds to the 60S ribosomal subunit, apparently prevents ribosome reassociation in the absence of an initiation complex (similar to eIF-1A, −3).
eEF-1	1	50 … 60	EF-Tu	Catalyzes GTP-dependent binding of aa-tRNA to ribosomes. Very abundant protein.
eEF-1	2 (β,γ)	48, 35	EF-Ts	Facilitates nucleotide exchange at eEF-1.
eEF-2	1	95	EF-G	Responsible for GTP dependent translocation step during elongation.
eEF-3	1	125		Ribosome-dependent nucleotidase (ATP and GTP). Found so far only in fungi.
eRF-1	2	49 … 55	RF-1, RF2	Release factors for peptide chain with a structure analogous to tRNAs.
eRF-3	1 (homo-dimer)	80	RF-3	eRF-3 binds GTP. The homology to bacterial release factors is only with respect to function.

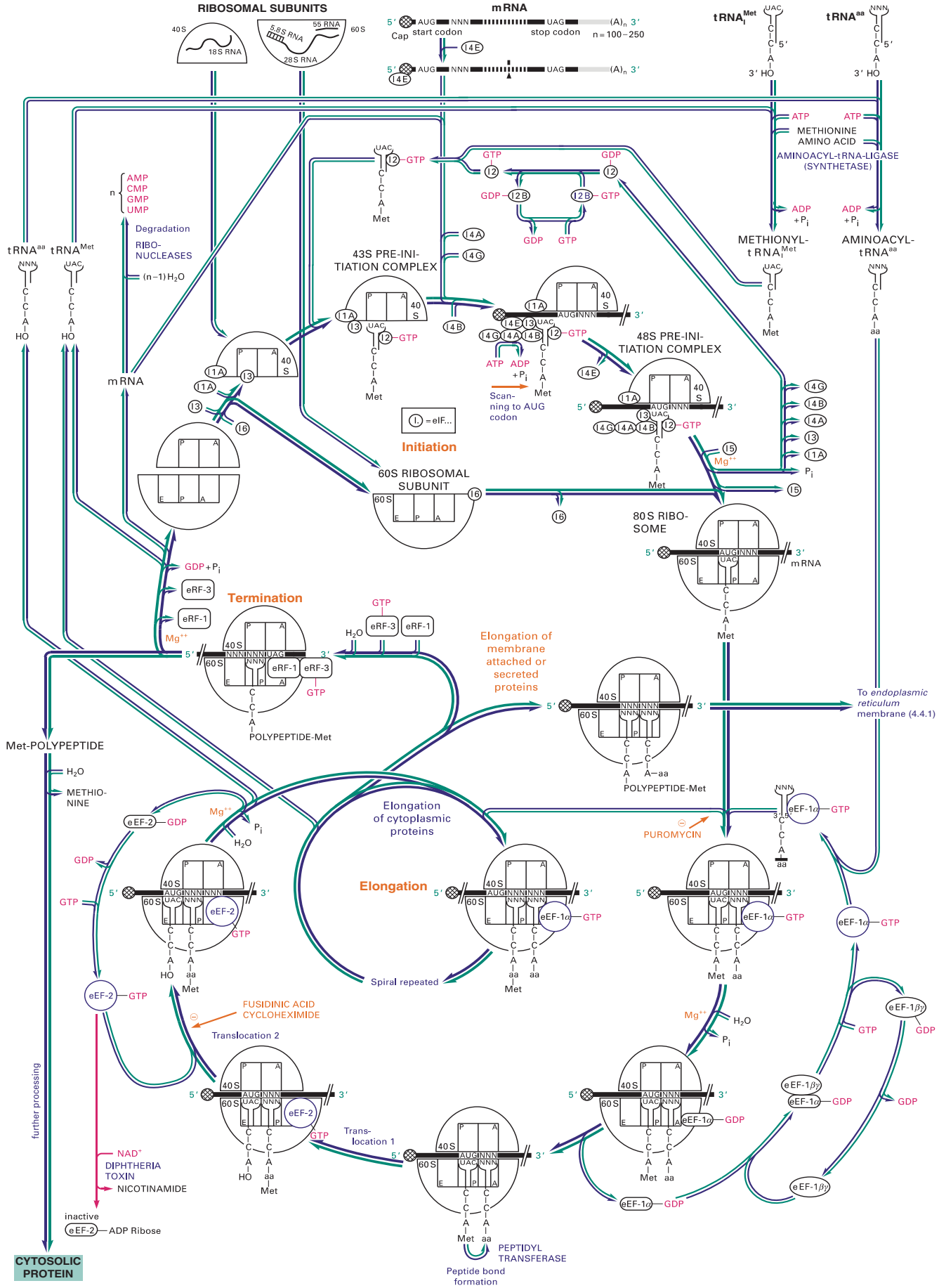


Figure 4.2.3-2. Eukaryotic Translation

tRNA can now bind to the free A-site, thus starting another elongation cycle.

The movements of the ribosome involve configuration changes. The energy comes from GTP hydrolysis which occurs after binding of eEF-2•GTP to the ribosome. After the translocation step, eEF-2•GDP is released. It reacts with free GTP to regenerate the complex for another reaction cycle. Cycloheximide stops the translocation step by inhibition of GTP binding, diphtheria toxin inactivates eEF-2 by ADP-ribosylation (compare 7.4.1).

Many glycosylated proteins (e.g., membrane-bound, secreted or lysosomal ones) are synthesized in the endoplasmic reticulum (4.4.1). The N-terminal parts of these proteins, which are still formed by free ribosomes in the cytoplasm, have the function of a signal sequence. This sequence is recognized by a signal recognition particle (SRP, 4.4.1.1), which targets the whole translation complex to the endoplasmic reticulum. The growing peptide chain is threaded into the lumen of this compartment, where post-translational modifications take place (4.4.1.1, also see below).

Termination: Like bacteria, eukarya use the stop codons UAA, UAG and UGA to signal the end of the coding region of a gene (4.1.3.3). There are two eukaryotic release factors, eRF-1 (codon specific) and eRF-3 (codon unspecific, binds GTP). Apparently the eRFs recognize the stop codon by mimicking a tRNA anticodon structure. This prevents further transfer of the polypeptide chain, which is then hydrolytically removed from the tRNA. Puromycin causes premature chain termination by acting as an analogue of aminoacyl-tRNA.

A high amount of free energy is spent on achieving the peptide bond formation and the reaction specificity (for details, see 4.1.3.3).

Ribosome recycling: The ultimate step of translation is different in prokarya and eukarya. In eukaryotic cells there is no ortholog of the prokaryotic ribosome recycling factor (RRF). Recent findings suggest that the separation of 80S ribosomes into 40S and 60S subunits is mediated by eIF3 and the release of the tRNA is supported by eIF1. The exact function of all involved factors and the respective timing are not yet known.

4.2.4 Translational Regulation

In contrast to bacteria, regulation of translation by repressor proteins is a common feature in eukarya. E.g., by acting as an ion sensor, iron response element binding protein (IRE-BP) regulates the production of the iron sequestering compound ferritin. When iron concentration is low, IRE-BP binds to ferritin mRNA; this inhibits ferritin translation. Iron excess causes its dissociation from the mRNA and resumption of translation. Similar mechanisms exist for stabilization of mRNAs.

Phosphorylation of initiation factors is a common way to regulate translation, e.g., for coordination of hemoglobin synthesis. In the absence of heme, reticulocytes generate the heme controlled repressor from a precursor (Fig. 4.2.4-1). This repressor phosphorylates the subunit of eIF-2, which then binds tightly to eIF-2B, thus stopping the regeneration cycle of eIF-2•GTP (Fig. 4.2.3-2) and thus the initiation of globin transcription. Heme inhibits the repressor formation and thus allows globin synthesis. Already phosphorylated eIF-2B can be reactivated by a phosphatase.

Upstream open reading frames (uORFs) in mRNA can impair translation initiation by ribosome stalling. Translation can re-start at sites downstream from the first AUG. This reinitiation process is a control mechanism often encountered in translation of important cellular genes like oncogenes, growth factor genes etc. Another method of regulation is the competition of different mRNAs for mRNA specific initiation factors. Also, temporary masking of mRNA by association with proteins occurs, especially in unfertilized eggs of several species.

In plants, translational inhibition can be mediated by miRNAs and siRNAs.

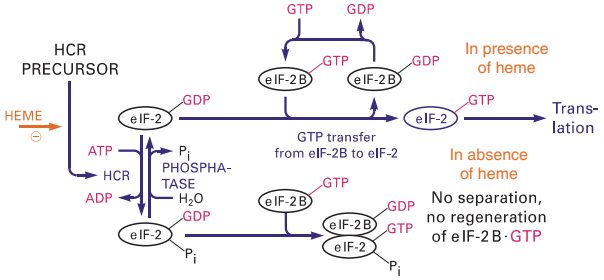


Figure 4.2.4-1. Control of Translation in Reticulocytes by Heme

4.2.5 mRNA Degradation

mRNA stability is an important control factor in eukaryotic protein synthesis. The half-life of eukaryotic mRNA can vary from a few minutes to months, while it is generally very short in bacteria. Structures which influence the degradation rate can occur in all parts of the molecule (cap, 5′ and 3′-UTR, coding region).

For degradation of a number of yeast mRNAs, the 3′ poly(A) tail, often covered by poly(A) binding protein 1 (Pab1), is shortened in the cytoplasm by action of deadenylase Ccr4. The association of the deadenylated RNA with the Lsm 1-7 complex promotes decapping by the Dcp1-Dcp2 protein complex, consisting of catalytic subunit Dcp2 and its coactivator Dcp1. The release of m^7GDP and a 5′ monophosphorylated RNA results in further degradation of the RNA by 5′-to-3′ exonucleases like Xrn1 (Fig. 4.2.5-1). The Dcp1-Dcp2 protein complex assembles with associated mRNA molecules and proteins of the decapping machinery into cytoplasmatic granules called processing bodies (PB). Whereas deadenylation does not always result in decapping (as mRNA can appear in a translationally repressed deadenylated form), decapping is also involved in the AU-rich element (ARE)-mediated mRNA degradation pathway, and in the nonsense-mediated mRNA decay (NMD). NMD is a translation-dependent degradation pathway specifically targeting mRNAs with premature termination codons as well as subsets of regular mRNAs.

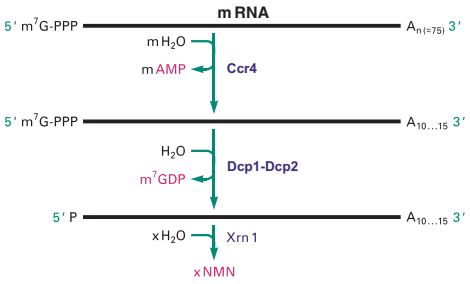


Figure 4.2.5-1. Deadenylation-mediated mRNA Degradation in Yeast

Degradation from the 3′-ends of RNAs often occurs by a highly regulated exonuclease complex called the RNA exosome. Eukaryotic nuclear and cytoplasmic exosomes share a core of common subunits. Rrp41, Rrp42, Rrp43, Rrp45, Rrp46, and Mtr3 are associated into a hexameric ring, associated to a cap trimer of additional proteins (Fig. 4.2.4-3).

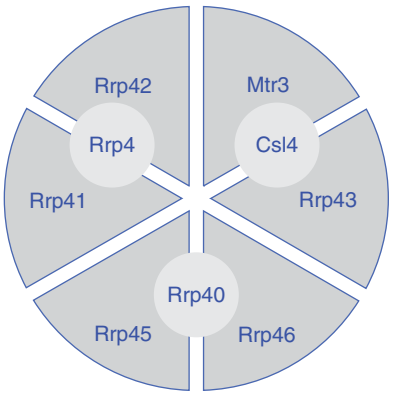


Figure 4.2.5-2. Schematic Structure of the Eukaryotic Exosome Complex

In addition to the above mechanisms, endonucleolytic cleavage can also cause exonucleolytic degradation, beginning at the unprotected end. Histone mRNAs are not polyadenylated, but possess a stem-loop structure at their 3′ ends. The initial decay event removes nucleotides at the 3′ end resulting in disruption of the stem-loop.

Literature:

Agris, P.F., Vendeix, F.A.P., Graham, W.D. J. Mol. Biol. 2007;366:1–13.

Alkalaeva, E.Z. *et al*. Cell 2006;125:1125–1136.

Belostotsky, D. Current Opinion in Cell Biology 2009;21:352–358.

Brodersen, P. *et al*. Science 2008;320:1185–1190.

Calvo, S.E., Pagliarini, D.J., Mootha, V.K. Proc. Nat. Acad. Sciences (USA) 2009;106:7507–7512.

Carter, C.W. Annual Review of Biochemistry 1993;62:715–748.

Fillman, C., Lykke-Andersen, J. Current Opinion in Cell Biology 2005;17:326–331.

Franks, T.M., Lykke-Andersen, J. Molecular Cell 2008;32:605–615.

Garza, A.S., Ahmad, N., Kumar, R. Life Sciences 2009;84:189–193.

Glisovic, T. *et al*. FEBS Letters 2008;582:1977–1986.

Houseley, J., Tollervey, D. Cell 2009;136:763–776.

Lykke-Andersen, S., Brodersen, D.E., Jensen, T.H. Journal of Cell Science 2009;122:1487–1494.

Moore, M.J., Proudfoot, N.J. Cell 2009;136:688–700.

Pain, V.M. Eur. J. Biochem. 1996;236:747–771.

Pisarev, A.V., Hellen, C.U.T., Pestova, T.V. Cell 2007;131:286–299.

Rebbapragada, I., Lykke-Andersen, J. Current Opinion in Cell Biology 2009;212:394–402.

Rodnina, M.V., Wintermeyer, W. Current Opinion in Cell Biology 2009;21:435–443.

Ross, J. Trends in Genetics 1996;12:171–175.

Schmid, M., Jensen, T.H. Trends in Biochemical Sciences 2008;33:501–510.

Shiu, S.H., Shih, M.C., Li, W.H. Plant Physiology 2005;139:18–26.

Tompa, P. FEBS Letters 2005;579:3346–3354.

4.3 Cell Cycle in Eukarya

Stefan Ries

The cell cycle describes the ordered process required for a cell to duplicate itself, and comprises the following major activities: growth of cell size, duplication of genetic information (chromosomes), segregation of chromosomes, and finally cellular division. While in prokarya the periods of DNA duplication and chromosome segregation overlap, in eukarya the cell cycle consists of four defined phases:

- **G_1 phase (gap),** where the cells prepare for DNA replication
- **S phase (synthetic),** in which a single copy of the genetic material is generated
- **G_2 phase (gap),** during which cells prepare themselves for the M phase
- **M phase (mitotic),** in which all the cellular components are distributed between two identical daughter cells

In addition, when cells cease proliferation, they can exit the cell cycle, and enter a resting **state called G_0**. Cells in G_0 account for the majority of the non-proliferating cells in the human body.

To ensure proper progression through the cell cycle, cells have developed a number of checkpoint controls that prevent them from entering into a new phase before the previous one is successfully completed. Furthermore, extracellular signals such as growth factors, nutrients or cell-matrix interactions can trigger and influence cell cycle progression.

Although some features of the cell cycle (e.g., the time required for the defined phases) may vary between cell types even from the same organism, the underlying principle and the components involved are largely conserved from yeast to human cells.

4.3.1 Core Components of the Cell Cycle Machinery (Tables 4.3-1 and 4.3-2)

Cyclin-dependent kinases (CDKs): Key components of the cell cycle machinery are the cyclin-dependent kinases, a group of serine/threonine kinases ranging from 30 to 36 kDa that show about 40% amino acid identity with one another. Only upon binding to their regulatory subunits, the cyclins, do they form active heterodimeric complexes. In general, the levels of CDKs are relatively constant throughout the cell cycle, while the levels of cyclins oscillate, and thereby provide the decisive control mechanism.

Cyclin-dependent kinase regulators: This group of proteins includes both activators and inhibitors of CDKs.

The **primary activator** of CDK activity is the cyclin subunit. Cyclin levels rise and fall during the cell cycle, and in this way they control CDK activity in a highly ordered fashion. Different CDK-cyclin complexes are required at different points of the cell cycle. Cyclins share a conserved sequence of 100 amino acids called the cyclin box, which interacts with the CDK subunit. This association causes a conformational change within the CDK, which leads to partial activation of CDKs. In mammals and yeast, cyclin levels are regulated at the transcriptional level (Fig. 4.3-1).

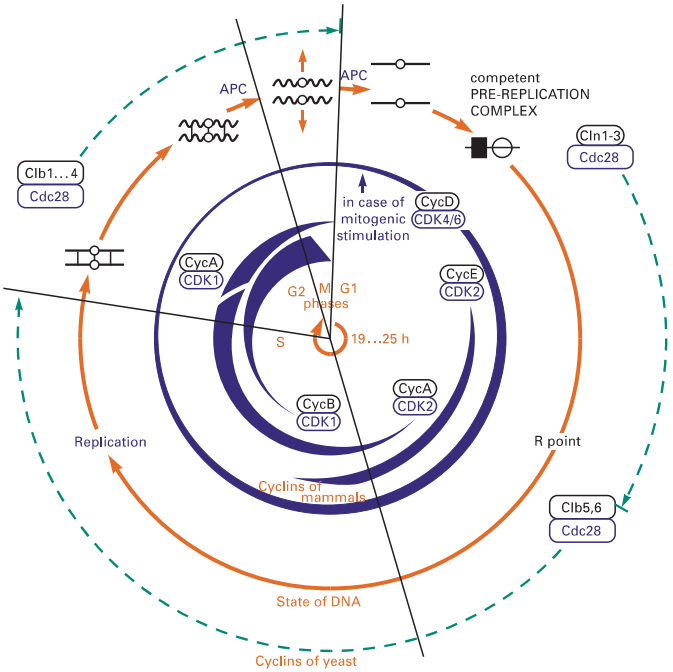


Figure 4.3-1. Eukaryotic Cell Cycle: State of DNA and Cyclin Concentration (inside: Mammals, outside: Yeast)

In mammals, there are four major classes of cyclins, each being defined by the stage of the cell cycle at which they bind to and activate CDKs:

- G_1/S cyclins: Cyclin E

  They activate CDKs in late G_1, and trigger cell cycle entry. During S phase, their levels drop significantly.
- S cyclins: Cyclin A

  After progression through the restriction point, they bind to CDKs and help to stimulate chromosome duplication. The levels of S cyclins remain elevated until mitosis.
- M cyclins: Cyclin B

  They activate CDKs that drive entry into mitosis at the G_2/M checkpoint.

Table 4.3-1. Mammalian Proteins Involved in Cell Cycle Regulation

Proteins	Expression / Activation and Inhibition	Partners / Targets
Cyclins and cyclin-dependent kinases:		
Cyclin A	Periodic expression, required for G_1/S transition and progression through S-phase, a role in G_2/M transition	CDK2 (CDK = cyclin-dependent kinase) in S-phase; CDK1 in G_2/M transition
Cyclin B (1, 2)	Periodic expression, essential for G_2/M transition	CDK1 (analogous to Cdc2 of *S. pombe*). It is involved in nuclear envelope breakdown, chromosome condensation, alignment of chromosomes on the metaphase plate.
Cyclin C	?	CDK8
Cyclin D (1, 2, 3)	Mitogen-responsive expression in G_1 phase; cell-type specific patterns of expression; cyclin D1 and CDK4 are overexpressed in many tumors	CDK4,6: The cyclin-CDK complex phosphorylates Rb protein
Cyclin E	Periodic expression, peak at G_1/S transition, stimulated by E2F, mitogen-independent	CDK2
Cyclin F	Periodic expression, peak at G_2/M transition	?
Cyclin G	Does not vary during cell cycle, transcription is activated by p53	?
Cyclin H	Protein levels do not vary during cell cycle	CDK7 (MO15): cyclin H-CDK7 is also termed 'CDK-activating kinase' (CAK). CAK phosphorylates CDK2, CDK4, CDK6
Regulatory proteins:		
E2F-DP1 (heterodimeric, several isoforms: E2F 1, 2, 3, 4, 5, 6; DP1, 2)	Phosphorylation of DP1 by cyclin A-CDK2 in S phase, inhibits E2F binding to DNA and thus activation of further transcription	Associated with Rb. Dissolution of this complex by Rb phosphorylation in late G_1 phase enables free E2F to activate transcription of genes encoding e.g., dihydrofolate reductase, thymidine kinase, thymidylate synthase, DNA-Pol-, CDK1, E2F-1, b-Myb, c-Myc, cyclins E and A etc. – E2F-1, 2, 3 bind to Rb, E2F-4, 5 bind to p107 and p130.
Rb (retinoblastoma tumor suppressor protein)	Phosphorylated by CDKs 4,6,2 in late G_1 phase. Remains hyperphosphorylated until reentry into G_1 phase. Connects cell cycle with transcriptional machinery. Mutations leading to loss of function as tumor suppressor protein are involved in tumorigenesis. Closely related molecules: p107, p130	Binds to and negatively regulates E2F-DP1, also binds to D-type cyclins, CDKs, Elf-1 [1], MyoD, PU1, ATF-2 [2], c-Abl [3], Mdm2 [4], UBF [5]
p53	Expression induced by, e.g., ionizing radiation	Tumor suppressor protein, activates expression of p21 (Cip1).
Cip/Kip family: p21 (Cip1)	Induced by tumor suppressor protein p53, e.g., after radiation. Potent and universal inhibitor of CDK activity, capable of inducing cell cycle arrest	Inhibits activated cyclin-CDK4, 6; CDK2 and CDK1 complexes
p27 (Kip1, kinase inhibitory protein 1)	Low level expression during cell cycle, increases upon mitogen deprivation. This induces cell quiescence (G_0 phase). Also induced by TGF, cAMP or contact inhibition.	Inhibits activated G_1 cyclin-CDK complexes by formation of a heterotrimeric complex.
p57 (Kip2)		Inhibits activated cyclin-CDK4, 6 and CDK2 complexes
Ink4 family: p16 (Ink4a), p15 (Ink4b), p18 (Ink4c), p19 (Ink4d)	Frequent mutations of the the tumor suppressor protein p16 play a role in *human* tumorigenesis	Inhibit cyclin D-CDK4, 6 by competition with cyclin D
APC (anaphase promoting complex)	Activated in anaphase via phosphorylation after activation of M-phase cyclins A, B-CDK1, active until accumulation of G1-CDKs	Essential for sister chromatid separation. Promotes ubiquitination of proteins containing a destruction box (e.g., M-phase cyclins, in *yeast*: Pds1, Cut2).

[1] Elf-1 = Lymphoid- specific transcription factor, association with Rb blocks G_0 exit; [2] ATF-2 = Activating transcription factor-2; [3] c-Abl = Transcripton factor, promotes the phosphorylation of the CTD domain of RNA Pol II; [4] Mdm2 = Mouse double minute-2; [5] UBF = Upstream binding factor, enhances Pol I transcription.

- G_1 cyclins: Cyclin D

  These types of cyclins, which are present in most cells, control the activities during the G_1 interval.

In addition to binding of the cyclin, full activation of the CDK holoenzyme requires phosphorylation on a threonine residue (T172 in CDK4 and T160 in CDK2) in the T-loop of the kinase subunit. This is carried out by the CDK7-cyclin H complex, also known as CAK (CDK activating kinase), a serine/threonine kinase (Fig. 4.3-2).

In contrast, phosphorylation of adjacent threonine and tyrosine residues (T14/Y15 in CDK1), which is mediated by dual-specific kinases (WEE1 and MYT1), inhibits CDK activity. CDC25 phosphatases can dephosphorylate these amino acid residues, and trigger entry into mitosis.

Degradation of cyclins (4.5.7) involves the ubiquitin-dependent proteolytic machinery, which requires the destruction box sequence near the N-terminus of mitotic cyclins.

Two distinct families of **CDK inhibitors (CKIs)** have been discovered:

- The INK4 family includes p15 (INK4b), p16 (INK4a), p18 (INK4c) and p19 (INK4d), and specifically inactivates G1 CDKs (CDK4 and CDK6) through formation of stable complexes with the CDK enzyme before cyclin binding.
- The Cip/Kip family includes p21 (Waf1, Cip1), p27 (Kip1, Cip2) and p57 (Kip2). These inhibitors form heterodimeric complexes with the G1/S CDKs. While p21 is induced through activation of p53 after DNA damage, the expression and activation of p15 and p27 increases in response to, e.g., transforming growth factor β (TGF-β).

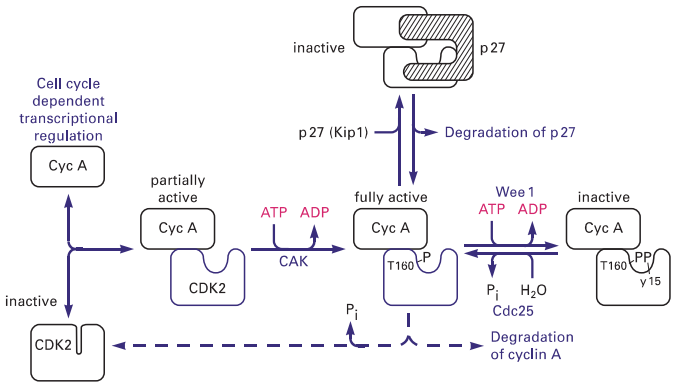


Figure 4.3-2. Activation and Deactivation of Cyclin-CDK Complexes (Example: Mammalian Cyclin A / CDK2)

Cyclin-dependent kinase substrates: Upon activation of CDKs, they phosphorylate target proteins on consensus sites, resulting in changes that are physiologically relevant for cell cycle progression. The primary substrates of CDK2 and CDK4/6 are members of the retinoblastoma family, RB, p107 and p130. Other CDK substrates include p27, nuclear protein mapped to the ATM locus (NPAT), WEE1, CDC25 and cytoskeletal proteins such as nuclear lamins, microtubules and vimentin required for correct mitosis.

4.3.2 Cell Cycle Regulation in Yeast (Fig. 4.3-3)

Regulation of G_1 to S-phase transition: The early steps of DNA replication have been best studied in yeast (*Saccharomyces cerevisiae*). Functional analogies to mammals are shown in Table 4.3-2.

Once the cell has grown beyond a critical size, which depends, e.g., on sufficient nutrient supply, a START signal is generated (equivalent to the restriction point in animal cells), resulting in the activation of the G_1 cyclin Cln1,2,3-cyclin-dependent kinase Cdc28 complex in the late G_1 phase. The activated kinase performs three functions:

- It phosphorylates and activates the transcription factors Sbf and later Mbf which promote transcription of enzymes necessary for DNA replication.

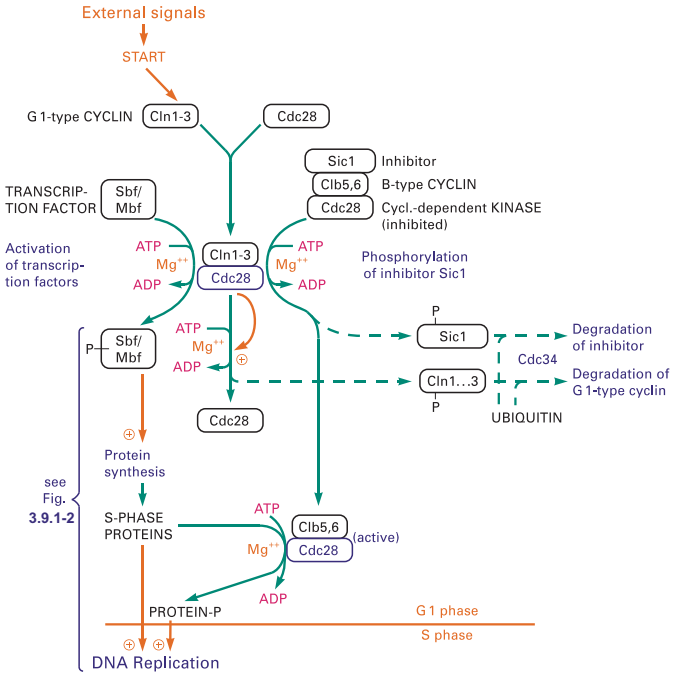


Figure 4.3-3. Regulation of DNA Replication by Cyclins in Yeast

Table 4.3-2. Functional Analogies/Homologies Between Cell Cycle Proteins

	Mammals	Yeast (*S. cerevisiae*)
Cyclin dependent kinases	CDK1/Cdc2, CDK2, CDK3, CDK4, CDK6 CDK7	Cdc28 (cell division cycle) = p34 (Cdc2 in *S. pombe*) Kin28
Cyclins	G_1 type cyclins: cyclin D, E S phase cyclins: cyclin E, A Mitotic cyclins: cyclin A, B	G_1 type cyclins: Cln1, 2, 3 B type cyclins: Clb5, 6 B type cyclins: Clb1, 2, 3, 4
Ubiquitin-protein ligase	Human Cdc34 APC	Cdc34 APC
Anaphase inhibitor	–	Pds1 (Cut2 in *S. pombe*)
DNA damage checkpoint (4.3.6)	p53 p21 ATM (kinase, ataxia telangiectasia mutated) ATR (FRP)	– – TEL1 (regulates telomere length) Mec1 (ESR1, SAD3)
Spindle assembly checkpoint	hsMad2 (*human* homologue of Mad2)	Mad2 (mitotic arrest deficient)

- It also phosphorylates Sic1 (p40), an inhibitor of S-phase CDK activities. This targets Sic1 for degradation by Cdc34 (ubiquitin-protein ligase), which conjugates it with activated ubiquitin (Ub) and thus initiates its degradation (Fig. 4.5.7-1). This way, the previously inhibited S-phase (B-type) cyclin complex Clb5,6–Cdc28 is activated. It phosphorylates and thus activates factors, which then enable the entrance into the S-phase.
- The G_1-phase cyclins Cln1, 2, 3 are marked analogously for degradation, which terminates the G_1 phase.

Control of the pre-replication complex assembly in yeast: The DNA replication process starts at autonomously replicating DNA sequences (ARS) by assembling a pre-replication complex (pre-RC). This complex is compiled at the end of M phase from the proteins Orc1 to 6 (which remain there throughout the cell cycle), the activating enzyme Cdc6 and several Mcm proteins. The presence of these compounds ('licensing factors') renders chromatin competent to replicate upon the arrival of the START signal. At the beginning of the S phase the complex dissolves, effected by the Clb5,6–Cdc28 activity. A new complex cannot be formed before the Clb5,6–Cdc28 and later the Clb1 to 4–Cdc28 activities cease (at the end of M phase). Thus replication starts only once during a cycle.

4.3.3 G_1 to S Transition in Mammalian Cells (Fig. 4.3-4)

Cells integrate external information such as growth factors or nutrient availability only during the first two-thirds of the G_1 phase. After passing the so called "restriction point" (R point), they proliferate independently of mitogenic stimuli. Otherwise, they exit from the cell cycle and enter the resting phase G_0. Therefore, the R point is considered to be the central event in normal cellular proliferation control.

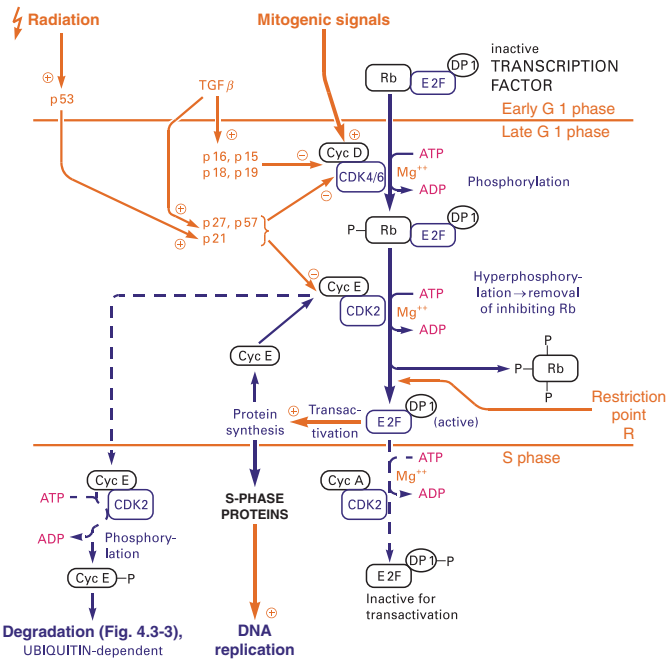


Figure 4.3-4. Regulation of DNA Replication in Mammals

Following generation of sufficient amounts of D-type cyclins (D1, D2 and D3) – induced by a variety of mitogenic growth stimuli –, CDK4/6-cyclin D complexes are formed, which partially phosphorylate Rb. While Rb in its hypophosphorylated form tightly binds and inhibits the E2F-DP1 transcription factors, partially phosphorylated Rb can no longer fully suppress their activity. Consequently, E2F-DP1 is able to trigger transcription of molecules required for G1/S transition, such as cyclin E. This, in turn, forms a complex with CDK2, which hyperphosphorylates Rb to render it completely inactive and therefore unable to continue to block cell cycle progression towards S-phase. Thus, DNA replication can take place.

In addition, Rb can be regulated by acetylation. Histone acetylases - associated to p300/CBP – which are under cell cycle control can prevent efficient phosphorylation by CDK2-cyclin E complexes.

Cells do not rely exclusively on CDK4/6 to fully activate CDK2. p27 is phosphorylated by CDK2, and undergoes proteasome-mediated degradation.

At the end of the G_1 phase, CDK2 activity decreases, which is mainly controlled by degradation of the E-type cyclins through the specific ubiquitin ligase CDC4 (AGO).

mPP1 phosphatase dephosphorylates Rb after mitosis and restores its growth-inhibitory function.

Control of G_1 to S transition via exogenous stimuli: Extracellular stimuli affect the decision to transit the R point and enter into the late G_1 phase. Although the three D-type cyclins expressed in mammalian cells seem to have redundant functions, they are controlled by different extracellular signals and signal transduction pathways, which gives cells enhanced flexibility of response:

- Cyclin D_1 is often induced by growth factors via the Ras-Raf-MAP kinase signal transduction pathway (Fig. 7.5-4). The cyclin D1 gene promoter carries binding sites for the transcription factors AP-1, Tcf/Lef, and NFκB (Table 4.2.2-2).
- Cyclin D_2 is positively regulated by cAMP levels, which are controlled by seven-membrane-spanning cell surface receptors (Fig. 7.4-2).
- Expression of the cyclin D_3 gene is responsive to STAT3 and STAT5 transcription factors, which respond to interleukin receptors (7.5.4).
- Cyclin E levels or the activity of cyclin E-CDK2 complexes are not directly modulated by extracellular stimuli, but indirectly via E2F-DP1 upon induction of cyclin D-CDK4/6 complexes (see above).

Besides positive regulation, extracellular signals can affect the cell cycle machinery negatively as well:

- Serum starvation leads to rapidly decreasing levels of cyclin D and induction of p27, which subsequently inhibits the action of various cyclin-CDK complexes. Terminally differentiated cells or cells trapped in G0 harbor high levels of p27 (and p21), which ensure that cyclin-CDK complexes remain inactive.
- Transforming growth factor β (TGF β) induces the expression of p15 (INK4b) and p16 (INK4a), which compete with D-type cyclins for CDK4/6 and thus inhibit their kinase activity.

Pathological aspects: In cancer, the control of cell division is fundamentally changed, resulting in unrestrained cell proliferation. Genetic alterations have been identified in a variety of genes encoding the cell cycle machinery, and include overexpression of cyclins (mainly cyclin D1 and cyclin E1) and CDKs (mainly CDK4 and CDK6), as well as loss of inhibitors of CKIs (p15, p16, p27) and Rb. These changes frequently result from chromosome alterations (amplifications, translocations, deletions) or epigenetic inactivation (e.g., methylation of the promoter controlling the p16 transcription). In addition, inactivating mutations have been identified in CDK4 and CDK6.

4.3.4 G_2 to M Transition in Mammalian Cells

In higher eukarya, passage through the G_2 phase to mitosis (M phase) requires two cyclins, cyclin A2 and cyclin B1. During the late stages of DNA replication, cyclin A2 activates CDK1 to enable the nuclear envelope breakdown. Thereafter, cyclin A is degraded, facilitating the formation of cyclin B-CDK1 complexes triggering the onset of mitosis.

Nuclear envelope breakdown (NEBD): Prior to mitosis the nuclear envelope needs to be disassembled to ensure that the DNA can be equally segregated between each daughter cell (‘open mitosis’). The nuclear envelope (NE) is formed by two juxtaposed lipid membranes (outer and inner nuclear membrane) perforated by holes, which are occupied by nuclear pore complexes (NPCs, 4.5.3). A lamina layer is located underneath the inner membrane.

The nuclear envelope breakdown comprises several distinct steps, including NPC disassembly and clearance of nuclear envelope membranes. These processes are controlled by mitotic kinases, which phosphorylate the NE proteins and thus disrupt their interaction. The lamina layer is depolymerized by phosphorylation of lamins by CDK1.

Besides CDK1, other kinases such as protein kinase C (PKC), Aurora A and polo-like kinase 1 (PLK1) contribute to NEBD.

4.3.5 Mitosis in Mammalian Cells (Fig. 4.3-5)

Mitosis is conventionally divided into several discrete stages, based on the morphology of the cell.

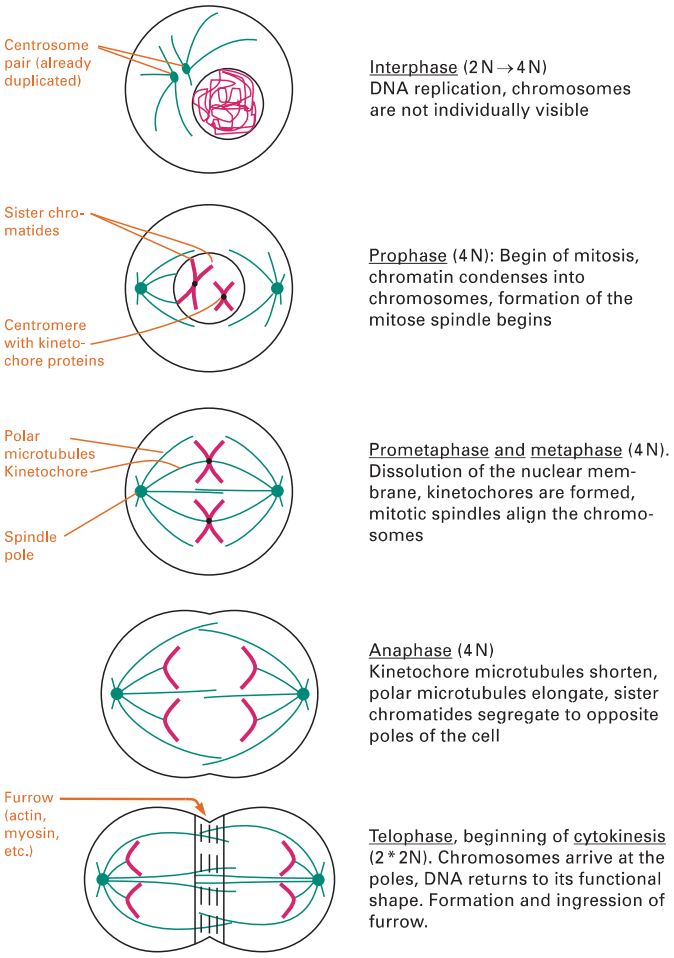


Figure 4.3-5. Summary of Mitotic Cell Division in Animals
Reactions are shown for a single, homologous pair of chromosomes (2N) of these diploid organisms.

The anaphase-promoting complex/cyclosome (APC/C), a large multiprotein E3 ubiquitin ligase (4.5.7) regulates the orderly progression through mitosis. It targets key mitotic regulators for destruction by the proteasome (similarly to Cdc34, 4.3.2). Besides its core components, the APC/C requires a member of the WD40 family for activity, Cdc20 or CDH1.

When cells are fully committed to mitosis, cyclin B-CDK1 complexes phophorylate the APC/C. Only this form is bound by Cdc20 and becomes thus activated. Since, however, CDH1 can bind to and activate both phosphorylated unphosphorylated APC/C, it is kept inactive in G2 phase due to a phosphorylation by cyclin G2-CDK. In animal cells, CDH1 is also sequestered and inactivated by EMI1, a cysteine-rich F-box protein that blocks the substrate binding site of CDH1.

Proteins that are targeted for degradation by APC/C include cyclin B1 (which is the major regulator of mitosis) as well as securin (Fig. 4.3-8). Breakdown of securin leads to activation of separase,

which subsequently cleaves the cohesion ends that attach the sister chromatids to each other. Besides cyclin B1, cyclin A, Nek2A kinase and HOXC10 become degraded by APC/C. The chromosomes attach to the spindle, and the spindle-assembly checkpoint (SAC) is activated to prevent cells from prematurely entering the anaphase of mitosis.

In somatic cells, the decline in cyclin B-CDK1 activity allows the APC/C to bind CDH1, which widens the recognition pattern of APC/C for substrates. One of these substrates is Cdc20 itself, leading to a shift from APC/C-Cdc20 to APC/C-CDH1. Consequently, the spindle assembly checkpoint can no longer turn off APC/C, which now degrades regulatory proteins and components of the mitosis- or cytokinesis-specific structures.

Spindle and chromosomal segregation (Fig. 4.3-6): The primary structural element of the spindle is an antiparallel array of microtubules composed of polymerized heterodimeric α/β tubulin subunits. The polymer is highly dynamic and switches stochastically between growing and shrinking phases, based on the binding and hydrolysis of GTP at the nucleotide exchangeable site in β tubulin (dynamic instability). Microtubules have a slow-growing minus end anchored at the spindle pole and a fast-growing plus-end.

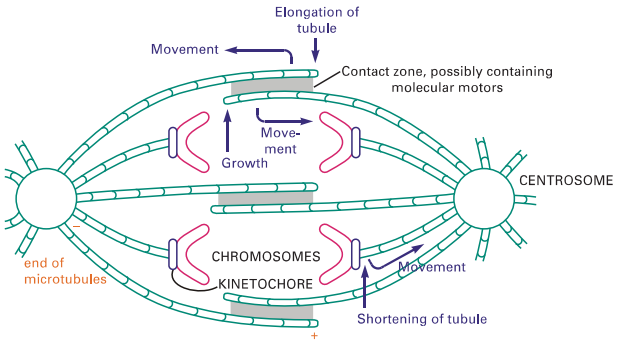


Figure 4.3-6. Elongation and Shortening of Microtubules During Anaphase

Three different kinds of microtubules exist within mitotic spindles:

- kinetochore microtubules, which attach with their plus end to the high-affinity site in a chromosome composed of multiple proteins (kinetochores),
- interpolar microtubules, that originate from opposite poles and stabilize the spindle,
- astral microtubules, which extend away from centrosomes and position the spindle within the cell.

Chromosome movements on the spindle are coupled to changes in the length of microtubules attached to the kinetochores. Several molecular motors contribute to this process:

- Cytoplasmic dynein, a minus-end-directed motor protein, drives poleward chromosome movement.
- CENP-E, a kinesin with large molecular mass and plus-directed motor activity, has been implicated in chromosome congression to the metaphase plate.
- MCAK, a member of the microtubule-destabilizing Kin I subfamily, destabilizes kinesin, which functions in depolymerizing microtubules proximal to the kinetochore as the chromosomes move towards the spindle pole.

In animals, the microtubule spindle also directs the accumulation and the alignment of contractile proteins (filaments of actin, myosin II and other proteins) around the metaphase plate underneath the plasma membrane by an unknown signaling process. The actomyosin contraction mechanism (partially resembling muscular contraction) pulls the membrane inwards (furrow ingression) As the furrow gets narrower, the filaments are gradually disassembled. Cell separation (cytokinesis) is achieved by the resolution of one membrane into two distinct membranes.

In plants and fission yeast, the division plane is filled with new membrane and cell wall components, starting at the center. They are transported there in vesicles from the Golgi apparatus along microtubules. The accuracy of cellular division in yeast is about 1 error/10^5 divisions.

4.3.6 Cell Cycle Checkpoints

The propagation of accurate copies of the genome is essential for survival and for avoidance of various diseases. Cells have evolved multiple mechanisms (3.9.1.6, 3.9.2) to ensure error-free transmission of the genetic material to subsequent generations. These are referred to as cell cycle checkpoints. These mechanisms allow cells to actively halt progression through the cell cycle until they can ensure that the previous process, such as replication of the DNA or mitosis, is successfully completed.

In some cases, damage to the cellular DNA is so substantial that it overwhelms the restorative power of the DNA repair mechanism. Elimination of the defective cell via apoptosis represents the only desirable response in this situation. Although the transcription factor p53 is a major governor of this process, the cell cycle machinery itself can also directly participate to trigger the apoptotic response (see below).

G_1 phase checkpoint (Fig. 4.3-7): The dominant checkpoint response to DNA damage in mammalian cells traversing through G_1 is the ATM/ATR - CHK2/CHK1 pathway, which targets two critical effectors operating in different branches of the G_1 checkpoint, p53 and Cdc25A phosphatase. ATM/ATR directly phosphorylate the p53 transcription factor within its N-terminal transactivation domain, particularly on serine-15. In addition, the ubiquitin ligase MDM2, that normally binds and degrades p53, is also targeted after DNA damage by ATM/ATR and CHK2/CHK1.

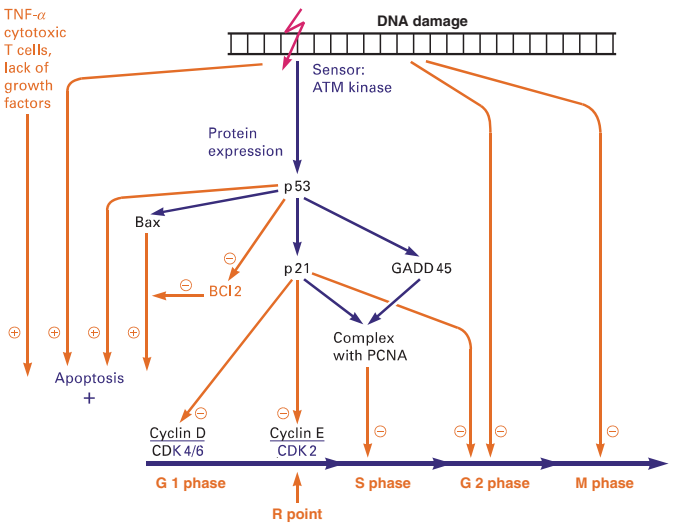


Fig. 4.3-7. DNA Damage Checkpoint in Mammals and Initiation of Apoptosis

Modification of both proteins leads to stabilization and activation of p53, which induces the expression of p21. p21 silences the activity of cyclin E-CDK2 complexes. This prevents the initiation of DNA synthesis and also preserves the Rb pathway in its growth-suppressing mode, causing a sustained G_1 arrest.

In late G_1, phosphorylation of Cdc25A by CHK1 and CHK2 enhances ubiquitination and degradation of Cdc25A and prevents the Cdc25A-mediated dephosphorylation of CDK2. This inhibits initiation of DNA synthesis, since it blocks the loading of Cdc45 (a protein required for recruitment of DNA polymerase α into pre-replication complexes) onto chromatin. This checkpoint pathway is implemented rapidly, operates independently of p53, but is relatively transient and delays the cell cycle progression for only a few hours.

S phase checkpoints: In S phase, three checkpoints exist to prevent genetic instability arising from difficulties with the replication process itself:

- The replication checkpoint is initiated when the progression of the replication forks become stalled in response to stresses such as collision of replication forks with damaged or aberrant DNA. The core component of this machinery comprises the replication protein A (RPA), the ataxia-telangiectasia and RAD3 related (ATR)-ATR-interacting protein (ATRIP) complex, claspin, RAD17 and the RAD9-RAD1-HUS1 complex. Besides inhibiting the DNA replication from yet unfired origins via blocking the cyclin-CDK and CDC7-DBF4 kinases, it protects the integrity of the replication forks allowing cell cycle progression after DNA repair.
- The S/M checkpoint ensures that cells do not divide prior to successful duplication of their entire genomes. Like the replication checkpoint, it depends on DNA replication. The key target of this checkpoint is cyclin B-CDK1.
- The intra-S-phase checkpoint is induced by DNA double strand breaks (DSBs), which are generated outside of active replicons. This checkpoint is independent of replication forks.

In contrast to the sustained G_1 arrest, none of the S phase checkpoints requires p53 for its activity.

G_2 checkpoint responses: This checkpoint prevents cells from initiating mitosis when they experience DNA damage during G_2. The key downstream target of the G_2 checkpoint is the mitosis-promoting activity of the cyclin B-CDK1 complex. Activation of cyclin B-CDK1 kinase is inhibited by ATM/ATR, CHK1/CHK2 and subcellular sequestration of Cdc25A, which normally activates CDK1 at the G_2-M boundary.

53BP1 and BRCA1 as well as other upstream regulators of Cdc25C and cyclin B-CDK1 (the Polo-like kinases) also contribute to G_2 checkpoint responses.

The p53-regulated checkpoint inhibitor p21 is most critical for the G1 checkpoint. Other transcriptional targets of p53, such as GADD45 (growth arrest and DNA damage) and 14-3-3α proteins, are important for G2 arrest. GADD45 binds to PCNA (a processivity factor for DNA polymerase δ or ε, 3.9.1.3, 3.9.2.3) and possibly depletes it in replication complexes. Thus, it becomes a potent inhibitor of cell proliferation. It also enhances nucleotide excision repair (3.9.2.3).

Spindle assembly checkpoint (Fig. 4.3-8): The onset of anaphase is tightly regulated by the spindle assembly checkpoint (SAC), which prevents premature advance to anaphase that would result in incorrect sister chromatid separation. The most important triggers of this checkpoint are unattached kinetochores, which recruit and process checkpoint components such as Mad1, Mad2, Bub1, Bub3 and MPS1 kinase, which, in turn, inhibit Cdc20-dependent recognition of cyclin B and securin by APC/C. Unattached kinetochores act as catalytic sites for the activation of Mad2, which synergistically with Mad3, associates with Cdc20, preventing it from activating the APC/C and degradation of securin and cyclin B1. Mad1 hyperphosphorylation, (perhaps directly by Mps1) may also be involved in this exchange.

After all the chromosomes are properly attached by kinetochore microtubules and aligned at the metaphase plate, the spindle assembly checkpoint is turned off (right panel). Activated Mad2 is no longer generated resulting in the activation of Cdc20-APC. Activated Cdc20-APC catalyzes the ubiquitination of securin, leading to its degradation by the proteasome, and causes the release of separin. Separin is then able to cleave the sister chromatid cohesion complex, triggering the separation of the sister chromatids and the onset of anaphase.

The spindle checkpoint is an integral part of each mitosis in somatic cells, and its elimination leads to aneuploidy (deviating chromosome numbers).

Apoptosis (Fig. 4.3-7, see 7.6 and 8.2.5 for more details): Apoptosis is the process of programmed cell death, which involves a series of biochemical events leading to a modified cell morphology (e.g., loss of membrane asymmetry, nuclear fragmentation) followed by death. p53 plays a major role in triggering this process due to upregulation of cell surface death receptor proteins including Fas/APO1 and DR5 (extrinsic apoptotic pathway, compare 8.2.5). p53 also transcriptionally activates the cytoplasmic pro-apoptotic proteins PIDD and Bid, as well as mitochondrial pro-apoptotic proteins, such as Bax, Bak, Puma and Noxa. (intrinsic apoptotic pathway, see 8.1.5). Bax activates the cysteine proteases caspases, which promote apoptosis by activating DNAses and other target proteins. In contrast, Bcl-2 antagonizes Bax effects. Consequently its expression is downregulated by p53 when Bax is induced. Bcl-2 shows homology with Bax and is able to form heterodimers with it, which inhibit the apoptosis mechanism. p53 upregulates APAF1, a protein involved in apoptosome formation. p53 can also induce apoptosis independent of its transcriptional activatory function.

Inappropriately high levels of the transcription factor E2F1 through S phase, which may occur after damage to the Rb control pathway (e.g., mutations in Rb, unrestrained proliferation signals), can also trigger an apoptotic response via p53. This process ultimately links the cell cycle machinery with the apoptotic program.

Pathological aspects: DNA damage is a major contributor to the development of cancer. While normal cells arrest or die in response to DNA damage, mutations in genes associated with apoptosis or cell cycle checkpoint proteins allow the survival and growth of cells with

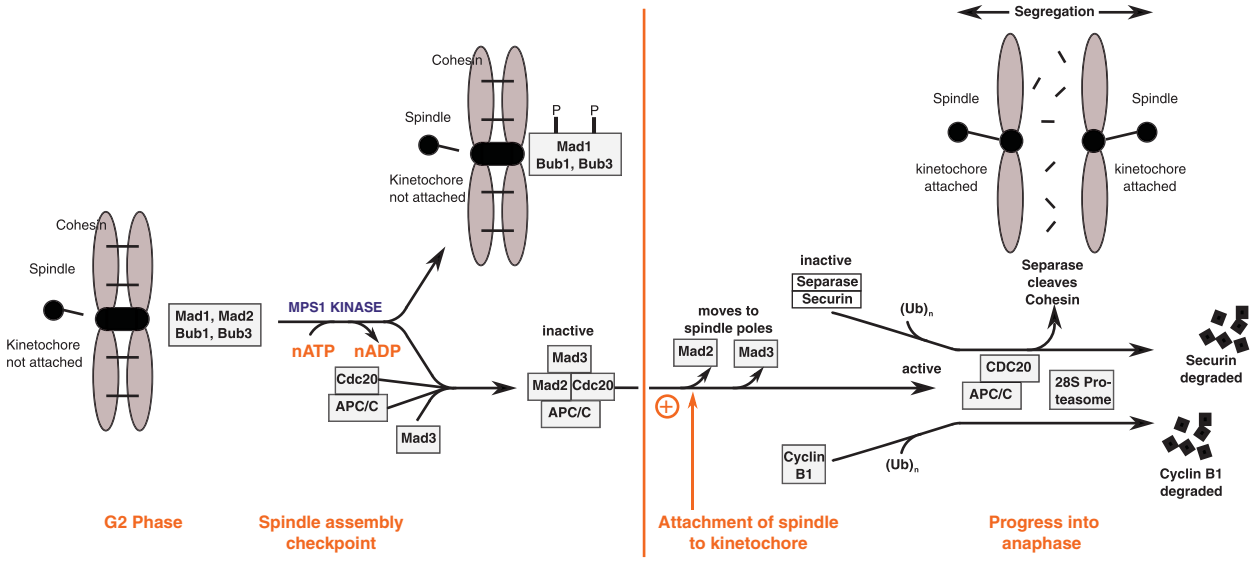


Figure 4.3-8. Spindle Assembly Checkpoint

genomic abnormalities. Therefore, alterations in cell cycle checkpoint proteins are frequently found in human tumors. The central role of p53 in protecting a cell from DNA damage is shown by the fact that mutations occur in more than 50% of all human tumors. Besides p53, mutations in BRCA1, BRCA2, ATM and other proteins linked to DNA damage response have been identified.

Chromosomal instability and aneuploidy (abnormal number of chromosomes), two major hallmarks of cancer, are prominent consequences of defective M-phase checkpoints. Several mutations in molecules that control correct sister-chromatid separation during mitosis have been identified in human tumors, e.g., BUB1, BUBR1, MAD1, MAD2, and all members of the ZW10-ROD-zwilch complex. Germline mutations in the *bub1b* gene locus are associated with a rare disorder called mosaic variegated aneuploidy (MVA), which is characterized by childhood cancer. Cells from individuals suffering from MVA show mosaicism for chromosomal gains and losses.

Protein degradation: The ubiquitin system provides controlled protein degradation in all eukaryotes (Fig. 4.3.4, see also 4.5.7). The enzymes Cdc34, APC (see above) and others ligate activated ubiquitin with the ε-amino group of lysine in the "condemned" protein. In most cases, poly-ubiquitin chains are formed. Protein degradation then occurs with a large complex, the 28S proteasome.

The Ub degradation apparatus recognizes appropriate proteins by their N-terminal amino acid (fast degradation with R, K, D, L or P-one letter amino acid abbreviations) or by internal PEST sequences (stretches of amino acids rich in P, E, S, T). In some proteins (e.g., mitotic cyclins) a destruction box (D-box) close to the N-terminus is present. Recognition of some substrates (Sic1, Cln2) takes place by a Cdc53-Cdc4-Skp1 complex, which is associated with Cdc34.

Literature:

Baker, D.J. *et al*. Cell. Mol. Life Sci. 20047;64:589–600.

Bartek, J. *et al*. Nat. Rev. Mol. Cell. Biol. 2004;5:792–804.

Güttinger, S. *et al*. Nat. Rev. Mol. Cell. Biol. 2009;10:178–191.

Kastan, M., Bartek, J. Nature 2004;432:316–323.

Kops, G.J.P.L. *et al*. Nat. Rev. Cancer 2005;5:773–785.

Lukas, J. *et al*. DNA Repair 2004;3:997–1007.

Lundberg, A.S., Weinberg R.A. Eur. J. Cancer 1999;35:1886–1894.

Malumbres, M., Barbacid, M. Nat. Rev. Cancer 2001;1:222–231.

Malumbres, M., Barbacid, M. Nat. Rev. Cancer 2009;9:153–166.

Massague,. J. Nature 2004;432:298–306.

Mathon, N.F., Lyod, A.C. Nat. Rev. Cancer 2001;1:203–213.

Musacchio, A., Salmon, E.D. Nat. Rev. Mol. Cell. Biol. 2007;8:379–393.

Nurse, P. Cell 2000;100:71–78.

Park, M.S., Koff, A. Curr. Prot. Cell. Biol. 1998;8.1.1-8.1.9.

Pines, J. Trends Cell Biol. 2006;16:55–63.

Przwelaka, M.R., Glover, D.M. Ann. Rev. Genetics 2009;43:439–465.

Sanchez, I., Dynlacht, B.D. Semin. Cell. Dev. Biol. 2005;16:311–321.

Sherr, C.J., Roberts, J.M. Genes. Dev. 2004;18:2699–2711.

Vermeulen, K. *et al.* Cell. Prolif. 2003;36:131–149.

Wittmann, T. *et al*. Nat. Cell. Biol. 2001;3:E28–E34.

Zhou, J. *et al.* J. Cell. Sci. 2002;115:3547–3555.

4.4 Posttranslational Modification of Proteins

Röbbe Wünschiers

4.4.1 Protein Processing in the Endoplasmic Reticulum

The endoplasmic reticulum (ER) is an important organelle of eukarya. With respect to proteins, its major functions are co- and post-translational glycosylation, folding, preparation for targeting and membrane attachment.

In eukarya, glycosylation is very frequent among membrane associated and secreted proteins. The main glycosylation machinery is located in the ER and the Golgi apparatus (4.4.2). The protein backbone is synthesized by ribosomes (4.2.3.1) attached to the ER membrane and threaded into the ER lumen. In case of N-glycosylation, the first steps of sugar attachment take place here (cotranslational glycosylation). This is continued in the Golgi apparatus.

With some secretory proteins, the protein chain is synthesized cytoplasmatically and then translocated through the ER-membrane in a way similar to mitochondrial import (4.5.4), followed by post-translational glycosylation (mostly in yeast, e.g., pre-pro-α-factor). One type of O-glycosylation, the addition of single N-acetylglucosamine residues to either a serine or threonine side chain, takes place in the cytosol, too. This modification occurs with cytosolic and nuclear proteins and is as abundant and dynamic as protein phosphorylation (2.5.2, 7.5).

Another protein modification, O-mannosylation, can be observed in bacteria, fungi and animals. Different from O-glycosylation, protein O-mannosylation is initiated in the ER by the transfer of mannose from dolichol monophosphate-activated mannose to serine or threonine residues of secretory proteins.

In N-linked glycosylation the carbohydrate moiety is attached to the amide nitrogen of the side chain of asparagine, which is a part of the consensus sequence Asn-X-Ser/Thr. This consensus sequence is a necessary but not a sufficient condition for N-glycosylation to occur. A proline residue at X prevents N-glycosylation while a proline residue in the sequence beyond Ser/Thr inhibits N-glycosylation.

In O-linked glycosylation the carbohydrate moiety is covalently linked to the hydroxyl oxygen of serine and threonine. In addition, O-glycosylation occurs as a primary modification of tyrosine and as a secondary modification of 5-hydroxylysine and 4-hydroxyproline. The motif Pro-X-X-Ser/Thr-Pro strongly favors O-glycosylation.

4.4.1.1 Protein Synthesis and Import into the Endoplasmic Reticulum (Fig. 4.4.1-1, upper part)

Most eukaryotic proteins to be glycosylated carry a signal sequence, containing several positively charged amino acids, followed by 7 … 20 hydrophobic amino acids and another charged sequence near the N-terminus. After translation of this ER-trafficking signal sequence at ribosomes within the cytoplasm, the signal recognition particle (SRP) binds to the translation complex. It arrests the polypeptide chain elongation by fixation of the signal sequence to its subunit p54. The whole complex of ribosome, mRNA, nascent polypeptide chain and SRP moves to the ER and 'docks' at the ER-membrane by interaction of the SRP with the SRP receptor. GDP bound to the SRP-receptor (and also to the SRP) is then replaced by GTP. This leads to separation of SRP both from the signal sequence and the ribosome. Thereafter, hydrolysis of GTP removes SRP also from the SRP receptor and enables it to start a new reaction cycle.

The ribosome with the nascent protein chain is transferred to the Sec61 complex, a heterotrimer that is the core component of the translocation system (translocon). The ribosome receptor is part of the Sec61αβγ protein complex. The signal sequence is bound to the transmembrane TRAM protein. Most likely, the Sec61 complex forms an aqueous membrane pore, through which the polypeptide chain is pushed. The protein synthesis is then resumed.

4.4.1.2. Location of ER proteins

- When secretory proteins pass through the pore, the signal sequence is cleaved off by a signal peptidase and the whole protein is translocated into the ER lumen to be exported later.
- The proteins can also be transferred to phosphatidylinositol membrane anchors.
- Transmembrane proteins contain additional stop transfer effector (STE) sequences, which mark the preceding hydrophobic areas to become transmembrane sequences. The STE sequences transiently stop the translocation and mediate the transmembrane integration of the protein (while the protein synthesis continues). In the case of multiple-spanning transmembrane proteins, this procedure is repeated.
  - Frequently, the signal sequence is located at the N-terminus of the nascent protein. When it is cleaved off by a signal peptidase,

the newly formed N-terminus is oriented towards the ER lumen, while the C-terminus remains at the cytosolic side of the membrane (type I membrane proteins).

– If an uncleavable signal sequence (signal anchor sequence) is present in the protein, it remains anchored in the membrane, while the nascent protein is pushed into the lumen. This yields a protein, in which the N-terminus is located in the cytosol and the C-terminus in the ER lumen (type II membrane proteins).

While correctly folded secretory and membrane proteins (4.5.1.2) leave the ER in transport vesicles (4.5.2), special mechanisms exist for proteins destined to remain in the ER:

- Most membrane-bound ER proteins are retained by interaction with cellular structures on the cytosolic surface, probably due to a Lys-Lys-X-X or X-Lys-X-X consensus motif (one or two lysines, 3-4 positions away from the C-terminus).
- Soluble ER proteins, like protein disulfide isomerase or binding proteins (BiP) have special amino acid sequences at the C-terminus, like Lys-Asp-Glu-Leu, Lys-Glu-Glu-Leu or His-Asp-Glu-Leu (in yeast). These sequences most likely interact with a receptor in the *cis*-Golgi network ('salvage compartment'), which returns them to the ER.

4.4.1.3. Synthesis of Dolichol-bound Oligosaccharides and N-Glycosylation (Fig. 4.4.1-1, lower part)

The structure of the glycoside chains plays an important role in the folding mechanism of glycoproteins (4.5.1), which also takes place in the ER.

The stepwise construction of glycosyl chains attached to membrane-bound dolichyl-P (3.5.3.3) up to Dol-PP-GlcNAc$_2$-Man$_5$ takes place on the cytosolic side of the ER-membrane. The 'flipping over' of this structure to the luminal side has been postulated, but not been shown so far. Dol-P-mannose carries mannose units through the ER-membrane, which are used to extend the glycosyl chains both of the dolichol and the glycosylphosphatidylinositol anchors (4.4.1.4). Generally, nucleoside diphosphate glycosides cannot pass through the ER membrane, although they are able to do this with the Golgi membrane by an antiport mechanism (4.4.2.1). The preformed oligosaccharides are transferred from dolichol to an asparagine residue of the protein chain. This amino acid has to be a member of the consensus sequence Asn-X-Ser/Thr (X may be any amino acid except proline). The oligosaccharide transferase complex is associated with the protein translocation machinery.

The structure of the glycoside chains plays an important role in the folding mechanism of glycoproteins (4.5.1), which also takes place in the ER.

4.4.1.4. Formation of Lipid-anchored Proteins in the ER (Fig. 4.4.1-1, left side)

Synthesis of glycosylphosphatidylinositol (GPI) in eukarya: The synthesis starts at the outside of the ER by adding N-acetylglucosamine residues from UDP-GlcNAc to phosphatidylinositol (3.4.3.2), followed by deacylation and acylation, mostly palmitoylation. A flippase brings the sugar moiety into the ER lumen. After triple-mannosylation, phosphoethanolamine is added from phosphatidylethanolamine (3.4.3.2). This structure may be modified by the addition, e.g., of more mannose residues, galactosyl side chains to the core carbohydrates, additional phosphoethanolamine or fatty acids to inositol.

Proteins to be loaded to the GPI anchor have a signal peptide near their C-terminus with three distinct regions: a signal-cleavage/GPI attachment site (including Ala, Asn, Asp, Cys, Gly and Ser, but without strict sequence requirement), a spacer region of about 7 … 14 amino acids and a stretch of ca. 20 hydrophobic amino acids. While the proteins are still attached to the membrane with their C terminus

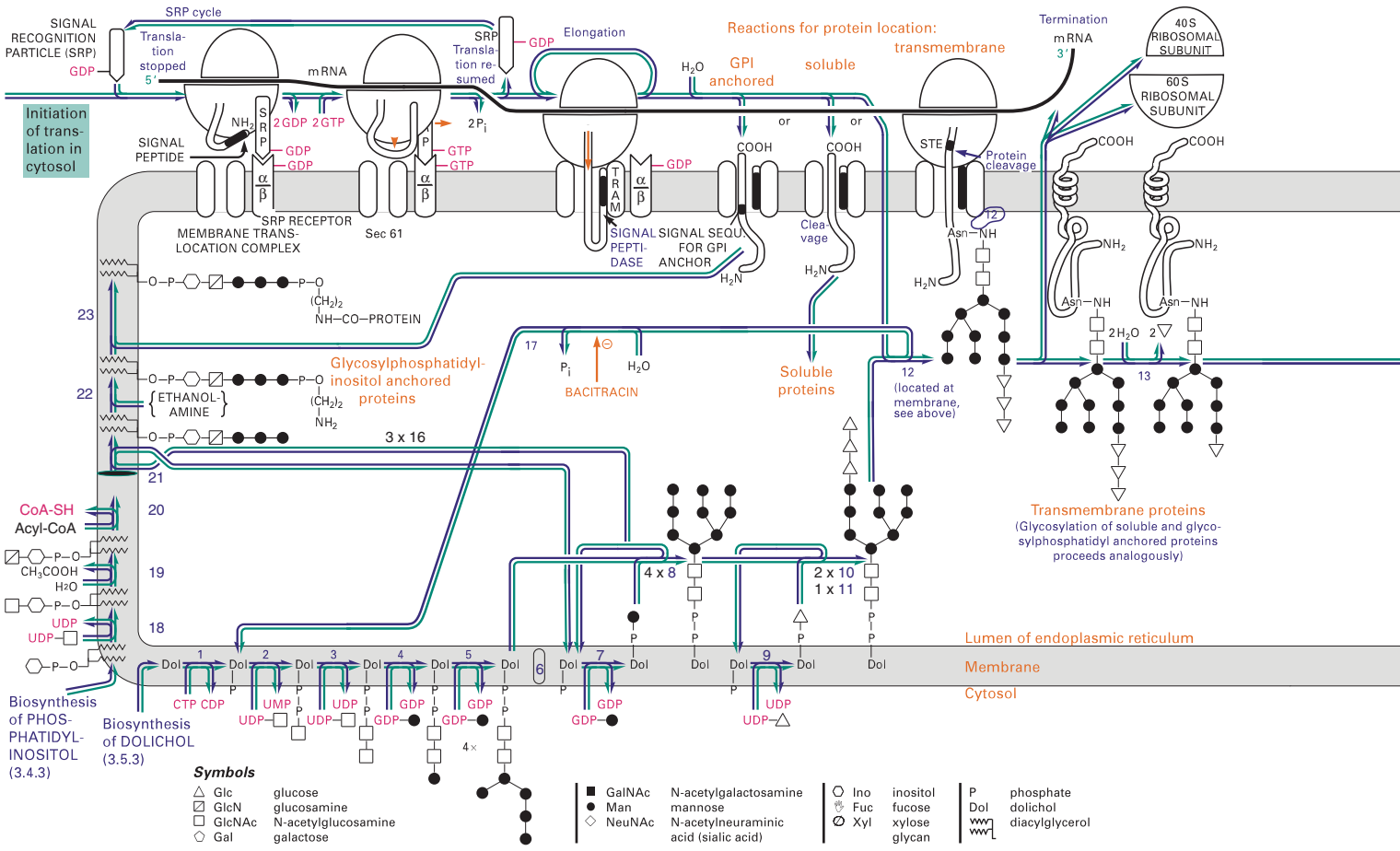


Figure 4.4.1-1. Protein Processing in the Endoplasmic Reticulum

after synthesis, the appropriate peptide bond is probably subjected to a nucleophilic attack by the NH_2 group of ethanolamine-GPI, resulting in transfer of the protein to the anchor.

GPI acts as an anchor for many proteins involved in cell adhesion, membrane signalling events, T-cell activation (e.g., LFA-3; 8.2.1), catalysis (e.g., human erythrocyte acetylcholinesterase, 3.4.3.5 etc.).

4.4.1.5 Acylation of Proteins

In about half of all proteins, acetyl- or myristoyl-CoA reacts with the N-terminus of the nascent polypeptide in an enzyme-catalyzed mode:

CH_3-CO-SCoA + H_2N (Ser/Ala/Met)-protein → CoA-SH + acetyl-protein

or

$C_{13}H_{27}$-CO-SCoA + H_2N(Gly)-protein → CoA-SH + myristoyl-protein.

Myristoylated proteins bind to specific membrane receptors and become membrane-attached this way. Palmitoylation (mostly of a cysteine close to the N-terminus), however, takes place posttranslationally at locations outside of the ER. This reversible reaction may play a role in regulation (e.g., of receptors and of G-proteins, 7.4.1).

Literature:

Burda, P., Aebi, M. Glycobiology 1998;8:455–462.

Casey, P.J. Science 1995;268:221–224.

Chen, W., *et al.* Proc. Natl. Acad. Sci. USA 1995;92:6229–6233.

Christlet, T.H.T., Veluraja, K. Biophys. J. 2001;80:952–960.

Doering, D.L., *et al.* J. Biol. Chem. 1990;265:611–614.

Hartl, F.-U., Hlodan, R., Langer, T. Trends in Biochem. Sci. 1994;19:20–25.

Table 4.4.1-1. Components of Protein Processing in the ER

Unit	Components	Function	Found in
Signal recognition particle (SRP)	associated peptides of 72, 68, 54, 19, 14, 9 kDa, 7S RNA (ca.300 nt)	54 kDa GTPase subunit binds signal sequence of the nascent protein	mammals, homologous to other eukarya, bacteria and plants
Membrane translocation complex (translocon)	SRP receptor (SR), dimeric (α, β/72, 30 kDa)	SRP binds to a GTPase subunit of SR	same
	Sec61 complex, trimeric (α, β, γ/54, 14, 8 kDa)	Integral membrane complex forming the translocation channel, binds ribosomes	same
	Translocating chain associating membrane protein (TRAM/36 kDa glycoprotein)	Spans membrane, transport of some proteins, location of signal sequence?	mammals
Signal sequence receptor complex	subunits α, β, γ, δ	Spans membrane at translocation site, receptor function?	mammals, birds, fish
	Sec62, Sec63	Transport of proteins, interaction with Bip*	yeast
Signal peptidase complex	peptides of 25, 23, 22, 21, 18, 12 kDa	Spans membrane, protease function, removes signal peptide within the lumen	mammals, yeast, plants, fungi
Oligosaccharyltransferase complex	48 kDa protein, ribophorins I (65 kDa) and II (63 kDa)	Spans membrane, transfers within the lumen oligosaccharides to nascent protein chain	mammals, yeast

* Bip = binding protein (Hsp70), ATPase function, present in the lumen of the *ER*.

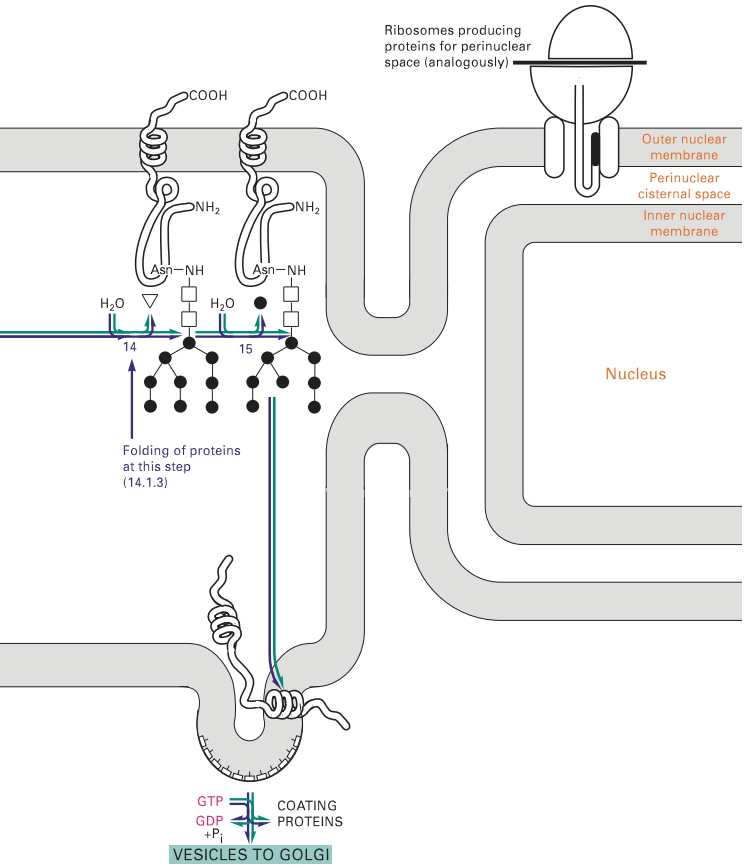


Table 4.4.1-2. Enzymes Involved in Glycosylation Reactions in the ER (EC numbers in parentheses)

1. Dolichol kinase (2.7.1.108)
2. UDP-N-acetylglucosamine-dolichyl-P N-acetylglucosamine-P-transferase (inhibited by tunicamycin, 2.7.8.15)
3. N-acetylglucosaminyldiphosphodolichol N-acetylglucosaminyltransferase (2.4.1.141)
4. Chitobiosyldiphosphodolichol α-mannosyltransferase (2.4.1.142)
5. Glycolipid mannosyltransferases (2.4.1.131, 2.4.1.257)
6. Flippase (encoded by ALG10)
7. Dolichyl-P β-D-mannosyltransferase (inhibited by amphomycin, 2.4.1.83)
8. Dolichyl-P-mannose-glycolipid α-mannosyltransferase (2.4.1.258-261)
9. Dolichyl-P β-glucosyltransferase (2.4.1.117)
10. Dolichyl-P-glucose-glycolipid α-1,3 glucosyltransferases (2.4.1.265, 2.4.1.267)
11. Dolichyl-P-glucose-glycolipid α-1,2-glucosyltransferase (2.4.1.256)
12. Dolichyl-diphosphooligosaccharide-protein glycotransferase (2.4.1.119)
13. Glucosidase I (mannosyl-oligosaccharide glucosidase, inhibited by castanospermine, 3.2.1.106)
14. Glucosidase II (glucan-1,3-α-glucosidase, inhibited by 1-deoxynojirimycin, 3.2.1.84)
15. Mannosyl-oligosaccharide 1,2-α-mannosidase (3.2.1.113)
16. Dolichyl-P-mannose-glycolipid-α-mannosyltransferase (2.4.1.-)
17. Dolichyl diphosphatase (inhibited by bacitracin, 3.6.1.43)
18. Phosphatidylinositol N-acetylglucosaminyltransferase (2.4.1.198)
19. N-acetylglucosaminylphosphatidylinositol deacetylase, (3.5.1.89)
20. Acyltransferase (2.3.-.-)
21. Flippase (uncharacterized)
22. Phosphatidylethanolamine transferase (2.7.-.-)
23. Glycosylphosphatidylinositol:protein transamidase

Helenius, J., *et al.* Nature 2002;415:447–450.

Kinoshita, T., Inoue, N. Curr. Opinion Chem. Biol. 2000;4: 632–638.

Lommel, M., Strahl, S. Glycobiol. 2009;19:816–828.

Low, M.G. Biochim. Biophys. Acta 1989;988:427–454.

Lütcke, H. Eur. J. Biochem. 1995;228:531–550.

Pattison, R.J., Amtmann, A. Trends in Plant Science 2009;14: 92–99.

Rapoport, T.A. Science 1992;258:931–936.

Reithmeier, R.A.F., in Vance, D.E., Vance, J. (Eds.) *Biochemistry of Lipids, Lipoproteins and Membranes. New Comprehensive Biochemistry* Vol. 31. Elsevier, 1996; p. 425–471.

Takeda, J., Kinoshita, T. Trends in Biochem. Sci. 1995;20: 367–371.

4.4.2 Glycosylation Reactions in the Golgi Apparatus

In eukarya, the Golgi apparatus continues the protein and lipid glycosylation that was initiated in the ER and effects the sorting of proteins to guide them to their destination. The Golgi apparatus is asymmetrically split into different compartments. The *cis*-face faces the endoplasmatic reticulum while the *trans*-face is oriented toward the cell membrane. Its parts have the following functions:

- *Cis*-Golgi network: recognition and recycling of proteins, which falsely escaped from the endoplasmic reticulum (ER, 4.4.1.2).
- *Cis*-, *medial*- and *trans*-Golgi compartments: glycosylation reactions.
- *Trans*-Golgi network: numerous sorting events. Proteolytic cleavage as a final processing step starts here with some proteins, and is continued in the vesicles (e.g., some hormones and neuropeptides, 7.1.5).

All compartments are characterized by different sets of enzymes (glycosidases and glycosyltransferases, Tables 4.4.2-1 and 4.4.2-2). Proteins are transported between the compartments by vesicles (4.5.2).

The carbohydrate chains shown in Figures 4.4.2-1 and 4.4.2-2 are a selection of basic structures. They occur in glycoproteins and proteoglycans (heavily glycosylated glycoproteins, 2.9.1) as well as in glycolipids (3.4.4). In glycoproteins, they are frequently further modified, often depending on the individual (e.g., blood group types, 4.4.3), the tissue type or the metabolic situation. In proteoglycans, the structures depicted form the linkage region between the protein and the extended, outer glycan chains (see Figs. 2.9-1, 2.9-2 and 4.4.2-2).

4.4.2.1 Synthesis of Glycoproteins (Fig. 4.4.2-1, Table 4.4.2-1)

The final product of protein glycosylation is determined by cooperation of different, specific enzymes. All sugar moieties in the Golgi are added stepwise directly from XDP-sugars or CMP-NeuNAc. Specific carriers exist in the Golgi membrane for all required XDP-sugars, which transport them from the cytosol into the Golgi lumen. Contrary to the ER, no oligosaccharide precursors are synthesized at anchors, such as dolichyl-phosphate.

N-Glycans: N-glycosylated proteins imported from the ER have a 'core' glycostructure (4.4.1.3), which is initially trimmed in the *cis*-Golgi compartment by action of glycosidases. The amount of this shortening determines the further fate of the compound. During the following journey through the *medial* and *trans* compartments, the various complex, hybrid and high mannose structures are formed (Pathways A … D in Fig. 4.4.2-1).

This process affects the microheterogeneity of carbohydrate structures, since only part of all possible reactions occur at an individual protein. This is (at least partially) caused by the particular protein structure. In addition, the degree of glycosylation depends on the transit time in the compartments, which is influenced by the translation rate and the secretion efficiency of the cell. The precise function of the individual structures is not always known.

In plants, the proximal mannose sugar is typically xylosylated after fucosylation.

O-Glycans: O-glycosylated proteins are a class of glycans with modified serine or threonine residues. Biosynthesis of O-glycans starts from the transfer of N-acetylgalactosamine (GalNAc) to serine or threonine. The resulting but intermediate GalNAc-peptide is called Tn-antigen. This is the precursor to T-antigen, which is formed after an additional galactosylation step. It has been estimated that in ca. 90% of all cancers, Tn- and T-antigens are expressed and visible to the immune system. Healthy cells do not accumulate Tn- or T-antigens but further elongate them. Depending on the sugars added, there are four common O-glycan core structures (I to IV). Less common core structures (not shown in Fig. 4.4.2-1) originating from Tn-antigen are cores V (GalNAcα1–3GalNAcα-Ser/Thr), Vi (GlcNAcβ1-6GalNAcα-Ser/Thr), VII (GalNAcα1-6GalNAcα-Ser/Thr), and VIII (Galα1–3GalNAcα-Ser/Thr).

In mammals, synthesis of O-glycans starts in the Golgi apparatus (Pathway E in Fig. 4.4.2-1). Here, the first sugar (usually N-acetylgalactosamine) is added to serine or threonine, in collagen to hydroxylysine. In yeast and green plants, however, the first sugar (mannose) already becomes attached in the ER and the chain is continued in the Golgi by adding up to three more mannose residues (Pathway F in Fig. 4.4.2-1).

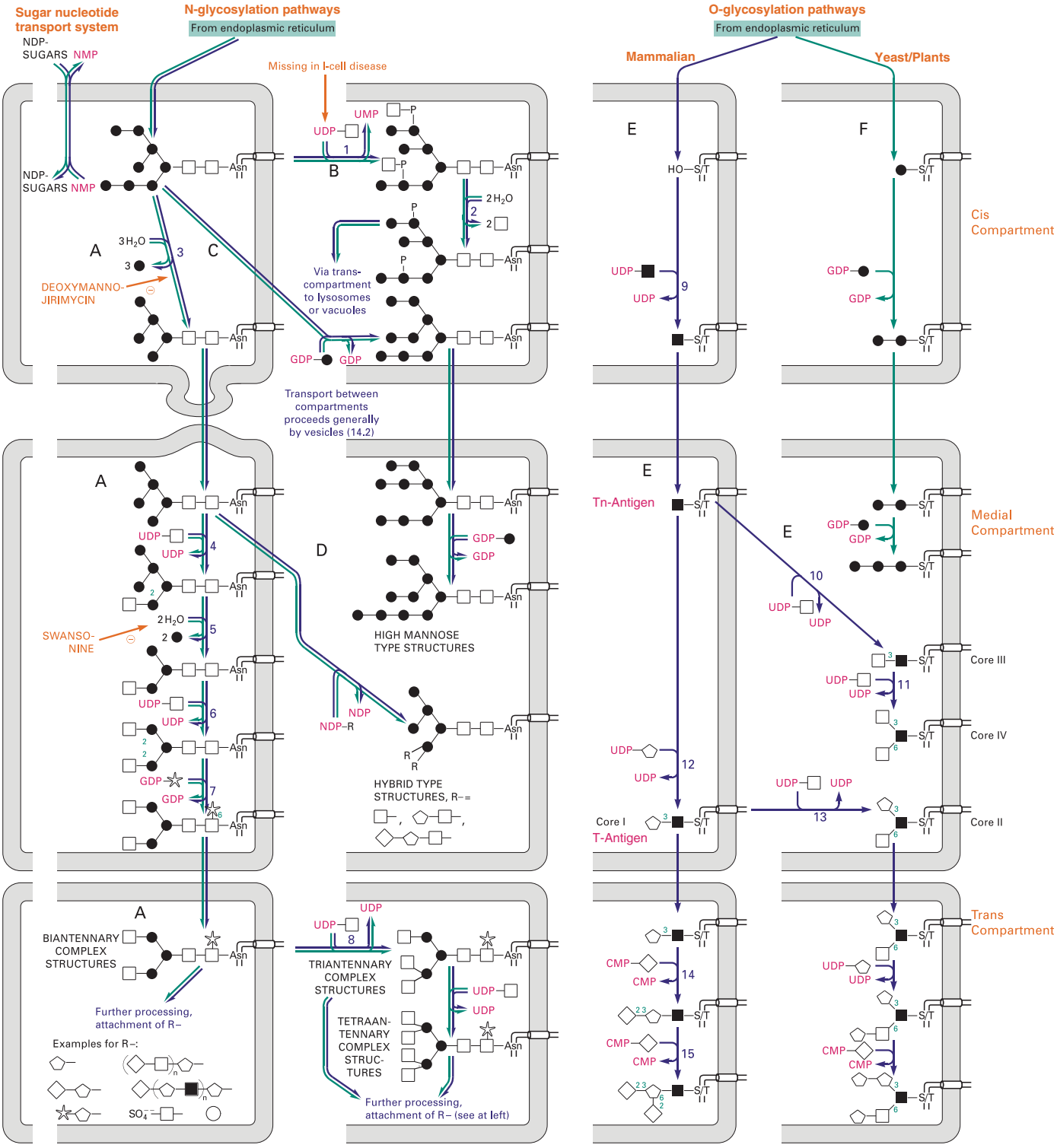
4.4.2.2 Synthesis of Proteoglycans

Proteoglycans are important components of the extracellular environment and function, e.g., as structural components, growth factor coreceptors, anticoagulants and clustering agents for neurotransmitter hydrolases. The short, inner branches of proteoglycans (see Fig. 2.9-2) have the structure of N- or O-glycoproteins and are synthesized the same way as described above. The synthesis of the repeating units of the long, outer glycosaminoglycan branches of proteoglycans also takes place in the Golgi apparatus, but the exact localization of the reactions is not known so far. This is also valid for the sulfatation reactions, which are not shown in Figure 4.4.2-1.

Proteoglycan synthesis often involves xylosylation catalyzed by protein xylosyltransferases (EC 2.4.2.26) that transfer formation in plants, β-1,2-xylosyltransferases (EC 2.4.2.38) which create immunogenic epitopes on N-glycans of all plants and some invertebrates, and an O-xylosyltransferase (EC 2.4.2.40) that modifies the plant cytokinin zeatin.

Table 4.4.2-1. Enzymes in Glycoprotein Synthesis (Accepted names and EC numbers in parentheses)

1. UDP-N-Acetylglucosamine-lysosomal enzyme N-acetylglucosamine-phosphotransferase (2.7.8.17)
2. N-acetylglucosamine-1-phosphodiester α-N-acetylglucosaminidase (phosphodiester glycosidase, 3.1.4.45)
3. Mannosidase I (mannosyl-oligosaccharide 1,2-α-mannosidase, 3.2.1.113)
4. N-acetylglucosaminyltransferase I (mannosyl-glycoprotein β-1,2-N-acetylglucosaminyltransferase, 2.4.1.101)
5. Mannosidase II (mannosyl-oligosaccharide 1,3-1,6-α-mannosidase, 3.2.1.114)
6. N-Acetylglucosamine transferase II (α-1,6-mannosyl-glycoprotein β-1,2-N-acetylglucosaminyltransferasae, 2.4.1.143)
7. Glycoprotein 6-α-L-fucosyltransferase (2.4.1.68)
8. N-Acetylglucosamine transferase III (β-1,4-mannosyl-glycoprotein β-1,4-N-acetylglucosaminyltransferase, 2.4.1.144)
9. Polypeptide N-acetylgalactosaminyl transferase (2.4.1.41)
10. Acetylgalactosaminyl-O-glycosyl-glycoprotein β-1,3-N-acetylglucosaminyltransferase (2.4.1.147)
11. Acetylgalactosaminyl-O-glycosyl-glycoprotein β-1,6-N-acetylglucosaminyltransferase (2.4.1.148)
12. Glycoprotein-N-acetylgalactosamine 3-β-galactosyltransferase (2.4.1.122)
13. β-1,3-Galactosyl-O-glycosyl-glycoprotein β-1,6-N-acetylglucosaminyltransferase (2.4.1.102)
14. β-Galactoside α-2,3-sialyltransferase (2.4.99.3)
15. α-N-acetylgalactosaminide α-2,6-sialyltransferase (2.4.99.3)



Symbols

- △ Glc glucose
- □ GlcNAc N-acetylglucosamine
- ◇ Gal galactose
- ■ GalNAc N-acetylgalactosamine
- ○ Xyl xylose
- ● Man mannose
- ◇ NeuNAc N-acetylneuram. acid
- ☆ Fuc fucose
- transmembrane domains

Figure 4.4.2-1. Synthesis of Glycoproteins in the Golgi Apparatus

Pathways of Protein Glycosylation (Letters A–F correspond to different pathways):
A) N-glycosylation pathway leading to complex structures. Additional glycoside units (besides Man and GalNAc) are added, leading to final structures shown (R-). They are also present in proteoglycans.
B) Pathway for glycoproteins destined for lysosomes (general) or vacuoles (yeast). Mannose units become phosphorylated and are able to react with the mannose 5-P receptor at the Golgi membrane.
C) Pathway to high-mannose type carbohydrate structures.
D) Pathway to hybrid structures.
E) O-glycosylation pathway in mammalian cells. These structures are also present in proteoglycans.
F) O-glycosylation pathway in yeast and plants.

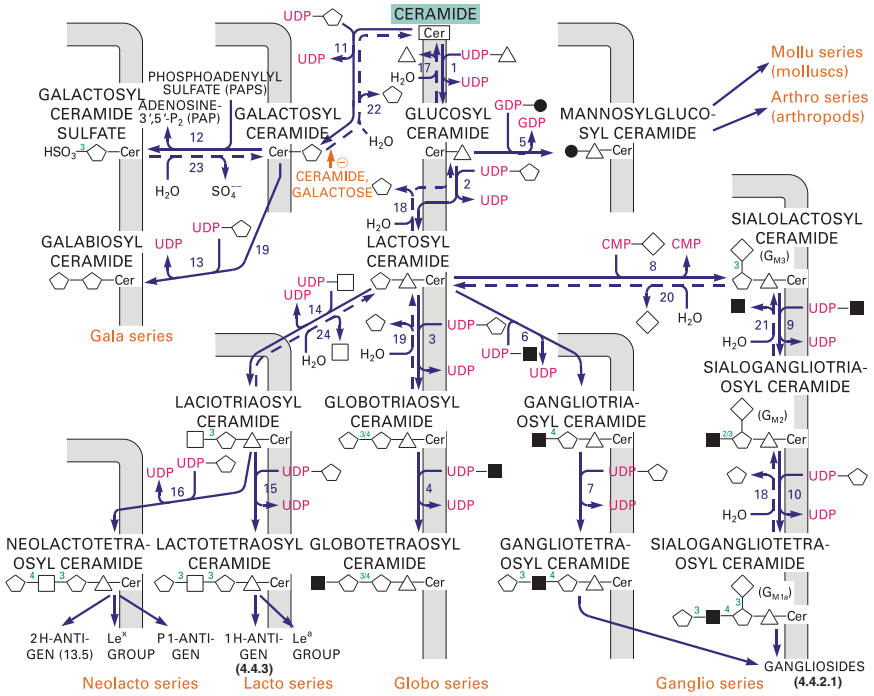


Figure 4.4.2-2. Synthesis of Glycosphingolipids in the Golgi Apparatus
(For symbols see Fig. 4.4.2-1)

Table 4.4.2-2. Enzymes in Glycosphingolipid Synthesis (EC-numbers are given where possible)

1. Ceramide glucosyltransferase (2.4.1.80)
2. UDP-galactose:glucosylceramide β-1,4-galactosyltransferase
3. Lactosylceramide β-1,3 (or 4) galactosyl-transferases (2.4.1.179)
4. Hexosyltransferases (2.4.1.-)
5. Hexosyltransferases (2.4.1.-)
6. Hexosyltransferases (2.4.1.-)
7. Hexosyltransferases (2.4.1.-)
8. Lactosylceramide α-2,3-sialyltransferase (2.4.99.9)
9. (N-acetylneuraminyl)-galactosylglucosylceramide N-acety-galactosaminyltransferase (2.4.1.92 or 165)
10. Ganglioside galactosyltransferase (2.4.1.62)
11. Ceramide galactosyltransferase (2.4.1.47)
12. Galactosylceramide sulfotransferase (2.8.2.11)
13. Hexosyltransferases (2.4.1.-)
14. Lactosylceramide β-1,3-N-acetyl-β-d-glucosaminyltransferase (2.4.1.206)
15. Glucosaminylgalactosylglucosylceramide β-galactosyltransferase (2.4.1.86)
16. β-1,4-galactosyltransferase (2.4.1.-)
17. Glucosylceramidase (3.2.1.45)
18. Ganglioside β-galactosidase (3.2.1.23)
19. α-Galactosidase (3.2.1.22 or 47)
20. Exo-α-sialidase (neuraminidase) (3.2.1.18)
21. β-N-Acetylhexosaminidase A (3.2.1.52)
22. Galactosylceramidase (lysosomal) (3.2.1.46)
23. Cerebroside sulfatase (arylsulfatase A) (3.1.6.8 or 1)

4.4.2.3. Synthesis of Glycolipids (Fig. 4.4.2-2, Table 4.4.2-2)

Glycolipids contain a lipid anchor (e.g., ceramide or acylglycerol, Fig. 3.4.4-1) in the membrane with an attached glycoside chain. The addition of the first sugar residue is supposed to take place at the cytosolic side of the Golgi membrane. The monoglycosylated ceramide is then transferred to the luminal side, where all the following reactions occur. Up to G_{M3} the glycosylations take place in the *cis* compartment of the Golgi, thereafter in the *medial* and *trans* compartments. Finally the glycolipids are transferred to the cellular surface, where they arrange themselves in clusters. Only the initial reactions, forming the 'root structures' are shown in Figure 4.4.2-2.

Literature:

Gamblin, D.P., *et al.* Chem. Rev. 2009;109:131–163.

Goldberg, D.E., Kornfeld, S.J. Biol. Chem. 1983;258:3160–3167.

Gomord, V., Faye, L. Curr. Opin. Plant Biol. 2004;7:171–181.

Hirschberg, C.B., Snider, M.D. Ann. Rev. Biochem. 1987;56: 63–87.

Montreuil, J., Vliegenthart, J.F.G., Schachter, H. *Glycoproteins. New Comprehensive Biochemistry* Vol. 29a. Elsevier, 1995.

Vance, D.E., Vance, J. (Eds.) *Biochemistry of Lipids, Lipoproteins and Membranes. New Comprehensive Biochemistry* Vol. 31. Elsevier, 1996.

Wilson, I.B.H. Cell. Mol. Life Sci. 2004;61:794–809.

4.4.3 Terminal Carbohydrate Structures of Glycoconjugates

The surface of eukaryotic cells carries many glycoproteins and glycolipids with characteristic arrangements of carbohydrate chains. Such arrangements can also be found on secreted proteins. These carbohydrate structures can be antigenic, the respective determinants are usually located in the terminal glycosidic groups. Some of them play a role in cell-cell recognition and interaction (e.g., sialylated Le^x groups as part of ligands for selectins, 8.4). Antigenic properties of cell surfaces can, however, also originate in protein structures.

4.4.3.1 Blood Groups (Fig. 4.4.3-1)

Blood group antigens are either sugars (e.g., ABO-system) or proteins (e.g., rhesus-system). They are attached to various components in the red blood cell membrane. A number of antibodies in human/animal serum are directed against antigens at the erythrocyte surface of other members of the same species (but also against similar antigens in other species, e.g., animals, microorganisms, viruses). The antigen-antibody reaction causes blood agglutination and cell lysis. This happens when foreign antigens are introduced into the circulation (e.g., by incompatible blood transfusion). The presence or absence of this reaction in individuals of the same species allows classification of blood groups.

The human blood group antigens are located in the layer of the extracellular matrix (2.2.4) not only at the surface of erythrocytes, but also of most other cells of the body. In about 80% of the population, they occur also in soluble form in the body fluids. The antigens at the surface of erythrocytes have been classified in (so far) 30 blood group systems. The ABO, Hh, Ii, P and Sid systems, e.g., are based on carbohydrates.

The cell-bound antigens of the ABO and other systems are glycoproteins and glycosphingolipids, whereas the soluble ones are only glycoproteins. The terminal antigenic determinants are carried by short chains (e.g., R = Cer-Glc-Gal, see 4.4.2.3) or by long sequences of (frequently repeated) glycosidic groups, which are bound via a carbohydrate core to a protein (e.g., anion transporter or glucose transporter of erythrocytes). The Lewis antigens (Le) are glycosphingolipids, which are not expressed at erythrocytes, but are acquired from plasma lipoproteins (HDL and LDL).

The pedigree of human blood group antigens is given in Figure 4.4.3-1. Starting from a basic structure (R-GlcNAc-Gal), fucose is either bound to the galactose moiety, leading to the ABO blood group substances and after a second fucosylation to the Lewis systems Le^b and Le^y, or to the N-acetylgalactosamine moiety, leading

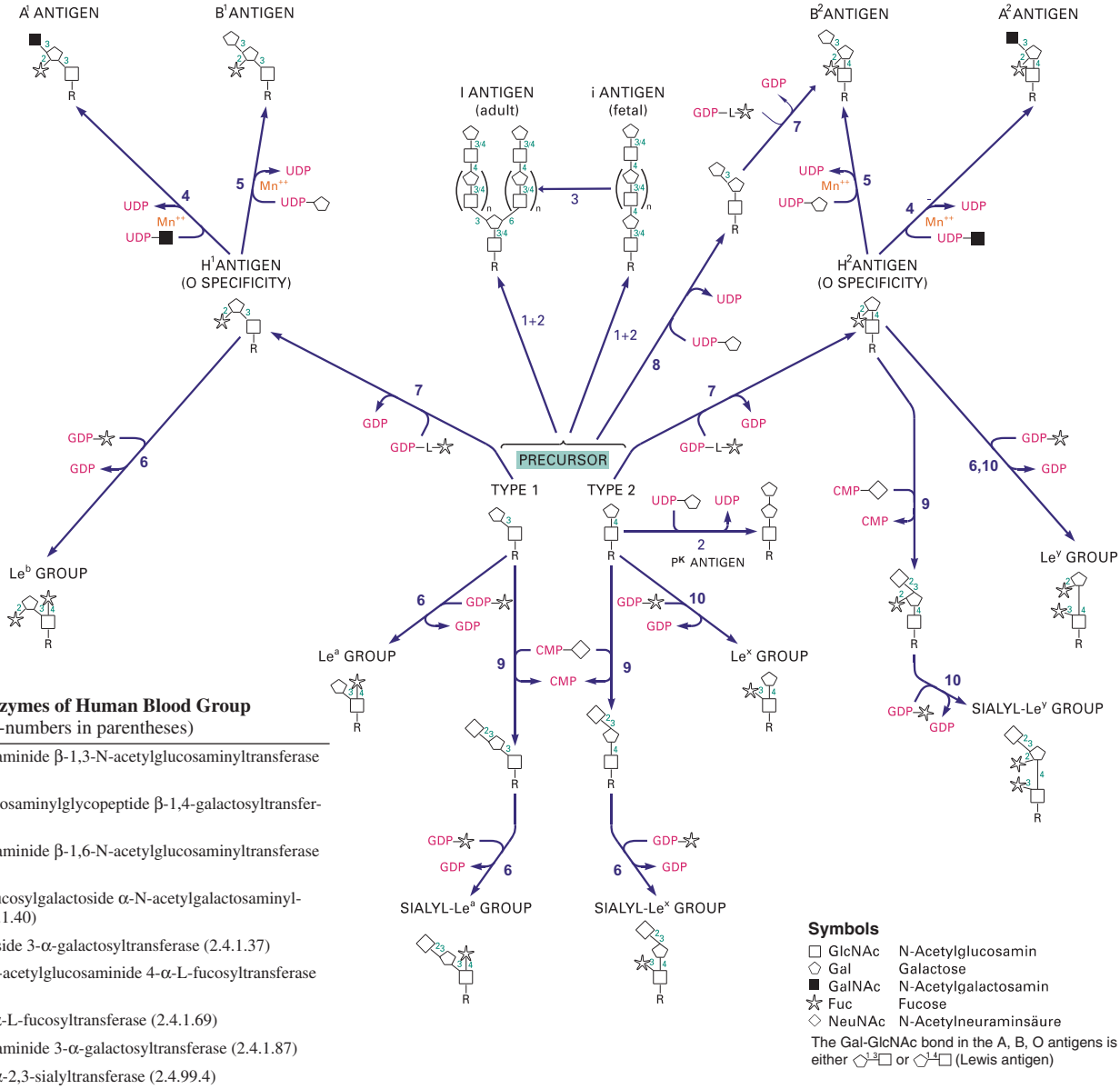


Figure 4.4.3-1. Terminal Carbohydrate Groups of Human Blood Groups

Table 4.4.3-1. Enzymes of Human Blood Group Biosynthesis (EC-numbers in parentheses)

1. N-Acetyllactosaminide β-1,3-N-acetylglucosaminyltransferase (2.4.1.149)
2. β-N-Acetylglucosaminylglycopeptide β-1,4-galactosyltransferase (2.4.1.38)
3. N-Acetyllactosaminide β-1,6-N-acetylglucosaminyltransferase (2.4.1.150)
4. Glycoprotein-fucosylgalactoside α-N-acetylgalactosaminyltransferase (2.4.1.40)
5. Fucosylgalactoside 3-α-galactosyltransferase (2.4.1.37)
6. 3-Galactosyl-N-acetylglucosaminide 4-α-L-fucosyltransferase (2.4.1.65)
7. Galactoside 2-α-L-fucosyltransferase (2.4.1.69)
8. N-Acetyllactosaminide 3-α-galactosyltransferase (2.4.1.87)
9. β-Galactoside α-2,3-sialyltransferase (2.4.99.4)
10. 4-Galactosyl-N-acetylglucosaminide 3-α-L-fucosyltransferase (2.4.1.152)

to the Lewis systems Le^a and Le^x. The human genome encodes at least 13 different α-1,3-fucosyltransferases, therefore the pathways of biosynthesis are not identical in different cell types. Depending on the genetic variant of the enzymes, A1/B1 or A2/B2 antigens are synthesized. Lewis structures are frequently masked by terminal sialylation. Other additions to the basic structure lead to the I/i, P^k/P/P_1 and other blood groups.

Literature:

Daniels, G. ISBT Sci. Ser. 2009;4:368–374.

Feizi, T. Nature (London) 1985;314:53–57.

Hakomori, S. Adv. Cancer Res. 1989;52:257–332.

Martinez-Duncker, I. *et. al.* Glycobiol. 2003;13:1C–5C.

Monteuil, J., Vliegenthart, J.F.G., Schachter, H. *Glycoproteins. New Comprehensive Biochemistry* Vol. 29a. Elsevier, 1995.

Muramatsu, T. Glycobiology 1993;3:291–296.

4.5 Protein Folding, Transport / Targeting and Degradation

Petra Dersch

4.5.1 Folding of Proteins

Although the tertiary structure of proteins (2.3.1) is determined by their amino acid sequence, not all proteins fold spontaneously into their final three–dimensional structure. Proteins also undergo stress-induced denaturation, and must be refolded to restore their function.

Polypeptides emerge from ribosomes (4.1.3.2, 4.2.3, 4.4.1) in an unfolded state, but they usually fold up rapidly and efficiently into a unique, compact and well-defined structure necessary to fulfill their function. Formation of secondary structures (α helices and β sheets, 2.3.1) is a spontaneous process that occurs very early in the folding process and leads to partially 'disordered' folding intermediates, the so-called 'molten globules'. They have extensive native-like

secondary structures, but are devoid of a regular tertiary structure. Some hydrophobic amino acids which are usually buried in the core of correctly folded soluble proteins are still exposed to the solvent in these globules, and this may lead to inter- and intramolecular aggregation, which sometimes inhibits further folding. Following further reorganisation they adopt a highly ordered well-folded state. Specialized proteins (chaperones) in bacteria and in eukarya assist in this folding process, in particular when larger proteins are newly synthesized. Apart from this function, they also maintain pre-existing proteins in a stable conformation, preventing their aggregation, under stress conditions or oppose stress-induced denaturation.

4.5.1.1 Protein Folding in Bacteria (Table 4.5.1-1, Fig. 4.5.1-1)

In bacteria, one group of molecular chaperones prevents undesired aggregation by binding and stabilizing unfolded polypeptide regions (DnaK/J, GrpE). The other group (chaperonins) provides an environment for the folding intermediate that allows a controlled and enhanced folding process (GroEL/ES complex).

Tabe 4.5.1-1. Molecular Chaperones and Enzymes Involved in Protein Folding in *E. coli* and Homologues in Eukarya

E. coli Components	Properties and Function	Homologue in Eukarya
Trigger factor	Ribosome-associated factor, facilitates protein folding of nascent protein chains, regulates access of SRP, keeps polypeptides in a translocation-competent state	Only present in bacteria and chloroplasts. Ribosome-associated systems Ssb/Ssz/ Zuo-triad (RAC) and NAC are specific for eukarya
DnaJ (41 kDa)	Binds to polypeptide, mediates DnaK binding through the J-domain, induces ATPase action of DnaK, resolves protein aggregates	Hsp40. – In ribosomes Zuo (MPP11). – In mitochondria Mdj1 (4.5.4)
DnaK (ca. 65 kDa)	Stabilizes nascent, unfolded or partially folded polypeptides, binds and hydrolyzes ATP, resolves protein aggregates, controls heat-shock response, required for ribosome biogenesis	Hsp70 [Hsc70, BiP, in ribosomes Ssb, Ssz (Hsp70-L1)]. – In mitochondria mtHsp70 (4.5.4)
GrpE (20 kDa)	ADP/ATP exchange factor, effects release of the polypeptide from the DnaK/J complex, resolves protein aggregates	– In mitochondria mtGrpE (4.5.4)
GroEL (14.57 kDa)	2 Homo-heptameric barrel-shaped stacked rings with a hydrophobic inner surface, binds GroES and ATP, catalyzes folding of cytoplasmic proteins, ATPase, facilitates acquisition of genetic variations	TCP-1 ring complex = TRIC (ca 10-fold hetero-oligomer). – In mitochondria mtHsp60
GroES (7.10 kDa)	Ring of 7 subunits capping one end of the GroEL 'barrel', encapsulation of the unfolded protein into GroEL cavity and acceleration of the folding process, its dissociation releases proteins from GroEL	– In mitochondria Hsp10
Hsp90	Flexible dimer, ATPase, more selective, cooperation with Hsp70/Hsp40, maturation of signalling kinases	Hsp90 – In ER Grp94
–	Co-chaperone of Hsp90, adaptor to link Hsp70 to Hsp90	Hop/StiI
ClpB (93 kDa)	Disaggregation and reactivation of insoluble protein aggregates, AAA+-superfamily, cooperates with DnaK	–In yeast Hsp104
Dsb proteins (A.E)	Formation and isomerization of –S–S– bonds, cytochrome c maturation	Protein disulfide isomerase (PDI, in ER)
Thioredoxin system TrxA,B,C (e.g. 8.1.4)	Reversible reduction of disulfide bonds (oxidative stress, DNA damage, iron metabolism)	Thioredoxin system TRR1, TRX1-2 (e.g., 3.6.1.4 and 3.12.2)
Glutathione system (Gor, Grx1-3, GshA,B)	Reversible reduction of disulfide bonds (oxidative stress, DNA damage, iron metabolism)	In yeast: Glutathione system (Glr1, Grx1-5, GSH1-2, GRL1)
RotA, SurA, FkpA (periplasmic)	Isomerization of proline. In eukarya, also an essential sub-unit of prolyl-4-hydroxylase	Peptidyl-proline *cis-trans* isomerase (in ER)

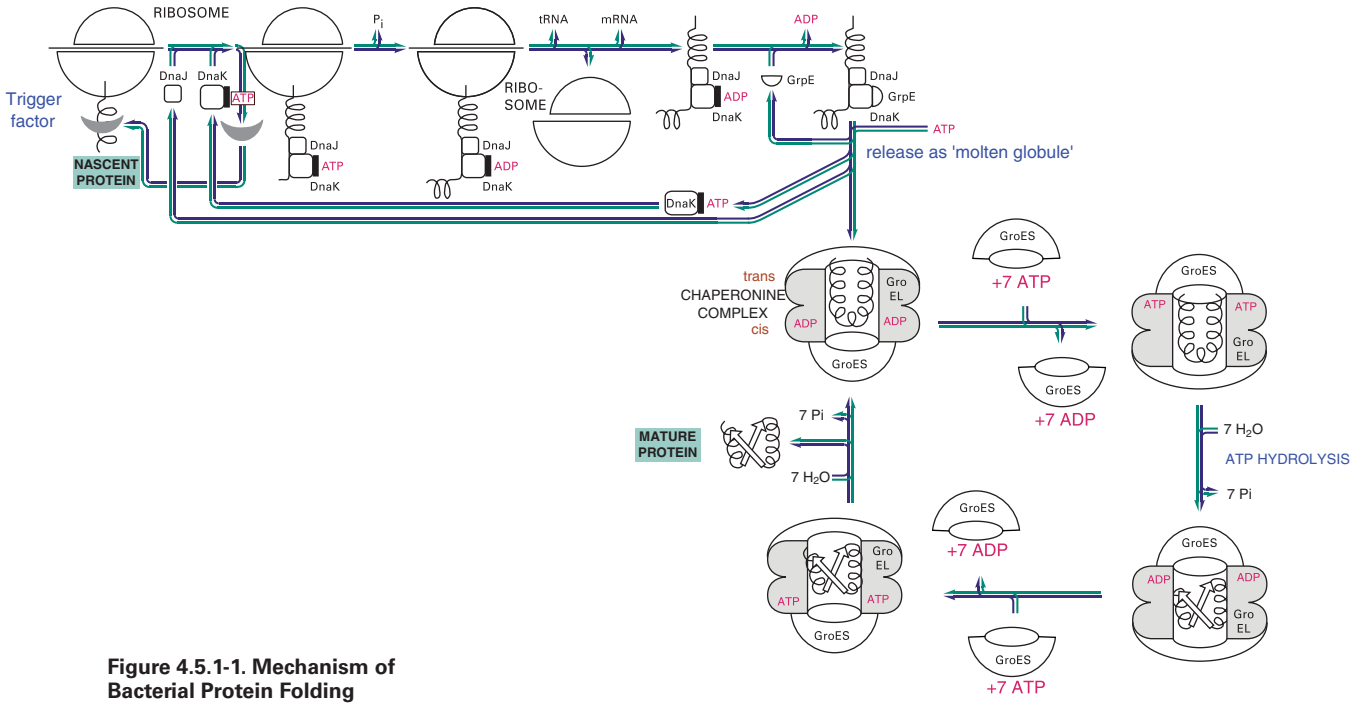


Figure 4.5.1-1. Mechanism of Bacterial Protein Folding

Both types of chaperones belong to the group of heat-shock proteins (Hsp), since they also respond to cellular stress situations. They do not only prevent aggregation of polypeptides, they also resolve aggregated polypeptides and protein complexes, which occur e.g., by thermal denaturation. DnaKJ/GrpE also repress excessive gene transcription of the heat shock proteins by binding the heat shock sigma factor σ^{32} (4.1.1.1).

The first protein to interact with nascent polypeptides is the ribosome-associated chaperone trigger factor (TF). TF arches over the exit of the ribosomal tunnel and forms a cradle that shields newly synthesized proteins and allows smaller proteins or protein domains to adopt their native structure. Nascent polypeptides from periplasmic and membrane proteins are then recognized and targeted to the membrane translocon by the signal recognition particle (SRP) or Sec-dependent pathways. Nascent polypeptides from larger cytosolic proteins, evolving from the ribosome, promote dissociation of TF and are subsequently bound by the co-chaperone DnaJ. This targets ATP-bound DnaK to the protein and accelerates ATP hydrolysis to enhance DnaK function. ATP hydrolysis stabilizes the polypeptide chain–DnaK•ADP–DnaJ complex. This prevents the polypeptide from aggregation with other unfolded polypeptides. The polypeptide is released as a molten globule after GrpE has resolved the complex by facilitating the ADP/ATP exchange. If the polypeptide still contains unfolded regions or has adopted an unfavorable structure, the cycle is repeated (Fig. 4.5.1-1).

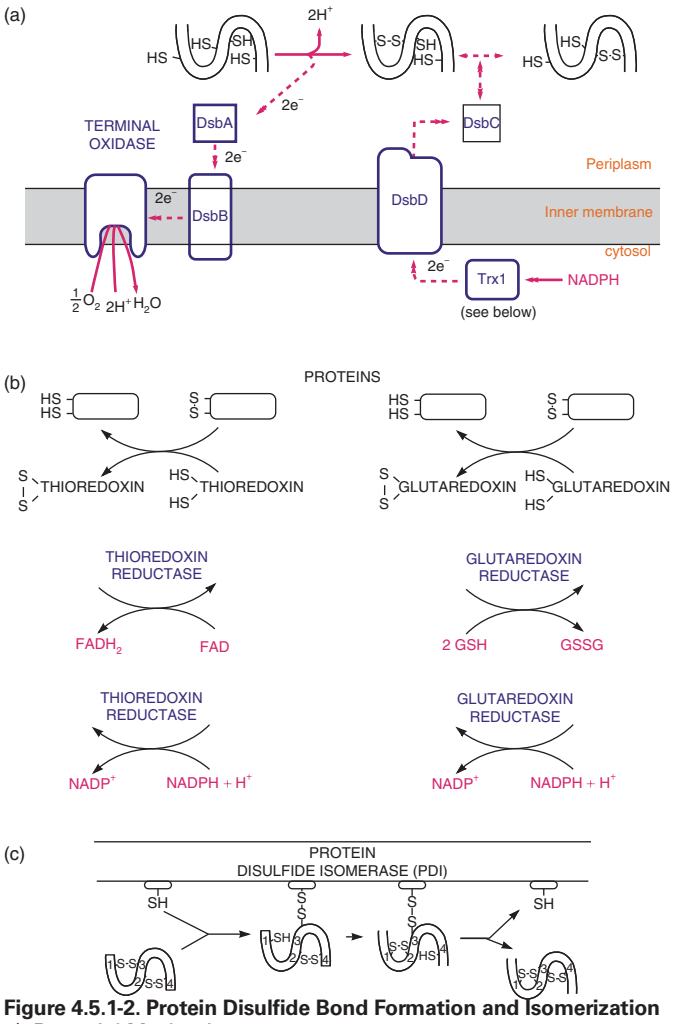


Figure 4.5.1-2. Protein Disulfide Bond Formation and Isomerization
a) Bacterial Mechanism
b) Thioredoxin and Glutaredoxin Systems
c) Protein Disulfide-isomerase

For many proteins, the consecutive folding of the molten globule into its final tertiary native structure might take place inside of the GroEL/ES chaperonin complex. This folding machine is a double-heptamer ring arranged back-to-back with a large central cavity in which the folding process takes place. Denatured proteins bind to hydrophobic residues exposed at the 'free' *trans* end of a GroEL molecule. Binding of ATP and the co-chaperonin GroES to this side of the molecule triggers the release and encapsulation of the peptide into the folding cavity (*cis*), where peptide chains fold by a sequence of binding-and-release steps until hydrophobic residues that cause aggregation are buried within the final folded structure. The substrate protein itself seems to influence the chaperonin structure and the reaction cycle. ATP hydrolysis triggers binding of new ATP molecules to the opposite *trans* ring and release of GroEL and the folded protein from the *cis* ring (Fig. 4.5.1-1).

Apart from the DnaK/GroEL family, a large number of distinct and specialized chaperones are produced which prevent premature folding of special proteins or protein groups for certain purposes, e.g., transport through inner and outer membrane, pili and flagella structures, or secretion systems. Furthermore, ClpB, an ATPase of the AAA+-superfamily with a hexameric structure is involved in disaggregation and reactivation of aggregated insoluble proteins. Together with DnaK it extracts polypeptides from the aggregates by forced unfolding and translocation through the central cavity of ClpB. After release, the polypeptide can then refold spontaneously or with help of other chaperones (not shown).

Disulfide bond formation (Fig. 4.5.1-2) helps to stabilize the proper folding of many secreted proteins. Bacteria possess special enzymes in the periplasm and/or membrane that act as protein folding catalysts (DsbAB) ensuring correct pairing of cysteine residues for disulfide-bond formation. DsbA is a thiol oxidase with a CXXC motif acting as a very active oxidant transferring its disulfide bond to translocated proteins. Reoxidation is accomplished by the inner membrane partner enzyme DsbB. As DsbA is not a specific oxidant, incorrect disulfide bonds in proteins with more than two cysteines have to be corrected by isomerization through the protein thiol-disulfide oxidoreductase DsbC. This protein is kept in the reduced state by the membrane protein DsbD, which itself is reduced by cytoplasmic thioredoxin (Fig. 4.5.1-2a). Many bacteria use these pathways of disulfide bond formation to produce virulence factors (adhesins, secreted toxins and enzymes, type III secretion systems, capsule biogenesis).

The thiol system consists of thioredoxin and glutathione pathways (Figre 4.5.1-2b). It uses cysteine residues to catalyze a thiol-disulfide exchange reaction to control the redox state of cysteines in cytoplasmic proteins. The components of the thioredoxin system (thioredoxins TrxA, TrxC, thioredoxin reductase TrxB) and the glutathione system (glutaredoxins Grx1(A), Grx2(B), Grx3(C), glutathione reductase Gor, γ-glutamylcysteine GshA and glutatione synthetases GshB) are kept reduced at the expense of NADPH oxidation. Analogous systems exist also in eukarya.

4.5.1.2 Protein Folding in the Eukaryotic Cytosol (Table 4.5.1-1)

Processing and folding of cytosolic proteins of eukarya is very similar to bacteria, often employing the respective chaperone homologues (which are highly conserved in all organisms). The evolving protein chain immediately binds chaperones of the Hsp40 (DnaJ homologue) and Hsp70 (DnaK homologue) family. The ribosome-associated chaperone complexes Ssb/Ssz-Zuo (RAC, a Hsp70/Hsp70/Hsp40 triad) and NAC (a heterodimer composed of an α- and a β-subunit) are the first to encounter the nascent chain. Many proteins need additional support by cytosolic chaperones of the Hsp70 (Ssa) family and enter the Hsp60-type TRIC/CCT structure (GroEL/ES homologue) for efficient folding. Several proteins require specific chaperones for their proper folding process, e.g., GimC/prefoldin is required to fold actin and tubulin. Disaggregation and reactivation of insoluble protein aggregates are mediated by Hsp104 (ClpB homologue). Similar mechanisms take place in folding of proteins imported into mitochondria and chloroplasts (4.5.4, 4.5.5).

4.5.1.3 Protein Folding in the Eukaryotic Endoplasmic Reticulum (ER, Table 4.5.1-2, Fig. 4.5.1-3)

Signal sequences of secreted proteins emerging from a ribosome are bound by the signal recognition particle (SRP), a complex of 7S RNA and six proteins. This slows down translation until the ribosome interacts with a proteinaceous channel formed by two Sec61αβγ heterotrimeric complexes located at the ER membrane. Then protein synthesis proceeds, the signal sequence is removed and the nascent chains are co-translationally injected into the ER lumen (Fig. 4.4.1-1). The ER lumen contains members of the Hsp70, Hsp40, Hsp90 and Hsp100 family. BiP/Grp78 is ubiquitously expressed in mammalian cells and is the master chaperone of the ER. It facilitates translocation, protein folding and oligomerization and prepares the dislocation of misfolded proteins from the ER to the cytoplasm. Several Hsp40-like ERdj proteins and GrpE-like proteins (BAP/Sil1, GRP170) assist protein folding by BiP/Grp78. GRP94 seems to associate with more advanced folding intermediates after release from BiP/Grp78.

The ER differs from the cytosol by providing an oxidizing environment (resulting in protein-S-S-bonds). Several proteins are subjected to special co- and posttranslational modifications, such as the formation of intra- and interchain disulfide bonds, addition of preassembled oligosaccharides (glycosylation by glycosyltransferases) and lipid anchors.

Accessory oxidoreductases from the protein disulfide isomerase family are involved in disulfide bond formation by thiol oxidation and in isomerization by reshuffling of disulfide bonds (Fig. 4.5.1-2c). Disulfide isomerase can also inhibit aggregation caused by intermolecular disulfide bonds. Peptidyl-prolyl *cis-trans*-isomerases (PPI) further catalyze prolyl *cis/trans* isomerizations.

The majority of proteins that use the secretory pathway receive multiple N-linked glycans (4.4.1.3). Preassembled glycans ($Glc_3Man_9GlcNAc_2$) are transferred by oligosaccharyltransferases from a lipid pyrophosphate donor in the ER membrane to nascent protein chains at Asn of the Asn-X-Ser/Thr motif. The glycan is then trimmed by the sequential action of glycosidases I and II to $Glc_1Man_9GlcNAc_2$ on the glycoprotein precursor (Fig. 4.4.1-1). The chaperone-like transmembrane protein calnexin (p88, IP90) and the soluble calreticulin then bind specifically to the remaining glucose residue to assist glycoproteins with proximal or distal glycans. Other chaperones (BiP/Grp78 and Grp94) cooperate in the folding mechanism. Accessory oxidoreductase ERp57, involved in disulfide bond formation, is also recruited. Upon completion of the folding process, the single remaining glucose is removed by glucosidase II, and the proteins are transported to the Golgi apparatus. However, incorrectly folded proteins undergo a reglucosylation cycle, catalyzed by a specific UDP-glucose: glycoprotein glucosyltransferase (GT), which keeps them in the ER until correct folding is achieved.

Table 4.5.1-2. Chaperones Involved in Protein Folding in the ER

Chaperones	Family	Function
Calnexin (p88: *mice*, IP 90: humans)	(Hsp90)	Membrane phosphoprotein, binds glycoproteins, binds Ca^{++}
Calreticulin	(Hsp90)	glycoprotein-dedicated chaperone
EDEM1, EDEM2, EDEM3	(Hsp90)	
ERp57, ERp72, PDI, PDIp, PDILT, P5	PDI-family	Oxidase/Isomerase, disulfide bond formation and isomerization
Cyclophilin B/CypB		Part of large ER multi-chaperone complexes
FKBP2/FKBP13		ER-stress-induced
FKBP7/FKBP23		Modulates BiP ATPase
FKBP10/FKBP65		Associates with BiP bound substrates
BiP/ Grp78	Hsp70	Soluble, ATPase, binds and retains proteins in the ER (DnaK homologue), involved in posttranslat. import
ERdj1/Mtj1 EEdj2/hSec63 ERdj3/HED/ERj3/ABBP-2 ERdj4/Mdj1 ERdj5/JPN1	Hsp40	BiP cofactors
BAP/Sil1, GRP170	GrpE-family	
Grp94	Hsp90	Endoplasmin, soluble, ATPase, binds proteins, dimeric, helps maturation of immunoglobulin heavy chains, integrins, TLRs
TorsinA	Hsp100	

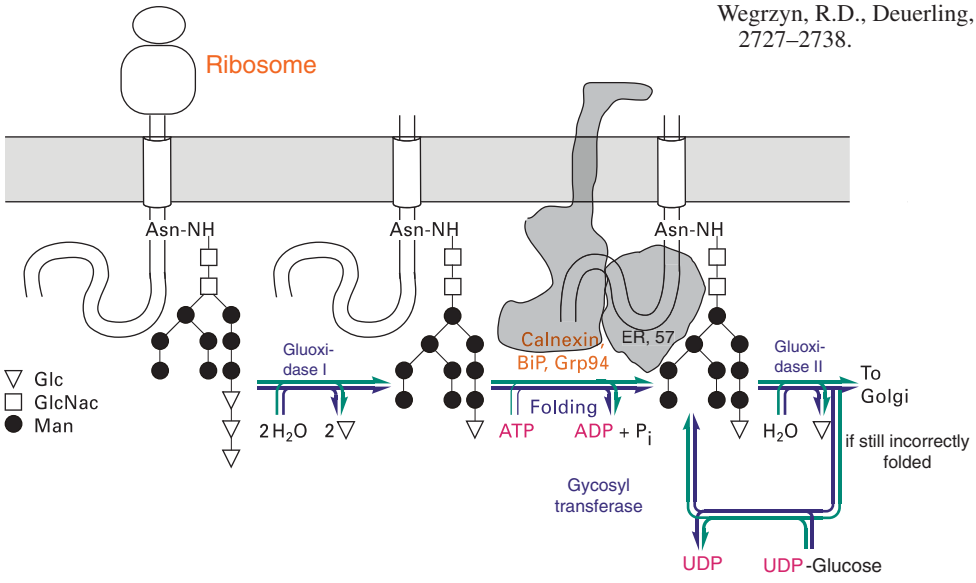


Figure 4.5.1-3. Protein Folding in the ER (compare Fig. 4.4.1-1)

Literature:

Clare, D.K. *et al*. Nature 2009;457:107–110.

Clarke, A.R. *et al.* Mol. Cell. 2008;24:165–171.

Dalle-Dome, I. *et al.* Trends Biochem. Sci. 2009;34:85–96.

Doyle, S.M., Wickner, S. Trends in Biochem. Sci. 2009;34:40–48.

Hebert, D.N., Molinari, M. Physiol. Rev. 2007;87:1377–1408.

Heras, B. *et al.* Nat. Rev. Microbiol 2009;7:215–225.

Inaba, K. Ito, K. Biochim. Biophys. Acta Mol. Cell Research 2008;1783:520–529.

Li, Y. *et al.* J. Biol. Chem 2009;284:4324–4331.

Liberek K. *et al.* EMBO J. 2008;27:328–335.

Lund, P.A. FEMS Microbiol. Rev. 2009;33:785–800.

Shaner, L.; Morano, K.A. Stress and Chaperones 2007;12:1–8.

Striebel, F. *et al.* Curr. Opin. Struct. Biol. 2009;19:209–217.

Tokuriki, N., Tawfil, D.S. Nature 2009;459:668–673.

Wandlinger, S.K. *et al.* J. Biol. Chem. 2008;283:18473–18477.

Wegrzyn, R.D., Deuerling, E. Cell. Mol. Life Sci. 2005;62: 2727–2738.

4.5.2 Vesicular Transport and Secretion of Proteins

4.5.2.1 Pathways of Transport (Fig. 4.5.2-1)

Proteins to be secreted from eukaryotic cells or to become members of the cellular membranes are transported by vesicles to their destination. This is a conserved mechanism in all eukarya. Membrane-coated vesicles control the architecture of intracellular compartments and communication of the cells with their environment. Major trafficking pathways are the inward flux of endocytotic vesicles from the plasma membrane back to the Golgi (4.4.2) or the ER (4.4.1) (endocytosis), and the outward flux of exocytotic vesicles including newly synthesized proteins formed in the ER and modified in the Golgi apparatus to the plasma membrane (exocytosis). For secretion, the proteins follow either the constitutive pathway for permanent secretion or the regulated pathway.

In this case, they are stored intracellularly in secretory vesicles, which fuse with the membranes and discharge their content only after receiving specific signals (e.g., in presynaptic vesicles of the nervous system, 7.2.3). All deviations from the general bulk-flow pathway are signal encoded. As an example, the mannose groups of glycoproteins destined for lysosomes are phosphorylated (Fig. 4.4.2-1). They become attached to a mannose-6-P-receptor, which causes the protein to enter the lysosomes. Thereafter, the receptor is recycled to the Golgi.

4.5.2.2 Transport Vesicles (Table 4.5.2-1, Fig. 4.5.2-2)

Different self-assembled multilayer coated vesicles participate in the transport processes. Transport vesicles include coat proteins (clathrin), coat protein complex I (COPI) and coat protein complex II (COPII), adaptor complexes (AP), soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) proteins, and Rab GTPases. Most yeast proteins are interchangeable with mammalian proteins in the various *in-vitro* transport assays.

Vesicles coated with COPI are used for the transport of cargo between the ER and the early secretory compartments and intra-Golgi transport processes (Fig. 4.5.2-2a). In yeast, the COPI complex is recruited to the membrane by the Sar1 GTPase, which can be inhibited by Brefeldin A. The COPI heptameric ring complex is composed of the two subcomplexes consisting of Sec26, Sec27, Sec21 (α, β´ and ε proteins in mammals) and of Ret1, Ret2, and Ret3 (β, γ, δ, and ζ proteins in mammals). The entire COPI complex interacts with transmembrane proteins harboring canonical KKXX/RXR recycling/retrieval sequences. The disassembly of the COPI coat is triggered by several GTPase activating proteins (GAPs) which may be used for vesicles with different cargo.

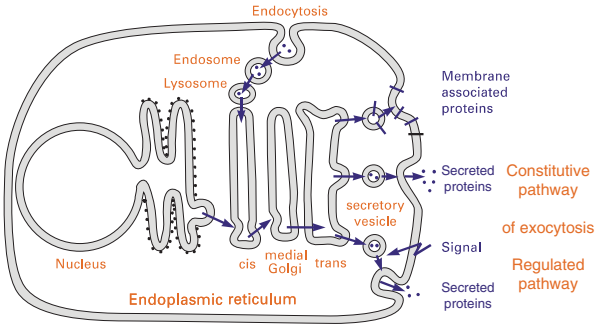


Figure 4.5.2-1. Pathways of Vesicular Transport

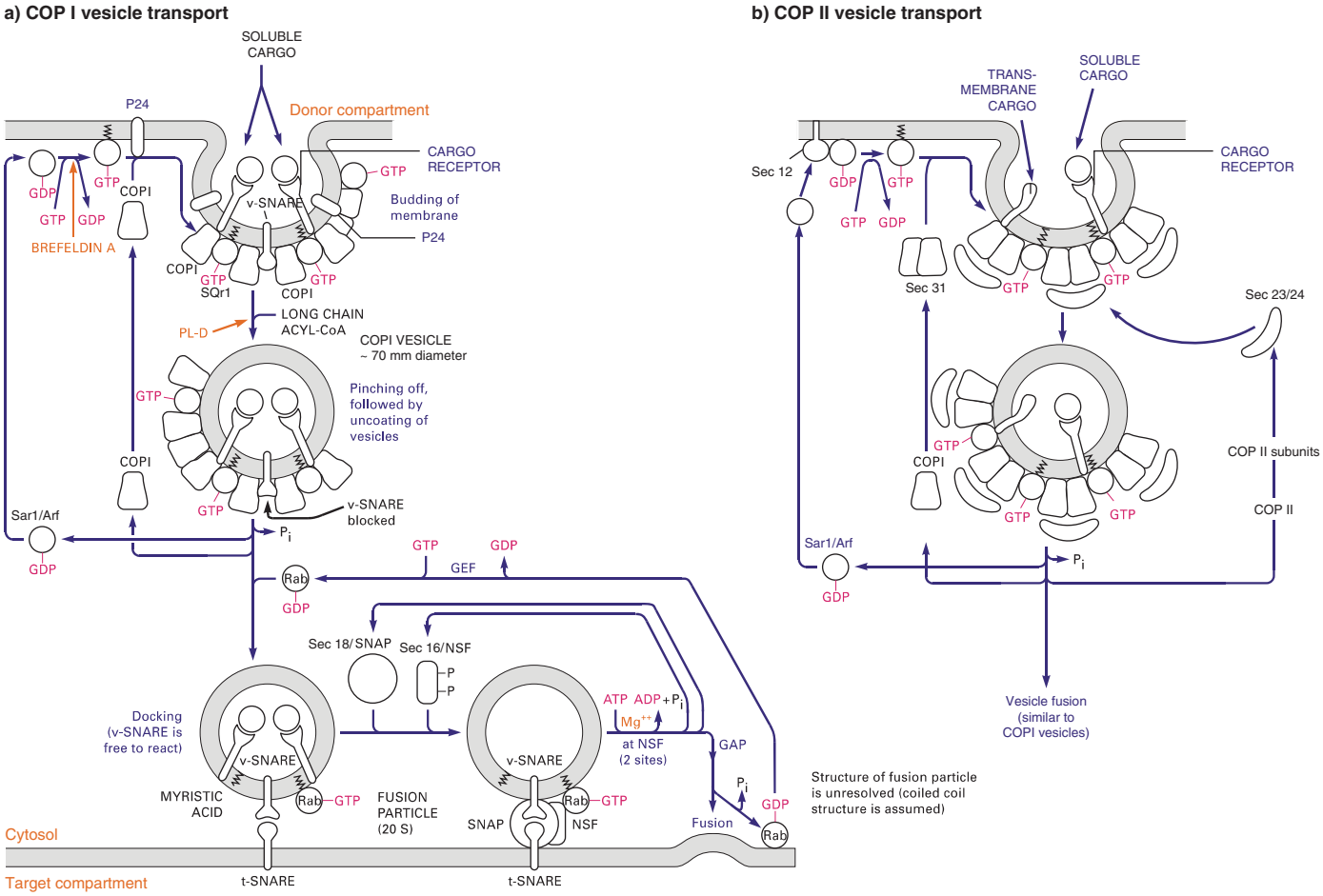


Figure 4.5.2-2. Mechanism of Intra-Golgi Vesicular Transport

Vesicles coated with COPII (Fig. 4.5.2-2b) are involved in trafficking from the ER to the Golgi. Their formation is initiated by recruitment and activation of the small GTPase Sar1 by the membrane-bound protein Sec12 (guanine nucleotide exchange factor, GEF). This results in the exposure and interaction of its N-terminal αα-helical domain and attaches the GTPase to the membrane. Membrane-associated Sar1-Sec12 recruits the Sec23-Sec24 adaptor complex, and combines further with the scaffolding Sec13-31 cage complex. The Sec23 protein is a Sar1-specific GTPase activating protein (GAP) and Sec24 binds the cargo through ER exit motifs.

The endocytotic pathway and the pathway from the trans-Golgi network to endosomes proceed via clathrin-coated vesicles (6.1.5).

Assembly, budding, tethering and vesicle fusion are controlled by Arf and Rab GTPases which cycle between the cytosol and the membranes and their modulators (GEFs, GAPs). Initial interaction of the vesicle with its target membrane involves complex formation between different Rab effectors and Rab exchange factors (tethering factors) on the vesicle and the target membrane. For instance, the TRAPPI, TRAPPII, Dsl1, COG and exocyst complex are required for secretion between ER and Golgi. Activated GTP-bound Rab proteins associate with the membrane whereas inactive GDP-bound Rabs bind to GDP-dissociating factors (GDIs). To remove Rab proteins from the GDIs and for Rab attachment to the membrane GDI-displacement factors (GDFs) are needed. Attachment to the membrane involves Rab binding to several specific Rab effector molecules which act as linker between the Rabs and membrane proteins/lipids (e.g., SNAPs).

Subsequent vesicular fusion to the target membrane is mediated by the SNARE proteins. Compartment-specific SNARE proteins of the vesicle (v-SNARE) interact with specific t-SNARE proteins on the target membrane and this causes the specificity of the membrane fusion.

Table 4.5.2-1. Proteins with Functions in the ER-Golgi Transport System

	Mammalian		Homology in Yeast
Arf1	ADP ribosylating factor (GTP/GDP binding)	++	Sar1
COPII complex	Membrane-bound		Sec12 (COPII)
	GTPase activating protein (GAP)		Sec23 (COPII)
	Cargo adaptor protein		Sec24 (COPII)
	Scaffolding cage complex: coatomer, helps to deform the membrane, stabilizes the polymerizing coat		Sec13/31 (COPII)
COPI complex	Coat proteins β, β′, γ, α, δ, ε, ζ	+	Sec26, 27, 21 (COP I) Ret1, Ret2, –, Ret3
p24	Transmembrane protein interacts with COPI and Arf-GTP		
GAPs	GTPase activating proteins ArfGAP1, ArfGAP2/3 for COPI, Sec23 for COPII		
GEFs	Guanine nucleotide exchange factor Gbf1, Sec7 domain proteins for COPI, Sec12 for COPII		
Rabs	GTP/GDP binding > 60 mammalian Rabs, Rab1, Rab3, Rab4, Rab5, Rab8, Rab9, Rab11, Rab26, Rab27		11 Rabs, Ypt1p, Sec4, Ypt5, Ypt7, Uso1p
GDI	Guanine dissociation inhibitor Yip3 for Rab9		
GDF	Membrane GDI-displacement factor		
Rab effectors	Rab binding proteins, act as linker proteins between Rabs on vesicles and protein/lipids in the target membrane, e.g., Rabphilin interacts with Rab3-27, Rim1/2, synapsin, Slp4, Slp2, Munc13-4		e.g., Yip1p interacts with Ypt1p
NSF	N-ethylmaleinimide sensitive protein		Sec18
SNAP	NSF attachment protein (α, β, γ)		Sec16 (homologous to SNAP α)
SNARE	SNAP receptor (v-SNARE = vescile-SNARE t-SNARE = target SNARE)	 * *	Sec22, Bos1, Bet1 Sec9, Sed5, Sso1

+ similar: COP II in transport from ER to Golgi
++ similar: Sar1p in transport from ER to Golgi
* similar: VAMP (synaptobrevin) in presynaptic vesicles
** similar: syntaxin in presynaptic membranes, stabilized by SNAP-25

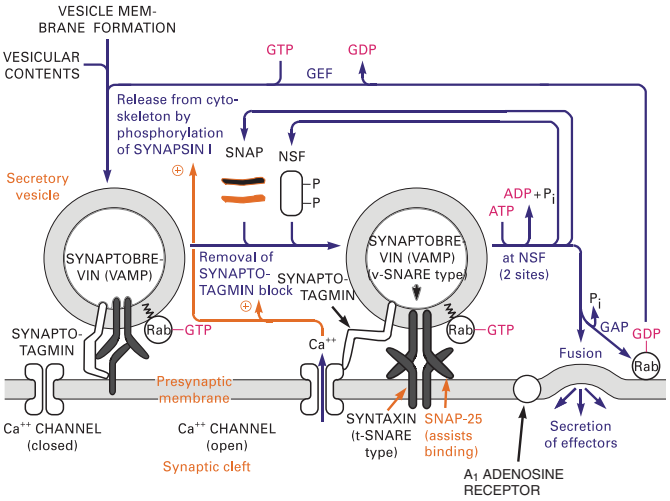


Figure 4.5.2-3. Action of Synaptotagmin in Nerve Terminals

Subsequent formation of the highly stable SNARE core complex, including the N-ethylmaleinimide sensitive protein (Sec18 in yeast, NSF in mammals) and the associated protein (Sec16 in yeast, SNAP in mammals) induces membrane fusion (proteins with analogous function in other transport steps are noted below in Table 4.5.2-1). ATP hydrolysis at Sec18/NSF supplies the energy for the fusion step.

In neurons, synaptic vesicles containing neurotransmitters (7.5.3) are stored near the presynaptic membrane after their synthesis and active transport to this site (6.1.5). They are attached to the cytoskeleton by synapsin I and to the presynaptic membrane by interaction between the membrane receptors syntaxin (t-SNARE analogue), the synaptosome associated protein of 25 kDa (SNAP-25) and the vesicle receptor synaptobrevin or VAMP (v-SNARE analogue). However, the vesicle membrane component synaptotagmin (p65) prevents formation of the fusion particle, possibly by blocking SNAP-25 access to syntaxin. Increase of intracellular Ca^{++} due to nerve stimulation releases the attachment to the cytoskeleton by phosphorylation of synapsin I. It also removes the synaptotagmin block, permits the formation of the fusion complex and concomitant release of the vesicle contents into the synaptic cleft (Fig. 4.5.2-3).

Literature:

Bethune, J. *et al.* J. Mem. Biol. 2006;211:65–79.
Cai, H. *et al.* Dev. Cell 2007;12:671–682.
Chaineau, M. *et al.* FEBS Letters 2009;583:3817–3826.
Fukuda, M. Cell Mol. Life Sci. 2008;65:2801–2813.
Gürkan, C. *et al.* Nat. Rev. Mol. Cell. Biol. 2006;7:727–738.
Kirchhausen T. Cell 2007;129:1251–1252.
Sato K., Nakano A. FEBS Letters 2007;581:2076–2082.
Spang, A. Curr. Biol. 2004;14:R33–R34.
Spang, A. Cell Mol. Life Sci. 2008;65:2781–2789.
Stagg, S.M. *et al.* Curr. Opin. Struc. Biol. 2007;17:221–228.
Stagg, S.M. *et al.* Cell 2008;134:474–484.
Strating, J.R.P.M., Martens, G.J.M. Biol. Cell 2009;101:495–509.
Stenmark, H. Nat. Rev. Mol. Cell Biol. 2009;10:513–525.
Vassilieva, E.V., Nusrat, A. Methods Mol. Biol. 2007:3–14.
Wu, H. *et al.* Trends Biochem. 2008;18:397–404.

4.5.3 Protein Transport into the Nucleus (Fig. 4.5.3-1, Table 4.5.3-1)

Both prokaryotic (3.10.2) and eukaryotic cells (4.5.2) possess protein export mechanisms that allow the translocation of polypeptides from the cytoplasm into the extracellular space. In addition, eukaryotic cells have translocation systems for proteins into the intracellular organelles such as the nucleus, the mitochondria (4.5.4) and the chloroplasts (4.5.5).

Import of proteins into the nucleus is a signal and energy dependent, saturable process, which is highly conserved in eukaryotic species. Examples for transported components are DNA/RNA polymerases, transcription and splicing factors, histones and viral proteins. In contrast to the transport of proteins across other membranes (e.g., endoplasmic reticulum, mitochondria, chloroplasts and also bacteria), it comprises several unique features:

- The carriers are soluble and shuttle back and forth
- Folded proteins can be transported
- Movement of proteins between the nucleus and the cytoplasm occurs via a nuclear pore complex.

The nuclear pore complex (NPC) is a protein complex of about 30 different proteins (nucleoporins) spanning the nuclear envelope which is large enough to allow passive diffusion of proteins < 60 kDa. The protein assembly is modular, with a stable structural scaffold supporting dynamically attached components. It disassembles during mitosis. The NPC of *Saccharomyces cerevisiae* has been classified in the TC system (see 6.1.1.2) under 9.A.14.1.1. The scaffold is made from multiple copies of the heptameric Y complex which itself is composed of heterotrimeric 134 kDa units of Nup84-Nup145C-Sec13.

Nucleoporins are classified in different families:

- Transmembrane nucleoporins (TM),
- FG/FN nucleoporins with FG repeats (GLFG, FXFG, FG, containing binding sites for karyopherins) or with FN repeats
- Nucleoporins with a WD repeat or a seven bladed propeller motif.

The total mass of the NPC is ca. 1.25 * 10^8 Da with an eightfold rotational symmetry axis perpendicular to the membrane, but with an asymmetrical structure across the membrane. It consists of three substructures

- the cytoplasmic filaments (50 nm),
- the central core,
- the nuclear basket made of the nuclear fibrils (100 nm) branching from the nuclear ring that merge at the distal end (terminal ring).

The outer diameter is ca. 125 nm; the channel diameter 8–10 nm and the length nearly 45 nm. A network of hydrophobic polypeptide nucleoporins lining the central channel acts as selective filter for small proteins. Larger molecules bind to the nucleoporins, the channel expands up to 26 nm and the active protein transport occurs much faster.

Nuclear transport is mediated by a family of transport receptors (karyopherins). The karyopherins for import (importins) and for export (exportins) interact directly with the transported proteins ('cargo').

The import machinery consists of

- cytosolic components: the transport mediators importin and Ran (a GTP binding protein), GTPase interacting factor and optionally, a Hsp70-homologue, and
- subunits of the nuclear pore complex (special O-glycosylated nucleoporins). A selection is listed in Table 4.5.3-1.

An important mechanism also exists for export of RNA, RNP (see 4.2.1.3–4.2.1.8) and proteins. Many details are still unknown. However, there are striking similarities to the import process.

Targeting mechanism (Fig. 4.5.3-1, upper part): Imported cargo molecules carry a signal for selective transport (nuclear localization sequence, NLS). This is a continuous sequence or consists of two clusters composed of positively charged amino acids with no obvious homologies (e.g., in SV40 antigen PKKKRKV and in nucleoplasmin KR –10 aa – KKKL). The NLS sequence of the cargo protein can form a complex with the adaptor molecule karyopherin-α/importin-α, which mediates protein import. However, most importins bind directly to cargo proteins and do not rely on the karyopherin-α adaptor. Promoted by importin, the importin/NLS-protein complexes bind to special nucleoporins (e.g., Nup358) on the cytosolic side of the NPC.

Ran-dependent transport mechanism (Fig. 4.5.3-1, lower part): Nuclear import of proteins with a NLS generally requires the 25 kDa cytoplasmic factor Ran and the hydrolysis of the GTP bound to it.

First, an import karyopherin-α/importin-β/NLS-protein complex is formed in the presence of the Ran-GTPase activating protein RanGAP1, the RanGAP activating protein RanBP1 and low concentrations of Ran-GTP in the cytoplasm. This complex docks to the RanBP2 nucleoporin (Nup358), a component of the cytoplasmic filaments.

Second, the cargo/importin-complex moves from the periphery of the nuclear pore complex to its central region (e.g., by bending of the fibrils) and from there to the nucleoplasmic side after release of the transport complex upon hydrolysis of GTP.

The subsequent transport into the nucleoplasm (= nuclear matrix) possibly occurs by multiple dissociation-association cycles of the cargo complex with the FG repeats of various nucleoporins lining the pore complex, which are mediated by binding of the nuclear transport factor Ntf2 to Ran-GDP. In the nucleoplasm, Nup153 (a nuclear basket protein) binds the transport complex by interacting with importin-β via its FG domains. Thereafter the nuclear basket component Trp interacts with the complex. This destroys the connection between the Nup153 and importin-β. The cargo/importin/Ran-GDP complex becomes disassembled and the imported protein is released into the nucleoplasm. RanGDP is transformed into RanGTP by the Ran guanine exchange factor (RanGEF), activated by RanBP3. The assisting import components are recycled into the cytoplasm (karyopherin-α in a complex with CAS/Ran-GTP and importin-β in a complex with Ran-GTP).

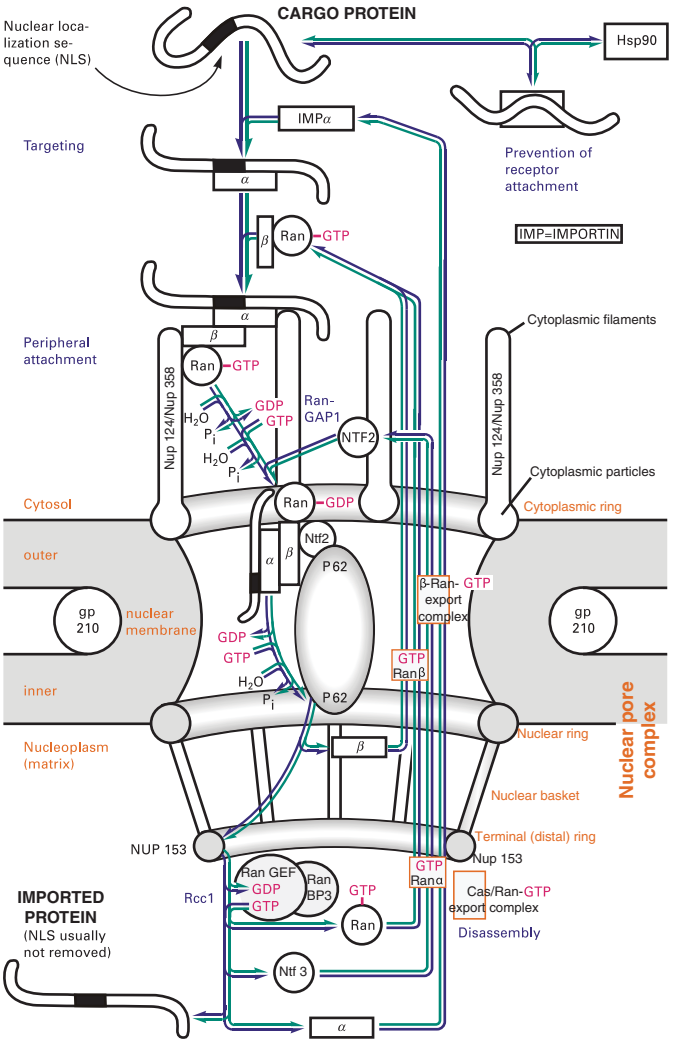


Figure 4.5.3-1. Mechanism of Protein Import into the Nucleus
The nomenclature for vertebrates is used.

Table 4.5.3-1. Components for Import into the Nucleus

Localization	Selection of Components (Species)	Characteristics, Function
Cytosol	Importin-α [karyopherin 60 (frog); p54/p56 (bovine, rat); Kap60 (yeast); NPI-1/Rch1 (human)]	Adaptor, NLS receptor, binds to NLS of the cargo and to importin-β. Enters nucleus together with cargo, Ran, NTF2
	Importin-β1 [karyopherin 90 (97 kDa, frog); p97 (bovine); Kap95 (yeast)]	Nucleus transport of many proteins with classical NLS in complex with karyopherin-α/ importin-α ?, promotes docking of cargo/importin-α complex to binding site at the NPC (NUP 358, NUP 214, NUP 153), at the nuclear envelope
	Importin-β2/Transportin-1 (Kap104, yeast) Transportin SR1 (Mtr10/Kap111, yeast) Transportin SR2 Importin 4 (Kap123, yeast) Importin 5/Karyopherin-β3 (Kap 121; yeast) Importin 7 (Nmd5/Kap119; Sxm1/Kap108; yeast)	Transport of hnRNPA1, histones, ribosomal proteins Transport of SR proteins HuR transport Histones, ribosomal proteins Histones, ribosomal proteins HIV RTC, glucocorticoid receptor, ribosomal proteins (TFIIS, Hog1, Lhp1, ribosomal proteins, yeast)
	Importin 8 Importin 9 (Kap114, yeast) Importin 11 Importin 13 Cas	SRP19 transport Histones, ribosomal proteins, UbcM2, rpL12 import Rbm8, Ubc9, Pax6 import Exports karyopherin-α
	Ran [TC4 (25 kDa, frog, mammals); Gsp1 (yeast)]	'Small' G-protein (Table 7.5-2). Binds as Ran•GTP to Nup358, hydrolyzes GTP. Mediates the committed step from binding NPC to translocation. Very abundant.
	RanGAP1; Rna1p (yeast)	Binds to Nup's. Enhances GTPase activity of Ran on bound GTP
	RanGEF/RCC1	Guanine nucleotide exchange factor
	Ntf2 (15 kDa, human); p10/B-2 (10 kDa, frog)	Transfer of the transport complex from docking site to gated channel (p62)
	Hsc70 (Hsp 70-like protein, human)	Chaperone, facilitates cargo binding to importin-α (not a general import factor)
Nuclear pore complex		
Cytoplasmic side	Nup 358/RanBP2, Nup214 (rat), Nup195 (yeast)	Nucleoporin, coactivator for RanGAP, bound by RanGAP1 localized to fibrils, contain FXFG repeat and four Ran-binding domains. Major component of the cytoplasmic fibrils, FG
	Nup214, hGle1, Nup88	Enriched in the cytoplasmic ring
Nucleoplasmic side	Nup98 (rat), NUP153 (rat), Nup50 (NPAP60), Tpr p265 (rat), NSP1 (yeast)	Enriched in the nuclear ring and the fibrils
	Nup62-complex: Nup62, Nup58, Nup 54 (rat, frog, mouse); NUP1/NUP2 (yeast)	Central region of the nuclear pore, special O-glycosylated nucleoporins, docking sites for cargo-NS receptoradaptor complex, involved in stepwise transport of the cargo through the NPC, proposed binding sites for Ran and importin-β.
	Nup35, Nup107, Nup75/Nup85, Nup93, Nup160, Nup133, Nup 96, Nup 155, Nup188, Nup 217(CAN), Nup 205, Rae1/Gle2, Seh1 (WD repeat)	Distributed symmetrically to the nuclear membrane
	Pom 121, GP210 (TM)	Pore membrane, anchors the NPC in the nuclear membrane
	Tpr (Nup60 yeast)	Nuclear
	ALADIN, Nup37, Nup43	? (WD repeat)
Nucleoplasm	RanGEF/Rcc1	GDP/GTP exchange factor for Ran

Localization of RanGAP1 in the cytoplasm and RanGEF in the nucleus ensures Ran complex formation with GTP mostly in the nucleus (high GTP concentration), while Ran-GDP complexes are prevalent in the cytoplasm. As a consequence, a Ran-GTP gradient is established that ensures the directionality of the transport.

The molecular mechanism of the transport steps (including interaction with the central channel, channel gating, translocation, disassembly and recycling) and the biophysical and energetic nature of translocation is still not known in all details. However, Hsp70-like proteins do not seem to have the same prominent role as in other systems.

Ran-independent transport mechanism: Several alternative translocation pathways have been described which are independent of Ran. For example, β-catenin, importin-β, transportin-1, exportin-t, calmodulin, kinase ERK, STAT proteins, PKI, the glucocorticoid receptor and others were all found to be Ran-independent. In some cases they do not require GTP/ATP hydrolysis (e.g., transportin-1) and/or seem to possess an affinity to certain nucleoporins (e.g., exportin-t and the STAT proteins).

Furthermore, additional functions for NPCs and their constituent proteins independent of classical transport are known. Nucleocytoplasmic transport is specifically regulated by enzymatic activities at the NPC which use it as a regulatory scaffold. Also, gene expression is controlled by nucleoporins by contacting chromatin.

Regulation of the nucleoplasmic transport: The nucleoplasmic distribution of certain proteins in response to environmental signals, growth, proliferation, and differentiation is hierarchically controlled at different levels. For instance, nuclear import can be regulated by modulation of the importin interactions with the substrate NLS. Intra- and intermolecular masking of the substrate NLS from recognition by the importin is a widespread mechanism. E.g., the transcription factor NFkB is synthesized as a p105 precursor molecule in which the NLS is inaccessible. Upon immune activation, p105 is phosphorylated, the C-terminal domain is degraded and the resulting p50 protein is now recognized by the karyopherin α/importin-β complex.

Intramolecular masking can also be caused by phosphorylation of the NLS sequence or results through conformational changes due to the formation of disulfide bonds. Intermolecular masking occurs by binding of the NLS-containing protein to other regulatory proteins or nucleic acids. E.g., BRCA1-binding protein attaches to BRAP2, the NLS of the transcription factor Gal4 of yeast overlaps with its DNA-binding region and IkB interacts with the NLS of the transcription factor NF-kB (Table 4.2.2-2). IKB is phosphorylated upon immune challenge. This results in its ubiquitinylation and degradation by the proteasome (Fig. 4.5.7-1), allowing the nuclear import of NFkB p. 65.

Alternatively, regulation of the transport can be controlled by binding of importins to the the NLS, e.g., by phosphorylation that strengthens the interaction of the NLS with the karyopherin-α/ importin-β complex. In addition, nuclear transport can be regulated by binding of the NLS to specific factors that retain the protein in the cytoplasm. For example, Hsp90 interaction with the steroid hormone

receptor (7.7) retains the protein in the cytoplasm, but nuclear transport is initiated upon binding of the hormone. It is also known that cotransport with other proteins and changing of the cargo-binding properties of the karyopherins as well as changes of the variety of importants and nucleoporins can modulate the efficiency of the transport process.

Literature:

Brohawn, S.G., Schwarz, T.U. Nat. Struct. Mol. Biol. 2009;16: 1173–1177.

Fernandez-Martinez, J., Rout, M.P. Curr. Opin. Cell Biol. 2009;21:603–612.

Fiserova, J., Goldberg, M.W. Biochem. Soc. Trans. 2010;38:273–277.

Hetzer, M.W., Wente, S.R. Dev. Cell 2009;17:606–616.

Lim, R.Y. *et al.* Int Rev. Cell Mol. Biol. 2008;267:299–342.

Peters, R. Biochim. Biophys. Acta. 2009;1793:1533–1539.

Sorokin, A.V. *et al.* Biochemistry 2007;72:1439–1457.

Terry, L.J. *et al.* Science 2007;318:1412–1416.

Wagstaff K.M., Jans, D.A. Traffic 2009;10:1188–1198.

Xylourgidis. N., Fornerod, M. Dev. Cell 2009;17:616–625.

4.5.4 Protein Transport into Mitochondria (Fig. 4.5.4-1, Table 4.5.4-1)

More than 90% of the mitochondrial proteins (about 10–15% of the nuclear genome) are encoded on the nuclear DNA and have to be imported posttranslationally after synthesis by cytosolic ribosomes. Examples are subunits of the respiratory chain (e.g., ATPase subunit $F_1\beta$, 3.11.4.5), translocators (e.g., ADP/ATP-translocase, 3.11.4.5) and matrix enzymes (e.g., alcohol dehydrogenase III, ornithine carbamoyltransferase, 3.2.9.1). Mitochondrial proteins are synthesized in the cytosol as precursors and associated with stabilizing factors (e.g., Hsp 70, Hsp90) or presequence binding factors (e.g., MSF) which keep them in a transport-competent conformation. The precursors are recognized by receptors on the surface of the mitochondria and special translocases in both mitochondrial membranes promote the import and subsequent intra-mitochrondrial sorting. The family of mitochondrial translocases in *Saccharomyces cerevisiae* has been classified in the TC system (6.1.1.2) under 3.A.8.

Targeting mechanism (Fig. 4.5.4-1, upper part): The typical imported protein is a precursor containing a N-terminal targeting presequence or prepeptides (10–80 amino acid residues), also called matrix targeting signal, which is removed enzymatically after import. These sequences can be identified by certain computer programs (MITOPRED, Mitoprot II, PSORT II, TargetP, see under 'Literature'). Some presequences, however, are not cleavable. Many precursors, however, lack N-terminal presequences and contain internal targeting sequences, e.g., outer membrane proteins. Presequences do not share sequence homologies, but show a high content of basic and hydroxylated amino acids and a lack of acidic residues and tend to form amphipathic helices with a hydrophobic and positively charged phase.

Targeting is achieved by specific binding of the presequence to receptors at the outer membrane. Additional targeting sequences can also target the protein to the final subcompartments (outer membrane, intermembrane space, inner membrane).

Transport mechanism (Fig. 4.5.4-1, lower part): The transport of the various proteins takes place through cooperation of a single type of import machinery located in the outer (TOM-complex) and inner (TIM-complex) membrane and proteins in the mitochondrial matrix (chaperones). The precursor protein can be transported simultaneously across both membranes at so-called translocation contact sites, where both membranes are in close vicinity. The process is thought to take place through an aqueous protein-conducting channel formed by interacting TOM and TIM-complexes (Table 4.5.4-1).

The presequences of the precursor proteins are recognized and become attached to receptors of the TOM-complex. It facilitates the release of cytosolic chaperones, contributes to the unfolding of protein domains, assists in insertion of some outer membrane proteins and transfers matrix and inner membrane polypeptides through pores in the outer membrane. Tom20 and Tom70 are the main receptors for precursor proteins on the outer (*cis*-) site, which differ in their substrate specificity. Many precursors with N-terminal prepeptides bind to the Tom20 receptor. Tom20 is connected to Tom22, which assists in binding and unfolding and mediates connection to the central translocation pore. Primary recognition of chaperone-bound hydrophobic precursors with internal signal sequences is achieved by binding to the Tom70/Tom37 complex. Binding of ATP triggers release of chaperones and insertion of the polypeptide into, and passage through, the

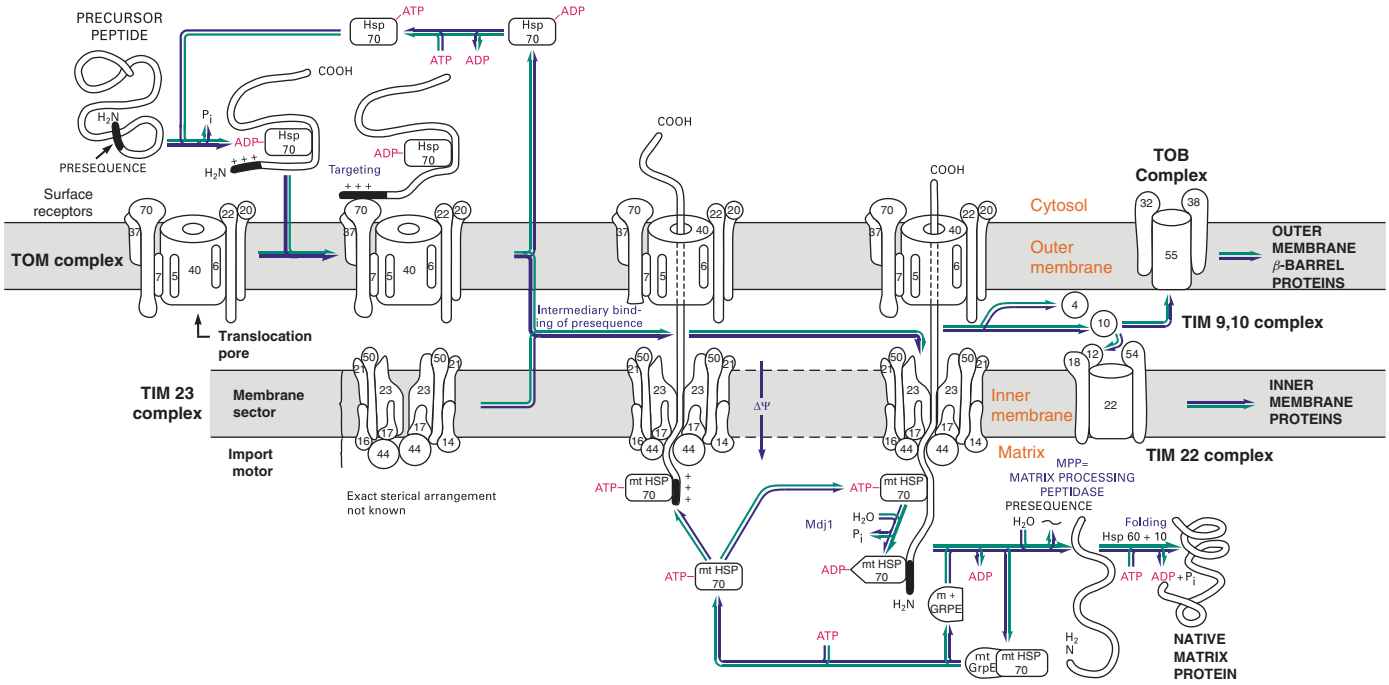


Figure 4.5.4-1. Mechanism of Protein Import into Mitochondria

Table 4.5.4-1. Components for Import into the Mitochondria

Location	Consensus Nomenclature	In yeast (*S. cerevisiae*)	Function, Characteristics (compare Table 4.5.1-1 for chaperones)
Cytosol		ctHsp70	Import competence, ATPase, DnaK homologue
		Vdj1	Import competence, DnaJ homologue
	MSF *(rat)*	(homol. to MSF)	Mitochondrial import stimulating factor, dimeric, targets precursor, dissolves aggregates, ATPase
Outer membrane	Tom70	Mas70	Surface receptor, recognizes MSF-bound precursors
	Tom71	Tom71	Associated with Tom70 and Tom37
	Tom37	Mas37	Associated with Tom70 and Tom71
	Tom22	Mas22	Surface receptor, essential for recognition of precursors
	Tom20	Mas20	Surface receptor, associates with Tom22
	Tom40	Isp42	Translocation channel
	Tom7	Mom7	Translocation channel regulator
	Tom6	Isp6	Translocation pore associated with Tom40
	Tom5	Mom8	Translocation pore associated with Tom40
	Tob55/Sam50/Tom50 Tob38/Sam35/Tom38 Mas37/Sam37/Tom37 Mdm10		Form the TOB/SAM complex for import of b-barrel outer membrane proteins
Inner membrane	Tim50 Tim23 Tim21 Tim17 Tim33 Tim11	Mim23 Mim17 Mim33	Form the membrane sector α complex, achieve transport of precursor across inner membrane. Tim21 might play a role in the interaction with the TOM complex, Tim50 interacts with incoming peptide chains, passes it to Tim23, which is important for substrate binding, Tim17 plays a role in gating the channel
	Tim44 Tim16 Tim14	Mim44 Mim14	Peripheral proteins form the import motor, transfer of precursor to mtHsp70, import motor
	Tim54 Tim22 Tim18 Tim9•10•12 complex		Form the TIM22 complex for import of inner membrane proteins
Matrix	mtHsp70		Translocation and sorting, ATPase, DnaK homologue
		Mdj1	Recycling of Hsp70, DnaJ homologue
	Mge1		Recycling of Hsp70, GrpE homologue
	?	?	Folding and assembly, GroEL/ES homologue
		Mas1/2	Enzymatic removal of targeting presequence
	Pam17		?
	Mmp37/Tam41		?

TOM pore complex, consisting of the central channel forming component Tom40 and the associated subunits Tom5–7. After transport across the outer mitochondrial membrane, another temporary binding of the presequence takes place at the inner (*trans*-) site of the outer membrane, and seems to contribute to vectorial translocation and to prevent aggregation.

The TIM23 complex is the major translocase in the inner membrane of the mitochondria. The energy for transport of the positively charged targeting sequence is supplied by the electric membrane potential and the hydrolysis of ATP (1.5.3 and 7.2.1). The proteins Tim21, Tim50, Tim23 and Tim17 form the inner-membrane core (membrane sector), which plays a critical role in substrate binding, permeability and gating of the pore. It can only transfer the matrix targeting signal by the proton motive force, then the import motor has to take over. The import motor consists of the peripheral components Tim44, Tim14, Tim16 and mtHsp70. Within the matrix, the protein is at first attached to Tim44. Then mtHsp70•ATP binds to the protein. According to the 'Brownian ratchet' model, mtHsp70 traps the protein and makes inward Brownian movements and spontaneous sliding reactions of the protein into the import machinery irreversible. In addition or alternatively, the ATP hydrolysis (assisted by Mdj1) could cause a conformational change of Hsp70, which acts as a lever arm and pulls the protein inward by a conformational change that occurs upon ATP hydrolysis ('translocation motor model', analogous to the 'power stroke' of myosin, 7.4.5). mtGrpE (Mge1) removes ADP from Hsp70 and causes its dissociation from the protein. Then the cycle begins again. Substrate binding and ATP-hydrolysis of mtHsp70 are controlled by DnaJ-like proteins Tim14 and Tim16. During the import step, the targeting sequence is removed by a specific peptidase. Folding and assembly of matrix proteins into complexes is thought to be facilitated by molecular chaperones (e.g., mt-Hsp70 and Hsp60/10), energized by ATP-hydrolysis.

Proteins destined for other subcompartments (outer membrane, inner membrane and intermembrane space) are sorted either during transport or are delivered by a second transport event after they have entered the matrix.

- Precursors of outer membrane β-barrel proteins are bound by small TIM proteins (Tim9•10 complex) after translocation through the TOM complex and are subsequently imported and assembled through the TOB complex (Tob55, Tob38 and Mas37) located in the outer membrane.
- For inner membrane proteins, three different import routes were identified:
  - Precursors with internal target sites use the Tim22 complex composed of Tim54, Tim22, Tim18 and the Tim9•10•12 complex.
  - Stop-transfer pathway: The precursors use the Tim23 complex and are laterally inserted into the membrane.
  - Conservative sorting pathway: The import occurs from the matrix. This sorting step is energized by the electron membrane

potential across the inner membrane and by matrix-ATP (see also chloroplasts, 4.5.5 and bacteria, 3.10).

- Proteins destined for the intermembrane space can be released from inner membrane precursor proteins
  - by proteolytic cleavage,
  - through rapid folding after translocation through the TOM complex (small proteins),
  - by high affinity binding with components of the intermembrane space.

Literature:

Ahmed, A.U., Fisher, P.R. Int Rev. Cell Mol Biol. 2009;273:49–68.

Balsera,, M. *et al.* Cell Mol Life Sci. 2009;66:1903–1923.

Bohnert, M. *et al.* FEBS Letters 2007;581:2802–2819.

Carrie, C. *et al*. FEBS J. 2009;276:1187–1195.

Chacinska, A. *et al*. Cell 2009;138:628–644.

Endo, T., Yamano, K. Biol. Chem. 2009;390:723–730.

Endo, T., Yamano, K. Biochim Biophys Acta 2009;390:723–730

Neupert, W., Herrmann, J.M. Ann. Rev. Biochem. 2007;76:723–749.

MITOPRED: http://bioinformatics.albany.edu~mitopred/; Mitoprot II: http://ihg.gsf.de/ihg/mitoprot.html; PSORT II: http://psort.ims.u-tokyo.ac.jp/; TargetP: http://www.cbs.dtu.dk/servies/TargetP/.

4.5.5 Protein Transport into Chloroplasts (Fig. 4.5.5-1, Table 4.5.5-1)

About 80% of the chloroplast proteins are encoded in nuclear DNA and have to be imported posttranslationally into the organelle after being synthesized on cytosolic ribosomes in precursor form. Examples are subunits of the photosynthetic machinery (3.12, e.g., subunits 17, 23 and 33 of photosystem II and plastocyanin) and stromal enzymes. Import occurs into the six distinct suborganellar compartments: outer envelope, intermembrane space, inner envelope, stroma, thylakoid membrane and thylakoid lumen. As a consequence, the internal transport route is a complex process. Envelope proteins usually utilize variations of the TOC/TIC import machinery, and proteins destined for the thylakoid membrane or lumen additionally employ one of four different targeting pathways. The family of chloroplast translocases in *Pisum sativum* has been classified in the TC system (6.1.1.2) under 3.A.9.

Targeting mechanism (Fig. 4.5.5-1, upper part): The typical imported protein is a precursor containing an N-terminal extension (transit peptide, targeting presequence; 35 … >100 amino acids), which directs the protein through a posttranscriptional targeting pathway and is removed after import. Some transit peptides, however, are not cleavable. Transit peptides do not share sequence homologies and vary in length (20 to >100 residues). However, they generally show a high content of aromatic and hydroxylated amino acids. They can also lack acidic residues and appear to form a random coil, similar to mitochondrial presequences (4.5.4). The random coil may be important to recruit cytosolic factors mediating the early stages in the import pathway or it could achieve a special structure upon contact with the outer envelope membrane. Cytosolic factors facilitate the passage of the precursors from the ribosome to the chloroplast surface and ensure that the proteins remain in an import-competent, unfolded state. Targeting into the organelle is then mediated by binding of the transit peptide to specific receptors in the outer envelope of the chloroplast (Table 4.5.5-1). Additionally, the precursor contains sequences affecting further targeting and sorting to the final subcompartments.

Transport mechanism (Fig. 4.5.5-1, lower part): The transport of the various proteins into the stroma takes place via a single type of import machinery, composed of two individual protein complexes located in the outer and inner envelope (TOC/TIC complex). The precursor protein can be transported simultaneously across both envelopes at translocation contact sites, where the envelopes are in close vicinity to each other. Transport presumably takes place through an aqueous protein-conducting channel. Energy is provided by hydrolysis of GTP and ATP. GTP binding proteins and Hsp70-like chaperones seem to play a central role in the recognition of the precursor transit sequence (e.g., proof-reading) and in translocation events.

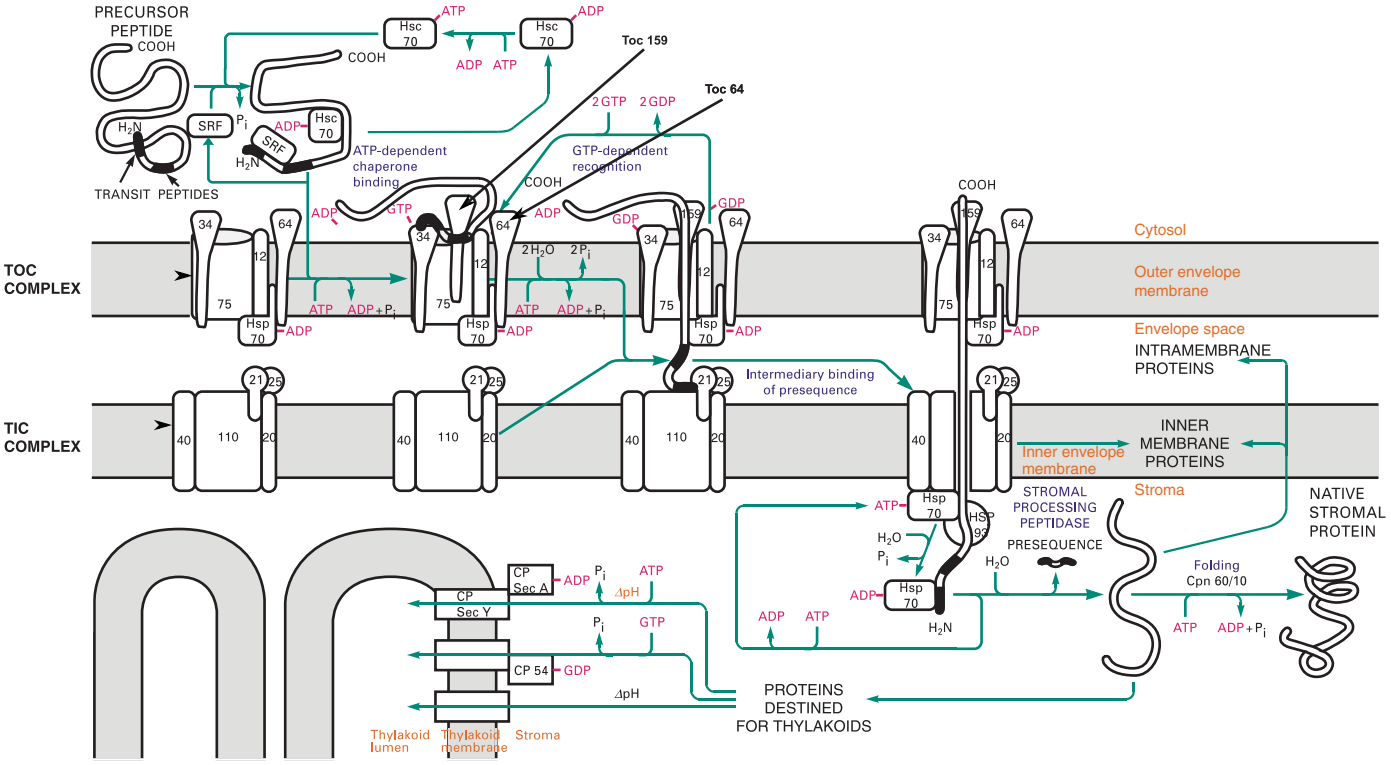


Figure 4.5.5-1. Mechanism of Protein Import into Chloroplasts

Table 4.5.5-1. Components for Import into Chloroplasts

Location	Consensus nomenclat.	Other names, Mol. Mass in kDa	Function/Characteristics (compare Table 4.5.4-1 for chaperones)
Cytosol	SRF		Signal recognition factor
	Hsc70		Import competence, Hsp homologue
Outer envelope	Toc159	*Arabidopsis* major isoform atTOC159, minor isoforms atTOC132 / atTOC120, atTOC90	Precursor recognition/GTP-binding part of the TOC core complex
	Toc34	OEP34/IAP34, *Arabidopsis* major isoform atTOC33 minor isoform atTOC34	Precursor recognition/GTP-binding part of the TOC core complex
	Toc75	OEP75, *Arabidopsis* major isoform atTOC75-III minor isoforms: atTOC75-IV, atTOC-I	Precursor recognition, β-barrel structure with polypeptide transport associated (POTRA) domains – channel formation and auxiliary function in assembly and transport, part of the TOC core complex
	Toc64	*Arabidopsis* isoforms atTOC64-III	Receptor, unknown function
	Toc12	*Arabidopsis* isoforms atTOC12	Co-chaperone; J domain
	Hsp70 (75 kDa)		Chaperone, Hsp70-homologue/integral membrane protein at *trans* side, ATPase
Inner envelope	Tic22	*Arabidopsis* atTic22-IV, atTic22-III	Protein recognition, TOC-TIC interaction
	Tic110	IAP100/Cim97/ClpC *Arabidopsis* atTic110	Precursor recognition, chaperone recruitment, β-barrel structure, channel formation, chaperone recruitment
	Tic20	*Arabidopsis* atTic20-IV, atTic20-II, atTic20-V	Homology to mitochondrial TIM machinery, four α-helical membrane spans, translocation channel
	Tic21	*Arabidopsis* atTic21	Translocation channel
	Tic40	*Arabidopsis* atTic40	Large stromal domain, co-chaperone, Sti1 domain
	Hsp93	*Arabidopsis* major isoform atHsp93-V, minor isoform at Hsp93-III	Hsp100 family, translocation motor
	Tic32	*Arabidopsis* atTic32-Iva, atTic32-IVb	Redox regulation, calcium regulation
	Tic55	atTic55	Redox regulation
	Tic62	*Arabidopsis* atTic62	Redox regulation
Stroma	Hsc70 (75 kDa)		Transport?, folding; Hsp70-homologue
	Cpn60/10		Chaperone, folding, assembly
	SPP (140 kDa)		Stromal processing peptidase (Zn protease), cleavage of transit sequence
	CPSecA		Transport into thylakoids, SecA homol. (bacterial)
	CPSecY		Transport into thylakoids, SecY homol. (bacterial)
	Tha4, Hcf106,		Homologues of the *E. coli* tatABC genes. Binding of substrate to the Hcf106–Tha4 core complex triggers recruitment of a separate Tha4 complex to form the active translocation complex
	CPtatC		
	CP54		transport into thylakoids, SRP homologue (ER)

The preprotein maintains a loosely folded conformation by interactions with cytosolic (Hsc70, SRF) and two different models have been proposed for their recognition by the TOC complex formed by outer envelope proteins (OEPs):

- Targeting model: The transit peptide binds specifically to cytosolic Toc159. The preprotein-Toc159 complex then binds to the Toc34 receptor in the outer envelope through interaction of the G-domains. This stimulates GTP hydrolysis and leads to the integration of the preprotein Toc159 into the TOC complex and the insertion and translocation of the preprotein;
- Motor model: the transit peptide is phosphorylated and then bound to Toc34. This stimulates GTP hydrolysis and dephosphorylation of the transit peptide as well as the transfer to Toc159. Subsequent GTP hydrolysis by Toc159 causes a large conformational change, forcing the preprotein into the Toc75 translocation channel.

When the transported preprotein becomes deeply inserted into the TOC complex and emerges at the *trans* side of the outer envelope, it binds to the molecular chaperone Hsp70.

Different isoforms of the TOC receptor proteins seem to be specialized for certain transport functions/targets. The major isoforms preferably bind highly-abundant photosynthetic proteins, while the minor isoforms mainly recognize non-photosynthetic housekeeping preproteins. This prevents the main flow of photosynthetic proteins from suppressing the import of less abundant housekeeping proteins. Furthermore, binding of the TOC receptors to GTP and thus to the preproteins seems to be controlled by phosphorylation through kinases located in the outer envelope membrane.

After translocation of the transit peptide through the TOC complex, contact with certain receptor proteins, e.g., Tic22 takes place. Then formation of the TOC-TIC supercomplex and opening of the membrane channel in the inner envelope membrane is initiated. Several proteins (Tic100, Tic21, Tic20) have been identified in the TIC complex, but their role in the import process is often unclear. It is possible that multiple TIC complexes exist with different channel components. However, it is widely accepted that the Tic110 protein binds transit peptides. This seems to favor the recruitment of the molecular Hsp93 chaperone and to ensure the unidirectional transport into the stroma. The chaperone seems to assist the import motor by stimulating the release of the transit peptide from Tic110. After transport of the protein into the stroma, the transit peptide is immediately removed by the stromal processing enzyme SPP. Progression through this step, folding and assembly into complexes is facilitated by molecular chaperones (Hsp70 and Cpn60/10 = Rubisco binding protein) and requires ATP, similar to folding after protein synthesis (9.5.1).

Targeting to the outer and inner membrane systems: For nonstromal proteins further sorting is again energy- and signal dependent (e.g., lumen targeting domain, LTD).

For outer envelope targeting of proteins, intrinsic targeting information in proteins is generally used and resides usually within a transmembrane domain. These targeting sequences ensure spontaneous insertions into the membrane bilayer, which is facilitated by its unique lipid composition. Other proteins seem to use Toc75 for outer membrane localization. The ankyrin repeat-containing protein (AKR2) acts as a key chaperone in this process which prevents aggregation of the target proteins.

Sorting into the intermembrane space is mostly unclear. There is evidence that two sorting mechanisms exist in which

- the TOC/TIC pathway is involved and enters the stroma, before relocation, or
- the translocation/processing does not occur in the stroma but within the envelope.

Two pathways exist for inner membrane proteins and both involve the TOC/TIC complex:

- the "stop transfer" pathway in which hydrophobic transmembrane regions block transport through the TIC complex, leading to lateral exit and integration into the inner membrane;
- the 'post-import' pathway, which is based on components, originating from bacteria, mediating the translocation of proteins from the stroma into the inner membrane.

For this import of proteins into the thylakoid lumen, apparently four ways exist:

- similar to the bacterial SecA dependent pathway (3.10), requiring ATP and ΔpH,
- similar to the Tat-dependent pathways, requiring the proton motive force (3.10),
- similar to the bacterial (3.10) and eukaryotic (4.4.1.1) SRP dependent pathways, requiring GTP, but acting post-translationally,
- spontaneous insertion, driven only by ΔpH.

Sec- and Tat-dependent thylakoid proteins possess signal peptides targeting into the lumen which promote interaction with the respective translocation complex and are cleaved by the thylakoid processing peptidase TPP. In addition, TOC/TIC-independent pathways have been recently identified.

Proteins using the co-translational endomembrane route are synthesized with signal peptides for translocation by the ER and Golgi which mediate glycosylation and subsequent translocation into the lumen of the chloroplast. Alternatively, proteins (i.e. Tic32, chloroplast envelope quinone oxidoreductase homologue: ceQORH) use a novel post-translational route in which internal non-cleavable targeting signals direct them to the inner membrane of the chloroplast. The exact mechanism and components involved in this transportation route are not yet clear.

Differences of protein import into chloroplasts as compared to mitochondria are:

- Almost no homology of envelope components involved in translocation
- Involvement of more than two different Hsp70-like proteins
- ATP-dependent binding of precursors to the import machinery
- Involvement of at least two GTP-binding proteins
- Requirement of NTP-hydrolysis in the intermembrane space
- No need for a membrane potential ($\Delta\Psi$) across the inner envelope.

Literature:

Ahmed, A.U., Fisher, P.R. Int Rev. Cell Mol Biol. 2009;273:49–68.

Balsera, M. *et al.* Cell Mol. Life Sci. 2009;66:1903–1923.

Benz, J.P. *et al.* FEBS Journal 2009;276:1166–1176.

Carrie, C. *et al.* FEBS J. 2009;276:1187–1195.

Inaba, T., Schnell, D.J. Biochem. J. 2008;413:15–28.

Jarvis, P. New Phytologist 2008;179:257–285.

Stengel, A. *et al.* Biol. Chem. 2007;388:765–772.

4.5.6 Protein Degradation

Peptidases are enzymes catalyzing the cleavage of peptide bonds (as defined by the IUBMB). Previously, the terms 'proteases' or 'proteinases' were equivalent to 'endopeptidases' and referred only to internal cleavage of longer peptide chains (proteins), but today 'proteases', 'peptidases', 'proteinases' and 'proteolytic enzymes' are all used interchangeably. All peptidases are assigned to families on the basis of amino acid sequence similarities (evolutionary relationships) and are collected in the MEROPS peptidase data base (http://merops.sanger.ac.uk/). Peptidases can either break specific peptide bonds (limited proteolysis), depending on the amino acid sequence of a protein, or digest a complete peptide or protein to amino acids (unlimited proteolysis). Degradation by the protease might only occur under certain circumstances (regulated proteolysis).

Terminology: Amino acid residues of the proteases are considered relative to the site of cleavage and are labeled Pi, ..., P3, P2, P1, P1′, P2′, P3′, ..., Pj from the N-towards the C-terminal end of the polypeptide substrate, and their respective binding subsites on the enzyme are labeled Si, …, S3, S2, S1, S1′, S2′, S3′, ..., Sj). The cleavage is catalyzed between P1 and P1′.

Substrate:	-P3-P2-P1 | P1′-P2′-P3′	(| scissile peptide bond)
Enzyme	-S3-S2-S1*S1′-S2′-S3′	(* catalytic site)

These enzymes serve a large number of functions in living organisms:

- Digestion of dietary proteins
- Intracellular protein turnover including termination of enzyme action, e.g., cyclins and other cell cycle factors (4.3)
- Extracellular proteolysis of cell surface factors, e.g., by secreted microbial peptidases during infection
- Maturation of newly synthesized proteins; e.g., removal of signal or targeting sequences (4.4.1.2, 4.5.3 … 5)
- Processing of precursor proteins and peptides, e.g., of organelle proteins or secreted proteins such as hormones (preproteins, e.g., 7.1.3, 7.1.5)
- Activation of zymogens (proproteins, inactive precursors of proteases, 7.1.9)
- Induction of a structural change to switch substrate specificity or gain a function, e.g., bacterial type III secretion systems (3.10.3).
- Participation in biological processes requiring cleavage of peptide bonds, e.g., coagulation (9.3), fibrinolysis (9.5), control of blood pressure (Fig. 7.1-9), complement activation (8.1.5), cell division (4.3.5), tissue disintegration and remodeling, apoptosis (7.6), sporulation, gene regulation, fertilization, tumor cell invasion, microbial infection processes etc.

Due to the wealth of their essential functions, peptidases are ubiquitous in all living cells and organisms.

4.5.6.1 Classification of Peptidases

- by Type of Catalyzed Reaction (Table 4.5.6-1): There are two groups of peptidases, comprising the exopeptidases (cleaving near the N- or C-terminus of the peptide) and the endopeptidases (cleavage of internal peptide bonds).
- by Cleavage Site Specificity (Table 4.5.6-2): Often, the cleavage specificity of peptidases is directed only towards the amino acid residue(s) immediately neighboring the cleavage site in the substrate protein. In other cases, the specificity depends on longer amino acid sequences ('restriction endoproteinases').
- by Catalytic Mechanism (Figs. 4.5.6-1 ... 3): Peptidases are classified according to their catalytic mechanism or the amino acids (or the ion) in the active center, which are involved in catalysis. There are currently four major categories of proteases which may be soluble or membrane-associated (Table 4.5.6-3).

The molecular mechanism for cleavage of a peptide bond involves the conversion of an amino acid residue (of serine, cysteine and threonine peptidases) or a water molecule (with aspartic and metallo peptidases)

Table 4.5.6-1. Groups of Peptidases

Group	EC-Subclass	Action (Symbols see below)	
Exopeptidases		N-terminus	C-terminus
Aminopeptidases	3.4.11	•-|-•-•-•-•…..	
Dipeptidyl-peptidases	3.4.14	•-•-|-•-•-•-•.....	
Tripeptidyl-peptidases	3.4.14	•-•-•-|-•-•-•-•.....	
Carboxypeptidases	3.4.16		-•.....•-•-•-•-•-|-•
Peptidyl-dipeptidases	3.4.15		-•.....•-•-•-•-•-|-•-•
Dipeptidases	3.4.13	•-|-•	
Omega Peptidases	3.4.19	☆-•-|-•-•-•-•..... or	.....•-•-•-•-•-•-|-•-☆
Endopeptidases (Proteinases)	3.4.21 ... 3.4.24, 3.4.99	.....•-•-•-•-•-|-•-•-•-•-•.....	

Symbols: •: acid residue, -: peptide bond, |: cleavage site, ☆ : residue bound in linkage other than peptide bond.

Table 4.5.6-2. Preferred Cleavage Specificity of Peptidases

Peptidases (Examples)	Cleavage Specificity (X: Any Amino Acid)
Trypsin (17.1.9)	...Lys-|-X... or ...Arg-|-X...
Chymotrypsin (17.1.9)	...hydrophobic-|-X..., e.g., ...Phe-|-X... or ...Tyr-|-X...
Thrombin (20.3)	...Arg-|-X...
Endoproteinase Glu-C from *Staphylococcus aureus V8*	...Glu-|-X... or ...Asp-|-X...
Asp-N-metallo-endoproteinase from *Pseudomonas fragi*	...X-|-Asp... or ...X-|-cysteic acid...
Coagulation Factor Xa (20.3)	...Ile-Glu-Gly-Arg-|-X...
Enteropeptidase (enterokinase)	...Asp-Asp-Asp-Asp-Lys-|-X...
Kexin (*yeast* KEX2 protease)	...(Lys or Arg)-Arg-|-X...

Table 4.5.6-3. Catalytic Mechanism of Peptidases

Class	Examples
Serine/threonine peptidases (Fig. 4.5.6-1)	trypsin, chymotrypsin, elastase, subtilisin, coagulation factors, t-PA, yeast carboxypeptidase Y, hepatitis C virus NS-3 protease, propyl oligopeptidase, ClpP, Lon, LexA
Cysteine peptidases (Fig. 4.5.6-1)	papain, bromelain, cathepsins B, H and L, calpain, hepatitis A virus 3C protease, bromelain, papain, cathepsins, caspases (important in apoptosis mechanisms)
Aspartate peptidases (Fig. 4.5.6-2)	pepsin, chymosinin, cathepsins D and E, HIV-1 protease, presenilins, signal peptide peptidase
Metallopeptidases (Fig. 4.5.6-3)	bacterial and mammalian collagenases, thermolysin, carboxypeptidase A, angiotensin, converting enzyme, S2P family

into an activated nucleophilic state, so that it can attack the carbonyl moiety of the substrate peptide bond. A possibility to generate a nucleophile is by a catalytic dyad (e.g., Lys and Ser) or triad (e.g., Asp, His and Ser), which enables the oxygen attack (or S attack in case of cysteine peptidases) on the carbonyl-C of the substrate. Formation of the tetraedric intermediate results in the consecutive cleavage of the C-N bond and this involves a concomitant proton binding to the substrate-N.

Many proteases are synthesized and secreted as inactive forms called zymogens and subsequently activated by proteolysis. This changes the architecture of the active site of the enzyme.

4.5.6.2 Reaction Mechanism of Serine Peptidases (Fig. 4.5.6-1)

More than one third of all known peptide bond hydrolases are serine peptidases. As an example for serine peptidases, the mechanism of chymotrypsin (an endopeptidase) is described. The reaction proceeds in two steps:

- Acylation step (upper line): The reaction starts with a nucleophilic attack of the serine 195-hydroxyl-O on the carbonyl-C of the substrate, while the proton is transferred to His 57. The resulting histidine cation is stabilized and held in a 'proton acceptor state' by Asp 102. Another transfer of this proton to the amide-N of the substrate causes the cleavage of the peptide bond. The amino component of the substrate is then released.
- Deacylation step (lower line): This reaction is a reversal of the acylation step, in which water plays the role of the amino component of the substrate, resulting in release of the carboxy component of the substrate.

The active site of serine proteases contains often the classical Ser-His-Asp triad and it is shaped as a cleft where the polypeptide substrate binds. Most serine proteases are sequence specific, and the enzyme is also able to cleave ester bonds by a similar mechanism. Its formation from an inactive precursor after secretion is described in 7.1.9. Regulatory serine peptidases contain an additional peptide chain, which interacts with cofactors essential for regulation. Trypsin-like serine proteases (clan PA peptidases) form the major constituent in the eukaryotic degradasome, whereas subtilisin-like (clan SB peptidases) and prolyl oligopeptidases (clan SC peptidases) are the most dominant serine proteases of prokarya, fungi and plants.

Other important serine peptidases use a dyad of Ser and Lys (LexA, Lon protease) as endoproteolytic catalytic machinery and are classified as clan SF and clan SJ peptidases, respectively. Unique to the SJ type proteases is that they depend on ATP and hydrolyze ATP during proteolysis of which the rate limiting step seems to be the ATP-dependent release of the product from the active site.

4.5.6.3 Reaction Mechanism of Cysteine Peptidases

The mechanism of cysteine peptidases (e.g., papain, cathepsin) is similar to serine peptidases, with cysteine-S^- and protonated histidine replacing the serine-OH and the nonprotonated histidine.

4.5.6.4 Reaction Mechanism of Aspartate Peptidases (Fig. 4.5.6-2)

The aspartic proteases utilize two adjacent aspartate residues in their active center. As a representative for aspartate peptidases, the catalytic mechanism of pepsin (an endopeptidase) is described. It depends on the γ-carboxyl groups of two aspartate residues in the active center (Asp 32 and Asp 215), one of them is dissociated. This environment activates a water molecule, the oxygen of which performs a nucleophilic attack on the carbonyl-C of the substrate, while the proton released from the water combines first with the ionized carboxyl of aspartate and then with the amide-N of the substrate, resulting in peptide bond cleavage. The other, initially undissociated aspartate carboxyl group acts first as proton donor and thereafter as proton acceptor. In contrast to the serine proteases, no covalent enzyme-substrate intermediate is formed.

4.5.6.5 Reaction Mechanism of Metallopeptidases (Fig. 4.5.6-3)

Some proteases require cations (e.g., Ca^{++}, Zn^{++}) for their proteolytic activity. Similar to the aspartate peptidases, these enzymes activate water directly and do not form covalent enzyme-substrate intermediates. Carboxypeptidase A is the prototype of a Zn^{++}dependent peptidase, which all share a zinc binding motif of the HXXE … H type. Zinc ligands are a histidine and a glutamate residue separated by two amino acids and a second histidine moiety located about 108–135 amino acids towards the C-terminus from the first histidine. The zinc binding residues in carboxypeptidase A are His69, Glu72 and His96.

Two alternative hypotheses exist for the catalytic mechanism of carboxypeptidase A:

- acyl pathway hypothesis with a covalent acyl enzyme intermediate or
- the promoted water pathway which involves a Zinc-promoted polarization of the substrate carbonyl group and activation of a bound water molecule.

Most experiments are currently in favor of the latter hypothesis. The activation of water enables the nucleophilic attack of its oxygen on the

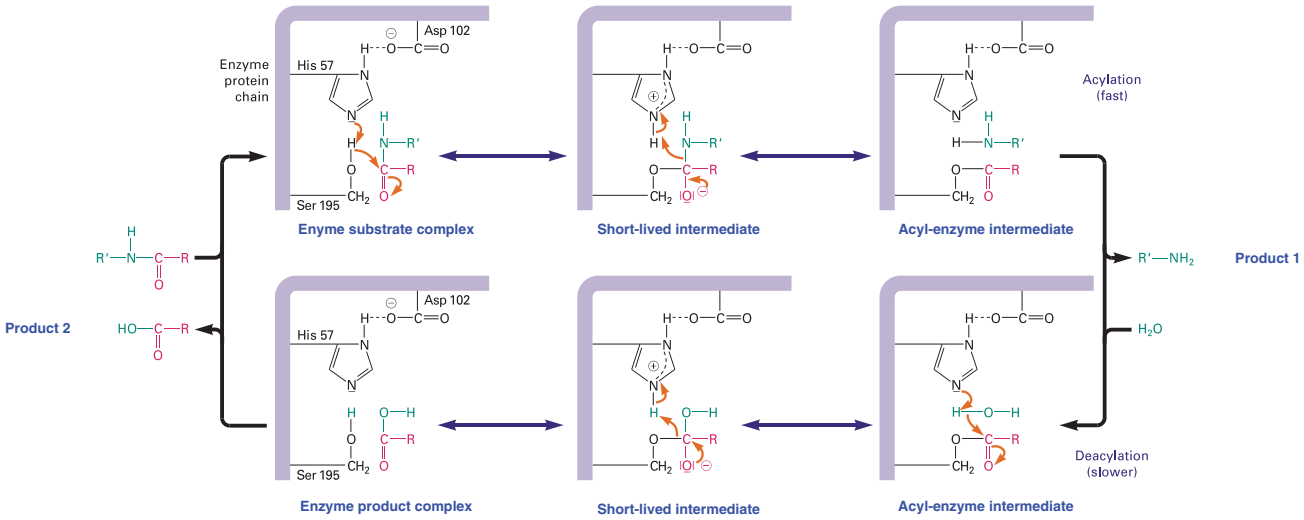


Figure 4.5.6-1. Reaction Mechanism of Chymotrypsin (Example of a Serine Peptidase)

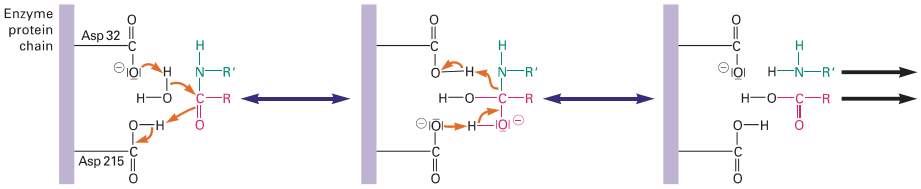


Figure 4.5.6-2. Reaction Mechanism of Pepsin (Example of an Aspartate Peptidase)

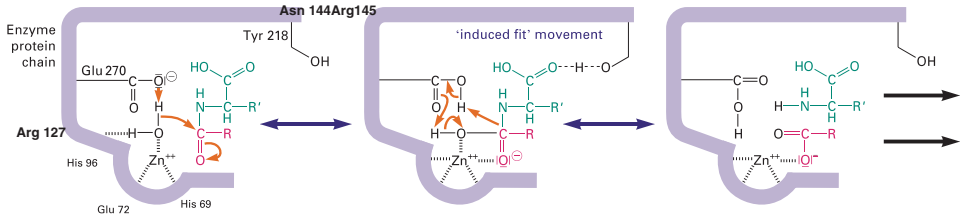


Figure 4.5.6-3. Reaction Mechanism of Carboxypeptidase A (Example of a Metallopeptidase)

carbonyl-C of the substrate. Several amino acids have been identified as being involved in substrate binding and proteolysis by the crystal structure of the molecule: S1´: Asn144, Arg145, Tyr248 (fixation of the terminal carboxylate group of the peptide), S1: Arg127, Glu270 (positioning of the scissile peptide bond), S2: Arg71, Ser197, Tyr198 and Ser199, S3: Phe279 and S4: Glu122, Arg124 and Lys128.

4.5.6.6 Peptidase Inhibitors

The function of peptidases is often regulated by specific inhibitors. Many natural (selected) and synthetic (designed) inhibitors are given in the BRENDA (http://www.brenda-enzymes.org) and the MEROPS database (http://merops.sanger.ac.uk). They are either active-site specific, which cause irreversible modifications, or special proteins, which act as competitive inhibitors (most natural inhibitors are of this type). The different inhibitory proteins are generally classified by the type of protease they attack (e.g., serine protease inhibitors: serpins) or by their inhibition mechanism. Unspecific inhibitors are, e.g., macroglobulin (acting by steric hindrance) or peptide aldehydes (non-cleavable peptide bond analogues). Class specific for serine peptidases are, e.g., organophosphates, coumarin derivatives (irreversible inhibition), α_1 proteinase inhibitor (α_1 antitrypsin, 9.3.4), antithrombin, α_1-antichymotrypsin, plasminogen activator inhibitor 1 (coagulation, fibrinolysis) and neuroserpin. Pepstatin (a bacterial peptide with an unusual amino acid) inhibits the aspartate peptidase class and metal chelators inhibit metallopeptidases. Aprotinin (Trasylol®), soybean trypsin inhibitor and serpins form tight, but reversible complexes with individual peptidases.

Use of peptidase inhibitors have frequently been proven to be useful for the treatment of different diseases, including cancer and infections. For example, special HIV-1 protease inhibitors (e.g., Saquinavir®, Ritonavir®, Indinavir®) were developed that target the active site, a long cylindrical cavity forming the substrate-binding groove of the HIV-1 protease. The lipocalin protein family are natural protease inhibitors which play a role in cell regulation and differentiation. Lipophilic ligands which are attached to these proteins have been shown to reduce tumor protease activity.

4.5.7 Protein Degradation by the Ubiquitin-Proteasome System (UPS, Fig. 4.5.7-1, Table 4.5.7-1)

This protein degradation system provides controlled proteolysis of proteins in all eukarya and exerts important regulatory functions in cell division (4.3.6), quality control of newly synthesized proteins (ERAD), DNA replication (3.9.1), transcription, receptor function (7.1.2), cell differentiation, apoptosis (7.6), immune response or pathologic processes including cancer, diabetes, arteriosclerosis, inflammatory responses and neurodegenerative diseases and aging. Structure and function of the proteosomal system is strongly conserved between fungi, plants and animals. It is a large multisubunit and multicatalytic degradation complex which is localized in the nucleus and in the cytosol.

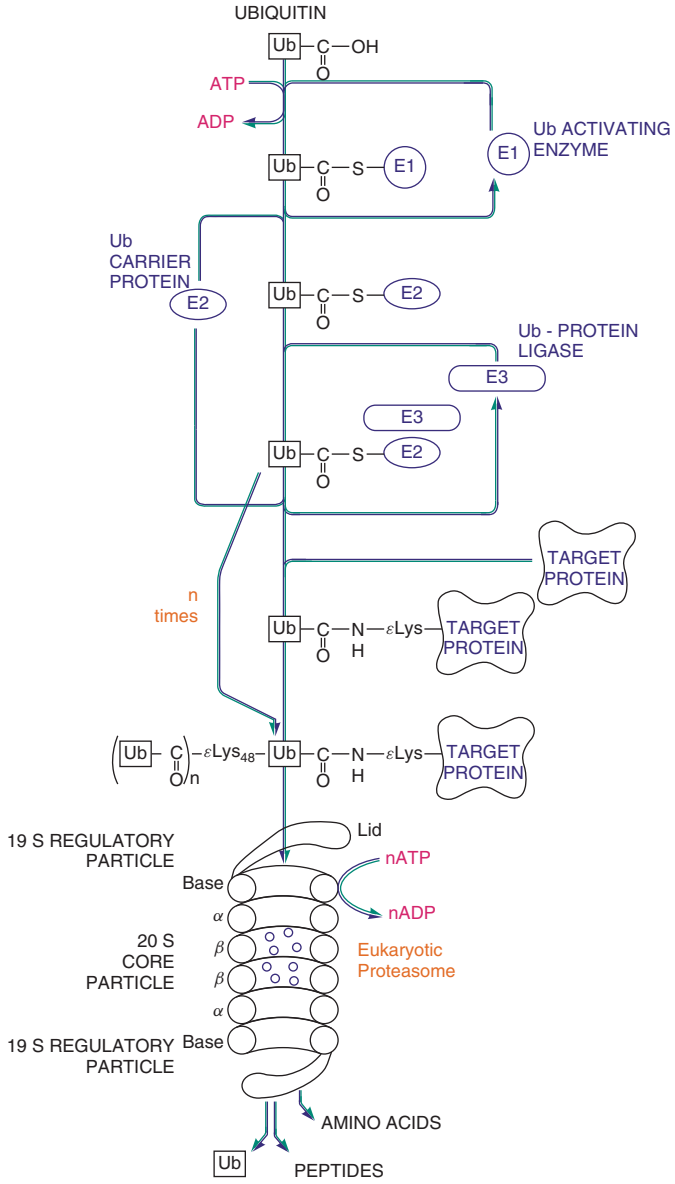


Figure 4.5.7-1. Regulated Proteolysis by the Ubiquitin System

Table 4.5.7-1. Enzymatic Reaction Partners Associated with Protein Degradation

Ubiquitylation	Components	Function
E1	Ube1, Ube12L	Ubiquitin activation
E2	UbcH1, UbcH2, UbcH3, UbcH4, UbcH5, UbcH6, UbcH7, UbcH8, UbcH10, UbcH13, Use1 etc.	Ubiquitin conjugation
E3	Anaphase promoting complex (APC), Cdc34, Ub-Camodulin ligase, UBR5 (EDD1), SUMO-dependent RNF4, mdm2, SOCS, E3A etc.	Ubiquitin ligase
E4	Hul5, UFD2a etc.	Ubiquitin ligase
Ddi2		Adaptor molecule
Dks2		Adaptor molecule
Rad23		Adaptor molecule
Usp/Ubp proteins	ubiquitin-specific processing protease superfamily	deubiquitylation enzyme (DUB)
Uch proteins	ubiquitin C-terminal hydrolyase superfamily	deubiquitylation enzyme (DUB)
Otu proteins	ovarian tumor (OTU) superfamily	deubiquitylation enzyme (DUB)
MjD proteins	Machado-Josephin domain (MJD) superfamily	deubiquitylation enzyme (DUB)
Cdc48/p97	conserved chaperone-like ATPase	Ubiquitylation platform

Ubiquitylation: Covalent attachment of multiple ubiquitin (Ub, 8.5 kDa) molecules, which are the most highly conserved eukaryotic proteins, function as a signal for proteolytic degradation by the proteasome. Ubiquitylation is mediated by an orchestrated action of several enzymes, including E1 (ubiquitin activation), E2 (ubiquitin conjugation) and E3 (recognition, transfer to the substrate and ubiquitin ligation). Different types of theses enzymes can be mobilized for the ubiquitylation process, e.g., two E1, more than 30 E2 and 500 E3 variants are present in humans and promote an accurate selection of the protein substrate.

Thioester-activated ubiquitin is almost always bound to the ε-amino group of a lysine residue in the protein to be degraded, forming an isopeptide. In a few cases conjugation occurs to N-terminal residues or cysteines. Polyubiquitin chains are frequently formed by reaction of the Ub C-terminal glycine with an internal lysine of another Ub molecule (in most cases Lys48, but also Lys11, Lys63 and others are possible). Many E3 and additional ubiquitin ligases (E4 enzymes) and DUBs (see below) can interact on a ring-shaped complex formed by Cdc48 ATPases (p97 in mammals) which acts as an ubiquitylation platform. It has also been shown that small ubiquitin-like protein modifiers (SUMO, which are covalently linked to Lys residues) can act as a signal for the recruitment of E3 ubiquitin ligases. Ubiquitylation is reversible and can by induced by cysteine proteases and metalloproteases [deubiquitylation enzymes (DUBs, e.g., Rpn11, Ubp6, Uch37) in yeast]. They are responsible for the maturation and the recycling of ubiquitin after protein degradation. In addition, the DUBs prevent blocking of the proteosomal complex in the case of degradation problems with a bound protein substrate. Most ubiquitylation and deubiquitylation enzymes are soluble but they can also form larger aggregates which act as protein-targeting machines.

Table 4.5.7-2. Proteasome Subtypes and Composition – Components of the UPS System

Catalytic 20S Proteasome	20S core particle CP, composed of:	2 outer α-rings (α1–α7) 2 inner β-rings (β1–β7)
Regulator	19S regulatory particle RP (PA700), composed of:	a) lid proteins: (Rpn 3, 5, 6, 7, 8, 9, 11, 12, 15) b) base proteins:(Rpt1-6), (Rpn1, 2, 10, 13)
Active proteasomes	30S proteasome	RP-CP-RP (19S-20S-19S)
	26S proteasome	RP-CP (19S-20S)

Proteasome (Fig. 4.5.7-1, Table 4.5.7-2): The enzymatically active proteasomes of eukarya consist of more than 30 subunits which form different subcomplexes:

- a barrel-shaped catalytic 20S core particle (CP) as a stack of two heptameric outer rings of α subunits and two inner rings composed of β subunits
- a 19S regulatory particle (RP or PA700) that contains multiple ATPase subunits. The RP recognizes the ubiquitylated target proteins and promotes unfolding and translocation of the substrate into the central cavity of the CP. In humans, the RP consists of the lid subcomplex composed of 9 non-ATPase subunits and the base subcomplex of 6 homologous AAA-ATPase proteins and 4 non-ATPase subunits.

The 26S and 30S proteasomes are found in all cells and are mainly responsible for degradation of regulatory and damaged proteins. They consist of the 20S CP and 19S subunits attached to one or both ends of the central core complex. However, different subtypes of the

proteasome complex also exist, which perform slightly different functions. Among them is the immunoproteasome (large multifunctional proteasome, 8.1.7) that is responsible for the production of major histocompatibility complex (MHC) class 1 ligands displayed on the cell surface.

The proteasomes of prokarya show a very similar structure. They consist of a central multimeric proteolytically active protease complex (ClpP) which is capped by a ring of ATPase molecules (e.g., ClpA or ClpX in *E. coli*).

Delivery and degradation: Ubiquitylated proteins are carried to the proteasome by certain adaptor molecules (Rad23, Dsk2, Ddi1) which interact with the polyubiquitin chains and RP subunits at the initiation site for degradation. In addition, some E3 enzymes (e.g., Ufo-1) can mediate binding to some adaptors (Ddi1) and to the proteasome. Target protein recognition by the proteasome complex system (and hence the degradation rate) also depends on a degradation initiation site (usually an unstructured region) located at the termini or within internal sequences of the protein (see also 4.3.6). It is likely that the initiation site binds to loops within the RP ATPase ring which acts as a translocation motor and pulls the substrate into the ring. This leads to an unfolding and translocation of the substrate into the CP complex where degradation occurs. Ubiquitylation-independent proteolysis of some substrates is also possible (e.g., ornithine decarboxylase, p53, c-Jun etc.). In these cases, unstructured regions within these proteins can function as very effective initiation sites with very high affinity to certain subunits of the RP complex (e.g., Rpn5, Rpn10, or Rpn13).

Although some bacteria use certain tags to modify proteins for degradation, bacterial proteins are usually targeted for proteolysis through signals located within their N- or C-terminal amino acid sequences. Furthermore, some adaptor molecules (e.g., SspB, ClpS and RssB in *E. coli*) are known which are used for protein degradation in prokaryotes.

Literature:

Barrett, A.J. *et al*. Biol. Chem. 2007;388:1151–1157.

Butler, G.S., Overall, C.M. http://www.ncbi.nlm.nih.gov/pubmed?term= Biochemistry 2009;48:10830–10845.

Deshaies, R.J., Joazeiro C.A.P. Ann. Rev. Biochem. 2009;78:399–434.

Jung, T. *et al*. Mol. Asp. Med. 2009;30:191–296.

Page M.J. *et al.* Cell Mol. Life Sci. 2008;65:1220–1236.

Rawlings, N.D. *et al.* Nucleic Acids Res. 2010;38:227–233.

Schrader, E.K. *et al.* Nat. Chem Biol. 2009;5:815–822.

Shindo, T., van der Hoorn, R.A. Mol. Plant. Pathol. 2008;9:119–125.

Tanaka, K. Proc. Jap. Acad. Ser. B 2009;85:12–36.

Viestra, R.D. Nat. Rev. Mol. Cell. Biol. 2009;10:385–397.

Wolfe, M.S. *et al.* Chem. Rev. 2009;109:1599–1612.

Zavasnik-Bergant, T., Turk, B. Biol. Chem. 2007;388:1141–1149.

5 Viruses

Klaus Klumpp

5.1 General Characteristics of Viruses

Viruses are infectious agents that require living host cells to replicate themselves. They contain nucleic acid genomes and form protective shells of protein or lipids and carbohydrates to spread between cells and hosts. A complete, free infectious virus particle existing outside of a cell is defined as a virion. While the evolutionary origin of viruses is uncertain, they are ubiquitous and, with an estimated number of 10^{31}, constitute the most abundant type of biological entity on earth. After the discovery of the first virus, Tobacco Mosaic Virus (TMV) in 1898, more than 5000 types of viruses have been extensively studied and described, and it is expected that even more remain undiscovered.

Viruses can infect all types of cellular living organisms, including animals, plants, fungi, bacteria and archaea. They redirect the metabolism of infected host cells to perform their own reproduction using a minimum of genetic information, ranging from as few as 4 genes for the simple RNA viruses MS2 and Qβ to as many as 911 genes for the largest virus, Acanthamoeba Polyphaga Mimivirus (APMV, Mimivirus). In addition to structural proteins to form mature, infectious virions, most viruses also encode certain non-structural proteins and enzymes that support replication of the virus's genetic material inside infected host cells. Most viruses range between 10 and 400 nanometers in diameter, approximately $1/10^{th}$ to $1/100^{th}$ of the size of bacteria.

Still smaller are viroids, infectious agents of plants, which consist of single stranded, structured, covalently closed RNA of approximately 250–400 nt. Viroid RNAs do not code for any specific protein but may have enzymatic activity (RNA cleavage) and are able to replicate themselves in the nuclei of infected plant cells by using host proteins (including RNA polymerase II, 4.2.1.1). Viroids transmit via infected pollen and seeds, via *trans*-encapsidation by carrier viruses, via insects (e.g., aphids), or mechanically (e.g., goat horns, contaminated tools). Pathogenic effects by viroids are likely caused through RNA interference mechanisms. Examples of economically detrimental viroids are Avocado Sunblotch Viroid (ASBVd) and Coconut Cadang-Cadang Viroid (CCCVd).

Prions are another unique form of ultramicroscopic infectious agents. They represent infectious proteins and are not capable of replication, which differentiates them from viruses and viroids. They constitute misfolded forms of host proteins and do not contain any genetic material. The misfolded prion proteins are capable of inducing misfolding of related host proteins. A small amount of prion protein as an infectious agent can induce misfolding of large amounts of host protein into stable, polymeric structures, called amyloid fibers. Amyloid fiber formation is a hallmark of fatal neurodegenerative prion diseases, collectively called Transmissible Spongiform Encephalopathies (TSEs).

5.1.1 Genomic Characteristics of Viruses

The first complete sequence for an RNA virus was that of bacteriophage MS2, obtained in 1976, followed in 1977 by the first complete sequence of a DNA virus, bacteriophage ΦX174. Since that groundbreaking work, an enormous variety of genomic structures have been observed among different virus families.

As a group viruses encompass more structural genomic diversity than animals, plants, bacteria or archaea. Based on genome structure and mode of replication, viruses can be divided into DNA-, RNA- and retroviruses. The genomic nucleic acids can be single- (ss) or double-stranded (ds). Viruses of the *Hepadnaviridae* family (Hepatitis B Virus and Duck Hepatitis B virus) are unique as they contain a genome that is partially double-stranded and partially single-stranded, and they replicate through an RNA intermediate, similar to retroviruses.

Viruses with ssRNA genomes can be subdivided into positive-strand (+) and negative strand (−) viruses. RNA of (+) viruses represents the coding strand of the genome, while that of (−) viruses represents the non-coding complementary sequence. Therefore, (+) strand RNA is ready for translation after infection of host cells, whereas (−) strand RNA requires at least one round of RNA replication intracellularly

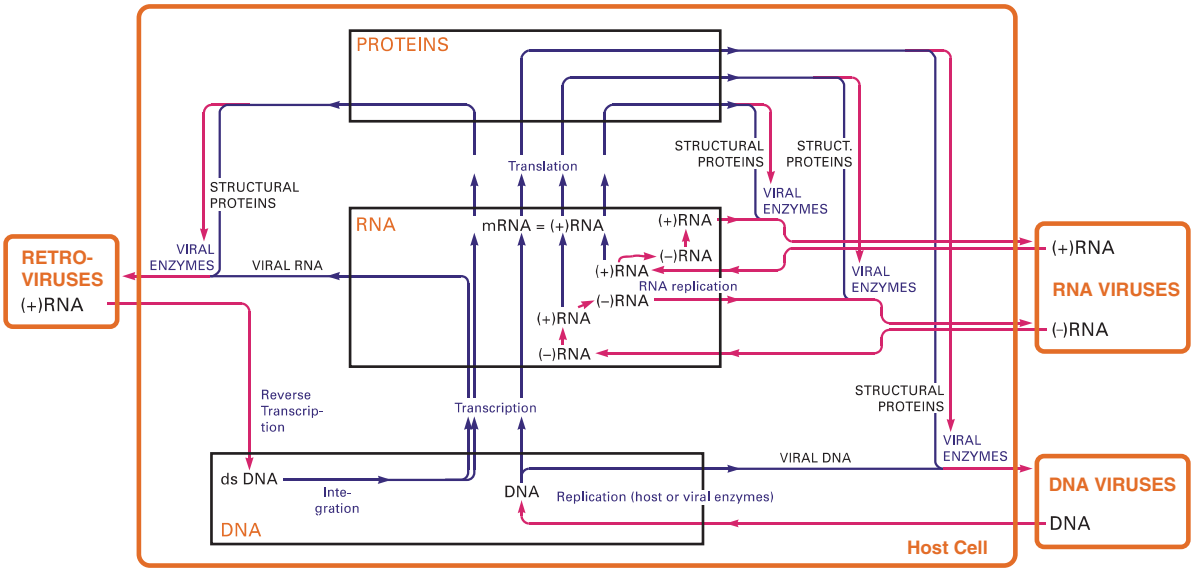


Figure 5.1-1. Flow of Information in Virus Reproduction
(red arrows = virus reactions, blue arrows = host reactions)

Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Second Edition. Edited by Gerhard Michal and Dietmar Schomburg.

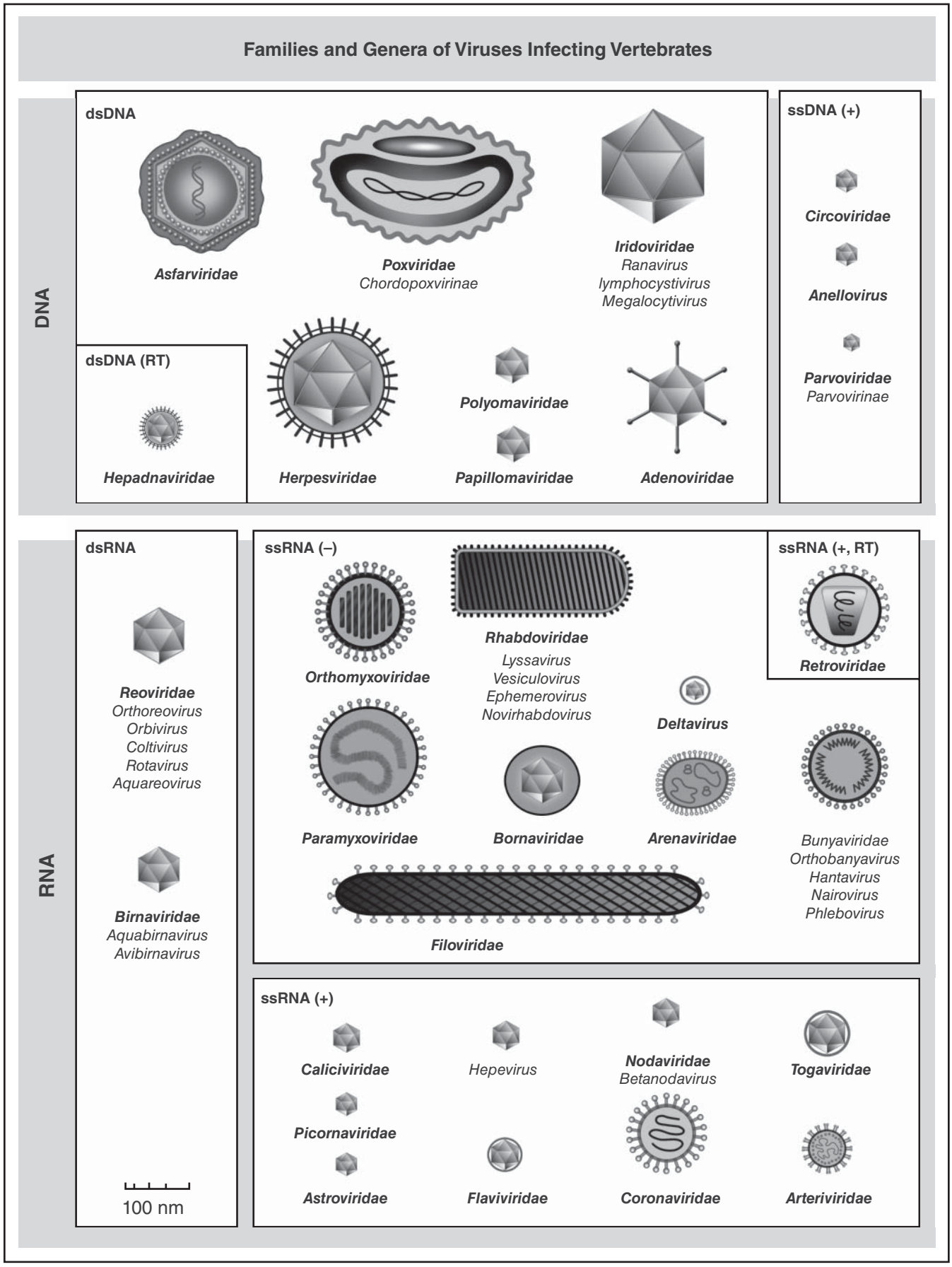


Figure 5.1-2. Diagrams of the Vertebrate Viruses
Diagrams are drawn to relative scale and grouped by genome type or replication strategy (reverse transcription). (after Knipe, D.M. *et al.*: Fields Virology 5th Ed. 2007)

prior to viral protein synthesis. ssDNA viruses are exclusively (+) strand viruses.

Accordingly, seven groups of viruses can be distinguished:

I. dsDNA viruses (e.g., Adenoviruses, Herpesvirus, Poxvirus, Papillomaviruse)
II. ssDNA viruses, (+)sense DNA, (e.g., Parvovirus)
III. dsRNA viruses (e.g., Reovirus)
IV. (+)ssRNA viruses, (+)sense RNA (e.g., Picornavirus, Flavivirus)
V. (−) ssRNA viruses, (−)sense RNA (e.g., Influenza Vírus, Rabies Virus)
VI. ssRNA retroviruses, (+)sense RNA with DNA intermediate in life-cycle (e.g., HIV)
VII. dsDNA viruses, replicated via pre-genomic RNA intermediate (e.g., Hepatitis B Virus)

The principle of virus reproduction for the different genomic classes is schematically shown in Figure 5.1-1.

The International Committee on Taxonomy of Viruses (ICTV) categorizes viruses into a single, genomics based classification scheme that reflects their evolutionary relationships, i.e. their individual phylogenies. The 2008 ICTV taxonomy places virus species in 5 orders, 82 families, 11 subfamilies and 307 genera.

Figure 5.1-2 provides an overview of vertebrate viruses according to their genomic structure, morphology and relative size. Due to the large variety of viruses, it is only possible to discuss a representative example in each of the following sections.

5.1.2 Structure

Usually, the genomic nucleic acid of a virion is surrounded by a protective coat of protein (capsid). This is often formed from multiple copies of one or a few proteins that are assembled into elongated helical cylinder, quasi-spherical icosahedron or conical shapes. Many viruses also acquire an outer membrane envelope during budding from the host cell. This membrane then contains proteins coded for by the viral genome, while the lipid membrane itself and any carbohydrates present originate entirely from the host. More complex capsid structures also exist. Poxviruses for example form complex nucleoid structures from a larger number of proteins surrounding the genomic DNA. In many cases, viral genome encoded enzyme proteins and host proteins are present within the virion and can be associated with the genomic nucleic acid. A few of the most important families of viruses are listed in Table 5.1-1.

Viruses and their hosts have co-evolved over many years, resulting in complex interactions. Host organisms have developed a variety of antiviral defense mechanisms to inhibit viral replication, e.g., restriction endonucleases in bacteria (4.1.4.2), basic immune systems including innate immune response in plants and invertebrates, and complex immune systems including adaptive immune response in vertebrates (3.6.1). Viruses have thus evolved strategies to replicate under the restrictive conditions provided by their hosts and to escape immune control mechanisms. A number of viral encoded proteins are therefore aimed at blocking immune sensing and immune stimulatory pathways and to interfere with the host antiviral defense systems (see e.g., Fig. 5.4-4).

Literature:

ICTV taxonomy: http://www.ictvonline.org.

Aguzzi, A., Calella, A.M. Physiol. Rev. 2009;89:1105–1152.

Breitbart, M., Rohwer, F. Trends. Microbiol. 2005;13:278–284.

Knipe, D.M., Howley, P.M. *et al.* (Eds.) *Fields Virology,* 5th Ed. Philadelphia: Lippincott, Williams & Wilkins, 2007.

Tsagris, E.M., Martinez de Alba, A.E. *et al.* Cell Microbiol 2008;10:2168–2179.

Table 5.1-1. Important Viruses and Their Families

Genome, Strands	Family	Species (Examples)	Genome Size (ca. kb), Structure	Virion Size (nm)	Enveloped	Morphology	Host	Disease
ds DNA	*Papillomaviridae*	Papillomavirus	9; closed circle, supercoiled	55	–	icosahedral	mammals	warts, associated with cervical cancer
	Adenoviridae	Adenovirus	36; linear	70–90	–	icosahedral	mammals, birds	respiratory infections
	Herpesviridae	Herpes simplex virus	200; linear	200	+	icosahedral	vertebrates	genital herpes, chicken pox
	Poxviridae	Vaccinia virus	200; linear, sealed ends	300	+	complex	vertebrates	cowpox
	Myoviridae (T4-like Bacteriophages)	T4	167 linear	65–95 (head); 17*115 (tail)	– –	icosahedral (head), helical (tail), 6 tail fibers	*E. coli*	bacteriolysis
	Siphoviridae (λ-like Bacteriophages)	λ	48, linear, cohesive ends	55 (head); 15*135 (tail)		icosahedral (head), helical (tail)	*E. coli*	bacteriolysis
(+)ss DNA	*Parvoviridae*	B19	5; linear	18–26	–	icosahedral	humans	erythema, anemia
	Microviridae (Bacteriophages)	ΦX174	5.4; circular	25	–	icosahedral	*E. coli*	bacteriolysis
ds RNA (+)ss RNA	*Reoviridae*	Rotavirus	25; 11 linear segments	60–80	–	Icosahedral	humans	gastroenteritis, infectious diarrhea
	Flaviviridae	Hepatitis C virus	10; linear	50	+	spherical	humans	Hepatitis C
	Picornaviridae	Poliovirus	7.4; linear	30	–	icosahedral	humans	poliomyelitis
	Leviviridae (Bacteriophages)	Qβ	4.2	26	+	icosahedral	*E. coli*	bacteriolysis
	Togaviridae	Semliki forest virus	13	60 – 70	–	spherical	vertebrates	encephalitis
	Unassigned	Tobacco mosaic vir.	6.4; helical	300*18		tubular rod	plants	Tobacco mosaic virus disease
(−)ss RNA	*Orthomyxoviridae*	Influenza virus	12–15; 8 linear segments	50–120	+	spherical or tubular	mammals, birds	Influenza
	Paramyxoviridae	Mumps virus	15; linear	75 *180	+	pleomorph.	humans	mumps
	Rhabdoviridae	Rabies virus	12	50–300	+	bullet shape	mammals	rabies
	Arenaviridae	Lassa fever virus	10–11; 2 linear segments		+	pleomorph.	humans	Lassa fever
(+)ss RNA→ DNA	*Retroviridae*	HIV	9.2; linear segments, diploid	100	+	conical capsid in spherical envelope	humans	AIDS
ds DNA-RNA[1]	*Hepadnaviridae*	Hepatitis B virus	3; circular, partially single-stranded	42	+	spherical	humans	Hepatitis B

[1] Replication via a RNA intermediate, which undergoes reverse transcription.

5.2 DNA Viruses

DNA viruses carry either double-stranded DNA (dsDNA) or single-stranded DNA (ssDNA) genomes. The central enzyme for virus replication is a DNA-dependent DNA polymerase. The generation of viral proteins proceeds via mRNAs, usually employing the host cell translation machinery.

Most viruses with single-stranded DNA genomes expand those to double-stranded DNA in infected cells in the process of viral replication. The replication of the DNA virus genome can occur in many ways, depending on the DNA structure. If the DNA has a circular structure, replication mechanism may be by the Θ- or by the σ- (rolling circle) mechanism, similar to the mechanism of bacterial genome replication. Figure 3.8.1-1 shows this mechanism with the example of bacteriophage T4 and ΦX 174 viruses, respectively. If the viral genome consists of linear DNA it will need a priming step for autonomous replication. Viruses with linear DNA genomes may have proteins attached to the DNA termini, which interact with the DNA polymerase (e.g., adenovirus). DNA viruses can also have their genomic DNA ends covalently sealed by a loop structure (e.g., poxvirus), while others have ends that interact through complementary sequences (so-called 'cohesive' ends), which are sealed into a loop after entering the host by a host enzyme (e.g., bacteriophage λ). The viral DNA can also be incorporated into the host genome. In multicellular organisms, this can cause cell transformation and tumorigenesis by interference with regulatory mechanisms of host gene expression (see 5.2.1, Papillomavirus and DNA tumor viruses, 7.5.1).

5.2.1 Papillomavirus

Papillomavirus is an example of a compact dsDNA virus that encodes only one viral enzyme (DNA helicase, E1) and thus uses host cell enzymes and processes for the largest part of viral replication. Several hundred species (types) of papillomaviruses form the family *Papillomaviridae* of non-enveloped DNA viruses that infect mostly mammals, birds and certain reptile species. Figure 5.2-1 shows an overview of human (HPV) and animal papillomavirus types known at present.

Papillomavirus virions are icosahedral structures with a diameter of 55 nm (Fig. 5.1-2). The double-stranded, circular DNA genome of papillomaviruses is approximately 8 kb in size and can be divided into three major regions: an early, a late, and a long control region (LCR or noncoding region [NCR]). The three regions in all papillomaviruses are separated by two polyadenylation (pA) sites: early pA (A_E) and late pA (A_L) sites (Fig. 5.2-2). The capsid consists of 360 copies of L1 protein arranged in 72 capsomeres that are associated with probably 12 copies of L2.

Transcription of **papillomavirus genes** is complex, involving splicing of bicistronic or polycistronic forms, polyadenylation at one of two poly(A) sites and regulation at the level of transcription (such as via DNA methylation) as well as at post-transcription steps (including RNA splicing, polyadenylation, and translation).

The early region of the papillomavirus genome occupies over 50% of the virus genome and covers eight designated open reading frames (ORFs) (E1–E8, see Table 5.2-1). The late region of papillomavirus genomes occupies approximately 40% of the virus genome and encodes L1 and L2 ORFs that form a major (L1) and a minor (L2)

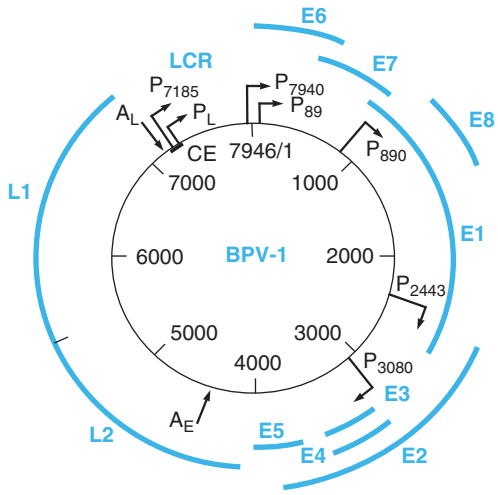


Figure 5.2-2. Genome Organization of Papillomavirus
(after Knipe, D.M. *et al.* (Eds.) *Fields Virology* 5th Ed. 2007)

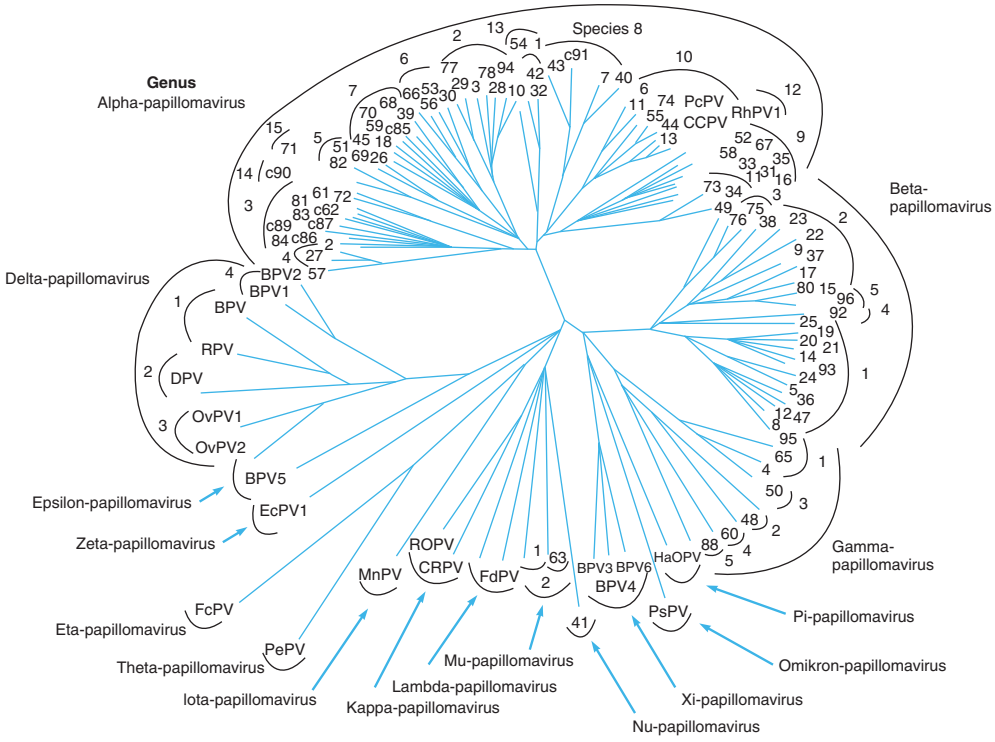


Figure 5.2-1. Phylogenetic Tree of Papillomaviruses Based on L1 ORF
(after de Villers, E.M. *et al.*: Virology 2004;324:17–27)

Table 5.2-1. Papillomavirus Proteins

Name	Expression	Function	Effects of protein
E1	Early	Helicase, part of the DNA replication machinery.	Needed for DNA upkeep as an episome Defect: DNA integration in host genome
E2	Early	Binds to the HPV upstream regulatory region, recruits E1 helicase to origin of replication	Needed for DNA upkeep as an episome Defect: Deregulated expression of E6 + E7 -> cell transformation, neoplasia
E3	Early	Function unknown	
E4	Early	Involved in DNA replication	
E5	Early	Involved in DNA replication	
E6	Early	Regulator, enables upkeep of infected cells	Binds pro-apoptotic proteins p53, Bax, Bak, prevents apoptosis of host cell, predisposing factor for cancer development
E7	Early	Regulator, enables DNA replication, contains the differntiation-dependent promoter within the ORF.	Binds to Rb, releases transcription factor E2F → expression of proteins (e.g., cyclin E) for virus replication, overrides S-phase checkpoint (4.3.6)
L1	Late	Capsid protein, 360 copies in virion. Contains splicing silencer, which leads to preferential synthesis of early transcripts in proliferating cells	Interaction with cell surface proteoglycan causes virus entry in cell. Assembles into capsomeres in plasma, which enter the nucleus for capsid formation
L2	Late	Minor capsid protein, approximately 12 copies in virion	Proteolytic cleavage enables virus escape from endosome. Recruits L1 for virus assembly / packaging

capsid protein. The LCR region, a segment of about 850 bp (10% of the HPV genome), does not encode any protein, but covers the origin of replication as well as multiple transcription factor binding sites that are important in regulation of host cell RNA polymerase II-initiated transcription from viral early and late promoters.

Life cycle – cell entry and establishment of papillomavirus infection: The mechanism of papillomavirus entry into its target cells, basal epithelial cells, is still not very well understood. Cell surface proteoglycans are involved in early interactions with the capsid protein L1 but interaction with at least one specific receptor is also required to prevent non-productive virus uptake. Cell entry proceeds through endocytosis in clathrin-coated vesicles (6.1.5). The minor capsid protein L2 is required to allow virus escape from this endosomal compartment in a process that is dependent on proteolytic cleavage of L2.

The viral DNA then has to reach the cell nucleus, where the virus typically maintains its genome as a low copy number (10–200) episome (2.2.1, independent from the host DNA), together with a low level expression of viral early proteins. In particular, E1 and E2 proteins are needed in order to maintain the viral DNA as an episome and to facilitate the correct segregation of genomes during cell division.

HPV DNA can also integrate into the host cell genome, which in most cases leads to loss of episomal DNA and loss of E2 function. Defects in the expression of E1 protein may initially facilitate the integration of viral genomes into the host cell chromosome. The consequence of E2 function loss is deregulated expression of the viral oncogenes E6 and E7, contributing to cell transformation and neoplasia. High-risk HPV type 18 appears to progress particularly fast towards integration into the host genome and loss of E2 function.

Replication phase (Fig. 5.2-3): In uninfected host epithelium, basal cells exit the cell cycle soon after migrating into the suprabasal cell layers and undergo a process of terminal differentiation. However, Papillomavirus requires dividing basal cells to allow its replication and multiplication. It therefore encodes proteins E6 and E7 that abolish the restraint on cell cycle progression and delay the normal terminal differentiation. Both E6 and E7 can associate with regulators of cell cycle to promote cell cycle progression. E7 binding to the negative regulator protein Rb displaces transcription factor E2F (Table 4.3-1) and leads to the expression of proteins required for DNA replication. The viral E6 protein complements the role of E7, and is thought to prevent the induction of apoptosis in response to unscheduled S-phase entry mediated by E7. E6 is known to associate with pro-apoptotic proteins such as p53, Bak and Bax to inactivate growth suppression and/or apoptosis (4.3.6). This process of deregulated cell division in basal cells and blocking of tumor suppressor genes provides an environment for optimal viral

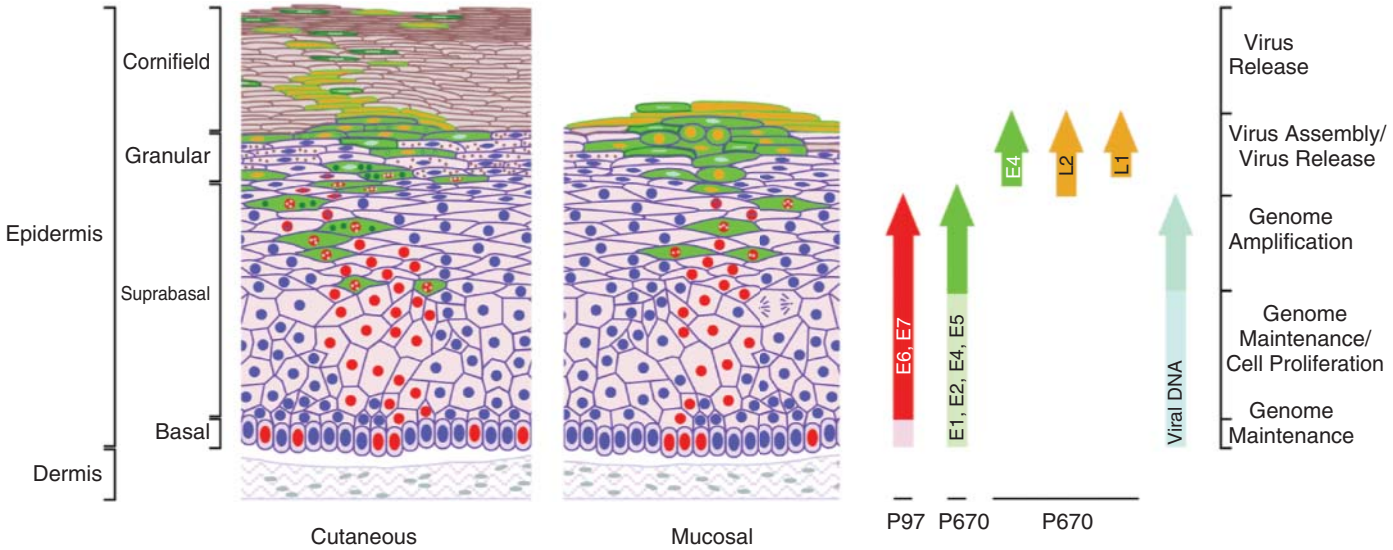


Figure 5.2-3. HPV-16 Gene Expression and Life Cycle is Linked to the Differentiation of Epithelial Cells
(after Doorbar, J.: Clinical Science 2006;110:525–541)

Table 5.2-2. Association of Alpha Papillomavirus Types with Cervical Cancer (after Knipe, D.M., Howley, P.M. *et al.,* see below)

Genus + Species	Type Species	Squamous cell Carcinoma Cervix	Adenocarcinoma Cervix	Category
Alpha 1	HPV32 , 42			low risk
Alpha 2	HPV3, 10,28, 29, 77, 78, 94			cutaneous
Alpha 3	HPV61, 72, 81, 83, 84			low risk
Alpha 4	HPV2, 27, 54			cutaneous
Alpha 5	HPV26	0.22 %		high risk
	HPV51	0.75 %	0.54 %	high risk
Alpha 6	HPV53	0.04 %		high risk
	HPV 56	1.09 %		high risk
Alpha 7	**HPV18**	**11.27 %**	**37.30 %**	**high risk**
	HPV45	**5.21 %**	**5.95 %**	**high risk**
	HPV59	1.05 %	2.16 %	high risk
	HPV39	0.82 %		high risk
	HPV68	0.37 %		high risk
Alpha 8	HPV7, 40, 43			cutaneous (mucosoal)
Alpha 9	**HPV16**	**54.38 %**	**41.62 %**	**high risk**
	HPV31	**3.82 %**	**1.08 %**	**high risk**
	HPV33	2.06 %	0.54 %	high risk
	HPV35	1.27 %	1.08 %	high risk
	HPV58	1.72 %	0.54 %	high risk
Alpha 10	HPV6, 11			low risk
Alpha 11	HPV34			high risk
	HPV 73	0.49 %		high risk
Alpha 13	HPV54			low risk
Alpha 14	C90			low risk
Alpha 15	HPV71			low risk

replication but also for the accumulation of chance errors in host cell DNA and is therefore a major predisposing factor in the development of human papillomavirus (HPV)-associated cancers (see below).

Virion formation, assembly and release: Although papillomavirus replication occurs in the basal cell layer of epithelia, no infectious virions are formed there. To trigger the expression of the late capsid genes L1 and L2, infected cells have to progress into upper skin layers. The trigger for the onset of late events is not yet fully understood, but appears to depend in part on changes in the cellular environment as the infected cell moves towards the epithelial surface. A critical component in virion formation is the up-regulation of the differentiation-dependent promoter, which for many HPV types is contained within the E7 ORF (p670 in HPV16; p742 in HPV31). A reduction in E7 and concomitant increase in E4 protein are hallmarks in this progression in the virus life cycle (Fig. 5.2-3).

The final stage in the papillomavirus productive cycle requires that the replicated genomes are packaged into infectious particles. The generation of capsid proteins (L1 and L2) is dependent on changes in mRNA splicing and thus on the generation of transcripts that terminate at the late (rather than the early) polyadenylation site (A_L, Fig. 5.2-2). The timing of capsid synthesis is regulated both at the level of RNA processing and at the level of protein synthesis. Negative regulatory elements that control RNA stability are present in the coding regions and in the late untranslated region of HPV, whereas a splicing silencer in the L1 gene leads to the preferential synthesis of early transcripts in proliferating cells. In addition to this, the pattern of codon usage within the HPV16 L1 and L2 genes is distinct from the pattern usually found in mammalian cells, which contributes further to the inhibition of capsid expression in the lower epithelial layers.

Packaging of the viral DNA occurs in the nucleus of differentiated epithelial cells, where L2 associates with PML bodies (sometimes also referred to as nuclear bodies, nuclear domains or nuclear dots), E2 and DNA. L1 assembles into capsomeres in the cytoplasm prior to nuclear relocation, where it then interacts with L2 to complete capsid formation. Papillomavirus replication and assembly therefore occur in regulated phases throughout the differentiation of epithelial cells.

Medical aspects: Papillomaviruses infect basal layers of body surface tissues, skin or mucosal epithelia of the genitals, anus, mouth or airways. Different types of papillomavirus are mostly specific for certain tissues, e.g., HPV type 1 tends to infect the soles of the feet, whereas HPV 2 is more likely found in the skin of the palms of hands, causing warts. Infection by most papillomavirus types is either asymptomatic or causes small benign tumors, known as papillomas or warts. In contrast, infections caused by some types, such as HPV 16, 18, 31, 45 and others are associated with an increased risk of cell transformation and cancer (high-risk type, Table 5.2-2). HPV infection can lead to cancers of the cervix, vulva, vagina, and anus in women, cancers of the anus and penis in men, as well as oropharyngeal cancers. HPV infections are associated with nearly all cervical cancer cases, an estimated 16% of all female cancers and 10% of all cancer cases.

Literature:

de Villiers, E.M. et al. Classification of papillomaviruses. *Virology* 2004;324:17–27.

Doorbar, J. Journal of Clinical Virology 2005;32:7–15.

Doorbar, J. Clin. Sci. (London) 2006;110:525–541.

Hamid, N.A. *et al.* Cell. Mol. Life Sci. 2009;66:1700–1717.

Kamper, N. *et al.* J. Virol. 2006;80:759–768.

Knipe, D.M., Howley, P.M. *et al*. (Eds.) *Fields Virology*, 5th Ed. Philadelphia: Lippincott Williams & Wilkins, 2007.

Lehoux, M. *et al.* Public Health Genomics 2009;12:268–280.

Munger, K. *et al.* Oncogene 2001;20:7888–7898.

Richards, R.M. *et al.* Proc. Natl. Acad. Sci. USA 2006;103: 1522–1527.

Wilson., V.G. *et al.* Virus Genes 2002;24:275–290.

Zheng, Z.M., Baker, C.C. Front. Biosci. 2006;11:2286–2302.

5.3 RNA Viruses

RNA viruses carry RNA as genetic material. Virions can contain single-stranded or double stranded RNA molecules. Many RNA viruses are associated with human diseases. Examples are mumps, measles, hepatitis C, influenza, polio and rabies.

RNA viruses encompass a large variety of RNA types and protein arrangements (5.1.1). Genomic RNA molecules can be negative-sense, positive-sense or ambisense. In all cases, the viral RNA is used as a template for its replication. (+)viral RNA is similar in sequence sense to viral mRNA and can often be immediately translated by the host cell into proteins necessary for virus replication. (–)viral RNA is complementary to mRNA and thus must first be converted to positive-sense RNA by a viral RNA-dependent RNA polymerase, before viral proteins can be synthesized. Bunyaviruses and arenaviruses carry genomic RNA segments with sections of both (+) and (–) sense (ambisense). The following section briefly summarizes the main features of Hepatitis C Virus biology as an example of a RNA virus life cycle.

5.3.1 Hepatitis C Virus

Hepatitis C virus (HCV) is classified in the family *Flaviviridae*, which comprises three genera, Flavivirus, Pestivirus and Hepacivirus, consisting of small, (+) strand RNA viruses. HCV is differentiated from the other members of the family by the genome organization of the structural protein ORFs and by the use of cap-independent translation initiation. HCV virions are spherical, enveloped particles of 50–60 nm diameter, carrying transmembrane glycoproteins E1 and E2, required for cell entry. The single strand RNA genome consists of a single RNA molecule of approximately 9600 nucleotides, surrounded by a protective shell of nucleocapsid protein (core). Virions obtained from the plasma of infected hosts are associated with host low density lipoprotein (LDL) and very low density lipoprotein (VLDL, 6.2). This additional lipid shell lowers the overall virion density and may play a role in virus envelope masking, immune evasion, as well as cell entry. HCV infects human hepatocytes and is a major cause of liver disease and hepatocellular carcinoma.

HCV genome (Fig. 5.3-1): It consists of a single, long ORF encoding a polyprotein that is then proteolytically cleaved into a set of distinct viral proteins. Translation of the HCV ORF is directed via an approximately 340 nucleotide long 5′ non-translated region (NTR) functioning as an internal ribosome entry site (IRES). The IRES is a highly structured RNA segment, which interacts with translation initiation factors and directs the assembly of ribosomal subunits in close proximity to the start codon of the ORF. The 5′-NTR is one of the most conserved segments of the HCV genome, a reflection of the importance of both RNA structure and sequence for this region.

The 3′ NTR is the RNA replication initiation site. It has a tripartite structure composed of a variable sequence (which follows immediately downstream of the stop codon of the ORF), a poly(U/UC) tract of heterogeneous length and a highly conserved 98 nucleotide sequence at the very 3′-end, essential for replication *in vivo*.

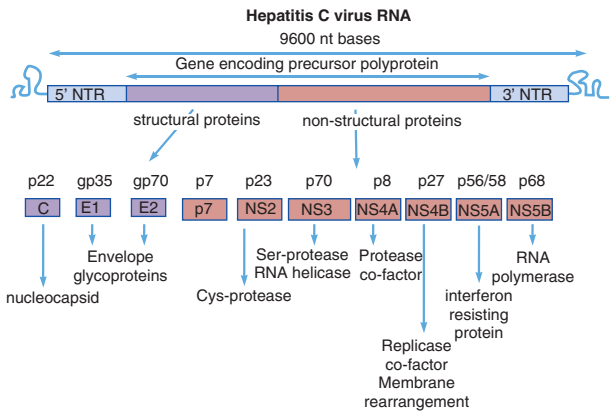


Figure 5.3-1. HCV Genome Organization

HCV life cycle (Fig. 5.3-2): HCV has a high rate of replication with approximately 10^{12} particles produced each day in an infected individual. Like many RNA viruses, HCV possesses a highly error-prone RNA replicase (polymerase), leading to a high mutation rate and the introduction of approximately one sequence change per replication cycle. The combination of high rate of virus replication and high mutation rate results in the formation of a virus quasispecies in infected hosts, which represents a large population of virion variants with heterogeneous genome sequences. The formation of a quasispecies of sequence variants in a single host allows the virus to respond rapidly to changes in selection pressures and thus facilitates immune evasion on a population level. The life cycle of HCV replication in human hepatocytes is characterized by six major phases, similar to many other (+)RNA viruses:

1. Receptor binding: Circulating HCV virions bind to receptors on the surface of hepatocytes. The main interactions during this phase have not been fully elucidated, but interactions with hepatocyte surface proteins LDL receptor, SRB-1, CD81 and CD209 (DC-SIGN) may play a role.
2. Fusion and uncoating (Capsid disassembly): HCV entry proceeds via receptor-mediated endocytosis (6.1.5). Acidic pH in the endosomal compartment is required to trigger membrane fusion, which require HCV envelope proteins and host receptor interaction.
3. Viral Protein Synthesis (Translation) and polyprotein processing: Host translation initiation factors and ribosomal subunits assemble on the genomic RNA and initiate the production of the HCV polyprotein, which is thereafter proteolytically cleaved.
4. Viral replicase assembly and RNA replication: Non-structural HCV proteins together with host proteins assemble into a virus replication machine (Replicase), which generates progeny virus genomes and viral RNAs for ongoing viral protein production.
5. Virion assembly: Structural HCV proteins are expressed and localize with genomic RNA to form nucleocapsid structures.
6. Virus Budding: The membrane bound structural proteins and nucleocapsids form new progeny virions that are released as independent, membrane enveloped particles.

Literature:

Lindenbach, B.D., Rice, C.M. Nature 2005;436:933–938.
Tanaka, T. *et al.* J. Virol. 1996;70:3307–3312.
Yamada. N. *et al.* Virology 1996;223:255–261.

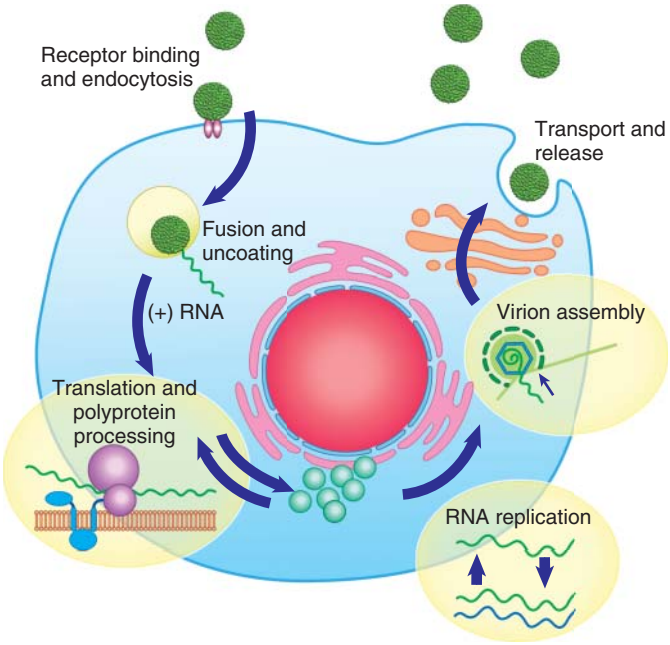


Figure 5.3-2. HCV Life Cycle
(after Lindenbach, B.D., Rice, C.M.: Nature 2005;436:933–938)

5.4 Retroviruses

Retroviruses, classified in the family of *Retroviridae* are characterized by the fact that virions carry genomic RNA, which is converted into DNA during infection by a process called reverse transcription. The DNA intermediate is then incorporated into the host genome by a virus-encoded integrase enzyme. Retroviruses are major pathogens affecting nearly all vertebrates. As an example, the intensively investigated human immunodeficiency virus (HIV-1) is described here. This virus is the causative agent of the acquired immunodeficiency syndrome (AIDS). The mode of retrovirus replication with an obligatory viral genome integration into the host genome is inherently mutagenic. Insertion of a retrovirus at a certain region in the host genome may activate previously silent proto-oncogenes or disrupt tumor suppressor genes. Retroviruses are therefore commonly associated with cell transformation. In addition, through the process of host genome integration and recombination, retroviruses can acquire full or truncated host genes. A number of retroviruses became acute transforming viruses, because they carry oncogenes that have the ability to transform normal cells into tumor cells. The best known example is Rous Sarcoma Virus, which carries a transforming version of the *c-src* gene. A number of other oncogenes have been observed in acute transforming retroviruses, including growth factors (*v-Sis*), growth factor receptors (*v-erbB*), kinases (*v-src, v-abl*) or transcription factors (*v-myc*). See Table 7.5-5.

5.4.1 Human Immunodeficiency Virus (HIV)

Genome structure of HIV-1 and protein processing (Fig. 5.4-1): The 9.7 kb dimeric (+)RNA genome of HIV-1 contains gag-pol-env genes coding for the essential enzymes and structural protein precursors (common to all retroviruses) and additionally, a number of regulatory genes. This RNA sequence is flanked by 3′ and 5′ enhancer and promoter sequences as well as untranslated (U3, U5) and repetitive (R) sequences, which are of importance in the reverse transcription step. All of the three possible reading frames are used in an overlapping way to derive a maximum of information out of a minimal genome size.

During transcription, differentially spliced mRNA's are synthesized to be translated into the regulatory protein factors and the Gag, Gag-Pol, Env polyprotein precursors. This is shown in the scheme below.

HIV replication cycle (Fig. 5.4-2): The HIV replication cycle is characterized by the requirement of reverse transcription and DNA integration into the host genome to enable viral protein expression. Main HIV target cells are CD4(+) T cells and macrophages (8.1.3). The replication cycle proceeds through eight major phases:

1. Attachment: The initial event in the replication cycle of HIV is the binding of the viral envelope glycoprotein gp120 to the CD4 receptor at the surface of host cells (CD4-positive target cells, lymphocytes, macrophages, dendritic cells).
2. Fusion and capsid disassembly: The HIV envelope protein gp120/gp41 will then interact with secondary co-receptors (e.g., the chemokine receptors CCR5 or CXCR4) to trigger membrane fusion and cell entry. In contrast to HCV described above, HIV entry occurs at the cell membrane rather than in the endosomal compartment, and does not require a pH change to trigger membrane fusion.
3. Reverse transcription (DNA synthesis): The process of reverse transcription takes place within a pre-integration complex (PIC) of nucleic acid, viral and host proteins, including reverse transcriptase, nucleocapsid protein, integrase and Vpr. It requires the coordinated enzyme activities of RNA-dependent DNA polymerase, RNAse-H and DNA dependent DNA polymerase (see p. 271). Reverse transcription of the (+)RNA is initiated with host t-RNA^{Lys3} binding to a conserved primer binding site (PBS) on the HIV genomic RNA, involves two obligatory strand transfer steps ('jumps') and results in double-stranded DNA. This process facilitates recombination events, which occur frequently during HIV

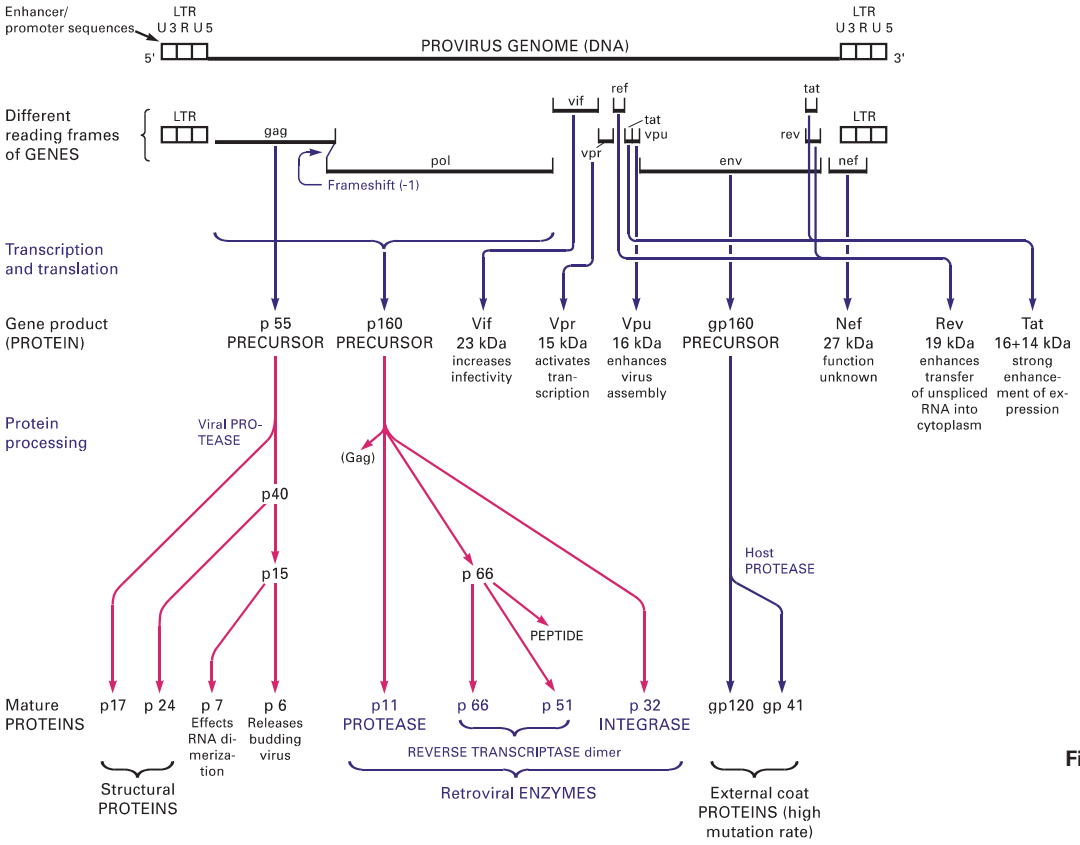


Figure 5.4-1. Gene Products and Protein Processing of HIV-1
(Red arrows: virus reactions, blue arrows: host reactions)

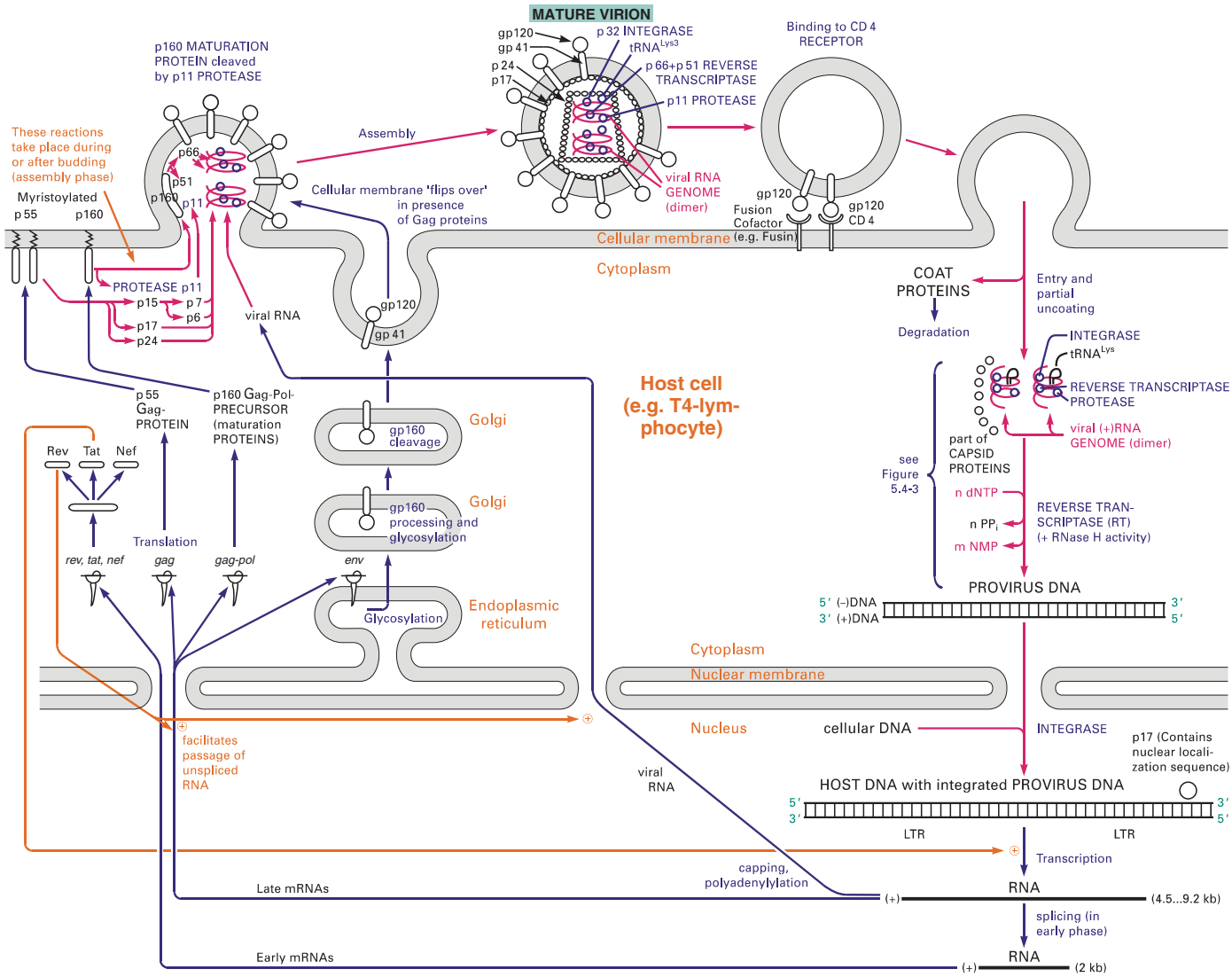


Figure 5.4-2. Mechanism of HIV Replication
(Red arrows: virus reactions, blue arrows: host reactions)

replication. Details of these reverse transcription steps are given in Figure 5.4-3 and also below.

① Viral t-RNA^{Lys3} primer is annealed to the primer binding site (PBS, length 18 nucleotides).

② (–)Strand DNA synthesis is started, the corresponding RNA strand gets afterwards degraded. The R region of the newly formed DNA binds to the R region at the 3′ end of the same or the second RNA strand (first jump).

③ The (–)DNA strand gets extended, most RNA is degraded.

④ The remaining RNA (PPT = polypurine tract) acts as primer for (+)strand DNA synthesis and is degraded afterwards, as well as the tRNA primer.

⑤ The (+)DNA strand binds to the other end of the (–)DNA strand at the PBS region (second jump).

⑥ Both linear strands are completed, the provirus formed contains long terminal repeats (LTR) at both ends.

4. Nuclear import and integration: After DNA synthesis is complete the PIC enters the host cell nucleus, where the viral DNA is integrated at random sites into the host DNA by the viral integrase enzyme. This phase is called the ‘proviral state’, as HIV can remain dormant for long periods (latent infection), unless transcription of the provirus DNA is initiated by the cellular transcription machinery (RNA polymerase II), often induced by gene activation signals, such as IL-2, T-cell mitogens, cytokines or TNFα.

5. Transcription (Viral protein expression): Transcription is mediated by host RNA polymerase II and associated transcription factors (the HIV-LTR contains three Sp1 binding elements, and two NF-κB and adjacent AP1 enhancer binding sites). Transcription of HIV mRNAs is complicated by the usage of multiple splice donors and splice acceptor sites, resulting in more than 30 alternatively spliced HIV-1 transcripts.

   Early after provirus activation, only highly spliced mRNA is generated, which codes for regulatory proteins (Nef, Rev, Tat). The presence of Rev is required for the suppression of secondary splice sites, allowing the production of partially spliced and unspliced gag-pol RNA. The *pol* reading frame encodes the HIV protease, reverse transcriptase and integrase proteins (Fig. 5.4 1).

6. Virus assembly: Viral proteins in the form of the polyprotein precursors Gag (p55) and Gag-Pol (p160) become located to the plasma membrane by post-translational addition of myristate moieties. Viral RNA becomes part of the assembly complex via a packaging signal sequence located between the 5′-LTR and the *gag* initiation site. The viral glycoprotein precursor gp160 is synthesized by ribosomes at the rough endoplasmatic reticulum. Cleavage to the external gp120 and the transmembrane gp41 glycoproteins is accomplished by a host protease.

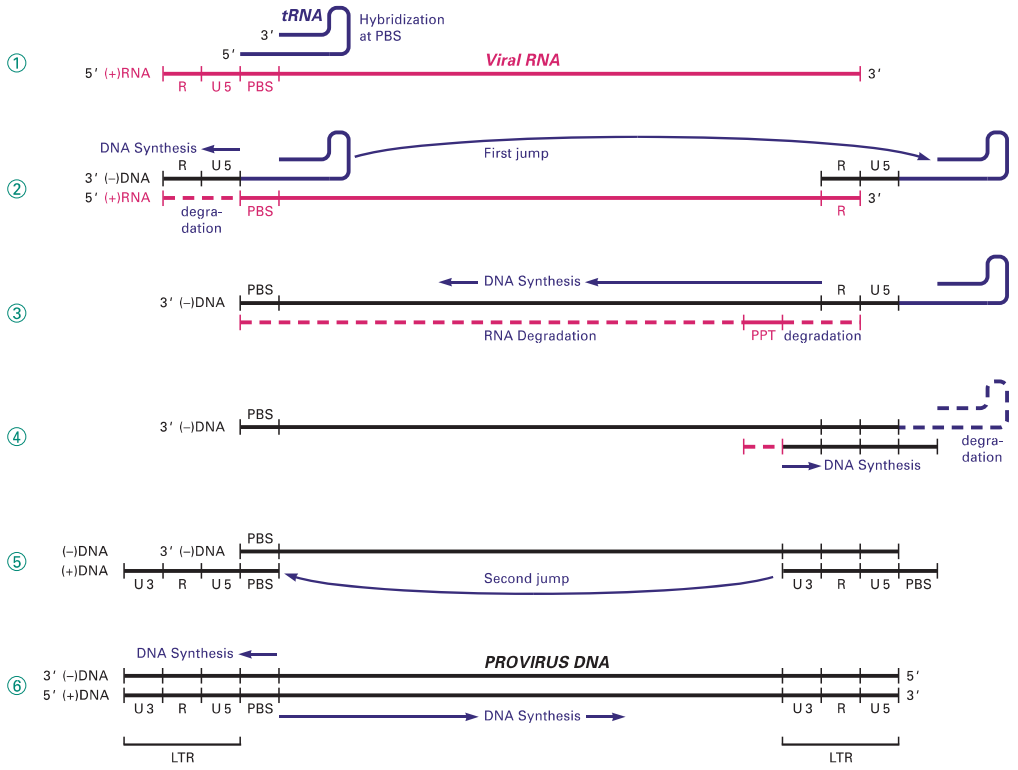


Figure 5.4-3. Formation of Provirus DNA from HIV-1 RNA (red: RNA, blue: $tRNA^{Lys3}$, black: DNA)

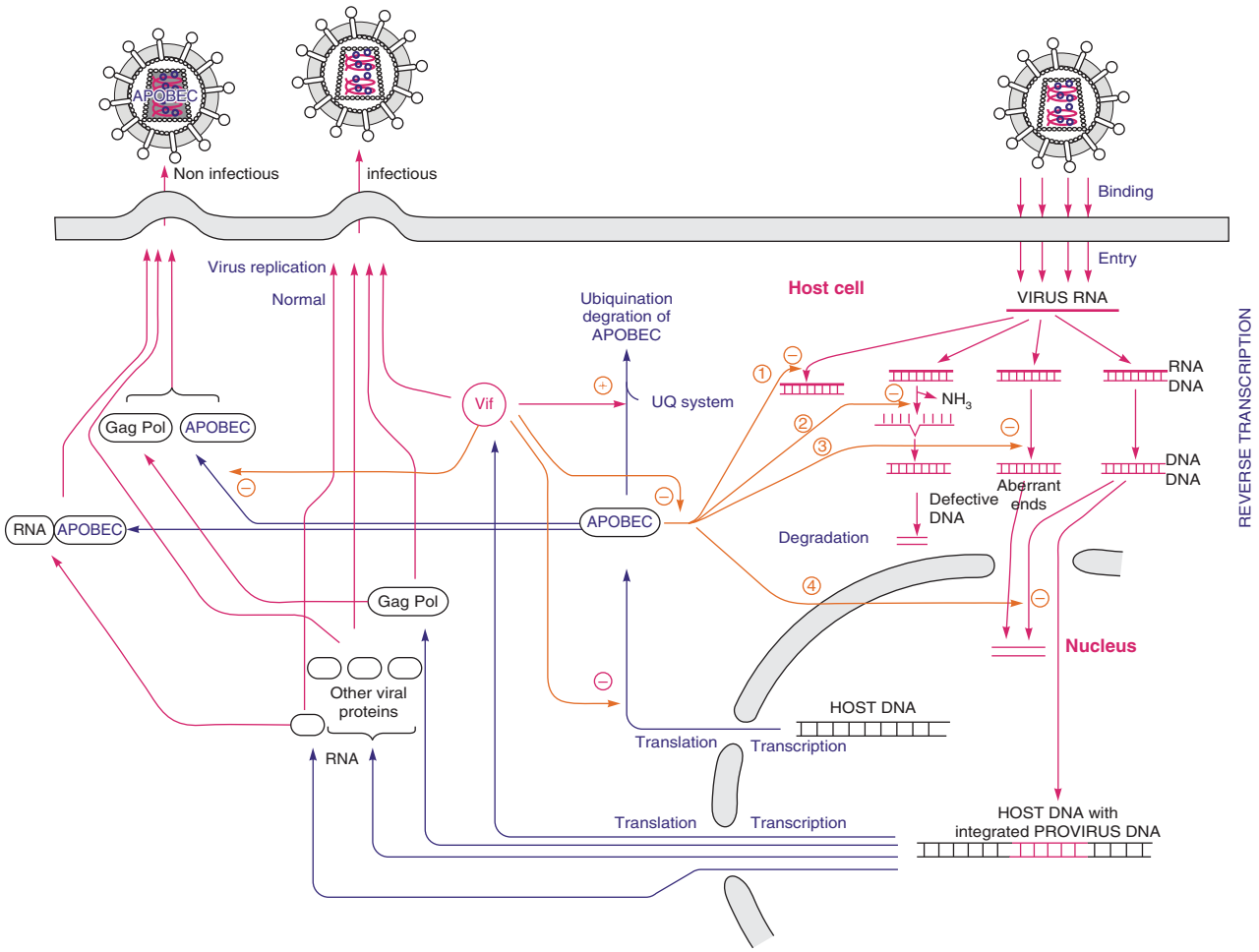


Figure 5.4-4. Schematic Representation of Vif and APOBEC3G Interactions during the HIV-1 Replication Cycle
Effects of APOBEC: ① Deaminase independent activity (inhibition of reverse transcriptase), ② Deamination, ③ Inducing defects in tRNA cleavage, ④ Inhibition of integration.

7. Virus budding: HIV uses cellular multiprotein complexes, termed 'endosomal sorting complexes required for transport' (ESCRT) to catalyze budding off of newly formed virions from the plasma membrane. The Gag protein contains a number of so-called late domains (PTAP, PPPY, YPDL) that interact with proteins in the ESCRT complexes, such as Tsg101. Inhibition of these interactions results in a block of virus budding from the cell membrane.
8. Virus maturation: After virus assembly and budding from the host cell membrane, cleavage of the Gag and Gag-Pol polyproteins by the HIV protease is required to complete the formation of a conical capsid structure and results in mature and infectious virions.

Mechanism of reverse transcription: Retroviral reverse transcriptases (RT) are multifunctional enzymes exhibiting three enzymatic activities: RNA-dependent DNA polymerase, ribonuclease H and DNA-dependent DNA polymerase which are provided by two active sites on the viral reverse transcriptase protein and are activated by protease action. These activities are necessary for copying the (+) RNA genome into (–)strand DNA, for removal of the RNA template and synthesis of the second (+)strand of DNA using the (–)DNA as a template. Reverse transcription is error-prone (*in vitro* ca. $1/10^4$), leading to introduction of replication errors and quasispecies populations of virion variants of heterogeneous sequences, including a large number of non-infectious virions. The quasispecies complexity is further increased by frequent recombination events.

Host cellular antiviral defense system (Fig. 5.4-4): As viruses and hosts co-exist, hosts have evolved antiviral defense mechanisms, while viruses continued to evolve ways to circumvent host defense mechanisms. APOBEC3G is an example of an antiviral host protein that interferes with replication of HIV and other retroviruses. APOBEC3G is a member of a larger family of cytidine deaminases, involved in mRNA editing and immunoglobulin diversification. As a cytidine deaminase, APOBEC3G can induce C to U mutations in the negative strand of the HIV. The uracil-containing DNA will then activate cellular uracil-DNA-glycosidase, leading to nucleotide excision and degradation of viral DNA (3.9.2.4). APOBEC3G is also incorporated into virions, reducing virion infectivity, and may also inhibit HIV replication through additional mechanisms independent of cytidine deaminase activity, e.g., APOBEC3G has been shown to inhibit the elongation activity of HIV-1 RT enzyme along the RNA template. APOBEC3G was also reported to inhibit viral DNA integration and provirus formation, probably by inducing defects in tRNA cleavage during plus-strand DNA transfer, leading to the formation of aberrant viral DNA ends. APOBEC3G can also interact with the integrase enzyme of HIV-1, which might interfere with the integrity of the integration complex, resulting in diminished integration rates.

HIV Vif protein apparently evolved specifically to counteract APOBEC3G function. By interacting with APOBEC3G, Vif triggers the ubiquitination and degradation of APOBEC3G via the proteasomal pathway (4.5.7). Figure 5.4-4 describes the interaction of APOBEC3G and Vif during the HIV-1 replication cycle.

Literature:

Cragie, R. J. Biol. Chem. 2001;276:23213–23216.

Gallo, S.A. *et al.* Biochim. biophys. Acta 2003;1614:36–50.

Goncalves, J., Santa-Marta, M. Retrovirology 2004;1:28.

Kaplan, A.H. AIDS Rev. 2002;4:104–111.

Wisniewski, M. *et al.* Proc. Natl. Acad. Sci. USA 2000;97: 11978–11983.

Malim, M. Phil. Trans. R. Soc. B 2009;364:675–687.

Wissing, S. *et al.* Mol. Aspects Med. 2010;31:383–397.

6 Transport

6.1 Transport Through Membranes

Wilhelm Just

6.1.1 Systems of Eukaryotic Membrane Passage

The eukaryotic cellular membrane is a permeability barrier to almost all molecules. The import of metabolites, nutrients, inorganic ions etc. into the cells and their export proceed through special structures. Many of the systems are described in detail in other contexts (quoted below). This section is a summary of the principles involved.

The important lipoprotein transport mechanisms are discussed separately in Section 6.2, while the different mechanisms of protein transport are the topic of Sections 4.5.2 … 5.

6.1.1.1 Energy Requirements (for details see 1.5.3):

If an uncharged compound A is to be imported into a cell against a concentration gradient ($[A_{inside}] > [A_{outside}]$, Eq. [1.5-9] and [1.5-9a]), expenditure of free energy

$$\Delta G\ [kJ * mol^{-1}] = R * T * 2.303 * \log \frac{[A_{inside}]}{[A_{outside}]} = 5.706 * \log \frac{[A_{inside}]}{[A_{outside}]}\ (\text{at } 25^{\circ}C)$$

is required. In the opposite direction, the same amount of free energy can be theoretically gained (it is lower under actual conditions).

For the import of charged molecules the electric potential also has to be considered (Eq. [1.5-12]):

$$\Delta G\ [kJ * mol^{-1}] = 5.706 * \log \frac{[A_{inside}]}{[A_{outside}]} + Z * 0.0965 * \Delta\Psi\ [mV]\ \ (\text{at } 25^{\circ}C)$$

where Z = charge number of the ion. The membrane potential of the cell at equilibrium can be calculated with the Goldman equation (7.2.1).

6.1.1.2 TC System

Membrane transport proteins are classified by the Transport Commission (TC) system analogous to the Enzyme Commission (EC) system for classification of enzymes (2.4.5, 10.3.1). The TC system provides description, TC numbers and examples of more than 600 family members of transport proteins. The TC database is operated by the Saier Lab Bioinformatics Group and consists of a comprehensive IUBMB-approved classification system. The transport systems are classified by the mechanism of transport as shown below. The categories 8 and 9 are reserved for accessory transport proteins and incompletely characterized transporters, respectively. The individual number, e.g., 1.A.1.2.10 indicates category – class – superfamily – family – protein.

1. Channels/Pores
   - 1.A. α-Type channels
   - 1.B. β-Barrel porins
   - 1.C. Pore-forming toxins (proteins and peptides)
   - 1.D. Non-ribosomally synthesized channels
   - 1.E. Holins
   - 1.F. Vesicle fusion pores
   - 1.G. Viral Fusion Pores
   - 1.H. Paracellular channels
2. Electrochemical Potential-driven Transporters
   - 2.A. Porters (uniporters, symporters, antiporters)
   - 2.B. Nonribosomally synthesized porters
   - 2.C. Ion-gradient-driven energizers
3. Primary Active Transporters
   - 3.A. P-P-bond-hydrolysis-driven transporters
   - 3.B. Decarboxylation-driven transporters
   - 3.C. Methyltransfer-driven transporters
   - 3.D. Oxidoreduction-driven transporters
   - 3.E. Light absorption-driven transporters
4. Group Translocators
   - 4.A. Phosphotransfer-driven group translocators
   - 4.B. Nicotinamide ribonucleoside uptake transporters
   - 4.C. Acyl CoA ligase-coupled transporters
5. Transport Electron Carriers
   - 5.A. Transmembrane 2-electron transfer carriers
   - 5.B. Transmembrane 1-electron transfer carriers
8. Accessory Factors Involved in Transport
   - 8.A. Auxiliary transport proteins
   - 8.B. Ribosomally synthesized protein/peptide toxins that target channels and carriers
   - 8.C. Non-ribosomally synthesized toxins that target channels and carriers
9. Incompletely Characterized Transport Systems
   - 9.A. Recognized transporters of unknown biochemical mechanism
   - 9.B. Putative transport proteins
   - 9.C. Functionally characterized transporters lacking identified sequences

Literature:

Busch, W., Saier, M.H. The transporter classification (TC) system. Crit. Rev. Biochem. Mol. Biol. 2002;37:287–337.

Hediger, M.A. *et al*. Pflugers Arch. 2004;447:465–468.

Saier, M.H., Tran, C.V., Barabote, R.D. TCDB: the Transporter Classification Database for membrane transport protein analyses and information. Nucleic Acids Res. 2006;34 (Database issue):D181–186.

International Union of Pharmacology. Pharmacol. Rev. 1994;46(2), 1999;51(2), 2005;57(4), 2008;60(3). See also http://www.iuphar-db.org/DATABASE.

Survey of the TCDB-system: http://www.tcdb.org/.

In the following only a few important groups are discussed in greater detail.

6.1.2 Channels / Pores

They are found in all types of organisms and allow the flow of solutes from higher to lower concentrations. They do not require any input of energy (passive transport). Many of them are used in a controlled way for conduction of electric signals or for initiation of metabolic activity (regulatory effects, Chapter 7), others function in compensation of osmotic or charge dysequilibria. Frequently, specificity for the ion or molecule passing through exists.

6.1.2.1 Voltage Gated Ion Channels (Table 6.1-1)

K^+ channels usually consist of homotetrameric structures with each α-subunit possessing six transmembrane spans (TMSs). There are four known K^+ channel families in mammals (humans):

1. The voltage dependent K^+ channels (K_v channels, in TC 1.A.1).
2. The two pore domain channels (K_{2P}, TASK, in TC 1.A.1.9.6).

Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Second Edition. Edited by Gerhard Michal and Dietmar Schomburg.

3. The calcium activated K^+ channels, (K_{Ca} channels, in TC 1.A.1).
4. The inward rectifier K^+ channels, (K_{ir}IRK channels, in TC 1.A.2).

Na+ and Ca++ channels: The α1 and α subunit, respectively, are about four times as large as the K^+ channel α-subunits.

6.1.2.2 Ligand-gated Ion Channels (Table 6.1-2)

Ligand-gated ion channels are found in TC 1.A.9. (Cys-loop family), TC 1.A.7. (P2X family), TC 1.A.10. (inotropic glutamate-gated ion channels) and TC 1.A.3. (Ryanodine-inositol-P_3 family). In addition, the IUPHAR nomenclature is used in Table 6.1-2 in order to describe the channel subunits.

Cys-loop family of neurotransmitter receptors (TC 1.A.9): The 5-HT_3 and acetylcholine receptors (cationic ion channels) and the $GABA_A$ and glycine receptors (anionic ion channels) generally depolarize or hyperpolarize the neuronal membrane, respectively. They contain an extracellular N-terminal ligand binding domain, which exhibits the receptor specificity. The receptor channels are probably hetero- or homopentameric.

E.g., the nicotinic acetylcholine receptor is comprised of a pentameric channel of $\alpha_2\beta\gamma\delta\varepsilon$ (mature muscle) subunit composition. The homologous subunits have four transmembrane α-helices (see 7.2.3; Figs. 7.2-4 and 7.2-5). At the amino-terminal extracellular region they exhibit a similar architecture in their ligand binding domains (containing a cysteine-loop).

Serotonin (5-hydroxytryptamine, 5HT) receptors are divided into seven classes, six are G-protein linked and one is a homo- or heteropentameric ligand gated non-specific cation channel (TC 1.A.9.2.1 and 2).

GABA (γ-aminobutyric acid) interacts with three classes of receptors, A, B and C. Classes A and C receptors are ligand-gated Cl^- channels (the A channel is described in Table 6.1-2). GABA binding to both Class A and C receptors opens the Cl^- channels, leading to increased membrane conductance. These two classes of receptors differ in their antagonist specificities and therefore are distinguished pharmacologically. Class B receptors activate other channels via G proteins.

The receptors consist of six types of subunits: α, β, γ, δ, ε, and π with variants (see Table 6.1-2). Usually the pentamer consists of 2α, 2β and 1γ (each with four putative TMSs), a long N-terminus and a short C-terminus, both extracellular.

ATP-gated P2X receptor family (TC 1.A.7): The various receptors differ in their sensitivities to ATP and in their inactivation kinetics. ATP binding initially causes opening of the non-selective cation channel, allowing Ca^{2+} entry. Prolonged exposure of slowly inactivating forms to ATP leads to dilation of the pore, making it permeable to larger molecules (up to 1000 Da). Then it functions as a cytolytic pore that is permeable to organic cations such as ethidium and N-methyl-D-glucamine.

Inotropic glutamate channel family (TC 1.A.10): These 'GIC channels' are divided into three types: (1) α-amino-3-hydroxy-5-methyl-4-isoxazole propionate (AMPA)-, (2) kainate- and (3) N-methyl-D-aspartate (NMDA)-selective glutamate receptors. The channels comprise homo- or heterotetrameric complexes, each containing three TMS + 1 P-loop. They all possess large N-terminal, extracellular glutamate-binding domains. There are six subfamilies: α, β, γ, δ, ε and ζ. The different channel (receptor) types exhibit distinct ion selectivities and conductance

Table 6.1-1. Voltage Gated Ion Channels

Ion specificity	Channel types	IUPHAR group, TC Family, Examples, Function	Occurrence (Examples)	Structure [1]
K^+ channels	Voltage gated	**K_v channels** (in TC 1.A.1, some in 8.A.5.): open or close in response to changes of the transmembrane voltage (see 7.2.2): Opened during membrane depolarization, outward flow → repolarization of membranes, 'delayed rectifier'. Inactivated near the equilibrium potential. Many voltage-sensitive K^+ channels function with β-subunits that modify K^+ channel gating. These nonintegral β-subunits are oxidoreductases that coassemble with the tetrameric α-subunits in the endoplasmic reticulum and remain tightly adherent to the α-subunit tetramer.	Brain, heart, skeletal or smooth muscle, (lung, eye, pancreas etc.)	$\alpha_4\beta$;, where α = 6 TMS + P Figures 7.2-2, 7.2-3
		K_A channels, (subtype): rapid, transient activation, opened during membrane depolarization, outward flow → repolarization of membranes	Neurons, secretory cells, muscles, heart	
		K_{SR} channels (subtype): strong voltage dependency, low selectivity for K^+/Na^+	Muscles	
		K_{ir} channels, 'inward rectifier' (in TC 1.A.2., e.g. K_{ir}IRK,) play a role in setting cellular membrane potentials. Closing of them upon depolarization effects long duration action potentials with a plateau phase (see 7.2.2). K^+ influx is preferred over outflux. Voltage-dependence may be regulated by external K^+, by internal Mg^{++}, by internal ATP and/or by G-proteins. Example Kir 6.2 (TC 1.A.2.1.7): High intracellular [ATP]/[ADP] ratio causes channel closing (response to energy supply). In pancreas → insulin release. The channel closing can also be achieved by sulfonylurea binding (oral antidiabetic drugs!). In brain: oxygen and glucose sensor.	Brain, skeletal muscle, (heart, kidney etc.), Pancreatic β-cells	minimal channel-forming structure: 4*(2 TMS + P)[2]
	Ca^{++} activated, voltage gated	**BK_{Ca} channels** (in TC 1.A.1.3.): large to small conductances, frequently activated by calmodulin-dependent mechanisms, frequently coupled with voltage-dependent Ca^{++} channels (TC 1.A.1.11), where Ca^{++} influx activates the BK_{Ca} channel. In neurons involved in after hyperpolarization.	Brain, skeletal and smooth muscle (liver, lung, prostate etc.)	4 * (7 TMS + P)[2]
Na^+ channels	Voltage gated	**Na_v channels** (in TC 1A.1.10, some in TC 8.A.17-18). They effect action potential initiation, repetitive firing and conduction in neurons, action potential initiaton and transmission in skeletal muscle. The α subunits are four times as large as the corresponding ones in K^+ channels.	Neurons, heart, skeletal muscle	$\alpha\beta_1\beta_2$, where α = 4* (6 TMS +P), Figures 7.2-2 (α-subunit) and 7.2-3 [2]
Ca^{++} channels	Voltage gated	**Ca_v channels** (in 1.A.1.11, some in 8.A.16). Several subtypes: L: activated by high voltage, large conductance, modulated by phosphorylation (PKA), inhibited by dihydropyridines (TC 1.A.1.11.2)	Muscle, neurons, endocrine cells, etc.	$\alpha_1\alpha_2\delta\beta$, α-subunit similar to Na^+- channel). [2] Figures 7.2-2 … 3 See 7.2.2, 7.4.4
		T: activated by low voltage, transient activation	Heart, neurons	
		N: activated by high voltage, large conductance, inhibited by ω-conotoxin	Neurons	
		P: activated by moderate voltage, moderate conductance, inhibited by ω-agatoxin	Neurons	
Cl^- channels (carry also other small anions)	Voltage gated	Opened by changes in membrane polarization	Neurons, epithelial cells	
	Ca^{++}-activated	Opened by increasing intracellular Ca^{++} concentration, also slightly voltage gated	Neurons, photoreceptors, secretory cells	

[1] TMS = transmembrane segments in the individual peptide chain, P = P domains, essential for pore structure.
[2] Contains also additional subunits.

Table 6.1-2. Ligand-gated ion channels. Nomenclature according to IUPHAR. Compare also Table 7.2-2

Receptor Family, TC number	IUPHAR Subunit	Function/Comments	Occurrence (Examples)
Cys-Loop Family			
5-HT_3 (serotonin) TC 1.A.9.2.1	5-HT3A-E	Cation-selective channel. Receptor activation triggers rapid depolarization; permeable for Na^+, K^+ and Ca^{++}; effectively excludes anions. Homo- or heteropentameric structure	Central and peripheral neurons
Nicotinic ACh TC 1.A.9.1.1	α_{1-10} β_{1-4} γ δ ε	Cation-selective channel. Mediates excitatory neurotransmission for muscle contraction at the vertebrate neuromuscular junction, across autonomic ganglia and selected synapses in brain and spinal cord: modulation of neurotransmitter release and neurotrophism. Homo- or heteropentameric structure	Peripheral neurons of ganglia, brain, spinal cord, skeletal muscle, neuromuscular junction, lymphocytes, fibroblasts, pulmonary neuroendocrine cells, spermatozoa, keratinocytes, granulocytes, chondrocytes, placenta
$GABA_A$ TC 1.A.9.5.2	β_{1-6} β_{1-3} γ_{1-3} δ, ε, π, (Θ, ρ)	Cl^--selective channel. Major inhibitory receptors in CNS mediating rapid phasic inhibitory synaptic transmission and also tonic inhibition. Pentameric structure; functional deficits occur in autism, anxiety, schizophrenia and epilepsy; target for benzodiazepines, barbiturates, neurosteroids and general anesthetics	Brain, spinal cord
Glycine TC 1.A.9.3.1	α_{1-4} β	Cl^--selective channel. Inhibitory neurotransmitter receptor; heptapentameric channel facilitating fast response. Mutations of the human α1 gene lead to hyperekplexia associated with severe muscular rigidity; α3 knockout leads to reduction in chronic pain sensitization; antagonistic effect of strychnine	Basal ganglia, substantia nigra, pontine regions, medulla oblongata, cervical spinal cord, retina
Zinc-activated	ZAC	K^+ permeable cation channel; Zn^{++} activates, tubocurarine inhibits spontaneous inward current	Pancreas, brain liver, lung, heart, kidney, skeletal muscle; absent in mouse and rat.
P2X Family			
ATP gated TC 1.A.7.1.1 … 3	P2X1-7	Non-selective cation channels (permeable for Na^+, K^+ and Ca^{++}) gated by extracellular ATP which has been released by exocytosis. A truncated version of the zebrafish P2X4.1 has trimeric structure. The channel is implicated in nerve transmission, pain sensation, control of blood pressure and response to inflammation	Widely expressed in mammalian tissues: central nervous system, smooth muscle neurons, other sensory neurons, immune system and hematopoietic organs.
Inotropic Glutamate Family / TC 1.A.10 Family			
AMPA TC 1.A.10.1.2 … 9	GluA1-4	Permeable primarily to monovalent cations, only low permeability to Ca^{++}. Implicated in the generation of the fast component of the excitatory postsynaptic potential (EPSP) in many central excitatory pathways and activity-dependent modulation of synaptic strength essential for synaptic plasticity (learning and memory)	Widespread in CNS
Kainate TC 1.A.10.1.1	GluK1-5	Permeable primarily to monovalent cations, only low permeability to Ca^{++}. Mediates excitatory synaptic signals and modulates presynaptic release of neurotransmitters; targets in the treatment of epilepsy and pain; homo- or hetero-oligomers	Specific CNS areas (e.g. stratum lucidum)
NMDA TC 1.A.10.1.3	GluN1 GluN2A-D GluN3A, B	Large conductance channels, highly permeable to monovalent cations and Ca^{++}. Heterotetramers, contain both GluN1 and GluN2 or GluN3 subunits; require co-agonism by glycine or serine; extracellular Mg^{++} blocks the channel at resting	Widespread in CNS, enriched in hippocampus and cerebral cortex
Orphan (GluD)	GluD1, 2	GluD1: Recombinant lurcher mutation reveals Ca^{++} permeability; TPA-induced PKC phosphorylation inhibits channel activity GluD2: Crucial in cerebellar function; mice lacking the gene display ataxia and impaired LTD (a model for cerebellar information storage)	GluD1: hair cells; transiently expressed in several brain regions (hippocampus, caudate, putamen) during early postnatal development GluD2: mainly in Purkinje cells (parallel fiber and climbing fiber synapses), moderately in olfactory bulb, hippocampus, cerebral cortex, midbrain, spinal cord
Ryanodine-Inositol-P_3 Family			
Inositol-P_3 receptors TC 1.A.3.2 family		Non-selective large conductance cation channel about sixfold selective for divalent over monovalent cations Intracellular channel, regulated by Ca^{++} and IP_3, the latter controlling the effects of Ca^{++}. Tetrameric structure; each subunit containing one IP_3 binding site	Highest density in brain cells, also in secretory vesicles, plasma membrane (of some cells). Subcellular localization: ER, nuclear envelope, Golgi
Ryanodine receptors TC 1.A.3.11 family		Intracellular Ca^{++}-induced Ca^{++} release channel; ryanodine from *Ryania speciosa* inhibits SR Ca^{++} release; tetrameric structure; poorly selective large conductance Ca^{++} channel regulated by Ca^{++}, Mg^{++}, ATP and caffeine	Expressed in a variety of tissues, highest density in striated muscle; subcellular localization: Sarcoplastic and endoplastic reticulum
Store-operated Ca^{++} Channels TC 1.A.4.3.1 TC 1.A.5.2.1.1		Mediate Ca^{++} release-activated Ca^{++} entry (store-operated Ca^{++} entry); the three homologues can form heteromultimeric complexes with high Ca^{++} selectivity; STIM1, an intraluminal Ca^{++} sensor, participates in the store-operated channel activation	Vascular endothelium, aortic and portal vein myocytes, pancreatic acinar cells, submandibular and parotide gland

properties (7.4.4, 7.4.5). Ligand (neurotransmitter) binding opens the transmembrane pore. Following activation, channels are desensitized although ligands remain to be bound.

Ryanodine and inositol 1,4,5-triphosphate receptor family (TC 1.A.3.): These channels in animal cells release Ca^{++} from intracellular storage sites into the cytoplasm and thereby regulate various Ca^{2+}-dependent physiological processes.

Ry and IP_3 receptors are homotetrameric complexes comprising C-terminal domains with six putative transmembrane α-helical spanners (TMSs). Ry and IP_3 receptor channels are regulated by phosphorylation of the regulatory domains, catalyzed by various protein kinases. Structure: 4 * (6 TMS + P), see 7.4.4, 7.4.5.

The cardiac and skeletal muscle sarcoplasmic ryanodine receptors are large ~2.3 MDa protein complexes. They include signalling proteins, such as FKBP12 protein, protein kinases, phosphatases, which modulate activity and binding of immunophilin to the channel. FKBP12 is required for normal as well as coupled gating between neighboring channels. PKA phosphorylation of RyR dissociates FKBP12 resulting in increased Ca^{++} sensitivity for activation.

Literature:

Miyazawa, A., Fujiyoshi, Y., Unwin, N. Nature 2003;423:949–955.

Reeves, D.C., Lummis, S.C. Mol. Membr. Biol. 2002;19:11–26 (review).

Witzemann, V. *et al*. Eur. J. Biochem. 1990;194:437–448.

Table 6.1-3. Solute Carrier Family

Carrier family, TC number (human carriers)	HUGO nomenclature (Alias)	Transport Specificity	Mechanism/Comments
High affinity glutamate and neutral amino acid transporter TC 2.A.23.2 … 3	SLC1A1-7 (Includes: EEAC1, EAAT, ASCT etc.)	L-Glu, L-Asp and D-Asp	Cotransport: 3 Na^+:1 H^+ / countertransport: 1 K^+ ASCT: Na^+-dependent exchange of Ala, Ser, Cys, Thr; involved in ischemia and ALS
Facilitative Glu transporter TC 2..A.1.1	SLC2A1-14 (GLUT1-12, HMIT)	Various hexoses- and H^+- *myo*-inositol cotransport (HMIT)	Role in control of glucose utilization, storage and sensing
Heteromeric amino acid transporters (heavy subunits) TC 8.A.9.1 … 3	SLC3A1-2 (rBAT, 4F2hc)	Amino acids, particularly Cys	Composed of a light (SLC7 members) and a heavy (SLC3 members) chain; rBAT / SLC7A9 responsible for apical Cys reabsorption in kidney
HCO_3^- transporter TC 2.A.31.1 … 4	SLC4A1-11 (AE1-3 etc.);	Cl^- / HCO_3^- exchanger (AE1-3); Na^+-coupled HCO_3^- transporters	At least eight of the SLC4 members transport HCO_3^- or CO_3^{2-}; role in CO_2 transport of erythrocytes
Sodium:glucose cotransporter TC 2.A.21.3 … 8	SLC5A1-12	Na^+: substrate (glucose, *myo*-inositol, iodide) cotransporters	Family of at least 220 members in animal and bacterial cells, e.g., Na^+: Cl^-: choline cotransporter; anion transporter, glucose-activated ion channel
Na^+-Cl^--dependent Na^+-neurotransmitter symporters TC 2.A.22.1 … 6	SLC6A1-20	GABA, glycine, proline, noradrenaline, dopamine, serotonin, betain, taurin, creatine	Regulation of extracellular solute concentrations
Cationic amino acid transporter; glycoprotein-associated TC 2.A.3.3 … 8	SLC7A1-14 (CAT (SLC7A1-4), gpaAT (SLC7A5-11))	CAT: cationic amino acids (facilitated diffusion). gpaAT: large and small neutral amino acids, negatively charged amino acid, cationic plus neutral amino acid	Associated glycoproteins form the SLC3 family
Na^+ / Ca^{++} exchanger TC 2.A.19.3	SLC8A1-3 (NCX1-3)	Na^+ influx / Ca^{++} efflux	NCX1 is expressed ubiquitously, NCX2 and NCX3 are limited to brain and skeletal muscle; NCX1 exchanges 3 Na^+ (extracellular) for 1 Ca^{++} (intracellular)
Na^+ / H^+ exchanger TC 2.A.36.1., -.8	SLC9A1-11 (NHE1-8)	Na^+ / H^+	Electroneural exchange of Na^+ and H^+ down their conetration gradient
Na^+: bile salt cotransporter TC 2.A.28.1.	SLC10A1-7 (NTCP, ASBT)	Na^+: bile salts	Na^+-dependent electrogenic uptake of bile salts
H^+-coupled metal ion transporter TC 2.A.55.2	SLC11A1-2 (NRAMP1, DMT)	H^+: various metal ions	NRAMP1 extrudes metal ions from phagolysosomes of macrophages and neurophils by H^+-metal ion cotransport; DMT1 is expressed widely and accepts a broad range of substrates
Electroneutral cation:Cl^- cotransporter TC 2.A.30.1 … -.5	SLC12A1-9	Na^+: K^+: Cl^-	Bumetanide-sensitive Na^+: K^+: Cl^- cotransporters and thiazide-sensitive Na^+: Cl^- cotransporters; electroneutral K^+: Cl^- cotransporter
Na^+: sulfate/carboxylate cotransporter TC 2.A.47.1	SLC13A1-5 (NaS, NaC, NaDC)	Na^+: SO_4^{2-}: mono, di- and tricarboxylates	Na^+-coupled symporters
Urea transporter TC 1.A.28.1	SLC14A1-2 (UT-A, UT-B)	Urea	Mediates passive urea uptake
H^+: peptide cotransporter TC 2.A.17.4	SLC15A1-2 (PEPT, PHT)	Di- and tripeptides	Electrogenic transporter utilizing the proton-motive force for uphill transport
Monocarboxylate transporter TC 2.A.1.13	SLC16A1-14 (MCT)	Monocarboxylates	H^+-linked transport of monocarboxylates (e.g. lactate, pyruvate, ketone bodies)
Vesicular Glu transporter TC 2.A.1.14	SLC17A1-9	PO_4^{3-}, organic anions	Mediate transport of organic anions, vesicular storage of Glu
Vesicular amine transporter TC 2.A.1.2.	SLC18A1-3 (VAChT, VMAT)	Positively charged amines	Utilizes an electrochemical gradient established by a vacuolar ATPase
Folate/Thiamine transporter TC 2.A.48.1	SLC19A1-3	Folate, thiamine	Transport energized by a transcellular H^+/ OH^- gradient
Type III Na^+: PO_4^{3-} cotransporter TC 2.A.20.2	SLC20A1-2 (Glvr, Ram, Pit)	Na^+, PO_4^{3-}	Electrogenic Na^+: PO_4^{3-} cotransport
Organic anion transporter TC 2.A.60.1	SLCO1-6 (OATP)	Amphipathic endogenous and exogenous organic anionic compounds	Na^+-independent transport (e.g. bile salts, steroid conjugates, thyroid hormones, drugs, toxins)
Organic cation / anion / zwitterions antiporter TC 2.A.1.19	SLC22A1-20 (OCT, OAT, OCTN2)		Participate in absorption and excretion of drugs, xenobiotics and endogenous compounds
Na^+-dependent ascorbic acid symporter TC 2.A.40.6	SLC23A1-4 (SVCT)	Na^+: ascorbic acid	Na^+-dependent vitamin C transport
Na^+: Ca^{++} + K^+ exchanger TC 2.A.19.4	SLC24A1-6 (NCKX)	Na^+ / Ca^{++} / K^+	Operates at a 4 Na^+ / 1 Ca^{2+} / 1 K^+ stoichiometry; normally extrudes Ca^{++} from cells
Mitochondrial carrier TC 2.A.29.1 … 16	SLC25A1-46 (MC)	Antiport for a wide variety of substrates	Gene defects cause e.g. Stanley syndrome, Amish microcephaly, hyperornithaemia-hyperammonaemia-hypercitrullinuria syndrome
Multifunctional anion exchanger TC 2.A.53.1 … 2	SLC26A1-11	Mono- and divalent anions (e.g. Cl^-, SO_4^{2-}, HCO_3^-, formate, oxalate)	At least three of them mediate electrogenic Cl^- / HCO_3^- and Cl^- / OH^- exchange
Fatty acid transport protein TC 9.B.17.1	SLC27A1-6 (FATP)	Long-chain and very long-chain fatty acids	Enhances the uptake of fatty acids into cells
Na^+-coupled nucleoside symporter TC 2.A.41.2	SLC28A1-3 (CNT)	Na^+: nucleoside	Na^+-dependent concentrative nucleoside transporter; critical roles in nucleoside salvage pathway

Carrier family, TC number (human carriers)	HUGO nomenclature (Alias)	Transport Specificity	Mechanism/Comments
Facilitative nucleoside transporter TC 2.A.57.1	SLC29A1-4 (ENT)	Purine and pyrimidine nucleosides and nucleobases	Equilibrative and concentrative nucleoside/nucleobase transport
Zn^{++} transporter TC 2.A.4.2 … 3	SLC30A1-10 (ZNT)	Zn^{++}	Zn^{++} efflux from the cytoplasm into subcellular compartments or across the plasma membrane
Cu^{++} transporter TC 1.A.56.1	SLC31A1 (CTR)	Cu^{++}	Cellular Cu^{++} uptake
Vesicular inhibitory amino acid transporter TC 2.A.18.5	SLC32A1 (VIAAT, VGAT)	GABA, Gly, H^+	Exchanges GABA or Gly for H^+
Acetyl-CoA transporter TC 2.A.1.25	SLC33A1 (AT-1, ACATN)	Acetyl-CoA	Substrate for acetyltransferases in Golgi
Type II Na^+- PO_4^{3-} cotransporter TC 2.A.58.1	SLC34A1-3 (NaPi-II)	Na^+, PO_4^{3-}	Major role in inorganic phosphate homeostasis
Nucleotide-sugar transporter TC 2.A.7.10 … 15	SLC35A-E	Nucleotide sugars	Transport of nucleotide sugars from the cytosol into the ER and Golgi
H^+-coupled amino acid transporter TC 2.A.18.5 … 8	SLC36A1-4 (LYAAT, PAT)	H^+, neutral amino acids (Gly, Ala, Pro)	Symport of H^+ and small neutral amino acids
Sugar- PO_4^{3-}/ PO_4^{3-} exchanger TC 2.A.1.4	SLC37A1-4 (G6PT)	Glucose-6-phosphate	Almost unexplored group of transmembrane sugar transporters
Na^+-coupled neutral amino acid transporter TC 2.A.18.6	SLC38A1-6 (SNAT)	Small aliphatic amino acids, particularly Gln	Na^+-coupled Gln transport; also H^+ countertransport (SNAT3 and SNAT5)
Metal ion transporter TC 2.A.5.4	SLC39A1-14 (ZIP)	Zn^{++}, Mn^{++}	Transport of metal ions from cell exterior or lumen of intracellular organelles into cytoplasm
Basolateral ion transporter TC 9.A.23.1	SLC40A1 (Ireg-1, ferroportin, MTP-1)	Fe^{++}	Efflux of Fe^{++} in a number of cell types
MgtE-like Mg^{++} transporter TC 9.A.19.4.	SLC41A1-3 (MgtE)	Mg^{++}	Prokaryotic MgtE homolog, Mg^{++} efflux
Rh ammonium transporter TC 1.A.11.4	(RhAG, RhBG, RhCG)	NH_4^+	Either NH_4^+ efflux or NH_4^+ / H^+ exchange, transporter is composed of Rh blood group (glyco)proteins
L-like amino acid transporter TC 2.A.1.44	SLC43A1-3	Neutral amino acids, Na^+-independent	Expressed in ileum, colon, cecum
Choline-like transporter TC 2.A.92.1	SLC44A1-5 (CTL)	choline	High affinity choline transport

6.1.3 Solute Carriers (Table 6.1-3)

This large group is distributed over several TC families, as shown in the Table. Comments on to a few important families are given below.

Solute : sodium symporter (SSS) family (TC 2.A.21): Members of the SSS family catalyze solute-Na^+ symport. The solutes transported may be sugars, amino acids, organo cations (e.g., choline), nucleosides, inositols, vitamins, urea or anions, depending on the system. They normally catalyze solute uptake via Na^+ symport.

Neurotransmitter : sodium symporter (NSS) family (TC 2.A.22): The members catalyze uptake of a variety of neurotransmitters, amino acids, osmolytes and related nitrogenous substances by a solute:Na^+ symport mechanism. These symporters have a critical role in regulating neurotransmission and are targets for psychostimulants, antidepressants and other drugs. The human dopamine transporter probably co-transports the positively charged or zwitterionic dopamine species with 2 Na^+ and 1 Cl^-. The human betaine:/GABA transporter cotransports 3 Na^+ and 1 or 2 Cl^- with one molecule of betaine or GABA. Two different glycine transporters, GlyT1 (TC 2.A.22.2.2) and GlyT2 (TC 2.A.22.2.6), cotransport glycine with 2 Na^+ and 3 Na^+, respectively as well as 1Cl^-.

Bile acid: sodium symporter (BASS) family (TC 2.A.28): Functionally characterized members of the BASS family catalyze Na^+:bile acid symport. These systems have been identified in intestinal, liver and kidney tissues of animals. In the human ileal intestine the symporter catalyzes the electrogenic uptake of bile acids with a stoichiometry of bile acid: Na^+ of 1 : 2.

Cation: chloride cotransporter (CCC) family (TC 2.A.30): Members of the family can catalyze NaCl:KCl symport, NaCl symport, or KCl symport depending on the system. The NaCl:KCl symporters are specifically inhibited by bumetanide while the NaCl symporters are specifically inhibited by thiazide. The proteins possess 12 putative transmembrane spanners flanked by large N-terminal and C-terminal hydrophilic domains.

Dicarboxylic amino acid:cation (Na^+ or H^+) symporter (DAACS) family (TC 2.A.23): The members of this family catalyze Na^+ and/or H^+ symport together with (a) a citrate cycle dicarboxylate (malate, succinate or fumarate), (b) a dicarboxylic amino acid (glutamate or aspartate), (c) a small, semipolar, neutral amino acid (Ala, Ser, Cys, Thr), (d) both neutral and acidic amino acids or (e) most zwitterionic and dibasic amino acids. These proteins possess between 10 and 12 hydrophobic segments per polypeptide chain. Possibly, the transporter consists of eight trans-menbrane segments, and one or two pore-loop structures that dip into the membrane.

Calcium/cation antiporter family (TC 2.A.19): The Na^+: Ca^{++} exchanger plays a central role in cardiac contractility by maintaining Ca^{++} homeostasis. Two Ca^{++}-binding domains, located in a large intracellular loop, regulate activity of the exchanger. Ca^{++} binding to these regulatory domains activates the transport of Ca^{++} across the plasma membrane.

Mitochondrial carrier (MC) family (TC 2.A.29): Members of the family are found exclusively in eukaryotic organelles although they are nuclearly encoded. Most are found in mitochondria, but some are found in peroxisomes of animals and other organelles. Permeases of the MC family (the human SLC25 family) possess six transmembrane β-helical spanners. They arose by tandem intragenic triplication. This event may have occurred less than 2 billion years ago when mitochondria first developed their specialized endosymbiotic functions within eukaryotic cells.

UDP-galactose/UMP antiporter family (TC 2.A.7.11): Nucleotide-sugar transporters (NSTs) are found in the Golgi apparatus and the endoplasmic reticulum of eukaryotic cells. They function by antiport mechanisms, exchanging a nucleotide-sugar from the cytoplasm for a nucleotide from the lumen. Thus, CMP-sialic acid is exchanged for CMP; GDP-mannose is preferentially exchanged for GMP, and UDP-galactose and UDP-N-acetylglucosamine are exchanged for UMP (or possibly UDP, 4.4.2). Also, other nucleotide sugars may be imported this way. The imported nucleotide-sugars are used for the synthesis of glycoproteins and glycolipids.

6.1.4 Primary Active Transport Systems (Table 6.1-4)

These ion transport systems are driven by energy derived from biochemical reactions. Thus they act as ion pumps increasing the concentration gradient across membranes. They are structurally and functionally different from ion channels. Ions bind to specific sites with high affinity. ATP hydrolysis energizes the 'uphill' transport by facilitating conformational changes that allow ion release at the side of higher concentration, e.g., Na^+/K^+ exchanging ATPase, TC 3.A.3.1.1. Flow in the opposite direction can be used for generating ATP (oxidative phosphorylation, photosynthesis, 3.11.4.5, 3.12.1).

The Na^+/K^+ exchanging ATPase is a P-type ATPase regulating the Na^+/K^+ ion gradients across the plasma membrane of all mammalian cells:

$$3\ Na^+_{inside} + 2\ K^+_{outside} + ATP = 3\ Na^+_{outside} + 2\ K^+_{inside} + ADP + P_i$$

Its major function is to reconstitute membrane polarization after opening of ion channels (7.2), to drive various secondary active transport processes (above) and to counteract the slight ion leakage at cellular membranes. In resting cells, the enzyme consumes 17 … 52% of the total energy turnover with the highest value found in the brain. The enzyme monomer consists of one α (112 kDa, 10 transmembrane helices) and one β subunit (35 kDa, 1 transmembrane helix). Three domains extend into the cytoplasm: the nucleotide binding (N), the phosphorus-binding (P) and the actuator (A) domain. Extension of α and β subunits into the extracellular space apparently has regulatory effects on the cation transport.

While phosphorylation of an aspartyl residue in the P domain only occurs in the presence of Na^+ (and Mg^{++}), dephosphorylation requires K^+. This results in a cyclic conformational change effecting

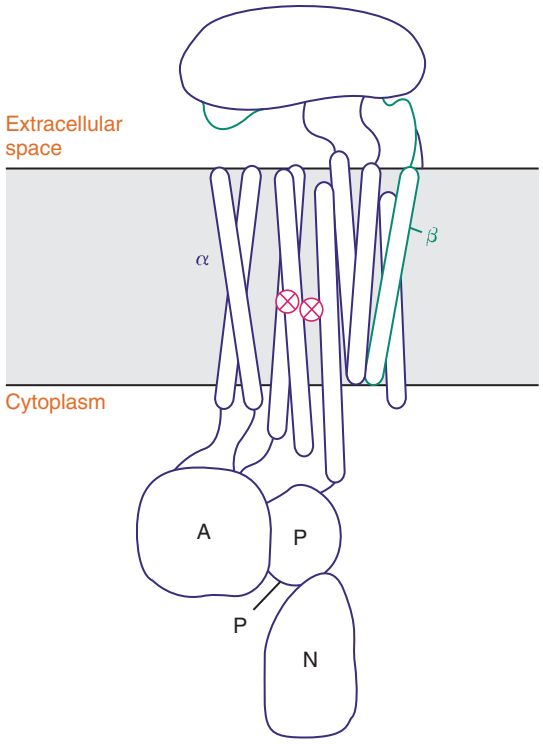


Figure 6.1-1. Structure of the Na^+/K^+-Exchanging ATPase, Simplified
The configuration marked with an asterisk in Figure 6.1-2 is shown.

Table 6.1-4. Primary-Active Transporters (Selection)

Transporter, TC number	Alias	Substrates	Function, comments
P-Type ATPases			
Na^+/K^+- ATPases TC 3.A.3.1.1 … 2		Na^+- K^+	See above and 7.2.1
Ca^{++}-transporting ATPases TC 3.A.3.2.25 TC 3.A.3.2.7 TC 3.A.3.2.5	 PMCA SERCA SPCA	Ca^{++}	PMCA: ATP-dependent removal of Ca^{++} from the cells, located in the plasma membrane (7.4.4). SERCA: ATP-dependent translocation of Ca^{++} from the cytosol into the sarcoplasmic/endoplasmic reticulum. (7.4.4) SPCA: Mg^{++} and ATP-dependent transport of Ca^{++}, located in the Golgi apparatus
Golgi aminophospholipid translocase TC 3.A.3.8.1		Phosphatidylserine, phosphatidylethanolamine	Aminophospholipid-transporting ATPase, flipping the phospholipid from the exo side to the cytosolic leaflet of membranes
H^+/K^+-transporting ATPase, gastric TC 3.A.3.1.2		H^+- K^+	In gastric parietal cells, effects H^+ efflux and K^+ uptake. 2 H^+ and 2 K^+ transported per 1 ATP hydrolyzed, ouabain-sensitive
H^+/K^+ transporting ATPases, non-gastric TC 3.A.3.1.4		H^+- K^+	Plasma membrane H^+/K^+ exchange, ouabain-sensitive
F-Type ATPases			
Mitochondrial F_1F_O ATP synthase, H^+ transporting. TC 3.A.2.1.3	F_1F_O ATPase	H^+, ADP, ATP	In mitochondria, chloroplasts and bacteria. By a rotary mechanism, it effects ATP synthesis from ADP and inorganic phosphate utilizing H^+ moving down an electrochemical gradient. Ratio 3 H^+_{in}: 1 ATP generated. Inhibited by oligomycin. See 3.11.4.5
V-Type ATPases			
Lysosomal ATP-ase, H^+ transporting TC 3.A.2.2.3		H^+	In eukarya and bacteria. Acidification of Golgi- derived and chlatrin-coated vesicles, liposomes, vacuoles etc. Structurally related to F-type ATPases, operating in reverse direction
ATP-Binding Cassette transporters			
Cholesterol/phospholipid flippase TC 3.A.1.211.1		Phospholipids	Specific organ distribution, cellular homeostasis of phospholipids and sterols
Multidrug resistance transporter TC 3.A.1.201.1 … 3	MDR, PGY1, TAP	Multiplicity of compounds	In animals, fungi, bacteria. Effects multidrug resistance by an export mechanism (MDR 1), secretion of phosphatidylcholine (MDR3) and bile salts (BSEP) etc. MDR: single chain, 6 + 6 TMC, 2 ABC units
ALD (subfamily) ABCD1-D4 TC 3.A.1.201.1 … 3	ALD, PXMP	Fatty acids, acyl-CoA	H^+ Half transporters, expressed in peroxisomes: Transport of fatty acids and acyl-CoA's; defects in ABCD1 convey X-linked adrenoleukodystrophy

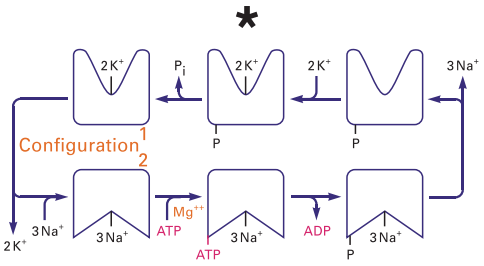


Figure 6.1-2. Postulated Mechanism of the Na^+/K^+-Exchanging ATPase

the transmembrane transport of the ions (Fig. 6.1-2). Similar phosphorylation/dephosphorylation cycles also occur with the related H^+/K^+-exchanging ATPase and the Ca^{++}-transporting ATPase (Table 6.1-4).

Cardiac glycosides (3.5.2.2) inhibit the enzyme and thus diminish the transmembrane Na^+ gradient impairing the Na^+-driven Ca^{++} export from the cytosol. The increased Ca^{++} concentration intensifies cardiac contractions.

Literature:

Barnard, E.A. Trends in Pharmacol. Sci. 1996;17:305–309

Fasolato, C. *et al.* Trends in Pharmacol. Sci. 1994;15:77–83.

Gadsby, D.C. Nature 2007;450:957–959.

Kuriyama, H., *et al.* Pharmacol. Reviews 1995;47:387–515.

Mintz, E., Guillain, F. Biochim. Biophys. Acta 1997;1318:52–70.

Shinoda, T. *et al.* Nature 2009;459:446–449.

See also references in the quoted chapters.

6.1.5 Import by Endocytosis and Pinocytosis

Endocytosis is the cellular uptake of macromolecules by formation of vesicles invaginating the plasma membrane. Uptake can also proceed with large cell fragments, microorganisms etc. (phagocytosis by specialized cells of multicellular organisms for defense purposes, 8.1.7) or with fluids (pinocytosis). An important example is the cellular uptake of lipoproteins. It is described in the next section (6.2).

Most animal cells use a receptor-mediated system for endocytosis. The extracellular macromolecules bind specifically to their cognate receptors, which assemble in clathrin coated pits at the plasma membrane. Clathrin is an association of three heavy and three light protein subunits to a 'three-legged' triskelion. The units combine to form a polyhedral structure and promote the fast budding of coated vesicles from the internal surface of the membrane. After leaving the membrane, the clathrin coat is removed and returns to the membrane. The ingested molecules are thereafter metabolized, while frequently the receptor is recycled to the membrane. This mechanism results in a thousandfold enrichment of imported molecules.

Endocytosis is also a means to terminate the action of receptors involved in signal transduction. In a number of cases both ligand and receptor are degraded thereafter (Figs. 7.5-2 and 7.5-3).

Pinocytosis is a constitutive process, which also starts at (receptor-less) coated pits. Pinocytotic vesicles are usually small.

Clathrin-coated vesicles are also used to transport enzymes from the *trans*-Golgi network to endosomes, while other intracellular transport steps use different coats. They are described in 4.5.2.2.

Literature:

Brodsky, F.M. Science 1988;242:1396–1402.

Doherty, G.J., MacMahon, H.T. Ann. Rev. Biochem. 2009;78: 857–902.

Kirchhausen, T. Curr. Opin. Struct. Biol. 1993;3:182–188.

Sandvig, K. *et al.* Histochem. Cell. Biol. 2008;129:267–276.

6.1.6 The Cytoskeleton as Means for Intracellular Transport and Cellular Movements in Eukarya

The cytoskeleton consists of

- microtubules (polymers of α- and β-tubulin, diameter 25 nm)
- actin filaments (polymerized β- and γ-actin, diameter 7 nm)
- intermediary filaments (diameter 8 … 15 nm).

Cytoskeletal proteins contribute to cellular movements by expansion and contraction and by providing tracks for molecular motors. Microtubules are present in all nucleated cells. They take part in cell division (4.3), intracellular transport and secretion, movement of flagella and cilia.

Structure of microtubules (Fig. 6.1-3): Microtubules consist of heterodimeric, globular α/β tubulin subunits (53 and 55 kDa, respectively), which are longitudinally aligned. 13 of these arrangements form a hollow tube with structural polarity (+ and – ends). Each subunit binds GTP, only the β-unit bound GTP is exchangeable. Microtubule associated proteins (e.g., MAPs 1 … 4 with different forms) are bound to the surface and stimulate the microtubule assembly. This activity can be modified by PKA-catalyzed phosphorylation. Dynein (MAP1C), kinesin and kinesin-like proteins are molecular motors moving along the microtubuli for transport of cellular components. Assembly of microtubules is inhibited by, e.g., colchicine, vinblastine, griseofulvin, Ca^{++}/calmodulin by binding to tubulin. Taxol, on the other hand, stimulates the assembly and stabilizes microtubuli.

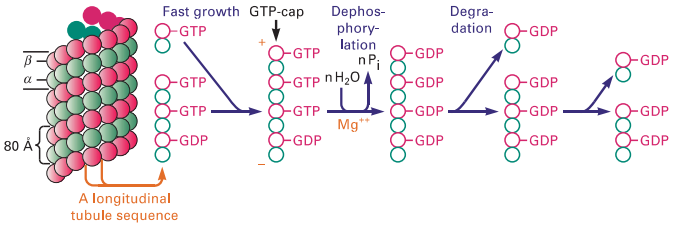


Figure 6.1-3. Assembly, Disassembly and Structure of Microtubules

Motion during mitosis: Microtubules support several steps in chromatide and cell separation (4.3.5). Controlled polymerization and depolymerization reactions (dynamic instability) are used to generate force. When GTP-carrying tubulin heterodimers assemble in presence of Mg^{++} (preferably at the + end), the GTP is slowly hydrolyzed. During fast assembly, several tubulin dimers at the + end still carry GTP (GTP-cap) protecting against disassembly. At decreased assembly rates the GTP cap disappears and the end of the microtubule disassembles. A number of molecular motors (dynein, kinesin-like proteins) associated with microtubules move polar microtubuli along each other and chromatides along kinetochore microtubuli (Fig. 4.3-6). Intracellular organelle distribution and motility are strongly dependent on these motors and microtubular tracks. However, many details are still unknown.

Intracellular movement along microtubules: Interphase microtubules are more stable than those formed during mitosis. ($t_{1/2}$ > 300 sec vs. < 50 sec.) Posttranslational modification of tubulin (acetylation, removal of C-terminal tyrosine) may play a role.

In various cells (e.g., neurons, 7.2.5) motor proteins recognize the polarity of the microtubules and move cargoes (vesicles, proteins, neurotransmitters) either towards the plus end (kinesin and most kinesin-like proteins, KLPs, in neurons towards the nerve terminal = anterograde) or towards the minus end (dynein, in neurons = retrograde).

The motor protein kinesin consists of two heavy and two light subunits. The structure of the head portion strongly resembles that of the myosin structure. Movements similar to myosin are energized by ATP hydrolysis and involve a cycle of conformational changes (Fig. 7.4-7). The kinesin heads move along the microtubules, while the tail is attached to the cargo.

Extracellular movements: Groups of cilia (hair like organelles) transport fluids, mucus or small particles (dust) along the cell surface

by wavelike movements, while the single, longer flagella propel, e.g., sperm cells and flagellates. The propelling units are similarly composed: Two central microtubules are surrounded by a ring of 9 fused microtubule doublets with attached dynein molecules. A major number of other proteins connect the structures. For movement, both dynein arms slide along the next microtubule, energized by ATP hydrolysis.

Actin filaments: Actin filaments, forming a network underneath the plasma membrane, are attached at focal adhesions. The structure of the filaments resembles those in muscle (7.4.5). Together with actin binding proteins, such as myosin, they modify cell shape and are implicated in forming spikes and invaginations.

Intermediary filaments: Intermediary filaments vary greatly in size and structure. Generally, they consist of a lengthy α-helical structure, two of which associate to form a coiled coil. Further associations lead to structures of higher order. The filaments carry a globular domain on each end. They mainly provide resilience. Known intermediate filament proteins are e.g., keratin (in outer epidermal layer, in hair, bird feathers), vimetin (in fibroblasts, endothelial cells etc.), nuclear laminin (underneath the inner nuclear membrane) and neuronal intermediary filaments (along the axon).

Literature:

Erickson, H.P., *et al.* Ann. Rev. of Biophys. Biomol. Struct. 1992;21:145–166.

Klymowski, M.W. Curr. Opin. in Cell Biol. 1995;7:46–54.

Przewloka, M.R., Glover, D.M. Ann. Rev. Genetics 2009;43:439–465.

6.2 Transport of Lipids in Plasma

Horst Klima

While the situation in humans is described here, a qualitatively similar situation exists in higher vertebrates.

Plasma lipoproteins are high molecular weight aggregates composed of lipids (mainly free or esterified cholesterol, triacylglycerols = triglycerides and phospholipids) and one or more specific apolipoproteins (Apo, Table 6.2-1, Fig. 6.2-1). Lipoproteins are the functional units of delivering water-insoluble lipids via the circulation to cells for utilization or storage. They play a decisive role in cholesterol homeostasis and in triglyceride transport. While they show a common basic structure (the polar groups of phospholipds, the OH-groups of cholesterol and the apolipoproteins at the outside, cholesterol esters and triacylglycerols in the core), they vary largely in size and composition.

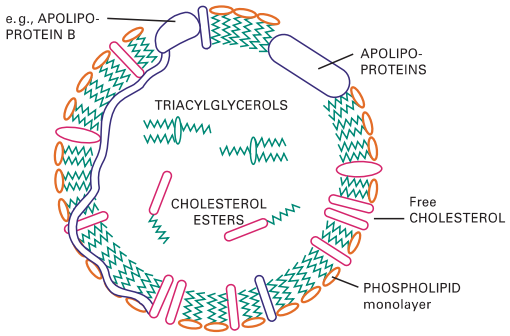


Figure 6.2-1. General Structure of Lipoproteins

6.2.1 Apolipoproteins (Apo)

There are also striking differences in the apolipoprotein composition of the lipoprotein classes. Tables 6.2-2 and 6.2-3 show the distribution and the properties of apolipoproteins.

6.2.2 Plasma Lipoprotein Metabolism (Fig. 6.2-2)

There are four major pathways of lipoprotein metabolism. The first two involve chylomicrons and VLDL, which contain Apo B as an essential, not exchangeable membrane constituent.

Chylomicron metabolism: Chylomicrons are transport vehicles for dietary lipids (mainly triacylglycerols) from the intestine to peripheral tissues and to the liver. They are assembled in the Golgi apparatus of intestinal mucosal cells from biosynthesized Apo B-48 and absorbed lipids. Following secretion, chylomicrons acquire Apo E and Apo C-II primarily from HDL. Apo C-II activates lipoprotein lipase (LPL), which originates mainly from adipose tissue and muscles. The majority of it is bound to heparan sulfate (2.9) at the cell membranes of the capillary vessel endothelium. LPL also exists in a free form in the circulation. The fatty acids liberated from triacylglycerols by action of membrane-bound LPL are directly transported into the adipose and muscle cells, while those generated by LPL in the bloodstream bind to serum albumin and are eventually taken up by heart and muscle cells. There, they can be degraded (3.4.1.5) or be utilized for synthetic reactions. As the size of the remnants decreases, a second (secreted) lipase, hepatic triacylglycerol lipase (HL), may play a role in triacylglycerol degradation. Eventually, the remnants are bound to remnant receptors of the liver (6.2.4) and internalized into liver cells similarly to the degradation of LDL (see below).

VLDL/LDL metabolism: VLDL transports lipids from the liver to peripheral tissues (such as muscle) and to adipose tissue. After assembly

Table 6.2-2. Distribution of Apolipoproteins in Plasma Lipoproteins (% of Protein)

Apolipoprotein	Chylomicrons	VLDL	LDL	HDL_2	HDL_3
A-I	33	trace	trace	65	62
A-II	trace	trace	trace	10	23
A-IV	14	0	0	?	trace
B	5	25	> 95	trace	trace
C-I, C-II, C-III	32	55	0 … 2	13	0
D	0	trace	trace	2	4
E	10	15	0 … 3	3	1
Other	6	5	0 … 5	4	5

Table 6.2-1. Classification, Properties and Composition of Plasma Lipoproteins (Data vary somewhat between authors)

	Chylomicrons	Remnants	Very low density lipoprotein (VLDL)	Intermediate density lipoprotein (IDL)	Low density lipoprotein (LDL)	High density lipoprotein (HDL_2)	High density lipoprotein (HDL_3)
Spec. mass (g/ml)	< 0.95	< 1.006	< 1.006	1.006 … 1.019	1.019 … 1.063	1.063 … 1.125	1.125 … 1.210
Diameter (nm)	80 … 500	> 30	30 … 80	25 … 35	18 … 28	9 … 12	5 … 9
Electroph. mobility [1]	origin	origin	pre-β	pre-β/β	β	α	α
Proteins (%)	1 … 2		5 … 10		20 … 24		**45 … 50**
Cholesterol (%)	0.5 … 1		6 … 8		5 … 10		3 … 6
Cholest. esters (%)	1 … 3		12 … 14		**35 … 40**		14 … 18
Triacylglycerols (%)	**86 … 94**		**55 … 65**		8 … 12		3 … 6
Phospholipids (%)	2 … 8		12 … 18		20 … 25		20 … 30

[1] Migration speed during electrophoresis as compared with α and β globulin.

Table 6.2-3. Properties and Function of Plasma Apolipoproteins

Apolipoprotein	Mass (kDa) ca.	Conc.[1] (mg/100 ml)	Major occurrence (in parentheses site of synthesis)	Function
A-I	29	130	HDL, chylomicrons (liver, intestine)	structural, cofactor of LCAT, ligand for HDL receptor
A-II	174		HDL, chylomicrons (intestine)	displaces Apo A-I from HDL particles, generates stable HDL complex
A-IV	45	?	HDL, chylomicrons (liver)	activator for LCAT, ligand for HDL receptor?
Apo (a)	300–800		Lp (a) (liver)	unknown, risk factor for coronary heart disease
B-100[2]	513	80	VLDL, IDL, LDL (liver)	structural, ligand for LDL receptor, cellular uptake
B-48[2]	241		chylomicrons, remnants (intestine)	structural, secretion of chylomicrons
C-I	6.6	6	HDL, chylomicrons, VLDL (liver)	displaces Apo-E from lipoproteins
C-II	9	3	chylomicrons, remnants, VLDL, HDL (liver)	activator of lipoprotein lipase
C-III	9	12	chylomicrons, remnants, VLDL, HDL (liver)	inhibits premature uptake of triacylglycerol-rich lipoprotein by the liver
D	20	10	HDL (liver, intestine, placanta, brain)	transports cholesterol and sterols, exchange of cholesterol esters
E	34	5	HDL (liver, macrophages)	Involved in transport of triglycerides, phosphorlipids, cholesterol and cholesterol esters in and out of cells

[1] In normal fasting plasma.
[2] Apo B-100 (4536 amino acids) and B-48 (2125 amino acids) are synthesized from a single Apo B gene. Apo B-100 is translated from the full length of mRNA in the liver, while Apo B-48 is synthesized from an Apo B post-transcriptional mRNA containing a premature translational stop codon introduced by mRNA modification in the intestinal mucosa (see 4.2.1).

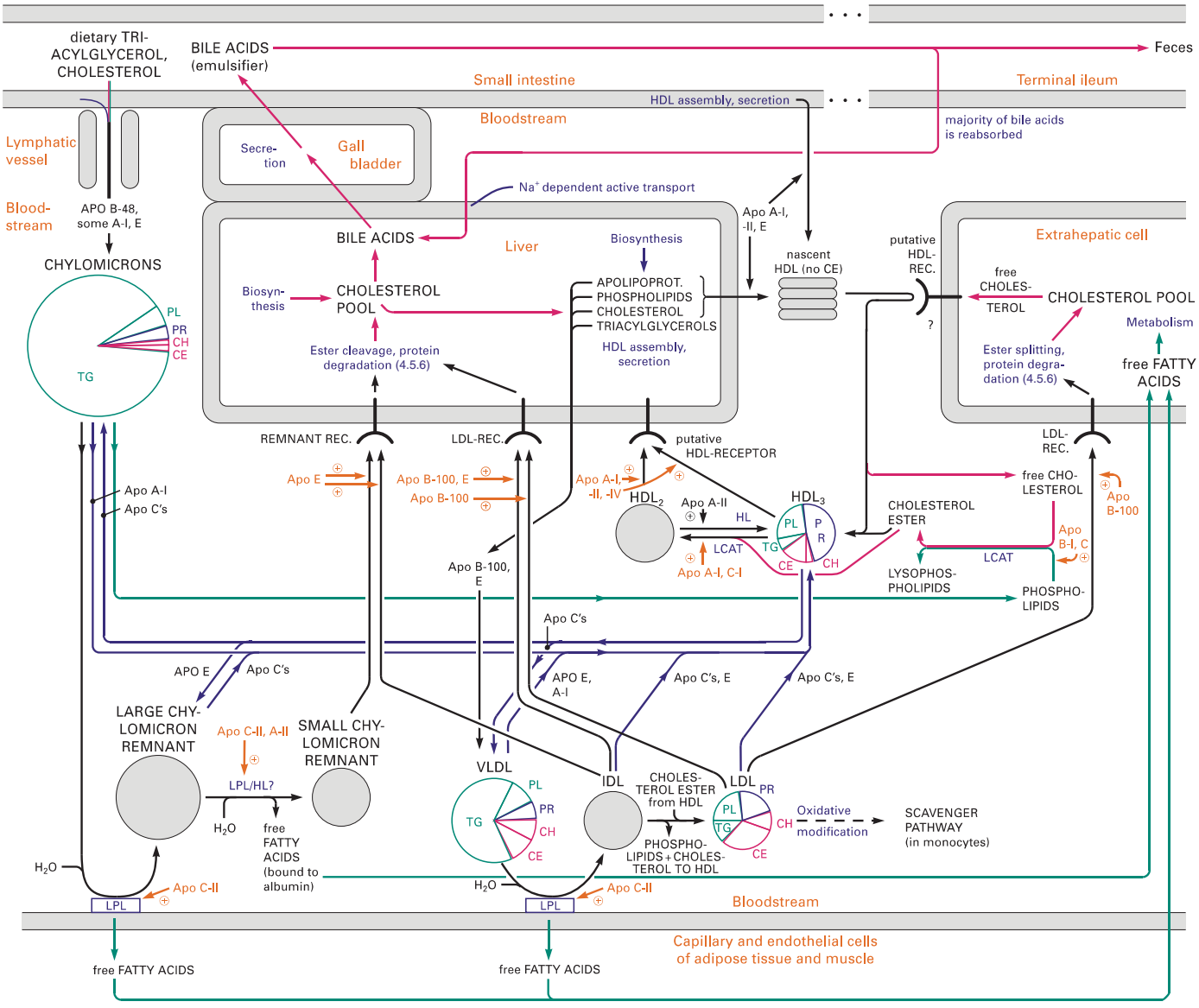


Figure 6.2-2. Metabolism of Chylomicrons, VLDL and HDL in Mammals
Contrary to the arrow colors elsewhere, their meaning is here: (black): lipoprotein metabolism, (blue): transfer of proteins (PR), (green): transfer of triacylglycerols (TG), phospholipids (PL) and fatty acids, (red): transfer of cholesterol (CH) and cholesterol esters (CE).

in the Golgi apparatus and secretion, triacylglycerol-rich VLDL from the liver (containing Apo B-100) associates with Apo C-II and Apo-E which have been dissociated from HDL. VLDL then becomes converted to IDL by endothelial and plasma lipoprotein lipase (LPL) analogous to the conversion of chylomicrons to remnants; the fate of the liberated fatty acids is also the same. About one half of the IDL is removed from plasma by interaction of Apo E and Apo B-100 at the particle surface with LDL- or remnant receptors. The other half is converted into LDL. HL is necessary for this step. Lipid components (via lipid exchange proteins) and apolipoproteins (except Apo B-100) are exchanged with HDL. LDL and IDL bind to LDL receptors of liver, adrenal and peripheral cells (including fibroblasts and smooth muscle cells), followed by endocytosis. For the intracellular metabolism, see 6.2.4.

In birds, oocytes express a special receptor for vitellogenin (an Apo B related protein at the surface of VLDL). This receptor directs VLDL to the ovaries in order to satisfy the lipid needs for egg production.

HDL metabolism: HDL seems to be involved in reverse cholesterol transport from peripheral tissues to the liver. Nascent HDL, composed primarily of Apo A-I, free cholesterol and phospholipid disks, appears to be assembled into stacked discoidal structures. These particles, synthesized by the liver, the intestine and macrophages, take up free cholesterol from extrahepatic cells including foam cells (The involvement of a specific HDL receptor is unclear). Cholesterol is esterified with fatty acid (from lecithin) by lecithin-cholesterol acyltransferase (LCAT). The esters enter the HDL core. Additional uptake of triacylglycerols occurs. The discoidal HDL changes into spherical HDL_3. Further uptake of triacylglycerols and cholesterol converts HDL_3 into a HDL_2 particle. (The reverse reaction is catalyzed by HL.)

The mechanism of HDL uptake and catabolism by liver cells is not completely known. In the liver, part of the cholesterol is converted into bile acids and thus removed from its interconversion cycles (3.5.9.2). This removal of body cholesterol from plasma explains the reversed relationship of HDL-cholesterol (especially HDL_2-cholesterol) and atherosclerosis risk.

Scavenger pathway (not shown): Blood monocytes, which adhere to the intact epithelium, may migrate into the subendothelial space and differentiate to macrophages (Fig. 8.1-5). Native LDL is not readily taken up by macrophages. However, oxidative modifications of LDL enhance the LDL uptake by the 'scavenger receptor' (6.2.4) of the macrophages, which causes massive deposition of cholesterol ester droplets, thus converting macrophages into foam cells. Lipoprotein (a) and possibly other Apo B-containing, oxidatively modified lipoproteins may also convert macrophages into foam cells. These components accumulate in atherosclerotic lesions and therefore may be 'risk factors' for atherosclerosis, especially for coronary heart disease. In contrast to the LDL receptor in liver, adrenal and peripheral cell macrophages are not able to downregulate the 'scavenger receptor' at the cell surface to control lipid uptake. As long as the macrophages are not too overloaded with cholesterol, excessive free cholesterol can be removed by HDL.

6.2.3 Lipid Transport Proteins

Cholesterol ester transport protein (CETP) is produced in the liver and secreted into plasma. Here it apparently effects a net transfer of triacylglycerols from chylomicrons and VLDL to HDL and concomitantly of cholesterol esters (but not of free cholesterol) in the opposite direction. The latter activity is possibly a risk factor, the actual metabolic role of CETP remains unclear. CETP also plays a role in intracellular lipid metabolism of adipocytes.

The ABCA1 lipid transporter enables cholesterol and phospholipids transfer from cells into the blood for synthesis of HDL. Mutations in the ABCA1 gene cause inability of cholesterol transport out of the cells, resulting in a deficiency of high-density lipoproteins in the circulation. This is a risk factor for coronary heart disease. Additionally, the build-up of cholesterol in cells can be toxic, causing cell death or impaired function. Both lead to the symptoms of Tangier disease.

6.2.4 Lipoprotein Receptors

The receptors effect the uptake of lipoproteins into the cells (6.2.2). The apolipoproteins at the lipoprotein surface act as ligands for specific receptors and thus have a 'targeting' effect.

LDL receptor (Fig. 6.2-3): This glycoprotein of 160 kDa binds Apo B and E as ligands. This results in endocytosis (6.1.5) and transport of LDL to lysosomes where the proteins (almost exclusively Apo B) are degraded; triacylglycerols and cholesterol esters are cleaved. Free cholesterol can react in various ways:

- It can enter the intracellular cholesterol pool (important for membrane synthesis and for regulatory effects).
- Excessive quantities are recycled by cytosolic acyl cholesterol acyltransferase (ACAT) to cholesterol esters, which are deposited as lipid droplets in the cell.
- In liver, it is converted to bile acids (3.5.9.1) and secreted via the gallbladder into the cystic ducts eventually serving in the intestine

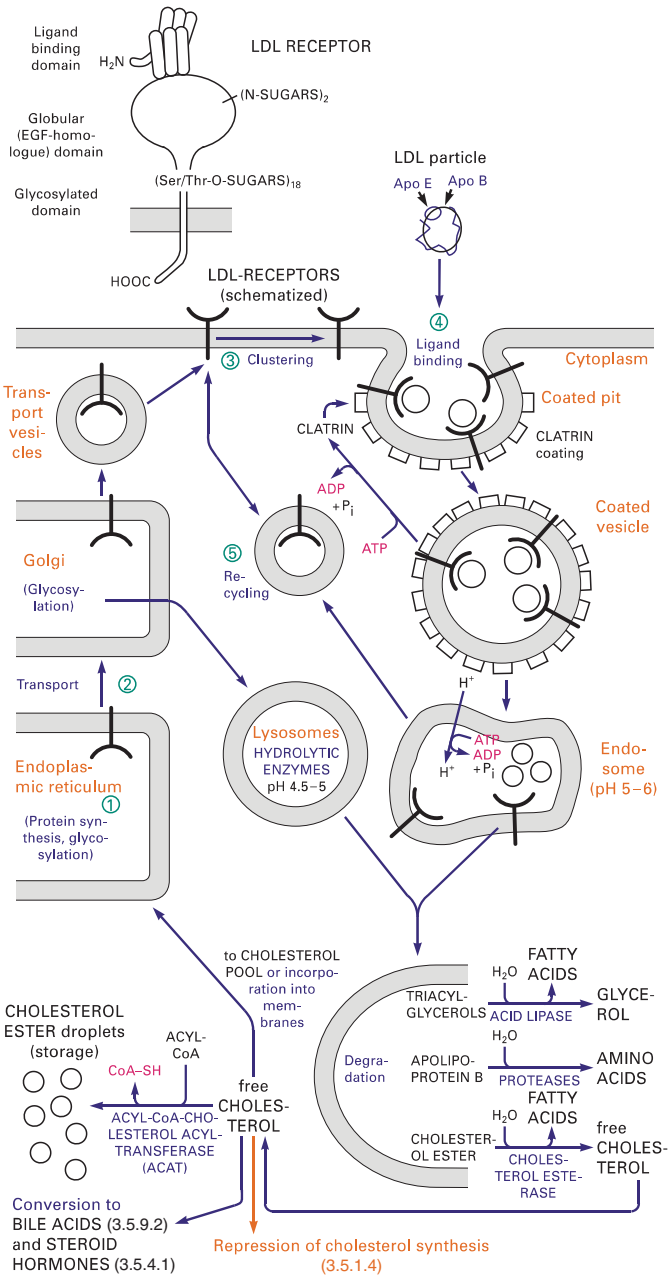


Figure 6.2-3. LDL Structure and Receptor Mediated Endocytosis of LDL

as a lipid emulgator. More than 80 % are later reabsorbed and returned to the liver, the rest is excreted in the feces.

- In specialized tissues (e.g., adrenal glands) cholesterol is converted to steroid hormones (3.5.4).

The LDL-receptor is synthesized at the ribosomes of the ER ①, transported to the Golgi apparatus and ② intercalated into the cell membrane. After clustering at the surface of coated pits ③ (6.1.5) and binding of LDL or IDL via Apo B-100 ④ the receptor-ligand complexes are internalized as coated vesicles. They are then converted into endosomes. Here, a more acidic pH exists, which leads to the dissociation of the ligand-receptor complex. The receptor is recycled to the surface of the cell ⑤. The endosomes combine with lysosomes. Subsequently, apolipoproteins, cholesterol esters and triacylglycerols are hydrolyzed. The resulting increase of the 'cholesterol pool' has several consequences:

- suppression of the transcription of the HMG-CoA reductase gene (3.5.1.4)
- acceleration of the degradation of the HMG-CoA synthase and reductase, geranyl transferase and squalene synthase (3.5.1.4)
- activation of acyl cholesterol acyltransferase (ACAT, 3.5.1.5)
- lowering of the concentration of mRNA coding for the LDL receptor (3.5.1.4)

The first three steps decrease the level of intracellular free cholesterol, while the fourth step leads to a lack of LDL-receptors on the cell surface and therefore to a dramatic decrease of cellular LDL uptake. As a consequence, plasma LDL rises. High plasma LDL leads to cholesterol deposits in skin and tendons (xanthoma), but especially in arteries (plaques). Thus, it is one of the main risk factors of atherosclerosis. In familial hypercholesterolemia, one of steps ① to ⑤ is defective (see 6.2.5).

Scavenger receptor: This trimeric integral membrane glycoprotein is composed of three 77 kDa subunits. It is located at the cell membrane of macrophages and mediates with broad specificity the uptake of chemically modified LDL (like oxidized LDL, glycated LDL, acetylated LDL etc.), which are poorly or not at all recognized by the LDL receptor.

Chylomicron remnant receptor: It is mainly responsible for the clearance of chylomicrons and recognizes Apo E as ligand. It has been presumed that the low-density lipoprotein receptor related protein (LRP) is this receptor. LRP appears to be a large relative of the LDL receptor; the extracellular domain of the LDL receptor is repeated about four times. LRP is definitely the receptor for α_2-macroglobulin. Whether or not it has a dual function is not known for certain at present.

HDL receptors: Their existence has not yet been completely proven. High affinity binding sites for HDL with specificity for Apo A-I, A-II and A-IV as ligands exist on many cell types. HDL receptor protein candidates have been described ranging from 58 to 120 kDa. Not much definite knowledge exists about the cholesterol removal mechanism from peripheral cells by HDL and its uptake by the liver ('reverse cholesterol transport').

6.2.5 Lipid Metabolic Disorders

Some thirty genetic dyslipoproteinemias are known. In most cases, the affected genes and their mutations are also well characterized. Best known are disturbances in synthesis and function of the LDL receptor, numbered ① to ⑤ in Figure 6.2-3. They cause familial hypercholesterolemia, a disease with extremely elevated levels of plasma cholesterol which usually leads to premature coronary heart disease.

Since atherosclerosis mainly seems to be caused by disturbances of lipoprotein metabolism (elevated LDL cholesterol, low HDL cholesterol, high triacylglycerols, foam cell formation etc.), lipid lowering therapy is the treatment of choice. If dietary means do not suffice, 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors, fibrates, nicotinic acid and derivatives, bile acid-binding ion exchangers, probucol (an antioxidant) and other drugs are applied. In extreme cases it may be necessary to remove LDL cholesterol by extracorporeal plasmapheresis.

A novel approach for lowering CHD risk is the development of drugs elevating HDL-C. A promising attempt is the inhibition of cholesterol ester transfer protein, CETP, leading to slower degradation of HDL-C particles. The potential benefits from CETP inhibition may not only be an increased HDL-C concentration, but also in qualitative changes of the particles which promote cholesterol efflux.

Literature:

Assmann, G. *Lipid Metabolism and Atherosclerosis*. Schattauer, (1982).

Brown, M.S., Goldstein, J.L. Science 1985;232:34–47.

Krieger, M., Herz, J. Ann. Rev. of Biochem. 1994;63:601–637.

Morton, I.L. J. Biol. Chem. 2007;282:21856–21865.

Rust, S. *et al*. Nature Genetics 1999;22:352–355.

Schettler, G., Habenicht, A.J.R. (Eds.) *Principles and Treatment of Lipoprotein Disorders*. Springer, 1994.

Shah, P.K. European Heart Journal 2007;28:5–12.

Vance, D.E., Vance, J. (Eds.) *Biochemistry of Lipids, Lipoproteins and Membranes*. New Comprehensive Biochemistry Vol. 31. Elsevier, 1996.

Survey on apolipoproteins: http://www.nlm.nih.gov/cgi/mesh/2007/MB_cgi?mode=&term=Apolipoproteins&field=entry#Tree D10.532.091).

6.3 Oxygen Transport by Hemoglobin

Gerhard Michal

6.3.1 Biosynthesis and Properties of Hemoglobin and Myoglobin

Hemoglobin (Hb) has the task of delivering oxygen to the organs of vertebrates. Invertebrates use different compounds: hemocyanin (contains Cu^{++}) or hemerythrin (contains nonheme-Fe). The O_2 transport by proteins is necessary since O_2 solubility in blood plasma is too low to satisfy the oxygen demand of the tissues (for an exception see 6.3.2). Myoglobin takes care of the oxygen storage and transport inside of rapidly respiring muscles. Mammals living in water also use myoglobin for oxygen storage. Hemoglobin and myoglobin contain 4 and 1 heme groups, respectively, which are surrounded by globulin.

Heme biosynthesis (3.3.1.1) is regulated in different, organ-dependent ways. In higher animals, ca. 85 % of the heme production takes place in reticulocytes (immature erythrocytes, Fig. 8.1-5) to provide hemoglobin, resulting in 280 million hemoglobin molecules per erythrocyte. Most of the rest is produced in the liver and is then converted into cytochrome P-450. This is the prosthetic group of hydroxylases, which are involved in detoxifying and in biosynthetic reactions (Fig. 3.5.4-2).

The iron supply for heme production is resorbed in the intestine and transported via protein carriers to these cells (Fig. 6.3-1), since permanent binding of Fe to proteins greatly reduces the generation of reactive oxygen species by free Fe (3.2.5.8). For intermediate storage, iron is bound to ferritin.

Globin biosynthesis: Hemoglobin is a tetramer of globin subunits. In humans, they change during life: embryonic state $\zeta_2\varepsilon_2$ or $\alpha_2\varepsilon_2$, fetal state $\alpha_2\gamma_2$ = HbF, adult state 97.5 % $\alpha_2\beta_2$ = HbA_1 and 2.5 % $\alpha_2\delta_2$ = HbA_2. Their genes are located at chromosomes 11 and 16, respectively, and are expressed at strictly balanced levels.

In reticulocytes heme enhances the rate of globin synthesis. The regulation takes place at the translational level (Fig. 6.3-2, see also 4.2.3.2). Glycosylated p67 (eIF-2 associated 67 kDa polypeptide) protects the translation factor eIF2α (Fig. 4.2.3-2) from eIF2α-kinase (EIF2αK1)-catalyzed phosphorylation and inactivation. Thus, it protects

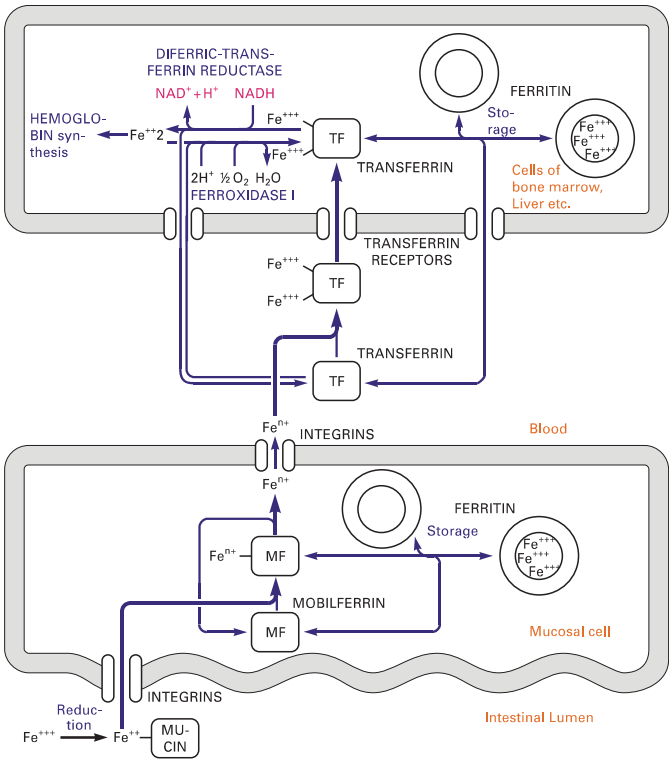


Figure 6.3-1. Transport and Storage of Iron in Higher Animals

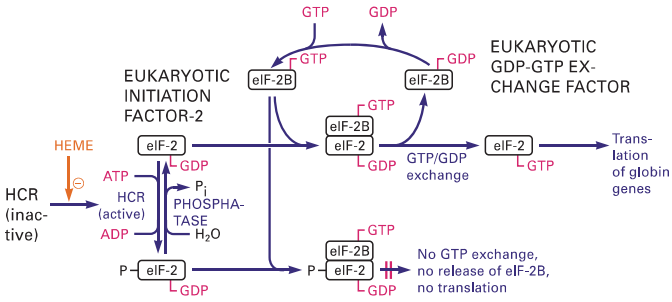


Figure 6.3-2. Control of Globin Synthesis by p67 Polypeptide

α-globin synthesis. Mitogens elevate the p67 concentration and increase protein synthesis further.

In heme deficient reticulocytes, p67 is deglycosylated and degraded. This causes phosphorylation of eIF2. The phosphorylated factor firmly binds the GDP-GTP exchange factor eIF-2B and thus prevents any further GDP-GTP exchange at eIF-2, which is necessary for its function in protein synthesis. Also, Met-tRNA$_f$ is bound to the 40S subunit. The initiation of globin gene transcription stops.

The synthesis of the globin β chain is controlled by the α chain. In β thalassemia, the β globin expression is reduced and an excess of unbound α chains is formed. The mild form (β thalassemia minor) affects only one of the two β globin alleles, the severe form (β thalassemia major) both. The surplus α chains form long insoluble precipitates. If the α-expression is likewise reduced, however, improvements take place. In α thalassemia, the excess β chains form unstable tetramers, which have abnormal oxygen dissociation curves.

Structure (Fig. 6.3-4): Hemoglobins are tetramers of roughly globular shape, containing 2 * 2 identical subunits (in adults $\alpha_2\beta_2$, 4 * 17 kDa), which are kept together by many noncovalent interactions. Each subunit contains one heme molecule, which is partially hidden in a 'pocket'. It is surrounded by 8 α-helices formed from mostly nonpolar amino acids except for two histidine residues on both sides of the heme ring ('proximal' and 'distal' histidine).

Myoglobin consists of a single chain with one heme molecule, which closely resembles a hemoglobin subunit (16.7 kDa).

In hemoglobin and in myoglobin, the coordination sites of Fe are occupied by the 4 porphyrin-N and the proximal histidine residue. The occupation of the sixth site and the oxidation state of Fe differ (Table 6.3-1).

Table 6.3-1. Environment of the Heme Group

Hemo-/myoglobin	Sixth Ligand	Oxidation state of Fe
Deoxy-	None	II
Oxy-	O_2 (loosely bound)	II
CO-	CO (firmly bound)	II
Met- (= Ferri-)	H_2O (firmly bound)	III

During hemoglobin synthesis, the ferrous ion of heme gets linked to the proximal histidine of the nascent globin chain, while the porphyrin ring is wedged in place by a phenylalanine on the globin chain. Heme binding also stabilizes the growing globin chain on the ribosome and promotes proper folding of the native tertiary structure of α-globin. This complex facilitates also the heme integration into β-globin. Excess α-globin is unstable, undergoes auto-oxidation and produces methemoglobin and superoxide (3.2.5.8). As a protection, it is stabilized intermediately by α-hemoglobin stabilizing protein (AHSP), which is later displaced by β-globin.

6.3.2 Oxygen Binding to Hemo- and Myoglobin

O_2 molecules slip freely between the distal histidine in the heme pocket and the ferrous iron. It binds with free heme and with hemo- or myoglobin at an angle to the axial direction. While CO binds to free heme likewise in axial direction, in hemo- or myoglobin it is forced by the distal histidine to obtain a more unfavorable tilted position. This reduces the binding affinity of CO considerably (from 25 000 times the affinity of O_2 to 200 times) and provides partial protection against the CO produced by heme degradation.

Under normal conditions, the transition from deoxy- to oxyhemo-/oxymyoglobin and back does not affect the Fe^{2+} oxidation state. Occasionally, the ferrous iron binds spontaneously to the distal histidine and is oxidized to ferric iron forming methemoglobin (in about 1 % of the circulating hemoglobin). In erythrocytes, the NADH-dependent methemoglobin reductase reconstitutes the Fe^{2+} state. The formation of methemoglobin is increased in mutations, which change amino acids in the neighbourhood of the heme-Fe and thus stabilize the Fe^{3+} state (methemoglobinemia).

Nitrogen monoxide (NO, 7.8.2) is oxidized by myoglobin-O_2 under formation of metmyoglobin:

$$\text{Myoglobin}(Fe^{++})\text{-}O_2 + NO = \text{metmyoglobin}(Fe^{+++}) + NO_3^-$$

This diminishes the effects of NO on vasodilatiation and its cardiac action. Metmyoglobin is reconverted to myoglobin by NADH-dependent metmyoglobin reductase.

Hemoglobin(Fe^{++})-O_2 can react in the same way with NO, while deoxygenated hemoglobin(Fe^{++}) can act as a carrier of NO (7.8.2) or act as a reductant on nitrite under formation of NO (The latter reaction is shared with other heme-containing proteins).

$$\text{Hemoglobin}(Fe^{++}) + H^+ + NO_2^- = \text{hemoglobin}(Fe^{+++}) + NO + OH^- .$$

Oxygen binding by the monomeric myoglobin takes place without major structural modifications. This is different in hemoglobin: Each of its four subunits can bind one oxygen molecule. This causes considerable modifications in the tertiary (folding state) and quaternary structure (association of subunits, 2.3.1).

Hemoglobin oxigenation involves a chain of events (Fig. 6.3-3):

- Fe, which was located ca. 0.06 nm outside the plane of the heme ring, moves to its center. This enables the contact of Fe with O_2.
- The proximal histidine is pulled closer to the heme plane.
- The 'F' helix, which contains the proximal histidine, is tilted and laterally moved about 0.1 nm across the heme plane.

- This structural change effects a movement of the subunits about 0.6 nm relative to each other along the $\alpha_1\beta_2$ and $\alpha_2\beta_1$ interfaces. This is a transition from one stable state to another (from T to R state, Fig. 6.3-4). In this way, the oxygen binding to one subunit of hemoglobin affects the configuration of the other subunits and gives rise to allosteric behavior (for details, see 2.5.2).

Dissociation curves: The O_2 dissociation curve for the monomeric myoglobin has the usual hyperbolic shape, while the curve for the tetrameric hemoglobin is sigmoid (Fig. 6.3-5).

For myoglobin, the following formulae can be derived analogously to the formulae for enzyme-substrate binding (1.5.4).

The dissociation constant K_D for the equilibrium state

$$Mb + O_2 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} Mb - O_2 \quad \text{is calculated as} \quad K_D = \frac{k_{-1}}{k_1} = \frac{[Mb]*[O_2]}{[Mb-O_2]}.$$

Using $[Mb_t]$ for the total myoglobin concentration [Mb] + [Mb–O_2], a derivation analogous to the Michaelis-Menten equation [Eq. 1.5-19], but without further turnover to products, yields

$$[Mb{-}O_2] = \frac{[Mb_t] * [O_2]}{[K_D + O_2]} \quad \text{or} \quad \frac{[Mb{-}O_2]}{[Mb_t]} = Y = \frac{[O_2]}{K_D + [O_2]}$$

where Y is the degree of myoglobin saturation. This results in a hyperbolic plot of Y vs. $[O_2]$. The term $[O_2]$ can be expressed as partial pressure pO_2.

The hemoglobin tetramer, however, shows allosteric behavior (2.5.2). The derivation of the equation has to consider the positive cooperativity: The binding of the second, third and fourth oxygen molecules proceeds gradually more easily than that of the first one. The most radical mathematical fomulation is the 'all or none'-principle, disregarding the intermediates Hb-O_2, Hb-$(O_2)_2$ and Hb-$(O_2)_3$. Then the dissociation constant K_D for the equation

$$Hb + 4O_2 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} Hb - (O_2)_4 \quad \text{is calculated as} \quad K_D = \frac{k_{-1}}{k_1} = \frac{[Hb]*[O_2]_4}{[Hb-(O_2)_4]}.$$

Using the substitution $[Hb_t]$ for the total hemoglobin concentration [Hb] + [Hb-$(O_2)_4$] analogously to above, the following equation results:

$$[Hb{-}(O_2)_4] = \frac{[Hb_t] * [O_2]^4}{K_D + O_2]^4} \quad \text{or} \quad \frac{[Hb{-}(O_2)_4]}{[Hb_t]} = Y = \frac{[O_2]^4}{K_D + [O_2]^4}.$$

The generalized case, considering the partially oxygenated intermediates is described by the Hill equation. The exponent 4 in case of tetrameric proteins is replaced by n. This Hill coefficient n is smaller than the number of binding sites. It can be taken as a 'measure of cooperativity'. For hemoglobin with 4 binding sites, Hill coefficients of 2.8 … 3.0 have been observed, indicating a fairly high cooperativity. Correspondingly, the share of partially oxygenated species is relatively small (Fig. 6.3-6). A more refined treatment of cooperativity is given by the symmetry and the sequential models (2.5.2).

Influence of pH and CO_2: The conformation change of hemoglobin, which is caused by oxygen binding, also effects a release of protons.

$$Hb{-}(O_2)_z + O_2 = Hb\text{-}(O_2)_{z+1} + y\ H^+ \quad (y \approx 0.6)$$

Therefore, a more alkaline pH (with lower proton concentration) 'pulls' the reaction towards higher oxygen binding, while more acid pH faciliates the oxygen release (Bohr effect, Fig. 6.3-5). Since capillary blood is more acidic due to the CO_2 production of the tissues and its conversion to HCO_3^- + H^+ by carbonic anhydrase, the oxygen

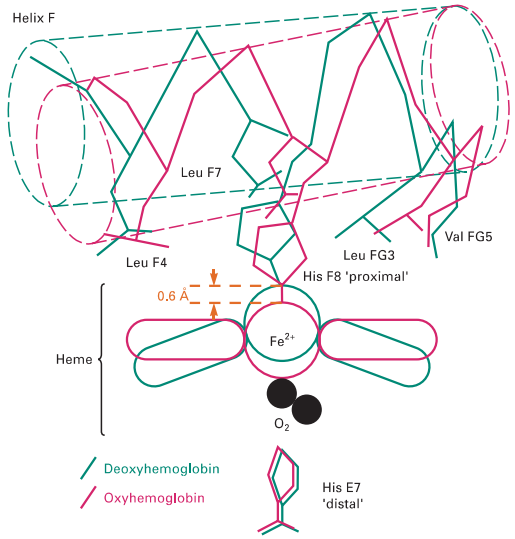


Figure 6.3-3. Structural Shifts During Oxygenation of Hemoglobin (green: T state, red: R state)

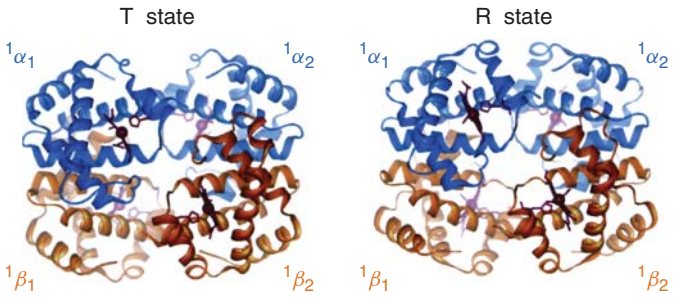


Figure 6.3-4. Structure of Hemoglobin (T and R States)
From Wikipedia: hemoglobin t-r states ani.gif.

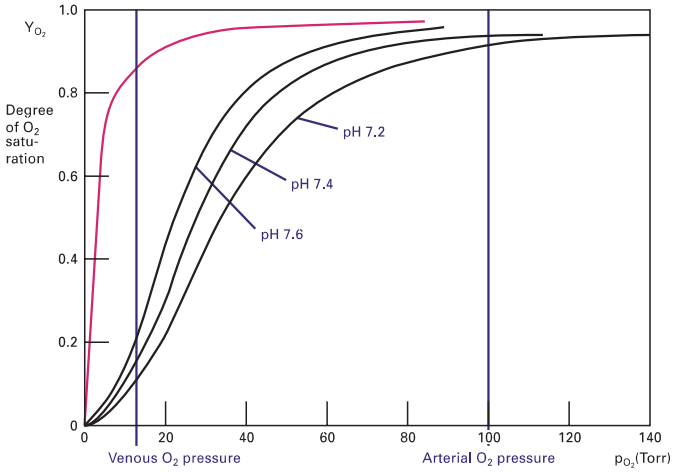


Figure 6.3-5. Oxygen Dissociation Curves for Hemoglobin —— and Myoglobin ——

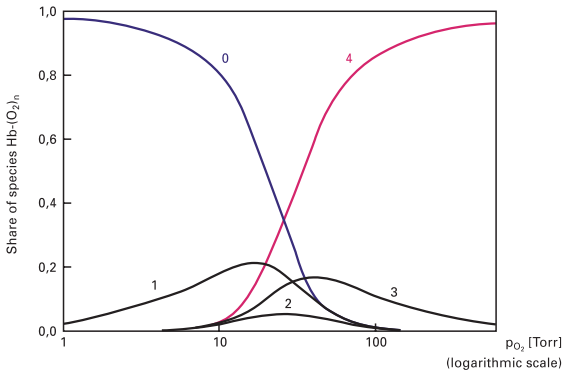


Figure 6.3-6. Share of the Hemoglobin Species With Different Degrees of Oxygen Saturation

transfer from the capillaries into the tissues is enhanced. On the other hand, CO_2 removal in the lung facilitates the saturation of hemoglobin with O_2.

In addition, CO_2 interacts directly with hemoglobin and causes conformation changes, which influence the O_2 binding in the same direction as the pH effect.

2,3-Bisphosphoglycerate (3.1-1), which is present in mammalian erythrocytes, binds more strongly to deoxygenated hemoglobin than to the oxygenated form. As a result, the hemoglobin saturation curve is shifted to higher O_2 concentrations. This allows easier dissociation of O_2 from hemoglobin. The same effect is achieved in birds by inositol hexaphosphate (7.4.4) and in fishes and most amphibians by ATP.

Additional understanding of the role of hemoglobin and myoglobin came from the study of antarctic icefishes (family *Channichthydiae*). They are living permanently in cold water with high oxygen content and have lost hemoglobin and myoglobin. Since the oxygen carrying capacity of blood dropped by 90 %, the supply rate has to be increased by larger diameter of the vessels, elevated flow rate and tissue vascularization. Also, the density of mitochondria is elevated as compared with other fishes.

The other extreme is the strong increase in oxygen storage capacity in, e.g., whales, which have to stop breathing when diving: the myoglobin content in muscles of sperm whale is about 80 g/kg muscle vs. 8 g/kg muscle in humans.

6.3.3 Hemoglobin Diseases in Humans

More than 300 genetic abnormalities of hemoglobin are known. Sickle cell anemia is caused by a mutation in the β-chains of hemoglobin ('Hb S', Glu 6 → Val). This does not affect the properties of oxygenated hemoglobin. However, in the shifted structure of deoxygenated HbS, hydrophobic interactions of the mutated β_2-Val 6 with β_1-Phe 85 and Leu 88 of another Hb molecule are possible. This decreases the solubility and leads to the formation of fibrous HbS strands, which change the erythrocyte shape into a 'sickle' form. The shorter lifetime of the affected erythrocytes apparently provides some protection against malaria. The disease is prevalent among the African and African-Americans population in Africa and America.

α and β thalassemias are disorders of hemoglobin synthesis (6.3.1), often caused by mutations in regulatory genes. The α or the β globin chains are either absent or are synthesized only at a reduced rate. α Thalassemia is found in people of West African descent, β thalassemia is prevalent in the Mediterranean area. Both diseases show some protection against malaria. Methemoglobinemias have been mentioned above.

Literature:

Allen, B.W. *et al*. Trends in Molecular medicine 2009;15:452–460.

Baldwin, J. Trends in Biochem. Sci. 1980;5:224–228.

Beale, S. in Neidhardt, F.C. *et al.* (Eds.) *Escherichia coli* and *Salmonella.* ASM Press, 1996; p. 731–748.

Dondekar, T. *et al.* EMBO J. 1991;10:1903–1909.

Flögerl, U., Merx, M.W. *et al*. Proc. Natl. Acad. Sci (USA) 2001;98:735–740.

Joshi, R. *et al*. Eur. J. Biochem. 1995;228:31–38.

Perutz, M.F. Ann. Rev. of Biochem. 1979;48:327–368.

Ray, M.K., Datta, B *et al*. Proc. Natl. Acad. Sci (USA) 1992;89: 539–543.

Sidell, B.D., O'Brien, M.O. J. Exp. Biol. 2006;209:1791–1802.

Tota, B *et al*. Current Medicinal Chemisry 2010;17:1915–1925.

Voon, H.P.J., Vadolas, J. Haematologica 2008;93:1868–1876.

Warren, M.J., Scott, A.I. Trends in Biochem. Sci. 1990;15:486–491.

Watowich, S.J. *et al.* J. Mol. Biol. 1989;209:821–828.

Wittenberg, J.B., Wittenberg, B.A. J. Exp. Biol. 2003;206:2011–2020.

7 Signal Transduction and Cellular Communication

Gerhard Niederfellner

Multicellular organisms require signal transduction systems to coordinate the cellular activities. This can take place the humoral way (hormones in plants and animals) or the neuronal way (only in animals). This chapter deals with the situation in vertebrates.

Both communication systems cannot be exactly separated; there are various connections between them. In the following, hormones are discussed first (7.1), acting mainly extracellularly both on non-neural cells and within the nerve system. Thereafter, neuronal signal transmission (7.2) and finally intracellular signal processing (7.3 … 7.7) are described.

7.1 Intercellular Signal Transmission by Hormones

7.1.1 General Characteristics of Hormones (Table 7.1-1)

Hormones are chemical signal transmitters, which are transferred humorally to their target cells. They are synthesized in

- hormone glands (glandular hormones, this section)
- various cells [non-glandular hormones, such as tissue hormones (7.1.9, 7.4.8, 9.2.2) and cytokines (8.1.8)]
- nerve cells (neurohormones, 7.2).

They act on the targets in different ways:

- autocrine: acting on the same cell
- juxtacrine: membrane bound hormones need direct contact to an adjacent target cell
- paracrine: acting on neighboring cells
- endocrine: acting on distant cells.

Hormones belong to various chemical classes. Peptide hormones are frequently synthesized as prohormones, which are later cleaved to yield the final hormone. Many hormones are stored in vesicles (secretory granula) and released upon arrival of appropriate signals, e.g., changes of the intracellular Ca^{++} concentration (7.4.4). The hydrophobic steroid and thyroid hormones are bound to carrier proteins for their transport in blood.

Hormone concentrations are usually very low (10^{7} … 10^{-12} mol/l). Frequently, a hormone cascade exists, where an endocrine hormone causes the generation of another one in a different tissue and so on. A functional grouping of hormones is given in Table 7.1-1.

7.1.2 General Characteristics of Receptors

After reaching the target cells, the hormones (being agonists) bind specifically and usually with high affinity to receptors, which thereafter initiate specific intracellular actions. These are described in the following sections. There are three types of receptors:

- transmembrane proteins, where binding of the hormone to the extracellular portion causes a conformation change, which effects a transmembrane signal transduction resulting in an intracellular response via second messengers (7.4) or via phosphorylation cascades (7.5).
- ligand-gated ion channels. After binding of the agonist, these transmembrane proteins either open or close integral ion channels. They play a central role in nerve conduction (7.2). The agonist may be an extracellular hormone or an intracellular agonist of the second messenger type, e.g., in regulation of the intracellular Ca^{++} concentration (7.4.4), sensing of light (7.4.6), odors and tastes (7.4.7).

Table 7.1-1. Hormones Grouped According to Function

	Group	Examples	Steroids, Lipids	Amino Ac.deriv	Polypeptides	Section	Acting via (Receptor = R)	Section
I	Cytokines (non-glandular growth and differentiation hormones)	PDGF, EGF, FGF, IGF-1+2, TGF-α			*	7.1.5	R Tyr kinases	7.5.3
		GM-CSF, G-CSF, EPO, interferons, interleukins, TNF			*	8.1.8	Tyr kinase associated receptor	7.5.4
II	Glandular growth and differentiation hormones	growth hormone (GH)			*	7.1.5	Tyr kinase associated receptor	7.5.4
		thyroid hormone		*		7.1.5	thyroid hormone R	7.7
		sex hormones, glucocorticoids	*			7.1.5 … 7	steroid hormone R	7.7
III	'Fast' hormones	insulin			*	3.1.1.4, 7.1.3	R Tyr kinases [1]	7.5.3
		glucagon			*	3.1.2.4, 7.1.3	G-protein R/cAMP	7.4.2
		catecholamines[2]		*		3.2.7.4, 7.1.4	G-protein R/cAMP	7.4.2
IV	Gastrointestinal hormones (digestion and resorption)	secretin			*	7.1.9	G-protein R/cAMP	7.4.2
		gastrin			*	7.1.9	G-protein R/PLCβ	7.4.3
V	Ca^{++} and phosphate metabolism regulating hormones	parathyroidhormone, thyreocalcitonin			*	7.1.7	G-protein R/cAMP	7.4.2
		vitamins D	*			3.7.11, 7.1.7	steroid hormone R	7.7
VI	H_2O and electrolyte metabolism regulating hormones	vasopressin			*	7.1.8	G-protein R/cAMP	7.4.2
		atrial natriuretic factor			*	7.1.8	guanylate cyclase	7.8
		angiotensin			*	7.1.8	G-protein R/PLCβ	7.4.3
		mineralocorticoids	*			3.5.8.3,7.1.8	steroid hormone R	7.7
VII	Releasing and inhibit. hormones	CRH, TRH; ACTH, TSH, FSH, LH, etc.			*	7.1.5 … 8	G-prot.R/PLC, cAMP	7.4.2, 7.5.3
VIII	Neurotransmitters	acetylcholine, dopamine, glutamate, etc.		*		7.2.3	various R	7.2
		endorphins, enkephalins			*	7.2.3	opiate R	7.2.3
IX	Others							

[1] Act also on cAMP phosphodiesterase antagonistically to glucagon (7.4.2).
[2] Also neurotransmitter (7.2.3).

Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Second Edition. Edited by Gerhard Michal and Dietmar Schomburg.

- intracellular proteins. The hormones, e.g., steroids, thyroid hormones and retinoic acid pass through the cellular membrane and meet their receptors intracellularly in the cytosol or in the nucleus (7.7). The receptor then modulates the transcription of genes.

In all cases, a deactivating mechanism is required for termination of the signal effect. This can be

- inactivation of the receptor by, e.g., specific phosphorylation/dephosphorylation
- desensitization of the receptor
- internalization of the receptor-ligand complex (similar to the lipoprotein-receptor complex, Figure 6.2-3, but with degradation of receptor and ligand.)

Kinetics of hormone binding: Similarly to enzymes (1.5.4), the hormone H and the receptor R can reach equilibrium with the hormone-receptor complex HR. This can be written as a concentration equation:

$$[H] + [R] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [HR]$$

with the dissociation constant $K_D = \frac{[H] * [R]}{[HR]} = \frac{k_{-1}}{k_1}$

Using $[R_t]$ for the total receptor concentration $[R] + [HR]$, a derivation analogously to the Michaelis-Menten equation [1.5-19] yields:

$$[HR] = \frac{[H] * [R_t]}{K_D + [H]} \quad \text{or} \quad \frac{[HR]}{[R_t]} = Y = \frac{[H]}{K_D + [H]}.$$

whereby Y is the part of the receptor bound to the hormone. This can be written as

$$\frac{[HR]}{[H]} = -\frac{1}{K_D}([HR] - [R_t]),$$

where $\frac{[HR]}{[H]}$ is the ratio of bound/free hormone.

A plot of this ratio [HR]/[H] versus the concentration of bound hormone [HR] yields theoretically a straight line with the slope $-(1/K_D)$, the abscissa intersection $[R_t]$ and the ordinate intersection $[R_t]/K_D$ (Scatchard plot, Fig. 7.1-1). This allows, e.g., the calculation of the dissociation constant and the receptor concentration. However, frequently the plot is not linear, e.g., if a population of receptors with different dissociation constants is present.

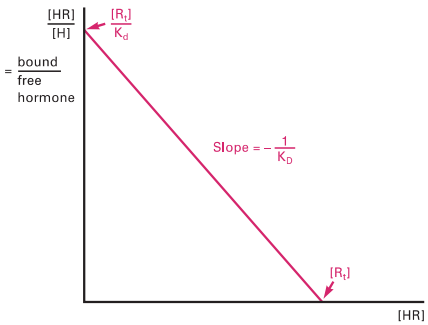


Figure 7.1-1. Scatchard Plot of a Hormone-Receptor Interaction

7.1.3 Insulin and Glucagon

Insulin is the most important anabolic hormone and acts primarily on muscle, adipose tissue and liver. Its binding to the insulin receptor induces the synthesis of glycolytic and represses the synthesis of gluconeogenetic enzymes (Table 3.1.1-1, 7.5.3). The fast metabolic effects of insulin (e.g., direct control of the blood glucose level, carbohydrate and fatty acid metabolism, 3.1.1.5) proceed by influencing the cAMP level, since insulin inhibits the adenylate cyclase and activates the 3′,5′-cyclic nucleotide phosphodiesterase. Details are described in 3.1.2.4, 3.4.1.2 and 7.4.2.

Insulin is synthesized in the b cells of Langerhans islets in the pancreas. Its expression is regulated by glucose concentration. Insulin is stored in secretory granula of the β cells. It is released, when the cellular glucose 6-phosphate concentration rises (sensor: ATP/ADP ratio, transmitted via the K_{ATP} channel and voltage-gated Ca^{++} channels, Table 6.1-1). Insulin secretion is enhanced by the gastric inhibitory peptide and decreased by epinephrine and somatostatin.

The primary transcription product (prepro-insulin, 104 … 110 amino acids, species dependent) is converted into proinsulin in the endoplasmic reticulum (removal of the signal peptide, formation of 3 S–S bonds) and into mature insulin in the Golgi apparatus (removal of the intermediary C peptide, Fig. 7.1-2).

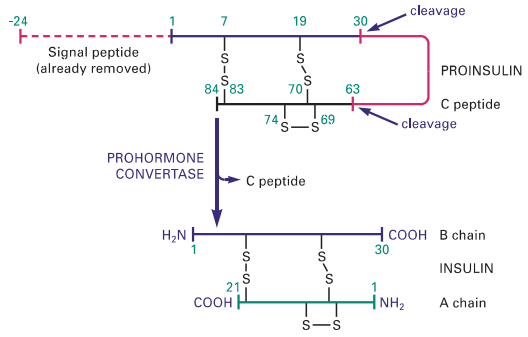


Figure 7.1-2. Structure of Insulin and Its Precursors

In diabetes type I (juvenile diabetes), the production of insulin in the pancreas is disturbed due to destruction of β-cells of Langerhans islets in the wake of autoimmune diseases or virus infections, while in diabetes type II at first there is impaired signal transduction by insulin receptors (insulin resistance) followed by progressive failure of the insulin release mechanism to respond properly to elevated blood glucose levels (insulin secretory dysfunction). Diabetes is characterized by elevated blood sugar, increased lipolysis, formation of ketone bodies (3.4.1.7), loss of electrolytes, and acidosis.

Glucagon, being a catabolic hormone, is the metabolic opponent of insulin. It is synthesized in a-cells of Langerhans islets in the pancreas. It is released when the glucose concentration drops. Upon binding to its receptor in the liver, it stimulates adenylate cyclase via G-proteins (7.4.2). The cAMP formed, in turn, activates a protein kinase, which increases the activity of phosphorylase and decreases the activity of glycogen synthase (3.1.2.4).

Analogous to insulin, the synthesis proceeds via the inactive precursor prepro-glucagon and several intermediates. In the intestine, prepro-glucagon is cleaved differently, yielding the glucagon-like peptides GLP-1 and GLP-2, which promote insulin secretion after food intake.

7.1.4 Epinephrine and Norepinephrine (Catecholamines)

These compounds are synthesized from tyrosine (3.2.7.4) in postganglionic nerve terminals and in the adrenal medulla (Fig. 7.1-3). The catecholamines are stored in specific granula as complexes with ATP and Mg^{++} and are released after arrival of neuronal signals, which are transmitted by acetylcholine. This takes place in physical and psychological stress situations.

The catecholamines act on several adrenergic receptor types, producing multiple effects (Table 7.1-2). The mechanism of action is described in detail in Sections 7.4.2 and 7.4.3.

The expression of the biosynthetic enzymes tyrosine 3-monooxygenase (hydroxylase) and dopamine β-hydroxylase is induced by neuronal signals, the expression of phenylethanolamine N-methyltransferase by glucocorticoids. Catecholamines, in turn, promote the synthesis of regulators of glucocorticoid formation (CRH, ACTH), so that the result is an upregulating circuit (Fig. 7.1-3).

7.1.5 Hypothalamus-Anterior Pituitary Hormone System (Fig. 7.1-4)

This complex system is arranged as a cascade: one hormone puts another into action and so on (by enhancing synthesis and/or release). Each step occurs in a different organ. The system regulates many

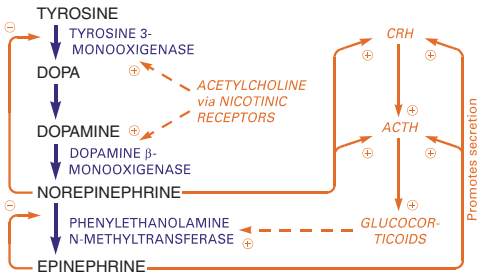


Figure 7.1-3. Regulation of Catecholamine Biosynthesis

Table 7.1-2. Metabolic Effects of Catecholamines

Effect	Adrenergic Receptors	Acting via
Glycogenolysis, gluconeogenesis	α_1; β_1	PLCβ ↑; adenylate cyclase ↑
Lipolysis (adipose tissue)	β_2, β_3	adenylate cyclase ↑
Increased contractility of the heart muscle	β_1	adenylate cyclase ↑
Vasodilatation in the skeletal muscle	β_2	adenylate cyclase ↑ [1]
Vasoconstriction in the intestinal area	α_1	PLCβ → Ca^{++} ↑
Decreased insulin secretion	α_2	adenylate cyclase ↓

[1] cAMP activates PKA, this promotes the uptake of Ca^{++} into the sarcoplasmic reticulum (7.4.4) and decreases the cytosolic Ca^{++} concentration, which controls contraction. Additionally, myosin light chain kinase is phosphorylated and deactivated (7.4.5).

important functions (growth, overall metabolic rate, reproduction, pain control etc.).

External and internal stimuli are processed in the hypothalamus and lead to the secretion of various releasing and inhibitory hormones (liberins and statins), which are transported by the bloodstream to the nearby anterior pituitary lobe (also called adenohypophysis). There, they bind to specific receptors, which regulate the formation and release of trophic hormones (tropins). These trophic hormones, in turn, act on receptors at peripheral tissues and stimulate them to synthesize and release hormones, which thereafter perform their respective functions in their target tissues. The hormones also control the synthesis or release of their own regulators in a feedback fashion (see Fig. 7.1-4).

The hypothalamus (and most of the pituitary) hormones are synthesized as long pre-prohormones and posttranscriptionally processed to the prohormone and further to the actual hormone (similarly as described above for insulin). In some cases, several hormones are cleaved from a common predecessor. All hypothalamus hormones are amidated at the C-terminus.

The melanocyte stimulating hormones (α- and β-MSH) are synthesized in the intermediate part of the pituitary gland. The posterior pituitary lobe (neurohypophysis) performs the final processing of oxytocin and vasopressin, which are synthesized in the hypothalamus (7.1.8). Some releasing hormones are additionally produced in other organs besides the hypothalamus.

Hypothalamo-pituitary-adrenal axis: 7 … 10 times a day, but increased during stress situations, the corticotropin releasing hormone (CRH) is secreted from the hypothalamus and acts on the pituitary gland. Here the proopiomelanocortin (POMC) is synthesized and processed into corticotropin (ACTH, stored in secretory granula), endorphins and α-, β- and γ-melanocyte stimulating hormone (MSH). The CRH and the ACTH secretions occur spasmodically and follow the same rhythm.

Cortisol: In the adrenal cortex, ACTH activates (via cAMP/protein kinase A, 7.4.2) the hydrolysis of cholesterol esters (3.5.1.5) and the *de novo* synthesis of cholesterol (3.5.1.1). It also induces hydroxylases involved in the conversion of cholesterol to cortisol (3.5.8.1) and stimulates cortisol secretion. This steroid is bound to transcortin during transport in blood.

Almost all cells of the organism are targets for cortisol action, where the hormone binds to the internal cortisol receptor (7.7). This induces transcription of genes coding for many catabolic enzymes in extrahepatic tissues (as an antagonist to insulin) and for gluconeogenetic enzymes in liver. It is also an immunosuppressor and a suppressor of inflammation.

The maximum of the wavelike diurnal cortisol secretion is in the morning, the minimum in the evening, with additional spikes corresponding to the CRH/ACTH secretion.

Endorphins and related neuropeptides: Several endorphine isoforms and Met-enkephalin are cleaved from POMC in the pituitary gland, but also in the stomach and intestine, placenta, lung etc. Dynorphin and neoendorphins are cleaved from other precursors. They act on opioid receptors (Table 7.2.2) and cause analgetic effects. They are also involved in body temperature regulation.

Other cleavage products of POMC are melanotropins (MSH, melanocyte stimulating hormones), which promote skin darkening in amphibia and fishes by activating tyrosinase, which increases melanin synthesis (3.2.7.3).

Hypothalamo-pituitary-thyroid axis: The HPT axis constitutes part of an endocrine network that regulates metabolism. The thyrotropin releasing hormone (TRH) of the hypothalamus and other organs stimulates the release of thyroid stimulating hormone (TSH, thyrotropin) from the pituitary gland. This heterodimer (whose α-subunit is identical to subunits of luteinizing hormone and follicle stimulating hormone) binds to the TSH receptor on the thyroid gland, which is of the heterotrimeric G protein associated type (7.4.1). The IP_3 mechanism (7.4.4) activates iodination of thyreoglobulin at multiple tyrosine residues in the thyroid follicles and the lysosomal degradation of the protein, which results in the liberation of thyroid hormones triiodothyronine (T_3) and thyroxine (T_4, for metabolism see 3.2.7.5). The release of these compounds is controlled by cAMP (7.4.2).

During their transport in blood, most of the thyroid hormones are bound to the thyroxine binding globulin (TBG) or other proteins. Only the free compounds are hormonally active. Inside the target cells, T_3 combines with its receptor (7.6, T_4 has to be deiodated first) and induces the expression of many enzymes and regulatory proteins.

In higher vertebrates, the agonist binding leads to enhanced glycolysis, gluconeogenesis and liponeogenesis. The activity of Na^+/K^+-ATPase (6.1) and of lysosomal enzymes is increased. The expression of growth hormones (STH, EGF, IGF, NGF, etc.) promotes growth and cellular differentiation. The increase in adrenergic β-receptors enhances catecholamine actions, e.g., yielding positive effects on heart muscle contraction (inotropic effects) and on heart rate (chronotropic effects), as well as increased thermogenesis.

Abnormal levels of thyroid hormones cause common diseases. Hypothyroidism can originate from genetic mutations at every level of the hormone cascade (causing serious physical and mental retardation of infants, cretinism) or develop postnatal by formation of autoantibodies against thyroid enzymes. Via feedback regulation lack of iodine supply leads to increased secretion of TSH and consecutively to hypertrophy of the thyroid gland (goiter). The most common cause of hyperthyroidism is the presence of autoantibodies, which permanently activate the TSH receptor (Basedow's disease).

Hypothalamo-pituitary-liver/bone axis: Glucose concentration in plasma is sensed by neural glucoreceptors. A decrease leads to secretion of the growth hormone releasing hormone (GRH) by the hypothalamus. This effects the secretion of somatotropin (STH, also named growth hormone, GH) from the pituitary gland (the greater portion during the night). The STH-expression is promoted by triiodothyronine (see above).

The release is antagonized by somatostatin (growth hormone release inhibiting hormone), which originates from many organs (including the hypothalamus) and also inhibits various other reactions (gastrointestinal secretions, release of TSH, insulin, glucagon, etc.).

In the liver, STH stimulates the synthesis and release of somatomedins (insulin-like growth factors, IGF I and II), which bind to specific receptors and promote growth of bone and cartilage, likely also of muscle and adipose tissue.

The IGF I receptor resembles the insulin receptor (a receptor tyrosine kinase), while the IGF II receptor is a tyrosine kinase-associated receptor. Both initiate reaction sequences leading to increased transcription (7.5.3).

Inactivating mutations can occur at every level of the cascade and cause dwarfism. Constitutive STH release occurs in patients with eosinophilic adenoma, a pituitary tumor that permanently activates the adenylate cyclase, and results in gigantism (in children) and in acromegalia (in adults).

Hypothalamo-pituitary-testis axis: The hypothalamus regulates both male and female sex hormone production by the formation and intermittent secretion of the gonadotropin releasing hormone (GnRH, also called luteinizing hormone releasing hormone, LHRH) about 12 … 18 times daily. This causes the release of the gonadotropins [follicle stimulating hormone (FSH) and luteinizing hormone (LH)] from the anterior pituitary lobe.

In the Leydig cells of male testes, LH stimulates the biosynthesis of testosterone and 5α-dihydrotestosterone (3.5.6.1). These hormones bind to intracellular receptors of Sertoli cells and promote the mitosis and meiosis steps of spermatogenesis, which are also supported by FSH. They also cause many other physical and psychological effects (Table 3.5.6-1).

An additional control factor is prolactin, a product of the anterior pituitary lobe, which potentiates the LH action on the Sertoli cells and supports the testosterone effects. Its secretion is promoted by endorphins and inhibited by GAP (another part of the LHRH prohormone).

Early in embryogenesis, the expression of the *sry* gene on the Y chromosome results in synthesis of the testis determining factor which, in turn, acts as a switch for sexual differentiation. It activates the transcription of the Müllerian inhibitory substance (MIS), which suppresses the development of female genitalia. Then testosterone and 5α-dihydrotestosterone induce the differentiation of the Wolffian ducts into the male reproductive organs. Later on, these

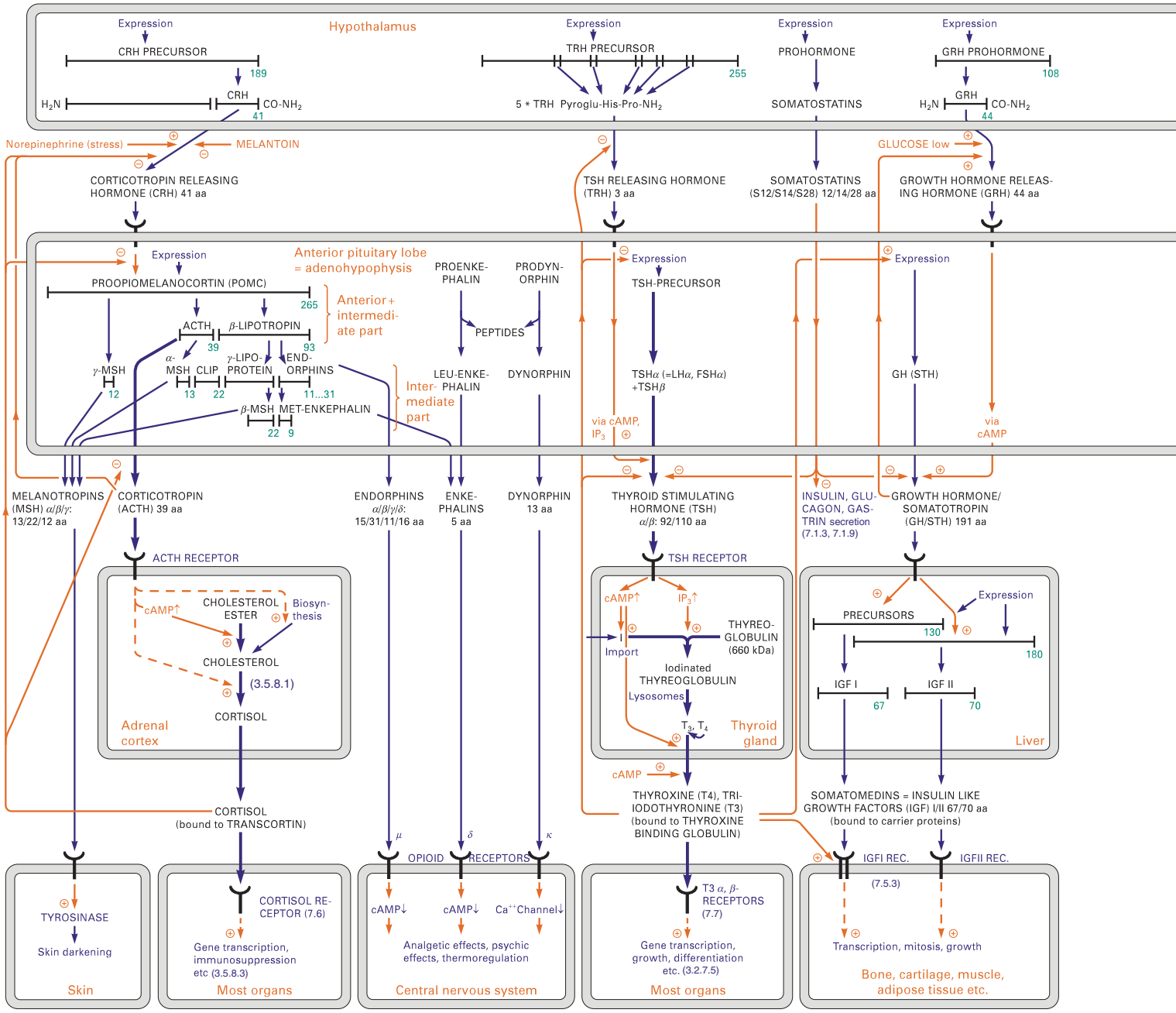


Figure 7.1-4. Hypothalamus-Anterior Pituitary Controlled Hormones

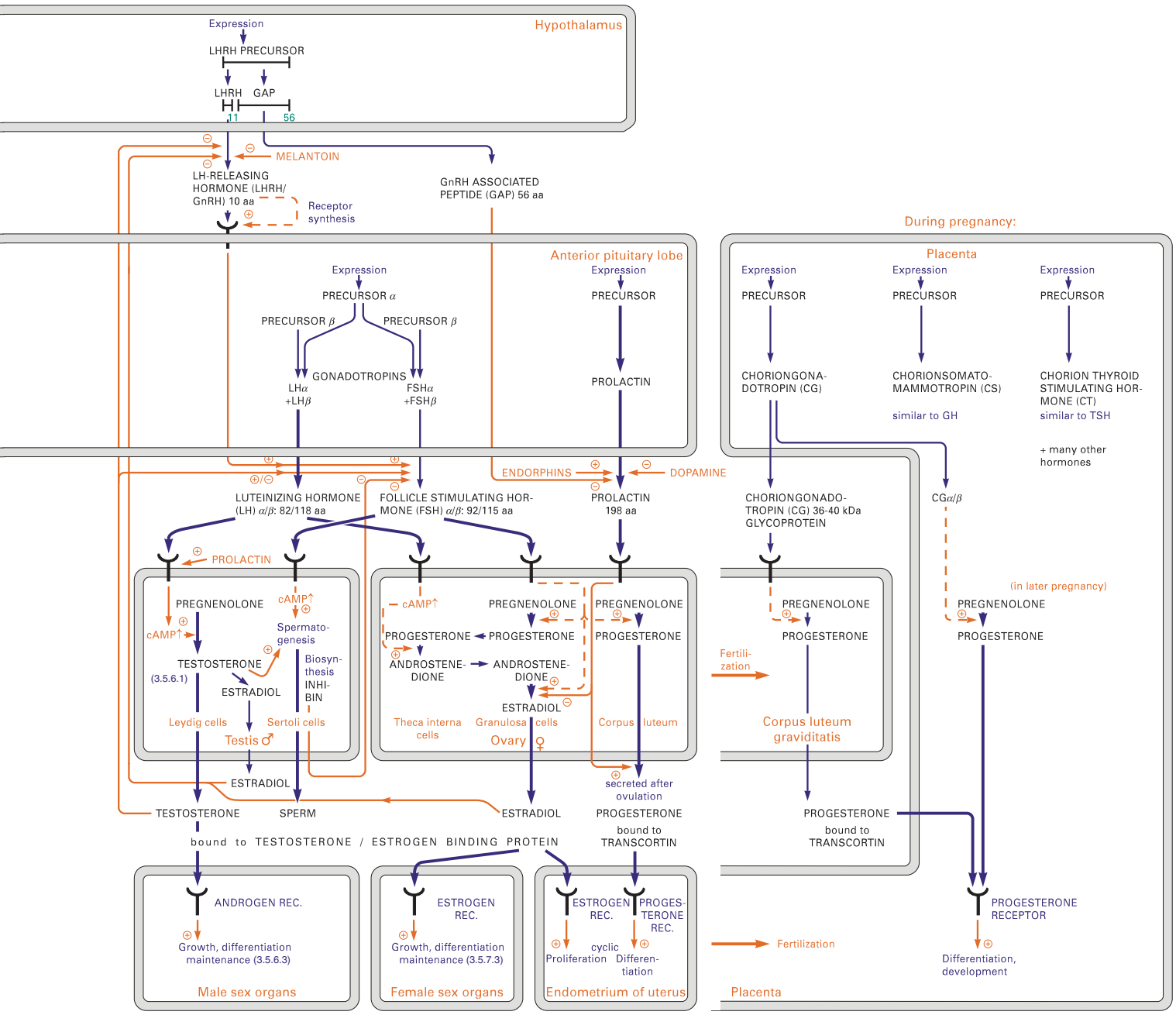
hormones direct the development of many other male sex characteristics (Table 3.5.6-1).

Hypothalamo-pituitary-ovary/uterus axis: The control of female sex hormone production depends also on the pulsed release of GnRH, LH and FSH (see above). Estrogens (estradiol, estrone) are synthesized via androgen intermediates. This takes place in the granulosa and the theca interna cells of the follicle, which are under control of FSH and LH, respectively (Figs. 7.1-4 and 7.1-5). Formulas and details of the metabolism are shown in Figures 3.5.6-1, 3.5.7-1 and Table 3.5.7-1.

Puberty marks the beginning of the menstrual cycle with a characteristic pattern of the plasma hormone fluctuations (Fig. 7.1-6), which are additionally modulated by the diurnal pulsating secretions. This periodicity is regulated by a complicated interplay of forward activations and feedback inhibitions within the hormone cascade.

Before ovulation, the granulosa cells of a follicle express first receptors for FSH and afterwards intracellular receptors for estradiol, progesterone, testosterone and cortisol (7.6). The binding of these hormones results in enzyme induction (e.g., of aromatase) and in cell proliferation. Other enzymes of estrogen biosynthesis are induced by LH binding to its receptor on the theca interna cells. The estrogen production increases and exerts a feedback inhibition on the FSH secretion, which blocks the maturation of additional follicles. The sharp peak of LH concentration at the midpoint of the cycle is associated with the formation of hydrolytic enzymes, which results in ovulation. Thereafter granulosa and theca interna cells form the corpus luteum, which is sustained by LH and prolactin hormones. The corpus luteum secretes mainly progesterone.

The estrogens also cause a proliferation of the endometrium and other responses of the uterus in the first phase of the menstrual cycle. After ovulation, progesterone modifies the endometrium. Without fertilization, the decrease of hormone production at the transition from the proliferative to the secretory state causes luteolysis and sloughing of the uterine lining (menstruation).



Generally, estrogens are responsible for formation and upkeep of female sex characteristics (Table 3.5.7-1). During the embryonic state, estrogens control the development of the female gonads. FSH and LH plasma levels increase slowly until puberty, when they commence the characteristic concentration changes of the menstrual cycle. After menopause, the decreasing estrogen production no longer exerts a feedback inhibition, FSH and LH levels rise without periodicity.

7.1.6 Placental Hormones

After fertilization, the corpus luteum increases progesterone output, but later on this hormone is provided by the placenta (3.5.5.2).

Estrogens prepare the uterus epithelium by induction of growth factors (e.g., EGF, TGF-α, etc.), cytokines and adhesion molecules (lectin-like compounds, integrins, etc., Table 8.4-1). The early embryo (blastocyte state) also produces various cytokines, adhesion molecules etc. Their interaction with receptors results in implantation of the fertilized ovum.

The developing placenta produces a series of hormones, which are analogous to hypothalamus or pituitary hormones:

Choriongonadotropin (CG) ⟷ LH
Chorionsomato(mammo)tropin (CS) ⟷ STH
Chorionthyrotropin (CT) ⟷ TSH.
Additionally it supplies LH-RH, TRH, somatostatin, etc.

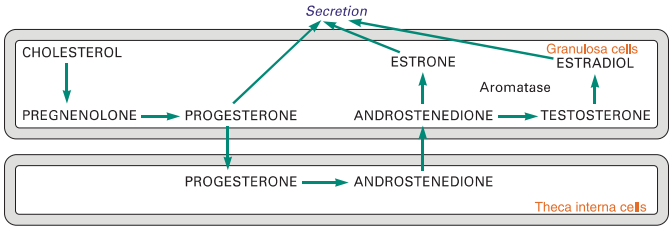


Figure 7.1-5. Synthesis of Estrogens in the Ovary

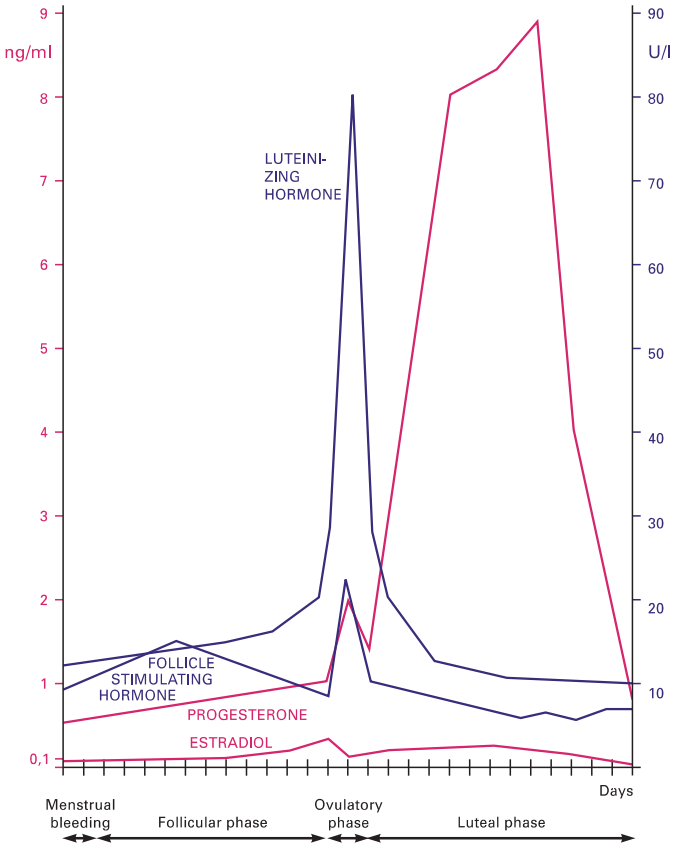


Figure 7.1-6. Hormone Concentration in Plasma during the Menstrual Cycle

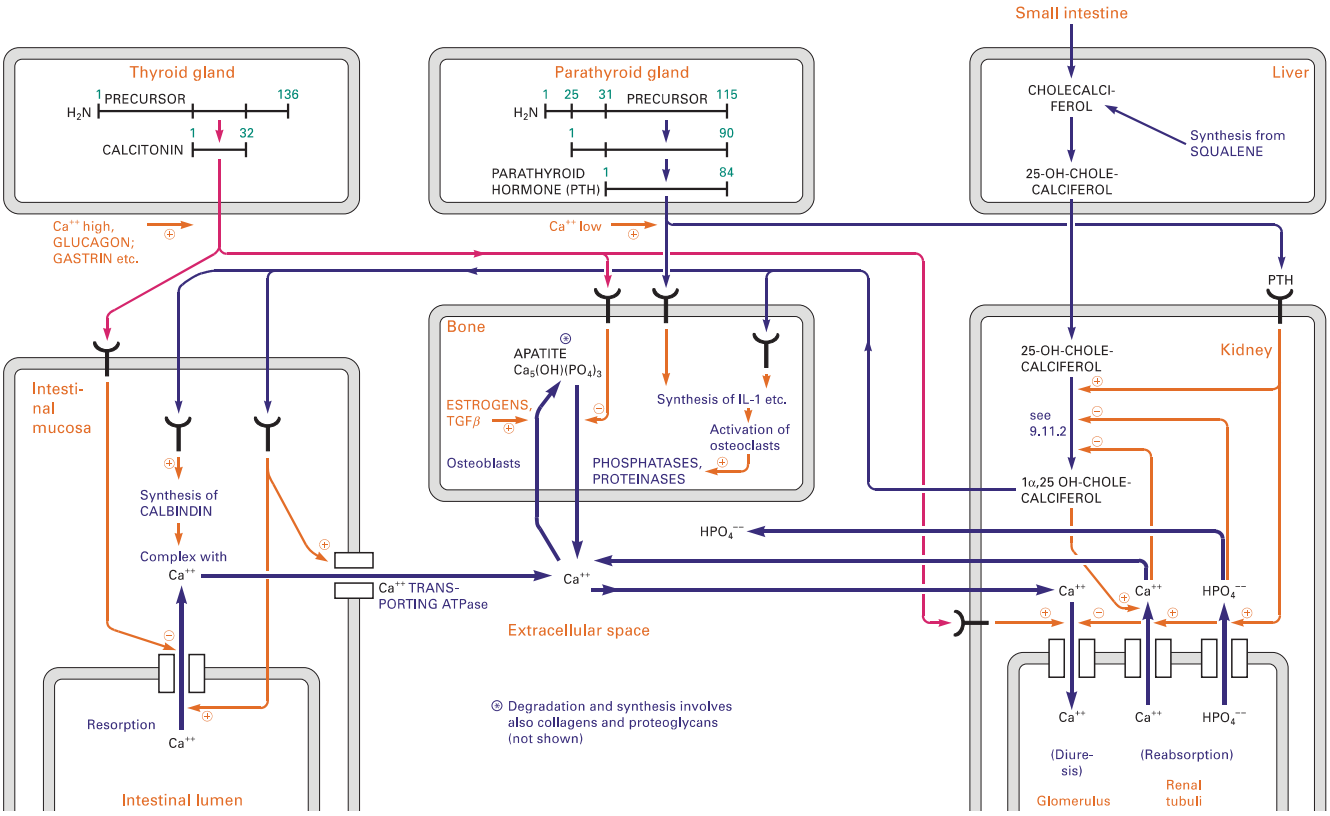


Figure 7.1-7. Regulation of the Ca^{++} Concentration in the Extracellular Space
The antagonistically acting calcitonin reactions are shown in red.

In the earliest phase of pregnancy, the blastocyte provides CG (human form: hCG). It prevents the regression of the corpus luteum and its progesterone production, until the placenta can take over. The hCG determination in plasma or urine can be used as a pregnancy test.

7.1.7 Hormones Regulating the Extracellular Ca^{++}, Mg^{++} and Phosphate Concentrations (Fig. 7.1-7)

Calcium is involved in many important biochemical functions, e.g., muscle contraction (7.4.5), nerve conduction (7.2) and blood coagulation (9). The intracellular mechanisms are described in the respective sections. The extracellular calcium concentration likewise has to be strictly controlled; even more so, since only a small portion of the total body calcium is in the liquid phase (99% is fixed in bones, teeth, etc., but can be mobilized) and therefore a large concentration gradient across membranes has to be maintained (7.4.4).

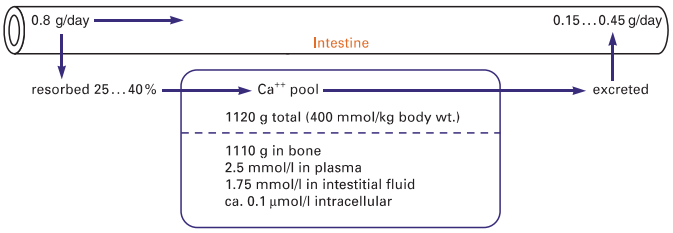


Figure 7.1-8. Flow Sheet of Calcium in an Adult Human (70 kg)

The Ca^{++} flow in an adult human is shown in Figure 7.1-8. The Ca^{++} metabolism is closely coupled to the phosphate metabolism. Solubility of the salts plays a major role.

Resorption, mobilization from the skeleton and excretion of Ca^{++} are regulated by the parathyroid hormone (PTH) and 1α,25-dihydroxycholecalciferol (which increases the concentration in the extracellular space) and by calcitonin (which decreases the concentration). The regulatory mechanisms and the biosynthesis of PTH and calcitonin are shown in Figure 7.1-7, while the biosynthesis of cholecalciferol and other compounds of the vitamin D group is dealt with in 3.7.11.1.

Magnesium is a cofactor of many reactions involving phosphate transfer, e.g., kinases (as Mg-ATP complex), phosphatases, nucleic acid synthesis etc. Little is known about the intestinal resorption mechanism and the concentration control in body fluids. During transfer through ion channels it acts as an antagonist to Ca^{++}.

7.1.8 Hormones Regulating the Na^+ Concentration and the Water Balance (Fig. 7.1-9)

While sodium is the most frequent cation in the extracellular space, potassium is prevalent in the intracellular space (Table 7.1-3). These disequilibria are produced by the ubiquitous Na^+/K^+-exchanging ATPase (Fig. 6.1-1). They are essential for cellular function (membrane potential, 7.2.1; Na^+ driven symport and antiport channels, Table 6.1-3). Na^+ also plays an important role in the osmoregulation of the cells (see below).

Table 7.1-3. Concentration of Na^+ and K^+ Ions

		Na^+	K^+
Plasma	mmol/l	135 … 145	3.5 … 5.5 (normal 4.0)
Interstitial fluid	mmol/l	144	4.0
Intracellular	mmol/l	10	ca. 150 (tissue dependent)

Na^+ turnover: The intestinal Na^+ resorption proceeds via a co-transport mechanism with glucose. The excretion mainly takes place in the urine and is the balance between glomerular filtration and tubular reabsorption. The reabsorption is an active transport process employing

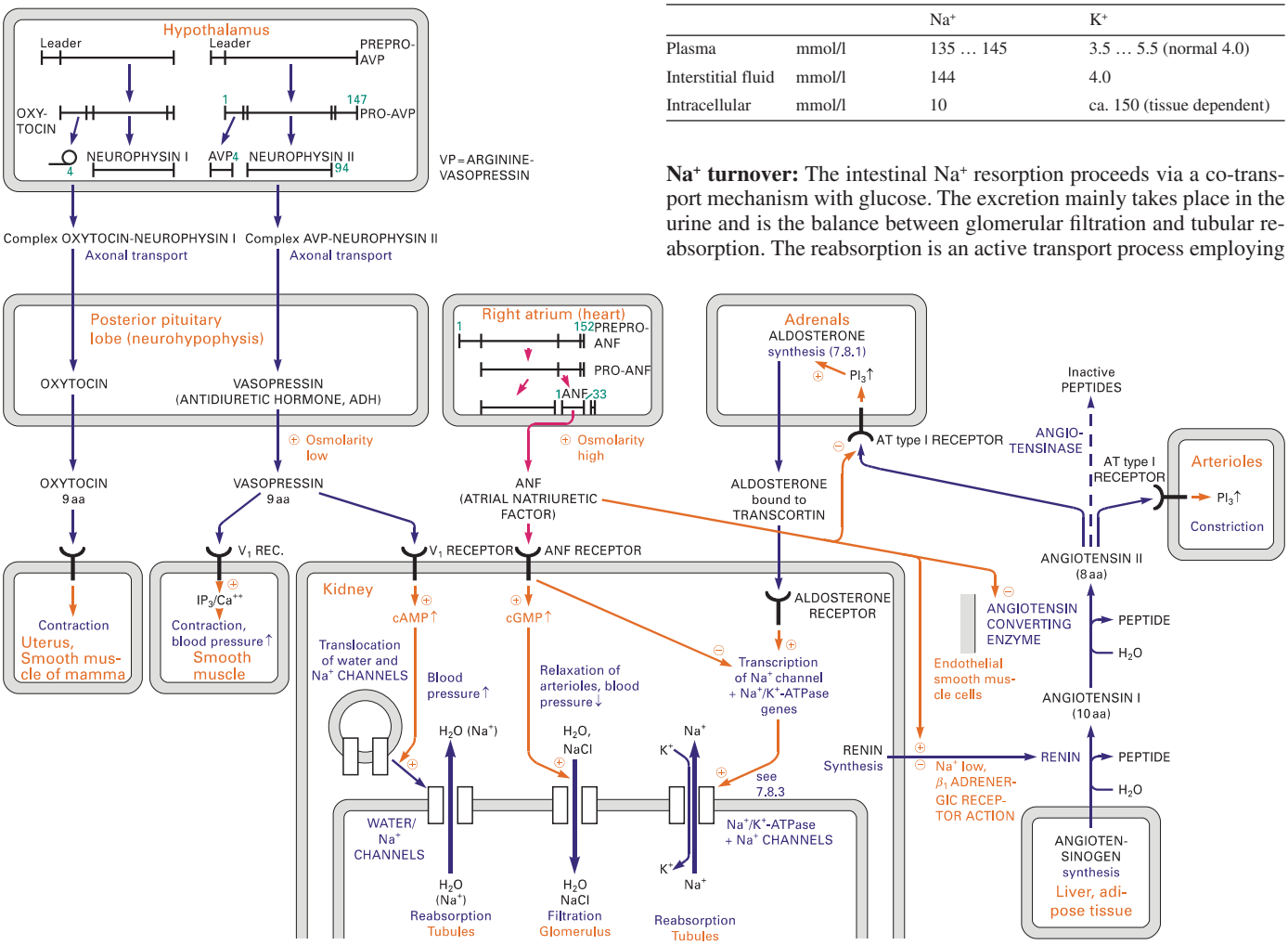


Figure 7.1-9. Regulation of the Na^+ Concentration and the Extracellular Water Volume
The antagonistically acting ANF reactions are shown in red.

both the Na^+/K^+-ATPase and Na^+ channels. The expression of both proteins is hormone controlled.

Water turnover: Water resorption from the intestine follows the osmotic gradient established by the Na^+ resorption. The major excretion occurs via the kidney, next most important are lung and skin. Renal reabsorption is facilitated by hormone-controlled translocation of water channels from internal vesicles to the membrane of epithelial cells of the collecting tubules.

The regulation of the extracellular volume and the interrelated control of blood pressure is performed by hormones: vasopressin and the renin-angiotensin system increase the blood pressure, the atrial natriuretic factor (ANF) antagonizes it. They respond either to the osmotic pressure or to the blood pressure as signals for their release. For biosynthesis and mechanism of action see Figure 7.1-9. Aldosterone (3.5.8.1) promotes Na^+ and water retention.

Oxytocin (Ocytocin, 9 aa), chemically closely related to vasopressin, contracts the uterus during labor and later on the breast smooth muscle (milk secretion).

7.1.9 Hormones of the Gastrointestinal Tract

Secretion of digestive fluids, secretion and reabsorption of water and electrolytes and intestinal movements are regulated by hormones. Some of them are neuronally controlled.

Gastric hormones (Fig. 7.1-10): Neuronal impulses (acetylcholine, or from the parasympathetic neurons: gastrin-releasing peptide), as well as signals from the stomach (high pH value, stretching) promote the secretion of gastrin, which causes release of HCl into the stomach. The same neuronal impulses inhibit the secretion of somatostatin, which blocks the HCl release.

Gastrin, cholinergic impulses or high acidity of the gastric milieu cause secretion of the digestive prohormone pepsinogen and of the epithelium-protective compound mucin. Pepsinogen cleaves itself autocatalytically at low pH into the active pepsin.

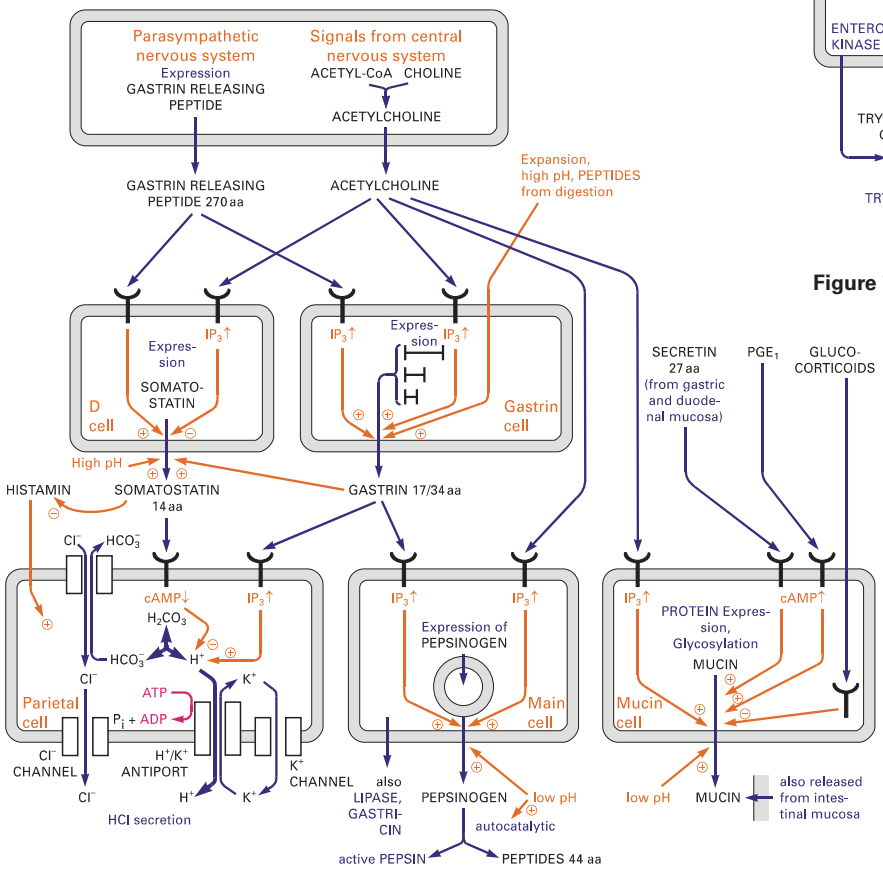


Figure 7.1-10. Regulation of Gastric Secretion

Pancreatic hormones (Fig. 7.1-11): The secretion of pancreatic enzymes and proenzymes is also under multiple control. Neuronal impulses are transmitted by acetylcholine. Hormonal signals are initiated by contents of the duodenum (peptides, amino acids, fatty acids), which cause release of cholecystokinin (CCK, also called pancreozymin, PZ) into the intestine. Other hormones are secretin and the related vasoactive intestinal peptide (VIP). They stimulate both the secretion of pancreatic (pro-)enzymes and of water. The endocrine hormones of the pancreas are described above (7.1.3).

Literature:

Barnard, E.A. Trends in Biochem. Sci. 1996;17:305–309.

Baxter, J.D. *et al.* Rec. Progr. Horm. Res. 1991;47:211–258.

Bezanilla F. Neuron 2008;60:456–468.

Brent, G.A. New Engl. J. Med. 1994;331:847–853.

Brown, E. Physiol. Rev. 1991;71:371–411.

Cooke, B.A., King, R.J.B., van der Molen, H.J. (Eds.) *Hormones and Their Activation. New Comprehensive Biochemistry* Vol. 18A, 18B. Elsevier, 1988.

MacDonald P.E., Joseph J.W., Rorsman P. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 2005;360:2211–2225.

Sekido R., Lovell-Badge R. Trends Genet. 2009;25:19–29.

Werner, M.H. *et al.* Trends in Biochem. Sci. 1996;21:302–308.

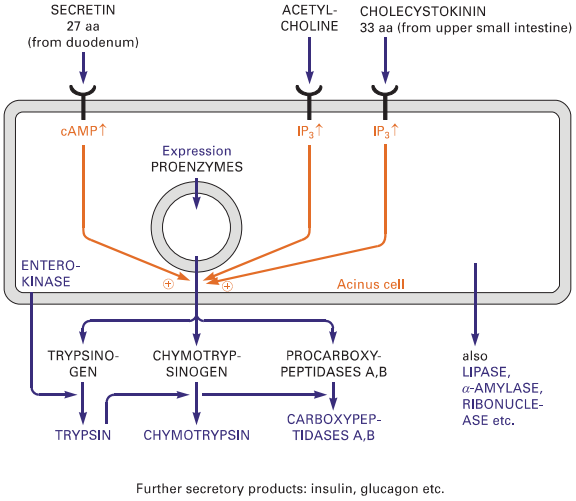


Figure 7.1-11. Regulation of Pancreatic Enzyme Secretion

7.2 Nerve Conduction and Synaptic Transmission

Nerve cells (neurons) receive, process, and transmit signals. While transport (= conduction) and processing within a single cell proceed electrically, the connection with other cells for transmission of signals at synapses may be electrical or chemical in nature.

7.2.1 Membrane Potential

Neuronal membranes contain channels which allow passage of ions at different rates. Due to this selectivity, the ion concentrations on both sides of the cellular membrane are different, causing an electric potential difference across the membrane.

The equilibrium potential $\Delta\Psi$ of a cell is calculated according to the Goldman equation (compare 1.5.3). This form is valid for singly charged ions. P = permeability coefficient, ca = cations, an = anions:

$$\Delta\Psi\ [mV] = \frac{1000 * R * T}{F} * 2.303 * \log \frac{\sum [P_{ca} * c_{ca\text{-}outside}] + \sum [P_{an} * c_{an\text{-}inside}]}{\sum [P_{ca} * c_{ca\text{-}inside}] + \sum [P_{an} * c_{an\text{-}outside}]}$$

R = gas constant (8.315 Joule * K^{-1} * mol^{-1}); T= temperature in Kelvin; F = Faraday's constant (96,485 Coulomb* mol^{-1}).

The equilibrium potential of mammalian cells is ca. –83 mV. It is composed mainly of the individual equilibrium potentials of K^+ = –91 mV and Na^+ = +55 mV. The Na^+/K^+-ATPase (6.1.4), however, moves Na^+ out of the cell and K^+ into it, causing a resting potential of an unstimulated cell of –65 mV. (The values for these parameters differ somewhat in the literature.) Opening of an ion channel allows the particular ion to approach its equilibrium state and therefore changes the resting potential towards the equilibrium potential of this particular ion (action potential).

Due to the capacitor effect of the membrane, only ca. $^1/_{2000}$% of the Na^+ ions present must diffuse to depolarize a cell body from –66 to +45 mV.

7.2.2 Conduction of the Action Potential along the Axon (Fig. 7.2-1)

An endogenic or exogenic stimulant shifts the membrane potential away from its resting state ①. If this change is beyond a threshold value, it is recognized by voltage sensitive structures in neighboring channel proteins and leads to a conformational change, causing the opening of closed Na^+ and K^+ channels ②. At first, the predominant influx of Na^+ shifts the membrane potential to positive values ③ (depolarization). After the closing of the Na^+-channel, the K^+ outflux through the still open K^+ channel reverses the potential, even beyond the resting potential ④ (hyperpolarization). As soon as the K^+-channel closes, the potential returns to the value of the resting potential ⑤. In this way, the action potential travels along the axon at constant speed to the synapse (ca. 0.5 m/sec in unmyelinated cells, up to 100 m/sec in myelinated cells, also depending on their diameter).

The general structures of the voltage-gated Na^+, Ca^{++} and K^+ channels are similar (Figs. 7.2-2 and 7.2-3, for TC classification see 6.1.2.1). The actual channel consists of 4 circularly arranged domains (K^+ channel: 4 separate protein chains) with 6 transmembrane segments each (S_1 … S_6). The P domains (hairpin loops between S_5 and S_6, with different amino acid sequences for each channel type) are part of the channel lining and likely determine the ion specificity of the channel. The segments S_4 with a regular arrangement of positively charged amino acids may act as voltage sensors, causing the conformation changes. Additional subunits regulate or locate the channel protein(s). Closing of the K^+ channel (K_A type) is apparently mediated by the contact of the N terminus of the protein chain with receptor structures at the intracellular mouth of the channel. The Na^+ channel may be closed when the short intracellular loop between S_6 (domain III) and S_1 (domain IV) covers the channel mouth like a lid.

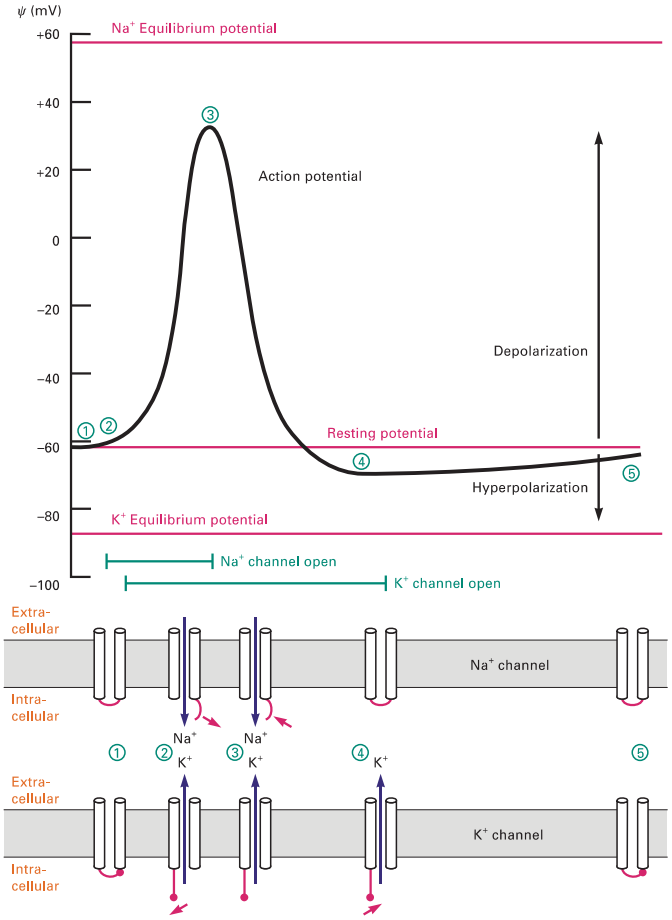


Figure 7.2-1. Generation of the Action Potential

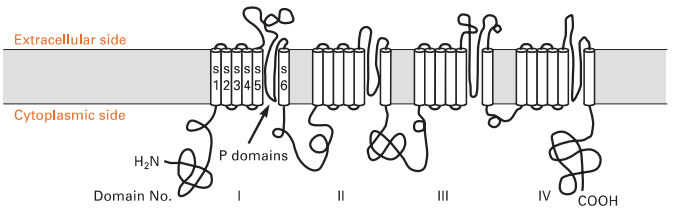


Figure 7.2-2. Protein Segments of Voltage Gated Na^+ and Ca^{++} Ion Channels (α subunit, additionally there are several other subunits). In the K^+ Channel there are four separate α protein chains

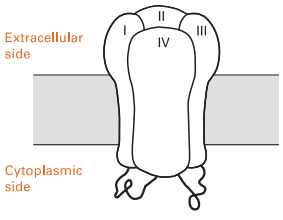


Figure 7.2-3. Structure of Voltage Gated Ion Channels (Na^+, Ca^{++}, possibly also K^+)

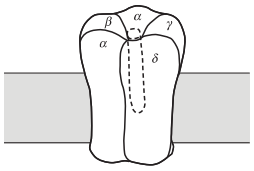


Figure 7.2-4. Structure of Direct Transmitter Gated Ion Channels (Type I, nicotinic acetylcholine receptor)

7.2.3 Transmitter Gated Signalling at the Synapse (Table 7.2-1)

Presynaptic reactions: Presynaptic vesicles are synthesized in the ER/Golgi system (4.4.1 and 4.4.2) and are transported by axonal transport mechanisms (7.2.6) to their location near the presynaptic membrane, where they stay attached to the cytoskeleton. During maturation, they

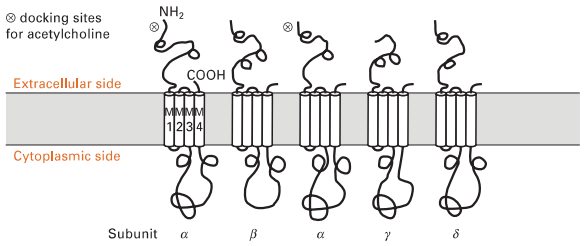


Figure 7.2-5. Protein Segments of Direct Transmitter Gated Ion Channels at the Postsynaptic Membrane (Nicotinic acetylcholine receptor)

pick up the respective neurotransmitters by specific, ATP-dependent uptake mechanisms.

When the action potential reaches the nerve terminal, it triggers the process of synaptic transmission. Opening of voltage-sensitive Ca^{++}-channels leads to Ca^{++}-influx into the active zone of the presynaptic membrane (Fig. 7.2-7). The increased Ca^{++} concentration causes phosphorylation of Ca^{++} sensors (e.g., synapsin) by protein kinases, which in turn effects vesicle separation from the cytoskeleton, fusion of the vesicle with the presynaptic membrane and neurotransmitter release into the synaptic cleft (for mechanism see 4.5.2). The neurotransmitters diffuse to specific postsynaptic receptors (7.2.5).

Postsynaptic reactions: Binding of the neurotransmitters to the receptors (Table 7.2-3) changes the excitability of the postsynaptic membrane. This leads to one of the following effects:

- Direct transmitter gating (receptors type I, ionotropic receptors, ligand-gated ion channels): Ion conducting channels open in response to neurotransmitter binding (For TC classification see 6.1.2.2).

  The ion channels (Figs. 7.2-4 and 7.2-5) consist of five separated subunits with four transmembrane segments each (e.g., nicotonic acetylcholine receptor: subunit structure $\alpha_2\beta\gamma\delta$, both α subunits bind acetylcholine). The five subunits are arranged to form a single membrane-spanning pore. The M_2 segments of the subunits are thought to line the inside of this pore. The amino acids contained therein (acidic or basic) determine the selectivity of the channel. The openings of the pore facing the outside and inside of the cell, respectively, are wide enough to let ions pass, but kinks in the three-dimensional structure of each subunit considerably narrow the inside of the pore making it impassable for ions. Binding of neurotransmitter molecules to the outside of the ion channel induces conformational changes in the pore's subunits. These result in rotational movements involving the area of the kinks and thereby widen the inside of the channel enough so that ions can pass it.

  The ionotropic glutamate receptors (subdivided into AMPA, NMDA, and Kainate receptors) also form multimeric assemblies of four (sometimes five) subunits as outlined above. However, their general structure differs from that of other ionotropic receptors in that the channel-lining M2 segment of each subunit forms a re-entrant loop. This completely changes the topology of all subsequent parts such that the long loop between the third and fourth transmembrane domain faces outwards, not inwards and forms part of the neurotransmitter binding site, while the carboxy terminus ends up inside the cell and mediates protein-protein interactions with signaling components.

- Indirect transmitter gating (receptors type II, metabotropic receptors): Heterotrimeric G proteins as second messenger systems are activated. Details are described in 7.4.

  The receptors (Fig. 7.2-6) usually contain seven transmembrane helices (e.g., muscarinic acetylcholine receptor, β-adrenergic receptor). Their activity can be modulated by phosphorylation.

Depending on the type of transmitters and receptors, these signals can be either excitatory (at the muscular endplate and in the central nervous system, CNS, causing depolarization) or inhibitory (only in the CNS, causing hyperpolarization).

The transmitter compounds finally return either by high- or low-affinity transport processes to the presynaptic vesicles (e.g., catecholamines) or are enzymatically inactivated (e.g., acetylcholine, 3.4.3.5); their metabolites are retrieved for resynthesis.

As an example of direct transmitter gating, synaptic excitation at the nerve-muscle synapse is mediated by acetylcholine, which is released from presynaptic vesicles out of synaptic buttons at the muscular endplate (Fig. 7.2-8). It binds to the postsynaptic nicotinic acetylcholine receptor, which causes the opening of Na^+ and K^+ channels and thus changes the end-plate potential (EPP). This is converted into an action potential (7.2.2) by voltage gated Na^+-channels located in the membranes of the muscle cell. The action potential propagates to the Ca^{++} stores. Release of Ca^{++} via voltage gated channels leads to muscle contraction (7.4.5).

Integration of signals: If signals from a number of afferent nerve fibers arrive at a neuron of the CNS (usually causing differences of less than 1 mV each) they are correspondingly added or subtracted, also taking into account time and space effects (neuronal integration). Only if the accumulated input causes membrane depolarization in excess of the threshold value of −55 mV (a change of +10 mV from the resting potential), an action potential is initiated. This is in contrast to the nerve-muscular synapse, where simultaneous action of 300 … 400 release sites of the single nerve and many postsynaptic receptors cause changes of ca. +70 mV, thus always surpassing the threshold value (except in pathological conditions).

Table 7.2-1. Types of Synaptic Transmission (For a List of Ion Channels, see Tables 6.1-1 and 6.1-2)

Type	Characteristics	Mechanism	Location
Direct gated transmission	Chemical: cleft 30 … 50 nm, presynaptic active zones, postsynaptic receptors, unidirectional transmission. Chemical neurotransmitters: acetylcholine, GABA, glycine, glutamate and aspartate.	The released neurotransmitters act directly on ionophoric postsynaptic receptors. Synaptic delay 1–5 msec. In case of excitatory effects, Na^+ and K^+ channels are opened. The depolarization, in turn, opens neighboring voltage gated channels, producing an endplate-potential in muscles and excitatory effects in the CNS. In case of inhibitory effects, voltage-gated Cl^- channels are opened, causing hyperpolarization.	Nerve-muscle synapse (nicotinic acetylcholine receptor). Aminergic, cholinergic and amino acid dependent neurons in the CNS.
Indirect gated transmission	Chemical synapses: cleft 30 … 50 nm, presynaptic active zones, postsynaptic receptors. Chemical neurotransmitters: norepinephrine, dopamine, serotonin, acetylcholine, neuropeptides.	The released neurotransmitters bind to specific postsynaptic receptors. This activates ion channels indirectly via heterotrimeric G-proteins, engaging second messengers. Synaptic delay usually >10 msec.	α- and β-adrenergic receptors, serotoninergic, dopaminergic, peptidergic and muscarinic acetylcholine receptors in the CNS (e.g., in the cerebral cortex).
Voltage gated transmission	Electrical synapses: cleft 3.5 nm, gap junction channels. Modulation by phosphorylation.	The action potential causes depolarization of the presynaptic cell. The ionic current (Na^+ and Ca^{++} influx, K^+ efflux) causes postsynaptic electrical changes. No synaptic delay, rapid bidirectional transmission.	Intracellular conduction of the action potential in neurons requiring rapid and synchronous synaptic transmission (neurons, heart, also in invertebrates).

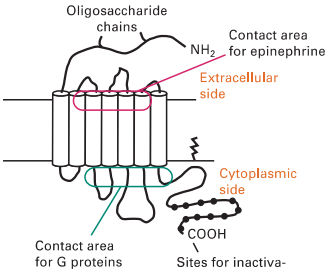


Figure 7.2-6. Protein Segments of an Receptor Type II (for Indirect Transmitter Gating) (β-Adrenergic receptor, circularly arranged)

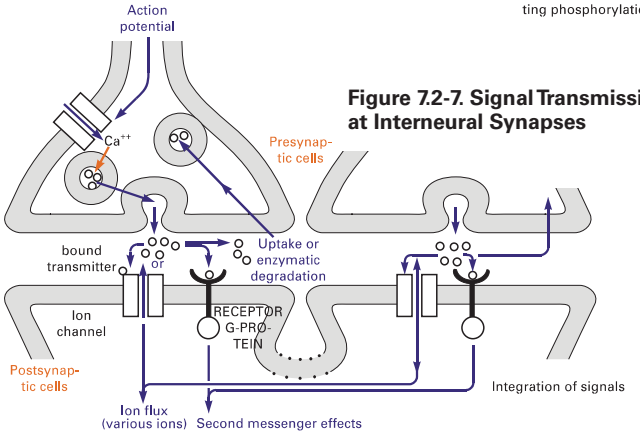


Figure 7.2-7. Signal Transmission at Interneural Synapses

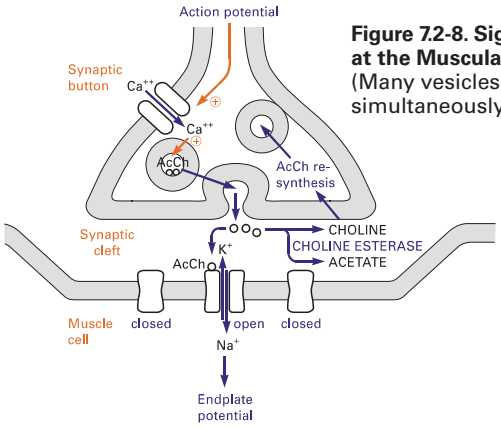


Figure 7.2-8. Signal Transmission at the Muscular Endplate (Many vesicles discharge simultaneously)

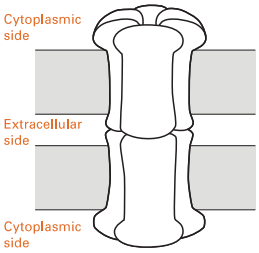
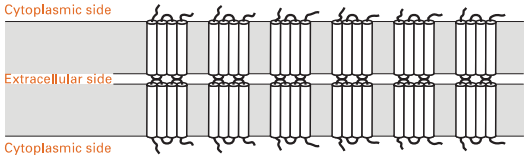


Figure 7.2-9. Structure and Protein Segments of Gap Junction Ion Channels



7.2.4 Voltage Gated Signalling at the Synapse (Fig. 7.2-9)

Voltage gated transmission is performed in the nervous system and the heart and especially in lower organisms (invertebrates) for fast and synchronous dissemination of the action potential.

Gap-junction channels appear as a pair of pores directly connecting the pre- and postsynaptic membranes. Both channels consist of connexins (six separated subunits with four transmembrane segments each); the corresponding ones in both membranes are in close connection.

The characteristics of neuronal transmission and of neurotransmitters are compiled in Tables 7.2-1, 7.2-2 and 7.2-3.

7.2.5 Postsynaptic Receptors (Table 7.2-3)

The receptors are divided into many subtypes. Agonists and antagonists perform transmitter-specific actions with different affinities to the receptors. This allows discrimination of the receptors in pharmacological and toxicological studies and development of specific pharmaceuticals. A selection of them, including synthetic compounds of high specificity, is listed below.

7.2.6 Axonal Transport (Table 7.2-4)

Maintenance of neuronal function requires continuous supply of metabolites from the cell body to the nerve terminal and vice versa. Axonal transport occurs along tracks formed by microtubules and by the smooth endoplasmic reticulum. Fast axonal transport is rapidly established by the microtubuli associated ATPases dynein (MAP1C) and kinesin, which act as motor molecules. Alkaloids, which disrupt microtubuli (e.g., vinblastin and colchicine) block this process immediately. In contrast to this, slow axonal transport involves physical translocation of the cytoskeleton. It provides mainly cytoskeletal elements for the nerve terminal.

Literature:

Barnard, E.A. Trends in Pharmacol. Science 1996;17:305–309.

Bajjalieh, S.M., Scheller, R.H. J. Biol. Chem. 1995;270:1971–1974.

Bennett, M.K., Scheller, R.H. Ann. Rev. of Biochem. 1994;63: 63–100.

Catteral, W.A. Ann. Rev. of Biochem. 1995;64:493–531.

Flavell S.W., Greenberg M.E. Ann. Rev. Neurosci. 2008;31:563–590.

Greer P.L., Greenberg M.E. Neuron 2008;59:846–860.

Hucho, F. (Ed.) *Neurotransmitter Receptors. New Comprehensive Biochemistry.* Vol. 24. Elsevier, 1993.

Kandel, E.R. in: *Principles of Neural Science,* 3rd Ed. Kandel, Schwartz and Jessel (eds.), Elsevier, 1991; p. 194–213.

Koester, J. *Ibid.,* p. 104–120.

Kennedy, M.B. Science 2000;290:750–754.

Lin R.C., Scheller R.H. Ann. Rev. Cell. Dev. Biol. 2000;16:19–49.

Neher E., Sakaba T. Neuron 2008;59:861–872.

Strange, P.G. *Brain Chemistry and Brain Disorders,* 1st Ed. Oxford University Press, 1992.

(Various authors). Pharmacol. Reviews 1994;45:121–224.

Watson, S., Girdlestone, D. *Receptors and Ion Channels Supplement.* Trends in Pharmacol. Science 1996;17.

7.3 Principles of Intracellular Communication

This section deals with the intracellular communication of higher multicellular organisms, mostly vertebrates.

The chemical information arriving at the cell surface by hormones or cytokines (7.1) or by discharge of neurotransmitter compounds (7.2) is recognized by specific transmembrane receptor proteins. Binding of these agonists causes a structural change in the intracellular portion of the receptors, which starts enzymatic reactions leading to the activation or release of second messengers. These transmit the information to the final site of action, either directly or via a cascade of further reactions. In the latter case, the additional steps allow modification and integration of various signals.

As an exception, steroid, thyroid, and retinoic acid receptors bind their agonists inside of the cell and act directly on gene transcription (see below).

Receptor antagonists compete with agonists for binding to the same site on the receptor, but they do not initiate a signaling cascade like agonists.

Table 7.2-2. Selection of Neurotransmitters and Receptor Subtypes. See also Table 6.1-2
(Some neurotransmitter compounds also act as hormones in the bloodstream. For G-protein dependent reaction mechanisms, see 7.4. Prostanoid and leukotriene receptors are dealt with in 7.4.8)

Name	Receptors (Receptor Type)	Location/Effects of the Neurotransmitters
Acetylcholine (3.4.3.5)	nicotinic receptor (I)	Binding to ionophoric receptors permeable to Na^+ and K^+ → → nerve-muscle synapse ('muscle type receptor'): activation of endplate potential → CNS ('neuronal type receptor'): excitatory and inhibitory actions
	muscarinic M_1 (neuronal), M_3 (glandular), M_5 receptors (II)	Binding to receptor linked with $G_{q\alpha}$-protein → activation of phospholipase C → formation of inositol-P_3 and diacylglycerol → → parasympathetic system: mediation of pre- and postganglionic synaptic transmission → sympathetic system: mediation of preganglionic synaptic transmission → CNS: predominantly excitatory transmitter, e.g., activation and maintenance of cognitive function in brain
	muscarinic M_2 (cardiac), M_4 receptors (II)	Binding to receptors linked with $G_{i2\alpha}$-protein[2] → inhibition of adenylate cyclase → decrease of cAMP and opening of Ca^{++} channels → heart: negative inotropic and chronotropic effects
Biogenic amines		
Norepinephrine, epinephrine (3.2.7.4)	adrenergic β_1, β_2, β_3 receptors (II)	Binding to β receptors linked with $G_{s\alpha}$-protein → activation of adenylate cyclase → increase of cAMP → → target organs of the sympathetic system: mediation of postganglionic action potential, metabolic effects (e.g., positive inotropic and chronotropic effects on the heart, relaxation of smooth muscles, vasodilatation) → CNS: excitatory actions, in brain: autonomic control of cardiovascular function, activation of cortical "arousal" behavior
	adrenergic α_1 receptor (II)	Binding to α receptors linked with $G_{q\alpha}$-protein → activation of phospholipase C → formation of inositol-P_3 → release of Ca^{++} from the endoplasmic reticulum → → target organs of the sympathetic system: metabolic effects
	adrenergic α_2 receptor (II)	Binding to receptor linked with $G_{o\alpha}$-protein[2] → inhibition of adenylate cyclase → decrease of cAMP → → target organs of the sympathetic system: mediation of postganglionic action potential antagonistic to β receptors, metabolic effects (e.g., vasoconstriction) → CNS: central regulation of blood pressure
Dopamine (3.2.7.3)	D_1, D_5 receptors (II)	CNS: binding to receptors linked with $G_{s\alpha}$-protein → activation of adenylate cyclase → increase of cAMP → excitatory cortical actions
	D_{2a} receptor (II)	CNS: binding to receptor linked with $G_{i\alpha}$-protein[2] → inhibition of adenylate cyclase → decrease of cAMP → inhibitory autoregulation of neuron firing rate and dopamine release
	D_{2b} receptor (II)	CNS: binding to receptor linked with $G_{q\alpha}$-protein → activation of phospholipase C → formation of inositol-P_3 and diacylglycerol → excitatory cortical actions
	D_3, D_4 receptors (II)	CNS: binding to receptors linked with $G_{q\alpha}$-protein → activation of phospholipase C → formation of inositol-P_3 and diacylglycerol → excitatory mesolimbic actions (weakly expressed)
Serotonin (5-hydroxytryptamine, 5-HT, 3.2.7.3)	5-HT_1 receptor (II)	CNS: binding to receptor linked with $G_{i\alpha}$-protein[2] → inhibition of adenylate cyclase → decrease of cAMP → inhibitory actions, e.g., sleep induction
	5-HT_2 receptor (II)	CNS: binding to receptor linked with $G_{q\alpha}$-protein → activation of phospholipase C → formation of inositol-P_3 and diacylglycerol → excitatory actions involved in emotional behavior
	5-HT_3 receptor (I)	CNS: binding to ionophoric receptor located on channel permeable to Na^+ and K^+ → different excitatory actions in emotional behavior and memory processing
Amino acid transmitters		
γ-Aminobutyrate (GABA, 3.2.2.2, major inhibitory transmitter in the CNS)	$GABA_A$ receptor (I)	CNS: binding to ionophoric receptor (similar to nicotinic acetylcholine receptor) → influx of Cl^- → hyperpolarization → inhibitory actions (e.g., of sensory transmission in the spinal cord), serotonin release, indirect dopamine release (Benzodiazepine and barbiturates increase Cl^- influx).
	$GABA_B$ receptor (II)	CNS: binding to receptor linked with $G_{i\alpha}$-protein[2] → inhibition of adenylate cyclase → decrease of cAMP → inhibitory actions
Glycine (3.2.4.2)	glycine receptor (I)	CNS: binding to ionophoric receptor (similar to GABA receptor) → influx of Cl^- → hyperpolarization → inhibitory action (e.g., of sensory transmission in the spinal cord)
Glutamate (3.2.2.2, major excitatory transmitter in the CNS) and aspartate (3.2.2.4)	NMDA (N-methyl-D-aspartate) receptor (I)	CNS: a) binding to ionophoric receptor located on channel permeable to Ca^{++}, K^+ and Na^+ → excitatory actions (e.g., contribution to long term potentiation for memory processing); b) activation of NO → synthetase activation of guanylate cyclase → increase of cGMP (7.8.2)
	kainate receptor[1] (I)	CNS: binding to ionophoric receptor located on channel permeable to K^+ and Na^+ → excitatory actions (e.g., activation of motor neurons)
	AMPA receptor (quisqualate receptor)[1,3] (I)	CNS: binding to ionophoric receptor located on channel permeable to K^+ and Na^+ → excitatory actions (fast depolarizing of excitatory synapses)
	$mGlu_{2..3}$ receptor (II)	CNS: binding to receptor linked with G-protein → activation of phospholipase C → formation of inositol-P_3 and diacylglycerol → opening of channel permeable to Na^+, K^+ and Ca^{++} → excitatory actions
Neuropeptides (Some examples of this very heterogeneous group)		
Enkephalins, dynorphine and β-endorphin	δ, μ, κ opioid receptors (II)	CNS: binding to receptors linked with $G_{i\alpha}$-protein → inhibition of adenylate cyclase → decrease of cAMP → inhibitory action, e.g., pain inhibition, inhibition of glutamate effects and of transmitters-endorphin (receptor: closing of Ca^{++} channels)

[1] Named after agonists for the glutamate receptor used in research.
[2] Effects also opening of K^+ channels, closing of Ca^{++} channels.
[3] α-amino-3-hydroxy-5-methylisooxazole-4-propionic acid.

Table 7.2-3. Agonists and Antagonists for Postsynaptic Receptors

Receptor (Selection)	Agonists (physiological agonists in **bold type**)	Antagonists
Acetylcholine receptors: nicotinic receptor	**acetylcholine,** nicotine	curare, α-bungarotoxin (muscle type rec.), κ-bungarotoxin (neuronal type rec.)
Muscarinic M_1 (neuronal), M_3 (glandural), M_5 receptors	**acetylcholine,** muscarine, carbachol	atropine, pirenzepine
Muscarinic M_2 (cardiac) and M_4 receptors	**acetylcholine,** muscarine, carbachol	atropine, methoctramine
Adrenergic β_1 receptor	**norepinephrine ≈ epinephrine;** isoproterenol	betaxolol, atenolol
Adrenergic β_2 receptor	**norepinephrine ≥ epinephrine;** isoproterenol	butaxamine
Adrenergic α_1 receptor	**norepinephrine, epinephrine**	
Adrenergic α_2 receptor	**norepinephrine, epinephrine**	prazosin
Dopamine receptors: D_1, D_5 receptors (II)	**dopamine**	SCH23390
D_2 receptors (II)	**dopamine,** quinpirole	(–) sulpiride, haloperidol
D_3, D_4 receptors (II)	**dopamine**	
Serotonin receptors: 5-HT_1 receptor (II)	5-carboxamidotryptamine > LSD > **serotonin**	spiperone, spiroxatrine
5-HT_2 receptor (II)	DOI, LSD, **serotonin**	ketanserin
5-HT_3 receptor (I)	**serotonin**	ondansetron
GABA receptors: $GABA_A$ receptor	**γ-aminobutyrate (GABA),** muscimol; benzodiazepine and barbiturates effect potentiation	biculluline, picrotoxin
$GABA_B$ receptor	**γ-aminobutyrate (GABA),** L-baclofen	2-hydroxy-S-saclofen
Glycine receptor	**glycine, β-alanine, taurine**	strychnine
Glutamate receptors: NMDA (N-methyl-D-aspartate) receptor	NMDA, **glutamate** (co-agonist **glycine**)	2-amino-5-phosphono-pentanoic acid
Kainate receptor	kainate, **glutamate**	
AMPA receptor (quisqualate receptor)	AMPA, quisqualate, **glutamate**	NBQX; 2,3-benzodiazepines
$mGlu_{2,3}$ receptor (quisqualate B receptor)	1-amino-cyclopentane-1,3 dicarboxylic acid (APCD), quisqualate, **glutamate**	2-amino-3-phosphonopropionic acid
Neuropeptide receptors: δ-receptor	**Leu-enkephalin, Met-enkephalin**	naltrindole
κ-receptor	**dynorphin A**	nor-binaltorphimine
μ-receptor	**β-endorphin,** morphine, L-polamidone	CTOP

Table 7.2-4. Components and Characteristics of Axonal Transport

Type		Mechanism	Speed [mm/day]	Transported Components
Fast axonal transport of membrane-associated organelles	Anterograde	saltatory transport along linear arrays of microtubules forming stationary tracks; motor molecules: ATPases, e.g., kinesin	200 … 400	synaptic vesicles and related metabolites, membrane associated proteins and lipids, neurotransmitters (i.e. aminergic transmitters), enzymes
	Mitochondrial	mitochondrial components transported along central microtubuli clusters	50 … 100	mitochondria and mitochondria-associated proteins and lipids
	Retrograde	components packed in large membrane-bound organelles. Motor molecules: dynein, MAP1C	100 … 200	lysosomes, growth factors (e.g., NGF)
Slow axonal transport of cytoskeleton- and cytoplasmic components	type A	slow axoplasmic flow. Components move in polymerized form with regulatory and cross-linking proteins	2 … 6	actin, chlathrin, spectrin, calmodulin, glycolytic enzymes
	type B	same	0.1 … 1	neurofilaments, microtubules

There are several different categories of generic signal transduction events:

- Direct transmitter gated ion channels change the intracellular ion concentration (Receptors type I, already described in 7.2.3). They are activated by ligands like acetylcholine, serotonin, amino acids etc.
- Members of a large family of seven transmembrane receptors (7TMRs) are coupled to heterotrimeric G-proteins (Type II, see 7.2.3 and 7.4.1). These G-protein coupled receptors bind hormones or neurotransmitters of the indirectly gating type, but can also be activated by physical stimuli, e.g., light.
- Signaling via heterotrimeric G-proteins can lead to subsequent:
  - activation of adenylate cyclase (7.4.2), e.g., in olfactory and gustatory sensing (7.4.7)
  - activation of phospholipase C (7.4.3)
    - activation causing activation of protein kinase C (7.4.3)
    - causing release of inositol phosphates (7.4.4), ultimately leading to responses like muscle contraction (7.4.5)
  - activation of cGMP phosphodiesterase, e.g., in the visual process (7.4.6)
  - activation of phospholipase A_2 (7.4.8).
- Many receptors for growth factors and cytokines either have tyrosine kinase activity themselves or can associate with and activate cytoplasmic tyrosine kinases. Tyrosine phosphorylation can not only regulate the conformation, catalytic activity, and stability of proteins, but also serves to build modular signaling complexes by regulating protein-protein interactions. At the end of such cascades of regulatory tyrosine phosphorylation events the signal is often transferred to changes in gene expression by modulation of transcription factors (7.5).
- The receptors binding steroid or thyroid hormones or retinoic acid act themselves as transcription factors (7.6). In contrast to other receptors they are not localized at the cell surface, but bind their ligands in the cytoplasm or nucleus.
- Guanylate cyclase can also be a hormone receptor. It activates various processes (7.8).

These reaction chains are interconnected in multiple ways. The general aspects of each pathway are described in Figure 7.3-1.

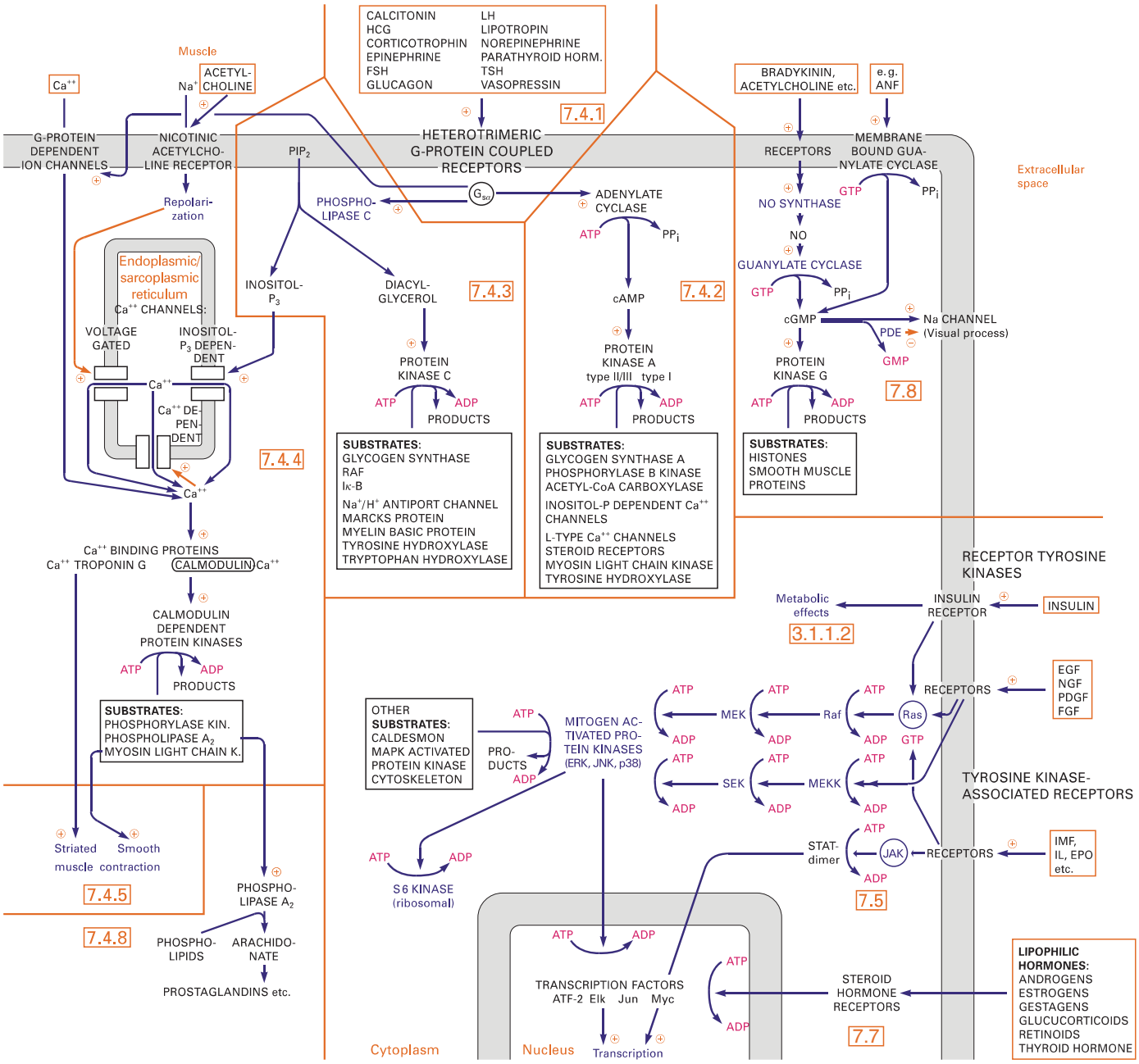


Figure.7.3-1. Receptors and Intracellular Transmission of Messages

7.4 Receptors Coupled to Heterotrimeric G-Proteins

An essential role in signal transduction through the cellular membrane of animals is played by receptors, which are coupled to heterotrimeric G-proteins (Table 7.4-1; they are also found in plants and precursors exist even in yeast). This GPCR family comprises many hundred receptors that share the presence of seven transmembrane passes of the glycoprotein chain as a common structural feature ('serpentine receptors', Fig. 7.2-6). Otherwise, this family shows a high degree of diversity between its members.

The agonists (ligands, which specifically activate the receptor) can be small biogenic amines (e.g., histamine, 3.2.8.2 and epinephrine, 3.2.7.4) or peptides (e.g., bradykinin, 9.2.2) as well as large glycoproteins (e.g., luteinizing hormone, 7.1.5). Even light (in the visual process) can cause receptor activation (7.4.6).

The stimulated receptors, in turn, activate G-proteins, which transmit the signal intracellularly via second messengers like cAMP. Several different receptors may activate the same G-protein type, thus integrating their effects.

The various heterotrimeric G-proteins consist of one each of α, β and γ subunits (40 … 46, 35 … 38 and 6 … 9 kDa, respectively). G_α carries the GTP-binding site and is reversibly membrane anchored by covalent lipid modification (myristoylation, frequently also palmitoylation, 4.4.1). G_γ is prenylated (3.5.3.4). The G_β and G_γ subunits are tightly, but non-covalently associated. In mammals, at least 17 genes for α, 5 genes for β, and 9 genes for γ subunits exist. Differential splicing and post-translational modification further increase the diversity of G-protein isoforms. Although only a moderate number of preferred combinations of heterotrimeric G-proteins and receptors exist, this system provides high flexibility of signal transmission.

The heterotrimeric G-proteins are a subgroup of the GTP-binding proteins (for other members of this group, see Table 7.5-2).

7.4.1 Mechanism of Heterotrimeric G-Protein Action (Fig. 7.4-1)

After forming a complex with the agonist, the receptor undergoes a conformational change. This allows binding of the heterotrimeric G-protein complex consisting of G_α•GDP-G_β-G_γ to its intracellular domain. The interaction with the receptor induces structural changes in the $G_{\beta\gamma}$ subunit. This 'open state' permits the escape of GDP from the G_α subunit, temporarily causing tight binding of the $G_{\alpha\beta\gamma}$ complex to the receptor. However, subsequent binding of GTP to G_α causes separation of G_α from the $G_{\beta\gamma}$ subcomplex leading to dissociation of the subunits from the receptor. While the free G_α and $G_{\beta\gamma}$ subunits propagate the signal, the receptor goes on to activate more G-proteins (signal amplification).

The duration of the signal generated by heterotrimeric G-proteins is controlled by GTP-hydrolysis. As with small monomeric G proteins (e.g., Ras and Rho) this step is regulated by a superfamily of RGS (Regulator of G-protein Signalling) proteins. Some of these proteins have GAP (GTPase-Activating Protein) or GEF (Guanine nucleotide Exchange Factor) function, while others act as guanine nucleotide dissociation inhibitors (GDIs). At the level of RGS proteins, signal integration and cross-talk occurs between GPCR and RTK (7.5.3) signaling. There is some evidence also that RGS proteins play a role in receptor-independent activation of heterotrimeric G-proteins by AGS proteins (Activators of G-protein signalling).

Once activated the free G_α and $G_{\beta\gamma}$ subunits act as activators or inhibitors of cellular components:

- **G_α subunits (Table 7.4-1):** $G_{s\alpha}$•GTP activates the cAMP pathway via adenylate cyclase (7.4.2), while $G_{i\alpha}$•GTP inhibits adenylate cyclase activity. $G_{q\alpha}$•GTP activates protein kinase C (7.4.3) and the inositol (7.4.4) pathways via phospholipase Cβ. Still other effects are caused by other Gα subunits (7.4.6 … 7).
- **$G_{\beta\gamma}$ subunits (Table 7.4-2):** Their effects include activation of adenylate cyclase II and IV, of phospholipase C subtypes β1, β2 and β3, of phospholipase A_2 (7.4.8), of K^+ and Ca^{++} channels, and of Raf and Ras. The latter two effector proteins are also signaling components of the tyrosine kinase pathways (7.5.3). The mechanism of coordination with the G_α effects is only partially known.

After autocatalytic hydrolysis of G_α•GTP to inactive G_α•GDP ($t_{1/2}$ of G_α•GTP is between 20 seconds and several minutes), the G_α subunit re-associates with $G_{\beta\gamma}$ to form the inactive heterotrimeric G protein, which then can enter another signal transduction cycle.

G-protein coupled receptor activity is terminated by a feedback inactivation mechanism, referred to as homologous desensitization (Fig. 7.4-1, bottom left). For instance, in case of the β-adrenergic receptor the $G_{s\alpha}$•GTP mediated increase in cAMP leads to activation of PKA and β-adrenergic receptor kinase (βARK), two serine/threonine kinases that negatively regulate receptor activity by phosphorylating it primarily at carboxyterminal sites. The $G_{\beta\gamma}$ complex helps sequestering βARK to the cell membrane by binding simultaneously with phosphoinositol lipids to its pleckstrin homology domain. βARK is a member of the small family of G-protein coupled receptor kinases (GRKs) that universally regulate all serpentine receptors by binding and phosphorylating them in a strictly ligand-dependent fashion.

While phosphorylation by PKA directly decreases the receptor's affinity for G-proteins, receptor desensitization by GRKs is mediated by binding of inhibitory adapter molecules, known as arrestins. This small family of proteins cooperates with the GRK family in stimulus-dependent downregulation of all 7TMRs by several mechanisms: Arrestins bound to ligand activated receptors not only block further recruitment of G-proteins, but also trigger receptor endocytosis by directing them to clathrin-coated pits (6.1.5). Moreover, β-arrestins quench cAMP signaling by recruiting phosphodiesterases (e.g., PDE4D) to the receptor. This leads to rapid degradation of the second messenger in the vicinity of the initial source of stimulation. However, in recent years it has become clear that the GRK/arrestin system is bifunctional, since it does not only provide desensitization mechanisms, but it also generates a separate set of intracellular signals that act in parallel to those mediated by G-proteins. These GRK/arrestin-dependent signaling pathways are mediated by effector molecules like c-src, MAP-kinases and AKT/PI 3-kinase and play a role in such diverse biological responses as cell survival/anti-apoptosis, chemotaxis, and cardiac contractility.

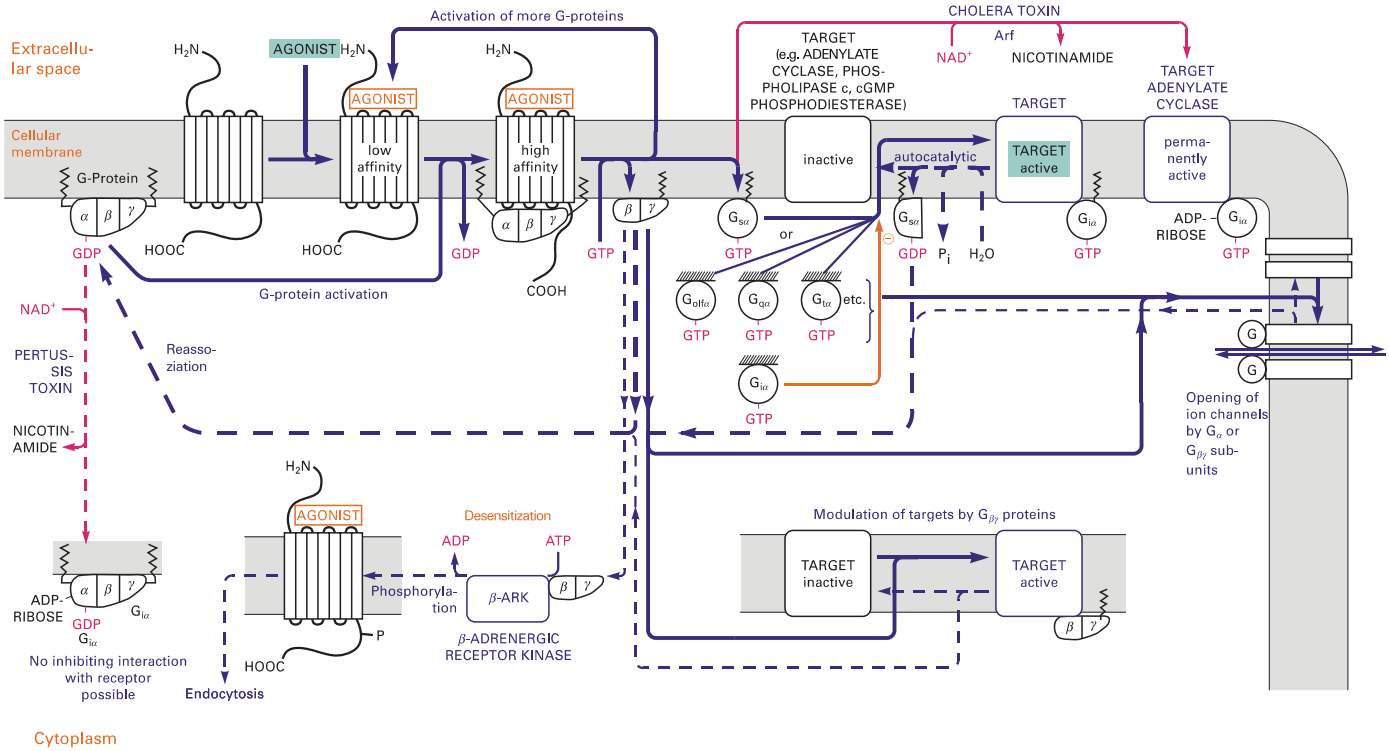


Figure 7.4-1. Activation Mechanism by G_α-Proteins

Table 7.4-1. *Mammalian* G_a Proteins (For indirectly gating nerve transmission systems, compare Table 7.2-2)

G_α Protein	Receptors (R, Examples)	Agonists (Examples)	Effects	Location
G_s Family (stimulatory G_a-proteins, 44 … 46 kDa, react with cholera toxin, palmitoylated at Cys)				
$G_{s\alpha}$ (Large or small splicing variants)	β-adrenergic R, ACTH R, FSH R, LH R, PTH R, TSH R.	norepinephrine, epinephrine ACTH, FSH, LH, PTH	adenylate cyclase type I and III ↑, II and IV (synergistically with $G_{\beta\gamma}$) ↑, L-type Ca^{++} channel ↑, Na^+ channel ↓	ubiquitous
$G_{olf\alpha}$	olfactory R.	odorants (aromatic essences)	adenylate cyclase type III ↑	olfactory epithelium (nose)
G_i Family [inhibitory G_a-proteins, 40 … 41 kDa, react with pertussis toxin (except $G_{z\alpha}$), myristoylated at N-terminal Gly, frequently also palmitoylated at Cys]				
$G_{i1\alpha}$	somatostatin R	somatostatin	adenylate cyclase ↓	ubiquitous
$G_{i2\alpha}$	acetylcholine R. (muscarinic type M_2, M_4)	acetylcholine	Ca^{++} channel ↓, K^+ channel ↑	ubiquitous
$G_{i3\alpha}$	prostaglandin EP_3 R	prostaglandin E	phospholipase A_2 ↑	nearly ubiquitous
$G_{o\alpha A}$ } (splice variants)	Met-enkephalin R	Met-enkephalin	Ca^{++} channel ↓, K^+ channel ↑	brain, nerves
$G_{o\alpha B}$ } (splice variants)	α_2 adrenergic R	norepinephrine, epinephrine	adenylate cyclase ↓, N-type Ca^{++} channel ↓	CNS, sympathetic nervous system
$G_{z\alpha}$	unknown	?	adenylate cyclase ↓, Ca^{++} channel ↓	brain, adrenals, platelets
$G_{t1\alpha}$ (transducin)	rhodopsin	light	cGMP phosphodiesterase γ ↑	rods of retina
$G_{t2\alpha}$ (transducin)	cone opsin	light	cGMP phosphodiesterase γ ↑	cones of retina
$G_{g\alpha}$ (gusducin)	gustatory R	food	adenylate cyclase ↓	taste buds
G_q Family (42 … 43 kDa, palmitoylated at Cys, no reaction with pertussis toxin)				
$G_{q\alpha}$ (formerly $G_p\alpha$)	α_1 adrenergic R	norepinephrine, epinephrine	phospholipase Cβ ↑	muscle, sympathetic nerves
	acetylcholine R (muscarinic type M_1, M_3, M_5)	acetylcholine	phospholipase Cβ ↑	CNS, sympathetic and parasympathetic nerve system
$G_{11, 14, 15, 16}$	serotonin R,IL-8 R,C5a R	serotonin (G_6), IL-8, C5a	phospholipase Cβ ↑	nearly ubiquitous
G_{12} Family (44 kDa, palmitoylated at Cys, no reaction with pertussis toxin)				
$G_{12\alpha, 13\alpha}$	?	?	phospholipase Cβ1 ... 3 ↑	ubiquitous (G_{13a} in platelets)

Table 7.4-2. Mammalian $G_{\beta\gamma}$ Proteins (Tightly Associated Complexes)

G_β Proteins (35 … 38 kDa)	G_γ Proteins (6 … 9 kDa)		Receptors (R, Examples)	Effects	Remarks	Location
β_1	γ_1	*	Rhodopsin	cGMP phosphodiesterase γ ↑	Action independent of activation by G_α	rods of retina
β_1	γ_2	#	β adrenergic R	adenylate cyclase type II/IV ↑	Action independent of activation by G_α	brain, adrenals
β_1	γ_2	#	muscarinic acetylcholine R: M_2	K_{IR}^+ channel ↑ (Table 6.1-1)		Heart
β_1	γ_3	#	somatostatin R	Ca^{++} channel (N type) ↓		brain
β_2	γ_2	#		adenylate cyclase type II ↑	Inhibited by G_α	brain, adrenals
β_2	γ_3	#		adenylate cyclase type II ↑		brain, testes, heart
β_3	not specif.			inhibits Gα activation		brain, testes, heart
not specif.	not specif.		α_1 adrenergic R	phospholipases Cβ2,3	Action independent of activation by G_α	muscle, sympathetic nervous system
not specif.	not specif.		α_{2A} adrenergic R	tyrosine protein kinase, Ras dependent (mechanism?)		CNS, sympathetic nervous system
not specif.	not specif.		β adrenergic R	$G_{\alpha s}$ activated adenylate cyclase type I	Inhibition only by surplus of $G_{\beta\gamma}$	ubiquitous
not specif.	not specif.		(locates receptor-specific kinases of β-ARK family)	phosphorylates muscarinic and β-adrenergic receptors	Inactivation of receptors	ubiquitous

Membrane anchored: * by farnesyl at C-terminal serine, # by geranylgeranyl at C-terminal leucine.
Also known: β_4, β_5, γ_4 (#, pairs with β_1 and β_2, widespread occurrence), γ_5 (#, pairs with β_1 and β_2, widespread occurrence), γ_7 (#, pairs with β_1 and β_2, widespread occurrence), γ_8 (#, in olfactory epithelium), γ_{10} (#, pairs with β_1 and β_2), γ_{11} (*, similar to γ_1, pairs with β_1, widespread occurrence).

Reestablishment of GPCR responsiveness after waning of stimuli (resensitization) also depends on how fast arrestins dissociate from the receptor, since that dictates the rate of receptor dephosphorylation by phosphatases and its subsequent recycling from the endosomal compartment to the plasma membrane.

Pathogenic reactions (Fig. 7.4-1): Cholera toxin or the heat labile enterotoxin of *E. coli* specifically transfer the ADP-ribose residue of NAD^+ via the ADP ribosylating factor (Arf) to the active, GTP-bound $G_{s\alpha}$ subunit. This modification prevents the autocatalytic hydrolysis to $G_{s\alpha}$•GDP and its successive re-association with $G_{\beta\gamma}$. Thus, $G_{s\alpha}$ remains locked in the active state and permanently stimulates adenylate cyclase leading to about a 100-fold increase in the intracellular cAMP level in epithelial cells of the intestine.

Pertussis toxin transfers the ADP-ribose residue of NAD^+ to the inhibitory subunit $G_{i\alpha}$•GDP in the trimeric complex with $G_{\beta\gamma}$. The modified $G_{i\alpha}$ subunit is no longer able to inhibit adenylate cyclase, which results in an elevation of the intracellular cAMP level.

Literature:

Gautam N. Structure 2003;11:359–60.

Hepler, J.R., Gilman, A.G. Trends in Biochem. Sci. 1992;17: 383–387.

Hamm, H.E., Gilchrist, A. Curr. Opin. in Cell Biology 1996;8: 189–196.

Hofmann K.P., *et al.* Trends Biochem. Sci. 2009;34:540–552.

Hubbard K.B., Hepler J.R. Cell. Signal. 2006;18:135–150.

Lefkowitz, R.J. Acta Physiol. 2007;190:9–19.

Neer, E. Cell 1995;80:249–257.

Oldham W.M., Hamm H.E. Q. Rev. Biophys. 2006;39:117-66.

Premont, R.T., Inglese, J., Lefkowitz, R.J. FASEB J. 1995;9: 175–182.

Ray, K., *et al.* J. Biol. Chem. 1995;270:21765–21771.

Reus-Domiano, S., Hamm, H.E. FASEB J. 1995;9:1059–1066.

Siderovski D.P. Willard F.S. Int. J. Biol. Sci. 2005;1:51–66.

Strader, C., *et al.* Ann. Rev. of Biochem. 1994;63:101–132.

7.4.2 cAMP Metabolism, Activation of Adenylate Cyclase and Protein Kinase A (Fig. 7.4-2)

cAMP (3′,5′-cyclic AMP, adenosine 3′,5′-monophosphate) is a messenger molecule in all kingdoms of life. In bacteria, it signals the nutritional status of the cell. In fungi, it regulates e.g., glycolysis, while in higher plants only a few regulatory functions have been demonstrated so far (e.g., Ca^{++} influx, enzyme induction during seed germination). In vertebrates, its major role as 'second messenger' in signal transfer is initiated by the G-protein system.

Adenylate cyclase: After leaving the receptor (7.4.1), $G_{s\alpha}$•GTP binds to adenylate cyclase and activates it to perform the reaction

$$ATP = cAMP + PP_i \qquad \Delta G'_0 = +6.7 \text{ kJ/mol.}$$

On the other hand, $G_{i\alpha}$ inhibits the adenylate cyclase (7.4.1). For other inhibitors and activators see Figure 7.4-2.

The slow hydrolysis of bound GTP by G_α terminates its active state. The second messenger cAMP itself is hydrolyzed by 3′,5′-cyclic nucleotide phosphodiesterase (PDE) to 5′-AMP. This enzyme is activated by insulin or by calmodulin + Ca^{++} and inhibited by methylxanthines, e.g., theophylline or caffeine.

There are eight known isoforms of mammalian adenylate cyclase with different properties. All forms contain 12 transmembrane domains and are activated by $G_{s\alpha}$ and by forskolin (a diterpene). Adenylate cyclase type I is also activated by calmodulin/Ca^{++} (7.4.4) but is strongly inhibited by $G_{\beta\gamma}$. Types II and IV are stimulated by $G_{\beta\gamma}$ in the presence of G_{sa}. Types V and VI are inhibited by $G_{i\alpha}$ and by Ca^{++} (Tables 7.4-1 and 7.4-2).

Protein kinase A (PKA): cAMP activates protein kinase A, which, in turn, regulates the activity of key enzymes in many essential metabolic pathways by phosphorylation. Major PKA substrates are found in the cytoplasm as well as in the nucleus (Table 7.4-3). Activation of PKA substrates is terminated when specific phosphatases, e.g., in glycogen metabolism phosphoprotein phosphatase-1 (PP-1, 3.2.4) dephosphorylate these enzymes. Some of the phosphatases are regulated themselves by phosphorylation (e.g., PP-1 by insulin stimulated protein kinase).

Inactive PKA from muscle consists of two regulatory and two catalytic subunits (C_2R_2). Binding of cAMP dissociates the complex by an allosteric mechanism. The free subunits C_2 are enzymatically active. Important substrates of protein kinase A are shown in Table 7.4-3. The substrate proteins are phosphorylated at Ser or Thr residues in the context of the consensus sequence: Arg-Arg-small amino acid-(Ser/Thr)-hydrophobic amino acid.

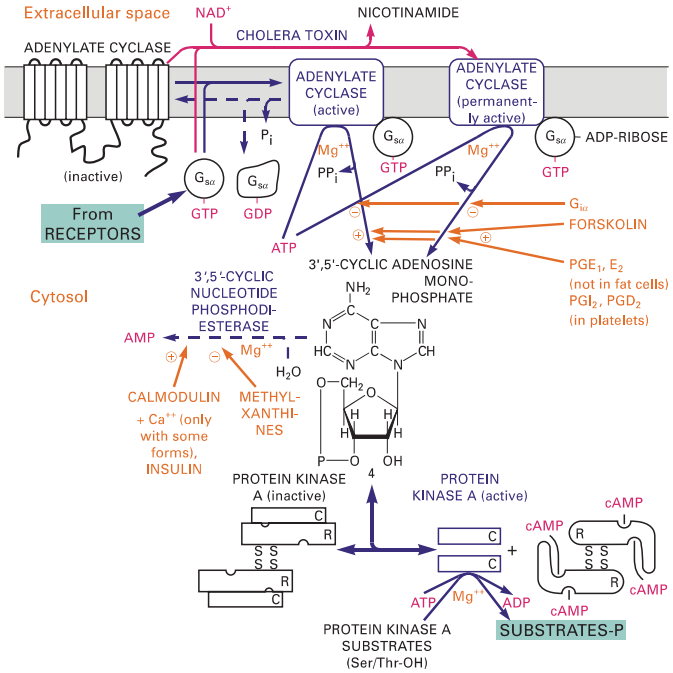


Figure 7.4-2. Metabolism and Effects of cAMP and Protein Kinase A in Vertebrates

Substrates	Effect of Phosphorylation	Section
Glycogen synthase a	synthase a → b, activity ↓	3.1.2.4
Phosphorylase kinase	activity ↑ (if Ca^{++} present)	3.1.2.4
Phosphatase inhibitor protein 1	activity ↑	3.1.2.4
Phosphofructo-2-kinase/fructose 2,6-bisphosphatase	kinase ↓/ bisphosphatase ↑	3.1.1.2
Acetyl-CoA carboxylase	activity ↓	3.4.1.2
Triacylglycerol lipase	activity ↑	3.4.2.2
Cholesterol esterase	activity ↑	3.5.1.5
Inositol-P dependent Ca^{++} channels	close	7.4.4
Phospholamban	Ca^{++} transporting ATPase ↑	7.4.4
L-Type Ca^{++} channels (*muscles*)	open	7.4.5
Steroid receptors (*nucleus*)	transcription ↑	4.2.2.3, 7.7
CREB (*nucleus*)	transcription ↑	4.2.2.3
Myosin light chain kinase	smooth muscle relaxation	7.4.5

Table 7.4-3. Substrates of Protein Kinase A (Examples)

Literature:

Fischer, E.H. Angew. Chem. 1993;105:1181–1188 (Nobel Prize Lecture).

Walsh, D.A., van Patten, S.M. FASEB J. 1994;8:1227–1236.

Gerits N, *et al.* Cell Signal. 2008;20:1592–1607.

7.4.3 Activation of Phospholipase C and Protein Kinase C (Fig. 7.4-3)

Phospholipases C (PLC): This family of enzymes is an important control center for intracellular signal processing. Most of these enzymes are located at the inner surface of the cellular membrane and become activated by binding of subunits of heterotrimeric G-proteins (7.4.1) in the presence of Ca^{++}. The activated G-protein subunit $G_{q\alpha}$•GTP binds to the isoenzymes of the PLCβ group and activates them to hydrolyze membrane bound 1-phosphatidyl-1D-myo-inositol (4,5)-P_2 (PIP_2), yielding two second messengers:

- Inositol (1,4,5)-P3 (IP_3) opens Ca^{++}-channels at the sarcoplasmic and endoplasmic reticulum, thus achieving an elevation of the cytosolic Ca^{++}-concentration (7.4.4).
- 1,2-Diacylglycerol (DAG, mostly 1-stearoyl-2-arachidonyl-*sn*-glycerol) remains in the plasma membrane and, in combination with phospholipid and Ca^{++}, activates membrane-bound protein kinase C (PKC). This enzyme is a connecting link to the signal transduction pathways mediated by tyrosine kinases (7.5). Diacylglycerol can also arise from other sources, e.g., from phosphatidylcholine by phospholipase C or by phospholipase D + phosphatidate phosphatase (3.4.3.2).

In addition to $G_{q\alpha}$•GTP, the $G_{\beta\gamma}$ subunits are also able to activate the isoenzymes PLC-β2 and -β3 (Table 7.4-2).

Other isoforms of phospholipase C: The activation of the isoenzymes PLC-γ1 and PLC-γ2 is independent of G-proteins and occurs via activated receptor tyrosine kinases (RTKs, e.g. PDGF and NGF receptors, 7.5.3), to which these cytosolic enzymes bind with their

src-homology 2 (SH2) domains. RTKs then phosphorylate several tyrosine residues on PLC-γ1/2, thereby activating these enzymes while they are membrane-attached and in close proximity to their substrate phosphatidylinositol 4,5-bisphosphate (PIP_2, Fig. 7.4-3, bottom left). In muscle, the degradation of PIP_2 by PLC-γ dissolves PIP_2-profilin complexes. The liberated profilin associates with actin and, in turn, disassembles actin filaments (7.4.5).

The activation mechanism and role of PLC-δ isoenzymes are yet unknown, but their structures resemble those of PLC-β forms.

Protein kinases C are a large family of serine/threonine kinases (so far 12 known in mammals, 68 … 84 kDa), which phosphorylate a wide range of substrates, including numerous receptor tyrosine kinases and transcription factors (Table 7.4-4). The 'conventional' isoforms PKCα, β, γ are activated by diacylglycerol (supplied by PLC), membrane phosphatidylserine and Ca^{++}. This involves a transfer from the cytosol to the membrane. The tumor promoter phorbol is an efficient activator in place of diacylglycerol.

Other isoenzymes of PKC: The 'novel' PKCδ, ε, η, μ, and σ only need diacylglycerol and phosphatidylserine for activation. It is assumed that these PKCs are stimulated by DAG originating from phospholipase D action and thus are not involved in the G-protein regulatory system. PKCε plays a role in the contraction of smooth muscles (7.4.5). During activation, both the 'conventional' and the 'novel' PKC isoenzymes are bound to the cytoplasmic side of membranes. Another group, the 'atypical' isoforms, PKCλ and ζ, remain cytosolic and do not require diacylglycerol and Ca^{++} for maximum activity. They are possibly involved in the MAPK cascade and function in the establishment of cell polarity (7.5.3).

Table 7.4-4. Substrates of Proteinkinase C (Examples)

Substrates	Effect of Phosphorylation	see Section
Glycogen synthase a	synthase a → b, activity ↓	3.1.2.4
Raf (upstream of MAPK cascade)	transcription ↑	7.5.3
Iκ-B (inhibitory protein)	release of κ-B, transcription ↑	4.2.2.3
Na^+/H^+ antiport channel	opens, cytoplasmatic H^+ ↑	Table 6.1-3
MARCKS protein (actin binding)	release of actin from membrane	
Tyrosine hydroxylase (brain)	activity	3.2.7.3
Tryptophan hydroxylase (brain)	activity	3.2.7.3

Literature:

Jaken, S. Current Opinion in Cell Biol. 1996;8:168–173.

Lee, S.B., Rhee, S.G. Current Opinion in Cell Biol. 1996;8:183–189.

Newton A.C. J. Lipid Res. 2009;50(Suppl. S2):66–71.

Newton A.C. Am. J. Physiol. Endocrinol. Metab. 2010;298: E395–402.

7.4.4 Metabolic Role of Inositol Phosphates and Ca^{++}

Calcium ions play a decisive role in the regulation of many cellular processes (Table 7.4-5). Inositol phosphates have a transmitter function in one of the Ca^{++} release mechanisms.

Ca^{++} concentration in cells: The concentration in the cytosol of all eukaryotic cells is very low (ca.10^{-7} mol/l), while on the outside of cells and in the smooth endoplasmic reticulum (ER, in muscles: sarcoplasmic reticulum, SR) it is above 10^{-3} mol/l. This gradient is preserved by a Ca^{++}/Na^+ antiport mechanism in the cellular membrane and also by Ca^{++}-transporting ATPases in the cellular membrane ① and in the ER/SR membrane ② of nerve and muscle cells (Fig. 7.4-4). In plants, Ca^{++} is accumulated in vacuoles.

In slow skeletal and in cardiac muscles, the pumping into the SR is positively regulated by phospholamban after phosphorylation by PKA or by calmodulin dependent protein kinase. Inside of the SR, Ca^{++} is sequestered by two binding proteins, calsequestrin (55 kDa, highly acidic, >40 Ca^{++} binding sites/molecule) and calreticulin.

Release of Ca^{++} from intracellular stores in animals (Fig. 7.4-4, left): The ER/SR membrane of various organs contains two types of Ca^{++} channels, which open after receiving appropriate signals and release Ca^{++} into the cytosol (see Table 6.1-2):

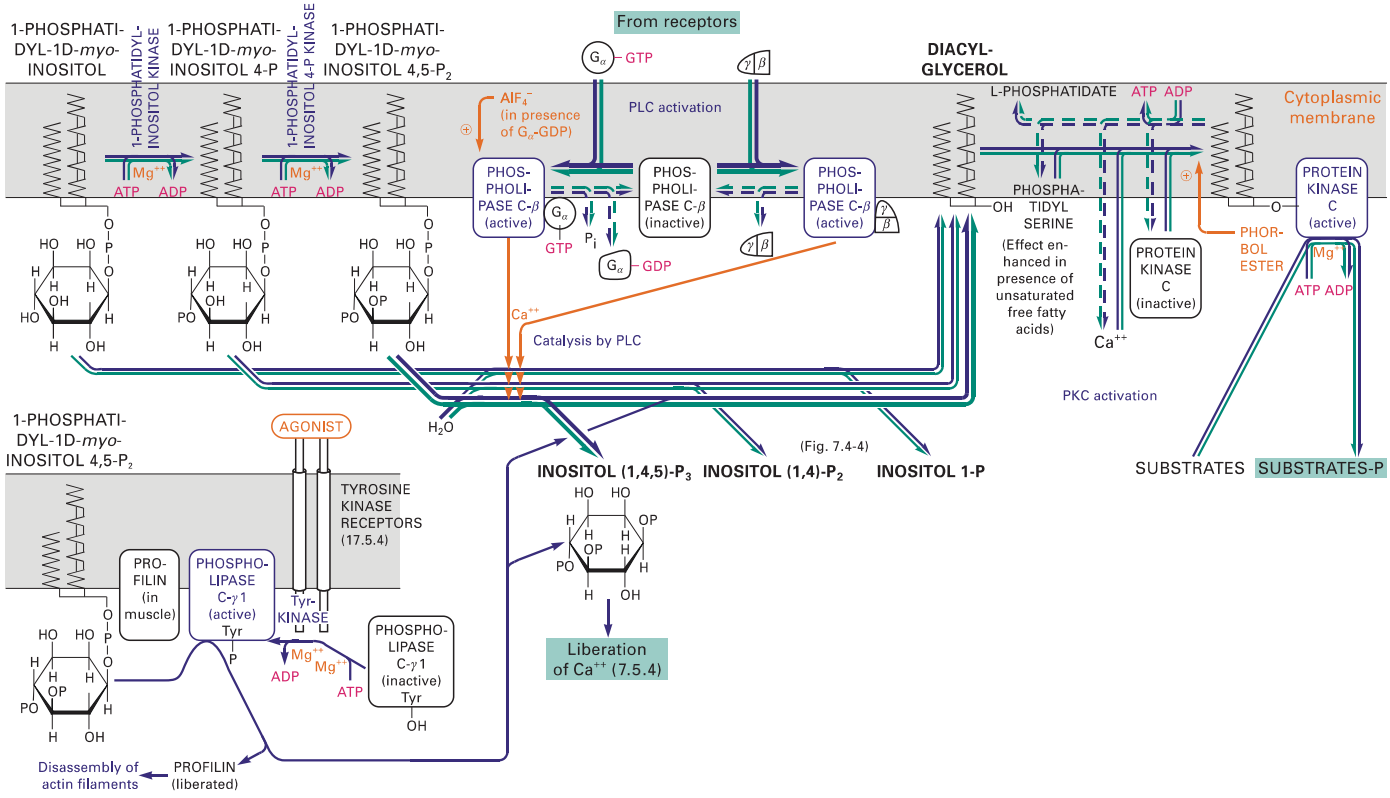


Figure 7.4-3. Activation of Phospholipase C and Protein Kinase C

- Inositol(1,4,5)-P3 receptor Ca^{++} channels (IP_3R, ③, e.g., in *Xenopus* oocytes): They open after binding of IP_3, which is generated by PLC action as described above (7.4.3).
- Ryanodine receptor (RYR) Ca^{++} channels open
  - after sensing a slight elevation of the cytosolic Ca^{++} concentration caused by Ca^{++} entry through voltage-gated dihydropyridine receptor Ca^{++} channels (DHPR) at the cellular membrane, ④a, e.g., in cardiac muscle) or
  - after sensing a conformation change in the DHPR Ca^{++} channels effected by depolarization of the cellular membrane ④b, e.g., in skeletal muscle).

  Ryanodine and dihydropyridine are compounds used in research to characterize the channel types.
- Both types of channels are present, e.g., in heart atrium, vascular smooth muscle, pituitary gland and neurons.

For both receptor types raising the cytosolic Ca^{++} concentration promotes channel opening, but higher concentrations (above ~300 nmol/l) are inhibitory and re-close the channels. This feedback mechanism causes the formation of Ca^{++} waves, which spread through the cytosol, into organelles (nucleus, mitochondria) and even into neighboring cells. Frequently, this occurs in a series of 'spikes'.

Both channel types are structurally and functionally closely related. They consist of four protein chains, each with eight transmembrane domains near the C terminus (forming the Ca^{++} channel) and a long N-terminal sequence located in the cytosol (important for signal sensing, IP_3 binding site).

Emptying of the intracellular Ca^{++} stores initiates a Ca^{++} influx through the plasma membrane. This is mediated either by a signal transfer from the receptors on the ER/SR membrane to Ca^{++} channels on the cellular membrane ⑤ ('capacitative' Ca^{++} influx) or through channels regulated by inositol (1,3,4,5)-P_4 (IP_4), the phosphorylation product of IP_3, and perhaps also by IP_3 itself ⑥. The Ca^{++}-transporting ATPase ⑦ immediately transfers the Ca^{++} from the cytosol back into the ER/SR and thus reconstitutes the low cytosolic Ca^{++} concentration.

Role of calmodulin: Calmodulin (CaM, 17 kDa) is a highly conserved, eukaryotic, acidic protein with four high affinity Ca^{++}-binding sites (10^{-6} mol/l). The Ca^{++}/CaM complex activates a large number of enzymes (Table 7.4-5, mostly CaM dependent Ser/Thr kinases). CaM can be described as a 'decoder' of the Ca^{++} information, which plays a central role in cellular regulation.

The four Ca^{++} binding sites of CaM are located in two pairs of helix-loop-helix motifs ('EF hands') at both ends of the molecule. Ca^{++} binding causes a conformational change in CaM. This exposes two hydrophobic patches which are suitable to bind to other proteins in a regulatory way, e.g., by removing autoinactivating protein segments from the active center of kinases. The Ca^{++} signal is terminated when the Ca^{++} concentration in the cytosol drops by active Ca^{++} export as described above. However, once activated the multifunctional CaM

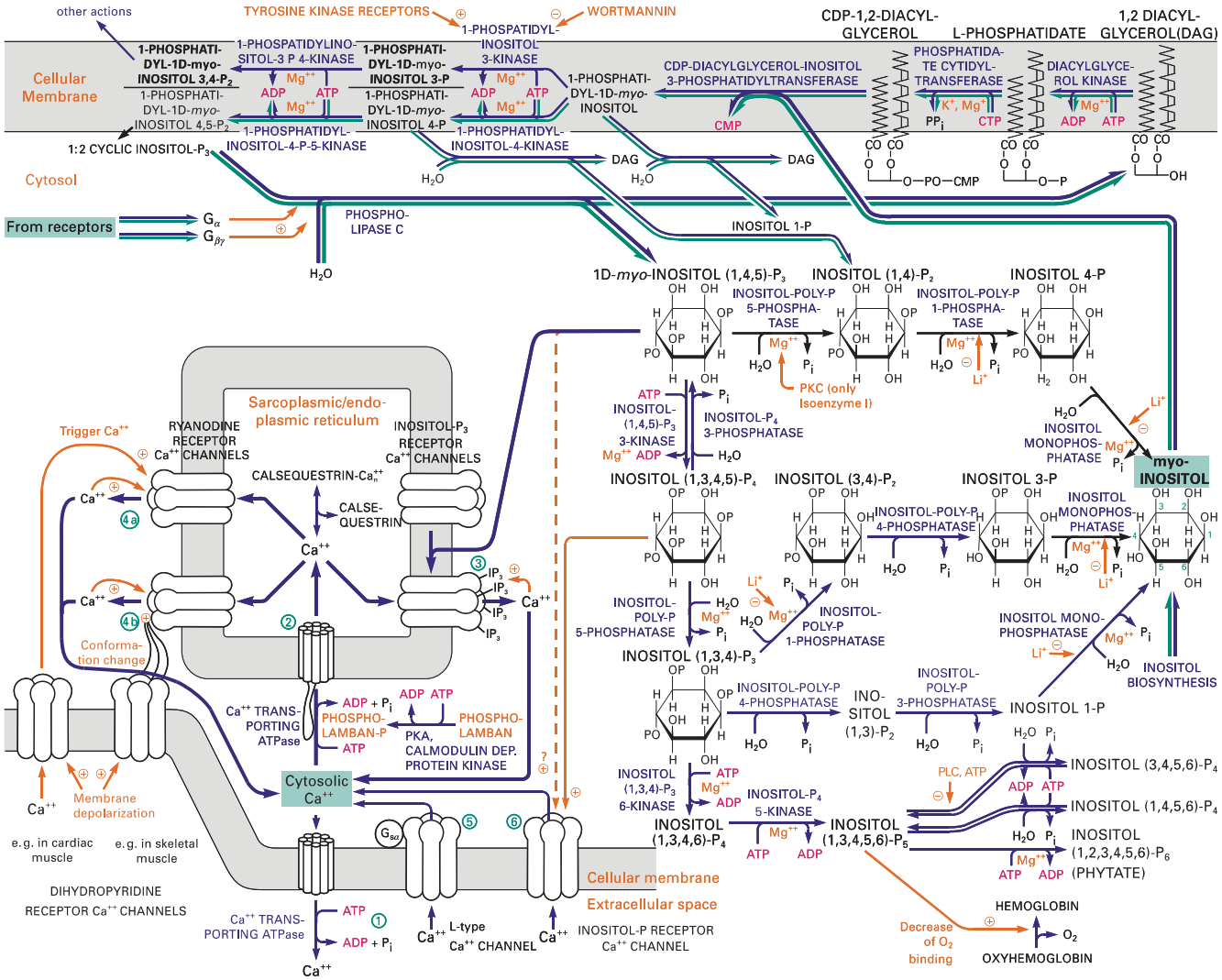


Figure 7.4-4. Metabolism of Inositol Phosphates and Release of Ca^{++} from Intracellular Stores

kinase II also phosphorylates itself, thereby extending its activity until dephosphorylation occurs ('molecular memory', Table 7.4-5).

Table 7.4-5. Targets Regulated by Binding of Calmodulin-Ca^{++}

Target (Examples)	Effect	Section
Phosphorylase kinase	glycogen degradation ↑	3.1.2.4
CaM kinase II, phosphorylates:		
Acetyl-CoA carboxylase	activity ↓	3.4.1.1
ATP citrate lyase		3.1.8.1
Glycogen synthase	activity ↓	3.1..2.4
Phospholipase A_2	activity ↑	7.4.8
N type Ca^{++} channels (neurons)	open, cytosolic Ca^{++} ↑	6.1.2.1
Inositol P_3 kinase	IP_4 + cytosolic Ca^{++} ↑	7.4.4
Ca^{++} pumping ATPases [1]	cytosolic Ca^{++} ↓	7.4.4
Myosin light chain kinase	smooth muscle contraction	7.4.5
CaM kinase IV, phosphorylates: CREB	transcription ↑	4.2.2.3
Protein phosphatases (e.g., calcineurin)	transcription ↑	7.5.4
Adenylate cyclase I (brain)	activated, cAMP ↑	7.4.2
cAMP phosphodiesterase	activated, cAMP ↓	7.4.2
Soluble NO synthase Type III	guanylate cyclase ↑	7.8.2

[1] Plasma membrane enzyme is directly activated, ER/SR enzymes via CaM kinase.

Metabolism of inositol phosphates (Fig. 7.4-4, right): The second messenger IP_3 loses its regulatory function within seconds by removal of the 5-phosphate group. Further dephosphorylation leads to myo-inositol. Different inositol phosphates are formed by phosphorylation and dephosphorylation steps. The pathways described here have been demonstrated mostly in animals. The metabolism in plants is less well known.

The action of inositol polyphosphate 1-phosphatase on inositol (1,3,4)-P_3 or (more weakly) on inositol (1,4)-P_2 as well as the action of inositol monophosphate phosphatase are inhibited by Li^+ ions. This might be the base for treatment of manic-depressive conditions by Li^+.

Higher inositol phosphates: Several phosphorylation steps, starting from inositol (1,3,4)-P_3 lead to inositol tetra-, penta- and hexaphosphates and compounds with additional pyrophosphate groups. Inositol (1,3,4,5,6)-P_5 and inositol-P_6 form the bulk of inositol in mammals, but are synthesized slowly. In birds, inositol (1,3,4,5,6)-P_5, similar to 2,3-bisphosphoglycerate in other animals, decreases the affinity of hemoglobin for O_2 (3.3.2.2). The details of the hexaphosphate (phytate) biosynthesis in plants are not fully known. The large quantities of phytate in seeds are a storage pool.

Reconstitution of phosphatidylinositol phosphates: Diacylglycerol is phosphorylated to L-phosphatidate and then converted to CDP-diacylglycerol. This activated compound reacts with myo-inositol to 1-phosphatidylinositol (3.4.3.2). Phosphorylation by a 4- and a 5-kinase to 1-phosphatidyl-1D-*myo*-inositol (4,5)-P_2 closes the inositol reaction circle (Fig. 7.4-4).

A phosphatidylinositol 3-kinase (PI 3-K) with a catalytic (p110) and a regulatory (p85) subunit is involved in mitogenic signalling via protein tyrosine kinase receptors (7.5.3) and receptors associated with tyrosine kinases (7.5.4). PI 3-K plays a role in cell motility/adherence, cell survival/anti-apoptosis, in vesicle trafficking (4.5.2), and in protein secretion (e.g., of insulin). In cancer, the PI 3-K pathway is often deregulated (e.g., by mutation of the *pten* gene that encodes the lipid/protein phosphatase that negatively regulates this pathway).

Literature:

Berridge, M.J. Biochim. Biophys. Acta 2009;1793:933–940.

Carafoli E. Trends Biochem. Sci. 2004;29:371–379.

Carnero, A., *et al.* Curr Cancer Drug Targets 2008;8:187–198.

Carpenter, C., Cantley, L.C. Curr. Opinion in Cell Biol. 1996;8: 153–158.

James, P., *et al.* Trends in Biochem. Sci. 1995;20:38–42.

Miyazaki, S. Curr. Opinion in Cell Biol. 1995;7:190–196.

Sasaki T., *et al.* Prog. Lipid Res. 2009;48:307–343.

(Various authors). Adv. in Sec. Mess. Phosphoprot. Res. 1995; 30:1–174.

7.4.5 Muscle Contraction

Although muscular contraction depends generally on an interaction between myosin and actin, controlled by the sarcoplasmic Ca^{++} concentration, the mechanisms differ in the various muscle types (striated muscle, cardiac muscle, smooth muscle).

Structure of striated (voluntary) muscles (Figs. 7.4-5, Table 7.4-6): Striated muscles of vertebrates and arthropods contain bundles of muscle fibers. Each fiber is a long, multinucleated cell which can extend through the whole length of the muscle. It is formed by fusion of precursor cells (syncytium). Within these cells there exists an arrangement of parallel myofibril bundles. These are composed of partially overlapping thick and thin filaments containing myosin and actin, respectively, as major components. They cause the striated (striped) appearance in the microscope. The movement of these filaments relative to each other effects the muscular contraction.

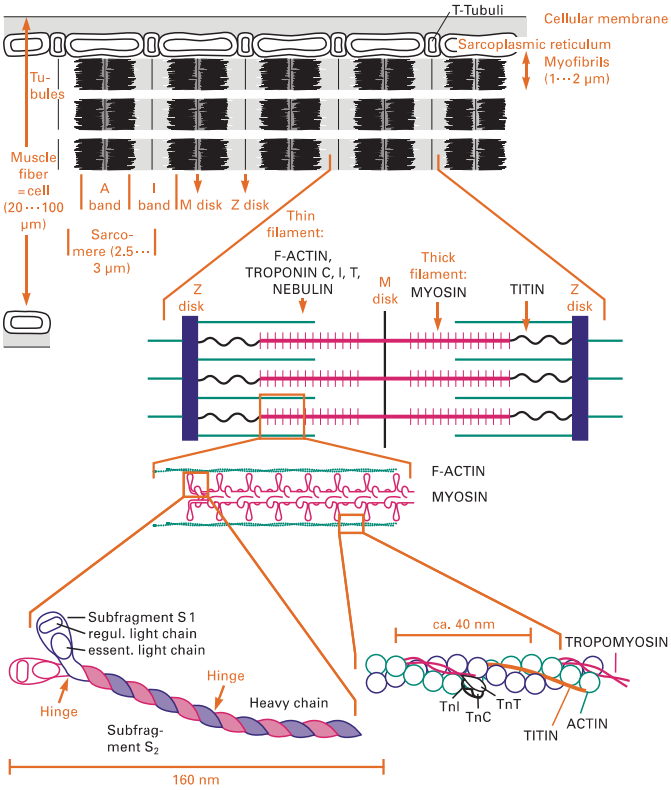


Figure 7.4-5. Structure of Myofibrils in Striated Muscle

Contraction of striated muscles (Fig. 7.4-6): The contraction starts upon arrival of a neuronal signal, which is transmitted in a Ca^{++} dependent way (see below). The contraction mechanism resembles a 'rowboat' movement with the following steps:

① ATP binding to the myosin head S1 forces open a cleft in its globular structure, which results in its separation from actin.

② The resulting change in myosin configuration triggers hydrolysis of the myosin-bound ATP. The myosin head is activated ('cocked') and turns more towards the Z disk.

③ The myosin head initially binds only weakly to a new actin site closer to the Z disk, however the release of P_i strengthens the actin-myosin binding by closing the actin cleft.

④ The 'power stroke' pulls the thick filament approximately 6 nm closer to the Z disk.

Table 7.4-6. Proteins of the Contractile System in *Vertebrate* Striated Muscle

Unit	Component	Subunits	Mol. Mass (kDa)	Structure, Function	Percentage in Myofibrils
Thick filament	myosin	heavy chain	2 * 230	Globular N-terminal head with long α-helical tail, which associates with a second chain into a coiled coil. Many coils aggregate into thick filaments. ATPase energizes muscular contraction.	44 … 50
		essential light chain	2 * 15 … 22	Associates with the globular head of heavy chain	
		regulatory light chain	2 * 15 … 22	Associates with the globular head of heavy chain	
	titin	heterodimer	1200/1400	Connects thick filament with Z-disk (spring function?)	9
Thin filament	F-actin	polymer of G-actin	n * 42	Polymerization with ATPase action. The polymer is helically arranged. Each monomeric unit binds 1 head of the myosin heavy chain near the cleft between the domains.	22
	tropomyosin	homodimer	2 * 33	Coiled coil of α-helices, wrapped around F-actin helix	5
	troponin	TnC	18	TnC: Ca^{++} binding, homologous to calmodulin. TnI: actin binding.	5
		TnI	23	TnT: tropomyosin-binding.	
		TnT	31	Together with tropomyosin they regulate the muscular contraction.	
	nebulin	3	3 * ca. 250	Actin-binding, coils around thin filament	3
M disk	C-protein	–	150	Assists assembly of thick filaments	1
	M-protein	–	100	Assists assembly of thick filaments	1
Z disk	α-actinin	homodimer	2 * 95	Interior of Z-disk, symmetrically binds the ends of F-actin	1
	desmin		50	Peripheral zone of Z-disk	
	vimetin		52	Peripheral zone of Z-disk	

⑤ Release of ADP enables another round of ATP binding and repetition of the cycle.

⑥ ADP is re-phosphorylated by phosphocreatine, which represents a large pool of energy-rich phosphates (3.2.9.2).

While isolated myosin has only low ATPase activity (turnover number 0.05 sec^{-1}), due to slow release of the reaction products, this rate increases to 10 sec^{-1} in the presence of actin. The two proteins form ion pairs between Lys residues of myosin and Asp and Glu residues of actin during the respective phases of the contraction cycle. Also hydrophobic areas in both molecules contact each other. The myosin light chains stabilize the α-helix in the heavy chain head.

Regulation of striated muscle contraction (Fig. 7.4-7): Ca^{++} plays a decisive role in muscular contraction. After the arrival of neuronal impulses at the neuromuscular endplate (7.2.3) the muscle cell depolarizes. This action potential propagates along the T-tubuli (membrane invaginations) towards contact sites of the membrane with the sarcoplasmic reticulum (SR). The voltage-sensitive DHP-receptors of the cellular membrane transmit the signal to the ryanodine receptor Ca^{++} channels of the SR membrane (Fig. 7.4-4). Ca^{++} is released from this store and greatly increases the cytosolic Ca^{++} concentration.

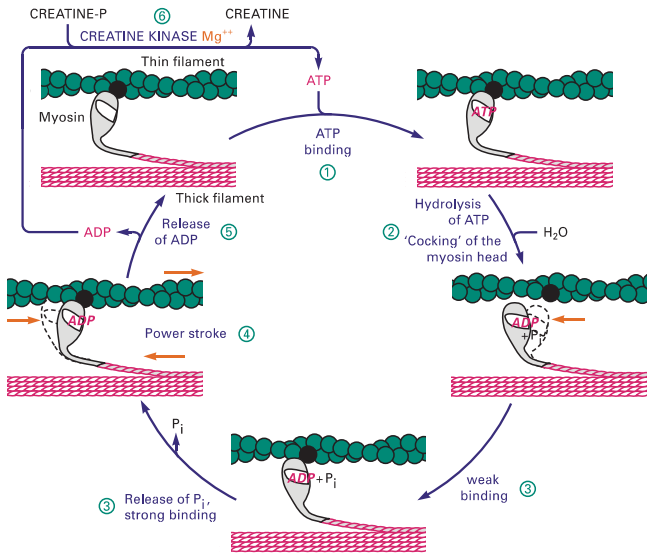


Figure 7.4-6. Mechanism of Muscle Contraction

The interaction of actin and myosin can only take place above a certain threshold concentration of Ca^{++}. Once this threshold is reached, Ca^{++} binds to troponin C causing it to allosterically interact with tropomyosin. By moving deeper into the groove of the actin helix tropomyosin uncovers the contact sites for myosin. After termination of the neuronal stimulus, Ca^{++} is pumped out of the cytosol again and the initial state is resumed.

Cardiac (heart) muscle: This muscle has a similar structure and contraction mechanism as voluntary striated muscles, although in this case the major signals for contraction originate in the heart itself.

The ryanodine receptor Ca^{++} channels of the SR membrane are Ca^{++} gated. They respond to the small influx of Ca^{++} into the cytosol that occurs after stimulation of the DHPR Ca^{++} channels by membrane depolarization (Fig. 7.4-4). The heart depends strictly on aerobic ATP synthesis via the respiratory chain (3.11.4.5), while striated muscle can temporarily provide extra ATP by glycolysis in addition to the aerobic metabolism (3.1.1.5).

Structure of smooth muscles: The muscle cells are spindle shaped and mononuclear. The filaments do not form myofibrils and are differently (e.g., diagonally) arranged. Since they contain no troponin, different mechanisms regulate the contraction. Smooth muscle myosin has only 1/10 of the ATPase activity of its striated muscle counterpart. It is suited for strong, but slow contractions.

Contraction of smooth muscles: Smooth muscles are controlled by the autonomic (involuntary) nervous system. Only one of the two principles of its contraction is Ca^{++} regulated.

- Calcium-dependent contraction (Fig. 7.4-8): Neuronal or hormonal stimulation of smooth muscle cells leads to an increase in cytosolic Ca^{++} by inflow from the extracellular space via voltage- or receptor-operated Ca^{++} channels (7.2.3) or from the sarcoplasmic reticulum

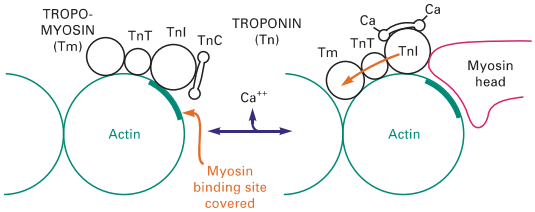


Figure 7.4-7. A Model for Triggering of Striated Muscle Contraction by Ca^{++}

via inositol-P_3-receptor Ca^{++}-channels (7.4.4). The Ca^{++} ions bind to calmodulin (7.4.4), which, in turn, becomes attached to myosin light chain kinase (MLCK) and activates it to phosphorylate the myosin light chain LC_{20}. This causes at first aggregation of myosin molecules and then their association with actin. The contraction cycle proceeds similarly as described above for striated muscle.

The ability of MLCK to interact with calmodulin is abolished by phosphorylation of the enzyme. This reaction is catalyzed by protein kinase A and thus depends on the epinephrine-cAMP pathway (7.4.2). Therefore, epinephrine can relieve asthmatic bronchoconstriction. A similar phosphorylation takes place by calmodulin dependent protein kinase II (not shown in Fig. 7.4-8) effecting desensitization at prolonged high Ca^{++} concentrations.

When the signals for elevation of the cytosolic Ca^{++} concentration cease, Ca^{++} is pumped out of the cytosol again (7.4.4). The decreasing cytosolic concentration causes dissociation of the Ca^{++}/calmodulin/myosin light chain kinase complex and inactivation of the kinase. The myosin light chains are dephosphorylated by myosin light chain phosphatases (e.g. $PP1_M$). Muscle relaxation ensues.

- Calcium-independent contraction: The mechanism is still not fully elucidated. However, a key role is played by Ca^{++} independent protein kinase C isoforms, especially PKCε (7.4.3). After activation by diacylglycerol as second messenger they either directly phosphorylate calponin or indirectly activate caldesmon via the Ras/Raf/MEK/MAPK phosphorylation cascade (7.5.3). Both proteins are associated with the thin filaments and can play a role in smooth muscle contraction.

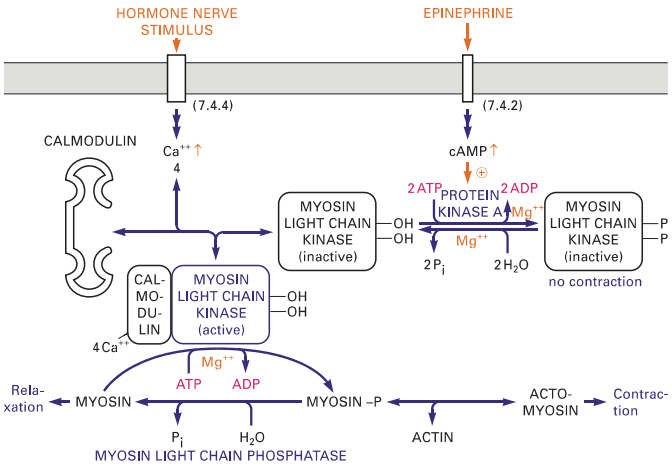


Figure 7.4-8. Regulation of Smooth Muscle Contraction

Nonmuscle cells: The contraction of actomyosin fibers involves myosin light chain phosphorylation and proceeds similarly as described above. Examples are the contractile ring formed during cell division (4.3.5), adhesion belts in myoepithelial and myofibroblast cells, and the cellular microfilaments.

Literature:

Allen, B.G., Walsh, M.G. Trends in Biochem. Sci. 1994;19:362–368.

Mesentean S., *et al.* J. Mol. Biol. 2007;367:591–602.

Rayment, I., *et al.* Science 1993;261:50–65.

Somlyo A.V., *et al*. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 2004;359:1921–1930.

(Various authors). Curr. Topics in Cell. Reg. 1990;31:1–271.

7.4.6 Visual Process

The eyes of vertebrates contain photopigments in stacked membranes (discs) inside of the photoreceptor nerve cells. In the retina of higher vertebrates, both rod cells (for vision at low light levels) and cone cells (responsible for color view) are present.

Rods (Figs. 7.4-9 and 7.4-10): The visual pigment in rods is rhodopsin, which consists of the 7 transmembrane helix protein opsin and 11-*cis*-retinal (3.5.3.2). Opsin itself is colorless; while free retinal, which in solution exists as a mixture of several freely interconvertible conformers, is faintly yellow due to the fact that it weakly absorbs light at the extreme blue end of the visible spectrum (380 nm). When covalently linked to opsin, retinal assumes a defined conformation with all bonds except for 6s-*cis* and 11-*cis* adopting a *trans* conformation. As a result of this, the adsorption spectrum of rhodopsin is shifted into the visible spectrum. Although the absorption peak can vary widely between species, vertebrate rhodopsins typically have their absorption maximum in the blue-green range around 500 nm ($\varepsilon_{500\ nm} = 4 \times 10^4\ cm^{-1} * mol^{-1}$, Fig. 7.4-11). Light isomerizes the retinal in rhodopsin within picoseconds to *all-trans*-retinal, resulting in a series of conformational changes of the protein moiety and leading to the photoexcited metarhodopsin II (R*).

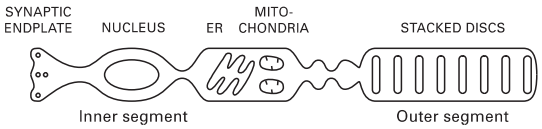


Figure 7.4-9. Structure of Retinal Rods

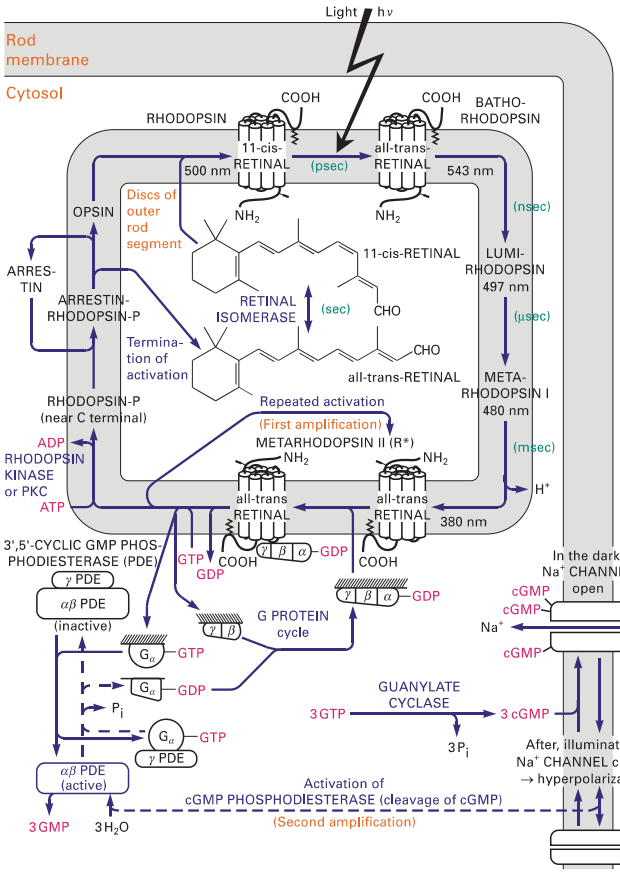


Figure 7.4-10. Visual Cycle in Retinal Rods
(500 nm etc: wavelength of absorption maxima)

Metarhodopsin II, like other serpentine receptors (7.4), activates a heterotrimeric G-protein, in this case transducin ($G_{t1\alpha}$•$G_{\beta\gamma}$). The liberated $G_{t1\alpha}$•GTP binds to the inhibitory γ subunits of cGMP phosphodiesterase (PDE) and removes them, thus initiating hydrolysis of cGMP.

In the dark, binding of cGMP keeps the Na^+ channels of the rod cells open (Table 6.1-1). The Na^+ influx, in turn, raises intracellular Ca^{++} levels by the opening of voltage-gated Ca^{++} channels, thus keeping the cells in a depolarized state. The hydrolysis of cGMP following stimulation of rhodopsin by light closes the Na^+ channels and stops the

Na^+ (and concomitantly also the Ca^{++}) inflow. This leads to hyperpolarization of the cells (see Fig. 7.2-1), which changes the transmission of glutamate-mediated neuronal signals.

Reconstitution of cGMP: Autocatalytic GTP hydrolysis of $G_{t1\alpha}$•GTP eventually terminates the activating input for PDE and thereby reduces the rate of cGMP degradation. Simultaneously, the low intracellular Ca^{++} concentration relieves inhibition of guanylate cyclase activating protein, an enzyme that stimulates guanylate cyclase (7.8.1), thereby increasing the rate of cGMP synthesis. Both mechanisms cooperate in rapidly re-establishing normal cGMP levels.

Reconstitution of rhodopsin: The photoexcited metarhodopsin II gets phosphorylated by rhodopsin kinase or PKC (7.4.3) and binds to arrestin, which stops its ability to activate transducin. The Schiff base structure connecting opsin and *all-trans*-retinal is deprotonated within a few seconds, resulting in the separation of both moieties. *All-trans*-retinal is reconverted in a dark reaction to 11-*cis*-retinal and then binds again to the protein. Illumination can now initiate another reaction cycle.

Cones: The chromophore compound is likewise 11-*cis*-retinal. However, depending on the amino acid sequence of the different opsin components used, the three different classes of cone photoreceptors have absorption peaks at 420 nm ('blue cones'), 530 nm ('green cones'), and 560 nm ('red cones') (Fig. 7.4-11). Despite this nomenclature, the absorption spectra of the three types of cones overlap considerably. Therefore, the perception even of basic colors, like blue (~450–490 nm), is always the result of differential stimulation of the three different classes of cones by light of a given wavelength rather than the color being preferentially or even exclusively perceived by one class of photoreceptor. The reaction mechanisms following illumination are similar to the one in rods.

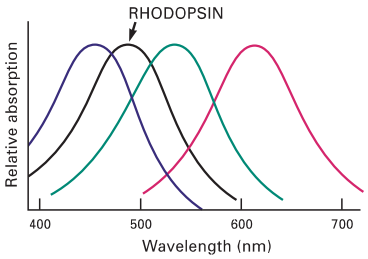


Figure 7.4-11. Absorption Spectra of Rhodopsin (in Rods) and of the Color Receptors (in Cones)

Rhodopsin is also used in the visual process of mollusks and arthropods. Higher plants, mosses and algae use the phytochrome system (with a tetrapyrrole chromophore) as light sensor for regulating metabolic processes. Phototaxis of bacteria and algae depends on retinal-related or flavin sensors.

Interconversion of 13-*cis*- and *all-trans*-retinal in bacteriorhodopsin takes place during proton pumping of halobacteria (archaea, Fig. 3.12-7), while higher plants and some bacteria use pyrrole derivatives for photosynthesis (Fig. 3.12-1).

Literature:

Hofmann K.P., *et al.* Trends Biochem. Sci. 2009;34:540–52.

Hubbell W.L., *et al.* Adv. Protein Chem. 2003;63:243–90.

Khorana, H.G. J. Biol. Chem. 1992;267:1–4.

Nathans, J. Neuron 1999;24:299–312.

Stryer, L. Ann. Rev. Neurosci. 1986;9:87–119.

Wald, G. Nature 1968;219:800–807 (Nobel Lecture).

7.4.7 Olfactory and Gustatory Processes

Olfactory process: Vertebrates sense odors by the olfactory epithelium of the nose. Cilia, which project from each sensory nerve cell into the nasal cavity, carry a multitude of different G-protein coupled olfactory receptors (GPCR) as well as receptor guanyl cyclases type D (GC-Ds, 7.8.1).

Contact of an odorant with olfactory GPCRs starts one of the following mechanisms:

- Adenylate cyclase is activated via $G_{olf\alpha}$ (Table 7.4-1). The cAMP produced causes opening of Na^+ ion channels and influx of Na^+ ions. This depolarizes the nerve cell and increases the frequency of action potentials.
- Phospholipase C is activated. Formation of inositol-P_3 (or inositol-P_4) opens Ca^{++} channels in the cellular membrane (7.4.4). The Ca^{++} influx hyperpolarizes the nerve cell and decreases the frequency of action potentials.

It is still unclear, how both mechanisms interact with each other and how different olfactory receptors respond to a particular odorant. Both mechanisms are effective only for short time intervals (roughly 50 msec). The receptor is desensitized by phosphorylation (7.4.1). The odorants are inactivated by degradation (mostly by P-450 catalyzed oxidation). The transmission of the action potentials to the brain is described in 7.2.2.

Insect olfactory receptors (ORs) lack sequence similarity to vertebrate GPCRs. Moreover, despite being also 7-TMRs (with 7 transmembrane helices), the topology of insect ORs is reversed compared to that of chemosensory GPCRs of vertebrates such that the amino terminus faces inside and the carboxy terminus outside of the cell. In contrast to the situation in mammals, where a given olfactory neuron expresses only one GPCR gene, functional insect ORs are heteromeric complexes that contain a variable subunit that binds the odorant and a constant Or83b co-receptor. The exact stoichiometry of the complexes is yet unknown. The Or83b co-receptor functions as a chaperone for the respective co-expressed odorant-binding subunit. These heteromeric complexes seem to act on the one hand as ligand-gated non-selective cation channels that trigger currents independent of GPCR-coupled second messenger pathways as initial response and on the other hand can apparently also mediate a prolonged cyclic-nucleotide-dependent response that peaks at 30–60 s. The molecular mechanism of hypersensitive recognition of specific odors by insects is incompletely understood, but perception of 11-*cis*-vaccenyl acetate, a male-specific lipid that acts as volatile pheromone in *Drosophila*, requires its binding by LUSH, an extracellular odorant-binding protein.

Literature on Olfaction:

Breer, H. (Ed.) Biology of Olfaction. Academic Press, 1994.

Farbman, A.I. Cell Biology of Olfaction. Cambridge University Press, 1992.

Liu, M., *et al*. Science 1994;266:1348–1354.

Nakagawa T., Vosshall L.B. Curr. Opin. Neurobiol. 2009;19: 284–292.

Touhara K., Vosshall L.B. Annu. Rev. Physiol. 2009;71:307–332.

Gustatory process: The gustatory perception of higher vertebrates is only able to differentiate between five basic tastes: sour, salty, bitter, sweet, and umami. The Japanese word umami means 'tasty' and refers to a taste quality that characterizes protein rich foods. Distinct nerve cells expressing unique membrane receptors for these five different taste qualities are present in the taste buds of the tongue. The umami receptors are activated by glutamate and this forms the molecular basis for the flavor enhancing action of sodium glutamate.

Ion channels mediate perception of the food qualities salty and sour. Salty tastes increase Na^+ flux through respective channels, while sour tastes cause a blockade of Na^+ or K^+ channels by H^+ ions. GPCRs of the T1R and T2R families transduce sweet, bitter, and umami tastes. Bitter taste causes a Ca^{++} release via the G-protein/PLC/IP_3 pathway (7.4.3), while sweet tastes increase the cellular cAMP level via adenylate cyclase (7.4.2). Gustducin, a heterotrimeric G-protein ($G_{g\alpha}$•$G_{\beta\gamma}$) of the G_i-family, is involved in signaling by receptors of the T1R and T2R families. The amino acid sequence of gustducin shows 80% identity with the transducin molecule of the visual process (7.4.6). In mice perception of sweet, bitter, and umami tastes is lost upon

knock-out of the genes for either gustducin or the phospholipase C isoform β2 or the ion channel TRPM5. Many other details of gustatory signal transduction are still unknown.

Literature on the Gustatory Process:

Chandrashekar, J. *et al.* Nature 2006;444:288–294.

McLaughlin, S.K. *et al*. Nature 1992;357:563–569.

Roper S.D. Pflügers Arch. 2007;454:759–776.

7.4.8 Arachidonate Metabolism and Eicosanoids

Arachidonic acid (*all-cis*-5,8,11,14-eicosatetraenoic acid, 1.4.1) belongs to the omega-6 class of polyunsaturated fatty acids (w-6 PUFA), because, counting from the non-carboxyl end of the molecule, the first of its four double bonds appears at the sixth carbon-carbon bond. The ω-6 as well as the ω-3 class of fatty acids are essential nutrients and have to be supplied from food sources, because the human body cannot synthesize them *de novo*. Furthermore, at the adult stage, humans have only a limited capacity to synthesize the longer chain variants of these fatty acids from their respective C_{18} precursors (linoleic acid in case of ω-6 PUFA and α-linolenic acid in case of ω-3 PUFA). The two biosynthetic pathways compete with each other particularly for the required desaturase activities.

There is good evidence now that the longer chain ω-3 fatty acids eicosapentaenoic acid (EPA; C_{20}) and docosahexaenoic acid (DHA; C_{22}), for which oily fish are a rich nutritional source, protect against heart failure. Moreover resolvins and protectins, the bioactive signaling molecules generated from EPA and DHA, generally exert anti-inflammatory effects. In contrast to this most, though not all, of the lipid mediators derived from arachidonic acid have pro-inflammatory functions (Tables 7.4-7 and 7.4-8). In all mammalian cells except erythrocytes, arachidonic acid is converted to the eicosanoide group of compounds, which exert important hormone-like effects.

Although these compounds are effective in very low concentrations, due to their short lifetime they can act only on the originating cell (autocrine) or on neighboring cells (paracrine). Thus, they are local mediators.

Release of arachidonic acid (Fig. 7.4-12): Located at the 2-position of glycerol, arachidonic acid is present in the membrane phospholipids phosphatidyl-inositol (7.4.3), -ethanolamine and -choline (3.4.3.1). It is released by different hormone-controlled phospholipases:

- The cytoplasmic phospholipase A_2 (cPLA$_2$, 85 kDa) is activated by a rise in the intracellular Ca^{++} level (7.4.4) and by MAPK-catalyzed phosphorylation (7.5.3) in response to stimulation by e.g., bradykinin (9.2.2) or angiotensin II (7.1.8).
- The non-pancreatic, secretory phospholipases A_2 (sPLA$_2$, 13 … 15 kDa) are secreted following a hormone stimulus (e.g. by TNF, IL-1, IL-6). sPLA$_2$ isoforms are activated upon secretion by the higher level of Ca^{++} in the extracellular milieu.
- Phospholipases C and D (3.4.3.2) can also release arachidonate.

Table 7.4-7. Series-2 Prostanoids: Prostaglandins, Prostacyclins and Thromboxanes[1]

Name[1]		Receptor	Receptor Occurrence	Transduction System	Major Metabolic Effects in humans
Prostaglandin G_2	PGG_2				(unstable)
Prostaglandin H_2	PGH_2				(unstable)
Prostaglandin D_2	PGD_2	DP	platelets, vascular smooth muscle, nerves	$G_{s\alpha} \rightarrow$ cAMP ↑	Inhibition of platelet aggregation, relaxation of smooth muscle, renal vasodilatation, increased water reabsorption in the small intestine
Prostaglandin E_2	PGE_2	EP_1	rodent smooth muscle, kidney, lung	Ca^{++} ↑ (not via IP_3?)	Inhibition of platelet aggregation
		EP_2	smooth muscle, thymus, lung, spleen, heart, uterus	$G_{s\alpha} \rightarrow$ cAMP ↑	Potentiation of platelet aggregation, relaxation of smooth muscle, vasodilatation, secretion by epithelial cells, inhibition of mediator release by inflammatory cells, activation of sensory nerves
		EP_3 (splice variants)	ubiquitous: smooth muscle, adipocytes, kidney, uterus, thymus, spleen, lung, brain	a) $G_{i\alpha} \rightarrow$ cAMP ↓	Inhibition of autonomic neurotransmitter release, of lipolysis in adipocytes, of water reabsorption in renal medulla, suppression of gastric acid secretion
				b) $G_{q\alpha} \rightarrow IP_3$↑	contraction of smooth muscle, potentiation of platelet aggregation
		EP_4	veins	$G_{s\alpha} \rightarrow$ cAMP ?	?
Prostaglandin $F_{2\alpha}$	$PGF_{2\alpha}$	FP	corpus luteum, smooth muscle, kidney, lung etc.	$G_{q\alpha} \rightarrow PI_3$	Increase of uterus muscle tone, contraction of bronchial smooth muscle, inhibition of water reabsorption in the small intestine
Prostacyclin I_2	PGI_2	IP	arterial endothelium, platelets, nerves	$G_{s\alpha} \rightarrow$ cAMP ↑	Vasodilatation, inhibition of platelet aggregation, increase of water reabsorption in the small intestine
Thromboxane A_2	TXA_2	TP	vascular tissue, platelets, bronchial muscle, thymus	$G_{q\alpha} \rightarrow PI_3$↑	Vasoconstriction, induction of platelet aggregation and release reactions, strong contraction of bronchial smooth muscle (asthma)

[1] The series and the index numbers indicate the number of double bonds in the molecule. The series-1 prostanoids/ thromboxanes are synthesized analogously from 8,11,14-eicosatrienoic acid (dihomo-γ-linolenic acid, DGLA), the series-3 prostanoids/thromboxanes from 5,8,11,14,17-eicosapentaenoic acid (EPA).

Table 7.4-8. Series-4 Leukotrienes and Precursors[1]

Name[1]		Receptor	Receptor Occurrence	Major Metabolic Effects in *humans*
5-hydroperoxyeicosatetraenoic acid	5-HPETE		leukocytes, mast cells, lung, spleen, brain, heart	Regulation of ion channels, histamine release, insulin and renin release
Leukotriene A_4	LTA_4		neutrophils, macrophages, monocytes, mast cells	(unstable)
Leukotriene B_4	LTB_4	LB_4	wide distribution	Contraction of vascular and intestinal smooth muscles, chemotaxis and adhesion of granulocytes, immunomodulation, liberation of lysosomal enzymes, initiation of inflammatory reactions, cAMP ↑
(Peptido-) Leukotriene C_4	LTC_4	LD_4 ? LC_4 ?	granulocytes, platelets, monocytes, mast cells, lung	SRS-A,[2] bronchoconstriction (immediate hypersensitivity, asthma), mucus secretion, inflammatory reactions, increase of vascular permeability
(Peptido-) Leukotriene D_4	LTD_4	LD_4	granulocytes, platelets, monocytes, mast cells, lung	same
(Peptido-) Leukotriene E_4	LTE_4	LD_4 ?	granulocytes, lung	Same

[1] The series and the index numbers indicate the number of double bonds in the molecule. The series-3 leukotrienes are synthesized analogously from 8,11,14-eicosatrienoic acid (dihomo-γlinolenic acid, DGLA), the series-5 leukotrienes from 5,8,11,14,17-eicosapentaenoic acid (EPA). They are less inflammatory than the compounds originating from arachidonic acid.

[2] Member of the slow reacting substances of anaphylaxis (SRS-A), acting at ca. 10^{-10} mol/l.

Biosynthesis of prostanoids by the cyclooxygenase pathway (Fig. 7.4-14, lower right): Prostaglandin H_2 synthase (PGH_2 synthase, 72 kDa) is a dimeric, heme-containing enzyme that produces the cyclic compound PGH_2. This enzyme is firmly anchored to the luminal side of the endoplasmic reticulum membrane and to the outer membrane of the nuclear envelope. It has dual enzymatic activity as a cyclooxygenase (→ PGG_2, Fig. 7.4-13 left) and as a peroxidase ($PGG_2 \rightarrow PGH_2$).

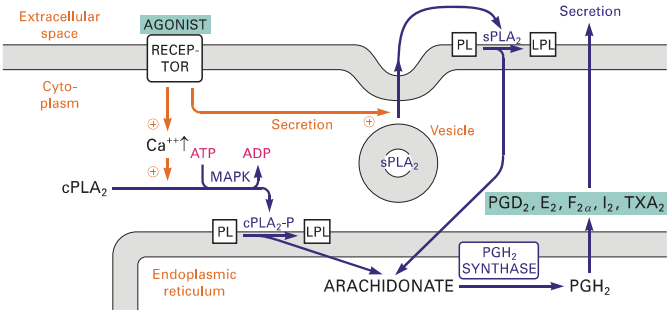


Figure 7.4-12. Liberation and Metabolism of Arachidonic acid (PL = phospholipid, LPL = lysophospholipid, MAPK = mitogen activated protein kinase)

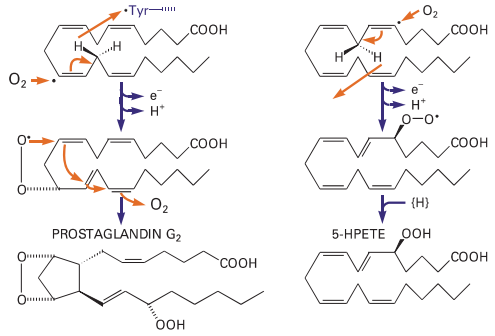


Figure 7.4-13. Conversion of Arachidonic Acid by the Cyclooxygenase Reaction (Left) and the Lipoxygenase Reaction (Right)

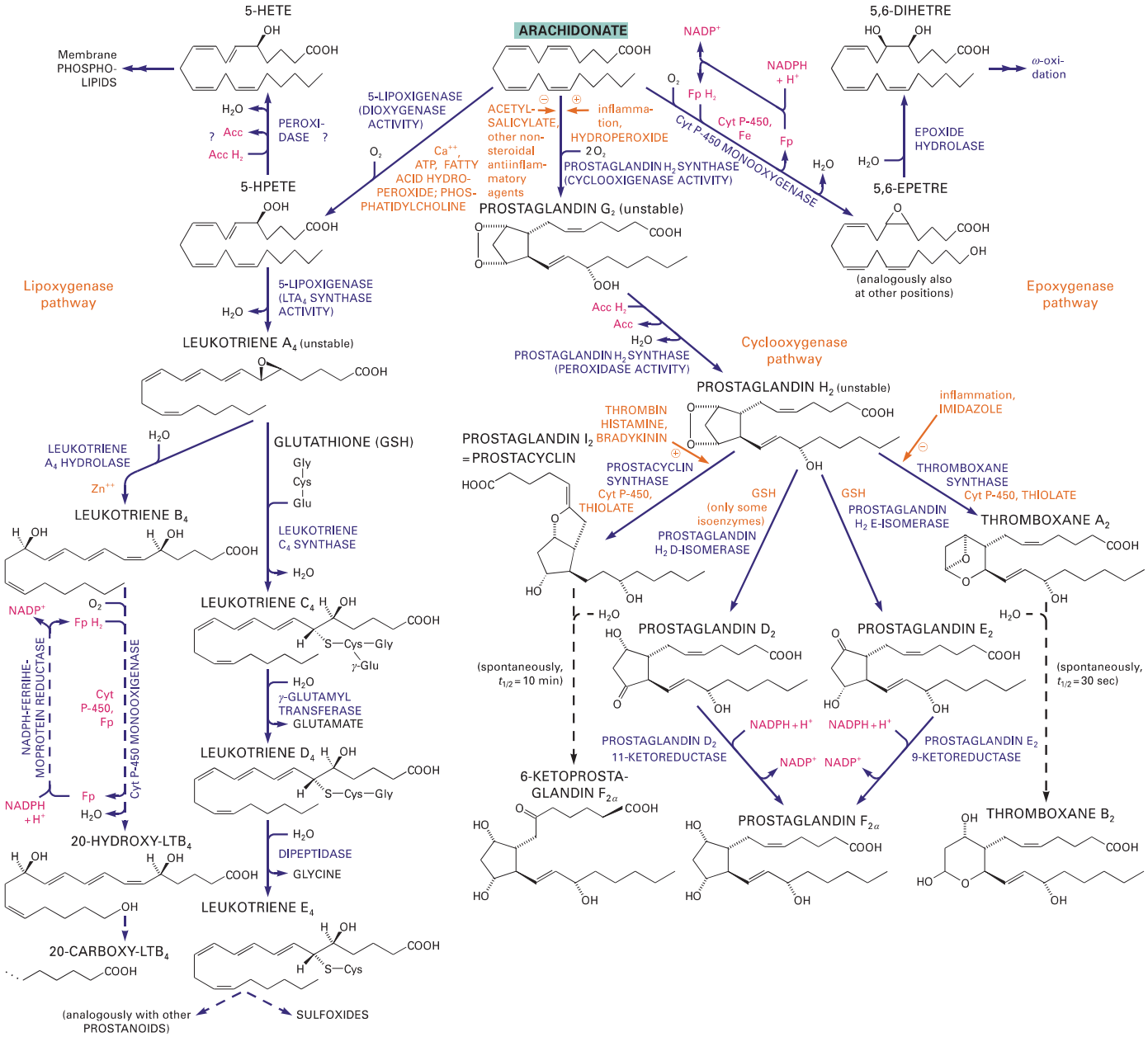


Figure 7.4-14. Biosynthesis of Eicosanoids

The cyclooxygenase reaction starts with abstraction of the 13-*pro-S* hydrogen, followed by attack of an oxygen molecule, rearrangements resulting in ring closure, and finally reaction of a second oxygen molecule with the resulting radical. There are two isoenzymes, of which only one is inducible. The activity of the enzyme is inhibited by steroidal and nonsteroidal anti-inflammatory drugs. Acetylsalicylic acid (Aspirin®) acetylates a serine residue in the active center, while indometacin, phenylbutazone, acetaminophen etc. compete with the substrate.

PGH_2 is the precursor of all series-2 prostaglandins (with two double bonds in the side chains), which are synthesized on the cytoplasmic side of the ER membrane. In cells of the vascular endothelium, PGH_2 is converted by prostacyclin synthase to prostacyclin I_2 (PGI_2) and in platelets and lung by thromboxane synthase to thromboxane A_2 (TXA_2), which exert opposite effects on the circulation (Table 7.4-7). Other prostanoids are formed by isomerization or reduction reactions. The prostanoids leave the cell via a carrier-mediated transport mechanism.

Prostanoids mediate their physiological effects via specific receptors at the cell surface, which have the usual seven transmembrane helix structure (Fig. 7.2-5) and are coupled to G-proteins (7.4.1).

The receptor mediated reactions are complex and can vary in different parts of the same organ and depending on different physiological conditions. There is also considerable variation of effects in different species.

Biosynthesis of leukotrienes (LT) by the lipoxygenase pathway (Fig. 7.4-14, left): Leukotrienes are important mediators of sustained inflammatory reactions. They play a key role in asthmatic and allergic reactions, since LTs also affect bronchoconstriction and increase vascular permeability. LTs are synthesized in granular leukocytes, monocytes and mast cells by the lipoxygenase pathway.

The bifunctional enzyme 5-lipoxygenase (5-LO) converts arachidonic acid first to 5(S)-hydroperoxy-6,8,11,14-eicosatetraenoic acid (5-HPETE, Fig. 7.4-13 right) by its dioxygenase activity and then to a highly unstable epoxy product, leukotriene A_4 (LTA_4), by its LTA_4 synthase activity. LTA_4 is the precursor of different other leukotrienes. It can be hydrolyzed by LTA_4 hydrolase to yield the dihydroxy compound LTB_4. The glutathione transferase activity of LTC_4 synthase converts LTA_4 into LTC_4. After removal of the terminal γ-glutamyl residue of LTC_4 by a transferase, LTD_4 is obtained.

5-LO contains a non-heme, non-FeS iron in its catalytic center. It is activated by Ca^{++} and is not inhibited by nonsteroidal anti-inflammatory drugs. '5-lipoxygenase-activating protein' (FLAP, 18 kDa) binds free arachidonic acid and presents it to 5-LO.

Alternatively to LTA_4 formation, the intermediate 5-HPETE can be converted by a two-electron reduction reaction into 5-hydroxy-6,8,11,14-eicosatetraenoic acid (5-HETE).

Other lipoxygenase family members are 8-lipoxygenases (in skin), 12-lipoxygenases (mostly in platelets) and 15-lipoxygenases (in granulocytes and epithelial cells etc). The products of these enzymes have only modest biological effects.

Epoxygenase P-450 pathway (Fig. 7.4-14, upper right): This term comprises all reactions which introduce a single oxygen molecule into arachidonic acid metabolites catalyzed by the various cytochrome P-450 monooxygenases (for mechanism compare 3.5.4.1).

5,6-, 8,9-, 11,12- or 14,15-epoxidation reactions yield various *cis*-epoxy-eicosatrienoic acids (EPETRES) besides monohydroxy-eicosatetraenoic acids (HETEs, which are also products of the lipoxygenase pathway, see above). This takes place in liver, kidney, lung, eye and other organs. These epoxy compounds are quickly converted by hydration to *vic*-diols (dihydroxyeicosatrienoic acids, DIHETREs). Some of the physiological effects of these nonclassical eicosanoids include peptide hormone release from endocrine cells, modulation of Ca^{++} release, vasodilatation, and inhibition of Na^+/K^+ ATPase activity.

P-450 dependent hydroxylation reactions of many eicosanoids are also part of their degradation pathways (e.g., in liver, Fig. 7.4-14, lower left).

Literature:

Brock T.G., Peters-Golden M. Scientific World Journal 2007;7: 1273–1284.

Coleman, R.A., *et al.* Pharmacol. Reviews 1994;46:206–229.

Harizi H., Corcuff J.B., Gualde N. Trends Mol. Med. 2008;14: 461–469.

Radmark O., *et al*. Trends Biochem. Sci. 2007;32:332–41.

Ratnayake, W.M.N., Galli, C. Ann Nutr Metab 2009;55:8–43.

Smith, W.L., Fitzpatrick, F.A. in Vance, D.E., Vance, J. (Eds.) *Biochemistry of Lipids, Lipoproteins and Membranes. New Comprehensive Biochemistry* Vol. 31. Elsevier, 1996; p. 283–339.

Waite, M. ibid., p. 211–236.

7.5 Receptors Acting Through Tyrosine Kinases

7.5.1 Regulatory Factors for Cell Growth and Function

Cell growth and differentiation in multicellular animals (metazoans) is regulated by a large number of extracellular polypeptide hormones (e.g., growth hormones, cytokines, 8.1.8), but also by steroids (dealt with in 7.7). Among the polypeptides (terminology of the human system) are:

- competence factors for transition from the resting phase G_0 to the G_1 phase of the cell cycle (4.3) or for bypassing G_0 altogether (e.g., EGF, TGF-α, FGF, PDGF and IL-2).
- progression factors for passage through the restriction (commitment) point in the G_1 phase of the cell cycle (4.3.2, 4.3.3) and initiation of DNA replication and cellular division (e.g., IGF-1 or insulin in high concentrations).
- antagonists of cell cycle progression, e.g., transforming growth factor-β (TGF-β), interferons or tumor necrosis factor α (TNFα).
- Compounds, which additionally cause different effects [e.g., insulin, regulating glucose levels (3.1.1.2 … 4) or cytokines, controlling chemotaxis (Fig. 8.1-19)].

The regulation of the cell cycle is described in 4.3.2 … 4.3.6. The discussion below focuses on the mechanism of signal transmission from these extracellular factors to the nucleus, where they primarily influence gene transcription (4.2.1).

These polypeptides bind as ligands to receptors that then start a signal cascade via protein phosphorylation by tyrosine kinases, finally leading to the nuclear effects on gene expression. This is, however, not a simple input-output sequence. Rather, there are manifold interactions of different members of the signal cascades (7.5.2), causing amplification, modulation and integration of external signals. A complex interplay of phosphorylation and dephosphorylation reactions is involved in the regulation of these signaling networks.

The downstream signal transduction mechanisms of T and B cell receptors are similar to those of other receptors dealt with here. For this reason, they are included in this section, although in contrast to other receptors they are activated by peptide complexes presented on the surface of antigen-presenting cells rather than by soluble hormones/cytokines (8.1.7, 8.1.8).

Oncogenes: Mutations in the genes for growth hormones, their receptors or downstream effectors of their signaling cascade can lead to malfunction of the respective gene products resulting in loss of proliferation control and possibly tumorgenesis. The pathologically modified, tumor-causing genes are therefore called oncogenes. They can be of cellular (c-oncogenes) or viral (v-oncogenes) origin. Their normal counterparts encoding unmodified, physiologically functioning gene products are named proto-oncogenes. Genes encoding regulatory proteins that counteract tumorgenesis are referred to as anti-oncogenes or tumor suppressor genes.

7.5.2 Components of the Signal Cascades

The essential components of the signal cascades (in vertebrates, homologues exist in other eukarya including *Drosophila* and yeast) are:

- membrane associated receptors, which contain an extracellular ligand binding domain, a transmembrane domain, and either an intracellular tyrosine kinase domain or a binding site for a cytosolic tyrosine kinase (Tables 7.5-1 or 7.5-3, respectively)
- protein tyrosine kinases (Table 7.5-3)
- protein serine/threonine kinases
- mixed lineage kinases (MLKs) phosphorylating tyrosine and serine/threonine residues
- protein phosphatases, counteracting the kinases (Table 7.5-5)
- small G-proteins (Table 7.5-2, different from heterotrimeric G-proteins, 7.4)
- GDP/GTP exchange and guanosine nucleotide releasing factors (Table 7.5-2)
- transcription factors (Table 7.5-6)

Each of these protein categories is composed of several families with many members. The multitude of these factors and their complex interaction patterns allow for a vast array of different cellular responses to extracellular signals. However, some general principles of signal transduction exist. The activities of protein kinases (518 human kinases) are counteracted by a similarly large number of cytosolic and transmembrane phosphatases (total of 555 in 5 sequenced eukaryotic genomes). As with kinases, there are phosphatase subfamilies specific for dephosphorylation of either only tyrosine (PTPases) or only serine/threonine residues or both (so-called dual phosphatases). The role of the transmembrane PTPase CD45 in B- and T-cell receptor signaling is particularly well understood (7.5.4).

7.5.3 Receptor Tyrosine Kinases (RTK, Table 7.5-1)

Receptor structure (Fig. 7.5-1): The receptors are usually monomers. After ligand binding, the receptors of vertebrates dimerize and undergo a conformation change, which enables the tyrosine kinase moieties of the two receptor molecules in each dimer to phosphorylate the intracellular portion of their adjacent partner at several key tyrosine residues (receptor autophosphorylation in *trans*). The different phosphorylated tyrosine sequence motifs generated in this way function as specific binding sites for the Src homology 2 (SH2) domains of various signalling proteins. Upon recruitment, these signaling molecules themselves become substrates for phosphorylation by the receptors resulting either in activation of their catalytic function (enzymes) or in formation of further SH2-binding sites (linker or scaffold proteins). In this way, large multimeric signaling complexes can be formed around activated receptor dimers. Examples for components of such signalling complexes are the linker protein Grb and GAP, a GTPase activating protein for the small G-protein Ras (see below), PLC-γ1 (7.4.3, connecting to the PIP_2 pathway), phosphatidylinositol 3-kinase (7.4.4), the protein-tyrosine kinase Src (Table 7.5-5), protein-tyrosine phosphatases (e.g., Shp1, Syp) and the SH-2 domain containing linker protein Shc (Table 7.5-2). In different tissues, the

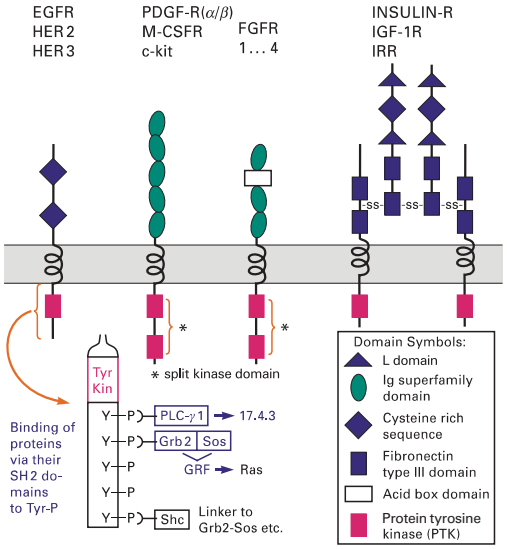


Figure 7.5-1. Structures of Some Receptor Tyrosine Kinases

Table 7.5-1. Vertebrate Receptor Families With Integral Protein-Tyrosine Kinase Activity (RTK). For Structures, see Figure 7.5-1

Important Receptor Families (14 known)	Examples for Members (R = receptor)	Ligands (GF = growth factor)	Targets of action (direct and indirect) – Effects of activated receptors	Oncogenes related to receptor or ligand genes
Epidermal growth factor (EGF) receptor family	EGFR (170 kDa), HER2 = Neu = ErbB2, HER3 = ErbB3; ErbB4	EGF, TGF-α, heparin-binding EGF-like GF	EGFR: Grb2 subunit of GRF (7.5.3) → Ras, Shc, Src, PLC-γ1 (7.4.3). – Mitogenic effects, influences Ca^{++} metabolism, activates transcription, cell division.	c-erbB, v-erbB encode truncated EGFR (without ligand binding site)
Insulin receptor family	insulin R (IR), insulin-like growth factor-1 R (IGF-1R), insulin relat.R (IRR)	insulin IGF-1 ?	IR: IRS-1 → Raf, PI 3-kinase, possibly others. – Induces expression of glycolytic and represses expression of gluconeogenetic enzymes. – Antagonizes cAMP and glucocorticoid effects. General anabolic activity. See 3.1.1.5.	
Platelet derived growth factor (PDGF)/macrophage colony stimulating factor (M-CSF) receptor family	PDGFR (α/β)	PDGF	Ras, Src, Shc, PLC-γ1, PI 3-kinase. – Mitogenic in smooth muscle cells, glial cells, fibrobl.	c-sis, v-sis encode homologue of PDGF
	M-CSFR	M-CSF	Src, PI 3-kinase. – Differentiation of macrophages	c-fms, v-fms encode homol. of M-CSFR
	c-kit (Steel R)	stem cell factor	PI 3-kinase – Melanocyte and hematopoiesis	
Fibroblast growth factor (FGF) receptor family	FGFR 1, 2, 3, 4 (4 genes + alternative splicing)	FGF (acidic, basic, keratinocyte GF, FGF4 … 9)	Raf-1 → ERK1+2, PLC-γ, S6 ribosomal kinase – Growth, differentiation (mesoderm formation), survival of cells, pH ↑, Ca^{++}↑ , PI turnover ↑	c-int 2 encodes homologue of FGF
Vascular endothelial cell growth factor (VEGF) receptor family	VEGFR/Fit 1 KDR/ Flk	VEGF	Mitogenic in endothelial cells	
Hepatocyte growth factor (HGF) receptor family	HGFR (extracellular + transmembrane subunits)	HGF/SF (scatter factor)	GAP, Src-related kinases, PLC-γ, PI 3-kinase. – Mitogenic in epithelial cells, stimulation of cell motility	
Neurotropin receptor family	TrkA, TrkB, TrkC	Neurotropins (NGF,NT-3, -4, -5, BDNF)	Raf, PLC-γ, S6 ribosomal kinase, Tyr hydroxylase. – Growth, differentiation, survival of neurons.	
	LNGFR		no enzymatic function. Role unknown.	
Eph-like receptor tyrosine kinases (largest family)	Eph, Elk, Eck, Eek, Er, Cek4/ HEK, Cek5	?	Cell adhesion?	
Axl receptor tyrosine kinase	Axl	?	Insulin receptor related	

same agonist/receptor pair can cause very different biological effects depending on the availability and functional interconnection of such signal transducers.

The insulin receptor molecule is composed of two disulfide bridged polypeptide chains: the entirely extracellular α chain that mediates ligand binding and the transmembrane β chain encompassing the kinase domain and the intracellular signaling capacity of the receptor. Contrary to other RTKs, the mature insulin receptor forms disulfide-linked α/β homodimers even in the absence of its ligand (Fig. 7.5-1). Recent crystallography data show that the insulin molecule has two different binding sites for its receptor. High affinity binding of insulin involves its cross-linking of two partial binding sites, one from each receptor monomer. This induces a conformational change of the dimeric receptor, which initiates transphosphorylation of several tyrosine residues per β-chain. These phosphotyrosine residues (P-Tyr) are docking sites for a family of large scaffolding proteins, the insulin receptor substrates 1–4 (IRS-1, see below).

Phosphorylation of Ser and Thr residues of the insulin receptor (e.g., by protein kinase A) decreases its tyrosine kinase activity. In insulin resistance and type II diabetes, serine/threonine phosphorylation of the insulin receptor is often abnormally increased, which renders the receptor less sensitive to activation by insulin and impairs its signal transduction capacity (see 7.1.3).

Two examples of receptor tyrosine kinase triggered signaling cascades are presented in some detail:

EGF receptor activation of the MAPK pathway (Figs. 7.5-2 and 7.5-4): The Ras/Raf/MAPK cascade is part of the signal transduction of many RTKs and plays a crucial role in regulating proliferation and survival of cells. Upon EGF stimulation a preformed complex of the adapter protein Grb2 and the GDP/GTP exchange factor Sos1 can be recruited to the EGF receptor either by direct binding of the SH2 domain of Grb2 to receptor autophosphorylation sites or by using Shc as adapter molecule. Shc can bind via its PTB domain to another autophosphorylation site of the EGF receptor. Subsequent phosphorylation of SHC by the EGFR provides three alternative docking sites for SH2 domains of Grb2. Once recruited to the membrane, the Grb2/Sos1 complex can convert Ras from its inactive GDP-bound to an active GTP-bound state by virtue of the GDP release activity of Sos1. Ras•GTP then binds with high affinity to the autoinhibitory domain of the serine/threonine kinase Raf. The resulting release of autoinhibition in conjunction with its recruitment to the plasma membrane (where it can be phosphorylated by a protein kinase) results in stable Raf activation. Activated Raf then phosphorylates and thereby activates the dual specificity kinases MEK1/2, which in turn activate by phosphorylation the so-called MAP kinases ERK1 and 2. Activated MAP kinases translocate to the nucleus, where they phosphorylate transcription factors of the Ets family (e.g., Elk1) and other substrates thereby causing changes in gene expression (4.2.2).

Ras returns to the inactive state by hydrolysis of the bound GTP to GDP. The intrinsic GTPase activity of Ras is accelerated by specific GTPase activating proteins (RasGAPs). The downstream components of the described kinase cascade are inactivated by membrane bound or soluble protein phosphatases, which are themselves regulated.

The protein serine/threonine phosphatases involved belong to the PPP family (Fe^{++}/ Mn^{++} or Zn^{++} dependent) or to the PPM family (Mg^{++}/Mn^{++} dependent). The dual specificity kinases MEK1/2 are counteracted by dual specificity (Thr-P, Tyr-P) phosphatases of the PP2A or PP2C type.

Insulin receptor activation of the PI 3K pathway (Figs. 7.5-3 and 7.5-4): A peculiarity of the insulin and IGF1 receptor is that they mainly use large scaffolding proteins of the insulin receptor substrate family (IRS-1 … 4) for assembling an intracellular signalling complex. IRS-1, one of the main members of this family, binds to phosphorylated Tyr residues of the insulin receptor via its PTB domain. After its recruitment, IRS-1 becomes a substrate for the tyrosine kinase activity of the receptor. The tyrosine phosphorylation sites of IRS-1 in turn act as docking sites for intracellular effector molecules, like the phosphatase SHP2, which can further recruit Grb2. However, most of the phosphotyrosine residues of IRS-1 provide docking sites for the regulatory p85 subunit of PI 3-kinase (see below). Activated PI 3-kinase triggers plasma membrane recruitment and activation of Akt (=PKB), a serine/threonine kinase, which is involved in initiating translocation of the glucose transporter GLUT4 from endosomal compartments to the plasma membrane. In addition, the Akt pathway is also involved in blocking apoptosis and in increasing protein synthesis via FRAP/mTOR. Uptake of glucose by the GLUT4 transporter in conjunction with phosphorylation events downstream of Akt lead to induction of glycolytic enzymes and repression of gluconeogenetic enzymes. The opposite effects are exerted by cAMP (7.4.2) and by glucocorticoids (3.5.8.3, 7.7). For details, see 3.1.1.3 and 3.1.2.4. Additionally, insulin

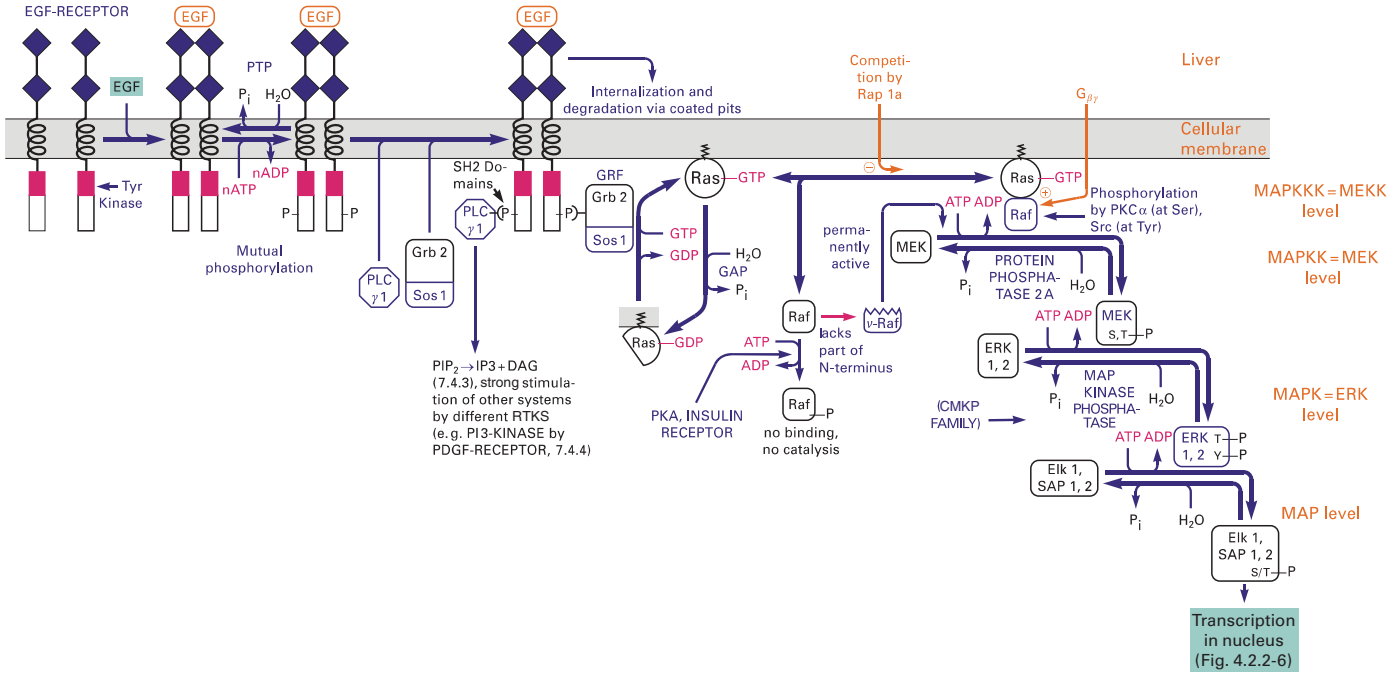


Figure 7.5-2. EGF Receptor Cascade (Terminology of the human system)

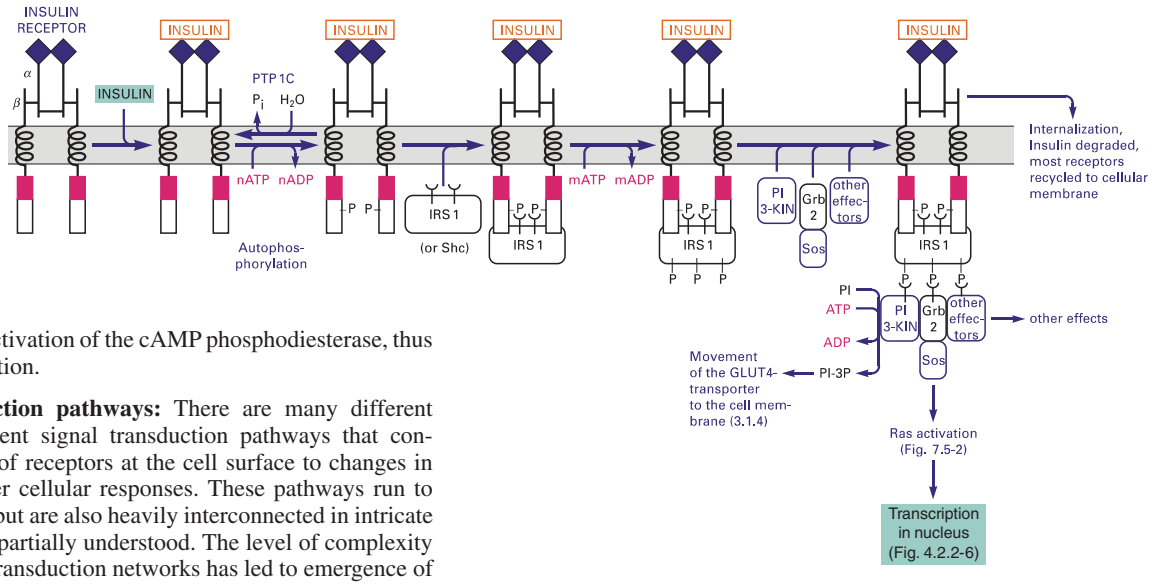


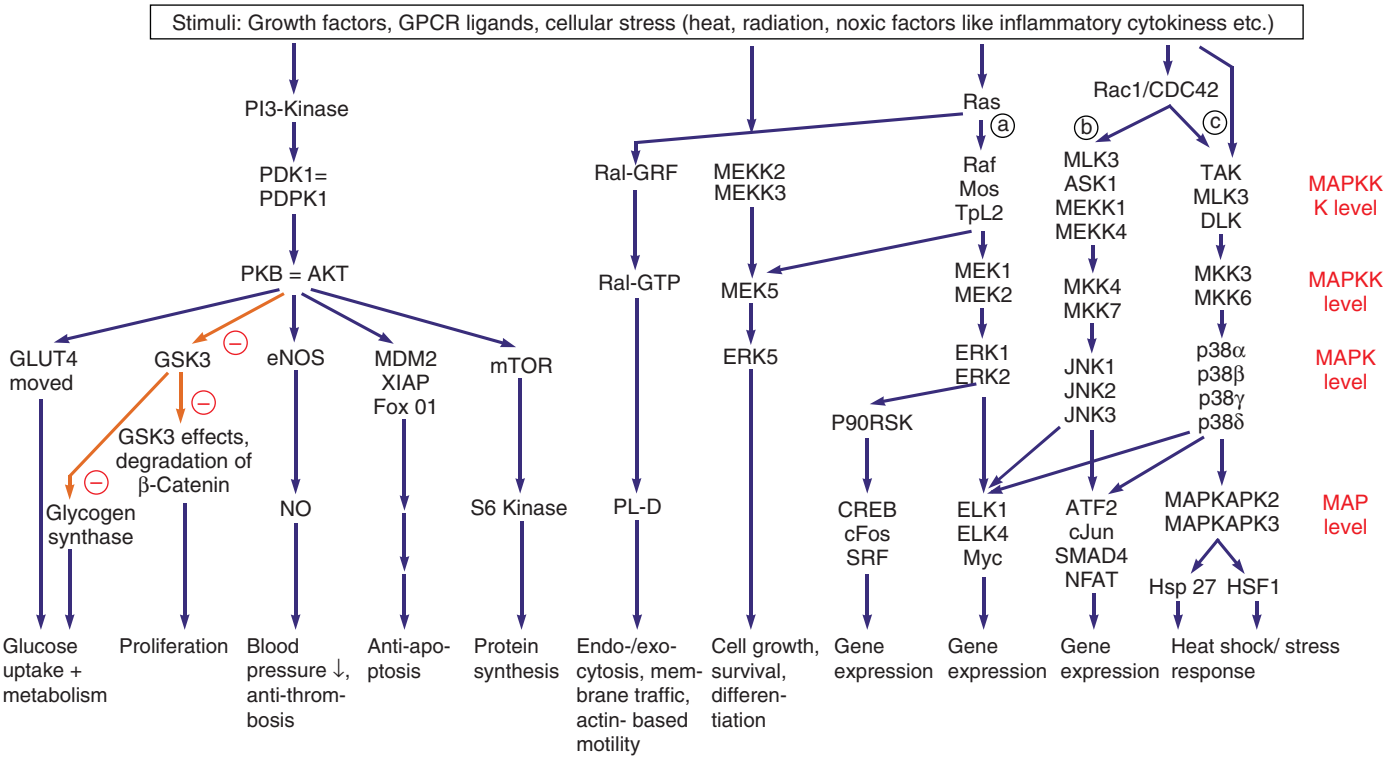
Figure 7.5-3. Insulin Receptor Cascade
(Terminology of the Human System)

antagonizes cAMP by activation of the cAMP phosphodiesterase, thus decreasing its concentration.

Other signal transduction pathways: There are many different phosphorylation-dependent signal transduction pathways that connect ligand stimulation of receptors at the cell surface to changes in gene expression or other cellular responses. These pathways run to some extent in parallel, but are also heavily interconnected in intricate ways, that are still only partially understood. The level of complexity of these cellular signal transduction networks has led to emergence of the field of systems biology as a discipline that tries to mathematically model such regulatory circuits. Figure 7.5-4 shows schematically some key signal transduction pathways, like the one governed by Ras. The different levels of cross-talk between these signaling cascades are not presented in this Figure. It has to be mentioned, that frequently several synonyma are used for the same compound.

Crosstalk with other signaling systems (Fig. 7.5-4): RTKs cross-talk to other signaling systems by affecting some of the same effector molecules. For instance, many receptor tyrosine kinases activate the phospholipase Cγ1 isoenzyme, thereby influencing pathways regulated by phosphoinositol and PKC (7.4.3, 7.4.4, Fig. 7.5-2). Many RTKs also activate phosphatidylinositol 3-kinase (PI 3-K, 7.4.4), which mediates mitogenic signals, but is also involved in vesicle trafficking, cell adherence and cell survival. As already mentioned, Ras governs the Raf/MAPK pathway, which is key for proliferation responses. Furthermore, Ras•GTP can bind to and activate a Ral-guanine nucleotide release factor, thereby promoting activation of Ral and, in turn, of phospholipase D, which initiates another sequence of lipid messenger signalling (Fig. 7.5-4).

Vice versa, RTK signalling is also affected by cross-talk from other signalling systems. For instance, Ras can be activated by G-proteins (7.4.1) and Raf can be activated by PKC-catalyzed phosphorylation at multiple Ser and Thr residues.



GLUT4 = Glucose transporter 4, Gsk3 = Glycogen synthase kinase, PKB /AKT = Proteinkinase B, mTOR = Mammalian target of rapamycin, S6 Kinase phosphorylates ribosomal protein S6. For other names see text. Some are derived from the gene names

Figure 7.5-4. Sequences in Vertebrate Ras Signalling Cascades

Ras related small, monomeric G proteins (Table 7.5-2): The members of this superfamily resemble the heterotrimeric G-proteins (7.4), since hydrolysis of bound GTP inactivates these proteins, while replacement of GDP by GTP transforms them into an active state. Individual families are named after key members and include the large Ras family (containing Ras, Rap, Ral, Rheb, Rin, and Rit), the Rac/Rho family (cytoskeletal effects), the Ran family (transport into the nucleus, 4.5.3), the Arf and Rab families (vesicle transport, secretion, and endocytosis, 4.5.2.1). The cycling between the GDP-bound 'off state' and the GTP-bound 'on state' of small G proteins is commonly regulated by specific interacting proteins of the GAP (GTPase-activating protein) and GEF (guanine nucleotide exchange factor) families. However, recently a subclass has been identified that is regulated by homodimerization. The signal recognition particle (SRP), septins, and dynamin belong to this category of G proteins activated by nucleotide-dependent dimerization (GADs).

7.5.4. Tyrosine Kinase-Associated Receptors (TKaR)

This group encompasses mostly cytokine receptors (8.1.8) and B and T cell receptors (8.1.6), all of which lack intrinsic enzyme activity but associate with cytosolic tyrosine kinases upon ligand stimulation. In contrast, the receptors for the cytokines colony-stimulating factor 1 (CSF1) and stem cell factor/kit-ligand (SCF/KL) belong to the RTK family (designated CSF1R and c-Kit, respectively).

Cytokine receptors (see Fig. 8.1-19): In many cases, these monomeric receptors undergo di- or oligomerization after ligand binding. This can occur by binding of dimeric ligands (e.g. Interleukin-10) or by binding of monomeric ligands with two binding sites (e.g., erythropoietin).

The type I cytokine receptor family (CRF1) consists of four subfamilies:

- The IL-2 receptor family (receptors for IL-2, IL-4, IL-7, IL-9, and IL-15) are heterotrimers of specific α and β subunits, and a common γ subunit. Their respective ligands bind to the specific α subunit, while the α,β, and γ subunits transduce the signal into the cell.
- The gp140 family (e.g., receptors for IL-3, IL-5, and granulocyte/macrophage colony stimulating factor). Like the gp130 family,

Table 7.5-2. "Small" G-proteins, GDP/GTP Exchange Factors and Linker Proteins

Name	Properties	Action
Ras (p21, 3 isotypes)	Well conserved from yeast to mammals. Membrane-anchored by prenylation. Proto-oncogene. v-ras or c-ras oncogene products hydrolyze GTP at a slower rate	Central role in RTK cascades. Activated by binding of GTP (effected by GRF). Activates Raf and other downstream effectors. GTPase function, enhanced by GAP
Rap1a	Member of Ras family	GTPase, acting as antagonist to Ras in Raf binding.
Rac, Rho, Ran, Arf, Rab, other Ras-related small G-proteins (ca. 21 kDa).		
GRF (GEF, GNRP)	Guanine nucleotide releasing factor (GDP/GTP exchange factor). Subunits: Grb2 and Sos, associated by SH3 domains.	Activated by EGFR, PDGFR, CSF-1R. After binding to these receptors it exchanges GDP with GTP at Ras. Similar factors act on Ras-related proteins
GAP	GTPase activation protein	Enhances hydrolysis of Ras-bound GTP. Similar GAPs exist for other proteins
Shc	No catalytic domain. 52 or 46 kDa (different translation initiation).	Binds to and becomes phosphorylated by activated RTK or by tyrosine kinases associated with receptors or via a G pathway. Linker function to GRF.
IRS1	No catalytic domain.	Gets highly phosphorylated by the activated insulin receptor, linker function

Table 7.5-3. 'Downstream' Protein Kinases, Counteracting Phosphatases and Transcription Factors

Level (K = kinase)	[1]	Abbreviation	Type, Properties	Phosphorylates and Activates	Counteracting Phosphatases
MAPKKK (S/T)	a	Raf	Activated by direct interaction with Ras•GTP (RTK activated) and with PKC (phorbol ester activated). Inactivated by PKA-catalyzed phosphorylation	MEK1,2	PP1, PP2A ?
	b	MEKK1,4	MEK kinase. Activated by direct interaction with Ras•GTP (which was activated by stress, heat shock, UV irradiation, translational inhibition, cytokines, ceramide etc.). Also activated by anisomycin, okadaic acid (phosphatase inhibitor), phorbol esters	MKK4,7	
	c	TAK, MLK3, DLK	Thylakoid membr. Protein kinase, mixed lineage kinase, delta- like protein	MKK3,6, MKK 4,7	
		ASK1	Apoptosis signal-regulating kinase. Activated by osmolarity changes, heat shock, lipopolysaccharides (?)		
MAPKK (T/Y)	a	MEK1,2	ERK activating kinase MAPK kinase	ERK1,2	PP2A
	b	MKK4,7	MEK related kinase (= SEK 1, JNKK1)	JNK/SAPK	
	c	MKK3,6	MEK related kinase	p38/RK	
MAPK (S/T)	a	ERK1,2	Extracellular signal regulated kinase = mitogen activated protein kinase. Migrates into the nucleus after being phosphorylated at Thr-Glu-Tyr (= TEY)	TCF = ELK1; ELK4. Myc, SAP1,2; Ets1,2; S6K; PLA2	CL100, PAC-1 (nuclear[3])
	b	JNK1,2,3/p54/ SAPKα,β,γ	c-Jun N-terminal kinase/stress activated protein kinase, ERK related. Migrates into the nucleus after being phosphorylated at Thr-Pro-Tyr (= TPY)	c-Jun, ATF2	CL100
	c	p38/RK/Mpk2	Reactivating kinase. Migrates into the nucleus after being phosphorylated at Thr-Gly-Tyr (= TGY). Related to yeast kinase HOG-1	MAPKAP kinase 2,3	CL100
AP[2]	a	1) ELK1,4; SAP-1,2; Ets-1,2, Myc	1) transcription factors: ternary complex factors (TCF, combine with SRF) = Elk-1, SAP-1,2, Ets-1,2 (Fig. 11.4-6)	Rsk phosphorylates phosphatase PP1G and Fos	
		2) p90-RSK	2) p90-Rsk = S6 kinase, insulin stimulating protein kinase, p90		
		3) PLA_2	3) PLA_2: formation of free arachidonic acid and lysophospholipids (7.4.8)		
	b	ATF2 = CREB1, Jun, ELK-2	Transcription factors (4.2.2.3 … 4, Fig. 4.2.2-6)		
	c	MAPKAP kinase 2,3	MAP-KAP kinase2 phosphorylates heat shock proteins and glycogen synthase (3.2.4)	Hsp25, Hsp27, HSF1	

[1] a, b, c designate different cascades, see Figure 7.5-4. – S,T,Y indicate phosphorylation at serine, threonine, tyrosine residues, respectively.
[2] Mitogen activated proteins (transcription factors and kinases).
[3] Before phosphorylated MAPKs migrate to the nucleus, they can be dephosphorylated by PP2A and by a phosphotyrosine phosphatase.

they also comprise of a unique ligand-specific α chain and a shared signalling competent β chain (gp140).

- The gp130 family (e.g., receptors for IL-6, IL-11, oncostatin M) are composed of a specific ligand- binding α subunit and a common signal-transducing β subunit (gp130).
- The growth hormone receptor family (e.g., receptors for growth hormone, EPO, Prolactin, granulocyte colony stimulating factor) have a single subunit and form homodimers upon ligand stimulation.

The type II cytokine/interferon receptor family (CRF2) comprises of 12 members (i.e., the receptors for interferon α,β,γ and τ as well as the receptors for IL-10, IL-19, IL-20, IL-22, IL-24, IL-26, IL-28, and IL-29). Upon ligand stimulation CRF2 members form heterodimers, in which both subunits function as signal transmitters into the cell.

The type III or TNF-receptor family contains a cysteine-rich common extracellular sequence and forms trimers after ligand binding.

The type IV cytokine (Ig superfamily/interleukin-1) receptor and Toll-like receptors. The former mediate cellular responses to interleukins produced by dendritic cells and macrophages/monocytes, while the latter recognize pathogen-specific molecules and, therefore, are also referred to as receptors for pathogen-associated molecular patterns. The presence of a TIR (Toll-Il-1-Receptor) domain is the defining structural feature that is common to the intracellular portion of all members of this superfamily as well as to their cytosolic adaptor molecules.

The family of the Interleukin-1 receptors/Toll-like receptors (TLRs, Fig. 8.1-3, Table 8.1-2) has a central role in mounting an innate immune defense against invading pathogens. Signaling by these receptors ultimately leads to activation of transcription factors, mainly of the NFkB (Nuclear Factor kappa B) and IRF (Interferon Regulatory Factor) families. TLRs form homo- and/or heterodimers, which are believed to exist as pre-assembled low-affinity receptors. Ligand binding induces a conformational shift that closely juxtaposes the two intracellular TIR domains. This is a structural prerequisite for adaptor recruitment. There are five known adaptor molecules: MyD88, MAL, TRIF, TRAM, and SARM. MyD88 mediates recruitment and activation of IRAK-4 (IL-1R Associated Kinase-4), which in turn recruits and activates IRAK-1. A key substrate of IRAK-1 is TRAF-6 (Tumor necrosis factor Receptor-Associated Factor 6), which recruits TAK1 (TGFβ-Activated Kinase 1), TAB2 (TAK1-binding Protein 2), and the ubiquitylating factors UEV1A and UBC13. This complex activates upstream kinases in the JNK and p38 pathways as well as NFκB signaling by degrading IKK (Inhibitor of NFκB Kinase). Alternatively the MyD88/ IRAK4/IRAK1/TRAF6 complex can recruit IRFs, like IRF7 which is phosphorylated and activated by IRAK1.

In case of all of the above described cytokine receptors, interaction with their respective ligand leads to recruitment and activation of kinases of the cytosolic janus kinase family (JAKs) and of membrane-associated Src family kinases (SFKs). The active JAKs and SFKs then phosphorylate themselves and the receptors, resulting in further recruitment of various SH2-domain-containing effector molecules to the complex.

Similar to RTKs, cytokine receptors can activate the Ras/Raf/MAPK pathway (Figs. 7.5-2 … -4) by recruiting the Grb2/Sos1 complex either directly or indirectly via the adaptor molecules Shc and SHIP (SH2-containing inositol 5-phosphatase). Additionally, cytokine receptors also often activate the PI 3K/AKT pathway by SH2-domain mediated recruitment of PI3K (p85). Tyrosine phosphorylation also generates docking sites for two key negative regulatory components of cytokine signalling, the hematopoietic cell phosphatase (SHP-1) and the suppressor of cytokine signalling (SOCS) family of proteins. Moreover, there is mounting evidence for important regulatory functions of different PKC isoforms in proliferation, apoptosis, and differentiation responses of hematopoietic cells (7.4.3).

Receptor	Phosphorylated by	Receptor	Phosphorylated by
IL-2R	Fyn, Lck, Syk	G-CSFR	Lyn, Syk
L-3R	Lyn, Tec?	GM-CSFR	Fes
IL-5R	Btk?	EPO-R	Fes
IL-6R	Btk, Hck, Tec	CD19 (B cell receptor)	Lck, Fyn, Lyn[1]
IL-7R	Fyn	CD4, CD8 (T cell receptor)	Lck[2]

[1] phosphorylate the ITAM region of the Igα/Igβ dimer (Fig. 8.1-11).
[2] phosphorylates the ITAM region of CD3 (Fig. 8.1-13).

JAK/STAT pathway (Fig. 7.5-5): An important component of cytokine signalling is the JAK/STAT pathway. After ligand binding and dimerization, the receptors for growth hormone, prolactin, erythropoietin, many interleukins and interferons associate at membrane-proximal sequences (boxes 1 and 2) with distinct protein tyrosine kinases of the Janus kinase family (JAK1 … 3, Tyk2). The name Janus kinase refers to the fact that these 120 … 140 kDa proteins have two carboxy-terminally located kinase domains, of which, however, the first one is catalytically inactive (pseudokinase domain). The interaction with the box 1/2 regions of cytokine receptors is mediated by the large, aminoterminal FERM domain. Once recruited to di-/oligomerized receptors, JAK proteins transphosphorylate and activate each other and also activate signal transduction by phosphorylating key tyrosine residues on the receptors. Several of these receptor phosphorylation sites represent binding motifs for the SH2 domain of a member of the STAT family (signal transducer and activator of transcription, seven known genes). In unstimulated cells STATs reside as inactive homodimers in the cytosol. Once recruited to ligand-stimulated cytokine receptors, STATs become substrates for the adjacent activated JAKs. Tyrosine phosphorylation leads to release of STAT molecules from the receptor complex as they rearrange into antiparallel dimers, in which the SH2 domain of one molecule binds to the tyrosine phosphorylation site of its partner and vice versa. These activated dimers translocate into the nucleus and directly activate transcription of their target genes.

As mentioned above, a number of two-chain cytokine receptors are composed of ligand specific α chains and common γ_c (or β_χ) chains,

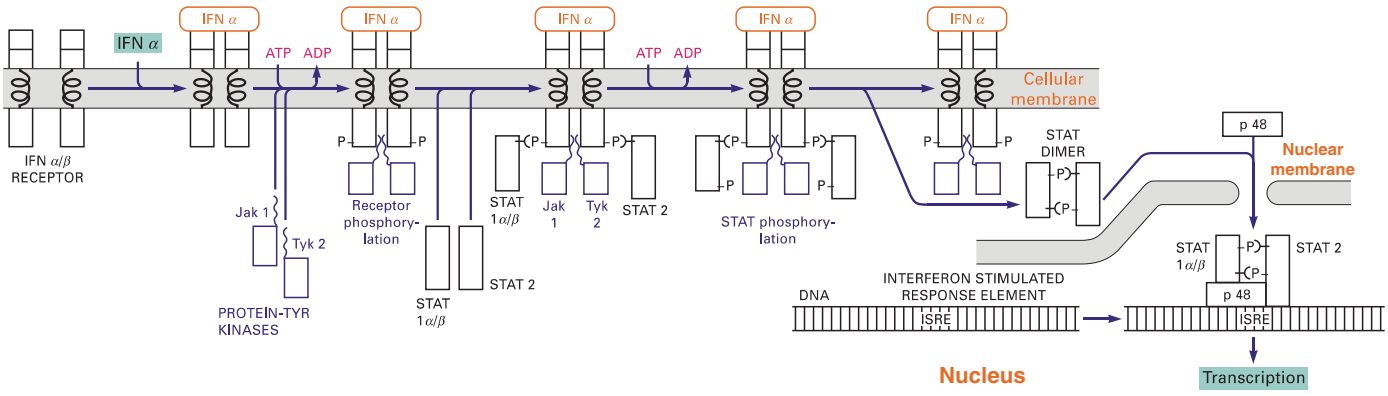


Figure 7.5-5. TKaR Receptor Signalling via STAT Proteins (Example: IFNα Receptor)
p48 = IRF9 (interferon regulating factor 9)

which associate with the same JAK after receptor dimerization (Fig. 8.1-19). This allows integration of the intracellular processing of different external signals and also explains some of the characteristic redundancy observed for different cytokines.

STAT activation is terminated by their dephosphorylation and nuclear export, as well as by feedback inhibition of JAKs through the above mentioned SOCS proteins, which also compete with STAT proteins for their docking sites on the receptor. SHP-2 and TC-PTP (T cell phosphatase) dephosphorylate nuclear STAT proteins. This seems to be a prerequisite for their nuclear export. Protein tyrosine phosphatases (PTPases) are structurally different from Ser/Thr phosphatases (7.5.3) and do not require metal ions. Catalysis is exerted by Arg and Cys residues. Many PTPases are themselves regulated by phosphorylation and some are transmembrane proteins, whose activity can be modulated by ligands (e.g., CD45, 17 … 5.5).

T cell receptors (TCRs; Figs. 7.5-6 and 8.1-13): The TCR consists of disulfide-linked α/β chain dimers and recognizes antigens, which are presented by MHC molecules (major histocompatibility complex, 8.1.5). However, in order to elicit cellular signalling events, TCRs have to form a complex with CD3 co-receptors (consisting of γε, δε, αβ and ζζ subunits), and, depending on the T cell type, either with the CD4 or the CD8 co-receptor. Such T cell receptor complexes are associated with Lck (or Fyn) kinases of the Src family through the cytoplasmic portions of CD4 and CD8. In resting cells, the associated Src family kinases (SFKs) are locked in an inactivate state by phosphorylation of a carboxy-terminal tyrosine residue (catalyzed by the Csk kinase). Intramolecular binding of the aminoterminal SH2 domain to this phosphorylation site obstructs access to the kinase domain and results in autoinhibition.

Upon contact of the TCR with peptide-loaded MHC molecules, the CD4 or CD8 co-receptors bind to monomorphic determinants of the MHC molecules class II or I, respectively, and in doing so move the associated SFK closer to the receptor. The transmembrane phosphatase CD45, which is also associated with the TCR, can then dephosphorylate the carboxy terminus of the SFK, thereby activating it. The activated SFK, in turn, phosphorylates tyrosine residues in the ITAM regions (immunoreceptor tyrosine-based activation motif) of CD3 co-receptors. These phosphate groups are docking sites for the SH2 domains of downstream effectors:

- Recruitment of the Grb2-Sos complex triggers the MAPK cascade (as described in 7.5.3), ultimately activating the AP-1 transcription complex (4.2.2.3). AP1 in cooperation with NF-ATp (see below) and NF-κB (nuclear factor κB) activates transcription of the IL-2 gene. The secreted IL-2 stimulates proliferation of T cells (see Fig. 8.1-20).
- After binding to the CD3ζ co-receptor via its tandem SH2 domains, the protein kinase ZAP-70 becomes activated by Lck mediated phosphorylation. ZAP-70 then phosphorylates the adapter molecules LAT and SLP-76. This triggers formation of a multimeric signalosome complex at the plasma membrane which also contains IL-2-inducible T cell kinase (Itk) and mediates activation of phospholipase Cγ1 (PLCγ1). However, the activity of Abl/Arg kinases is also required for TCR-stimulated phosphorylation of a certain tyrosine residue of ZAP70 and the downstream activation of LAT and SLP-76.
- PIP_2 is cleaved by PLC γ1 into inositol-P_3 (IP_3) and diacylglycerol (DAG). IP_3 induces Ca^{++} release from the ER which triggers the entry of extracellular Ca^{++} through calcium release-activated Ca^{++} channels (CRAC). Increased Ca^{++} leads to the formation of the Ca^{++}-calmodulin complex (7.4.3, 7.4.4). This complex attaches itself to the calcineurin dimer (PP2B) and enables it to dephosphorylate and activate the transcription factor NF-ATp (nuclear factor of activated T cells, preformed). Upon translocation to the nucleus, NFATp promotes gene transcription (i.e., IL-2 and other genes). Immunosuppressive drugs like cyclosporin A, which binds to the peptidyl-prolyl-*cis-trans* isomerase cyclophilin, interfere with this pathway by blocking the phosphatase activity of calcineurin.

Diacylglycerol (DAG), the other product of PLCγ1 catalysis, activates protein kinase C (PKC, 7.4.3), which, among other effects, also activates the MAPK pathway via Raf phosphorylation and NFκB which in conjunction with additional co-stimulatory signals (8.2-1) lead to IL-2 transcription (7.5.4).

B cell receptors (BCRs; Fig. 7.5-7, see also Fig. 8.1-11): The BCR has a similarly complex structure as the TCR. It consists of an antigen-binding subunit, the membrane immunoglobulin (mIg), non-covalently associated with a signalling subunit, which is composed of a disulfide-linked Igα/Igβ heterodimer. The co-receptor CD19 is in close contact with the glycoprotein CD21 (complement receptor 2, Table 8.1-6) and with the SFK Lyn. Upon contact of the receptors with antigen, activation of Lyn via CD 45-catalyzed dephosphorylation takes place (analogously to the T cell receptor mechanism). In the case of the BCR, the phosphorylation target is the immunoreceptor tyrosine-based activation motif (ITAM) in each of the Igα and Igβ co-receptors, which together act as signal transducers (similar to the CD3 subunits of the TCR complex). However, while the positive role of Lyn can be substituted by other SFKs (Fyn, Lck), it also has a unique inhibitory function due to its ability to phosphorylate immunoreceptor tyrosine-based inhibitory motifs (ITIMs) present in the inhibitory co-receptors CD22 and FcγRIIb. Once phosphorylated, these recruit PTPases which suppress BCR signalling.

Analogous to the role of Zap-70 for the TCR, the kinase Syk is activated upon binding via its tandem SH2 domain to the ITAMs of Igα and Igβ. Phosphorylation of several effectors by Syk is crucial for coupling BCR activation to downstream signaling. One of its key substrates is the B cell linker protein (BLNK), which not only recruits PLCγ2, but also activates the Ras/Raf/MAPK pathway via Grb2/Sos and Nck/Sos complexes (Table 7.5-4b). Activation of the PI-3K pathway upon BCR stimulation occurs in several parallel ways:

- by recruitment of the regulatory p85 subunit to tyrosine phosphorylation sites in the co-receptor CD19.
- by binding of the SH3 domain of Lyn to proline rich regions of p85.

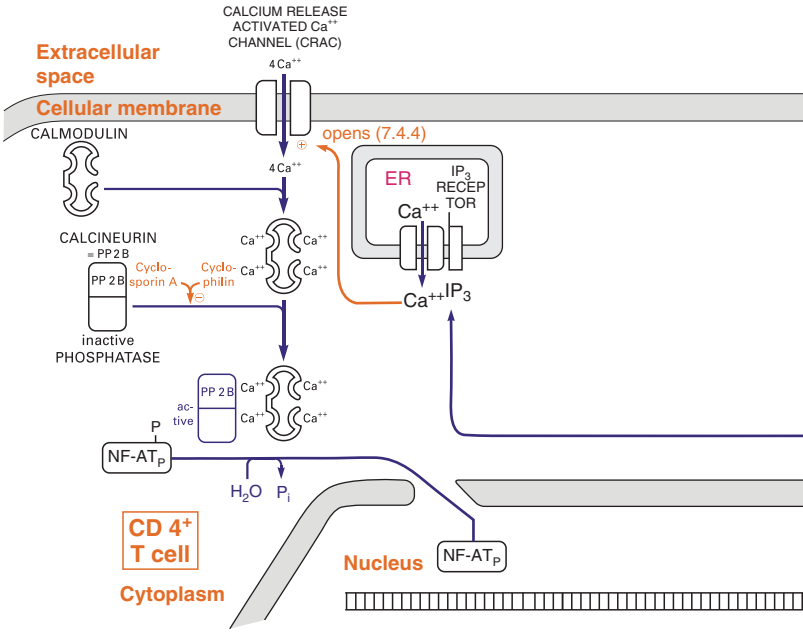


Table 7.5-4. Receptors Associated With Protein-Tyrosine Kinases (TKaR)

a) Cytokine Receptor Family. For Structures, see Figure 8.1-19

Type	Receptor Family	Group 1 Receptors	Group 2 Receptors	Group 3 Receptors
Ia	Hematopoietin rec. family	interleukin (IL)-2 R., IL-3 R., IL-4 R., IL-5 R., IL-7 R., IL-9 R.	granulocyte-macrophage colony stimulating factor (GM-CFS) R. (Fig. 8.1-5)	erythropoetin R., growth hormone (GH) R., prolactin R. (Fig. 7.1-4)
Ib		IL-6 R., IL-12 R.	granulocyte colony stimulating factor (G-CSF) R. (Fig. 8.1-5)	oncostatin M R., leukemia inhibitory factor (LIF) R.
II	Interferon rec. family	interferon R.; IL-10 R.		
III	TNF rec. family			tumor necrosis factor (TNF) R.; nerve growth factor (NGF) R.
IV	Interleukin-1 receptors / Toll-like receptors	IL-1 R.		

b) T and B Cell Receptors. For Structures, see Figures 8.1-11 and 8.2-1

Receptor	Co-receptor, associated with protein kinase	Co-receptor, assoc. with phosphatase, which activates this kinase	Target of phosphorylation	Ligands (kinases) docking at phosphorylated target
T cell R	CD4 + Lck (Lyn, Fys)	CD45	CD3ζ	ZAP-70, binds PLC1 and Grb2-Sos
T cell R	CD8 + Lck (Lyn, Fys)	CD45	CD3ζ	ZAP-70, binds PLC1 and Grb2-Sos
B cell R	CD19 + Src-type	CD45	Igα, Igβ	Syk, binds PLC1 and Grb2-Sos

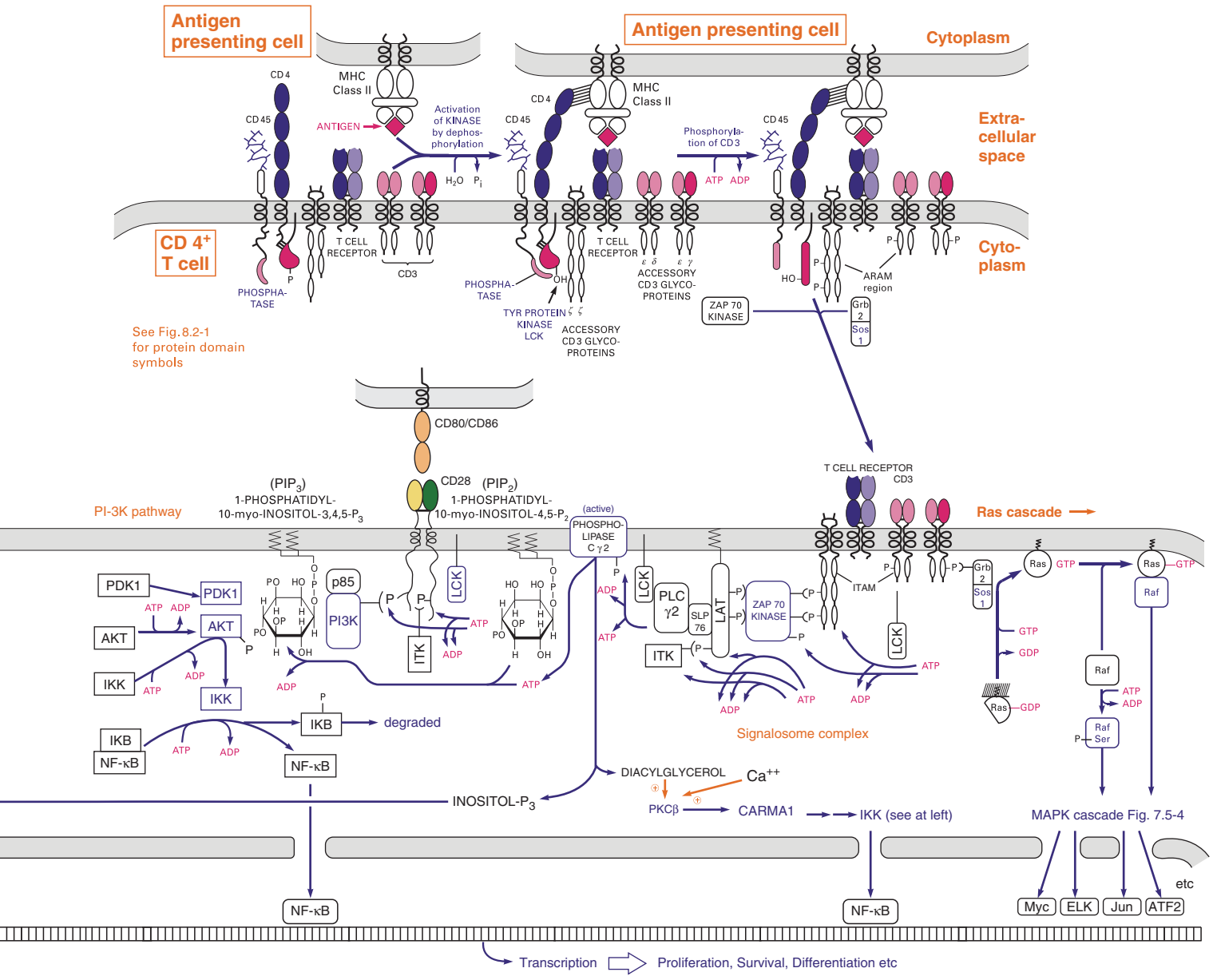


Figure 7.5-6. Signalling Pathways after Receptor Activation in T Cells

Many of the extra features shown here occur also in B-cells (Fig. 7.5-7). Names of components (for others, see text)

AKT = PKB, a serine-/threonine protein kinase (Fig. 7.5-4)
CARMA = Caspase recruitment domain-containing protein
IκB = Regulatory protein, which inhibits NF-κB by complex formation
IKK = IκB kinase complex
PDK1 = phosphoinositide-dependent protein kinase-1
PI3-K = Phosphoinositide 3-kinase

Table 7.5-5. Protein-Tyrosine Kinases Associated With Receptors

Kinase Family	Members	Mol. Mass (kDa)	Properties
Src family	Blk, Csk, Fgr, Fyn, Hck, Lck, Lyn, Src, Yes, Yrk Oncogene variants of Fes (v-, c-), Fgr (c-), Src (v- in Rous sarcoma virus), Yes (c-) are known.	53 … 64	SH2 and SH3 domains, kinase domain. Membrane anchored via myristoylation. They are activated by Tyr-kinase associated receptors and also by receptor tyrosine kinases. Lck and Fyn kinase are deactivated by phosphorylation (Csk kinase) and activated by dephosphorylation (CD45 phosphatase) in T and B lymphocytes.
Jak family	Jak1 … 3, Tyk2, hopscotch (*Drosophila*)	ca. 130	Tyrosine kinase- and tyrosine kinase-like domain, no SH2 and SH3 domains. They are directly activated by IFN receptors and phosphorylate them. This allows binding of STAT proteins to the receptor and their phosphorylation.
Syk family	Syk, ZAP-70	72, 70	2 SH2 domains, not myristoylated, intermediates in B and T cell receptor cascades
Other	Abl, Arg, Btk, Fak, Fer, Fes, Tec, Tsk Oncogene variants of Abl (v-, bcr- in leukemia)	50 … 150	Nuclear Abl promotes the phosphorylation of the CTD domain of RNA Pol II (4.2.1.2). Its activity is regulated during the cell cycle by Rb protein (4.3.3)

Table 7.5-6. Some Specific Transcription Factors (see also Table 4.2.2-2 and Fig. 4.2.2-6)

Name	Function	Characteristics	Oncogenes related to the encoding genes
Elk-1, SAP-1	Binds with SRF as ternary complex factor (TCF) to serum response element (Fig. 4.2.2-6)	member of Ets family	
Ets	Binds as a monomer to the Ets motif (C/AGGAA/T), binds to the TCR enhancer, possibly involved in general transcription on TATA-less core promoters (4.2.1)	helix-loop-helix (Table 4.2.2-1)	v-ets (avian retroviral oncogene)
Fos	Binds heterodimeric with Jun to AP-1 binding site (Fig. 11.4-6)	leucine zipper (Table 4.2.2-1)	c-fos, v-fos (retroviral oncogene)
Jun	Binds homodimeric (or heterodimeric with Fos) as AP-1 to the AP-1 binding site (Fig. 4.2.2-6)	leucine zipper (Table 4.2.2-1)	c-jun, v-jun (retroviral oncogene)
Myc	Binds to Myc control site, induction of cell division	helix-loop-helix (Table 4.2.2-1)	c- and v-myc (lead to cell immortalization)
STAT (1 … 4)	Bind to phosphorylated IFN receptors, dimerize and translocate to the *nucleus* after being phosphorylated by Jak kinases	Mol. mass (kDa): 1a: 91; 1b: 81; 2: 113; 3 88; 4 86	

- by recruitment of the adapter molecules BCAP (B cell adapter for PI3 kinase), and Gab1/2 (Grb2-associated binding protein), all of which are Syk substrates and contain phosphotyrosine binding motifs for the SH2 domains of p85 (not shown in Fig. 7.5-7).

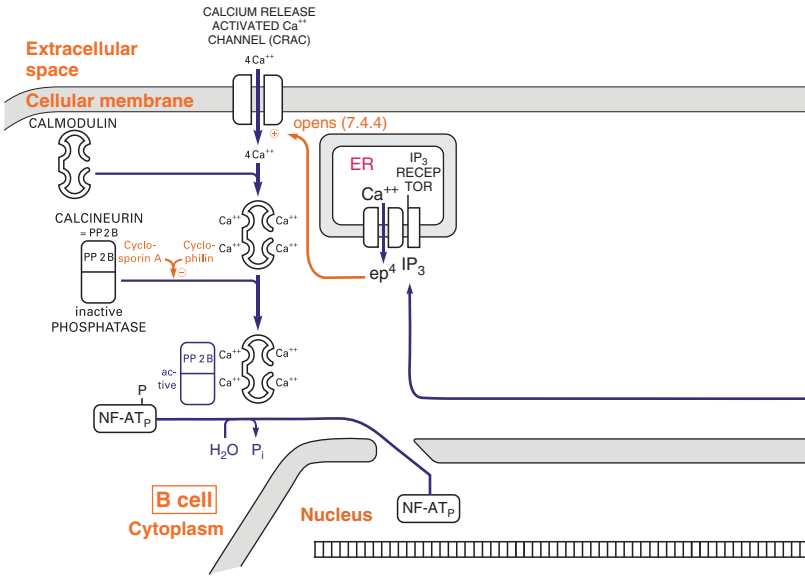
Phosphatidylinositol 3,4,5-triphosphate production by PI 3-K, in turn, is a prerequisite for plasma membrane recruitment and activation of Bruton's tyrosine kinase (Btk) via its pleckstrin homology (PH) domain. The importance of this activation step for BCR signalling is underscored by the fact that mice with X-linked immunodeficiency, which is characterized by impaired B-cell maturation and responsiveness, have a single point mutation in the PH domain of Btk that abolishes its binding to PI(3,4,5)P_3. Dual phosphorylation of PLCγ2 by Syk and Btk is necessary for full lipase activation and efficient generation of second messengers (IP_3, Ca^{++}, and DAG).

Literature:

Brown M.D., Sacks D.B. Cell Signal. 2009;21:462–469.
Cohen, G.B., Ren, R., Baltimore, D. Cell 1995;80:237–248.
Gasper R., *et al.* Nat. Rev. Mol. Cell. Biol. 2009;10:423–429.
Geng L., Rudd C.E. Br. J. Haematol. 2002;116:19–27.
Howe, L.R., Weiss, A. Trends in Biochem. Sci. 1995;20:59–64.
Karin M., Hunter, T. Current Biology 1995;5:747–775.
Leonard W.J. Int. J. Hematol. 2001;73:271–7.
Macara, I.G., *et al.* FASEB J. 1996;10:625–630.
O'Neill, L.A.J., Bowie, A.G. Nat Rev Immunol 2007;7:353–364.
Pawson, T. Eur. J. Cancer 2002;38(Suppl 5):3–10.
Pawson, T., Kofler, M. Curr. Opin. Cell Biol. 2009;21:147–153.
Stahl, N., *et al.* Science 1995;267:1949–1953.
Taniguchi, T. Science 1995;268:251–255.
Treisman, R. Current Opin. in Cell Biol. 1996;8:205–215.
van der Geer, P., *et al.* Ann. Rev. Cell Biol. 1994;10:251–337.
(Various authors). Trends in Biochem. Sci. 1994;19:439–513.
White, M.F., Kahn, C.R. J. Biol. Chem. 1994;269:1–4.

7.6 Programmed Cell Death (Apoptosis)

Maintaining tissue homeostasis and eliminating certain subpopulations of cells in a strictly controlled fashion are crucial tasks in the development and regeneration of multicellular organisms. Therefore, a tightly regulated and highly ordered cell death program known as apoptosis (or anoikis when triggered by loss of cell adherence) has evolved. The apoptotic process consists of the ordered, sequential activation of a series of specific cysteine proteases, the so-called caspases,



and endonucleases. This ultimately results in complete disintegration of all cellular structures. Apoptosis can be triggered by many different stimuli that all work through one of two signaling pathways: the intrinsic (or mitochondria-mediated) and the extrinsic (or receptor-mediated) pathway (Fig. 8.2-7).

Intrinsic apoptosis pathway: The intrinsic pathway is described in more detail in Chapter 8.2.5. It is commonly activated in response to cell damage (e.g., by radiation) or to dysregulation of normal cell function (e.g., overriding of cell cycle checkpoints, see 4.3.6). A key node that relays signals from sensors of cellular stress or damage to effectors of the intrinsic pathway is the tumor suppressor protein p53 (Fig. 4.3-7).

Extrinsic apoptosis pathway: While the extrinsic pathway ultimately mediates activation of the same caspase cascades, it is triggered not from inside the cell, but by the presence (in the case of death receptors) or by the absence (in the case of dependence receptors) of extracellular ligands. Besides that, cytotoxic T-cells and natural killer cells can extrinsically trigger apoptosis of target cells by releasing enzymes like granzymes A and B that enter the cytoplasm with the help of perforin molecules (for details see Chapter 8.2.5 and Fig. 8.2-8). The extrinsic and intrinsic pathways are interlaced at many levels and therefore mutually influence each other.

Receptor-induced apoptosis: There are two types of receptors that mediate extrinsic cell death signals. The so-called death receptors activate apoptotic pathways only when stimulated by the appropriate ligand, whereas the so-called dependence receptors act non- or anti-apoptotic in the presence of their ligand, but pro-apoptotic in its absence.

Death receptors (Table 7.6-1): Two apoptosis-inducing ligands are known in man: Fas Ligand (FasL) and TNF-related apoptosis-inducing ligand (TRAIL/Apo2L, Fig. 8.2-8). Both are type II transmembrane proteins and hence are normally membrane-bound ligands, unless cleaved by extracellular proteases (e.g., MMP7). The receptors for these ligands, the so-called death receptors, all belong to the tumor necrosis factor receptor superfamily (TNFRs). For FasL there is one functional receptor (Fas) known. Recently, the TNFR superfamily member 6b (TNFRSF6B or DCR3) was found to act as decoy receptor for FasL, since it can bind FasL, but lacks the intracellular signaling capacity to elicit apoptosis. For TRAIL there are two functional human death receptors (DR4, DR5), and two decoy receptors (DCR1 and DCR2).

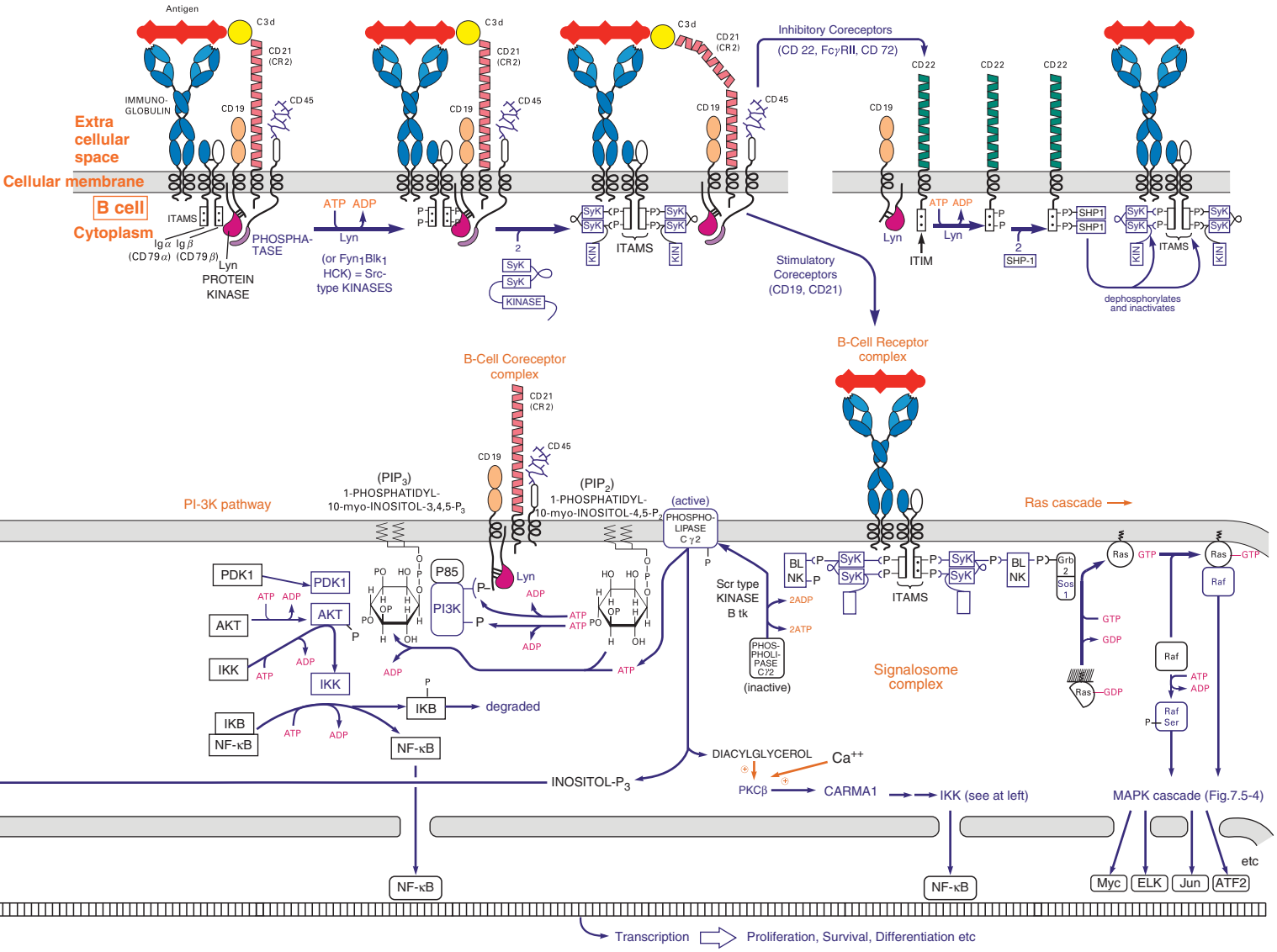


Figure 7.5-7. Signalling Pathways after Receptor Activation in B Cells
Many of the extra features shown here occur also in T-cells (Fig. 7.5-6).
SHP1 = Src homology phosphatase type 1
Syk = spleen tyrosine kinase

Table 7.6-1. Receptors Mediating Apoptotic Responses

Name	Superfamily (member)	Function in the presence (and absence) of ligand	Ligand
Death Receptors:			
Fas	TNFRSF (6)	Apoptosis induction	FasL
DcR3	TNFRSF (6B)	Decoy receptor	FasL
DR4	TNFRSF (10A)	Apoptosis induction	TRAIL/Apo2A
DR5	TNFRSF (10B)	Apoptosis induction	TRAIL/Apo2A
DcR1	TNFRSF (10C)	Decoy receptor	TRAIL/Apo2A
DcR2	TNFRSF (10D)	Decoy receptor	TRAIL/Apo2A
Dependence Receptors:			
$p75^{NTR}$	TNFRSF (16)	Neurotrophic receptor	Nerve growth factor (NGF)
Patched (Ptc1and 2)		Organ and limb morphogenesis, Epidermal development (Tumor suppressors)	Hedgehog family members (Shh, Ihh, Dhh)
Deleted in colorectal carcinoma (DCC)	Immunoglobulin superfamily	Axon guidance (Tumor suppressor)	Netrin-1
UNC5(A-D)		Axon guidance	Netrin-1
Androgen receptor (AR)	Nuclear receptor superfamily (NR3C4)	Steroid hormone dependent gene expression	Androgens
RET	Protein kinase superfamily (RTK subfamily)	Proto-oncogene (development of nervous system and kidneys, spermatogenesis	Glial cell line-derived neurotrophic factor (GDNF)
Integrins ($\alpha v\beta 3$, $\alpha 5\beta 1$)		Cell-matrix contacts	Extracellular matrix proteins

Like other TNF ligands, TRAIL/Apo2L and FasL form homotrimers. Hence ligand binding results in receptor oligomerization on the cell surface. The carboxy-terminal portions of the receptors contain so-called death domain motifs that serve as docking sites for recruitment of FADD (Fas-associated protein with death domain). A death effector domain (DED) in FADD can in turn recruit caspases 8 and 10. In this way, ligand-induced receptor oligomerization triggers formation of a multimeric death-inducing signaling complex (DISC). Within this complex, caspase 8 and 10 first autoactivate and then activate caspase 3, which can degrade many cellular proteins. Cleavage of gelsolin by caspase 3 causes fragmentation of actin filaments – a prerequisite for one of the phenotypic hallmarks of apoptosis, the disintegration of the cell into apoptotic bodies. After nuclear translocation caspase 3 also degrades ICAD/DFF45, the inhibitor of caspase-activated DNase/DFF40. This leads to internucleosomal cleavage of the chromosomal DNA and thereby generates the characteristic DNA laddering pattern of fragments, which are multiples of 200 bp length (Fig. 8.2-7). Apart from mediating caspase 3 activation, caspase 8 can also cleave Bid (Bcl2 homology domain 3-Interacting domain Death agonist). In its cleaved form, Bid triggers the intrinsic apoptosis pathway by interacting with various Bcl2 family members (e.g., Bax and Bak). In some cell types this is the primary mechanism of apoptosis induction by death receptors, while in others it serves as an amplification mechanism.

Dependence receptors (Table 7.6-1): The family of dependence receptors is functionally defined as receptors that transmit completely opposite intracellular signals depending on whether their ligand is available or not. In the presence of ligand these receptors transduce signals for cell survival, differentiation or migration, while in the absence of ligand they trigger apoptosis. Since these receptors signal in a positive fashion in the presence of ligand, but in a negative fashion in its absence, cells that express such receptors depend upon ligand availability for their survival. The concept of dependence receptors was established, when it was found that expression of the neurotrophic receptor $p75^{NTR}$ induced apoptosis in the absence of nerve growth factor. Like the above mentioned death receptors, $p75^{NTR}$ belongs to the TNFR superfamily. However, other family members show no sequence similarity to the TNFRs and in fact display a wide variety of different structural features. Currently, more than 10 other receptors are known to possess the defining properties of dependence receptors: the sonic hedgehog receptor patched, several netrin-1 receptors (DCC, UNC5A-D), the androgen receptor, RET (a receptor tyrosine kinase), and various integrins ($\alpha_v\beta_3$, $\alpha_5\beta_1$).

How dependence receptors mechanistically induce cell death, when not bound by ligand, is not understood in much detail, but all dependence receptors require caspase activity (Table 4.5.6-3). Most of the dependence receptors are themselves caspase substrates and cleavage by caspases is required for their pro-apoptotic action. Another common feature is the presence of Addiction/Dependence Domains (ADDs) that mediate apoptosis induction. The current working model is that the unligated receptors are recognized and cleaved by caspases (possibly also by other active proteases). Proteolytic cleavage either releases a pro-apoptotic intracellular fragment or unmasks an ADD thereby triggering apoptosis. Ligand binding inhibits the pro-apoptotic activity by blocking receptor cleavage and, at least in some cases, in addition by activating classical survival signal transduction pathways. Many of the dependence receptors, like the ones for netrin-1, are best studied in their functional roles of shaping nervous system development by regulating axon outgrowth, axon pathfinding, and neuronal migration. However, mounting evidence also implicates them in normal tissue homeostasis and suggests that their dysfunction plays a role in tumorigenesis. The netrin-1 receptor DCC (Deleted in Colorectal Carcinoma) has been proposed to be a tumor suppressor, because its expression is lost or strongly reduced in colorectal and other cancers.

Literature:

Chowdhury I., Tharakan B., Bhat G.K. Cell Mol Biol Lett. 2006;11:506–525.

Mehlen P., Thibert C. Cell Mol Life Sci. 2004;61:1854–1866.

Stupack D.G. Cell Death Differ. 2005;12:1021–1030.

7.7 Receptors for Steroid and Thyroid Hormones, for Retinoids and Vitamin D

Although steroid (3.5.4) and thyroid (3.2.7.5) hormones, retinoids (3.5.3, 3.7.1) and vitamin D (3.7.11) are structurally very different, they all represent small, hydrophobic molecules. For transport in blood, this type of hormone has to be bound to protein carriers. Since in this way they remain in the circulation for extended periods of time (hours to days), their action is long-lasting.

Once released from their carrier protein, the hydrophobicity of these hormones enables them to simply diffuse across the cytoplasmatic membrane. Contrary to other hormones, they bind and activate their cognate receptors intracellularly. The receptor/hormone complexes

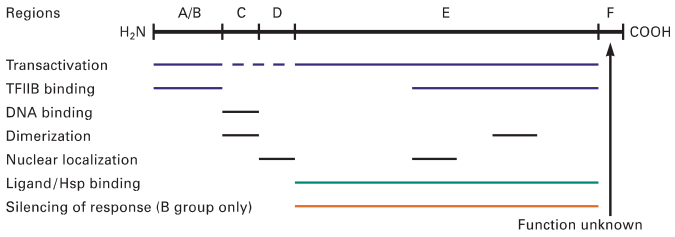


Figure 7.7-1. General Structure of Nuclear Receptors

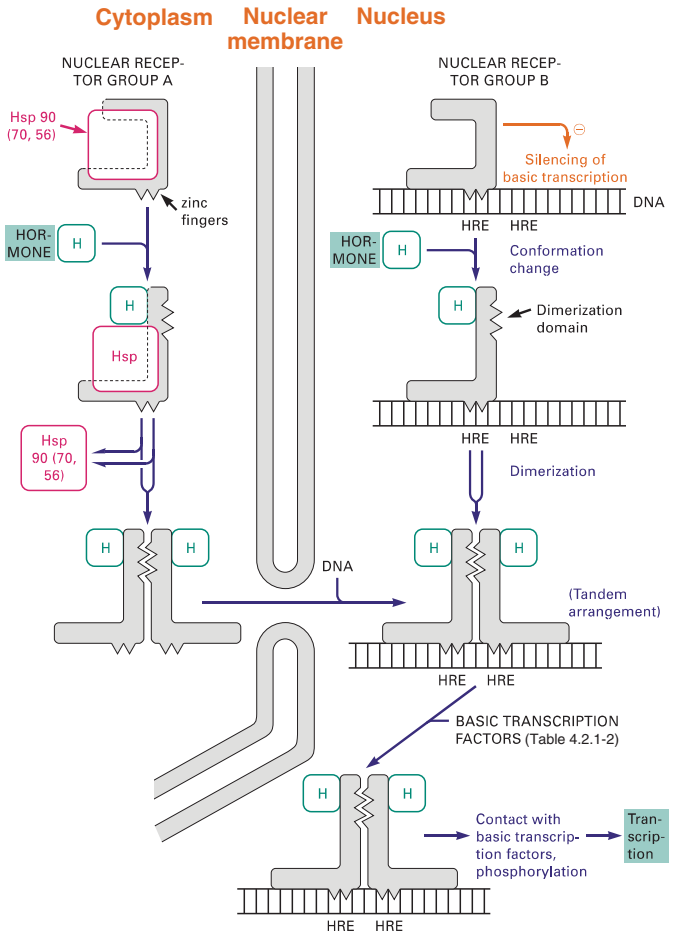


Figure 7.7-2. A Model for the Activation Mechanism of Nuclear Receptors (Structure is Schematized)

Table 7.7-1. Superfamily of Nuclear Receptors (There are many isoforms and additionally > 40 'orphan' receptors with unknown ligands)

Group		Name (R = Receptor)	Binds to[1]	Effects, see Chapter
A (homodimeric)	GR	glucocorticoid R	AGAACA	3.5.8.3, .4.2.2, 7.1.5
	MR	mineralocorticoid R	AGAACA	3.5.8.3, 7.1.8
	AR	androgen R	AGAACA	3.5.6.3, 7.1.5
	PR	progesterone R	AGAACA	3.5.5.2, 7.1.5
	ER	estrogen R[2]	AGGTCA	3.5.7.3, 4.2.2, 7.1.5
B (heterodimeric)	TR	thyroid hormone R	AGGTCA	4.2.2, 7.1.5
	RAR	all-*trans*-retinoic acid R	AGGTCA	7.3.2
	RXR	9-*cis*-retinoic acid R	AGGTCA	7.3.2
	VDR	vitamin D R	AGGTCA	4.2.2, 7.1.7
	E EcR	ecdysone R (insects)[3]	AGGTCA	7.2.3

[1] DNA half site. The receptor dimers bind to palindromic sequences or to direct repeats with various length of spacers.
[2] Specific receptor, binds only estrogens (Other A group receptors are less selective).
[3] Related receptor: ultraspiracle R, preferred dimerizing partner to EcR.

formed in the cytosol then translocate to the nucleus and bind to specific hormone response elements (HRE) on DNA, where they act directly as enhancers (or sometimes also as repressors) of transcription (4.2.1).

These receptors share a common structure principle (Fig. 7.7-1) and constitute the superfamily of nuclear hormone receptors (48 known human members).

Generally, these receptors consist of a C-terminal hormone binding and dimerization domain (E) that is connected via a hinge region (D) with a DNA binding domain (C, containing 2 Cys_2-Cys_2 zinc fingers, Table 4.2.2-1) and an N-terminal transcription activating domain (A/B). The DNA binding and a good portion of the dimerization domains are highly conserved, while the transcriptional activation domains differ greatly between receptors. Two receptor groups are to be distinguished (Table 7.7-1, Fig. 7.7-2):

- Group A receptors: In the absence of a hormone, an inhibitory protein (e.g., the heat shock protein Hsp90, possibly also Hsp70 and Hsp56) binds to the group A receptors and covers the DNA binding/dimerization domain. Alternatively, intramolecular blocking of the active sites in the absence of hormones has been proposed. Hormone binding causes a conformational change of the receptor, which results in dissociation of the inhibitory protein and enables formation of homodimers that then enter the nucleus. Subsequent high affinity binding of activated receptor complexes to palindromic DNA response elements and interaction of their transactivation domains with TFIIB accelerates formation of the pre-initiation complexes of transcription on target gene promoters (4.2.1).
- Group B receptors: They are not associated with heat shock proteins and bind to DNA even in the absence of a ligand. In this non-activated state, the group B receptors interfere with the formation of a pre-initiation complex and cause transcriptional silencing. A conformational change occurring upon hormone binding reverses the functional role of the receptor from a transcriptional inhibitor to a transcriptional activator. The opposing effects on gene expression are due to mutually exclusive recruitment of two different sets of interaction partners that act as adapters for the basal transcription machinery: corepressors and coactivators. Frequently, the activated receptors have to cooperate with additional protein factors for modulation of transcription (e.g., AP-1, 4.2.1). Receptor phosphorylation possibly also plays a role in its functional regulation.

In many cases, the activated receptors induce only the transcription of so-called primary response genes. Some of the expressed proteins then initiate the transcription of secondary response genes, while others repress further transcription of the primary response genes and thus serve to limit their activation.

Literature:

Evans, R.M. Science 1988;240:889–895.
Han S.J., *et al.* Trends Endocrinol Metab. 2009;20:8–15.
Leid, M., *et al.* Trends in Biochem. Sci. 1992;17:427–433.
Robinson-Rechavi M., *et al.* J Cell Sci. 2003;116:585–586.
Tsai, M.J., O'Malley, B.W. Ann. Rev. of Biochem. 1994;63:451–486.
(Various authors) in: Cooke, B.A., *et al.* (Eds.) *Hormones and Their Actions* Part I. *New Comprehensive Biochemistry* Vol. 18A. Elsevier, 1988.

7.8 Cyclic GMP Dependent Pathways and Effects of Nitric Oxide (NO)

In several vertebrate organs, hormones activate guanylate cyclases (GC) to synthesize 3′,5′-cyclic GMP (cGMP) according to the reaction

$$GTP = cGMP + PP_i.$$

Guanylate cyclase also plays a role in lower animals (e.g., chemoattraction of amoebae). There are indications of cGMP occurrence in plants, archaea and bacteria.

As a second messenger, cGMP acts either directly on targets or activates protein kinase G (PKG) that then phosphorylates various substrates and thus modifies their actions.

7.8.1 Membrane Bound Guanylate Cyclases (Fig. 7.8-1)

These enzymes are monomeric glycoproteins with different extracellular ligand-binding domains. They act as receptors.

- GC-A and -B in brain, heart, smooth muscles and kidneys can be stimulated by the atrial natriuretic factor (ANF, 7.1.8). Binding of this ligand causes phosphorylation and activation of the enzyme. The formed cGMP mediates (via PKG) Na^+ and water secretion by the kidneys and vasodilation. ATP or ADP potentiate this effect.
- GC-C in the intestinal mucosa can be stimulated by the peptide guanylin (15 aa) and by structurally similar bacterial (*E. coli*) enterotoxins. This leads via cGMP and PKG to changes in ion transport mechanisms and is the molecular basis of enterotoxin-caused diarrhea.
- In retinal rods, GC sustains the cGMP level, which keeps the Na^+ channels open (7.4.6).

  Upon stimulation of retinal cells by light, cGMP is hydrolyzed and the ion channels close, therefore, the intracellular Ca^{++} concentration decreases. The Ca^{++} binding protein recoverin (23 kDa), acting as a sensor, changes its conformation at low Ca^{++} levels and thus becomes able to activate GC. The resulting increase in cGMP level opens the ion channels again (recovery phase).

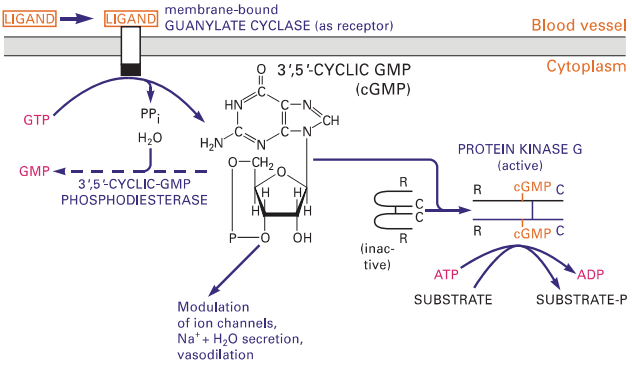


Figure 7.8-1. cGMP Formation and Protein Kinase Activation by Membrane Bound Guanylate Cyclase

7.8.2 Soluble Guanylate Cyclases and Their Activation by Nitric Oxide (NO)

NO metabolism: The NO radical is a gas, which diffuses freely through membranes and is quickly oxidized ($t_{1/2}$ = 5 … 30 sec). Therefore it acts as such only on short distances. However, in form of SNOs (S-nitrosothiols, X-S-N=O, where X is, among others, the β93 Cys of deoxygenated hemoglobin-Fe^{++}, Cys of membrane proteins or of soluble low molecular weight thiols), it can be stored for some time and also be transported to their site of action.

NO is synthesized in vertebrates by NO synthases from arginine in an unusual 5-electron reaction (Fig.7.8-2). All NO synthases are homodimers and contain FAD, FMN, tetrahydrobiopterin and heme (protoporphyrin IX). 3 isoforms of NO synthase are known (Table 7.8-1).

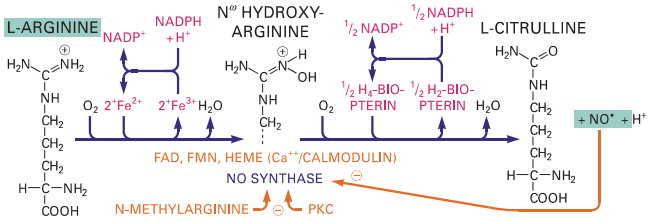


Figure 7.8-2. Synthesis of Nitric Oxide

The NO formed by the constitutive type I and III enzymes mainly activates soluble guanylate cyclase. These enzymes can be upregulated by physiological stimuli (e.g., hypoxia). The main effect of NO synthesized by the inducible type II enzyme is cytotoxicity.

NO synthase activity has also been found in gram positive bacteria, e.g., *Staphylococcus aureus.* Here, it has a protective function against oxidative stress, diverse antibiotics and host immune repressors.

In blood, NO is oxidized by oxygenated hemoglobin to nitrate (and nitrite), while deoxygenated hemoglobin can reconstitute this compound (NO cycle, 6.3.2). Details are still under discussion.

The receptors, which activate the type I and III NO synthases, are stimulated by neurotransmitters (via an increase of intracellular Ca^{++}). Even shear forces in the bloodstream are effective.

The synthesis of NO synthase type II is induced by bacterial lipopolysaccharides and cytokines (IFNγ). The resulting NO destroys bacteria and tumor cells and thus is part of the cellular defense system (8.2). Nitroglycerol and other nitro drugs are metabolized to NO• and thus mimic its physiological effects. Similar effects are ascribed to carbon monoxide (CO) and to the hydroxyl radical (OH•, 4.5.8), which also activate guanylate cyclase.

Soluble guanylate cyclases (GC-S, Fig. 7.8-3): These enzymes are heterodimers of α_1 and β_1 or β_2 subunits. The enzymes contain heme as prosthetic group. Binding of NO activates the enzyme (contrary to most other heme containing enzymes).

The cGMP formed activates protein kinase G (7.8.3). It also regulates some cAMP phosphodiesterases (7.4.2).

7.8.3 Protein Kinase G (PKG, Fig. 7.8-1)

The cGMP formed by membrane bound or soluble GC can activate protein kinase G.

This enzyme is a serine-threonine kinase (ca. 80 kDa). It contains a regulatory (R) and a catalytic (C) domain on each peptide chain

Table 7.8-1. Types of Eukaryotic NO Synthase

NO synthase type		Regulation	Occurrence	Effects of formed NO
nNO synthase (neuronal, Type I)	soluble	Activated by norepinephrine, acetylcholine, vasopressin, oxytocin, cytokines, glutamate via Ca^{++}/calmodulin	a) CNS (postsynaptic)	Retrograde diffusion → activation of presynaptic GC → modulation of signal transmission. Possibly involved in brain development and learning, but also in stroke damage and neurodegenerative disorders.
			b) peripheral neurons (presynaptic)	Activation of postsynaptic GC → muscle relaxation
eNO synthase (endothelial, Type III)	membr. bound	Activated by hormones via Ca^{++}/calmodulin, shear forces	endothelial cells	Activation of GC in underlying muscle → relaxation and vasodilation, angiogenesis, negative inotropic effect. Increased mitochondria content in muscles, heart, brain, kidney etc. Diffusion to platelets → decrease of adhesion and aggregation.
iNO synthase (inducible, Type II)	soluble	Synthesis induced by bacterial lipopolysaccharides and by cytokines. Not stimulable by Ca^{++} (permanently saturated)	macrophages, hepatocytes, etc.	Diffusion to neighboring tissues and cells → blocks Fe centers in the respiratory chain, aconitase, ribonucleotide reductase, causes DNA damage, acts bactericidal (8.2.4–8.2.6) and tumoricidal. Misregulation leads to autoimmune diseases, septic shock. etc.

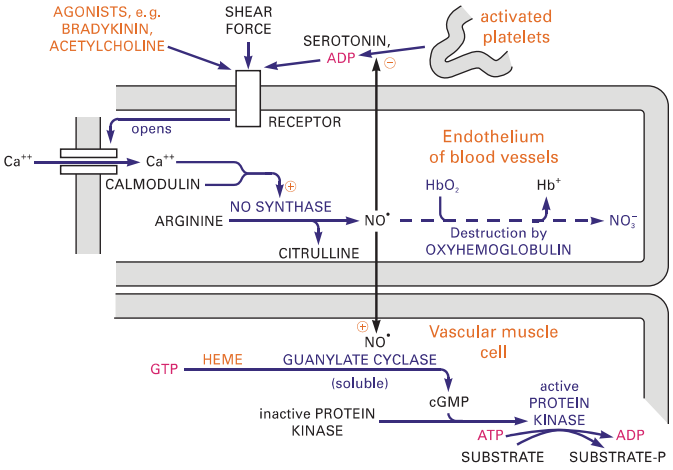


Figure 7.8-3. Vasodilatory and Antiaggregatory Effects of NO

of the homodimer. The structure of both domains resembles the two subunits of PKA (7.4.2, which, however, are located on different peptide chains). Binding of cGMP causes a conformational change, which relieves autoinhibition of the active domain without causing dissociation of the R and C subunits. High activities of PKG are found in the smooth muscles and the brain of mammals.

Phosphorylation by this enzyme

- stimulates the ion pumps which keep the intracellular Ca^{++} concentration low, causing muscle relaxation and vasodilatation.
- keeps Na^{+} channels open, resulting in Na^{+} and water secretion and in neurological effects.

Literature:

Allen, B.W. *et al*. Trends in Molecular medicine 2009;15:452-460.

Bredt, D.S., Snyder, S.H. Ann. Rev. Biochem. 1994;63:175–195.

Conway, E.J., Collen, D., Carmeliet, P. Cardiovasc. Res. 2001;49:507–521.

Chinkers, M., Garbers, D.L. Ann. Rev. Biochem. 1991;60: 553–575.

Garthwait J. Front. Biosci. 2005;10:1868–80.

Gusarov, I. *et al*. Science 2009;325:1380–1384.

Ishii, K., *et al*. J .Appl. Cardiology 1989;4:505–512.

Jeffrey, S.R., Snyder, S.H. Ann. Rev. Cell Dev. Biol. 1995;11: 417–440.

Kuhn M. Circ. Res.2003;93:700–709.

Nisoli, E. *et al*. Proc.Natl.Acad.Sci (USA) 2004;101:16507–16512.

Schmidt, H.H.H.W. FEBS Letters 1992;307:102–107.

Tota, B. *et al*. Current Medicinal Chemistry 2010;17:1915–1925.

8 Immune System

Ernst Peter Rieber

The immune system has developed to defend higher organisms against infectious agents such as bacteria, viruses, parasites and worms. In addition, it contributes to the homeostasis of the organism by eliminating dead, damaged, senescent and malignantly transformed cells. It closely interacts with the two major control systems of the body, the central nervous system (CNS) and the endocrine system. In the following, structure and function of the human immune system are described.

8.1 Components of the Immune System

Pathogenic microbes are prohibited from entering the body by the epithelial cell layers of the skin and the mucous membranes of the respiratory, gastrointestinal and genitourethral tract. This physical barrier is strengthened by antimicrobial substances such as defensins produced by the keratinocytes of the skin, lactic acid and fatty acids secreted by sebaceous and sweat glands, and lysozyme and lactoferrin contained in the mucus.

Pathogenic microbes are highly diverse and proliferate rapidly after entering an unprotected host. Therefore, an efficient defense system must be able both to immediately react against the invaders and to specifically recognize the multitude of different pathogens. Most importantly, it has to cope with the subtle strategies pathogens develop to evade host defense mechanisms. To fulfill these requirements the immune system is composed of two strategic arms:

- the immediately active, but less specific innate (natural) immune system and
- the highly specific, adaptive or acquired immune system which needs some time to establish an effective defense against newly encountered pathogens.

Both systems are composed of soluble factors and mobile cells. Whereas the innate immune response always follows the same inborn pattern of recognition and elimination of pathogens, the specific immune response is highly diverse, flexible, and by establishing a specific memory after the first contact with a pathogen it can quickly and more efficiently act at the second encounter and at further contacts with the pathogen (see Fig. 8.1-1).

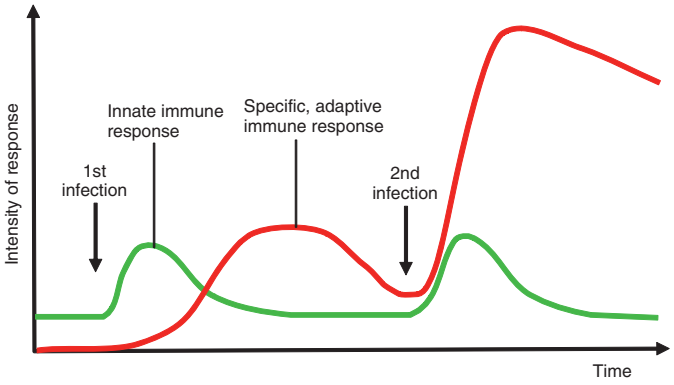


Figure 8.1-1. Innate and Adaptive, Specific Immune Response at First and Second Encounter with an Infectious Agent

Table 8.1-1. Components of the Non Adaptive Immune System

Soluble factors	Antibiotic peptides (lysozyme, defensins, serprocidins, bactericidal / permeability increasing protein (BPI)
	Factors of the complement system (see 8.1.5)
	Soluble extracellular receptors for conserved pathogen structures: acute phase proteins (APP) [C-reactive protein (CRP), serum amyloid A (SAA)], collectins (mannan-binding lectin (MBL), surfactant-apoproteins A and D)
	Soluble intracellular receptors for conserved pathogen structures: NOD-like receptors (NLR), RIG-like receptors (RLR)
	Iron-binding proteins (lactoferrin, transferrin)
	Cytokines (IL-1, IL-6, TNF-α, IFN-α, IFN-β), chemokines
Membrane-bound receptors	Mannose receptors, galactose receptors, scavenger receptors Receptors for soluble factors (complement receptors, collectin receptors, pentraxin receptors) Toll-like receptors
Cells	Phagocytes (macrophages, neutrophilic granulocytes)
	Dendritic cells
	Eosinophilic and basophilic granulocytes, mast cells
	Natural killer cells (NK cells)

8.1.1 Innate, Non Adaptive Immune System

8.1.1.1 Soluble Factors

These factors can either directly attack pathogens or indirectly promote their elimination.

Direct effect: The so called antibiotic peptides such as lysozyme, defensins, serprocidins and the bactericidal/permeability increasing protein (BPI) are directly cytotoxic to pathogenic microbes. Lysozyme cleaves the peptidoglycan of bacterial cell walls. Defensins are amphiphilic peptides that integrate into the phospholipid layer of the bacterial cell wall. The bactericidal Cathepsin G, elastase and proteinase-3 (PR-3) belong to the serprocidins. They are homologous to serine proteases (4.5.6.2) and are mainly found in the azurophilic granules of neutrophilic granulocytes which also contain the bactericidal/permeability increasing protein (BPI), a LPS-binding protein with increased cytotoxicity against Gram-negative bacteria (3.10.1).

Indirect effect: Other factors such as collectins or pentraxins bind to conserved surface structures on the pathogens. On the one hand, this results in the activation of the complement cascade leading to the deposition of C3b on the surface of the pathogens and to the assembly of the lytic membrane attack complex (see complement, 8.1.5). On the other hand, when these factors are bound to pathogens they are recognized by specific receptors on phagocytic cells. This increases the cellular uptake of pathogens, a process called opsonization (Fig. 8.1-2).

The acute phase proteins (APP) C-reactive protein (CRP) and serum-amyloid A (SAA) are pentraxins. Their production in the liver is induced by the cytokines IL-1 and TNF-α that are secreted by activated macrophages. CRP binds to phosphorylcholine on the surface of pneumococci. Limited proteolysis of SAA can lead to insoluble amyloid fibrils that are deposited in organs and tissues and can result in amyloidosis, a complication encountered in chronic infections. The serum concentration of APP is used as a diagnostic marker of infectious disease activity.

Collectins are molecules with a collagen-like tail and a C-type lectin domain (carbohydrate recognition domain, CRD) at the C-terminus that binds to complex carbohydrates of microbes. The collectin family includes the mannan-binding lectin (MBL) and the surfactant proteins A and D (SP-A, SP-D). MBL induces the lectin pathway of the complement cascade. It is important in the defense against influenza A viruses in the first 18 months of life when the carbohydrate-specific antibody formation is still reduced.

Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Second Edition. Edited by Gerhard Michal and Dietmar Schomburg.

Iron-binding proteins such as lactoferrin and transferrin (6.3.1) reduce the concentration of iron that is required by proliferating microorganisms.

Microbes can directly activate the alternative and the lectin pathway of the complement system. The generated C3b is an important opsonizing factor (Fig. 8.1-2).

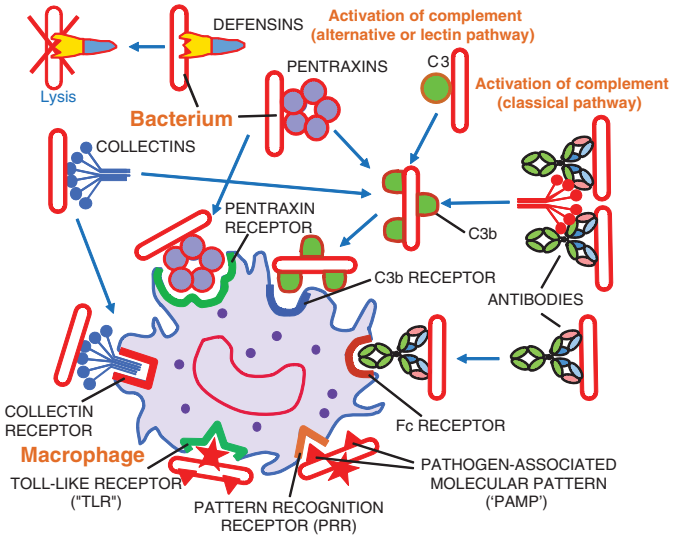


Figure 8.1-2. Soluble Factors and Receptors on Phagocytic Cells in the Defense Against Pathogens

8.1.1.2. Cells of the Innate Immune System

Monocytes/macrophages, neutrophilic granulocytes and dendritic cells (DC) develop along the myelo-monocytic differentiation line of hematopoiesis (see Fig. 8.1-5). The main function of monocytes/macrophages and neutrophilic granulocytes is phagocytosis and intracellular destruction of pathogens, mainly by reactive oxygen species (ROS, 3.2.5.8) and enzymatic degradation. Dendritic cells take up infectious material and process it for the presentation to specific T cells (see 8.2.1). Hence, they are pivotal for the induction and programming of the specific, adoptive immune response.

Natural killer cells are cytotoxic and destroy virus-infected cells. By secreting cytokines they are involved in the regulation of the innate and adaptive immune response.

Basophilic granulocytes, mast cells and eosinophilic granulocytes are important effector cells in the defense against parasites and in allergic reactions. Following antigen-mediated crosslinking of surface receptor-bound IgE and IgG molecules they release proinflammatory mediators such as histamin stored in granules (see 8.3.1).

8.1.1.3 Danger Receptors (Fig. 8.1-3)

Cells of the innate immune system recognize pathogens through a variety of evolutionary conserved receptors. These are either expressed

- on the cell surface or
- on endosomal membranes or they occur
- extracellularly as soluble receptors or
- within cells in the cytosolic compartment.

Part of the receptors interacts with molecules that are common to various classes of pathogens. These structures are essential for the survival of the microbes and are, therefore, not mutated. They are not represented in the host organism and are called 'pathogen-associated molecular patterns, PAMPs'. Another group of receptors react with microbial structures that are not distinguished from the analogous structures of the host, yet occur at locations where they are not found physiologically such as DNA in the cytoplasm during viral infection. Other receptors recognize endogenous structures that appear in the course of cell stress or cell damage and are therefore called damage-associated molecular patterns, DAMPs. Examples are heat shock proteins, monosodium urate crystals and reactive oxygen species (ROS). Altogether, the receptors of the innate immune system are called pattern recognition receptors (PRR).

Some receptors bind to PAMPs on microorganisms and increase their uptake by phagocytic cells. These opsonizing receptors are either soluble as the pentraxins and collectins (see 8.1.1.1) or membrane-bound like the mannose and galactose receptors or the scavenger receptors that bind to LPS, peptidoglycan, teichoic acid and phospholipids of the bacterial wall (3.10.1).

One important function of PRRs is the generation of alarm signals that activate the cells and stimulate the production of proinflammatory cytokines. Furthermore, they increase the capacity of antigen-presenting cells to activate T cells and to program a specific adaptive immune response. Since these receptors indicate danger to the organism and call for an immediate immune defense they are also called danger receptors. They recognize PAMPs and DAMPs in the extracellular space as well as within cells. Extracellular surveillance is mediated by the Toll-like receptors (TLRs), whereas the interior of cells is supervised by the Nod-like receptors (NLRs) and the RIG-like receptors (RLRs).

Toll like receptors (TLR) are homodimeric transmembrane proteins with an extacellular ligand-binding domain containing leucine rich repeats (LRRs, Fig. 8.1-3 ①) and a cytosolic signal-transducing domain (Toll/IL-1 receptor domain, TIR) ②. The 10 TLRs identified to date in humans, their localization and their ligands are listed in Table 8.1-2.

Table 8.1-2. Toll-Like Receptors

Receptor	Localisation	Ligand
TLR 1	PM	triacyl lipopeptides, TLR1 combines with TLR2
TLR 2	PM	triacyl- and diacyllipopeptides, peptidoglycans, lipoteichoic acid, lipoproteins, glycolipids, lipoproteins, zymosan, HSP70
TLR 3	E, L	double-stranded RNA, poly-IC
TLR 4	PM, E, L	lipopolysaccharides (LPS), lipoteichoic acid, heat shock proteins
TLR 5	PM	flagellin
TLR 6	PM, E, L	diacyl lipoproteins, lipopeptides, zymosan, combines with TLR2
TLR 7	E, L	single-stranded RNA
TLR 8	E, L	single-stranded RNA
TLR 9	E, L	single-stranded DNA, plasmodial hemozoin
TLR 10	?	unknown

PM = plasma membrane, E = endosome, L = lysosome.

TLR1, 2, 4, 5 and 6 are expressed on the plasma membrane and recognize cell wall components of bacteria and fungi. TLR4 requires additional proteins such as LPS-binding protein (LBP), CD14, MD2 and RP105 in order to interact with lipopolysaccharide (LPS) of Gram-negative bacteria ③. After binding of LPS the TLR4-MD2-LPS-complex is translocated into the endosomal compartment ④. TLR3, 7, 8 and 9 are located in the endosomal membrane and bind nucleic acids which are either endocytosed together with apoptotic material from virus-infected cells ⑤ or are translocated from the cytoplasm into the endosome by a process called autophagy ⑥. Nucleic acids are found in endosomes preferentially in danger situations such as viral infection and cellular stress. TLR3 binds dsRNA, TLR7 and 8 interact with ssRNA and TLR9 recognizes ssDNA. TLRs associate with variable combinations of adapter proteins via the TIR domains. Most TLRs interact with the myeloid differentiation primary response gene 88 (MyD88) ⑦, the endosomal TLR3 and TLR4 associate with the 'TIR-containing adapter inducing IFN-β' (TRIF) ⑧. The signal cascades initiated by the adapter proteins lead to the production of type I interferons and proinflammatory cytokines (see 7.5.4).

Intracellular receptors: The soluble intracellular danger receptors comprise mainly the Nod-like receptors (NLR) and the RIG-like

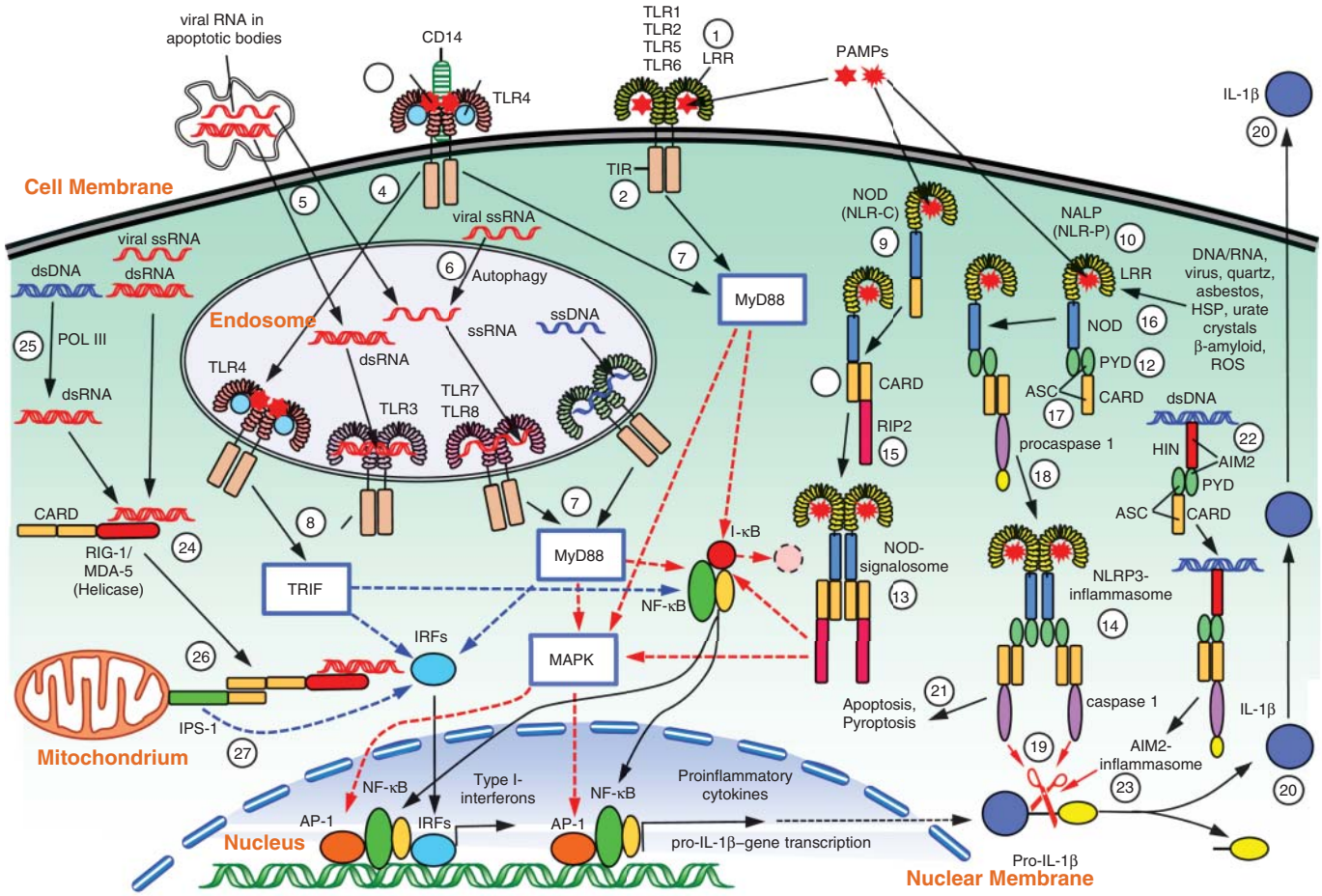


Figure 8.1-3. Danger Receptors

receptors (RLR). These receptors sense replicating microorganisms and cell damage or cellular stress. They stimulate the production of type I interferons and proinflammatory cytokines and, in addition, they are able to induce programmed cell death (apoptosis, see 8.2.5) of infected or stressed cells.

Nod-like receptors (NLR): NLRs include the NODs (NLR-C) ⑨ and the NALPs (NLR-P) ⑩. They consist of

- a C-terminal ligand-binding domain containing LRRs,
- a central nucleotide-binding oligomerization domain (NOD), and
- a protein interaction domain at the N-terminus which, in NODs, comprises one or two caspase recruitment domains, (CARD) ⑪, and in NALPs a pyrin-domain, PYD ⑫.

After binding of ligands NLRs oligomerize and recruit additional proteins to assemble into large complexes. NOD complexes are called 'NOD-signalosomes' or 'nodosomes' ⑬, NALP complexes are referred to as 'inflammasomes' ⑭. By a homophilic CARD-CARD interaction NODs associate with the serine/threonine kinase 'receptor interacting protein-2' (RIP2) ⑮ which activates NF-kB and the MAPK signaling pathway (7.5.3) resulting in the production of proinflammatory cytokines. NALPs are activated by a number of structures occurring in the cytosolic compartment during infection and cell stress such as DNA, RNA, heat shock proteins, urate crystals, β-amyloid and ROS ⑯. Through their pyrin domains NALPs associate with the adapter protein 'apoptosis-associated speck-like protein' (ASC) which contains a CARD domain ⑰. Via CARD-CARD interaction NALPs bind procaspases resulting in autocatalytic caspase activation ⑱ (see 8.2.5). The caspases, particularly caspase-1, process cytokine precursors such as pro-IL-1β ⑲ and thus promote the release of the active proinflammatory cytokines ⑳. In addition, procaspase-1 can induce apoptosis by activating nucleases ㉑. In contrast to other forms of apoptosis this apoptotic pathway is accompanied by inflammation and is, therefore, called 'pyroptosis'.

Double-stranded DNA is found in the cytosol during infection and at cell damage. It is recognized by various cytosolic sensors such as the 'DNA-dependent activator of IFN-regulatory factors' (DAI) and 'absent in melanoma2' (AIM2) ㉒. Binding of DNA induces the association of AIM2 with ASC (see above) and procaspase-1 resulting in the AIM2-inflammasome ㉓ which also processes pro-IL-1β and leads to the secretion of active IL-1β.

RIG-like receptors (RLR): Viral double-stranded and single-stranded RNAs are recognized by the two cytosolic 'RIG-1-like helicases' (RLH), 'retinoic acid-inducible gene 1' (RIG-1) and 'melanoma differentiation-associated gene-5' (MDA-5) ㉔. Cytosolic double-stranded DNA can be transcribed by polymerase III into dsRNA ㉕, which is then recognized by RLHs. After binding of RNA the helicases interact via CARD domains with the adapter molecule 'IFN-β promoter stimulator-1' (IPS-1), which is located in the external membrane of mitochondria ㉖. IPS-1 activates interferon responsive elements, IRFs and NF-κB ㉗ and thus induces the production of type-I interferons and proinflammatory cytokines.

The early recognition of pathogens and the induction of type I interferons and proinflammatory cytokines by danger receptors are pivotal for the first line defense against microbial infection and for the initiation of a specific immune response.

8.1.2 Specific, Adaptive Immune System

Soluble factors assigned to the specific, adaptive immune system are the various classes of specific antibodies (immunoglobulins) and a number of cytokines that regulate the immune response. The cells of the adaptive immune system are the B lymphocytes which produce the antibodies and T lymphocytes, which regulate the immune response or are cytotoxic effector cells.

Table 8.1-3. Components of the Specific, Adaptive Immune System

Soluble factors	Immunoglobulins (antibodies)
	Cytokines (IL-2, IL-4, IL-5, IFN-γ, IL-10, TGF-β, IL-17, IL-23)
Cells	B lymphocytes
	T lymphocytes

The specific adaptive immune system has several characteristic features:

- The recognition structures of the adaptive immune system (antigen receptors) are extremely diverse and provide the capacity to react specifically against all conceivable molecular structures.
- The diversity of antigen receptors is autonomously generated during ontogenesis of lymphocytes by somatic recombination of a limited number of inherited genes. It is increased after specific antigen activation of lymphocytes by point mutations in the DNA coding for the binding sequences of the receptors.
- The huge spectrum of receptor specificities is largely present at the first contact of the organism with an antigen ('anticipatory immune system'). Each individual lymphocyte carries receptor molecules of only one specificity (clonal distribution of receptors). A nominal antigen activates the lymphocytes presenting the best fitting receptors (clonal selection). The activated lymphocytes proliferate and differentiate into clones of specifically reactive effector lymphocytes (clonal proliferation) (Fig. 8.1-4).
- For the effective defense against extracellular and intracellular pathogens and their soluble products, activated lymphocytes differentiate into cells with different functions. B lymphocytes produce antibodies which can neutralize toxins and opsonize extracellular pathogens either via Fc receptors on phagocytic cells or by activating the complement pathway. T lymphocytes recognize intracellular pathogens. As cytotoxic cells (8.2.5) they destroy cells in which pathogens proliferate and as so called T helper cells they support the antibody production by B lymphocytes and assist macrophages in the elimination of internalized pathogens. A subset of T cells controls the extent of the specific immune response (regulatory T cells).
- After the first contact with a pathogen a specific memory is established resulting in a more efficient reactivity at later encounters with the pathogen (see Fig. 8.1-1).
- In general, the specific immune response is not directed against self antigens (autotolerance). As a rule, a primary specific immune response is induced by dendritic cells (DC) only when they are activated by 'danger motifs' of pathogenic origin (see above). Quiescent DC presenting self antigens cannot activate lymphocytes. A second reason for autotolerance is the negative selection of autoreactive T cells during their differentiation in the thymus.

8.1.3 Development and Maturation of the Cellular Components (Fig. 8.1-5)

The mobile cells of the immune system develop from common precursor cells in the bone marrow, the pluripotent stem cells. Colony stimulating factors, CSF (growth and differentiation factors) provided by bone marrow stromal cells, influence the formation of colony forming

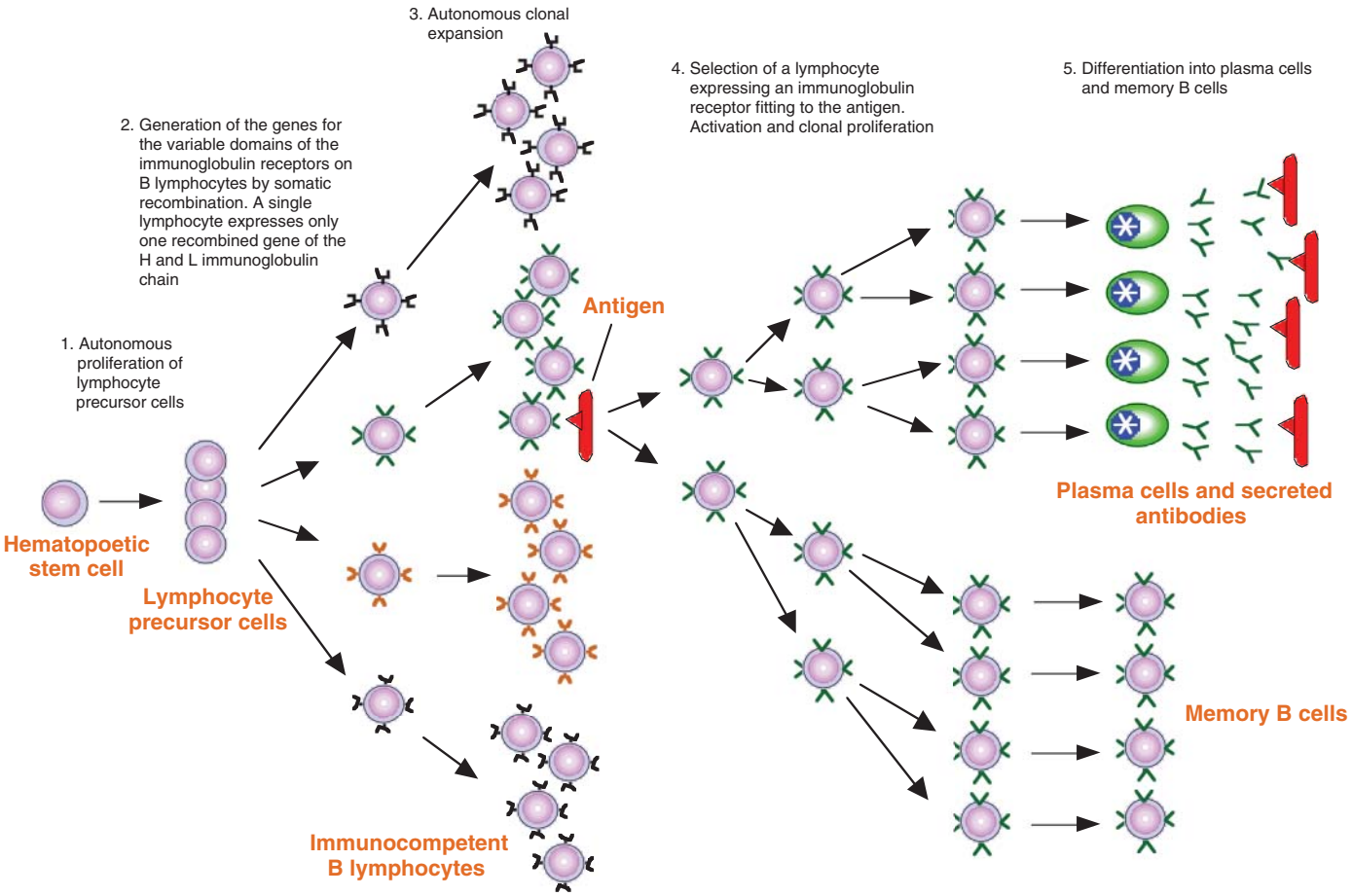


Figure 8.1-4. Clonal Selection

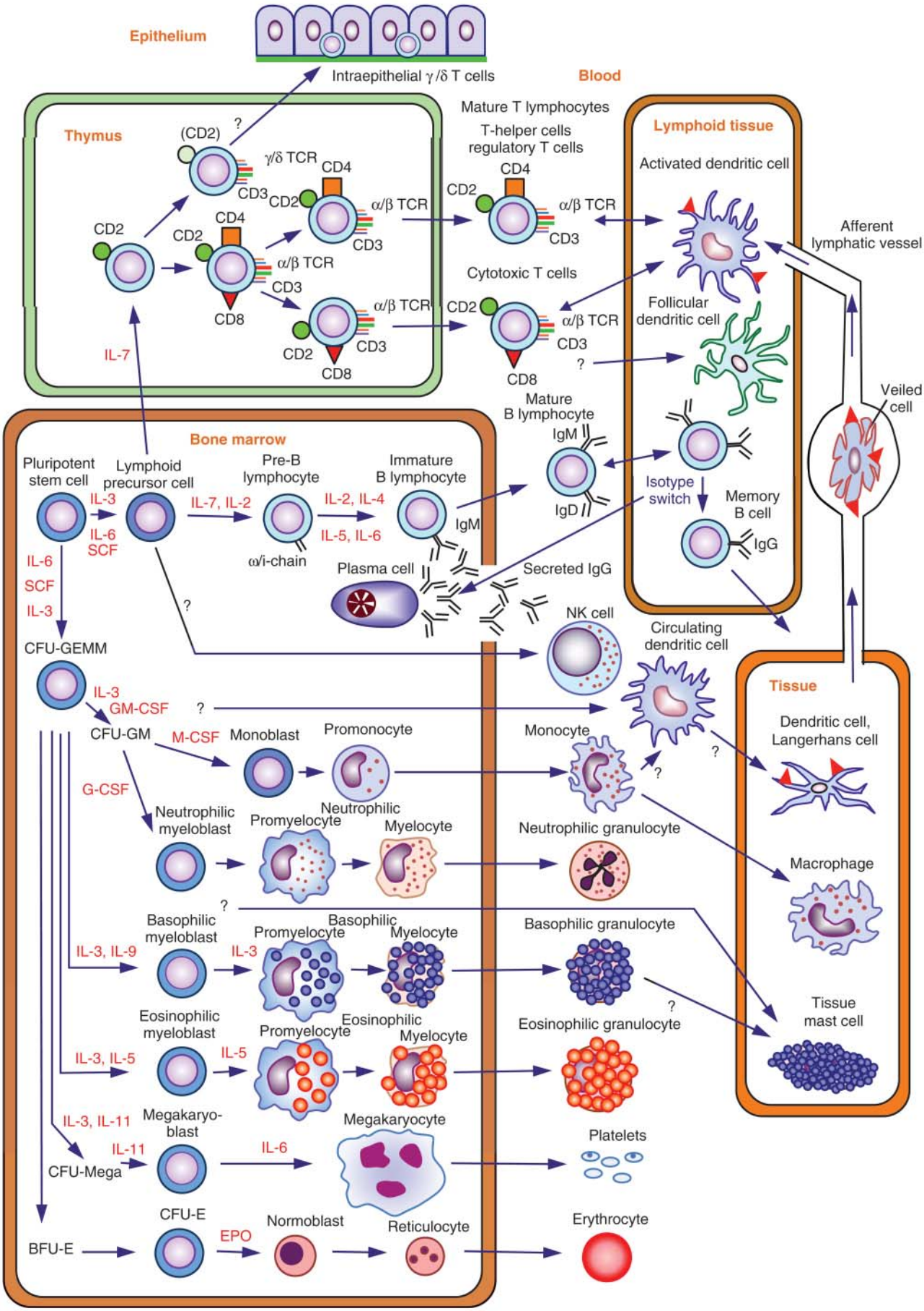


Figure 8.1-5. Development of Blood Cells

units, CFU, which are precursors of the various differentiation lines. They are named by the initial letter of the final cells. There are four principal differentiation lines:

1. the myelomonocytic line that gives raise to monocytes/macrophages, neutrophilic, basophilic and eosinophilic granulocytes and dendritic cells;
2. the line ending up with the formation of blood platelets;
3. the line resulting in red blood cells (erythrocytes);
4. the lymphoid differentiation pathway as the source of B and T lymphocytes and most probably NK cells.

After maturation in the bone marrow, the hematopoetic cells migrate into the circulatory system. As an exception, the T lymphocytes leave the bone marrow as immature cells and migrate into the thymus for further differentiation. The various differentiation lines and steps of leukocytes are characterized by expression of a certain spectrum of surface molecules, which can be detected with monoclonal antibodies and are designated with CD-numbers (cluster of differentiation). These molecules serve specific functions in cell-cell contact and as receptors in the intercellular communication and antigen recognition.

T lymphocytes are characterized by the CD2 molecule, which appears early during the T cell maturation in the thymic cortex. There most T cells also express simultaneously the CD4 and CD8 coreceptor molecules and the α/β T cell receptor chains, which are associated with the CD3 polypeptide chains. Before leaving the thymic cortex, separation into two different lines takes place. They carry either the CD4 molecule (T helper cells and regulatory T cells) or the CD8 molecule (mainly cytotoxic T cells). A third population maturing in the thymus expresses the γ/δ T cell receptor (γ/δ T cells). These cells preferentially settle as cells of the first line defense in the epithelia of the skin and of the gastrointestinal, respiratory and urogenital tract.

B lymphocytes mature in the bone marrow. The various differentiation steps are characterized by the expressed isotypes of immunoglobulins. Mature B lymphocytes carry simultaneously IgM and IgD on their surface. After activation by antigen, a switch to expression of other Ig isotypes (e.g., IgG, IgA, IgE) can take place (see 8.2.4).

8.1.4 Antigen Receptor of B Lymphocytes, Antibodies

Immunoglobulins are the antigen recognition structures of B lymphocytes. They occur either as membrane receptors or as secreted soluble antibodies. Immunoglobulins are composed of heavy (H-, 53 … 75 kDa) and light (L-, 23 kDa) polypeptide chains, which are folded into so-called immunoglobulin domains (Fig. 8.1-6). Disulfide bonds stabilize the domain structure or connect the polypeptide chains. Each domain of ca. 100 amino acids forms two β-pleated sheets (Fig. 2.3-2) consisting of 3 … 5 antiparallel β-strands, which are stabilized by a disulfide bond. Polypeptide chains, which are folded according to this principle, are the basic structure of a large number of membrane molecules. Their coding genes compose the immunoglobulin supergene family.

Variable (V) domains: In immunoglobulins, the domain closest to the N-terminus of each polypeptide chain contains 3 hypervariable sequences, which are different in individual immunoglobulin

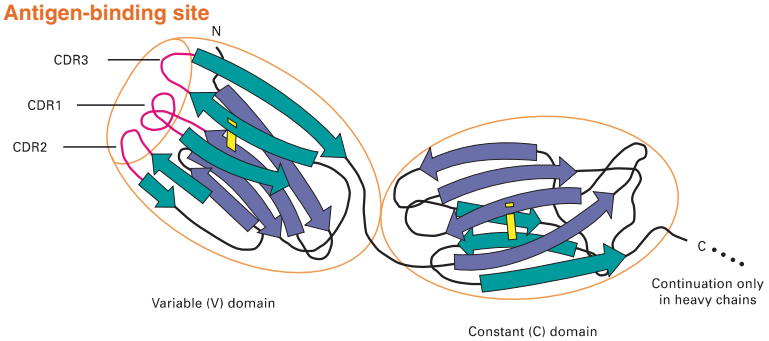


Figure 8.1-6. Structure of Immunoglobulin V and C Domains

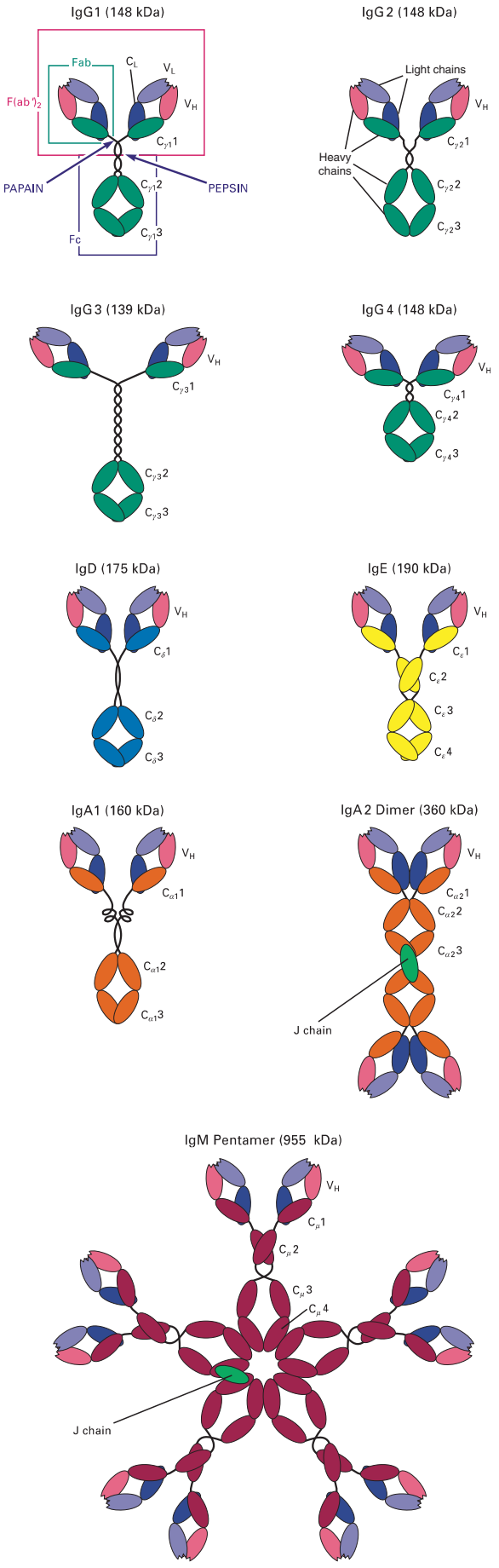


Figure 8.1-7. Human Immunoglobulin Isotypes

molecules. These complementarity determining regions (CDR) of an H and an L chain are close to each other in the completed immunoglobulin molecule and form the specific binding site for an antigen.

Constant (C) domains: The other domains of the Ig polypeptide chains are constant (C_H1 … C_H3 or C_H4, respectively, in heavy chains, C_L in light chains). They determine the biological properties of the Ig molecules and the membership in a particular immunoglobulin class (isotype). While IgG, IgD and IgE are bivalent (possess 2 binding sites), secreted IgM forms oligomers, mainly a decavalent pentamer, and IgA can aggregate to a tetravalent dimer (Fig. 8.1-7). IgM and dimeric IgA are stabilized by an extra covalently bound J chain (ca. 20 kDa). The notation of the isotype heavy chain domains is $C_\mu1$ … 4 for IgM, $C\gamma_1$1 … 3 for IgG1, etc.

Both H polypeptide chains of IgG are connected by disulfide bridges, which are located between the C_H1 and C_H2 domains in the so-called hinge region. This region provides flexibility to the immunoglobulin molecule for bivalent binding to antigenic determinants at variable distances. The size of this region is different in various Ig isotypes (see Fig. 8.1-7). In IgM and IgE heavy chains, the hinge region is replaced by an extra domain. During limited proteolysis of IgG with papain, the H-chains are split at the N-terminal side of the disulfide bridges. This yields Fab-molecules (fragment antigen binding), which contain the V_H and the C_H1 domain of a heavy chain and a complete light chain. Pepsin splits the IgG heavy chains at the C-terminal side of the disulfide bridges, resulting in a $F(ab')_2$ molecule, which is able to bind an antigen bivalently. It does not, however, show the biological properties of the native molecule, such as binding to Fc receptors (Fig. 8.1-2) and activation of complement (8.1.5).

Structure and recombination of immunoglobulin genes: The huge diversity of V domains with different antigen-binding specificity originates

- from somatic recombination of gene segments during B cell development.
- from somatic point mutations in recombined genes following antigen activation of B lymphocytes.

The genes for light chains (κ or λ chain) and for heavy chains are located on different chromosomes. In downstream direction (5′ → 3′, Fig. 8.1-8 for heavy chains) they consist of

- a number of V segments (87 in genes for heavy chains in humans + 24 on 2 other chromosomes, about half of them functional), each preceded by a L (leader peptide) gene segment
- (only in H chain genes): a number of D (diversity) segments (25 in humans)
- a number of short J (joining) segments (9 in genes for heavy chains in humans, 6 of them functional)

  The V gene segments code for the major part of the V domain, while the (D and) J gene segments code for the CDR3 region of the V domain.
- (in κ light chain genes): a number of V segments (40 in humans), a number of J gene segments (5 in humans) and a gene segment coding for the constant domain C_κ.
- (in λ light chain genes): a number of V segments (30 in humans), a number of J gene segments (4 in humans) and several C_λ gene segments, each of them preceded by a single J gene segment

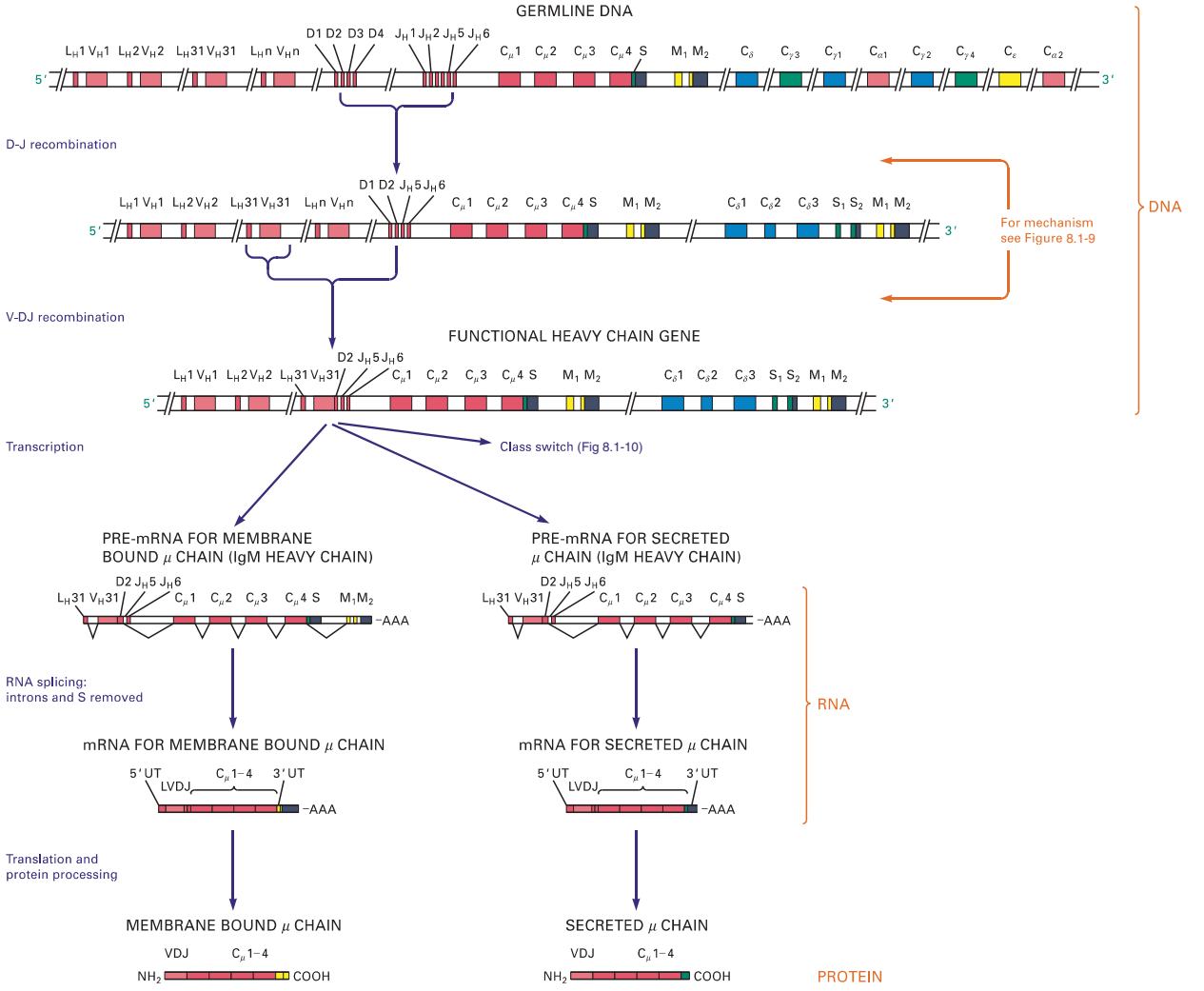


Figure 8.1-8. Generation of a Functional Heavy Chain by Somatic Recombination

– (only in H chain genes): gene segments coding for the constant domains (in 5′-3′ direction) of IgM ($C_\mu 1 \dots 4$), IgD ($C_\delta 1 \dots 3$), IgG3 ($C_{\gamma 3} 1 \dots 3$), IgG1 ($C_{\gamma 1} 1 \dots 3$), IgA1 ($C_{\alpha 1} 1 \dots 3$), IgG2 ($C_{\gamma 2} 1 \dots 3$), IgG4 ($C_{\gamma 4} 1 \dots 3$), IgE ($C_\varepsilon 1 \dots 4$) and IgA2 ($C_{\alpha 2} 1 \dots 3$)

– a sequence in secreted Igs (S)

– the transmembrane and cytoplasmic parts of Ig receptors (M_1, M_2)

Generation of diversity (Fig. 8.1-8): The recombination of immunoglobulin genes proceeds in a fixed order of events. In a maturing B cell the gene segments coding for the H chain are rearranged on one chromosome at first. One of the D gene segments is combined with one of the J gene segments.

This rearrangement is catalyzed by specific recombinases and is dependent on signal sequences, which are located 5′ and 3′ to each D and 5′ to each J gene segment (Fig. 8.1-9). They consist of highly conserved 7 bp and 9 bp sequences, separated by spacers of 12 bp or 23 bp length, which are not conserved. The signal sequences combine in such a way, that the 12 bp spacer is located on one side, the 23 bp spacer on the other side of a loop (12/23 spacer rule). The recombinase system likely recognizes this palindromic structure, removes the DNA loop and ligates the D and J gene segments.

A sequence modification takes place by the introduction of a variable number of nucleotides between the cuts by the template-independent enzyme terminal deoxynucleotidyl transferase, TdT (so-called N region formation), by removal of nucleotides or by variation in the crossover points.

After the D-J recombination, the DJ segment is combined with a V segment by the same mechanism resulting in a functional heavy chain gene. This is transcribed into a pre-mRNA, which also contains downstream the genes for the constant domains of the μ chain (in case of IgM; for other isotypes see below). After splicing, the mRNA is translated into the μ polypeptide chain. The functional gene for the light chain is formed analogously.

If the recombination of the heavy chain gene segments results in a non-functioning gene, e.g., by generation of a stop codon, the recombination is repeated at the other chromosome. With the light chain gene, the V-J recombination takes place at first at the κ chain locus. If this recombination on either chromosome did not result in a functioning gene, then recombination of the V and J gene segments at the λ chain locus takes place. For heavy chains as well as for light chains, the information of only one chromosome is used for production of the immunoglobulin polypeptide chain (‘allelic exclusion’).

After activation of mature B lymphocytes by antigens, many point mutations may occur in the V segments of immunoglobulin genes (somatic hypermutation, several orders of magnitude more frequent than in other genes). This way, the affinity of antibodies for a certain antigen can increase (or decrease) during the immune response (maturation of affinity) (see 8.2-4).

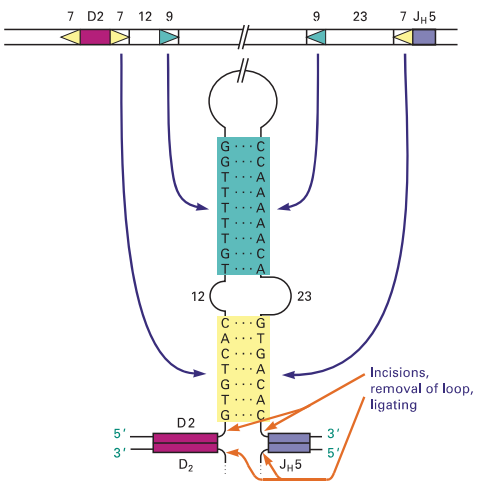


Figure 8.1-9. D-J Gene Recombination by Enzyme-Dependent Looping Out of Intervening DNA

Thus, the huge diversity of binding specificity by the V domains of antibodies is achieved by the following mechanisms:

- Selection of a single gene segment out of each of the V, D, J gene complexes (which yields many different VDJ combinations),
- modification of the nucleotide sequence in the recombination area (junctional diversity),
- combination of various H and L chains. The V domains of both contribute to the structure of the antigen binding site.
- Somatic point mutations in the V segments of immunoglobulin genes occur after activation of mature B-lymphocytes.
- Membrane-bound and soluble immunoglobulins: Immunoglobulins exist either as membrane-based receptors or as soluble antibody molecules.

The DNA for the heavy chain of immunoglobulin M (Fig. 8.1-8) contains also:

- two gene segments (M_1, M_2) coding for the transmembrane region and the short cytoplasmatic portion of the H-chain,
- one gene segment (S) coding for the hydrophilic C-terminal region of the soluble form of the H-chain.

After transcription, the mRNAs for both types of expression are provided by different splicing mechanisms. Either the S segment or the M segments are removed, yielding, after translation, membrane bound or soluble proteins respectively.

Classes of immunoglobulins: Immunocompetent B lymphocytes, which have not yet been specifically activated by antigen, simultaneously carry both IgM and IgD on their surface.

- **IgM**: The reactions leading to IgM molecules have been described above.
- **IgD** (Fig. 8.1-10, left): The connection of the recombined VDJ sequence with the sequences for the constant domains of the δ heavy chain (C_δ) takes place post-transcriptionally by differential splicing of the long pre-mRNA. Hereby, the VDJ segment of the RNA becomes linked with the 3 RNA segments coding for the constant domains of the δ chain. The intermediate segments coding for the constant domains of the μ chain are eliminated.
- **IgG, IgA, IgE**: The genes for these immunoglobulin isotypes are generated by further DNA recombination procedures at the H-chain locus (Fig. 8.1-10 right, shows the case of the γ1 chain gene). The recombined VDJ gene segment for the V_H domain, which is at first connected with the gene segments for the constant domains of the μ chain ($C_\mu 1 \dots 4$, coding for IgM), gets connected with the gene segments coding for the constant domains of other H chain isotypes ($C_{\gamma 1 \dots 4} 1 \dots 3$ for IgG′s, $C_{\alpha 1 \dots 2} 1 \dots 3$ for IgA′s, $C_\varepsilon 1 \dots 4$ for IgE). These class switch recombinations (CSR) are directed by repetitive, GC-rich intron sequences (switch or S regions), which are located 5′ of each H chain gene segment, except the δ gene segments. The DNA between S_μ (as donor sequence) and $S_{\gamma 1 \dots 4}$, $S_{\alpha 1 \dots 2}$ or S_ε (as acceptor sequences) forms a loop which is cut out (looping-out deletion). CSR is preceded by the transcription of a short exon (I exon) located 5′ to the S intron sequence together with the exons coding for the constant domains of a heavy chain (‘switch transcript’, e.g., Ig1 transcript). Due to several stop codons this transcript is not translated into a functional protein, rather it is assumed to direct the switch recombination machinery. The switch transcript is induced by transcription factors binding to the promoter region of the I exon. Through this, the switch recombination is regulated by a certain combination of signals. The switch to IgE, for example, is activated by signals generated by the CD40 receptor and the transcription factor STAT6 which is activated by the cytokines IL-4 and IL-13. The functional heavy chain gene is then transcribed into a primary RNA, which is processed into mRNA and translated into a polypeptide chain (4.2.1, 4.2.3).

CSR and somatic hypermutation (8.2.4) in the Ig-V genes depend on a functional ‘activation-induced cytidine deaminase’ (AID). This enzyme

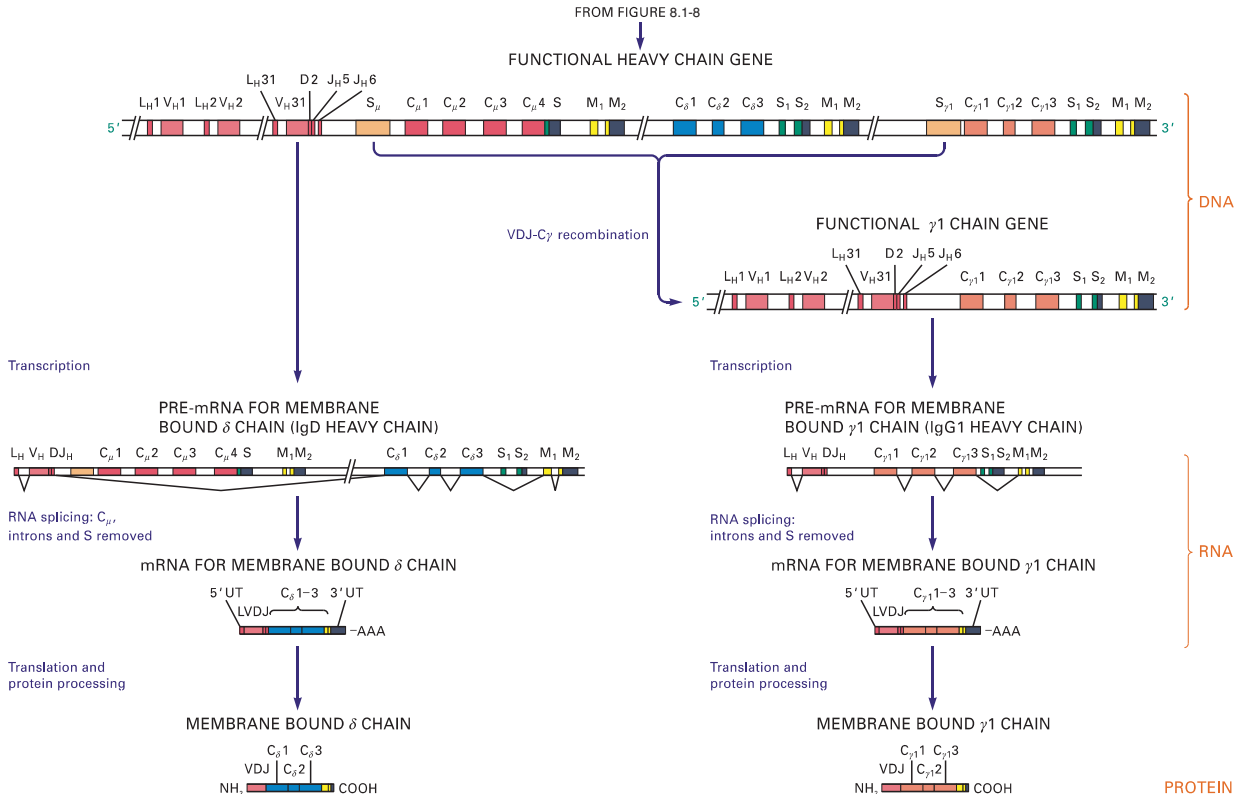


Figure 8.1-10. Class Switch by Different Splicing of Long Transcripts (Left) or by DNA Recombination (Right)

is specifically expressed in activated B lymphocytes in germinal centers. It deaminates cytidine residues to uridine only in single-stranded DNA. The required unwinding of the DNA occurs during immunoglobulin gene transcription. The base excision repair enzyme uracil-DNA glycosylase (UNG, 3.9.2.3) removes the pyrimidine base and the 'apurinic/apyrimidinic endonuclease 1′ (APE1) excises the ribose forming a nick in the single stranded DNA. The different enzyme activities can result in point mutations in the V genes, whereas double-strand breaks following single-strand nicks are essential steps in CSR.

Antigen receptor complex on B lymphocytes (Fig. 8.1-11): The antigen receptor complex on the surface of B lymphocytes contains the membrane form of an immunoglobulin and a heterodimer of the polypeptide chains Igα (CD79a, not to be confused with IgA!) and Igβ (CD79b), which are responsible for the signal transduction leading to gene activation after antigen binding (7.5.4). They contain in their cytoplasmic part a conserved sequence called 'immunoreceptor tyrosine-based activation motif' (ITAM). It is composed of two YXXL/I motifs separated by 6-9 amino acids. The tyrosine residues are phosphorylated by Src-family tyrosine kinases following interaction of the receptor with its ligand. They serve as binding site for the SH2 domains of downstream tyrosine kinases (see Fig. 8.2-3). Closely associated with the B-cell receptor complex is the co-receptor molecule CD19. On its cytoplasmatic side, it is associated with a tyrosine kinase of the Src family (Table 7.5-5).

If the antigen exists as an immune complex with C3d molecules (8.1.5), its binding to the B-cell surface via the immunoglobulin receptor can be augmented by the complement receptor 2 (CD21).

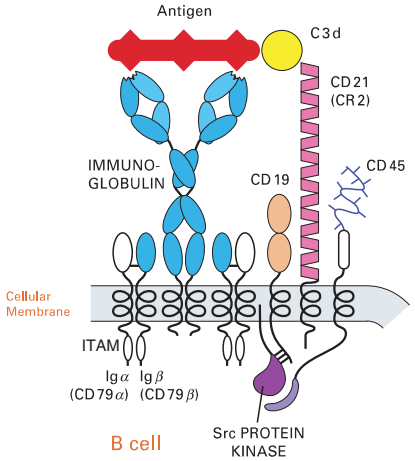


Figure 8.1-11. Antigen Receptor Complex on B Lymphocytes

Biological Function of Human Immunoglobulin Isotypes (Table 8.1-4)

- **IgM** is the first immunoglobulin produced by B lymphocytes in the course of an antigen-specific immune response. Secreted IgM forms pentamers that are stabilized by a J chain and has 10 antigen binding sites (see Fig. 8.1-7). Multimeric binding results in a high avidity interaction of the IgM molecule with the antigen even when the affinity of a single antigen-binding site is low ('bonus effect' of multimeric binding). This is particularly important for the so called 'natural antibodies'. These are IgM molecules which circulate at a low concentration in the body before the onset of a specific immune response. Their binding sites recognize a broad spectrum of antigens, although with low affinity. They essentially contribute to the first line defense against infectious agents such as bacteria. IgM has an outstanding capacity to activate complement (see 8.1.5).
- **IgD** is expressed together with IgM as antigen receptor on naive B lymphocytes (μ^+/δ^+ B lymphocytes).
- **IgG** has the highest serum concentration of all immunoglobulin classes (see Table 8.1-4). IgG antibodies can opsonize antigenic particles by binding to Fc receptors on macrophages (see Fig. 8.1-2). When bound to antigens on the surface of a cell IgG antibodies can induce the 'antibody-dependent cellular cytotoxicity', ADCC, by activation of NK cells via Fc receptors (see 8.2.5). The IgG subclasses have different capacities to activate complement (IgG3 >> IgG1 >> IgG2).

Table 8.1-4. Human Immunoglobulin Isotypes

Isotype	Mol. W. (kDa)	Mean concentrat. in serum (mg/ml)	Half-life in serum (days)	Complement activation	Opsonization ADCC	Mast cell activation
IgM	970	1.5	8	+++	–	–
IgD	180	0.03	3	–	–	–
IgG1	150	9	21	++	+++	+
IgG2	150	3	20	+	(+)	(+)
IgG3	170	1	7	+++	++	(+)
IgG4	150	0.5	21	+	+	(+)
IgA1	160	3	6		–	–
IgA2	160	0.5	6		–	–
IgE	190	0.0001	3	–	–	+++

- **IgA** has two subclasses, IgA1 and IgA2. Soluble IgA can form dimers that are stabilized by a J-chain (see Fig. 8.1-7). An important role in the defense against pathogens is played by the 'secretory IgA'. It is produced as a dimer by plasma cells of the lamina propria of the mucosa. It binds to a poly-Ig receptor (receptor for polymerized Ig) which consists of 5 Ig-like domains and is expressed by the epithelial cells of the mucosa at their basolateral membrane. Bound to this receptor the IgA is internalized, transcytosed and expressed at the apical cell membrane. There, the poly-Ig receptor is cleaved off; its five domains remain attached to the IgA dimer as the so called secretory piece. It protects the secreted IgA from proteases derived from bacteria or contained in the mucus. Secretory IgA efficiently inhibits invasion of pathogens such as viruses and bacteria by preventing them from contact with epithelial cells.
- **IgE** circulates in blood and tissue at a very low concentration (its concentration in serum is at least 10.000 times lower than that of IgG). Yet, it can efficiently arm basophilic granulocytes, mast cells and eosinophilic granulocytes because it binds with high affinity to the Fcε receptor on these cells. IgE antibodies are principally involved in the defense against intestinal parasites such as helminths. IgE-armed eosinophilic granulocytes can destroy parasites. IgE-coated mast cells reacting with parasite-derived antigens immediately release mediators like histamine which stimulate contractions of the intestinal smooth muscles resulting in the expulsion of the parasites. IgE is the antibody responsible for the immediate type hypersensitivity (allergic reaction) (see 8.3.1).

Literature to Sections 8.1.1–8.1.4:

Beutler, B.A. Blood 2009;113:1399–1407.

Cerutti, A. Nature Rev. Immunol. 2008;8:421–434.

Chaudhury, J., Alt, F.W. Nature Rev. Immunol. 2004;4:541–552.

Dale, D.C., Boxer, L., Liles, W.C. Blood 2008;112:935–945.

Grawunder, U., West, R.B., Lieber, M.R.: Curr. Opin. Immunol. 1998;10:172–180.

Muramatsu, M., Kinoshita, K. *et al*. Cell 2000;102:553–563.

Murphy, K., Travers, P., Walport, M. *Janeway's Immunobiology,* Garland Science, 7th Ed., 2008.

Sansonetti, P.J. Nature Immunology 2006;7:1237–1242.

Zhu ,J., Emerson, S.G. Oncogene 2002;21:3295–3313.

8.1.5 Complement System (Fig. 8.1-12)

Complement is a system of plasma proteins in vertebrates that received its name from its property to interact with bound antibodies and to 'complement' the antibacterial activity of antibodies. Three pathways can activate the effector systems of complement:

- the phylogenetically older alternative pathway (part of the innate immune system)
- the mannan-binding lectin pathway (part of the innate immune system)
- the classical pathway (part of the adaptive humoral immune response).

The main effects of complement are:

- enhancement of phagocytosis of pathogens ('opsonization'),
- induction of inflammation, activation of inflammatory cells,
- lysis of pathogens and cells,
- solubilization of immune complexes.

The activation of the complement factors proceeds via a proteolytic cascade (similar to 9.1), each member cleaves and activates in this way the next member of the sequence. The activated factors join the complex of their predecessors at the pathogen surface. Except for the classical complement factors C4, C2 and C3, the numbering reflects the sequence of action. Added lower-case letters (e.g., C4a) indicate cleavage products.

Classical pathway: The classical complement pathway is activated by IgM or IgG. A prerequisite for activation is the binding of the antibodies to antigenic determinants at the surface of a pathogen. The C1 component is a complex consisting of the C1q molecule and two of each of the serine proteases C1r and C1s. The activation is initiated by the binding of at least two of the six globular heads of the C1q molecule to the Fc domains of bound IgM or IgG. Whereas a single pentameric IgM molecule can provide the two interaction sites for C1q, at least two bound IgG molecules at a distance of less than 40 nm are required. This means that many IgG molecules must be bound to the surface of a pathogen in order to activate the classical complement pathway whereas considerably fewer IgM molecules are required. The binding of C1q leads to autoactivation of C1r which cleaves and activates the associated C1s molecule. The activated serine protease C1s cleaves C4 resulting in the release of C4a and the exposition of an internal thioester bond at C4b which is either hydrolyzed (inactivation) or mediates covalent binding of C4b to a nearby hydroxyl or amino group exposed at the surface of the pathogen or on a bound antibody. After attachment, C4b binds C2 which then becomes susceptible to cleavage by C1s, resulting in the release of C2a and activation of the serine protease C2b.

The C4b,2b complex is the classical pathway C3 convertase which cleaves large numbers of C3 molecules into C3b and the proinflammatory fragment C3a (anaphylotoxin). C3b can covalently bind as a potent opsonin to the surface of the pathogen by the same mechanism as C4b. The complex C4b,2b,3b is the classical pathway C5 convertase which then cleaves C5 into C5b and the proinflammatory fragment C5a (anaphylotoxin).

Lectin pathway: The mannan-binding lectin (MBL) is a collectin (8.1.1.1) circulating as a complex with the 'MBL-associated serine proteases 1 and 2' (MASP1 and MASP2). When the complex interacts with mannose residues on a pathogen, the serine proteases are activated and cleave C4 and C2 resulting in the C3 convertase C4b,2b. The following steps are identical to the classical pathway.

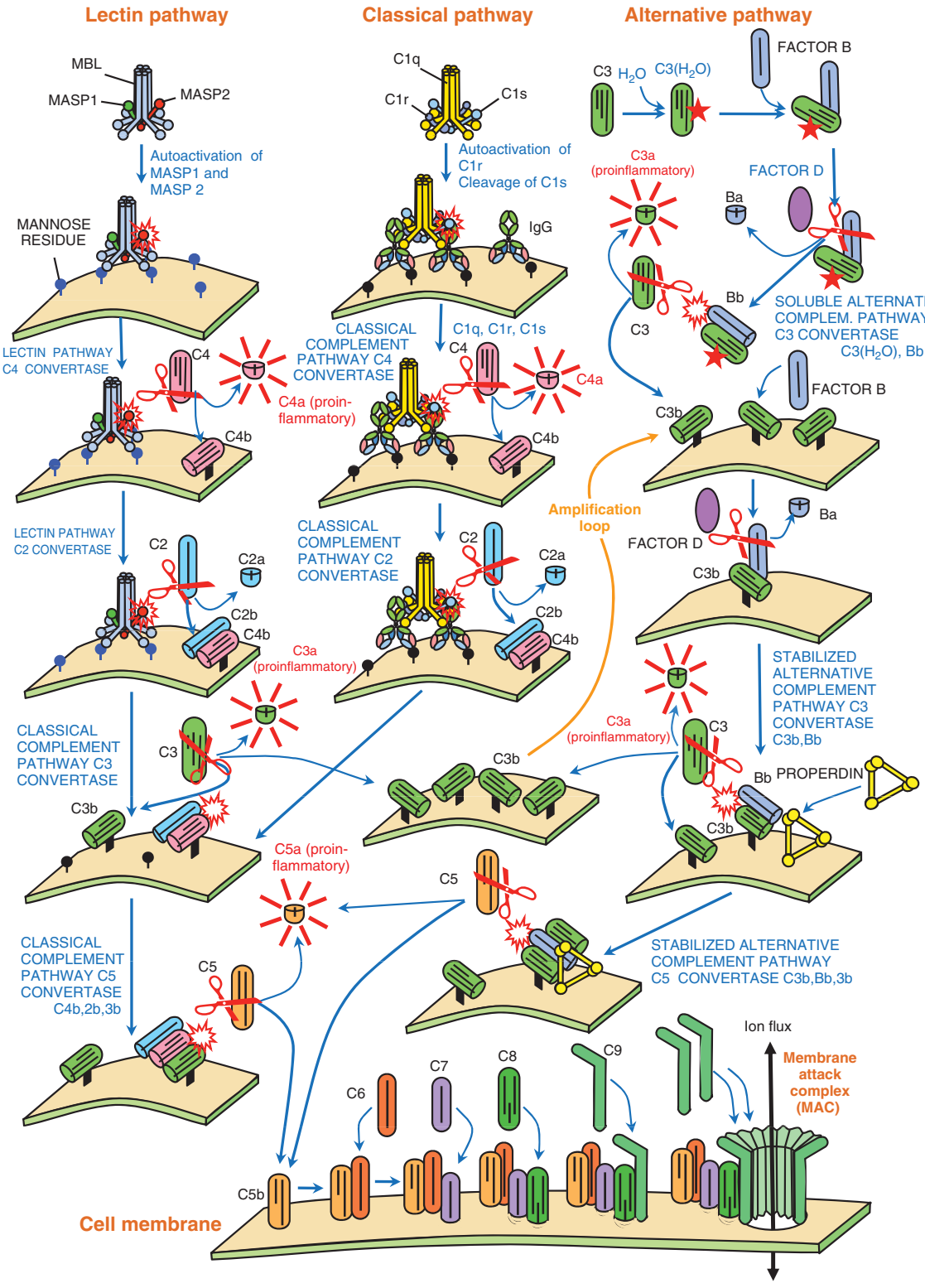


Figure 8.1-12. Pathways of the Complement Activation

Alternative pathway: The alternative pathway starts either with the hydrolysis of C3 in the fluid phase or with C3b bound to the surface of a pathogen. The thioester bond of C3 is continuously hydrolysed at a low level to form C3(H_2O) ('tick over'), which is able to bind factor B. Factor B in the complex is then cleaved by factor D into Ba and Bb. The resulting complex C3 (H_2O),Bb is the alternative (fluid-phase) C3 convertase. Also C3b generated by the classical C3 convertase C4b,2b and bound to the surface of a pathogen can attach Factor B which is then cleaved into Bb and Ba by factor D resulting in the formation of the C3b,Bb complex (alternative pathway C3 convertase). This convertase acts on another C3 molecule, leading to the release of the anaphylatoxin C3a and its association with the C3b formed. This results in the complex C3b,Bb,C3b (alternative pathway C5 convertase) which then cleaves C5. The activity of the C3 convertase C3b,Bb generates extra C3b, which saturates rapidly the surface of the pathogen and is the seed of new C3b,Bb complexes. This mechanism is called 'amplification loop'.

Homology between classical and alternative pathway (Table 8.1-5): There is a striking homology between both pathways, starting from covalent binding of C4b and C3b respectively, and leading to the formation of C5 convertase.

Table 8.1-5. Homology Between the Classical and Alternative Pathway

Steps	Classical Pathway	Alternative Pathway
Deposition of covalently bound complement protein	C4b	C3b
followed by binding of further complement protein	C4b, 2	C3b, B
This complex is cleaved by a serine protease	C1s	D
resulting in C3 convertase	C4b, 2b	C3b, Bb
Binding of C3 takes place	C4b, 2b, 3	C3b, Bb, 3
Activation, formation of C5 convertase	C4b, 2b, 3b	C3b, Bb, 3b

Formation of the 'membrane-attack complex' (MAC), lysis of pathogens and cells (Fig. 8.1-12): The cleavage of C5 by the classical or alternative C5 convertase initiates the assembly of the terminal complement components to form the membrane-attack complex C5b … 9. Binding of C6 to C5b enables the binding of C7. This exposes a hydrophobic site on C7 leading to insertion in the lipid bilayer of the membrane. C8 binds to C5b … 7 on membranes and induces the polymerization of approx. 10 to 16 C9 molecules which form a transmembrane ring-shaped structure with a hydrophilic inner channel (about 10 nm diameter). This channel allows free passage of water and other small molecules across the lipid bilayer, leading to cell lysis.

Other effects of the complement system:

- Solubilization of immune complexes: Insoluble antigen-antibody complexes in the blood plasma can be solubilized by activation of the classical complement pathway and the covalent deposition of C4b and C3b. Via the deposited C4b and C3b immune complexes can bind to the complement receptor CR1 on erythrocytes. They are transported into the liver and spleen where they are removed from the erythrocytes and taken up by macrophages.
- Opsonization of pathogens: The deposition of covalently bound C3b and, to a much lesser extent, C4b at the surface of pathogens is very important. This plays a major role in their uptake and destruction by phagocytic cells. Pathogen-attached C3b (C4b) and some of their cleavage products bind specifically to the receptors present on various phagocytic cell types (opsonin activity, 8.1.4, Table 8.1-6). This interaction between opsonins and their receptors is a prerequisite for efficient engulfment and uptake of pathogens. However, additional signals are required to initiate phagocytosis. In addition, binding of immune complexes via complement receptors to the surface of follicular dendritic cells in lymphoid organs results in long lasting antigen deposition for continuous B cell stimulation (8.2.4).
- Induction of inflammation: The C5a, C3a and C4a fragments (also called 'araphylotoxins') which are released during the complement activation steps directly induce a local inflammatory reaction, including smooth muscle contraction, increased vascular permeability and chemotactic recruitment of inflammatory cells such as neutrophilic granulocytes and dendritic cells into tissues. The increase in tissue fluid enhances the movement of pathogens to local lymph nodes. In addition, C5a (the most active one) acts on mast cells, which are induced to release a number of inflammatory mediators (see Fig. 8.3-1).

Control mechanism of the complement system: Inherent amplification mechanisms of the complement system require a tight control in solution and on host cells. The activated components are rapidly inactivated unless they bind to a surface. Additionally, complement control proteins exist in plasma and as membrane proteins. If, e.g., the C3/C5 convertases (C4b,C2b; C4b,C2b,C3b or C3b,Bb; C3b,Bb,C3b, respectively) are attached to host cells instead of pathogens, they become inactivated by the control proteins listed in Table 8.1-7.

Medical aspects: Deficiency of C1INH causes hereditary angioneurotic edema. A glycophosphoinositol tail attaches DAF and CD59 to the cellular membrane. If this cellular fixation mechanism is defective, paroxysmal nocturnal hemoglobinuria occurs due to erythrocyte lysis. Epstein-Barr virus binds to CR2 expressed at B cells as part of the infection cycle, resulting in infectious mononucleosis.

8.1.6 Antigen Receptor of T Lymphocytes

While immunoglobulins specifically interact with a wide range of different structures on the surface of soluble and insoluble molecules [e.g., linear peptide sequences, conformation determinants, small haptene molecules and carbohydrate structures (4.4.1)], T lymphocytes mainly recognize, with their antigen receptors, complexes of small peptides and proteins coded for by the Major Histocompatibility Complex (MHC molecules, Fig. 8.1-14) at the surface of nucleated cells. These peptides originate from intracellular processing of intracellular or of exogenous proteins (see below).

Structure of the T cell receptor (TCR, Fig. 8.1-13): The antigen receptor of the majority of T cells is a covalently bound heterodimer of an α- (40 … 60 kDa) and a β-chain (40 … 50 kDa) constituting the TCR2. A subpopulation of T lymphocytes expresses a receptor dimer consisting of γ- and δ-chains (TCR1) instead (so-called γ/δ T cells). Each chain comprises a variable domain and a constant domain next to the membrane. They are structured analogously to the immunoglobulin domains (Fig. 8.1-6). The variable domains each contain 3 complementarity-determining regions (CDR), which are responsible for the binding specificity of the T cell receptor.
The genes for the four different T cell receptor chains are located at different chromosomes. Analogously to immunoglobulins, the functional

Table 8.1-6. Complement Receptors

Receptor	CD Nr.	Specificity for	Present at	Function
CR1	35	C3b, C4b	erythrocytes, macrophages, monocytes, neutrophils, B cells	stimulation of phagocytosis, decay of C3b and C4b, transportation of immune complexes (erythrocytes)
CR2	21	C3d, C3dg, iC3b, EBV	B cells	EBV receptor, B cell stimulation
CR3	11b/18	iC3b	macrophages, monocytes, neutrophils	stimulation of phagocytosis
CR4	11c/18	iC3b	macrophages, monocytes, neutrophils	stimulation of phagocytosis
C1q Rec.		C1q	macrophages, monocytes, B cells, platelets	binding of immune complexes to phagocytic cells

Table 8.1-7. Complement Control Proteins

Factors	Function
Plasma Proteins	
C1 Inhibitor (C1INH)	binds to activated C1r and C1s, leading to displacement from C1q (limits duration of activation)
C4 binding protein (C4bp)	binds to C4b, displacement of C2b, cofactor for C4b cleavage by factor I
Factor H	binds to C3b, displacement of Bb, cofactor for C3b cleavage by factor I
Factor I	cleavage of C3b (cofactors H, MCP or CR1)
Membrane Proteins	
Complement-receptor 1 (CR1, CD 35, see above)	binds to C4b, displacement of C2b, binds to C3b, displacement of Bb, cofactor for factor I
Decay-accelerating factor (DAF, CD 55)	displacement of Bb from C3b, displacement of C2b from C4b
Membrane cofactor protein (MCP)	cofactor for C4b and C3b cleavage by factor I
CD 59	prevention of MAC formation on homologous cells

genes for the V domains are formed by somatic recombination of V, J and D segments (in case of β and δ chains) or of V and J segments (in case of α and γ chains). The same mechanisms as in immunoglobulins (8.1.4) are involved in the generation of the multiple receptor specificities. In contrast to immunoglobulin receptors, frequent somatic point mutations have not been observed with T cell receptors.

Like the B cell receptor, the T cell receptor is associated with polypeptide chains (Fig. 7.5-6), which are primarily responsible for signal transduction (Fig. 8.2-1) and are required for the expression of the TCR in the cytoplasmic membrane. These chains form the CD3 complex, which is composed of homo- or heterodimers of δ, ε, γ, ζ and η chains (Fig. 8.1-13). They contain in their cytoplasmic tail one or more ITAM sequences. In a wider sense, CD4 and CD8 molecules are also considered as part of the T cell receptor complex. They contribute as co-receptors to the activation of T cells after recognition of antigens (see below).

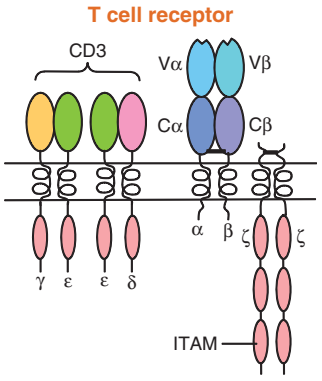


Figure 8.1-13. T Cell Receptor Complex

8.1.7 Antigen Presentation by MHC Molecules

In addition to the antigen receptors of B and T lymphocytes, the adaptive immune system comprises as third antigen-binding structure the MHC molecules (major histocompatibility complex). In humans, they are named HLA molecules (human leukocyte antigen). Two types of MHC molecules are involved in antigen-binding (Fig. 8.1-14):

- MHC Class I molecules consist of an α-chain with 3 domains (ca. 44 kDa) and of β_2 microglobulin (β_2m, 12 kDa), which is not covalently bound to the α-chain and serves to stabilize the class I molecule. MHC Class I molecules are expressed on almost every nucleated cell of the organism. The antigens presented by them are recognized by $CD8^+$ T lymphocytes (cytotoxic T cells, Fig. 8.2-1).
- MHC Class II molecules comprise an α- and a β-chain (33 and 28 kDa, with 2 domains each), which are not covalently linked. They occur primarily on cells, which present extracellular protein antigens to the T lymphocytes, such as dendritic cells, monocytes / macrophages and B lymphocytes (see below). The antigens presented are recognized by $CD4^+$ T lymphocytes (helper cells, Fig. 8.2-1).

The domains of the MHC molecules, which are adjacent to the membrane and the β_2 microglobulin domain show a structure similar to immunoglobulins (Figs. 8.1-6, 8.1-7). The sections distant to the membrane form a groove with a β-pleated sheet structure at the bottom and an α-helix at each wall. The antigen peptides are bound within the groove (Figs. 8.1-15, 8.1-16 and 8.1-18). The specificity of the binding is determined by the side chains of the amino acids composing the β-pleated sheet and the helices, which interact with the side chains of the peptides. The T-cell receptor binds both to the peptide and to the side chains of the α-helices.

About 200 different genes and pseudogenes are localized in the major histocompatibility complex (Fig. 8.1-17). In humans, it is localized on the short arm of chromosome 6 and is subdivided into 3 regions. Closest to the centromere is the class II region, which contains the genes for the α and β chains of the HLA-DP, DQ and DR molecules. Then follows the class III region, which contains the genes for the complement factors C2, C4 and Bf (8.1.5), the cytokines tumor necrosis factor (TNF), lymphotoxin (LT) and others. Finally there is the class I region, which encompasses the genes for the α chains of the HLA-A, HLA-B and HLA-C molecules as well as for the α chains of the HLA-E, HLA-F and HLA-G molecules. The genes for the MHC chains are exceedingly polymorphic (Fig. 8.1-17). For each of the MHC genes a different number of alleles exist, which are co-dominantly transmitted. They vary mainly in the sequence coding for the peptide-binding groove. Therefore:

- MHC molecules derived from different alleles will bind a given peptide with different affinity

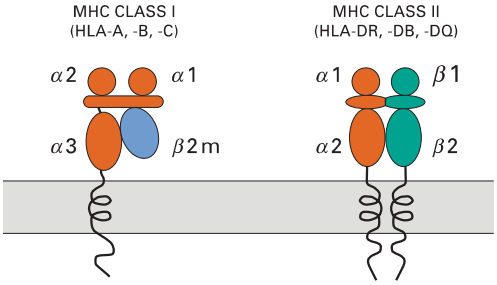


Figure 8.1-14. Structure of MHC Class I and Class II Molecules

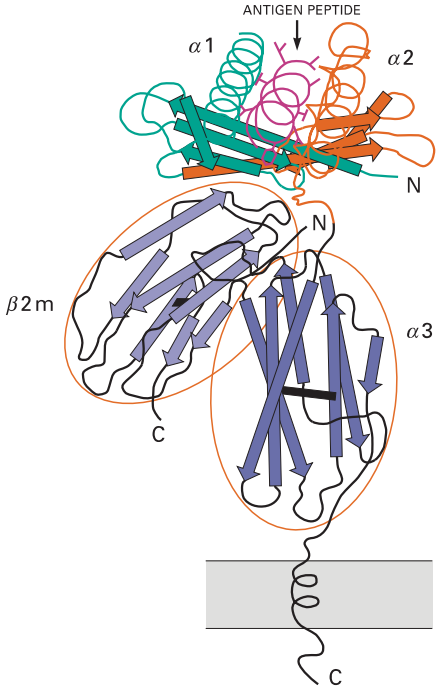


Figure 8.1-15. MHC Class I Molecule with Bound Peptide

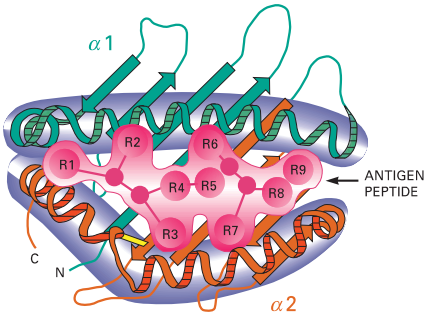


Figure 8.1-16. MHC Class I Molecule with Bound Peptide, Top View

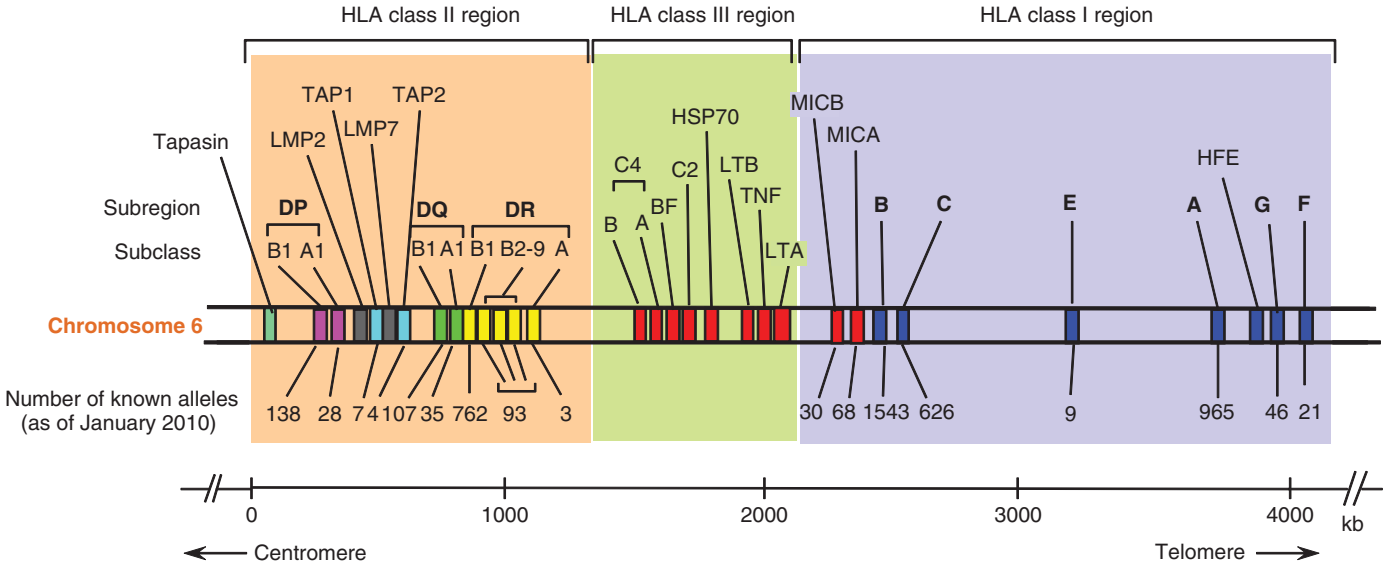


Figure 8.1-17. Human Major Histocompatibility Complex

- antigen-presenting cells from individuals differing in their HLA haplotype will present different sequences from a given protein to the T lymphocytes.

Processing and presentation of antigens for T lymphocytes (Fig. 8.1-18): Class I MHC molecules are charged with peptides immediately after synthesis of the α-chain and their complexing with β_2 microglobulin in the endoplasmic reticulum. Only peptides with a length of 9 … 10 amino acids fit into the groove. They originate mainly from intracellular proteins, which are cleaved into peptides by a protease complex (large multifunctional proteasome, LMP ①, see 4.5.7) present in the cytoplasm. Then they are transported into the endoplasmic reticulum by special transport proteins ("transporter associated with antigen processing", TAP1 and TAP2 ②, where they bind to MHC class I / β2 microglobulin and stabilize the complex. (The genes for LMP and the TAP molecules are also located in the MHC, Fig. 8.1-17). The peptide/MHC complex is then transported via the Golgi apparatus ③ and transport vesicles to the cellular membrane ④. Thus, class I MHC molecules allow T lymphocytes to recognize all intracellularly synthesized proteins in form of peptides. This is of special importance for the recognition of virus-infected cells by cytotoxic ($CD8^+$ T lymphocytes, 8.2.5) which see peptides derived from newly synthesized viral proteins.

The primary task of Class II MHC molecules is the presentation of peptides derived from extracellular proteins. These enter the so-called antigen-presenting cells (APC, e.g., dendritic cells, monocytes/macrophages and B lymphocytes) via phagocytosis or endocytosis ⑦, 6.1.5) and are split into peptides in phagolysosomes/endosomes ⑧, 4.5.6). The fragments bind to class II MHC molecules, which are present at the endosome membrane ⑨. The MHC-peptide complexes are transported to the cellular membrane, where they can be recognized by $CD4^+$ lymphocytes ⑩.

After synthesis in the endoplasmic reticulum, the class II MHC molecules form a complex with the invariant chain Ii ⑤, which prevents the premature binding of peptides by its CLIP segment (class II-associated invariant chain peptide) and directs the transport of the class II MHC molecules into the endosomes. The class II MHC-Ii complex reaches the endosomes via the Golgi apparatus. There, the CLIP segment is split off and the rest of the Ii chain dissociates ⑥. Thus, the groove is free for binding of external peptides. These peptides (15 … 24 amino acids) are longer than peptides bound by class I MHC molecules, since the class II groove is open on both sides.

To ensure that T lymphocytes on the one hand mainly recognize peptides bound to autologous MHC molecules and on the other hand do not react against peptide sequences derived from self proteins they undergo a complex selection procedure during their differentiation in the thymus. As a consequence, mature T cells leaving the thymus and settling in the lymphoid tissues preferentially recognize peptides, which are bound to autologous MHC molecules. Concomitantly, the majority of T cells, which recognize peptide sequences of self proteins, are eliminated.

γ/δ T cells: γ/δ T cells colonize the mucosal epithelia and the epithelium of the skin as sentinel cells at the border to the outside world. Unlike α/β T cells described above they can bind unprocessed antigens such as soluble protein antigens or pyrophosphates and alkylamines exposed on bacterial and eukaryotic cell surfaces. Particularly, they recognize antigens expressed on cells following stress, during infection and inflammation, or after malignant transformation like heat shock proteins and the MHC-class-I-like antigens MICA and MICB (see 8.2-5). Since they do not require antigen processing they can react much faster than α/β T cells and can thus essentially contribute to the first line defense.

8.1.8 Cytokines, Chemokines and Receptors (Fig. 8.1-19)

Cytokines comprise a great number of proteins and glycoproteins, which, being mainly soluble messengers, are responsible for the intercellular communication. They bind to specific receptors at the cell surface and can induce or inhibit the activation of genes. In this way, they influence growth, differentiation and activation of cells. 'Chemokines' are a subgroup of cytokines which control the migration of mobile cells. Here only those cytokines are discussed that are known to be particularly important for the functions of the immune system.

Special properties of cytokines are

- a particular cytokine can act on different cells and hereby exert different functions (pleotropism)
- several cytokines can effect the same results (redundancy)
- The cytokine effects can be autocrine (acting on the producing cell), paracrine (acting on cells in immediate contact to the producing cell) or endocrine (acting on target cells after reaching them via the circulatory system).

With respect to basic structure, cytokines are subdivided into three major groups, consisting of four short or four long α-helices or of antiparallel β-strands whereas chemokine molecules contain an α-helix and three β-strands and are stabilized by two disulfide bonds.

Cytokine receptors are also subdivided according to their structure:

- Type I or hematopoietin receptors, also called the cytokine receptor superfamily, are the largest group. They consist generally of one or two polypeptide chains which are responsible for the specificity of cytokine binding and of another chain which effects the signal transduction. After binding of their ligand, they associate with the 'common' (c, since it is used by several receptors) signal

transduction sequence. Frequently this chain starts with tyrosine kinases of the Jak type (17.5.4).

- Type II or interferon receptors also consist of 2 polypeptide chains 8.5.4, (Fig. 7.5-5).
- Type III or TNF receptors contain the cystine-rich sequences of the nerve growth factor receptor family. They form trimers after binding of cytokines (see Fig. 8.2-7).
- Type IV or Ig superfamily receptors contain cytoplasmic sequences with tyrosine kinase activity in a number of cases (7.5.3).
- TGF (transforming growth factor) family receptors contain cytoplasmic sequences with serine-threonine kinase activity. They can associate with proteoglycan molecules for the enhancement of cytokine binding.
- The chemokine receptors possess a structure with seven membrane passes, typical of G-protein coupled receptors (7.4.1).

Immunologically important cytokines are listed in Figure 8.1-20, their effect on the differentiation of hematopoietic cells is shown in Figure 8.1-5.

Chemokines represent a family of small structurally related molecules that are secreted by different cells either constitutively or in the course of an inflammatory reaction. Chemokines produced in the steady state situation regulate the differentiation and maturation of cells and the recirculation of dendritic cells and of lymphocytes, particularly their homing in the secondary lymphoid organs. For this, they are called homeostatic chemokines. Chemokines released during inflammation are called inflammatory chemokines. They determine the constitution of cellular infiltrates in injured and inflamed tissues. The family of chemokines comprises two major and two minor subgroups which are defined by the amino acid sequence in the area of the two N-terminal cysteine residues. CXC chemokines represent the first major subgroup. They have an intervening amino acid between the two cysteine residues, whereas the CC chemokines, the second major subfamily, the cysteine residues are adjacent to each other. The minor subfamilies CX_3C and C chemokines have three intervening amino acids or only one cysteine residue, respectively. The current chemokine nomenclature which replaces a confusing multitude of different acronyms uses CC or CXC followed by L for "ligand" and a number. Accordingly the chemokine receptors are classified by CXC or CC followed by R for receptor and a number. Chemokine receptors belong to the family of seven transmembrane receptors coupled to heterotrimeric G proteins (7.3, 7.4). Immunologically important chemokines and their receptors are listed in Figure 8.1-21. Homeostatic chemokines are highlighted by green color, inflammatory chemokines by red color.

Literature to Sections 8.1.5–8.1.8:

Charo, I.F., Ransohoff, R.M. N. Engl. J. Med. 2006;354:610–621.

Davis, M.M., Bjorkman, P.J. Nature 1988;334:395–402.

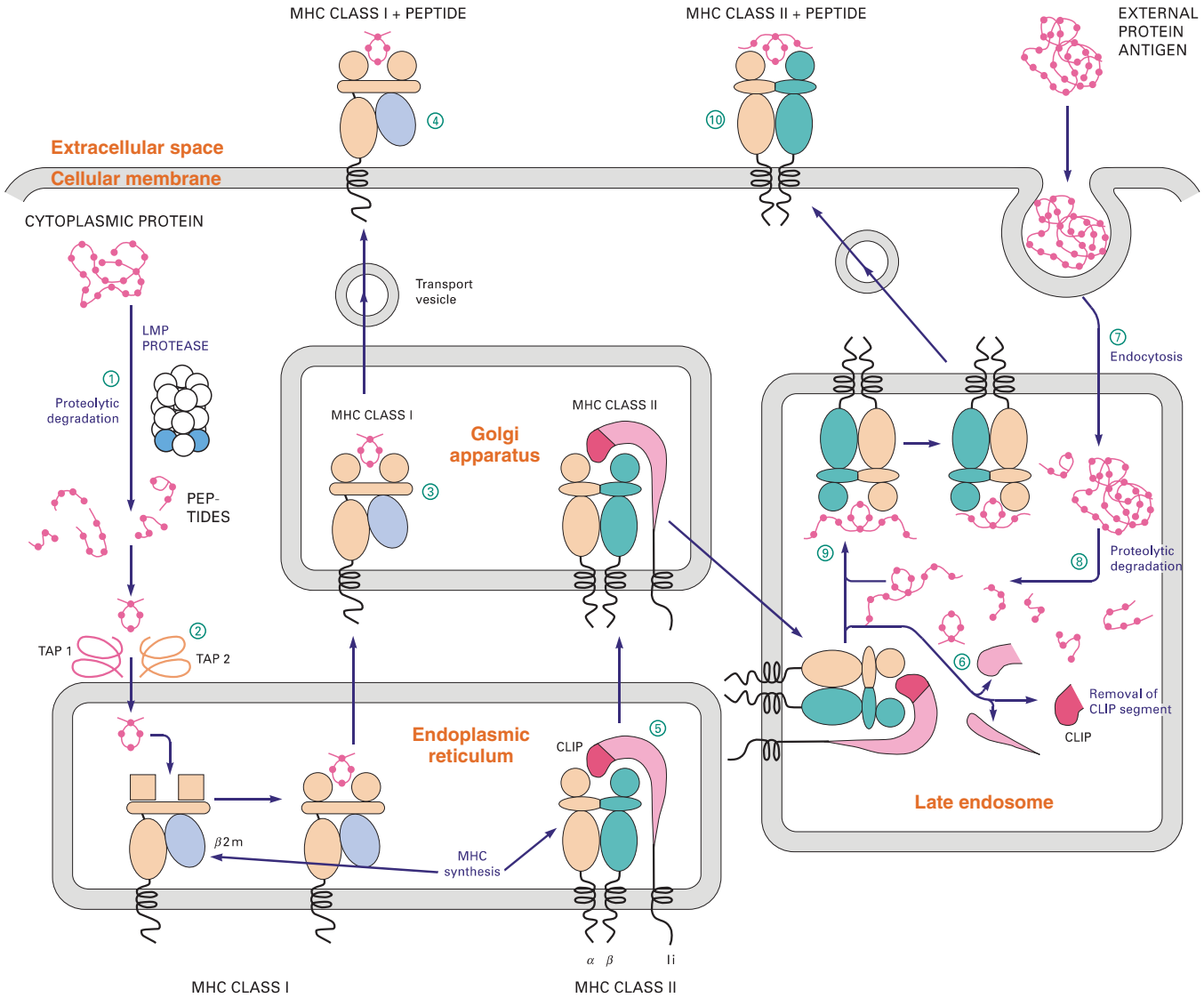


Figure 8.1-18. Antigen Processing and Presentation

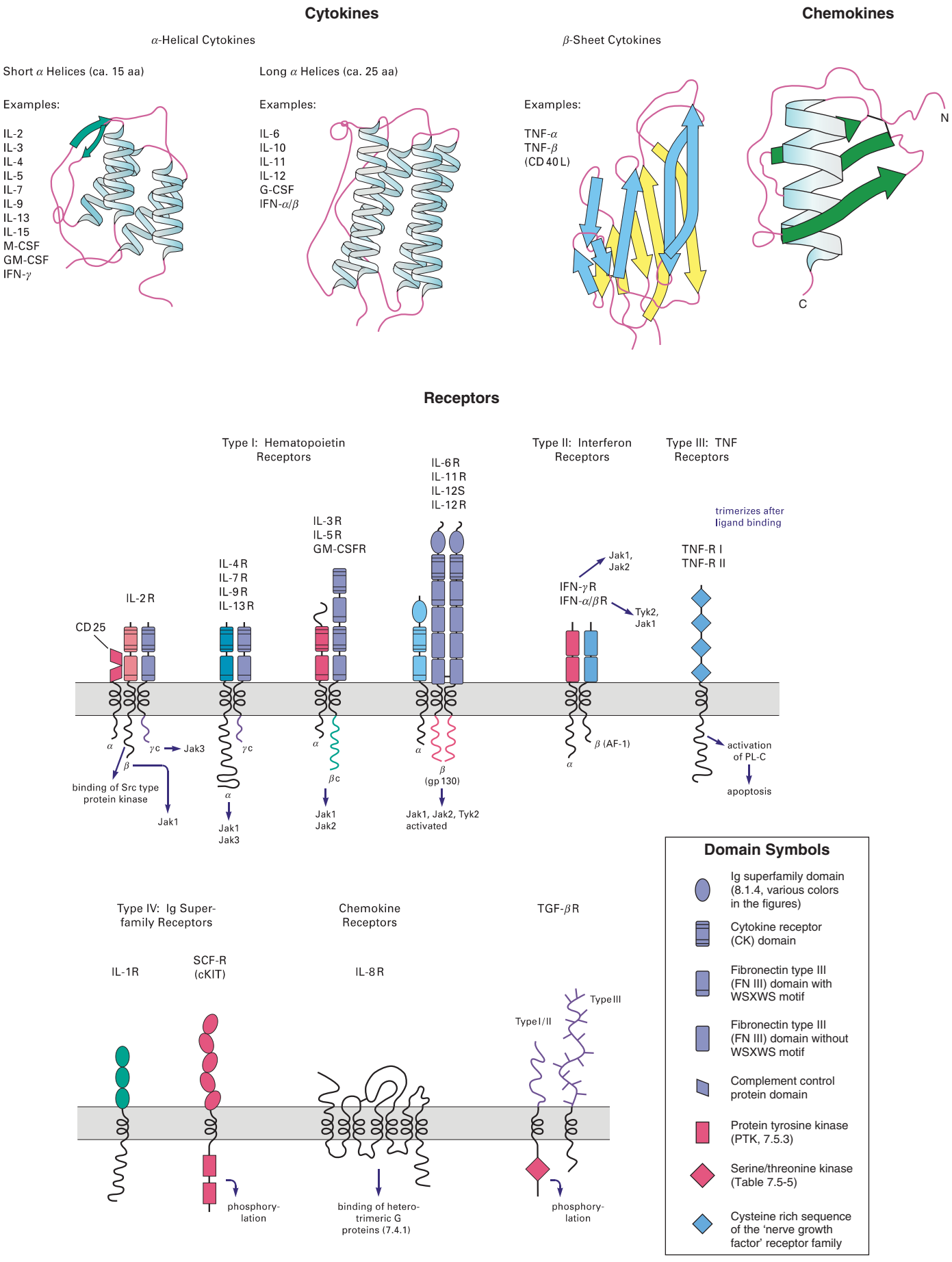


Figure 8.1-19. Cytokines, Chemokines and Their Receptors

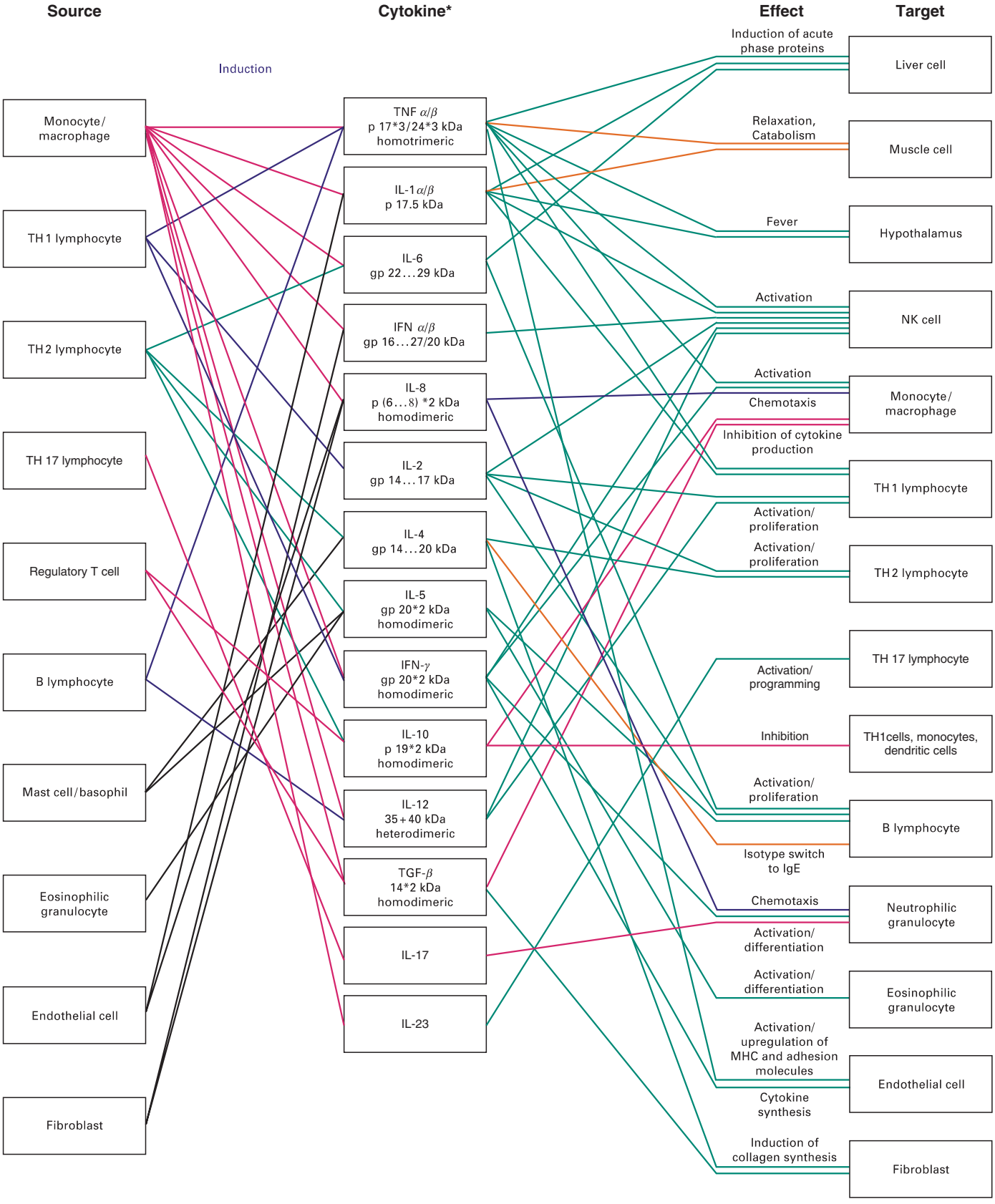
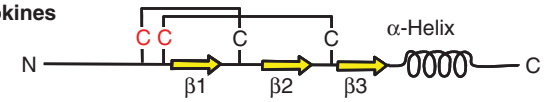


Figure 8.1-20. Immunologically Important Cytokines, Their Sources and Main Effects on Target Cells

Line colors at the left side are for differentiation only
* gp = glycoprotein, p = protein, IFN = interferon, IL = interleukin, TGF = transforming growth factor, TNF = tumor necrosis factor

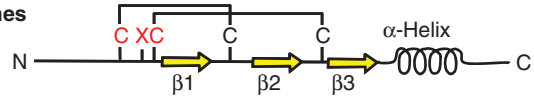
Line colors (right side):
green = activation
red = inhibition
blue = chemotaxis
orange = other

CC-Chemokines



Systematic name	Historical ligand name	Produced by	Effect on N	Effect on Eo	Effect on Mph	Effect on T	Receptor
CCL2	MCP-1,5	Mph, act. E	−	−	++	++	CCR2B
CCL3 / 4	MIP-1 α / β	T, act. Mph	−	+	+	+	CCR1, 5
CCL5	RANTES	Mph, act. E, T	−	+	+	+	CCR1, 3, 5
CCL11	Eotaxin-1	Mph, akt. E, Epi, T	−	++	−	−	CCR3
CCL19	MIP-3 β, ELC	Lymphatic tissue	−	−	−	++	CCR7
CCL20	LARC,MIP-3α	Gut, Epi, lymphatic tissue	immature DC				CCR6
CCL21	SLC	Lymphatic tissue	−	−	−	++	CCR7

CXC-Chemokines



Systematic name	Historical ligand name	Produced by	Effect on N	Effect on Eo	Effect on Mph	Effect on T	Receptor
CXCL1,2,3	Gro-α, β, γ	Mph, act. E	+	−	++	−	CXCR2
CXCL4	PF-4	Megakaryocytes	+	−	+	+	CXCR3B
CXCL8	Interleukin-8	Mph, act. E, N	++	−	−	+	CXCR1, 2
CXCL10	IP-10	Mph, akt. E, N	−	−	−	+	CXCR3
CXCL12	SDF-1α / β	Stroma cells, lymphatic tissue	−	−	−	++	CXCR4
CXCL13	BCA-1	Lymphatic tissue		B Lymphocytes			CXCR5

Mph = monocytes / macrophages
N = neutrophilic granulocytes
T = T lymphocytes

Eo = eosinophilic granulocytes
DC = dendritic cells

Epi = epithelial cells
E = endothelial cells

MCP-1, 5 = monocyte chemoattractant protein-1, -5
MIP = macrophage inflammatory protein
RANTES="regulated upon activation and normal T cell expressed and secreted"
LARC = liver and activation-regulated chemokine

SLC = secondary lymphoid organ chemokine
PF-4 = platelet factor-4
IP-10 = interferon-inducible protein-10
BCA-1 = B cell attractant-1

Figure 8.1-21. Chemokines and Chemokine Receptors

Fernandez, E.J., Lolis, E. Ann. Rev. Pharmacol. Toxicol. 2002;42:469–499.

Fernandez, K.A., O'Neill, L.A.J., Gearing, A.J., Callard, R.E. *The cytokine factsbook* 2nd Ed., London: Academic Press, 2001.

Klein, J., Sato, A. The HLA system. First of two parts. N. Engl. J. Med. 2000;343:702–709.

Murphy, K., Travers, P., Walport, M. *Janeway's Immunobiology*, 7th Ed., Garland Science, 2008.

Ono, S.J., Nakamura, T. *et al*. J. Allergy Clin. Immunol. 2003;111:1185–1199.

Rudolph, M.G., Stanfield, R.L., Wilson, I.A. Ann Rev Immunol. 2006;24:419–66.

Volonakis, J.E., Frank, M.M. *The Human Complement System in Health and Disease*. New York: Dekker, 1998.

Walport, M.J. Complement. First of two parts. N. Engl. J. Med. 2001;344:1058–1066.

Walport, M.J. Complement. Second of two parts. N. Engl. J. Med. 2001;344:1140–1144.

Zipfel, P.F., Skerka C. Complement regulators and inhibitory proteins. Nature Rev. Immunol. 2009;9:729–740.

8.2 Generation of a Specific Immune Response

8.2.1 Activation of T Cells (Fig. 8.2-1)

While the T cell receptors interact mainly with the peptides presented by the MHC molecules (8.1.5) and with polymorphic determinants of the MHC molecules themselves, the co-receptors bind to monomorphic determinants on the MHC molecules. The CD4 co-receptor attaches to the β2 domain of the MHC class II molecule and the CD8 co-receptor to the α3 domain of the MHC class I molecule. Since exclusively either CD4 or CD8 molecules are expressed on mature T lymphocytes, it follows that:

- CD4⁺ T lymphocytes are activated by antigen peptides, which are presented by MHC class II molecules.
- CD8⁺ T lymphocytes recognize peptides, which are bound to MHC class I molecules.

Some proteins are able to simultaneously bind to the β_2 domain of class II MHC molecules and the Vβ domain of the TCR2, thus linking the TCR to the class II MHC molecule (Fig. 8.2-2). As a consequence these proteins can even activate at low concentrations a large number of T lymphocytes independently of their receptor specificity. The polyclonally activated T cells secrete large amounts of cytokines leading to clinical symptoms of acute intoxication. The proteins are

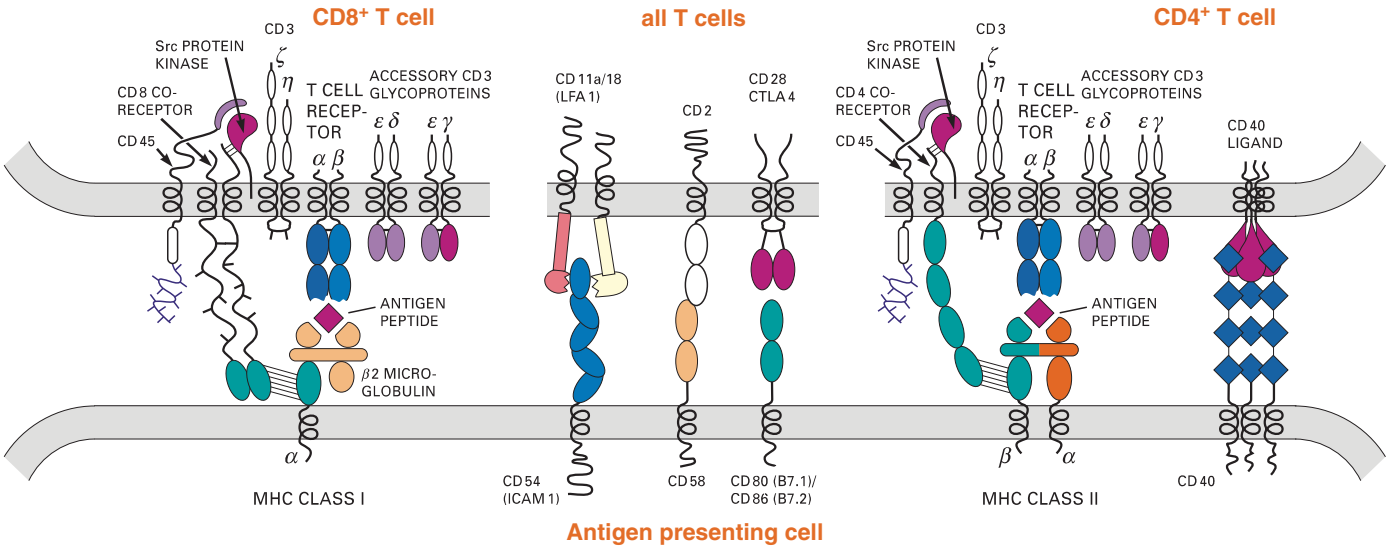
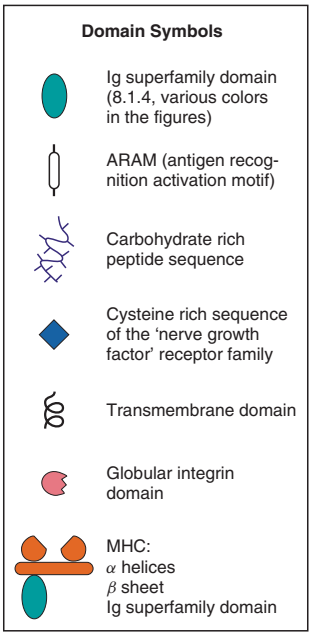


Figure 8.2-1. Membrane Molecules Involved in the Activation of T Lymphocytes

Table 8.2-1. Membrane Molecules Involved in the Activation of T Lymphocytes

Name	Structure	Molecular mass (kDa)	Family [1]	Function
T cell receptor	dimeric (αβ)	α: 50, β: 45	Ig	Binds to peptides presented by MHC class I or II molecules
CD2 (LFA-2)	monomeric	45 … 58	Ig	Cell adhesion molecule, binds to CD58 (LFA-3)
CD3	dimeric (γ, δ, ε, ζ, η)	γ: 25 ... 28, δ: 20, ε: 20, ζ: 16, η: 22	Ig	Associated with T cell receptor, required for its expression, signal transduction
CD4	monomeric	55	Ig	Co-receptor for MHC class II molecules, associates with Lck (p56) protein kinase (7.5.4). Receptor for HIV-1 and -2
CD8	dimeric (αα or αβ)	α: 30, β: 32	Ig	Co-receptor for MHC class I molecules, associates with Lck (p56) protein kinase (7.5.4)
CD11a (LFA-1)	assoc. with CD18	180	integrin	Cell adhesion molecule, binds to CD50 (ICAM-3), CD54 (ICAM-1), CD102 (ICAM-2)
CD18	assoc. with CD11	95	integrin	Associates with CD11a (= LFA-1), CD11b and CD11c
CD28	homodimeric	44 × 2	Ig	Receptor for CD80 (B7.1) and for CD86 (B7.2)
CTLA-4	homodimeric		Ig	Receptor for CD80 (B7.1) and for CD86 (B7.2)
CD 40 ligand (CD154)	trimeric	39	TNF	Binds to CD40
CD40 (gp50)	mono-/trimeric	40	NGFR	Receptor for CD40 ligand (7.5.4), trimerizes after ligand binding, mediates co-stimulatory effects
CD45 (LCA)		180 … 240		Contains phosphotyrosine phosphatase at cytoplasmic side, activates Lck (p56) protein kinase (7.5.4).
CD54 (ICAM-1)		85 … 110	Ig	Cell adhesion molecule, binds to CD11a/CD18 (= LFA-1) etc.
CD58 (LFA-3)		55 … 70	Ig	Cell adhesion molecule, binds to CD2
CD80 (B7.1)		60	Ig	Ligand to CD28 and CTLA-4, provides co-stimulatory signals for T cells
CD86 (B7.2)		80	Ig	Ligand to CD28 and CTLA-4, provides co-stimulatory signals for T cells

[1] Ig = immunoglobulin superfamily, NGFR = nerve cell growth factor receptor superfamily, TNF = tumor necrosis factor superfamily.



therefore called superantigens. Examples are streptococcal and staphylococcal exotoxins, the 'toxic shock syndrome toxin' (TSST) from *Staphylococcus aureus* and proteins derived from viruses and mycoplasms.

The cytoplasmic section of the CD4 and CD8 co-receptors are associated with a tyrosine kinase of the Src family (7.5.4). This kinase is activated by dephosphorylation, which is catalyzed by the cytoplasmic section of the CD45 molecule. CD45 is expressed in all nucleated hematopoietic cells, it occurs on various cells in different isoforms.

Also, other membrane molecules take part in the activation of T lymphocytes (Fig. 8.2-1, center). For the primary antigen-specific activation of T cells, two signals are necessary. While the binding of the α/β TCR chains to the MHC/peptide complex generates the first signal, the second signal is contributed by interaction of the CD28 molecules at the T lymphocyte surface with the B7.1 (CD80) and B7.2 (CD86) molecules, respectively, at the surface of activated antigen-presenting cells [APC, mainly dendritic cells (DC)]. Therefore, they are named co-stimulatory molecules. According to the 'two signal theory' of T cell activation T cells require both signals mediated by the T cell receptor and by CD28 (Fig. 8.2-3). In the absence of the co-stimulatory signal T cells become anergic or die by apoptosis.

Another molecule of the CD28 family is the 'cytotoxic T cell antigen 4' (CTLA4) which is expressed on T cells after activation with some delay. In contrast to CD28 it generates inhibitory signals and thus contributes to the control of the immune response (see 8.2.7). DCs express co-stimulatory molecules when they are activated through pattern recognition receptors (PRR) such as Toll-like receptors (TLR) (see 8.1.1).

Of importance is the enhancement of the cellular contact between T lymphocytes and antigen presenting cells (Fig. 8.2-1, center). This is achieved by:

- binding of the LFA-1 molecule (an integrin molecule, consisting of the CD11a and CD18 chains) to the ICAM-1 (CD54) molecule on the APC
- interaction of the CD2 molecule with the CD58 molecule on the APC.

Simultaneous signaling through the TCR and the co-stimulatory receptor leads to activation of the genes for interleukin-2 (IL-2) and for the α-chain of the IL-2 receptor (CD25, Fig. 8.2-3). The major signal pathways leading to IL-2 gene activation are described in detail in 7.5.4 and Figure 7.5.6. Association of the IL-2 receptor α-chain with the constitutively expressed receptor β and γ_c chains confers high affinity binding of IL-2 to the receptor. This allows autocrine stimulation of activated T cells and their clonal expansion even at low IL-2 concentrations.

8.2.2 CD4⁺ T Effector Cells, Regulation of the Immune Response (Fig. 8.2-4)

Following antigen-specific activation T lymphocytes proliferate and differentiate into effector T cells. CD8⁺ T cells develop into cytotoxic T cells which are able to destroy, e. g., virus-infected cells expressing viral peptides bound to MHC Class I molecules on their surface. CD4⁺ T cells can give rise to at least four functionally different effector cells depending on the quality of the so-called signal 3 which is mainly provided by cells of the innate immune system at an early stage of the immune response. It consists of a certain pattern of cytokines and of molecular interactions during contact with the antigen-presenting cell. The functional CD4⁺ T cell subsets are defined by the profile of secreted cytokines.

Th1 cells: T helper cells 1 (Th1 cells) develop from antigen-activated naive T cells when stimulated by signal 3 consisting of IFN-γ and IL-12. These cytokines are secreted by dendritic cells in response to contact with pathogen-associated molecular patterns (PAMPs, see 8.1.1.2). Th1 cells have two major functions: one is to support macrophages in eliminating intracellularly growing bacteria such as mycobacteria and to stimulate cytotoxic cells such as CD8⁺ T cells and NK cells for the lysis of virus-infected cells ("cell-mediated immunity"). The other principal function is to stimulate the production of opsonizing and complement-fixing antibodies, particularly IgG, by B lymphocytes (see Fig. 8.2-11). Th1 cells mainly produce the cytokines IL-2, IFN-γ and TNF-α, which activate macrophages, cytotoxic T cells and NK cells. A transcription activator characteristic for Th1 cells is T-bet which is induced by external IFN-γ. It activates the genes for IFN-γ and the signaling subunit of the receptor for IL-12.

Th17 cells: The signal 3 for the differentiation of Th17 cells is provided by IL-6 and TGF-β in the absence of IL-12 and IL-4. They induce the transcription factor $ROR_\gamma T$ which activates the gene for the IL-23 receptor. IL-23 is required for further proliferation and differentiation of Th17 cells. By secreting IL-17 these cells recruit neutrophilic granulocytes in acute infection and essentially contribute to tissue inflammation in various disorders such as autoimmune diseases and asthma.

Th2 cells: The principal signal 3 for the differentiation of T helper cells 2 (Th2 cells) is IL-4. It induces the transcription factor GATA-3 which regulates the production of cytokines characteristic of Th2 cells such as IL-4, IL-5, IL-10 and IL-13. Via this positive feedback loop IL-4 sustains its own production and the differentiation of Th2 cells. The primary source of IL-4 is still under discussion. Possibly, IL-4 and GATA-3 genes are primarily activated by signals provided by dendritic cells such as activation of the Notch receptor on T cells by Notch ligands expressed by dendritic cells. Th2 cells support the production of IgM, IgA and IgE antibodies by B cells.

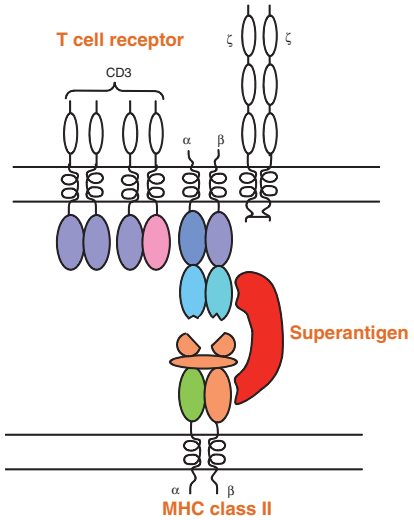


Figure 8.2-2. Superantigen Binding

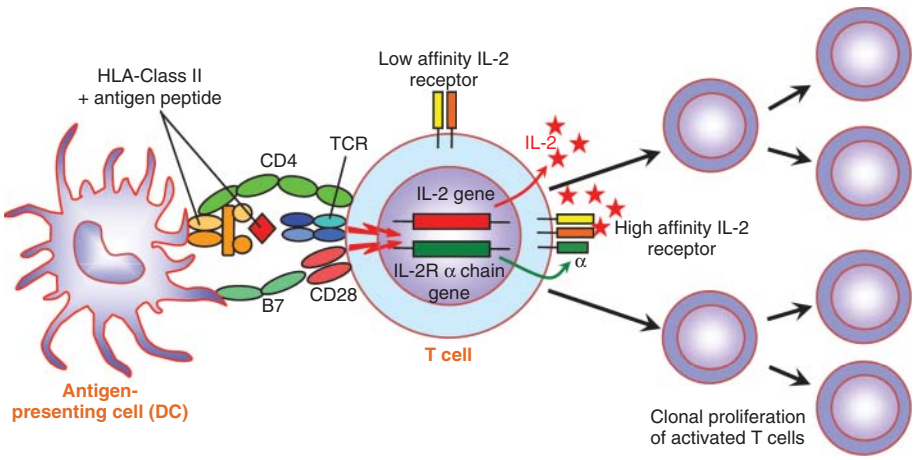


Figure 8.2-3. First Steps of Antigen-Induced Clonal T Cell Proliferation

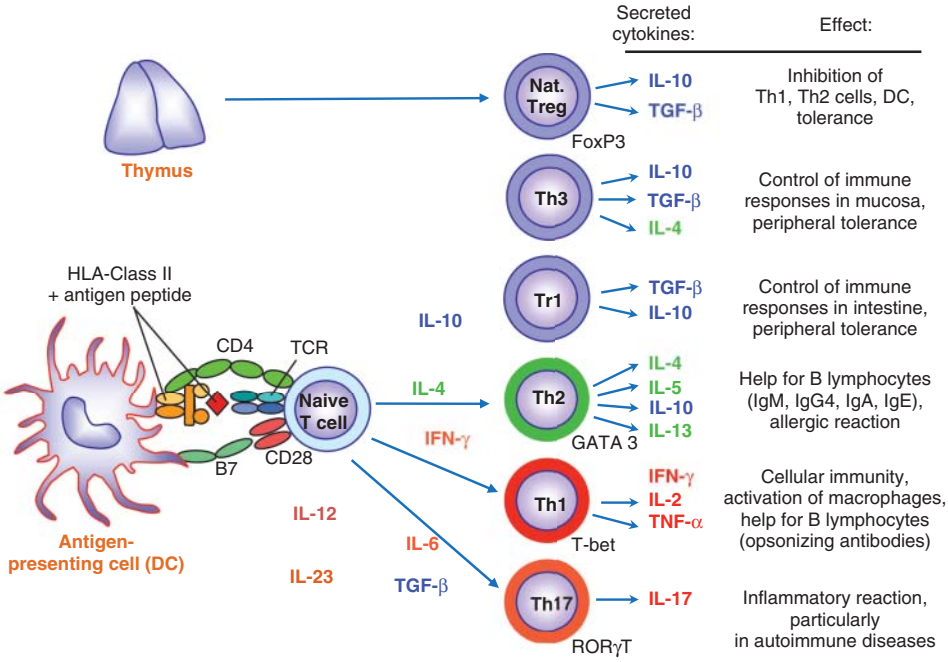


Figure 8.2-4. Generation of Functionally Different CD4⁺ Effector Cells

Regulatory T cells: The CD4⁺ regulatory T cells comprise two major subsets:

1. Natural Treg that are generated during T cell development in the thymus. They express the α chain of the IL-2 receptor (CD25) and the transcription factor FoxP3 which prevents the activation of the IL-2 gene. Natural Treg are committed in the thymus medulla by interaction with dendritic cells activated by thymic stromal lymphopoetin (TSLP). They secrete IL-10 and TGF-β and inhibit proliferation of Th1 and Th2 cells and production of IL-12 by dendritic cells.
2. Adaptive regulatory T cells comprise two subsets: the Th3 cells secreting IL-4, IL-10 and TGF-β are mainly found in the intestinal mucosa, and the Tr1 cells which develop under the influence of IL-10 and produce TGF-β. These cells are also represented in the intestine. Both types of adaptive regulatory T cells control the immune response against ingested antigens and sustain peripheral tolerance.

8.2.3 Activation of B Cells

Analogously to T lymphocytes, usually two signals are required for activation of B lymphocytes, while a third specifies the product:

- Signal 1: binding of antigen to the Ig receptor on their surface.
- Signal 2: interaction of the CD40 receptor with the CD40 ligand, which is expressed on activated T-cells following activation.
- Signal 3: cytokines determining the selection of isotypes.

The activation signal 1 is transduced by phosphorylation of the ITAM motifs of the Igα and Igβ chains associated with the immunoglobulin receptor (see Fig. 8.1-11). The ensuing signal cascade is similar to that in T lymphocytes (see 7.5.4 and Fig. 7.5-7). The co-stimulatory signal 2 is provided by the interaction of the CD40 molecules on B cells and the CD40 ligand (CD40L, CD154) expressed on activated CD4⁺ T helper cells. For this, the T helper cells must come into close contact with the B cells. This is achieved by the binding of the antigen receptor on the T cell to the specific antigen peptide presented by MHC class II molecules on the B lymphocyte with the co-operation of adherence molecules. With their high affinity receptors B lymphocytes can enrich endocytose and process antigens even if these are present only at a low concentration. The signal 3 delivered by cytokines determines which constant heavy chain gene is recombined with a functional VH gene during the class switch recombination.

This coupled, dual recognition of protein antigen molecules by B and T lymphocytes ensures that B lymphocytes produce antibodies only against those antigens which they recognize themselves and whose peptide fragments are recognized by specific T lymphocytes. This is necessary, because B lymphocytes can change the specificity of their receptors via somatic point mutations (8.1.4) and the control of the B cell antibody production by T lymphocytes can prevent the secretion of antibodies against endogenous structures.

Under certain circumstances B lymphocytes are activated and produce antibodies without T cell help. The so-called T independent antigens (e.g., carbohydrates and repetitive determinants) induce only IgM antibodies, since the isotype switch in B cells requires T cell assistance. These antibodies are essential for the defense against bacteria such as Pneumococci that are surrounded by a capsule consisting of polysaccharides that do not provide T cell epitopes.

8.2.4. Lymphocyte Circulation and Generation of Cellular and Humoral Immune Responses in Lymphoid Tissue (Fig. 8.2-6)

The specific adaptive immune response is initiated in the lymphoid tissues which provide the structural prerequisites that T lymphocytes and B lymphocytes need to meet the antigens for which they have specific receptors. Given the low frequency of lymphocytes specifically recognizing a distinct antigen it is necessary that as many cells as possible travel through an area where antigens are concentrated. The most sophisticated structure that has developed to facilitate successful contacts of lymphocytes with their specific antigens is the lymph node (Fig. 8.2-5).

This organ is surrounded by a fibrous capsule that is perforated by numerous afferent lymphatic vessels conducting the lymph from tissues. The lymph enters the subcapsular sinus and reaches the efferent lymphatic through the trabecular and medullary sinuses. The flow rate of the lymph is markedly reduced by the widening of the lymph channels allowing the efficient filtration of the lymph by the numerous phagocytic cells contained in the sinuses. The lymphoid tissue of a lymph node is divided into a cortex and a medulla. The cortex consists of primary and secondary lymph follicles harboring B lymphocytes and the paracortical area where T lymphocytes are concentrated. The medulla is composed of medullary cords containing mainly plasma cells.

Antigens that have entered peripheral tissues reach the draining lymph node either conveyed in soluble or particulate form by the afferent lymph or they are taken up by dendritic cells which then

migrate as so called 'veiled cells' via the afferent lymphatic vessel into the paracortical areas of the lymph node where they present the processed antigen to T cells (Fig. 8.2-6, ①).The migration of lymphocytes and dendritic cells is tightly controlled by chemokines. To reach the paracortical T cell area activated dendritic cells express the chemokine receptor CCR7 ①.

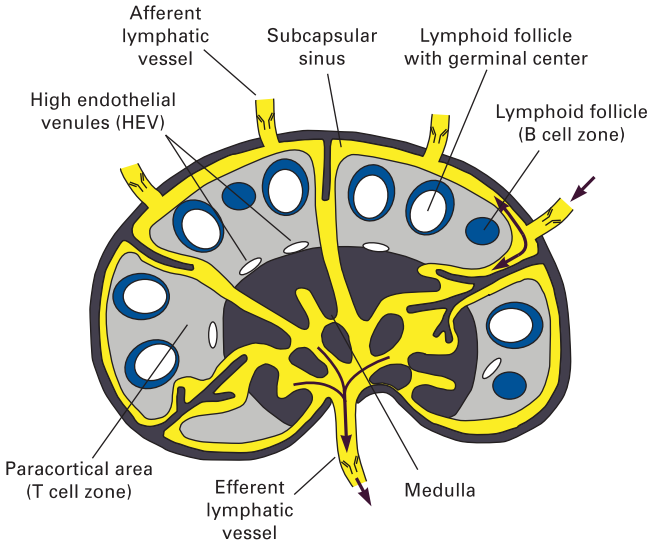
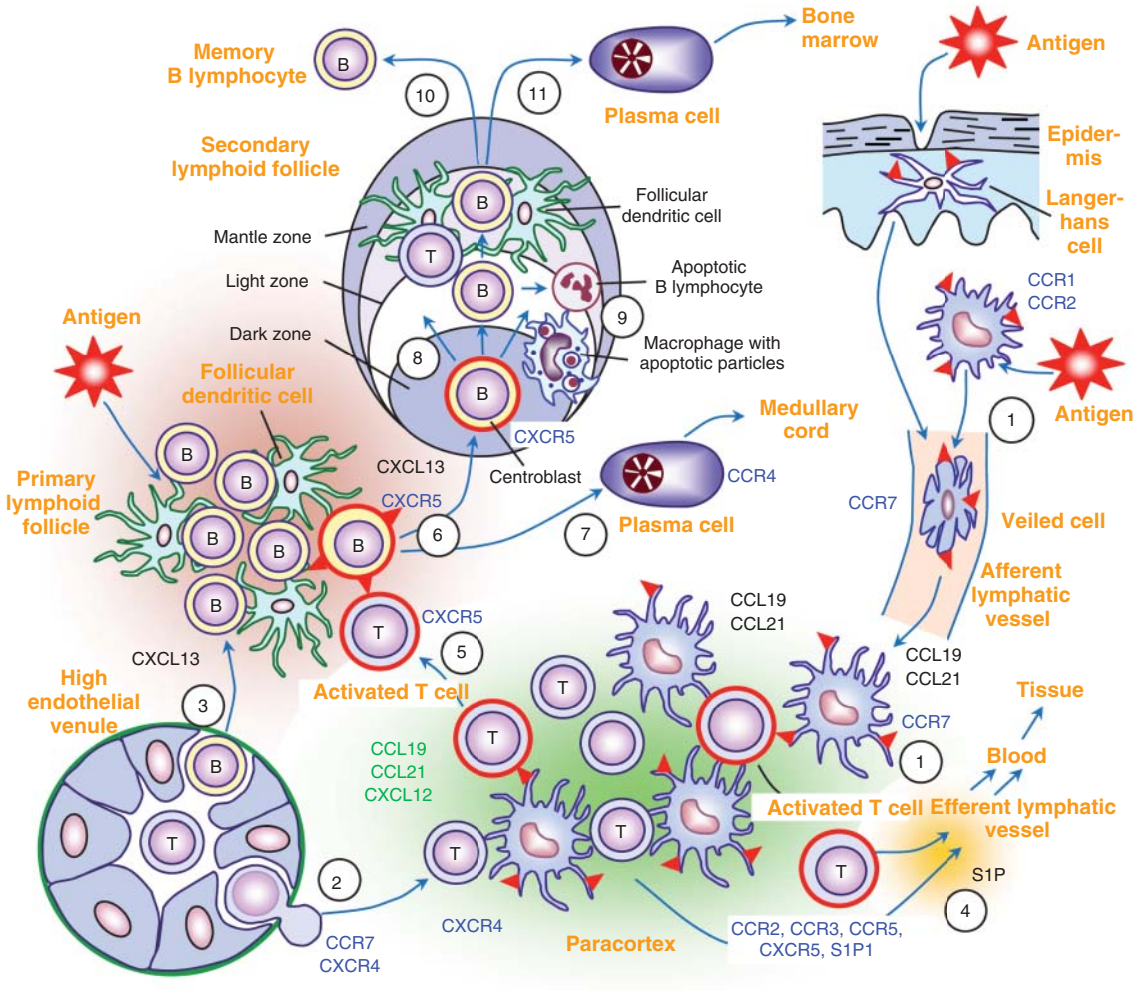


Figure 8.2-5. Schematic Diagram of a Lymph Node

T and B lymphocytes circulating in blood enter the lymph node via the so called high endothelial venules (HEV) at the border between cortex and medulla. The first contact is mediated by homing receptors on the lymphocytes and addressin molecules on the endothelial cells (8.4). Immigration into the lymphoid tissue is directed by chemokines. CCL19 and CCL21 concentrated in the paracortical area attract T lymphocytes expressing the chemokine receptors CCR7 and CXCR4 ② whereas CXCL13 directs CXCR5-expressing B lymphocytes into the primary lymph follicles ③. The high mobility of lymphocytes in the lymphoid tissue increases the chance to meet the fitting antigen. When the antigen-transporting dendritic cells have been activated by pathogen-associated molecular patterns (PAMPs, see 8.1.1) and express the co-stimulatory molecules CD 80/CD 86 (8.2.1), antigen-recognizing T lymphocytes are activated and then proliferate and differentiate into effector cells. They leave the lymph node via the efferent lymphatic vessel to reenter the blood circulation and to immigrate into tissues. The emigration out of the lymph node is governed by a concentration gradient of sphingosine-1-phosphate (S1P) which is recognized by the S1P receptor ($S1P_1$) on T cells ④. Antigen-specific activated T lymphocytes express the chemokine receptor CXCR5 guiding them to the B cell area which has an increased concentration of CXCL13 ⑤. Here they recognize the cognate antigen on B lymphocytes and provide these cells with the activation signal 2 through the CD40L-CD40 interaction (see 8.2.3). To meet the T cells, B lymphocytes on their part upregulate the chemokine receptor CCR7, which guides them towards the T cell area ⑥.



The color densities in the paracortex and the primary lymphoid follicle areas indicate the concentration of the respective chemokines.

Figure 8.2-6. Generation of a Cellular and Humoral Immune Response in Lymphoid Tissue

The activated B cells proliferate and can further differentiate into antibody-producing plasma cells which migrate into the medullary cords ⑦. Most of the proliferating B cells give rise to a secondary lymphoid follicle which is histologically oriented versus the subcapsular sinus and can histologically be subdivided into different zones ⑧. The area of proliferating B cells (centroblasts) oriented towards the medulla is called dark zone, followed by the light zone, both comprising together the germinal center which is surrounded by the mantle zone consisting of small, quiescent B lymphocytes. During their proliferation and differentiation B cells require continuous stimulation by T cells and by antigen which is presented on the follicular dendritic cells (FDC). These cells store antigens bound to complement or Fc receptors on their surface for prolonged periods of time. In this phase somatic hypermutations occur in the genes for the variable domains of the immunoglobulin receptors under the influence of the 'activation-induced (cytidine) deaminase' (AID) (see 8.1.1.4). As a result the affinity of the Ig receptors on a B cell can either increase or decrease. When the affinity is reduced the receptor-induced activation signal for the B cell is insufficient and the cell dies by apoptosis (8.2.5). Apoptotic cell particles are phagocytosed by so called 'tingible body macrophages' ⑨. As a consequence of the diminishing antigen concentration only B cells that have generated high affinity receptors by somatic hypermutation get the activation signals required for further proliferation and differentiation ('affinity maturation'). An important effect of the activation signals provided by T cells is the switch of antibody isotypes in B cells which is also dependent on AID activity (see 8.1.4). The differentiation pathway of B lymphocytes results either in memory B cells ⑩ which leave the lymph node and recirculate as quiescent cells in blood and lymphoid tissues or in plasma cells which migrate into the bone marrow and home to niches where they survive and secrete antibodies for prolonged periods of time (long-lived plasma cells) ⑪.

Cessation of an adaptive immune response: When a pathogen has effectively been eliminated by an adaptive immune response the number of effector cells becomes markedly reduced, mainly for two reasons: a) the specific antigenic stimulus has ceased and the cells die by apoptosis ('death by neglect'). b) Activation-induced cell death (AICD). After antigen-specific activation, T lymphocytes express, with some time lag, negatively regulating receptors of the CD28 family such as CTLA4 (see 8.2.1) that competitively block CD28 activation, and in addition, generate inhibitory signals. Further inhibitory receptors are the receptor 'programmed death-1' (PD-1, CD279), which interacts with the ligands B7-H1 (CD274) and B7-DC (CD273), the Fas receptor (CD95) and the Fas ligand (FasL, CD178) expressed on activated cells. The signal generated by these receptors induces apoptosis resulting in clonal elimination of antigen-specific T cells, a phenomenon that is observed as a decline in frequency of virus-specific cytotoxic T cells following the acute phase of a virus infection. This decreased antigen-specific reactivity of T cells is called 'terminal differentiation' or 'T lymphocyte exhaustion'. Some of the specifically activated lymphocytes survive as memory cells. They are responsible for the immediate and increased memory response following a second encounter with the pathogen.

Literature to Sections 8.2.1–8.2.4:

Alouf, J.E., Muller-Alouf, H. Int. J. Med. Microbiol 2003;292: 429–440.

Bromley, S.K., Mempel, T.R., Luster, A.D. Nature Immunol. 2008;9:970–980.

Greenwald, R.J., Freeman, G.J., Sharpe, A.H. Ann. Rev. Immunol. 2005;23:515–548.

Jiang, H., Chess, L. N. Engl. J. Med. 2006;354:1166–1176.

Klein, U., Dalla-Favera, R. Nature Rev. Immunol 2008;8:22–33.

Lee, Y.K., Mukasa, R., Hatton, R.D., Weaver, C.T. Curr. Opin. Immunol. 2009;21:274–280.

Schwartz, R.H. Ann. Rev. Immunol. 2005;23:23–68.

Vignali, D.A.A., Collison, L.W., Workman, C.J. Nature Rev. Immunol. 2008;8:523–532.

Vignali, D.A.A., Collison, L.W., Workman, C.J. Nature Rev. Immunol 2008;8:523–532.

Wilson, B., Rowell, E., Sekimata, M. Nature Rev. Immunol. 2005;9:91–105.

8.2.5 Cellular Cytotoxicity and Apoptosis (Fig. 8.2-7)

Cytotoxic cells comprise $CD8^+$ cytotoxic T cells and natural killer cells (NK cells). They are able to kill their target cells by induction of apoptosis (programmed cell death). Apoptosis is a natural process by which the organism gets rid of cells without causing damage to the tissue. It is extensively used during development and tissue remodelling and is essential for the function of the immune system (see also 7.6). Morphologically it is characterized by nuclear condensation, plasma membrane blobbing and separation of cellular fragments into apoptotic bodies. Major apoptotic pathways in a cell are shown in Figure 8.2-7.

A crucial step in apoptosis is the sequential activation of caspases (cysteine proteases specific for aspartic acid) which exist in the cytoplasm as inactive procaspases. Apoptotic pathways start by activation of initiator caspases such as caspase 8, 9 and 10 which then activate effector caspases (caspase 3 and 6). Effector caspases act on many substrates. Caspase 6, for example, cleaves lamins resulting in the disassembly of the nuclear lamina, whereas caspase 3 activates the endonuclease 'caspase-activated DNase' (CAD). In addition to the caspase-dependent pathways there is also direct activation of apoptosis-mediating endonucleases. Examples are the DNase NM23-H1 which is activated by granzyme A or the DNA-cleaving enzymes Endo G and apoptosis-inducing factor (AIF) which are released from mitochondria (see below).

There are two major apoptotic pathways: the extrinsic pathway which is induced by external stimuli and the intrinsic or mitochondrial pathway decreased growth stimuli, starvation or DNA damage. The two pathways are linked and they mutually influence each other.

Extrinsic pathway: It is either induced by binding of a ligand to a death receptor (more details in 7.6) or by enzymes like granzyme A and B which are released from granula of cytotoxic cells (see Fig. 8.2-8) and enter the cytoplasm with the help of perforin molecules.

- Receptor-induced apoptosis: Following binding of a ligand such as TNF, FAS ligand or TRAIL (TNF-related apoptosis-inducing ligand) to its specific trimeric death receptor ① adaptor molecules like the Fas-associated death domain protein (FADD) bind to the intracytoplasmic death domains (DD) of the receptor chains. By recruiting the procaspase 8 the death-inducing signalling complex (DISC) is formed ②. The procaspase 8 is autocatalytically cleaved and the resulting caspase 8 activates the effector caspase 3 ③ which initiates the fragmentation of DNA ④ by the endonuclease CAD (caspase-activated DNase). Following nuclear translocation caspase 3 liberates the inhibitor I-CAD from the CAD-ICAD complex ④. The active CAD cleaves the chromosomal DNA between nucleosomes thus generating DNA fragments of 180 to 200 bp or multiples thereof which are seen as the typical ladder in agarose gel electrophoresis. In addition, Caspase 3 cleaves gelsolin which induces the fragmentation of actin filaments leading to the disintegration of the cell into apoptotic bodies.
- Granzyme-induced apoptosis: Granzyme A directly activates the DNase NM23-H1 (also called GAAD = granzyme A-activated DNAse) by cleaving the inhibitor SET complex ⑤. NM23-H1 makes nicks into single stranded DNA. In addition, it destroys the base excision repair endonuclease Ape1 which is contained in the SET complex, breaks down lamin and damages mitochondria. Granzyme B activates the effector caspase 3 ⑥. Furthermore, it cleaves Bid which stimulates cytochrome c release from mitochondria ⑦ (see below).

Intrinsic (mitochondrial) pathway: The pivotal step in this pathway of apoptosis is the release of cytochrome c from mitochondria induced by cellular stress signals such as heat shock, starvation, gamma or UV irradiation, free radicals, damage to the DNA or toxic agents. Cytochrome c associates with the protein Apaf-1 (apoptotic

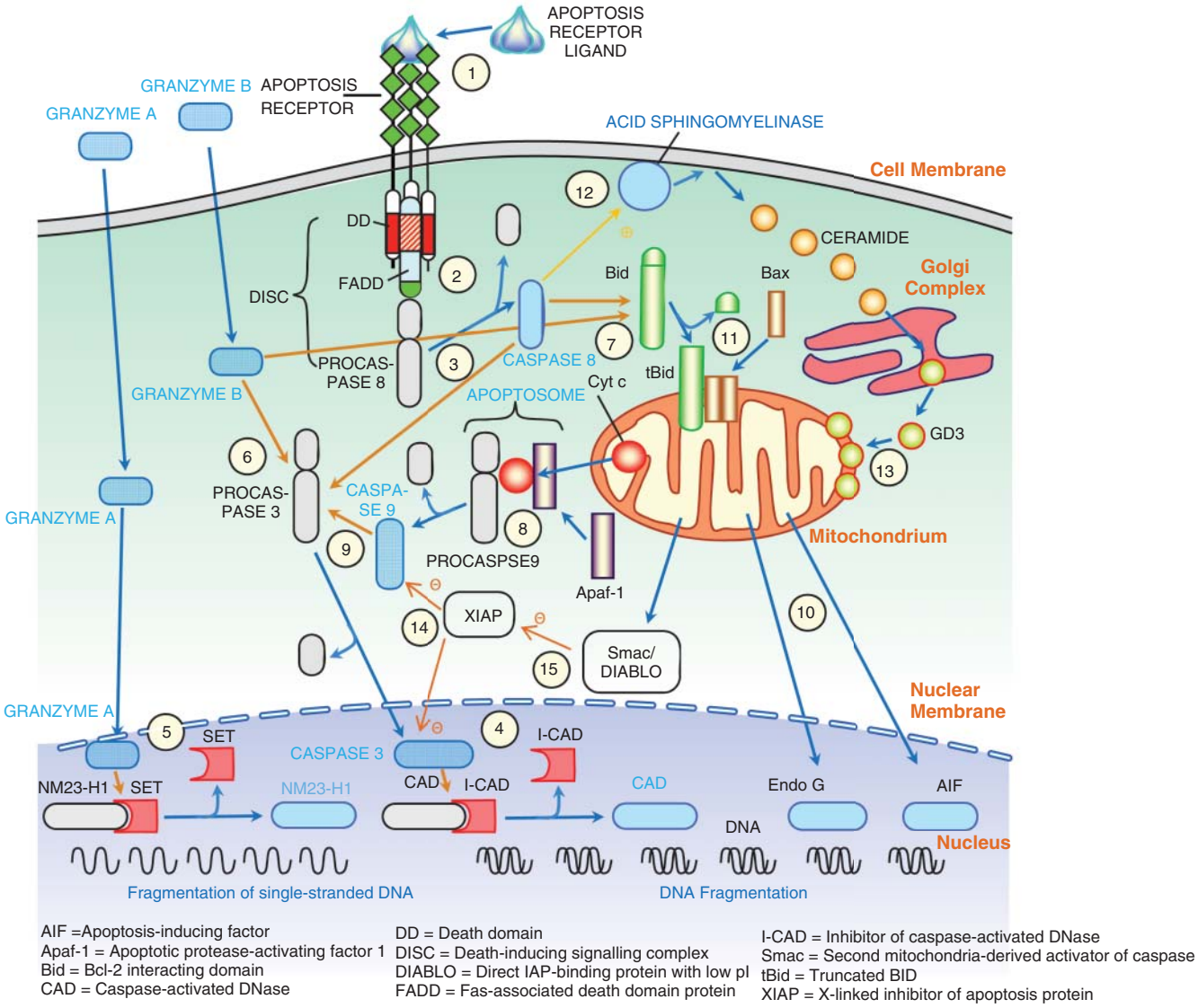


Figure 8.2-7. Apoptotic Pathways
Blue boxes = active enzymes, grey boxes = inactive proenzymes.
For details of the granzyme mechanism, see Figure 8.2-8.

protease-activating factor 1) and the procaspase 9, building a macromolecular complex called apoptosome ⑧. The procaspase 9 autocatalytically generates the active caspase 9 which activates the effector caspase 3 ⑨. The final sequence is shared with the extrinsic pathway.

A further consequence of an impaired mitochondrial function is the release of endonuclease G and the apoptosis-inducing factor (AIF) which are also capable of fragmenting DNA ⑩. The efficiency of the extrinsic pathway of apoptosis can be increased by the additional activation of the intrinsic pathway. Caspase 8 when activated in the DISC can cleave the protein Bid which belongs to the Bcl-2 (B lymphoma-2) family of apoptosis-regulating proteins and generate the truncated Bid (tBid) which is then translocated into the mitochondrial membrane ⑪. Together with the proapoptotic protein Bax it promotes the release of cytochrome c. In addition, the active caspase 8 can activate an acid sphingomyelinase (3.4.3.4) located in the cytoplasmic membrane which cleaves sphingomyelin ⑫. The generated ceramide is modified in the Golgi complex into GD3 which is able to cause the release of cytochrome c from the mitochondria ⑬.

The apoptotic pathways are controlled by members of the Bcl-2 (B lymphoma-2) protein family. Proapoptotic proteins such as Bax, Bad and Bid are counteracted by antiapoptotic proteins like Bcl-2 and Bcl-xL. In addition, there are several 'inhibitor of apoptosis proteins' (IAPs) such as survivin and the 'X-linked inhibitor of apoptosis protein' (XIAP). XIAP inhibits caspase 9 and the effector caspase 3 ⑭ and is itself blocked by the Smac/DIABLO complex which is released from mitochondria ⑮.

Cytotoxic CD8⁺ T cells (Fig. 8.2-8): Cytotoxic CD8⁺ T cells are crucial for the defense against viruses and intracellular bacteria. They recognize the intracellularly generated microbial peptides bound to MHC class I molecules on the surface of infected cells and eliminate the cells by inducing apoptosis. The cytotoxic T cells obtain their 'license to kill' during antigen-specific activation by dendritic cells that have been activated by TLR signals (see 8.1.1) or by CD40-CD40L-mediated stimulation in contact with CD4⁺ T lymphocytes.

There are several ways in which cytotoxic cells induce apoptosis in a target cell:

- One important way is the release of apoptosis-inducing substances such as perforin and granzymes that are preformed and stored in cytotoxic granules as complexes with the proteoglycan serglycin serving as a scaffold (A in Fig. 8.2-8). This multimeric complex is released in response to activation signals delivered from the T cell receptor after recognition of the specific antigen peptide on the target cell. It is thought that the perforin molecules that resemble the

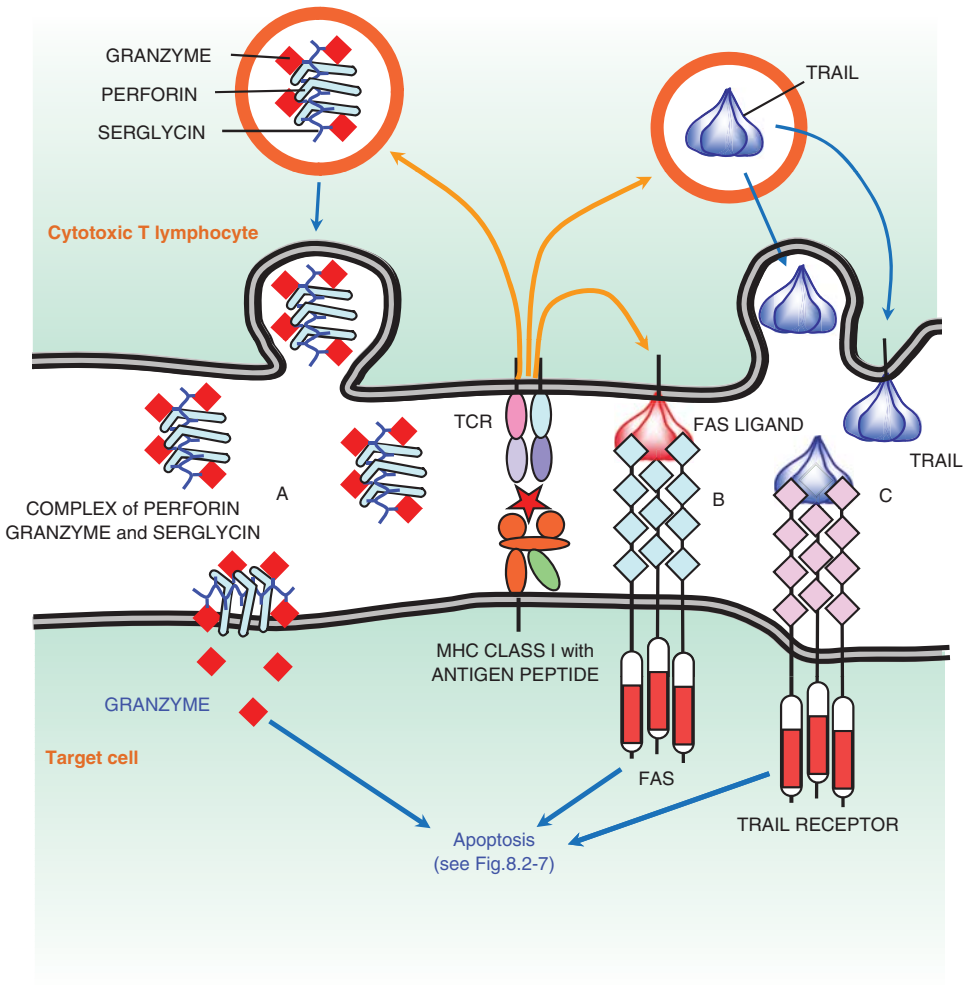


Figure 8.2-8. Mechanisms of Cellular Cytotoxicity

pore forming C9 molecules of the complement system (see 8.1.5) integrate the complex into the target cell membrane and allow the granzyme molecules to enter the cell. There, granzyme B induces the apoptotic pathway by cleaving and activating procaspase 3. Granzyme A activates the DNase NM23-H1.

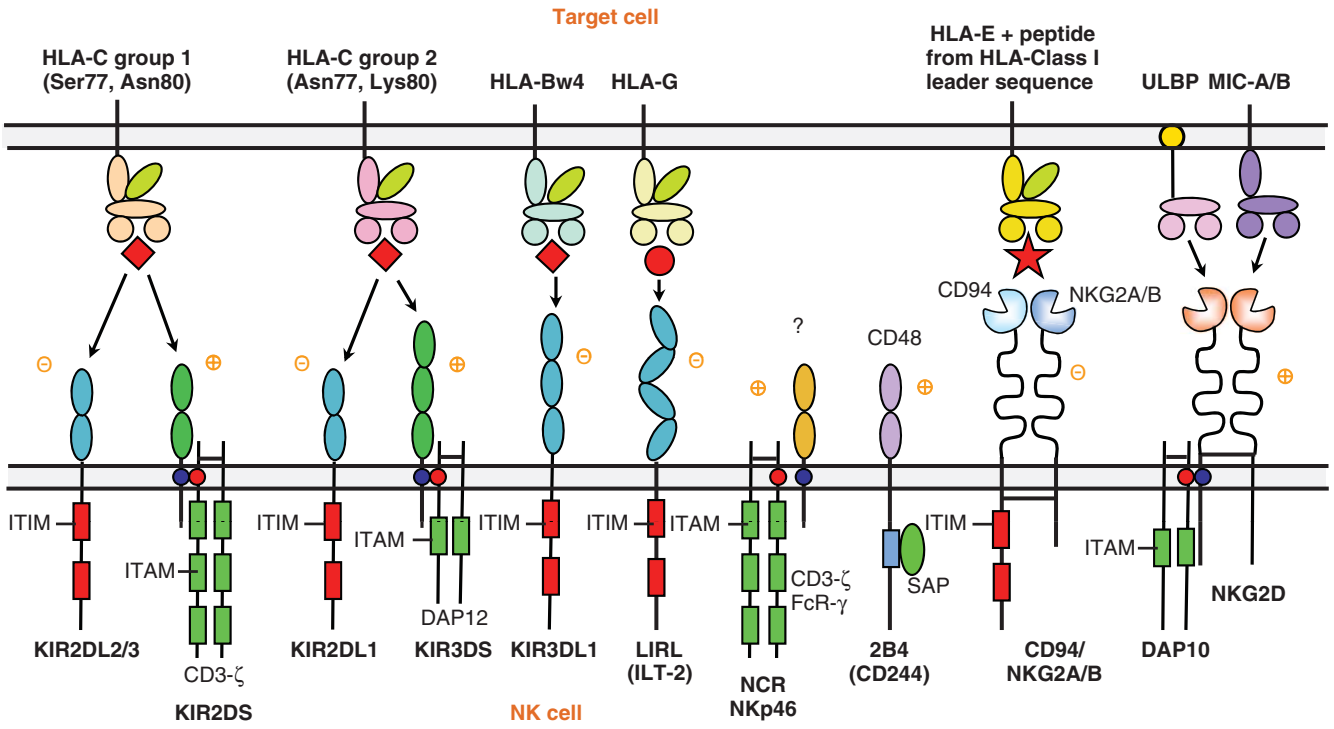
- Antigen-activated cytotoxic T cells can express (B) the FAS ligand (FASL, CD178) which binds to FAS (APO-1, CD95), a receptor belonging to the TNF receptor family which is found on many target cells. By activating the procaspase 8 the extrinsic apoptotic pathway is induced.
- A ligand structurally related to FASL is TRAIL (TNF-related apoptosis-inducing ligand) which can be released in soluble form from the cytotoxic cell or can be expressed inserted in the cell membrane (C). It activates the TRAIL receptor analogously to the stimulation of FAS by the FASL.

Natural killer cells (NK cells, Fig. 8.2-9): The second population of cytotoxic cells are the NK cells. Their cytotoxic effector mechanisms are almost identical to those of cytotoxic T cells. They belong to the innate immune system and their crucial task is to recognize and to destroy cells infected by viruses and by intracellular bacteria at an early stage of infection long before the cytotoxic T cells become specifically activated. NK cells express a limited range of invariant recognition structures whose specificity is genetically determined. The receptors which belong either to the immunoglobulin superfamily or to the C type lectin family are specific for either classical and non-classical MHC class I molecules or MHC class I-like molecules.

NK cells are thought to recognize

- The absence of classical MHC molecules ('missing self') e.g., on infected cells where viruses downregulate the expression of MHC class I molecules.
- Ligands expressed as a consequence of cell stress such as MIC molecules ('MHC class I chain related antigens') ('induced self').
- Modified peptides presented e. g. on HLA-C molecules (8.1.7, 'modified self').

The specificity of some receptors is not yet well defined. The signals mediated by NK cell receptors can either activate or inhibit the cytotoxic activity of NK cells as indicated in Figure 8.2-9. Negative signals are generated by immunoreceptor tyrosine-based inhibitory motifs (ITIM) in the cytoplasmic part of the receptor whereas activating receptors mediate their signals through associated signal transduction chains containing immunoreceptor tyrosine-based activation motifs (ITAM). When a tyrosine in the ITIM sequence [I/V] XYXX[I/L] is phosphorylated, inhibitory phosphatases are recruited via their SH2 domains. These phosphatases dephosphorylate proteins that have been phosphorylated by tyrosin kinases in the activation signaling pathway (7.5.4). How the decision about activation or inhibition of a NK cell after contact with a target cells is made is still elusive. According to one hypothesis it depends on the overall balance of signals provided by both activating and inhibitory receptors involved. One special inhibitory receptor is the CD94 / NKG2A/B lectin type receptor which recognizes the nonpolymorphic HLA-E molecule



KIR = killer-cell immunoreceptor
ITAM = Immunoeceptor tyrosin-based activation motif
ITIM = Immunoeceptor tyrosin-based inhibition motif
ILT = Ig-like transcripts (LILR = leukocyte Ig-like receptor, CD 85a-m)
NCR = natural cytotoxicity receptors
ULBP = UL-16 binding protein

Figure 8.2-9. Receptors of NK Cells and Their Ligands

presenting a peptide derived from the leader sequence of HLA class I molecules. Presence of this HLA-E/peptide complex indicates that the cell is synthesizing MHC class I molecules. Thus, NK cells are prevented from destroying noninfected cells. When the MHC class I synthesis is blocked by viral infection this inhibitory signal ceases. On the other hand the activating NKG2D receptor recognizes molecules like MIC-A/B that are expressed as a result of cell stress or malignant transformation.

Cytotoxic activity of NK cells is also stimulated by the immunoglobulin receptor Fcγ receptor III. It interacts with antibodies that are bound to cell surface antigens. This engagement of NK cells in the destruction of cells marked by antibodies is called 'antibody-dependent cell-mediated cytotoxicity' (ADCC). It is considered to be a crucial effector mechanism in the therapy of tumors with monoclonal antibodies.

8.2.6. Interactions between the Immune System and the Neuroendocrine System (Fig. 8.2-10)

Both the cells of the immune system and the cells of the central nervous system (CNS) are capable of producing cytokines, neuropeptides and other neurotransmitters and both groups of cells express receptors for these substances allowing intensive reciprocal interactions. In addition, immune organs like lymph nodes, spleen, thymus and bone marrow are connected with the CNS by the autonomic nervous system via sympathetic and parasympathetic nerve fibers.

Pathogens entering the body cause a local inflammatory reaction which is reported to the CNS through humoral and neural routes:

- Inflammatory cytokines such as IL-1 and TNF-α are transported in the blood to the brain and cause elevation of the body temperature (fever) in the hypothalamic thermo-regulatory center.
- Inflammatory mediators generate signals in the sensory parasympathetic nerve fibers (sensory vagus) which then activate the cholinergic efferent pathway (motor vagus). It suppresses the production of inflammatory cytokines such as TNF-α and IL-1 by acetylcholine which binds to nicotinic acetylcholine receptors on macrophages and lymphocytes.

Furthermore, sympathetic neurons are activated leading to the release of noradrenaline and adrenaline which results in the increased production of IL-10 and in the inhibition of inflammatory cells. Stimulation of hypothalamic nuclei activates the hypothalamic pituitary adrenal axis (7.1.5): corticotropin-releasing factor (CRF) is liberated from the hypothalamus activating the production of adrenocorticotropic hormone (ACTH) in the anterior pituitary. ACTH is transported in the blood to the adrenal gland where it stimulates the production of glucocorticoids. These hormones inhibit inflammatory cells mainly by interfering with the activation of NF-κB. The immediate regulation of inflammatory reactions by the CNS is important since it can prevent at an early stage the excessive production of inflammatory cytokines that can lead to life-threatening conditions like sepsis. On the other hand, a CNS-induced excessive IL-10- and TGF-β-mediated immunosuppression can be critical in situations such as chronic stress or sepsis.

Also fat cells contribute to the regulation of the immune system. They release the proinflammatory cytokine leptin that directly activates Th1 cells and blocks the production of CRF in the hypothalamus thereby reducing the negative regulatory effect of glucocorticoids.

8.2.7. Immunological Tolerance

Immunological tolerance is defined as a specific, actively induced nonreactivity of an immunocompetent host against an antigen. Of great importance is the antigen-specific tolerance of T cells, since antigen receptors on B lymphocytes can change their specificity through somatic hypermutation (see 8.2.3). Furthermore, B lymphocytes depend on the help of T cells recognizing the same antigens as the respective B cells. Two types of tolerance are distinguished:

Central (recessive) tolerance, is mainly responsible for the unresponsiveness against self antigens (autotolerance). It is achieved by

negative selection of self-reactive cells during lymphocyte development (clonal deletion). T cells are controlled during their differentiation in the thymus, B lymphocytes are screened in the bone marrow.

Peripheral (dominant) tolerance is the induction of tolerance in mature, immunocompetent cells. Important mechanisms are:

- Clonal anergy. When the specific antigen peptide/MHC complex is presented to a T lymphocyte by a dendritic cell that does not simultaneously provide the co-stimulatory signal (8.2.1), the T cell becomes anergic or dies by apoptosis. To express co-stimulatory molecules the dendritic cell must be activated by 'danger signals' such as pathogen-associated molecular patterns or by inflammatory cytokines.

  In addition, dendritic cells that express the indolamine-2,3 dioxygenase (IDO) can cause a local decrease in the concentration of tryptophan which is required by proliferating T cells.
- Immunological ignorance. Low levels of antigens or antigen presentation in so-called immunologically privileged sites such as brain, anterior chamber of the eye and testis fail to activate an immune response.
- Suppression by regulatory T cells. Regulatory $CD4^+$ T lymphocytes comprise the native $CD25^+$ Treg and the adaptive Th3 and Tr1 cells (see 8.2.2). They are able to inhibit the proliferation of T cells by negatively regulating cytokines such as IL-10 and TGF-β. Th3 and Tr1 cells are regarded as being responsible for tolerance induced by oral application of antigens.

Since the options for a therapeutic tolerance induction are still limited, nonreactivity of the adaptive immune system as required in allogeneic organ or stem cell transplantation and in the treatment of autoimmune diseases must be achieved by various immunosuppressive drugs such as:

- Drugs that block cell proliferation, e.g., Azathioprine.
- Cyclosporine A and FK 506 inhibiting the activation of the transcription factor 'nuclear factor of activated T cells' (NF-AT) (Fig. 7.5.6).
- Rapamycin, an antibiotic that specifically blocks IL-2-dependent T cell proliferation.
- Glucocorticoids inhibiting activation of T lymphocytes by interfering with activation of the transcription factor NFkB.
- Monoclonal antibodies against the pan T cell antigen CD3 that markedly reduce the number of circulating T cells.

8.2.8 Induction of Specific Immune Responses against Pathogens

The essential steps during induction of cellular and humoral immune responses against bacteria and viruses are compiled in Figure 8.2-11.

The first step for induction of the adaptive immune response is the uptake and processing of antigens and their presentation to T cells by antigen-presenting cells (primarily dendritic cells, DCs ①. The uptake can be augmented by various mechanisms: e.g., bacteria can be attached directly to the cell membrane of DCs via lectin receptors such as mannose receptors. Also, bacteria can be opsonized by 'natural antibodies' (8.1.4). Binding and uptake of opsonized bacteria occur either via Fcγ receptors or via complement receptors when complement fragments such as C3b are bound to their surface (8.1.5). The direct or indirect interaction of pathogens with surface receptors on dendritic cells leads to activation of these cells and to the expression of co-stimulatory molecules for T cells.

The internalized antigens are processed and those antigen peptides that fit into the binding grooves are bound by class II MHC molecules (Fig. 8.1-15). After translocation to the cellular membrane, the complexes can be recognized by $CD4^+$ T lymphocytes ② which, in the presence of co-stimulatory signals ③ are stimulated to proliferate and to differentiate into Th1 or Th2 effector cells which are pivotal in the coordination of the defense mechanisms ④. Release of IL-12 by DC following TLR stimulation favors the formation of Th1 cells, while IL-4 guides towards Th2 cells ⑤. Autocrine stimulation by IL-2 drives the proliferation of T cells after stimulation ⑥. Increased concentrations of IL-6 and TGF-β favor the development of inflammatory Th17 cells which is supported by the DC-derived IL-23 ⑦. Th17 cells efficiently recruit and activate granulocytes into inflammatory tissue infiltrates ⑧. Under circumstances which are not yet known in detail, antigen-activated $CD4^+$ T cells develop into adaptive regulatory T cells that suppress the specific immune response by secreting the cytokines IL-10 and TGF-β ⑨ (see 8.2.2).

$CD4^+$ T helper cells recognizing their specific antigen in processed form on the surface of B lymphocytes support the activation and differentiation of B-lymphocytes by presenting the co-stimulatory molecule CD40L and by the release of cytokines (8.2.3). IL-4, IL-5 and IL-6 ⑩ are provided by Th2 cells and direct the production of IgG4, IgA and IgE antibodies whereas the Th1-derived IFN-γ induces an antibody class switch to the opsonizing and complement-fixing isotypes IgG3, IgG1 and IgG2 ⑪. IL-6 especially supports the differentiation

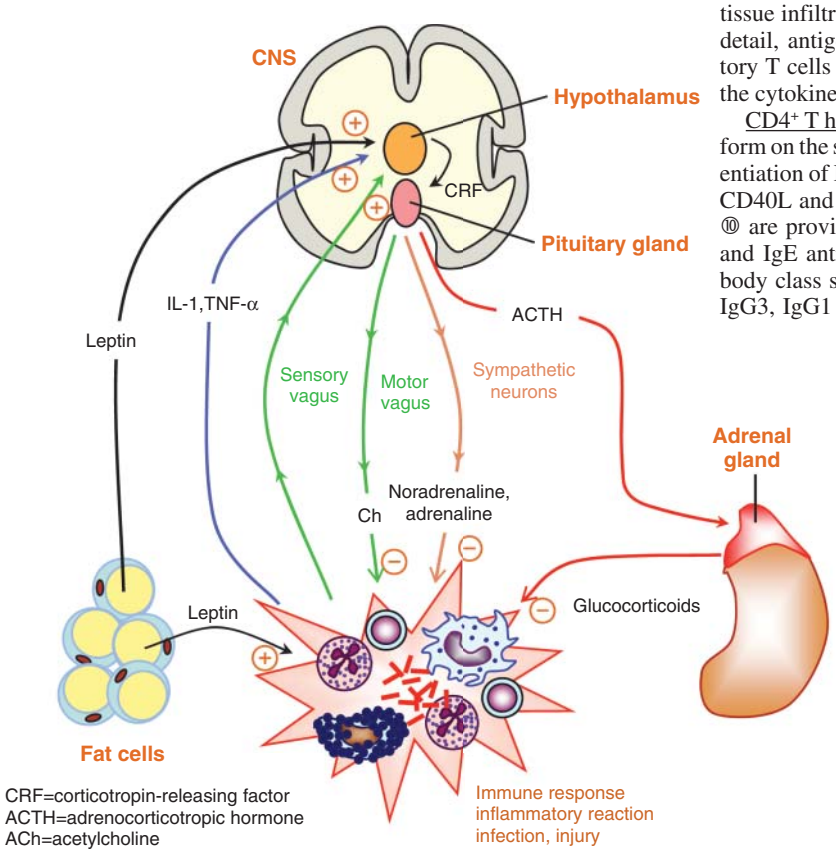


Figure 8.2-10. Interaction Between the Immune System and the Neuroendocrine System

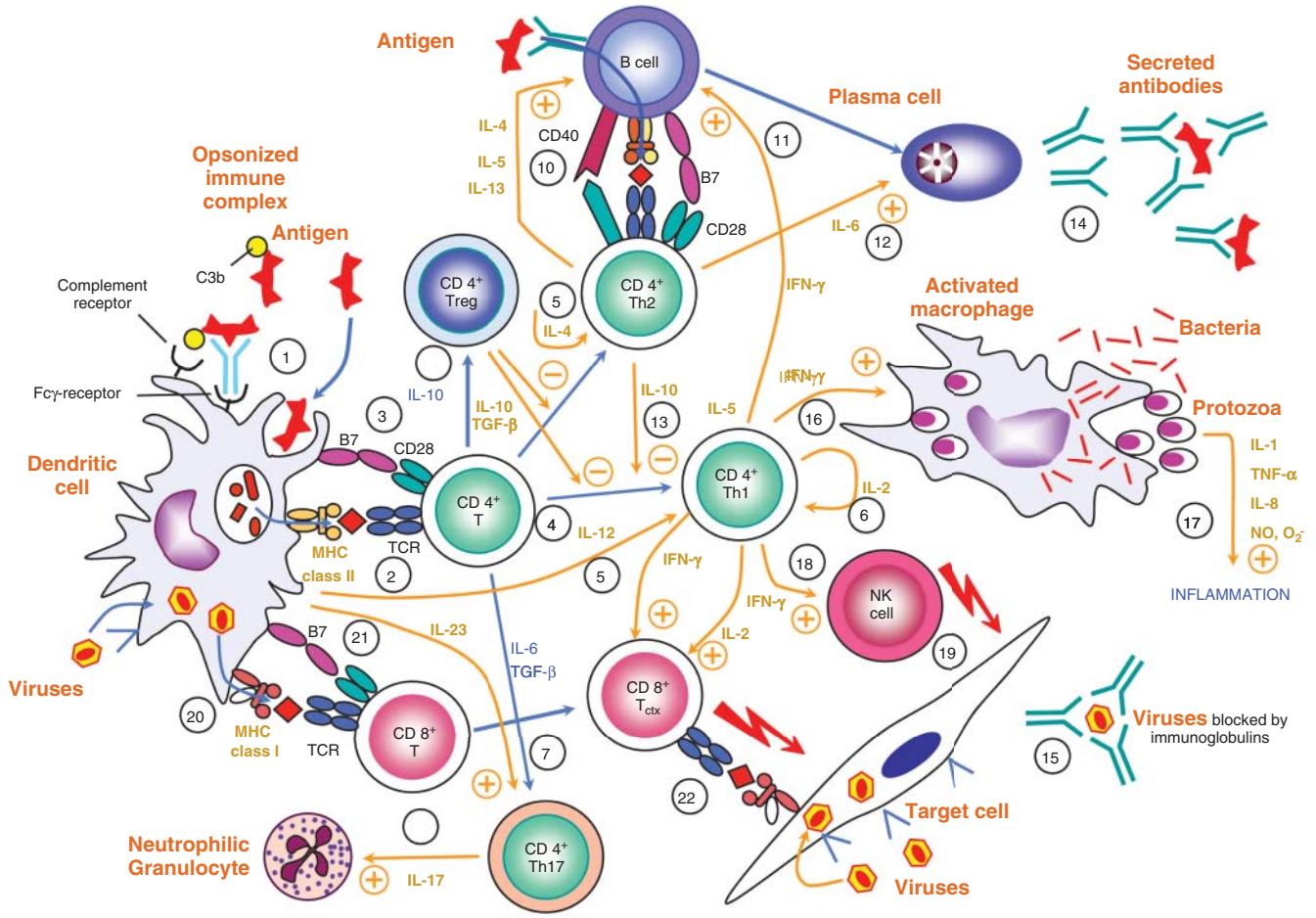


Figure 8.2-11. Induction of Specific Immune Responses Against Pathogens

into plasma cells ⑫. IL-10, which is formed by Th2 cells, inhibits the development of Th1 cells ⑬.

Plasma cells secrete specific antibodies ⑭ which have various effects on infectious microorganisms depending on their isotype. They agglutinate or opsonize extracellular bacteria or viruses ①, ⑭ and prevent the attachment of viruses to their specific receptors on the surface of target cells ⑮. They induce the lysis of bacteria by activation of the complement system (8.1.5) and neutralize bacterial toxins. Besides IL-2, IFN-γ is the most important cytokine secreted by activated Th1 cells. One major function is the activation of macrophages to enable them to kill intracellularly growing microorganisms ⑯. Activated macrophages liberate the cytokines IL-1 and TNF-α which cause local and systemic inflammatory reactions ⑰. The formation of cellular infiltrates is increased by chemokines. IFN-γ also stimulates the cytotoxic activity of natural killer cells (NK) ⑱ against virus-infected cells or tumor cells ⑲.

The immune response against viruses is also initiated by antigen-presenting cells (DC). After infection, they present viral peptides (bound to HLA class I molecules) and stimulate $CD8^+$ T lymphocytes to differentiate into cytotoxic effector cells ⑳. Second co-stimulatory signals required for activation are also provided by interaction between CD80 or CD86 and CD28 (㉑, see 8.1.7) or by IL-2, which is produced by $CD4^+$ T cells. The major function of $CD8^+$ T lymphocytes (cytotoxic T cells) is the recognition and killing of virus infected cells ㉒.

Literature to Sections 8.2.5–8.2.8:

Elmore, S. Apoptosis: A review of programmed cell death Toxicol. Pathol. 2007;35:495–516.

Goodnow, G.G. Proc. Natl. Acad. Sci. USA 1996;93:2264–2271.

Lanier, L.L. Nature Immunol. 2008;9:495–502.

Mueller, D.L. Nature Immunol. 2010;11:21–27.

Tracey, K.J. Nature 2002;420:853–859.

Vivier, E., Tomasello, E. *et al*. Nature Immunol. 2008;9:503–510.

Von Boehmer, H., Melchers, F. Nature Immunol. 2010;11:14–20.

Webster, J.I., Tonelli, L., Sternberg, E.M. Ann. Rev. Immunol. 2002;20:125–163.

8.3 Pathologic Immune responses

8.3.1 IgE-Mediated Hypersensitivity of the Immediate Type

The allergic reaction of the immediate type is a pathological, disease-causing hypersensitivity where the natural, IgE-dependent defense against intestinal parasites such as helminths is directed against harmless substances of the environment. It proceeds in two phases (Fig. 8.3-1):

- During the induction phase, the formation of IgE antibodies against a definite antigen is stimulated. How an allergen promotes the decisive isotype switch to IgE in a B cell remains elusive. It seems to depend on the selective activation of Th2 cells and the release of the cytokines IL-4 and IL-13. The IgE antibodies bind to the high affinity IgE receptors (FcεR I) located on the membrane of mast cells in tissues and basophilic granulocytes in the blood.
- During the effector phase, the binding of multivalent antigens to the receptor-bound IgE antibodies leads to cross-linking of the Fcε receptors. This generates a signal, which causes the instant

liberation of mediators by exocytosis, mostly histamine (4.8.2), tryptase, kallikrein and chemotactic factors for neutrophilic and eosinophylic granulocytes. Simultaneously, synthesis of prostaglandins, leukotrienes (7.4.8), platelet activating factor (3.4.3.3) and cytokines is induced; their secretion proceeds with some delay. The essential effects of the various mediators are shown in Figure 8.3-1. Allergic reactions in the airways can lead to allergic asthma. Here, the principal effector cells are eosinophilic granulocytes that are activated by interaction of allergens with Fcε receptor-bound IgE and by the Th2-derived IL-5. The inflammatory mediators released by eosinophilic granulocytes induce swelling of the mucosa, increased mucus production, airway obstruction and airway hyperreagibility.

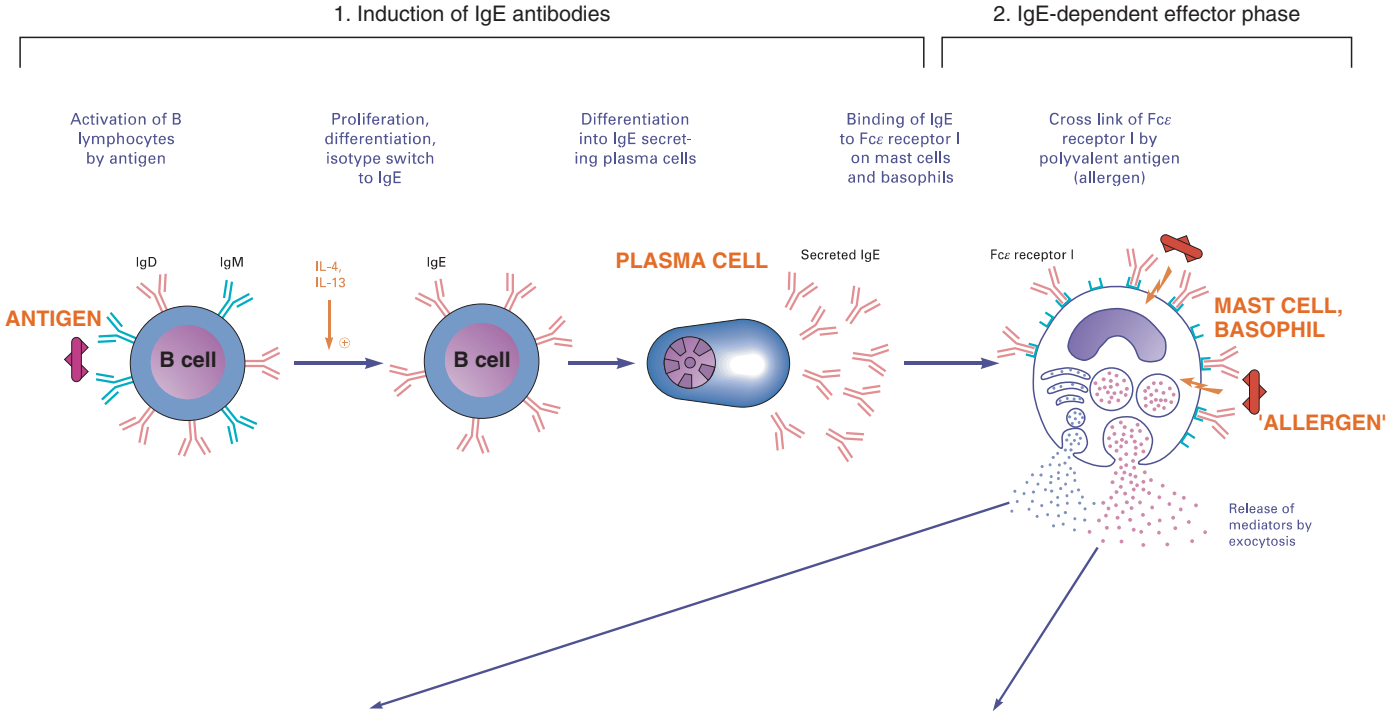
8.3.2 Autoimmunity

Autoimmune diseases develop when tolerance against self antigens breaks down and an adaptive immune response against the body's own cells and tissues is generated. Autoimmune diseases can be organ-specific or systemic. A number of factors have been identified which influence the initiation of autoimmune diseases:

- Genetic predisposition. Genetic risk factors are well documented. The most important ones have been attributed to MHC alleles. There are monogenetic autoimmune diseases that are caused by a defect in a gene coding for a protein involved in negative regulation such as CTLA4 (8.2.8), FoxP3 (8.2.2) or Fas (8.2.8).
- Environmental factors. The impact of environmental factors is well known although the mechanisms by which they promote autoimmune diseases are far from being understood.
- Hormonal influence. The observed difference between females and males in the occurrence of certain autoimmune diseases indicates a hormonal influence.
- Incomplete elimination of autoreactive lymphocytes. Among mature lymphocytes self-reactive cells are found, since the elimination of autoreactive clones during the development of lymphocytes is not complete. In addition, mature B cells can generate new antibody specificities by somatic hypermutation (see 8.2.4).
- When immunologically privileged sites (8.2.7) are injured or inflamed, sequestered antigens become exposed and can evoke an autoimmune reaction in the immunostimulatory environment.
- Epitope spreading. T cells being specific for a protein in a complex antigen can help B lymphocytes to make antibodies against other proteins in the complex.

Although self-reactive T lymphocytes are strictly controlled by thymic selection, induction of anergy or death, and by regulatory cells, there are various models suggesting how T cell tolerance can be bypassed in autoimmunity:

- Modification of self proteins. Modification of proteins by small molecules such as drugs can result in altered peptides binding to MHC molecules. These may represent neoantigens for T lymphocytes which are then activated and provide help to B cells producing antibodies to the original self protein.
- Activation of autoreactive T cells in inflammatory environment. Autoantigen-presenting dendritic cells that are normally quiet and induce anergy or apoptosis to self-reactive T cells become activated



Newly synthesized mediators	Biological effect	See Section
Leukotriene C_4, Leukotriene D_4	bronchoconstriction, edema, hypersecretion, inflammation	7.4.8
Leukotriene B_4	chemotaxis	7.4.8
Thromboxane A_2, PDG_2	bronchoconstriction, edema, hypersecretion, vasodilatation	7.4.8
Platelet activating factor (PAF)	bronchoconstriction, vasodilatation, aggregation and activation of platelets	3.4.3.3, 9.4
Cytokines	activation of inflammatory cells	8.1.8

Preformed mediators	Biological effect	See Section
Histamine	vasodilatation, bronchoconstriction	3.2.8.2
Heparin	inhibition of blood coagulation	9.3
Tryptase	activation of complement factor C3	8.1.5
Kallikrein	generation of kinins, vasodilatation, edema	9.2.2
Eosinophilic chemotactic factor A (ECF-A)	chemotaxis of eosinophils	
Neutrophilic chemotactic factor (NCF) = IL-8	chemotaxis of neutrophils	

Figure 8.3-1. Two Phases of the IgE-Mediated Allergic Immune Response

and express co-stimulatory molecules in an inflammatory environment. This is frequently observed in the course of an infection.

- Molecular mimicry. When a T lymphocyte is activated against a pathogen that shares an amino acid sequence with a self protein the activated T lymphocyte can react against this antigen and support B lymphocytes in the production of antibodies against the autoantigen. This type of cross reactivity is called molecular mimicry. A multitude of sequences shared between pathogens and self proteins have been identified so far.
- Superantigen stimulation. Bacterial or viral superantigens polyclonally activate T lymphocytes (8.2.1). Some of the activated T cells might possess receptors specific for autoantigenic peptides. When the stimulatory signals are sustained, a damaging autoimmune reaction can ensue.

Literature:

Goodnow, C.C., Sprent, J. *et al*. Nature 2005;435:590–597.

Gould, H.J., Sutton, B.J. Nature Rev Immunol 2008;8:205–217.

8.4 Adhesion of Leukocytes

Anton Haselbeck

Cell adhesion is a fundamental feature of multicellular organisms including their defense mechanisms. In the latter case in mammals, leukocytes play a central role. They bind bacteria, parasites, viruses, tumor cells etc (8.1.1.2). Furthermore, their interactions with the endothelium are of special importance. Two situations are relevant in this respect:

- During an inflammation or an immune reaction (8.2), specialized leukocytes (neutrophilic and eosinophilic granulocytes, monocytes) adhere to and pass through the endothelium of the blood vessels and the underlying matrix.
- Generally, lymphocyte adhesion and passage from the bloodstream to the lymphatic system occurs in the high endothelial venules (HEV) of the lymph nodes (Fig. 8.2-6). This way, lymphocytes can communicate with each other in the lymphatic system and search for foreign compounds after their recirculation to the bloodstream, thus fulfilling their role in the immune system.

In both cases, the reaction passes through the following steps (Figs. 8.4-2 and 8.4-3):

Rolling: The flow of cells is slowed down by first making contacts to the endothelium. The interaction takes place between P-, E- and L-selectins and their receptors (Table 8.4-1, Fig. 8.4-1); it is weak and reversible. Involved are:

- P-Selectin. It is released from granules of endothelial cells to their surface within minutes after irritation. Mediators for release are, e.g., thrombin, histamine or complement proteins. It recognizes the sialyl-Lewis^x-structure (4.4.3) at the mucin PSGL-1 of leukocytes.
- E-Selectin. It is expressed between 3 to 6 hours after irritation of the endothelium by *de novo* synthesis. Mediators of the expression are, e.g., interleukin-1, tumor necrosis factor-α (TNF-α) or bacterial lipopolysaccharides. It recognizes sialyl-di-Lewis^x structures and their sulfated derivatives at the leukocyte surface.
- L-Selectin. It is constitutively expressed on leukocytes and is shed into the serum. It recognizes sialyl-Lewis^x structures or sulfated derivatives (4.4). The receptors for L-selectin at the endothelial cells are predominantly PSGL-1 and glycoproteins bearing sialyl-Lewis^x structures. In the lymphatic system, L-selectin binds to MadCAM and to the mucins CD34 and GlyCAM.

Adhesion: After activation of leukocyte integrins, firm contacts are established between them and endothelium molecules of the immunoglobulin superfamily. These interactions are irreversible. Involved are the integrins LFA-1 (binds to ICAM-1 and 2), Mac-1

Table 8.4-1. Molecules Involved in Leukocyte Binding (Examples)

	CD No.	Mol. mass kDa	Present in	Notes	Ligand Name	Family	Mol. mass kDa	Present in
Selectins								
P-selectin = PADGEM	62P	140	endothelial cells, platelets	released within minutes from α-granula	PSGL-1 (sialyl-Le^x-structure)	sialomucin	220 dimer	leukocytes
E-selectin = ELAM-1	62E	105 ... 115	endothelial cells	expressed between 3–6 hours	sialyl-di-Le^x-structures = CLA	glycoprotein		granulocytes, memory T- cells
L-selectin	62L	75 ... 110	neutrophils, monocytes, eosinophils, etc.	constitutive, shed from cells after leukocyte activation	PSGL-1, sialyl-Le^x-structures, also sulfated	glycoprotein		endothelial cells
Same			lymphocytes		CD 34 MAdCAM,GlyCAM (secreted)	sialomucin mucinous immunoglobulins	105 ... 120 58–66 50	periph. lymph nodes lymphoid high endothelial venue
Integrins[1]								
LFA-1 ($\beta_2\alpha_L$)	11a	180 (α) + 95 (β)	leukocytes	activated by cytokines, chemokines, etc.	ICAM-1 (CD 54) ICAM-2 (CD 102)	immunoglob. immunoglob.	90 ... 115 55–65	endothelial cells endothelial cells
Mac-1 = CR3 ($\beta_2\alpha_M$)	11b/18	170 (α) + 95 (β)	leukocytes	activated by cytokines, chemokines, etc.	ICAM-1 (CD 54) fibrinogen (soluble)	immunoglob. fibrous prot.	90 ... 115	endothelial cells bridge to endothelial cells
					iC3b	inactivated complement factor		attached to endothelial cells
p150,95 ($\beta_2\alpha_X$)	11c/18		leukocytes		fibrinogen (soluble)	fibrous prot.		bridge to endothelial cells
					iC3b	inactivated complement factor		attached to endothelial cells
VLA-4 ($\beta_1\alpha_4$)	49d/29	150 (α) + 130 (β)	lymphocytes, monocytes etc. not neutrophils	activated by cytokines, chemokines, etc.	VCAM-1 (CD106) fibronectin	immunoglob. fibrous prot.	90–110	endothelial cells bridge to endothelial cells
$\beta_7\alpha_4$ = MLA	49d	150 (α) + 120 (β)	lymphocytes		MAdCAM-1	mucinous immunogl.	58–66	lymphoid high endothelial venule
Immunoglobuin								
PECAM-1	31	120 ... 130	platelets, monocytes, granulocytes, T cells	constitutively expressed	PECAM-1 (see at left)	immunoglob.	120 ... 130	endothelial cells

[1] The suffixes indicate the type of the subunits, not their numbers.

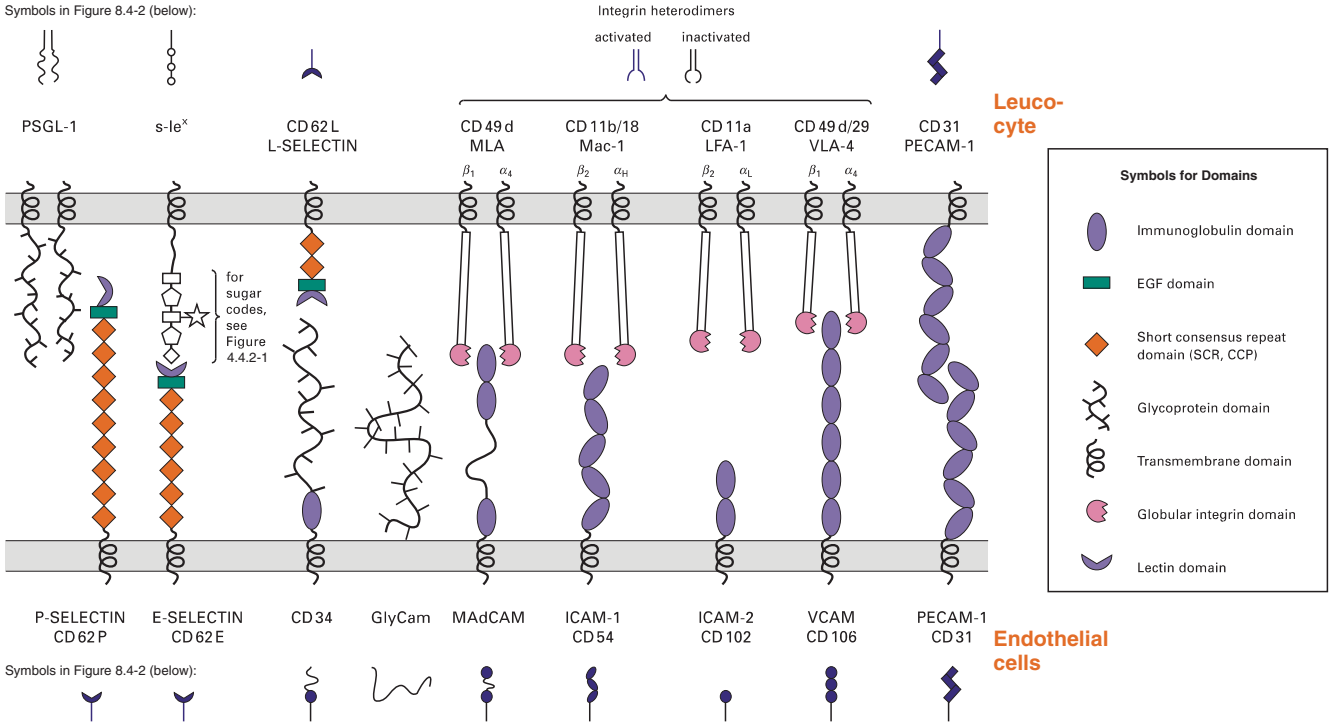


Figure 8.4-1. Structure of Adhesion Molecules

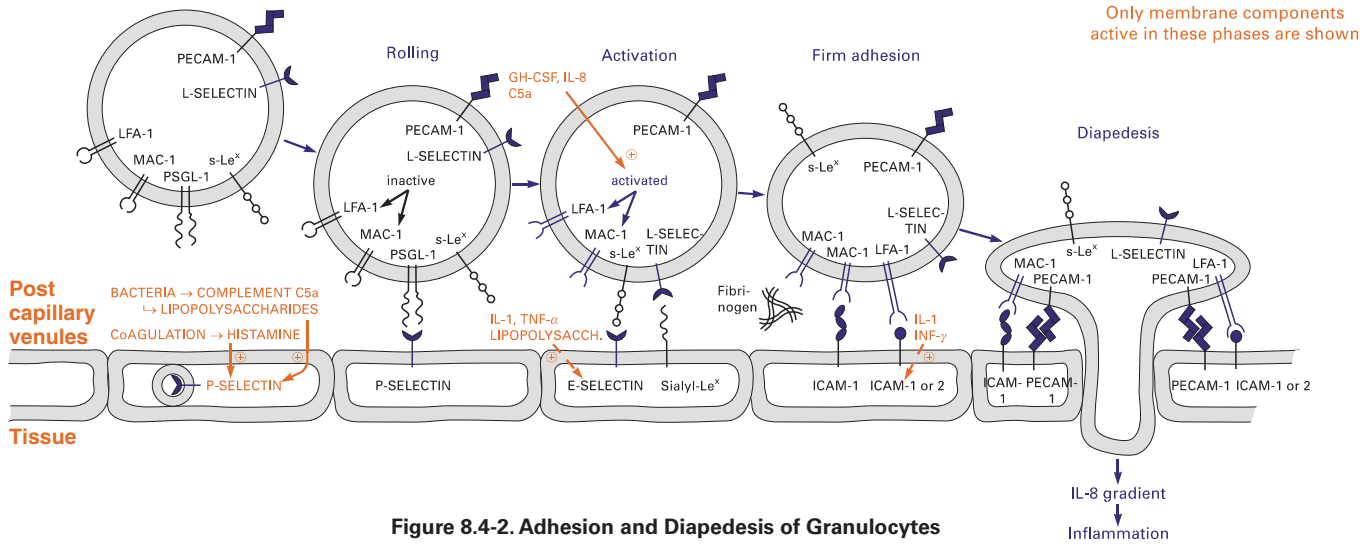


Figure 8.4-2. Adhesion and Diapedesis of Granulocytes

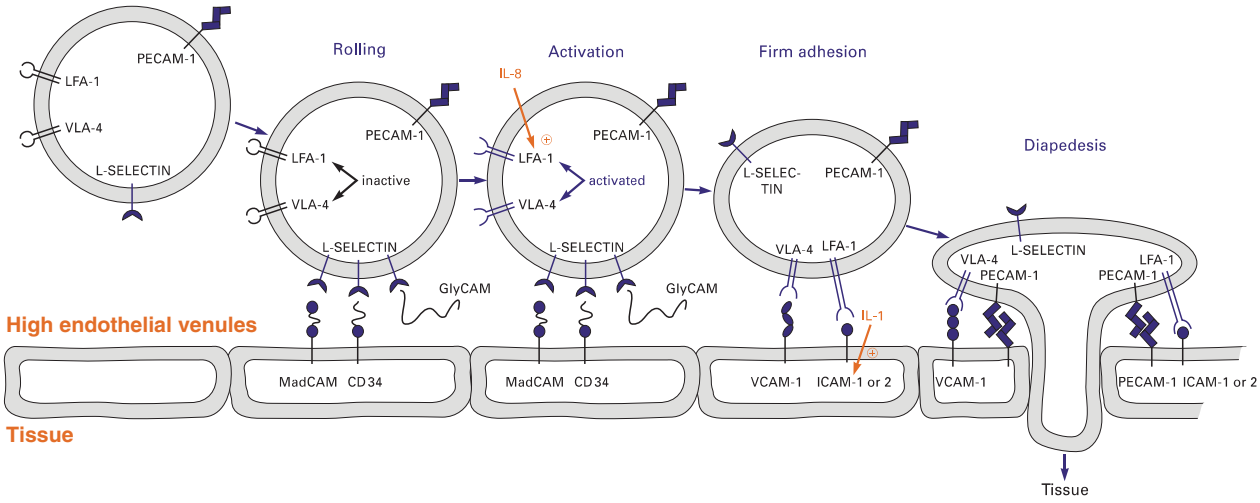


Figure 8.4-3. Adhesion and Diapedesis of Lymphocytes

(binds to ICAM-1) and VLA-4 (binds to VCAM-1 or MadCAM-1). Additionally, Mac-1 also binds to fibrinogen (forming network bridges) and to the inactivated complement factor iC3b (therefore, Mac-1 is also named complement receptor 3, CR-3, 8.1.5). VLA-4 also binds to fibronectin which is expressed on the endothelium.

Leukocyte activation and expression of endothelial factors are regulated. Activation of integrins is effected by cytokines (e.g., GM-CSF), chemokines (e.g., IL-8) and chemoattractants (e.g., complement factor C5a). Increase in transcription of the endothelial ligands ICAM-1 and VCAM-1 is caused by the same agents as for E-selectin.

Flattening of the cells and diapedesis: Adhering leukocytes crawl to an intercellular junction of the endothelium and then transmigrate to, or even through, the intercellular matrix. This is mediated by a homophilic interaction of PECAM (platelet-endothelial cell adhesion molecule 1; CD31). It is expressed both on endothelial cells (concentrated at the junctions) and on granulocytes and monocytes. During inflammation, granulocytes follow an IL-8 gradient in the tissue towards the focus of infection.

The transmigration of lymphocytes to the lymphatic system is mediated by interaction of L-selectin or $\beta_7\alpha_4$-integrin (at the lymphocytes) with MadCAM (on the cells of the high endothelial venules).

Other adhesion reactions: Similar cell-cell interactions as described for leukocytes and their modification also play a role in:

- Embryonic development for segregation of different groups of cells and dispersion/aggregation of migratory cells, including the formation of the nerve system. Besides similar molecules (as listed in Table 8.4-1 for leukocytes), cadherines are also involved. They occur in most cells and mediate Ca^{++} dependent homophilic adhesion (that is, between the same cell type). For dispersion of cells, their expression is downregulated.
- Maintaining correct contacts between cells or with the extracellular matrix during postnatal life.
- Aggregation of platelets during blood coagulation.
- Tumor cell migration. Extravasation of migrating tumor cells into certain tissues is mediated by a similar cascade of events as described here for leukocytes. Additionally, in several invasive tumor cell types, the levels of E-cadherin or specific integrins are reduced (apparently causing separation from the tissue of origin), while the levels of other adhesion molecules are elevated (needed for invasion?).

Literature:

Carlos, T.M., Harlan, J.M. Blood 1994;84:2068–2101.

Garrido-Urbani, S., Bradfield, P.F., Lee, B.P., Imhof, B.A. Biochem. Soc.Trans. 2008;36:203–211.

Lasky, L.A. Ann. Rev. of Biochem. 1995;64:113–139.

Patel, T.P. *et al.* Biochemistry 1994;33:14815–14824.

Rosen. S.D., Bertozzi, C.R. Current Opinion in Cell Biology 1994;6:663–673.

Smith, W.C. J.Allergy Clin.Immunol. 2008;121:2, 375–379.

van der Merwe, P.A., Barclay, A.N. Trends in Biol. Sci. 1994;19:354–358.

9 Blood Coagulation and Fibrinolysis

Peter Müller

9.1 Hemostasis

The word 'hemostasis' is derived from the Greek expressions for 'blood' and 'stagnation'. Hemostasis is the process that stops bleeding from injured blood vessels by means of fibrin polymerization and formation of a blood clot containing a fibrin-network, that covers the injured site. In addition, the hemostasis process includes the important limitation and eventual dissolution of the fibrin clot. The major hemostasis functions are:

- Limiting blood loss by sealing leaks in blood vessels
- Sealing of non-physiological surfaces and injured walls within blood vessels
- Initiating inflammation and wound healing processes
- Dissolving wound-sealing clots after healing.

In order to fulfill these tasks, the hemostasis system has to react quickly, effectively, and in a precisely controlled manner. Both insufficient and excessive clotting has dire life threatening consequences. Hemophilia, an inheritable disease causing deficiencies in blood coagulation, can lead to uncontrolled bleeding when left untreated, causing death either by the depletion of the blood or by tissue damage through the internal accumulation of spilled blood (hematoma) in essential organs such as the brain. On the other hand, a loose fibrin clot (embolus) in the circulation can clog smaller blood vessels downstream and may cause infarction that can have fatal consequences when essential organs such as the heart, brain or lung are affected. As a consequence, a highly intricate system has evolved to control both coagulation and fibrinolysis. Major components encompass cells and platelets, membrane bound and soluble proteins, and low molecular weight factors.

Individual components are present as latent precursors which allow a quicker response in comparison to signalling circuits that rely on the regulation of transcription factors. High sensitivity and effective action are achieved by a cascade of proteolytic activation steps (9.2, 9.3). The product of each activation step can specifically activate multiple molecules of the next downstream factor, leading to an avalanche effect: even a few molecules activated at the injured site lead to macroscopic clots that can stop bleeding within minutes. It is an essential feature of this system that not only the fibrin-forming clotting cascade is activated, but simultaneously regulatory measures are taken that limit clot formation and eventually lead to its dissolution, such as the activation of the fibrin-degrading enzyme, plasmin.

Regulation takes place at various levels:

- The coagulation cascade is initiated only at surfaces that are not protected by endothelial cells (non physiological surfaces, subendothelial binding sites exposed after damage to the vessels or at the surface of blood clots themselves).
- The inner endothelial cell lining of blood vessels and inhibitors in the plasma are essential to suppress unwanted clot formation.
- The coagulation cascade is controlled by numerous feedback inhibition mechanisms and by mutual regulatory circuits with the clot resolving pathway.

After an injury, the following reactions take place:

- Platelets and the injured vessel wall secrete vasoconstrictive compounds (9.4). The smooth muscle cells of the vessel wall contract for about 1 minute. This decreases the vessel diameter and curbs the blood flow, thereby limiting blood losses. Thereafter, vasodilatory and vascular permeability increasing compounds (e.g., bradykinin) are liberated. This allows a slower blood flow that supports clot feeding and also activation of inflammation processes by immune and wound-healing cells (e.g., granulocytes and monocytes, 8.3) that migrate through the vessel wall (extravasation).

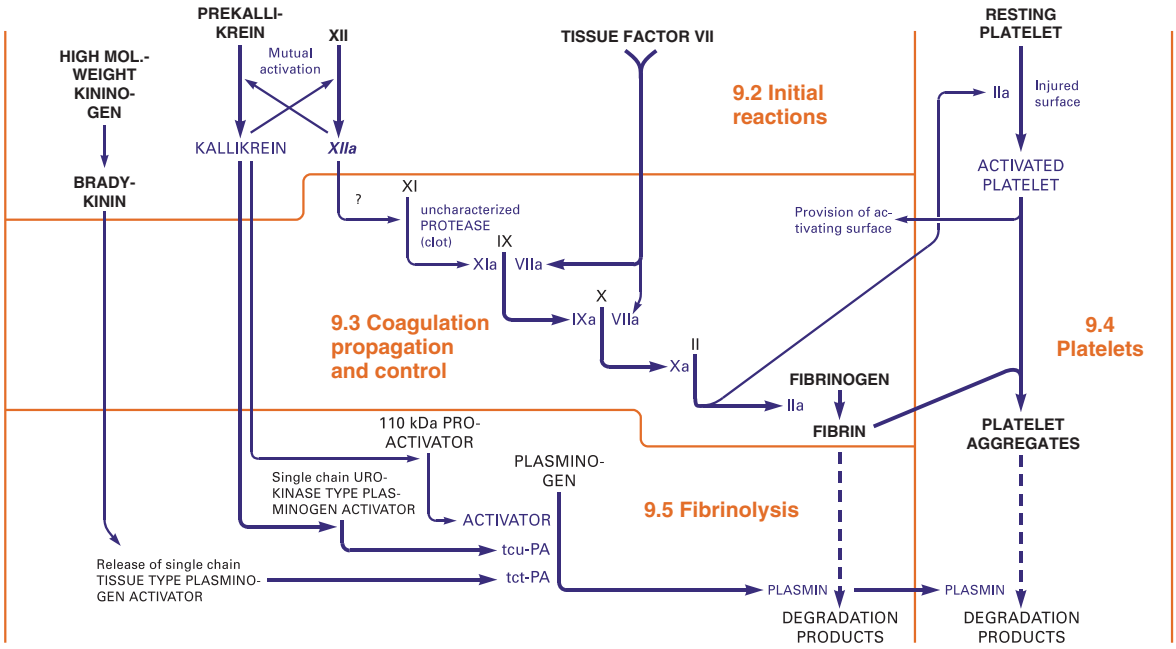


Figure 9.1-1. Systems of Coagulation and Fibrinolysis

Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Second Edition. Edited by Gerhard Michal and Dietmar Schomburg.

- Platelets adhere to the collagen layers of the subendothelium, which become exposed after injury of the endothelial cell layer. They form aggregates (9.4) together with other cells and proteins, especially with fibrinogen.
- Starting at the injured surface and the adhering platelets, coagulation cascades culminate in the activation of thrombin, which then converts fibrinogen into insoluble fibrin. This and platelet reactions lead to the formation of a wound-sealing clot (9.3).
- After the initial phase of clot formation, recent research has yielded strong evidence of a distinct second phase of clot expansion.
- Formation of the clot is limited by coagulation inhibitors. After wound healing, fibrinolysis dissolves the clots (9.5).

A schematic survey on coagulation and fibrinolysis is given in Figure 9.1-1 (surfaces and inhibitors are not shown). Most coagulation factors involved in hemostasis are denoted by Roman numerals; 'a' denotes the activated form. The numbering is based on the order of discovery of the respective factors. The factors acting towards the end of the cascade occur in larger concentrations in the blood.

9.2 Initial Reactions

Usually, two principal events are known to lead to coagulation:

- exposure of tissue factor to blood ('extrinsic pathway')
- coagulation propagation and control. This includes the important mechanism traditionally known as 'contact activation' or 'intrinsic pathway', which is responsible for coagulation initiation specifically *in vitro* (see below for details).

9.2.1 Reactions Initiated by the Tissue Factor (Fig. 9.2-1)

The tissue factor (TF, thromboplastin, factor III, CD142) is constitutively expressed as an integral membrane protein on the surface of most cells surrounding the blood vessels underneath the endothelial cell layer. Therefore, TF gets into contact with plasmatic factors only in case of a mechanical lesion (including, e.g., myocardial infarction). During pathological conditions (e.g., inflammation, sepsis) inflammatory mediators can induce the expression of TF at the surface of endothelial cells and monocytes and thereby cause systemic coagulation without the prerequisite of a mechanical lesion.

After binding of the plasmatic factor VII to TF, factor VIIa can be generated by autoactivation (noticeable proteolytic activity in contrast to XII and PK) or by attack from other coagulation factors (e.g., thrombin, Xa or hepsin, a membrane-associated protease). TF does not only localize the reaction, but also triggers structural changes in VII as a prerequisite for proteolytic attack. For full activity, Ca^{++} and a special phospholipid composition (PL) surrounding TF is required. Formation of the TF-VIIa complex seems to be the main cause of starting the coagulation cascade (9.3.3).

9.2.2 Contact Activation (Fig. 9.2-2, top)

Active factors (α-XIIa or KK) docking at a lesion site (LS) activate reciprocally inactive prekallikrein (PK) to kallikrein (KK, with HK as cofactor) and factor XII to α-XIIa. Kallikrein, in turn, splits α-XIIa to non-binding β-XIIa and also high molecular weight kininogen (HK) to bradykinin, releasing peptide fragments.

Although contact activation is well documented *in vitro*, details of the *in vivo* activation of the initiating factors and of the contact activating surface (at the lesion site, see below) are still unclear.

The major functions of contact activation appear to be the initiation of inflammatory processes and the propagation of clot formation. The contact activation reaction can contribute to occlusive thrombus formation, which is however counteracted by the activation of fibrinolysis (9.5) involving the following reactions (Fig. 9.5-2):

- Bradykinin strongly stimulates the secretion of sct-plasminogen activator (tPA) from internal storage pools in endothelial cells.
- Plasma kallikrein, in collaboration with plasmin, converts the weakly active scu (single-chain urokinase-type)-plasminogen activator into the highly active tcu (GPI-anchored two-chain urokinase)-plasminogen activator.
- Plasma kallikrein also activates a 110 kDa proactivator, which consecutively converts plasminogen into plasmin.

The *in vivo* importance of contact activation for coagulation (as derived from *in vitro* experiments) must be reconsidered. Individuals with a deficiency in factor XII, HK or prekallikrein (PK) show no signs of abnormal bleeding. Rather, some patients with factor XII defects suffer from thrombotic episodes, indicating a possible role of factor XII in fibrinolysis.

XI can be activated independently of XII and activation of XI by thrombin has only been documented *in vitro*. Contact activation may also start the complement system by attack of β-XIIa on complement component $C1qr_2s_2$ (8.1.5). It remains to be proven that *in vivo* a feedback loop exists whereby thrombin activates XII and the resulting XIIa is a physiologically relevant trigger for activation of XI for the subsequent activation of the coagulation cascade (9.3).

Receptor dependent binding of XI to platelets or binding of XI to growing clots appears to be the essential prerequisite for its activation. XI deficiency in animal models impairs arterial thrombus formation but does not lead to increased bleeding. There is evidence that XIa may protect clots against degradation. In accordance with the finding that XI deficiencies cause mild hemorrhage this seems to be of importance in particular in tissues with high fibrinolytic activity

9.2.3 Generation of Binding Surfaces (Fig. 9.2-2, bottom)

Localized coagulation requires the presence of factor specific binding surfaces for activation of the various factors. They are generated

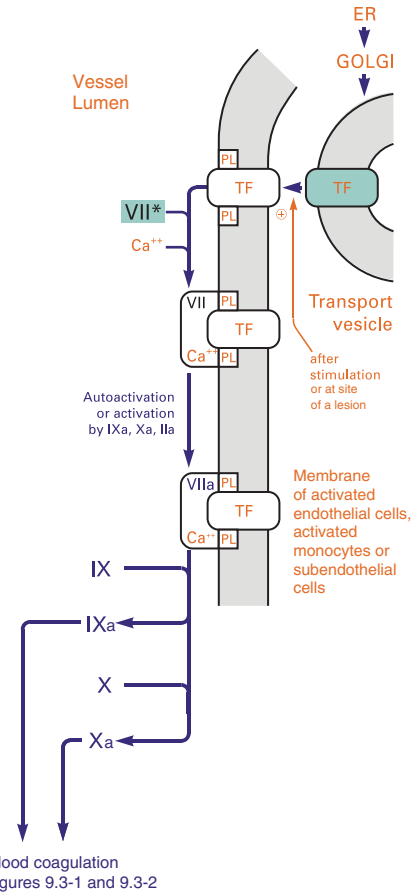


Figure 9.2-1. Reactions Initiated at the Tissue Factor

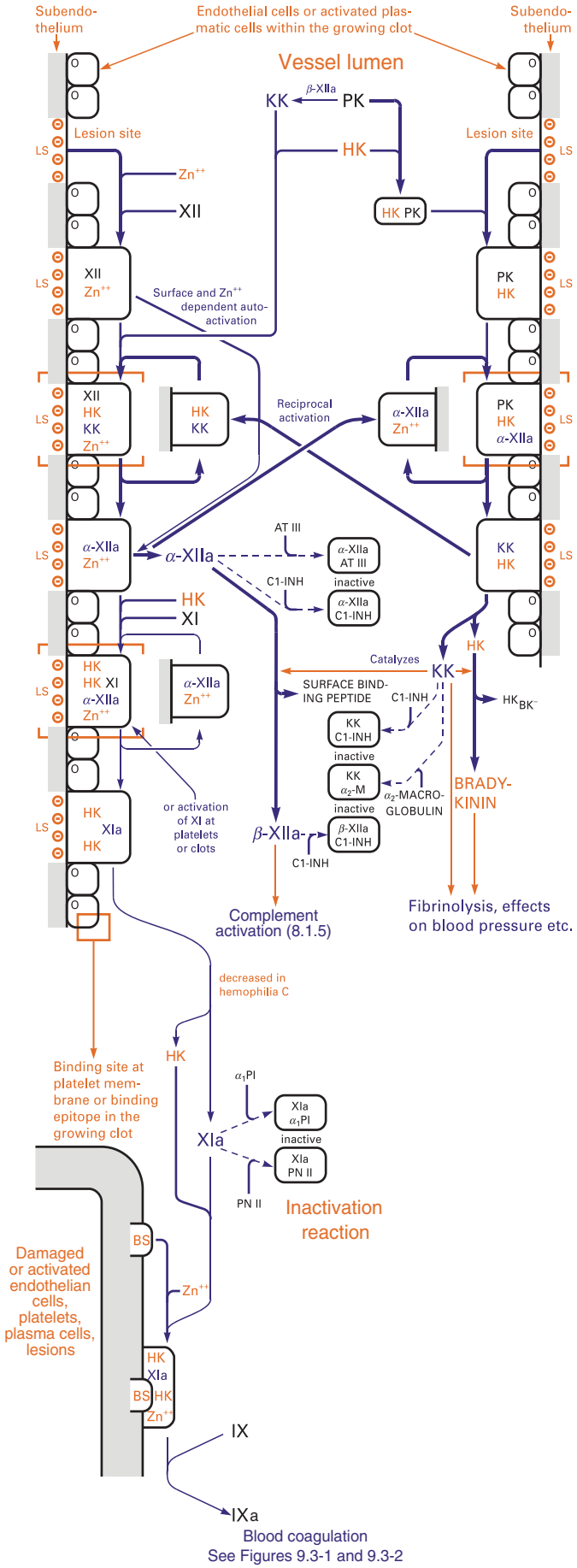


Figure 9.2-2. Contact Activation

by damage to, or by activation of endothelial cells, platelets or other plasma cells (lesion sites). These surfaces are composed of:

- Accumulated negatively charged phospholipids. Their main component is phosphatidylserine (3.4.3.2), that is actively accumulated on the inner surface of all living cells by a specific, ATP driven flip-flop mechanism. It reaches the outer surface by inhibition of this flip-flop mechanism
  - in cells in the case of effector signaling, injury or death
  - in microparticles (inside-out particles) discharged from, e.g., stimulated platelets
- Proteins (Table 9.3-2), which form complexes with the phospholipids:
  - TF, an integral membrane protein (9.2.1) and/or
  - Cofactors VIII/VIIIa and V/Va (9.3) and/or
  - other integral membrane proteins, e.g., thrombomodulin or 'receptors' for XI/XIa, IXa, protein C and protein S Calcium enables interactions between binding sites and coagulation factors.

In the case of contact activation, the nature of this surface and the mechanism of its generation are under discussion. The binding surface is possibly an exposed collagen layer. There is some evidence, that the contact activating surface may also be composed of glycosaminoglycans (2.9) present in the membrane of plasmatic cells, which are activated by inclusion into the growing clot. *In vitro*, coagulation can be started at e.g., sulfated glycoproteins or kaolin.

9.3 Coagulation Propagation and Control (Fig. 9.3-2)

The main purpose of these reactions is to provide thrombin for coagulation and other effects and to regulate its formation. The proteolytic activation mechanism (9.1) occurs in definite cascade complexes, which allow localization, activation and limited protection of the cascade reactions from inhibition.

9.3.1 Requirements for Protease Activity

All of the cascade complexes assemble at binding surfaces (9.2.3). In the complex, aided by cofactors, the conformation of the protease and the substrate orientation are optimized (Table 9.3-1).

Many factors involved in the coagulation cascade (VII, IX, X, prothrombin, as well as the coagulation control factors PC and PS, indicated in Fig. 9.3-2 with *) contain domains with 9 … 12 Ca^{++}-binding amino acids γ-carboxyglutamate (Gla, Fig. 1.3-3). This compound is formed posttranslationally by carboxylation of glutamate in presence of vitamin K (3.7.13). Vitamin K deficiency or presence of K antagonists (e.g., dicoumarol, Warfarin®) prevents the reaction. Lack of γ-carboxyglutamate seriously impairs binding of the factors to the binding site and their activation. This causes severe bleeding.

Table 9.3-1. Activity Increase by Coagulation Complex Formation

Protease	Substrate	Cofactors	Effectivity[1]
VIIa	X	TF	3.0×10^5
IXa	X	VIIIa + platelets	1.7×10^7
Xa	II	Va + platelets	6.5×10^6

[1] Ratio of activities (k_2/K_M values) with and without cofactors (1.5.4).

9.3.2 Pathways Leading to Thrombin

Several interconnected and regulated pathways lead to thrombin activation (a survey is given in Fig. 9.3-1).

9.3.3 Key Events (Fig. 9.3-2)

Factor X is activated to Xa by the TF-VIIa complex ①. (At the same time, TF-VIIa activates IX ②, see below). The consecutive thrombin formation from prothrombin (II to IIa by Xa) requires the presence

of cofactor Va, which has to be activated from Va by thrombin. Thus, thrombin exerts a feedback activation mechanism. Therefore the 'first' thrombin molecule is generated in a suboptimal way.

TFPI (tissue factor pathway inhibitor) quickly inactivates Xa by forming a complex, which, in turn is a feedback inhibitor of the activation of X and IX by TF-VIIa ①,②. This avoids starting the coagulation cascade, if the triggering cause is too weak.

Although the concentration of free TFPI in plasma is very low, TFPI is almost unlimitedly available (liberation of a membrane-bound form by heparin, secretion from endothelial cells). Therefore most – if not all – of the VIIa-TF complexes are blocked this way.

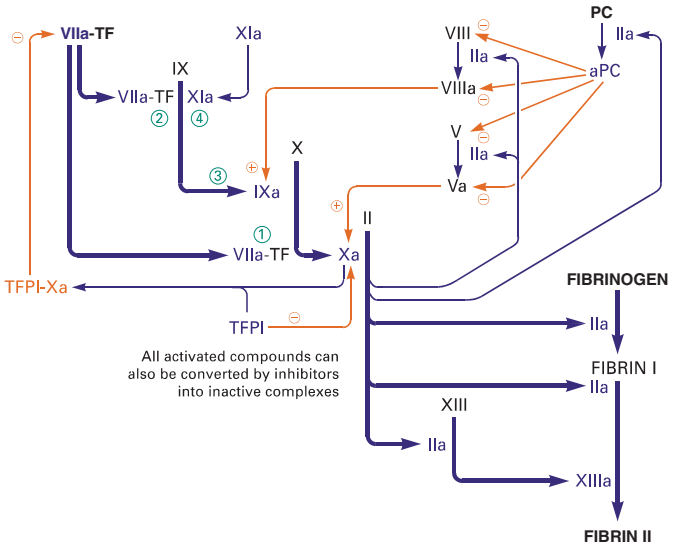


Figure 9.3-1. Coagulation Cascades

9.3.4 Controlled Propagation (Fig. 9.3-2)

The progress of the coagulation cascades depends essentially on the activation of sufficient factor IX to IXa by TF-VIIa ②), before the inhibition of this reaction by Xa-TFPI becomes effective. IXa, in turn, causes activation of X to Xa ③ and thus circumvents the TFPI inhibition of X activation. Insufficient activity of this pathway causes bleeding: In hemophilia A, cofactor VIII and in hemophilia B, factor IX is lacking.

In later stages of coagulation, generation of more thrombin is essential for proper clot formation. This is accomplished by a supplementary pathway for IX activation by XIa ④. The physiological activator of XI is unknown (9.2).

Another important feedback control system of coagulation is initiated by thrombin itself (Fig. 9.3-2, bottom right): Binding to thrombomodulin (TM, present at the membrane of intact endothelial cells) converts thrombin into an anticoagulant. It proteolytically activates protein C (PC) to aPC ⑤, which in a complex with protein S (PS) inactivates the essential cofactors of the coagulation cascade VIIIa and Va as well as their unactivated precursors. These inhibitory reactions of coagulation are also controlled: the quantity of available aPC is limited by complex formation with PC-inhibitor, α_1-proteinase inhibitor or α_2-macroglobulin ⑥. PS can be proteolytically inactivated by thrombin.

A recently discovered mutant of cofactor V (Factor V Leiden) cannot be inactivated by aPC. In humans with the factor V Leiden, thrombin generation is less restricted. About 30% of all thrombotic events are caused by this mutant. High levels of homocysteine (3.2.5.4, causing damage to the endothelial cell layer) are another source of thrombotic events.

Proteins C, S and others also play a role in, e.g., inflammation and cell growth.

Thrombin can be inactivated by complex formation with heparin cofactor II (HC-II) or AT III ⑦. These reactions are mediated by dermatan sulfate or heparan sulfate, or by heparin, respectively.

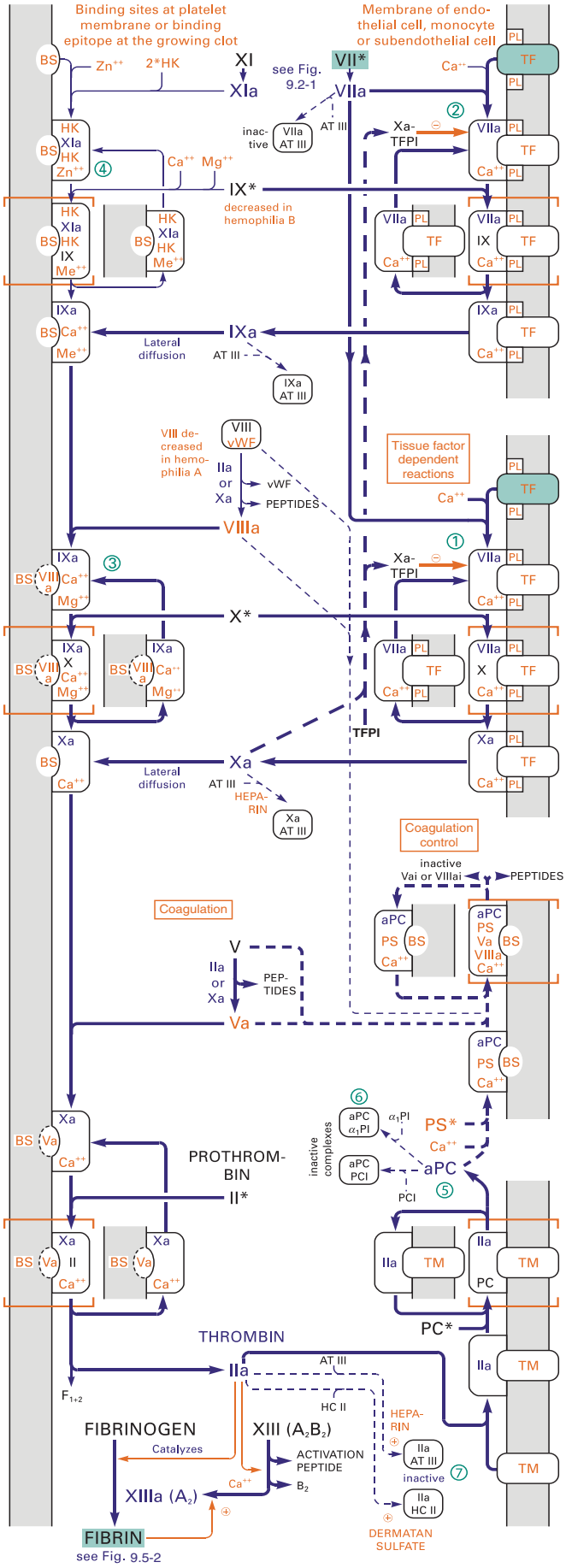


Figure 9.3-2. Coagulation Propagation and Control

The presence of TM and PS receptor at the surface of intact endothelial cells is the cause of their anticoagulant properties. If lesions or noxious agents disturb these cells, procoagulant effects come into action: establishment of binding sites, secretion of TF (9.2), von Willebrand factor and PAI-1 (9.5), loss of thrombomodulin and PS-receptor.

9.3.5 Generation of Fibrin (Figs. 9.3-3, 9.3-4 and 9.5-2)

Fibrinogen consists of 2 sets of triple protein helices $[(A\alpha)\ (B\beta)\ \gamma]_2$, which are symmetrically joined in the center by disulfide bonds. To effect coagulation, thrombin has to release four strongly negatively

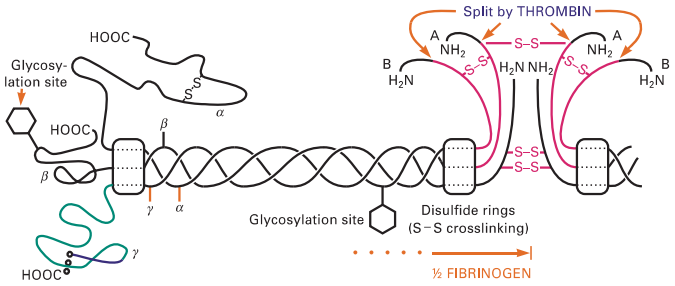


Figure 9.3-3. Fibrinogen Structure

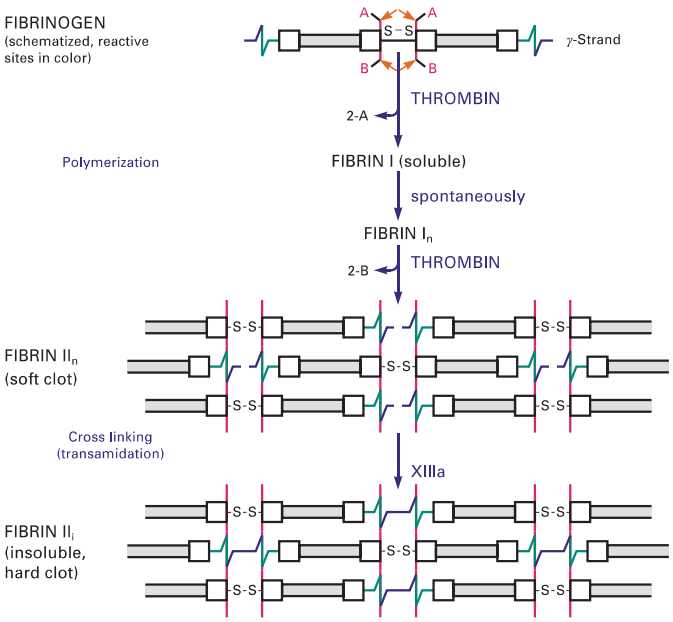


Figure 9.3-4. Formation of Fibrin Polymers

Table 9.3-2. Properties of Coagulation Factors

Abbreviation	Factor	Trivial Name	Mol.Wt. (kDa)	Conc. (mol/l)	Acts (in activated form) on	Pathology
Proteases and Precursors						
II/IIa	factor II/IIa	(pro-) thrombin	72/37	1.4	V, VIII, Fibrinogen, XIII, PC	hypoprothrombinemia
VII/VIIa	factor VII/VIIa	(pro-) convertin	50/50	0.01	IX, X	hypoproconvertinemia
IX/IXa	factor IX/IXa	Christmas factor	57/46	0.09	X; VII	hemophilia B
X/Xa	factor X/Xa	Stuart-Prower-factor	59/46	0.17	II; VII, V, VIII	Stuart disease
XI/XIa	factor XI/XIa	plasma thromboplastin	160/160	0.04	IX	hemophilia C
XII/α-XIIa/β-XIIa	factor XII/α-XIIa/β-XIIa	Hageman factor	80/80/30	0.4	PK, 110 kDa proactivator; XI	Hageman trait
PK/KK	plasma (pre) kallikrein	Fletcher factor	86/86	0.6	XII, HK, scu-PA	Fletcher trait
PC/aPC	protein C		62/56	0.07	Va, VIIIa	thromboembolism
Other Enzymes and Precursors						
XIII/XIIIa	factor XIII/XIIIa	fibrin stabilizing factor	320/160	0.1	fibrin	hemorrhagic diathesis
Protease Inhibitors						
α_1PI	α_1-proteinase inhibitor	α_1-antitrypsin	55	50	XIa, aPC	
α_2M	α_2-macroglobulin		720	0.003	IIa; KK; aPC	
AT III	antithrombin III		58	2 … 3	IIa, Xa; IXa,α-XIIa, VIIa-TF	thromboembolism
C1-INH	complement factor C1 inhib.		110	1.8	XIIa(α+β), KK; XIa	hereditary angioedema
HC II	heparin cofactor II	antithrombin BM	66	1.2	IIa	thromboembolism (rare)
PCI	protein C inhibitor (= PAI-3)		56	0.1	aPC	
PN II	protease nexin II		51	(platelets)	XIa; IXa?	
TFPI	tissue factor pathway inhibitor	lipid associated coagulation inhibitor	43/41/34	0.003	Xa, VIIa	disseminated intravascular coagulation
Cofactors and Receptors						
V/Va/Vai	factor V/Va/Vai	proaccelerin	330/180	0.02	Xa	APC-resistance
VIII/VIIIa/VIIIai	factor VIII/VIIIa/VIIIai	antihaemophil. globulin A	285/165	0.0007	IXa	hemophilia A
BS	'binding site'		–	–		
Fg/Fb	fibrinogen/fibrin, factor I/Ia		340/333	7	thrombospondin, XIII, scu-/tcu-PA, plasminogen	dysfibrinogenemias
HK/HKi	high mol.-weight kininogen	Fitzgerald factor	110	0.7	PK/KK, XI/XIa	Fitzgerald trait
LS	lesion site					
PS	protein S		75	0.3	aPC	thromboembolism
TF	tissue factor	thromboplastin	45	(membrane)	VII/VIIa	
TM	thrombomodulin		100	(membrane)	IIa	[lethal genetic defect]
	von-Willebrand factor (vWF)	Ristocetin cofactor	(2 … 100) * 250		platelets/endothelial cells	vW's disease

charged peptides (2 * A and 2 * B) from this central area. This way, electric repulsion is removed and (with the new N-termini) high-affinity ligands for the binding epitopes at both ends of the fibrinogen molecule are generated. This allows self-assembly of fibrin and formation of polymers (fibrin I_n and thereafter II_n). Then cross-linking of specific lysine and glutamate residues at the C terminals of the γ-strands is effected by thrombin-activated factor XIIIa (which is not a protease, but a transglutaminase). This converts the clot into an insoluble form and causes contraction. Fibrin and fibrinogen also have binding epitopes for other proteins for crosslinking with platelets (9.4) and other cells. This crosslinking is the central mechanism to rapidly stop bleeding.

9.4 Platelets (Thrombocytes)

Platelets (Fig. 9.4-1) play a central part in the hemostatic process. When injuries to the vessel wall cause a tearing of the endothelial layer, the subendothelial collagen gets exposed. Platelets adhere to it:

- directly via integral membrane glycoproteins (GP) acting as receptors, e.g., GPIa–IIa, GPIIIb and GPVI (Table 9.4-2).
- indirectly via linker proteins present in blood plasma (e.g., von Willebrand factor, fibronectin) or by release from the endothelium (e.g., von Willebrand factor, fibronectin) or from platelets after activation (e.g., thrombospondin, TSP, Table 9.4-1).

The contact to the subendothelium activates the platelets. The granule membranes fuse with the platelet membrane, causing release of the contents of secretory organelles (dense bodies, α-granules) and integration of α-granule membrane proteins (e.g., P-selectin) into the platelet membrane. This has multiple consequences (Fig. 9.4-1):

- Vascular effects: ADP, serotonin (from dense bodies) and epinephrine quickly cause vasoconstriction, limiting the blood flow and assisting the thrombus formation.

  They also induce the production of NO (7.8.2, formerly called endothelium-derived relaxing factor, EDRF), which has strong vasodilatory effects, thus limiting the primary actions.

- Change of platelet shape: The microtubule bundles (which keep up the normal platelet form) depolymerize. Binding of ADP to its receptors effects release of Ca^{++} ions from dense bodies into the cytosol, where they interact with calmodulin (IP_3 pathway, 7.4.4). This causes activation of the myosin light chain kinase (7.4.5), which, in turn, activates myosin. The activated myosin interacts with the actin filaments of the cytoskeleton, changing the platelet shape from a discoidal to a globular form with long pseudopods.

- Platelet aggregation and clot formation**:** The released factors (e.g., ADP, thromboxane A_2 and platelet activating factor, PAF, 3.4.3.3) enhance the activation of surrounding platelets. By formation of fibrinogen or thrombospondin bridges to other platelets via the activated GPIIb–IIIa receptor complex, platelets aggregate and provide the primary hemostatic plug.

  On formation of the first traces of thrombin by the coagulation cascade (9.3), these fibrinogen-platelet aggregates are converted into a stable fibrin-platelet network that also includes other blood cells. Even small amounts of thrombin activate factor XIII (from plasma and from α-granules), which catalyzes the covalent crosslinking inside of the fibrin clot.

- Microparticle formation: Platelet activation leads to the disintegration of the platelet membrane and to the formation of microparticles. In platelets, phosphatidylserine appears at the outer surface due to inhibition of the flip-flop mechanism (9.2.3) following activation. Factor V (from plasma and from α-granules) as well as vitamin K dependent coagulation factors (factors II, IX, X from plasma) bind at the procoagulant surface (9.2-3) of these particles. This provides effective progress of the coagulation and assures quick sealing of the vessel wall after injury.

- Control of coagulation: Not only the coagulation progress, but also its feedback inhibition via protein C and protein S (9.3.4) depend on the modified phospholipid bilayer structures (9.2.3) present in, e.g., activated platelets and microparticles. In addition, the intact endothelium synthesizes substances that inhibit platelet function and prevent unbalanced platelet activation. The most important inhibitory factors are prostacyclin I_2 and D_2 (7.4.8), ADP-decomposing nucleotidase and nitric oxide (NO, 7.8.2). All these inhibitors cause an increase of the intracellular cAMP concentration. This effects the sequestration of Ca^{++} ions into intracellular storage areas (7.4.4). This way, all processes based on calcium ions are inhibited.

The major contents of the platelet vesicles and the receptors at the membrane are listed in Table 9.4-1 … -2.

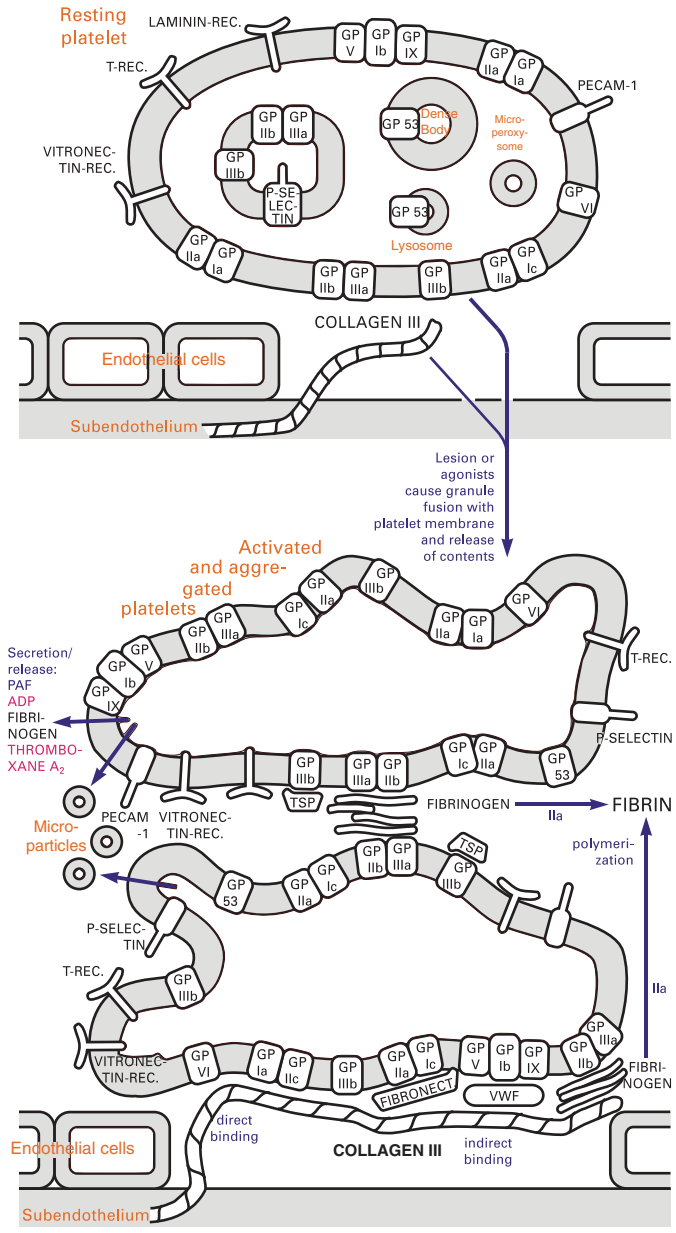


Figure 9.4-1. Role of Platelets in Coagulation

Table 9.4-1. Compounds Stored in Vesicles of Thrombocytes

Storage Vesicles	Compounds Involved in Aggregation	Compounds Causing Vasoconstriction	Other Compounds
α-Granules	fibrinogen, fibronectin, thrombospondin, von Willebrand factor (vWF), factors V XIII, multimerin (stabilizes factor V)		platelet factor 4 (PF4, suppresses anticoagulation), β-thromboglobulin (β-TG), α_2-antiplasmin
Dense bodies	Ca^{++}, ADP	ADP, serotonin, epinephrine	
Lysosomes			hydrolyzing enzymes (involved in inflammation)
Microperoxisomes			catalase (involved in inflammation)

Table 9.4-2. Receptors on the Platelet Membrane
The glycoproteins usually have large extracellular domains. These are not shown in detail in Figure 9.4-1.

Platelet Glycoprotein (Complexes)	CD-No.	Family	Mol. Mass kDa	Copies (resting platelet)	Receptor for	Physiological Function
GP Ib/IX/V	42b,c/a/d	Leu-rich glycoproteins (LRG)	143 (Ib), 22 (Ibβ)/22/82	Ib:25,000 IX:25,000 V:11,000	vW factor (by GP Ib_a), thrombin (by GP Ib_a)	shear stress induced adhesion of unactivated platelets to the subendothelium of injured vessels, modulation of thrombin induced activation, activation of GP IIb/IIIa for binding of vWF
GP Ia/IIa	49b/29	$\alpha_2\beta_1$ integrin	153/130	Ia:2000 IIa:7000	collagen, laminin in fibroblasts and endothelial cells, fibronectin	collagen receptor, collagen induced adhesion (Mg^{++} dependent), activation of tyrosine kinase adhesion of platelets to fibronectin
GP Ic/IIa	49e/29	$\alpha_5\beta_1$ integrin	160/130	Ic: ? IIa:7000		
GP Ic*/IIa		$\alpha_6\beta_1$ integrin		Ic[1]: ? IIa:7000	laminin	adhesion of platelets to laminin
GP IIb/IIIa	41/61	$\alpha_{IIb}\beta_3/\alpha_v\beta_3$ integrin	125(IIba), 23 (IIbb)/95	40,000 ... 80,000	fibrinogen, vW factor, fibronectin, thrombospondin, vitronectin	after activation: irreversible attachment of platelets to the subendothelium via vWF and fibronectin, aggregation of platelets, signal transmission
GP IIIb	36	scavenger receptor (6.2-4)	88	12,000 ... 19,000	collagen type I, thrombospondin, oxidized LDL	binding between platelets, endothelial cells and monocytes, signal transmission, receptor for oxidized LDL
GP VI			62	?	collagen types I, III, V, VI, VIII (fibrillar collagens)	collagen receptor, collagen-induced signal transduction
GP 53	63					?
Laminin receptor			67		laminin	adhesion of platelets to laminin
Thrombin receptor		G-protein (7.3)	47	1800	thrombin	initiation of platelet aggregation and secretion, activation of PL C, inhibition of adenylate cyclase
Vitronectin receptor		$\alpha_\omega\beta_3$ integrin		500	vW factor, fibronectin, vitronectin, thrombospondin	adhesion of cells to the extracellular matrix, mediates cell spreading, phagocytosis and intercellular interactions (virus entrance into platelets and endothelial cells?)
PECAM-1	31	immunoglobulin (8.3)	130	5000	PECAM-1 (auto-association)	
P-Selectin	62 P	selectin (8.3)	140		neutrophils granulocytes, monocytes	mediates contact of activated platelets and endothelial cells, to other cells (incl. tumor cells)
ADP receptor	not yet identified		43 ?		ADP	G_{ia} inhibition of adenylate cyclase
Thromboxane receptor(s)		possibly 2 receptors	37		thromboxane A_2	G protein activation of PLC and Ca^{++} channel
Receptors for adenosine, PGI_2, PGE_2, PGD_2					adenosine, PGI_2 PGE_2, PGD_2	G protein activation of adenylate cyclase

[1] different from GP 1c.

Table 9.5-1. Properties of Fibrinolytic Factors

Abbreviation	Factor	Mol.Wt. (kDa)	Acts (in activated form) on
Precursors and Proteases			
Sct-PA/tct-PA	single/two chain tissue type plasminogen activator	70	plasminogen
scu-PA/tcu-PA	single/two chain urokinase type plasminogen activator	55/34	plasminogen
Plg/Pl	plasminogen/plasmin	90/78	fibrin (plasmin dissolves)
Protease Inhibitors			
α2AP	α2-antiplasmin	ca. 70	plasmin
α2M	α2-macroglobulin	720	plasmin
PAI-1	plasminogen	52	Sct-PA, tct-PA
PAI-2	activator	60 (glycosyl.)	tcu-PA
PAI-3	inhibitors	56	tcu-PA
PN II	protease nexin II	51	sct-PA, tct-PA
HRGP	histidine rich glycoprotein	60	plasminogen
Cofactors and Receptors			
BE	'binding epitope' at fibrin surface		

9.5 Fibrinolysis

The fibrinolytic system has the task of removing fibrin deposits and thrombo-embolic fibrin plugs in the blood vessels. It fulfills important control functions and forms a dynamic equilibrium with the coagulation system. Fibrinolysis is controlled by activation and inhibition processes. Specific activators convert the inactive proenzyme plasminogen into the active serine protease plasmin, which degrades the fibrin clot (Fig. 9.5-1).

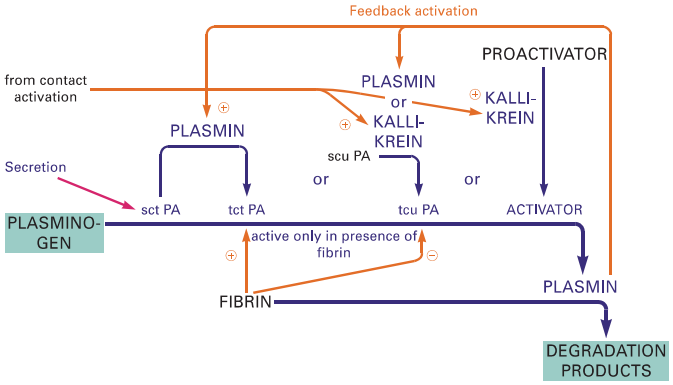


Figure 9.5-1. Scheme of Fibrinolytic Activation

9.5.1 Pathways of Plasminogen Activation (Table 9.5-1, Fig. 9.5-2)

Plasminogen can be activated via several pathways:

- Small quantities of tissue plasminogen activator (t-PA) are permanently liberated from the endothelial layer of the blood vessels. Furthermore, acute stimuli (e.g., vascular congestion due to thrombosis) cause secretion of t-PA from intracellular depots. t-PA requires fibrin for effective conversion of plasminogen into plasmin. Therefore it is inoperative in the absence of a fibrin clot.

  The physiological concentration of plasminogen is about 1.5 µmol/l. The K_M of t-PA for plasminogen in the absence of fibrin is 65 µmol/l and 0.15 µmol/l in its presence. Thus, the presence of fibrin ensures quick activation.

  There is strong evidence that t-PA is essential for starting the fibrinolytic system. At first, small quantities of plasminogen are converted into plasmin. This causes cleavage of some fibrin chains and exposition of lysine residues, which are attachment sites for more plasminogen and t-PA. This potentiates the formation of plasmin. Although t-PA is already fully active in its single chain form (sct-PA), it is converted by plasmin into the two-chain form (tct-PA).

- Urokinase type plasminogen activator (u-PA) occurs in the circulation only in a weakly active single-chain form (scu-PA). Its physiological source is unknown. In cell cultures it was found in the supernatant of endothelial, epithelial and tumor cells. Plasmin converts scu-PA into the much more active 2-chain form tcu-PA. Both scu-PA and tcu-PA act only on plasminogen, which is bound to fibrin.

  Presumably, the intravascular u-PA pathway becomes operative after some plasmin has been obtained by the tPA pathway. It presumably acts as an amplification system for the essential tPA pathway.

  Additionally, urokinase is involved in many extravascular processes, e.g., wound healing (removal of the fibrin layer, neovascularization, development of organs, tumor invasion/metastasis and degradation of the extracellular matrix by plasmin).

  Streptokinase (a bacterial toxin from *Streptococcus haemolyticus*, which is therapeutically used) forms a complex with plasminogen, which also acts as a plasminogen activator.

- The contact activation of coagulation (9.2) represents another system of initiating fibrinolysis. It generates kallikrein, which converts a (still not completely characterized) 110 kDa proactivator

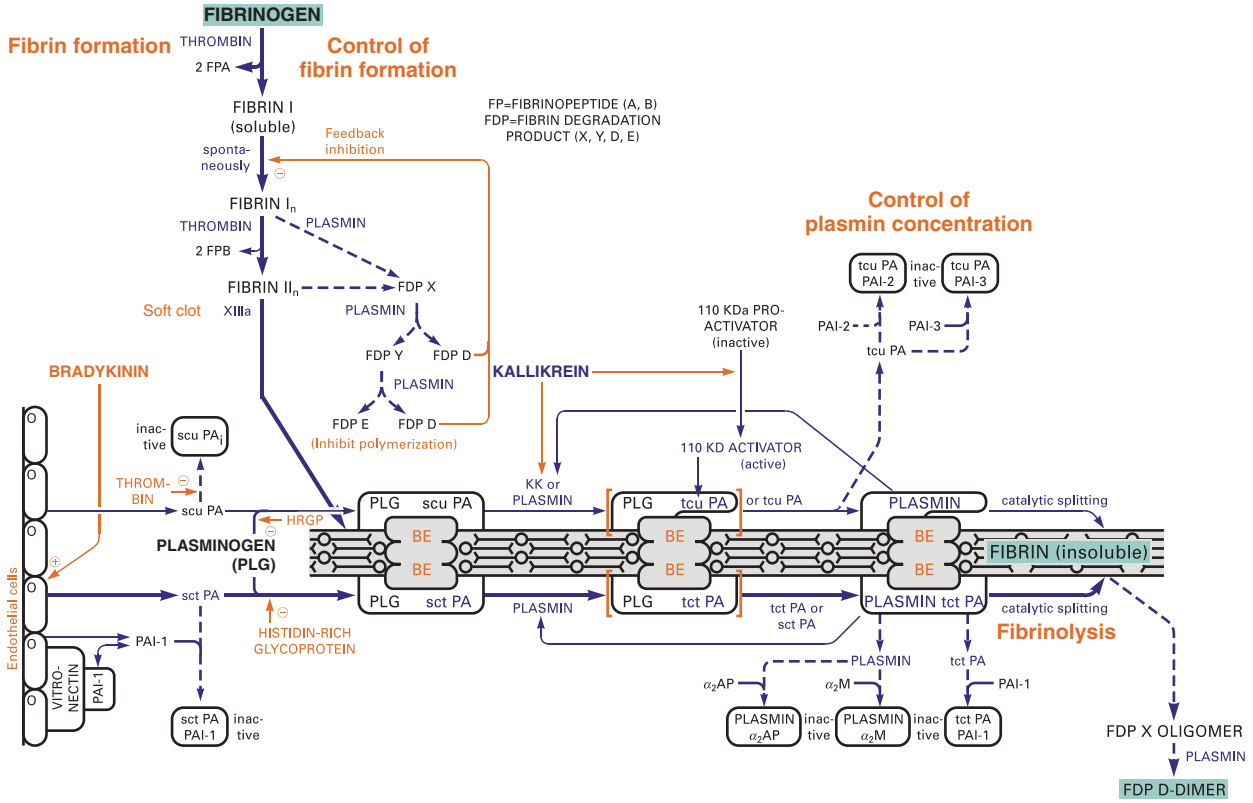


Figure 9.5-2. System of Fibrinolysis

into an active plasminogen activator, resulting in plasmin formation. Additionally, kallikrein activates scu-PA to tcu-PA, while bradykinin (9.2.2) stimulates the liberation of t-PA from intracellular depots in the endothelium. Thus this system acts as another amplification pathway of fibrinolysis.

9.5.2 Control of Fibrinolysis (Fig. 9.5-2)

Although plasmin has no definite substrate specificity (similar to trypsin), under physiological conditions its fibrinolytic activities are strictly confined to fibrin. Likewise, t-PA requires the presence of fibrin in order to generate the fibrin-degrading plasmin.

Circulating free plasmin and free activators (except scu-PA) are effectively controlled by the respective inhibitors. These inhibitors (PAI-1, -2, -3; α_2-antiplasmin, α_2-macroglobulin) form inactive complexes with unbound fibrinolytic proteases.

It is unlikely that in plasma only PAI-1 is responsible for the control of t-PA. PAI-2 occurs in plasma only in special situations, such as pregnancy or certain malignancies. The active form of PAI-1 is conserved for extended periods when bound to vitronectin (Table 9.4-2). Otherwise, it is quickly converted into the latent form. This form can be reactivated by binding to anionic phospholipids, e.g., from microparticles.

Binding of proteases to fibrin protects the enzymes from inhibition. Therefore, fibrin is not only the target of the fibrinolysis, but also the central cofactor, which determines initiation, location and extent of this event.

Summary: Recent results initiated a profound rethinking of old models of hemostasis. They indicated that the contact pathway initiated by factor XII, kallikrein and factor XI, which determines the outcome of the *in vitro* coagulation assays, plays a quite different role *in vivo.* This refers to the propagation phase of blood clot formation and possibly to inflammatory and defense processes in response to pathogens. Interestingly, such a role is evolutionarily related to the antimicrobial role of the blood coagulation system in horseshoe crabs.

Literature:

Blockmans, D. *et al.* Blood Reviews 1995;9:143–156.

Bombeli, T. *et al.* Thrombosis Hemostasis 1997;77:408–423.

Collen, D. Thrombosis Hemostasis 1980;43:77–89.

Colman, R.W. *et al.* Blood 1997;90:3819–3843.

Degen, J.L. *et al.* J Thromb Haemost. 2007;5 Suppl 1:24–31.

Kalafatis, M. *et al.* Biochimica et Biophysica Acta 1994;1227: 113–129.

Lane, D.A. *et al*. Thrombosis Hemsostasis 1996;76:651–662.

Lorand, L., Mann, K.G. (Eds.) *Methods in Enzymology* Vol. 222. Academic Press, 1993.

Opal, S.M, Esmon, C.T. Crit Care. 2003;7:23–38.

Petersen, L.C. *et al.* Thrombosis Research 1995;79:1–47.

Renné, T. *et al.* J Thromb Haemost. 2009;7 Suppl 1:79–83.

Renné, T. Gailani, D. Expert Rev Cardiovasc Ther. 2007;5:733–741.

Rijken, D.C. Lijnen, H.R. J Thromb Haemost. 2009;7:4–13.

Schmaier, A.H. J Clin Invest. 2008;118:3006–3009.

Zwaal, R.F.A. *et al*. Blood 1997;89:1121–1132.

10 Biochemical Networks, Bioinformatics and Systems Biology

Dietmar Schomburg

This book is intended to provide a summary of biochemistry with special focus on metabolic pathways. More detailed information can be obtained from a number of international databases concentrating on enzymes or metabolic pathways. A survey is given below.

10.1 Systems Biology and Networks

Whereas the assignment of specific capabilities to organisms in a specific environment has been the topic of experiments mostly concentrating on a single biochemical reaction, in recent years the advent of the '-omics' techniques (genomics, transcriptomics, proteomics, metabolomics, etc.) together with the tools of bioinformatics has led to an entirely new approach in the biological sciences. Many biochemical functions are treated as what they really are: Reactions of complicated biological networks and structures, which adapt to environmental conditions by the use of genetic information and tightly controlled mechanisms.

The methods of bioinformatics allow the prediction of biochemical functions of proteins, making use of sequence similarities, sequence patterns and genome positions. Whereas for some organisms like *E. coli* this automatic genome annotation has now reached a high accuracy, in many other organisms the function of about one third of the products can be predicted with high reliability, another third with limited accuracy, and one third of the functions are still unknown. In some organisms like archaea the situation is worse still. Nevertheless it is now possible to recognize which pathways are really present in which organism classes.

The methods of bioinformatics are accompanied by extremely fast progress in many analytical and experimental high-throughput techniques. Whereas current genome sequencing methods allow the sequencing of a microbe in half a day, even faster machines are under development, which claim to sequence human genomes in a stunning three minutes. Transcriptome analysis or deep sequencing techniques allow the measuring of mRNA levels and the determination of how the organism responds at the regulatory level to any change in its environment. Proteome analysis tells us which protein repertoire is currently present in the cell. Metabolome analysis measures the concentrations of several hundred metabolites in parallel and may inform us about missing enzymes and missing links in our predicted networks. By making use of isotope labeling, a calculation of metabolic fluxes, at least in central metabolism, is possible.

10.2 Modeling of Metabolic Fluxes

Whereas the experimental methods allow an ever deeper insight into the cellular contents and capacities, the real understanding and even prediction of cellular processes definitively require the construction and application of network models, making use of simulation techniques similar to the ones engineers use to predict the functions of airplanes or complicated electronic devices.

Model construction is always based on the analysis of the genome and corresponding experiments. For metabolic models the enzymes are predicted from the genome, supplemented by experimental data provided in one of the databases, connected with chemical reactions, and finally used for analysis and simulation.

Once a rather reliable and complete model of a biological system (regulatory pathway, metabolic pathway, full genome-scale network, or in the future finally the whole cell) is available, three different levels of modeling/simulation approaches are possible:

- Topological analysis of the network or Boolean modeling. Here a signal or a reaction is either in the 'on' or 'off' state. Critical points in regulation including possible drug targets or metabolic bottlenecks can be predicted with this method.
- Quantitative metabolic flux analysis making use of reaction stoichiometries, the assumption of a temporary flux equilibrium and biological assumptions, e.g., on fitness. Metabolic fluxes, capacities and reactions to environmental and genetic changes can be predicted.
- Fully time-resolved quantitative simulation of biochemical networks. This method is the most sophisticated one making use of ordinary differential equations but the tens of thousands of parameters for larger networks will not be available for the foreseeable future. Nevertheless for smaller systems the method is already used.

All modeling approaches are used to design experiments that can refine the models and finally be used for the development of new drugs, for biotechnological product optimization or even for 'synthetic biology', the construction of new biological tools for many different applications.

The task of putting together the huge metabolic network described in this book was in the past a cumbersome but fascinating manual piecework. Now the information is used for the automatic construction of biological models and will allow a much higher degree of sophistication in biological research.

10.3 Biochemical Pathways Information Resources

Meanwhile many hundreds of different databases have been developed worldwide. Whereas many of these have only a limited lifetime, a number of long-term stable ones exist. A complete enumeration is beyond the scope of this chapter. Therefore the selection presented here necessarily reflects a subjective priority list. The European Bioinformatics Institute (EBI) houses a number of databases on sequences, structures, domains, interactions, etc.

10.3.1 Overview

With respect to the areas covered in this book a number of different databases are of interest:

- Sequence databases: The results of sequencing projects are stored as gene or whole genome sequences in the primary databases in an international collaboration of the EMBL nucleotide database with GenBank (USA) and the DNA Database of Japan (DDBJ). Each of the three groups collects a portion of the total sequence data reported worldwide, and all new and updated database entries are exchanged between the groups on a daily basis. Only very basic information is added to the sequence. The corresponding protein sequences are stored in UniProt, where the TrEMBL section is basically a translation of the nucleotide sequence and the Swiss-Prot section provides an annotated view on the sequence (including sequence related protein features, e.g., active centers or glycosylation sites). Functional protein domains are calculated and stored in databases like Pfam or Interpro.
- Protein function and property databases:
  - Enzymes are classified according to the EC-system, based on their catalyzed reaction, by a commission organized by the IUBMB. The structure of this system is presented in Section 2.4.5. This system is used by a number of databases (BRENDA, ExplorEnz, IntEnz, SIB-ENZYME Nomenclature Database, etc.).

Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Second Edition. Edited by Gerhard Michal and Dietmar Schomburg.

  - Membrane transport proteins are classified by the Transport Commission's TC-system, a system analogous to the EC system. Its setup is described in Section 6.1.1.2.
- Special protein function databases exist for enzymes: BRENDA deals with all enzyme classes, MACiE covers enzyme mechanisms, Merops proteases, Rebase restriction enzymes, CAZy carbohydrate-active enzymes and others.
- Protein structure databases: Coordinates or protein 3D structures are stored in the PDB. The CATH and SCOP databases provide a classification scheme for protein 3D structures
- Pathway databases: Metabolic pathways are stored and provided in the KEGG and the MetaCyc databases. The Reactome database stores information on a wide variety of biological pathways.
- Small molecule information: Chemical structures of small organic molecules and information on their biological activities are provided in the databases PUBCHEM and Chebi. Whereas the first one is user-driven, the second is a manually curated dictionary of compounds of biological interest.
- Genetic diseases: A compendium of human genes and genetic diseases is given in OMIM.
- Protein family databases: Numerous enzyme databases on the web are specialized in the analysis of protein and gene sequences for enzyme groups. Examples are:
  - The ESTHER Database is dedicated to the analysis of protein and nucleic acid sequences belonging to the superfamily of α/β hydrolases which are homologous to cholinesterases http://bioweb.ensam.inra.fr/ESTHER/definition),
  - PeroxiBase aims to compile most of the peroxidase superfamilies encoding sequences, to follow the evolution of peroxidases among living organisms, and to compile the information concerning putative functions and transcription regulation (http://peroxibase.isb-sib.ch/index.php),
  - KinBase holds information on over 3,000 protein kinase genes found in the genomes of humans and many other sequenced genomes. It explores the functions, evolution and diversity of protein kinases, the key controllers of cell behavior with a focus on the kinome, the full complement of protein kinases in any sequenced genome (http://kinase.com/).

10.3.2 Detailed Description of Some Databases

GenBank / EMBL nucleotide database / DDBJ

The contents of all three databases are practically identical, due to a daily exchange of data. GenBank is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences. As of August 2009, there are approximately 106,533,156,756 bases in 108,431,692 sequence records in the traditional GenBank divisions and 148,165,117,763 bases in 48,443,067 sequence records in the WGS division. In addition to the sequence, the source organism, its chromosomes and literature references are listed, but very little additional information is given. The very fast growth of sequence information can be seen from Figure 10.3-1.

UniProt protein sequences

The Universal Protein Resource (UniProt) is a comprehensive resource for protein sequence and annotation data. The UniProt databases are the UniProt Knowledge Base (UniProtKB), the UniProt Reference Clusters (UniRef), and the UniProt Archive (UniParc). The UniProt Metagenomic and Environmental Sequences (UniMES) database is a repository specifically developed for metagenomic and environmental data (Fig. 10.3-2).

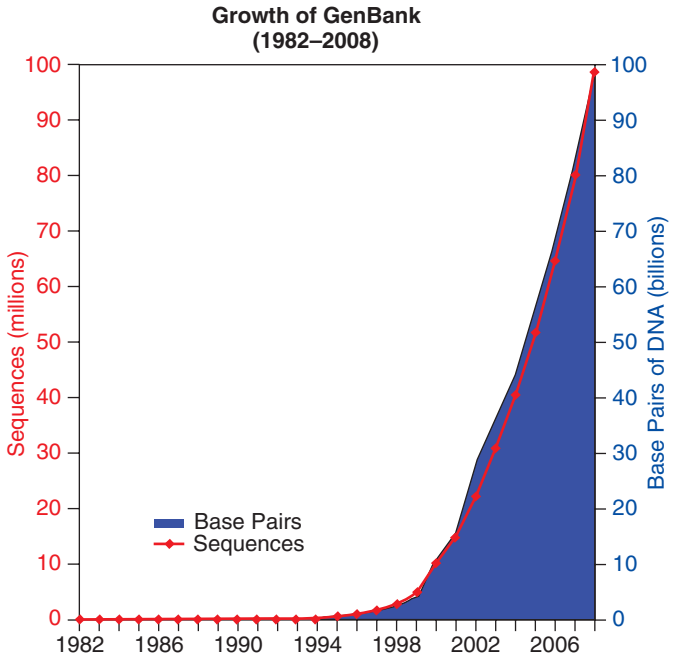


Figure 10.3-1. Growth of Nucleotide Sequence Databases

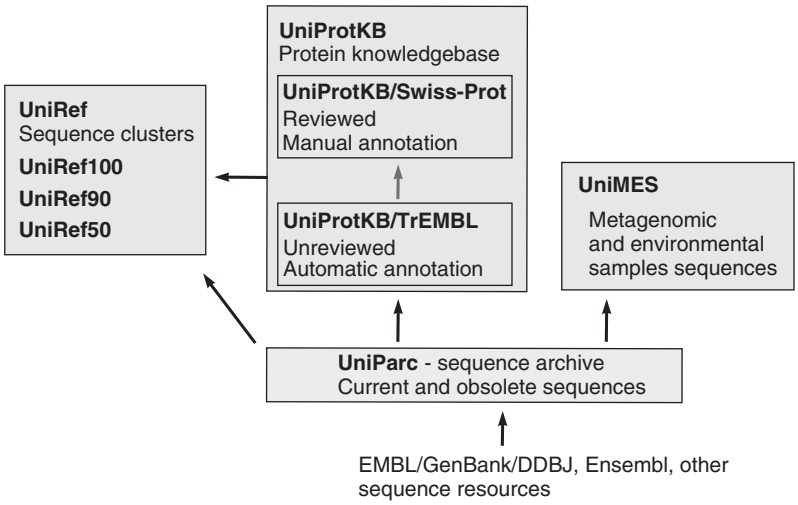


Figure 10.3-2. UniProt Overview

UniProt is a collaboration between the European Bioinformatics Institute (EBI), the Swiss Institute of Bioinformatics (SIB) and the Protein Information Resource (PIR). In the three institutes close to 150 people are involved in performing different tasks such as database curation, software development and support.

UniProt is divided in two sections, TrEMBL and Swiss-Prot. The TrEMBL section of UniProtKB was introduced in 1996 in response to the increased dataflow resulting from genome projects. It was already recognized at that time that the traditional time- and labor-consuming manual annotation process (which is the hallmark of Swiss-Prot) could not be broadened to encompass all available protein sequences. Publicly available protein sequences obtained from the translation of annotated coding sequences in the EMBL-Bank/GenBank/DDBJ nucleotide sequence database are now automatically processed and entered in UniProtKB/TrEMBL where they are computed-annotated in order to make them swiftly available to the public.

UniProtKB / TrEMBL contains high quality computationally analyzed records that are enriched with automatic annotation and classification. The UniProtKB/TrEMBL unreviewed entries are kept separate from the UniProtKB/Swiss-Prot manually reviewed entries so that the high quality data of the latter is not diluted in any way.

Release 2010_08 of Swiss-Prot contains 518,415 sequence entries, comprising 182,829,264 amino acids abstracted from 190,192 references. Protein existence assumptions were based on evidence at the protein level 13.5%, evidence at the transcript level 12.9%, inferred from homologies 70.5% and predicted or uncertain 3.1%.

12,146 species are presently represented in UniProtKB/Swiss-Prot. Among them are about 1000 organisms which have more than 100 listed sequences. 4% of the sequences are from archaea, 62% from bacteria, 31% from eukarya, and 3% from viruses.

A position-dependent annotation is included in Swiss-Prot. It includes, among others, the following features:

- Propeptide: Part of a protein that is cleaved during maturation or activation
- Beta strand, Turn or Helix: Secondary structure regions within the experimentally determined protein structure
- Cross-link: Residues participating in covalent linkage(s) between proteins
- Disulfide bond: Cysteine residues participating in disulfide bonds
- Glycosylation: Covalently attached glycan group(s)
- Lipidation: Covalently attached lipid group(s)
- Coiled coil: Positions of regions of coiled coil within the protein
- Region: Region of interest in the sequence
- Natural variant: Description of a natural variant of the protein
- Alternative sequence: Amino acid change(s) producing alternate protein isoforms
- Modified residue: Modified residues excluding lipids, glycans and protein crosslinks
- Binding site: Binding site for any chemical group (co-enzyme, prosthetic group, etc.)
- Metal binding: Binding site for a metal ion
- Active site: Amino acid(s) directly involved in the activity of an enzyme
- Motif: Sequence motif of biological interest
- DNA binding: Position and type of a DNA-binding domain
- Nucleotide binding: Nucleotide phosphate binding region
- Signal: Sequences targeting proteins to the secretory pathway or to the periplasmic space
- Zinc finger: Position(s) and type(s) of zinc fingers within the protein
- Calcium binding: Position(s) of calcium binding region(s) within the protein
- Domain: Denotes the position and type of each modular protein domain
- Transmembrane: Extent of a membrane-spanning region
- Topological domain: Location of non-membrane regions of membrane-spanning proteins
- Intramembrane: Extent of a region located in a membrane without crossing it
- Peptide: Extent of an active peptide in the mature protein
- Transit peptide: Extent of a transit peptide for organelle targeting

InterPro

InterPro is an integrated documentation resource for protein families, domains, regions and sites. InterPro combines a number of databases (referred to as member databases) that use different methodologies and a varying degree of biological information on well-characterized proteins to derive protein signatures. By uniting the member databases, InterPro capitalizes on their individual strengths, producing a powerful integrated database and diagnostic tool (InterProScan).
The member databases use a number of approaches:

- ProDom: provider of sequence-clusters built from UniProtKB using PSI-BLAST.
- PROSITE patterns: provider of simple regular expressions.
- PROSITE and HAMAP profiles: provide sequence matrices.
- PRINTS provider of fingerprints, which are groups of aligned, unweighted Position Specific Sequence Matrices (PSSMs).
- PANTHER, PIRSF, Pfam, SMART, TIGRFAMs, Gene3D and SUPERFAMILY are providers of hidden Markov models (HMMs).

While all of the methods share a common interest in protein sequence classification, some focus on divergent domains (e.g., Pfam), some focus on functional sites (e.g., PROSITE), and others focus on families, specializing in hierarchical definitions from superfamilies down to subfamily levels in order to pin-point specific functions (e.g., PRINTS). TIGRFAMs focus on building HMMs for functionally equivalent proteins. PIRSF uses HMMs over the full length of a protein and domain HMMs to gather family members. HAMAP profiles are manually created by expert curators. These profiles identify proteins that are part of well-conserved bacterial, archaeal and plastid-encoded proteins families or subfamilies. PANTHER builds HMMs based on the divergence of function within families. SUPERFAMILY and Gene3D are based on structures using the SCOP and CATH superfamilies, respectively, as a basis for building HMMs.

IUBMB Enzyme Nomenclature and ExplorEnz

The classification of enzymes according to the rules of enzyme nomenclature is the responsibility of the Enzyme Commission of the IUBMB (International Union for Biochemistry and Molecular Biology). The outcome of the decisions made by the commission is deposited in the Enzyme Nomenclature List which is made accessible by several websites (IUBMB website, ExplorEnz, SIB-ENZYME, IntEnz, BRENDA, etc.). They provide forms for searching the enzyme's accepted name, the systematic name, some synonyms, the reaction, cofactors, and literature references.

The Enzyme Commission is the curator of the ExplorEnz database. Main topics are classification and nomenclature. In a concise way it contains the basic data for all classified enzymes. Changes to the enzyme list, e.g., corrections in names, references or reactions are displayed on a separate website. The compilation of new enzyme classes issued by the NC-IUBMB is followed by a period of public review. Enzymes undergoing this process are displayed on the ExplorEnz websites, where scientists can add their comments or request changes. The contents of the ExplorEnz database are also displayed, often together with reaction diagrams, on the Enzyme Nomenclature pages of the IUBMB (http://www.chem.qmul.ac.uk/iubmb/enzyme/). In addition, this site gives detailed information on the rules for naming enzymes and on the nomenclature of biochemical molecules.

IntEnz

IntEnz also contains enzyme data that are curated and approved by the Nomenclature Committee. Enzyme names and reactions are taken from the enzyme list of the NC-IUBMB (see above), sequence data from UniProt.

BRENDA

BRENDA holds by far the largest amount of enzyme- and reaction-related biochemical data. Not counting sequence data, about 3 million data manually extracted from more than 100,000 literature references and organized in about 50 information fields can be found, covering: Classification and nomenclature, reaction and specificity, enzyme ligand data, functional parameters, organism related information, enzyme structure, isolation and preparation, application and engineering, enzyme-disease relationships, literature references.

The database has a highly sophisticated and flexible query engine. The large number of entries is displayed in Table 10.3-1. All data in BRENDA are linked to the original reference papers.

The section on 'Classification and Nomenclature' is based on the enzyme names as defined by the NC-IUBMB and is supplemented with all synonyms found in the literature references which have been manually annotated so far. 'Reaction and Specificity' covers the complete range of natural and artificial substrates accepted by a particular enzyme. Many enzymes have a wider substrate specificity and accept different substrates. Additional sections provide lists of inhibitors, cofactors, metal ions, and activating compounds. A huge number of kinetic parameters is stored (Figs. 10.3-3 and 10.3-4). Since in biological sciences trivial names are very often used instead of IUPAC (International Union of Pure and Applied Chemistry) nomenclature, many compounds are known with a variety of names.

Table 10.3-1. Overview on Enzyme Data in BRENDA as of end of 2010

		Values	References
General	EC Class	5127	
	Proteins	62,720	
	References	1,03,140	
	Synonyms	1,04,954	
	Organisms	10,498	
	Ligands	1,46,842	
Reaction & Specificity	Reactions	1,60,490	2,19,300
	In vivo reactions	41,437	67,469
	Inhibitors	1,58,618	1,80,321
	Cofactors	19,492	34,512
	Metals & Ions	29,011	41,057
	Activating Compounds	23,284	28,717
Functional Parameters/ Kinetics	K_M Values	1,04,850	1,11,191
	Turnover Numbers	41,642	42,781
	Specific Activity	40,494	43,864
	K_I Value	27,493	28,113
	IC50 Values	25,028	25,214
	pH Optimum	31,832	36,892
	Temperature Optimum	17,845	22,299
Application/ Variants	Application	8620	12,576
	Mutants	40,200	44,868
Occurrence	Source Tissue	72,238	1,02,408
	Localization	24,971	38,566
Enzyme structure	Molecular Weight	29,978	33,574
	Posttranslational Modification	4813	7112
	Quaternary Structure	24,830	33,579
	PI Valus	3539	3666
Stability	different stability values	40,524	42,692
Isolations/ preparation		52,469	89,167

Thus even simple molecules may have a dozen or more names. Brenda is equipped with a thesaurus for ligand names based on the INChI (IUPAC International Chemical Identifier) codes for 138,000 different compound names amounting to ca. 46,000 different chemical entities.

Enzyme-catalyzed reactions and compounds interacting with the enzyme protein (cofactors, inhibitors, activating compounds etc.) can be viewed as graphical representations. A tool for substructure searches can be used for drawing a molecule and searching this or its more complex derivatives in the database (Fig. 10.3-5). The molecular structures are also stored as 'Molfiles' enabling a wide range of bioinformatic and cheminformatic usages.

All data are stored in a relational database system. The user can choose from nine search modes:

- 'Quick Search' can be used for a direct search in one of the 54 data fields providing a fast and direct access, e.g., via enzyme names or metabolites.
- 'Fulltext Search' performs a search in all sections of the database, including commentaries.
- 'Advanced Search' allows a combinatorial search for text or numerical data fields.
- 'Substructure Search' is a tool for drawing a molecule which then is searched in the database. The results are exact matches or any molecule containing the plotted structure.
- 'TaxTree Explorer' enables a search for enzymes or organisms in the taxonomic tree.
- 'EC Explorer' can be used to browse or search the hierarchical tree of enzymes.
- 'Sequence Search' is useful for enzymes with a known protein sequence.
- 'Genome Explorer' connects enzymes to genome sequences. The location of classified enzymes is displayed in their genomic context.
- 'Ontology Explorer' allows a simultaneous search in all biochemically relevant ontologies, among them BTO (BrendaTissueOntology).

Functional data are often context-dependent. Since every laboratory carries out experiments on enzyme characterizations under individually defined conditions, and since the results depend on the given experimental know-how, methods and technical equipment available, raw data for the same enzyme are not comparable. In order to account for these differences, BRENDA very often includes the experimental conditions together with the data.

All data in BRENDA are connected to the biological source of the enzyme, i.e. the organism, the tissue, the subcellular localization and to the protein sequence (if available). For the organisms in BRENDA, the taxonomy-lineage is given if the respective organism can be found in the NCBI taxonomy database. Using the TaxTree search mode the user can search for enzymes along the taxonomic tree. The BRENDA tissues are grouped into a hierarchical tissue ontology (Brenda Tissue Ontology, BTO) which was developed by the BRENDA team. It is available from OBO and is meanwhile used by a large number of different groups.

The large number of registered values in BRENDA allows a detailed statistical analysis of enzyme functional data. Two examples are given in Figures 10.3-3 and 10.3-4.

AMENDA/FRENDA

AMENDA (Automatic Mining of ENzyme DAta) and FRENDA (Full Reference ENzyme DAta) are supplements to BRENDA. AMENDA contains a large amount of enzyme data which are automatically extracted from ca.18 million PubMed abstracts using modern optimized text-mining procedures. FRENDA aims at providing an exhaustive collection of literature references containing organism-specific enzyme information. The use of these databases is restricted to the academic community.

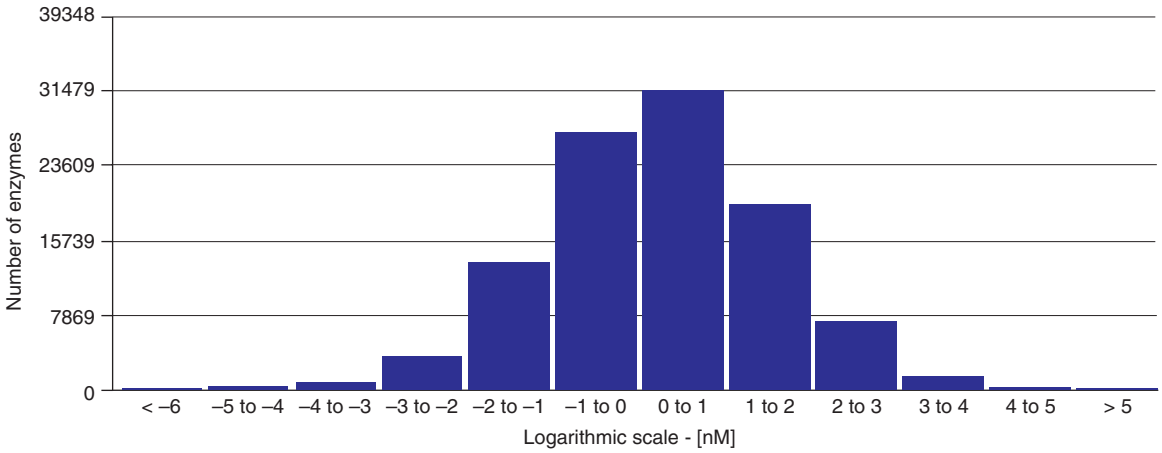


Figure 10.3-3. Distribution of K_M-Values in the BRENDA Database, Showing a Maximum Between 1 and 10 mM

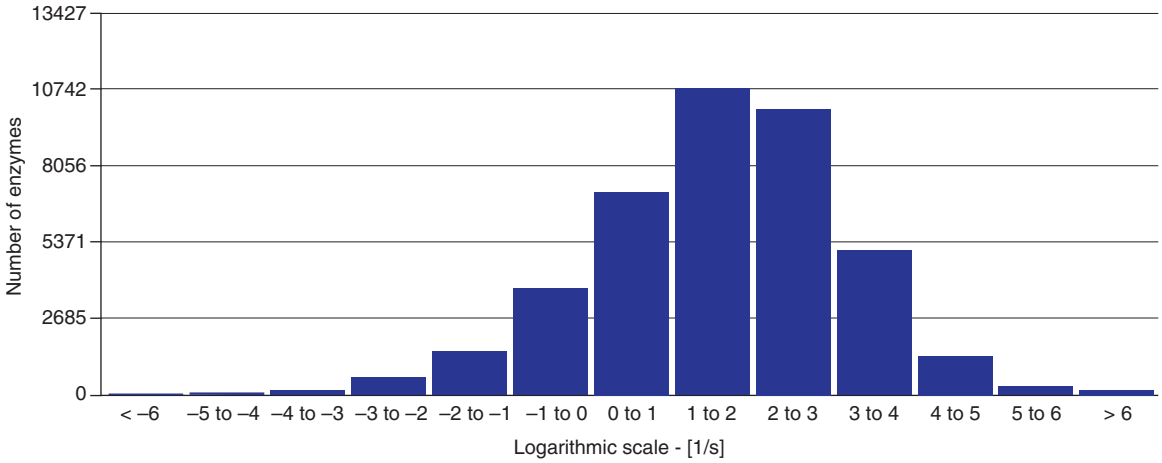


Figure 10.3-4. Distribution of Turnover Numbers in the BRENDA Database, Showing a Maximum Between 10 and 1000 s^{-1}

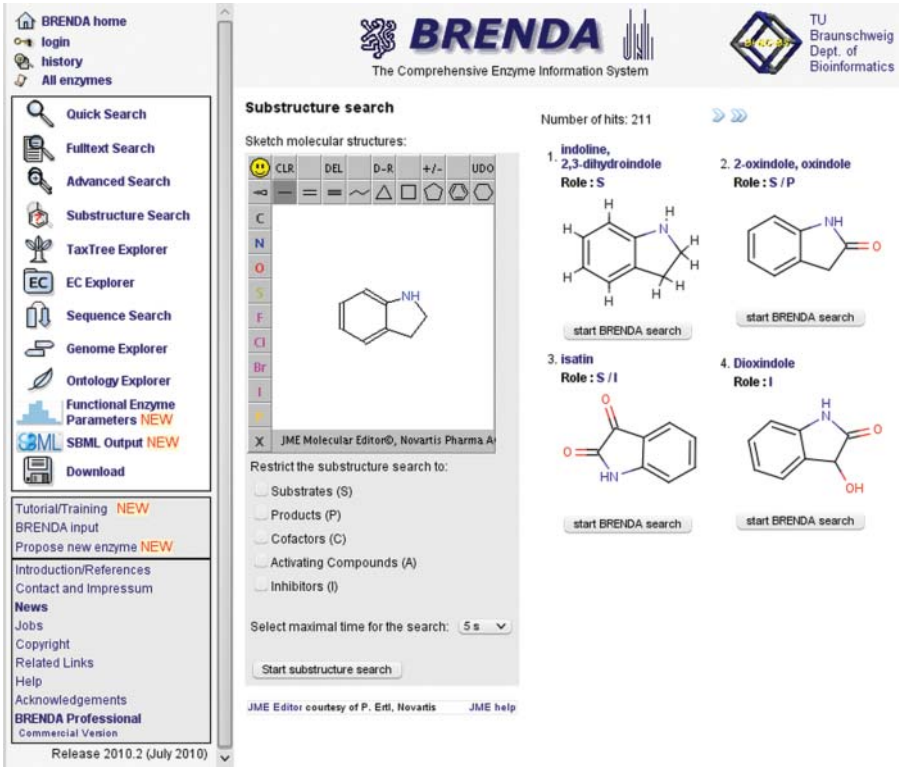


Figure 10.3-5. Example for BRENDA Substructure Search

MEROPS
The MEROPS database is a manually curated information resource for peptidases, their inhibitors and substrates. Peptidases and protein inhibitors are arranged in the database according to a hierarchical classification. The classification is based on sequence comparisons of the domains known to be important for activity. A protein that has been sequenced and characterized biochemically is chosen as a representative ('holotype'). All sequences that represent species variants of the holotype are grouped into a 'protein species'. The sequences of statistically-significant related protein species are grouped into a 'family'. Families that are believed to have had a common ancestor (either because the tertiary structures of the proteins are similar or - in the case of peptidases - active site residues are in the same order in the sequence), are grouped into a 'clan'.

For any peptidase with more than ten known cleavage types a display is presented that gives an indication of the preferred amino acids at its substrate binding sites. Details of the amino acid sequences near the cleavage sites are displayed in the 'Specificity Matrix'. In addition to a graphical display characterizing sequence specificity, a short text description is also given. Artificial or model substrates are summarized in text-sheets, including literature references.

REBASE
REBASE is a comprehensive database of information about restriction enzymes, DNA methyltransferases and related proteins involved in the biological process of restriction and modification. It contains fully referenced information about recognition and cleavage sites, isoschizomers, neoschizomers, commercial availability, methylation sensitivity, crystal and sequence data. Experimentally characterized homing endonucleases are also included. All newly sequenced genomes are analyzed for the presence of putative restriction systems and these data are included in REBASE.

CAZy (Carbohydrate-Active Enzymes)
The CAZy database describes the families of structurally-related catalytic and carbohydrate-binding modules (or functional domains) of enzymes that degrade, modify, or create glycosidic bonds. The NC-IUBMB Enzyme nomenclature of glycoside hydrolases is based on their substrate specificity and occasionally on their molecular mechanism. Such a classification does not (and was not intended to) reflect the structural features of these enzymes. A classification of glycoside hydrolases in families based on amino acid sequence similarities was proposed a few years ago.

TRANSFAC
TRANSFAC is a database of eukaryotic *cis*-acting regulatory DNA elements and *trans*-acting factors. It covers the whole range from yeast to humans. It contains transcription factors and their genomic binding sites, consensus binding sequences and binding profiles. TRANSFAC is a commercial database sold by the Biobase company. Older versions are publically available. The current public version (containing entries of the 2005 commercial version) comprises the following tables:

SITE	7,915 entries
GENE	2,397 entries (1,504 entries with SITE links)
FACTOR	6,133 entries
CELL	1,307 entries
CLASS	50 entries
MATRIX	398 entries.

SITE gives information on (regulatory) transcription factor binding sites within eukaryotic genes. GENE gives a short explanation of the gene that a site (or group of sites) belongs to. FACTOR describes the proteins binding to these sites. CELL gives brief information about the cellular source of proteins that has been shown to interact with the sites. CLASS contains some background information about the transcription factor classes, while the MATRIX table gives nucleotide distribution matrices for the binding sites of transcription factors.

Transporter classification database
The database contains a comprehensive IUBMB approved classification system for membrane transport proteins known as the Transporter Classification (TC) system. The TC system is analogous to the Enzyme Commission's EC system for classification of enzymes, except that it incorporates both functional and phylogenetic information. Descriptions, TC numbers, and examples of over 600 families of transport proteins are provided. Transport systems are classified on the basis of five criteria, and each of these criteria corresponds to one of the five numbers or letters within the TC number for a particular type of transporter.

Thus a TC number normally has five components as follows: V.W.X.Y.Z. V (a number) corresponds to the transporter class (i.e., channel, carrier/porter, primary active transporter or group translocator); W (a letter) corresponds to the transporter subclass which in the case of primary active transporters refers to the energy source used to drive transport; X (a number) corresponds to the transporter family (sometimes actually a superfamily); Y (a number) corresponds to the subfamily in which a transporter is found, and Z corresponds to the substrate or range of substrates transported. The classes are listed in 6.1.1.2.

PDB (Brookhaven National Laboratory, USA)
The Protein Data Bank contains 3D coordinates of macromolecular structures, crystallographic and NMR structure models. In the spring of 2010, 59,000 protein and 2,100 nucleotide 3D structures were deposited.

The Worldwide Protein Data Bank organization was formed to ensure that the PDB archive is and will be freely and publicly available to the global community. The wwPDB members [RCSB, PDB, Protein Data Bank in Europe (PDBe), Protein Data Bank Japan (PDBj), and the BioMagResBank (BMRB)] host deposition, processing, and distribution centers for PDB data and collaborate on a variety of projects and outreach efforts.

Table 10.3-2. PDB entries as of end of 2010

Exp.Method	Proteins	Nucleic Acids	Protein/NA Complexes	Other	Total
X-ray	53,916	1212	2530	17	57,675
NMR	7,409	917	159	7	8,492
Electron Microscopy	202	17	77	0	296
Hybrid	21	1	1	1	24
Other	125	4	4	13	146
Total	61,673	2151	2771	38	66,633

CATH
The CATH database is a hierarchical domain classification of protein structures, present in PDB. The hierarchy levels are: class, architecture, topology (fold family) and homologous superfamily. In 2010 128,000 domains exist, organized in 10,000 sequence families, 2600 superfamilies and 1288 folds. All classification is performed on individual protein domains. Multidomain protein structures are split into their constituent domains.
CATH Hierarchy and Classification:

- Class (C-level) is determined according to the secondary structure composition and packing within the structure. Three major classes are recognized; mainly-alpha, mainly-beta and alpha-beta. This last class (alpha-beta) includes both alternating alpha/beta structures and alpha+beta structures. A fourth class is also identified which contains protein domains which have low secondary structure content.
- Architecture (A-level): This describes the overall shape of the domain structure as determined by the orientations of the secondary structures but ignores the connectivity between the secondary structures. It is currently assigned manually by using a simple description of the secondary structure arrangement, e.g., 'barrel' or '3-layer sandwich'.

- Topology (Fold family, T-level): Structures are grouped according to whether they share the same topology or fold in the core of the domain, that is, if they share the same overall shape and connectivity of the secondary structures in the domain core. Domains in the same fold group may have different structural decorations to the common core.
- Homologous Superfamily (H-level): This level groups together protein domains which are thought to share a common ancestor and can therefore be described as homologous. Similarities are identified either by high sequence identity or structure comparison.

SCOP

A manually curated structural classification of protein domains. In 2010 there are 38,221 PDB entries. 110,800 domains are organized in 3900 families, 1900 superfamilies and 1200 folds. The SCOP database aims at providing a detailed and comprehensive description of the structural and evolutionary relationships between all proteins whose structures are known, including all entries in the Protein Data Bank (PDB). It is available as a set of tightly linked hypertext documents which make the large database comprehensible and accessible. The SCOP classification of proteins has been constructed manually by visual inspection and comparison of structures, but with the assistance of tools to make the task manageable and help provide generality.

Classification: Proteins are classified to reflect both structural and evolutionary relatedness. The principal levels are family, superfamily and fold, described below. The different major levels in the hierarchy are:

- Family (Clear evolutionarily relationship): Proteins clustered together into families are clearly evolutionarily related. Generally, this means that pairwise residue identities between the proteins are 30% and greater.
- Superfamily (Probable common evolutionary origin): Proteins that have low sequence identities, but whose structural and functional features suggest that a common evolutionary origin is probable, are placed together in superfamilies.
- Fold (Major structural similarity): Proteins are defined as having a common fold if they have the same major secondary structures in the same arrangement and with the same topological connections. Different proteins with the same fold often have peripheral elements of secondary structure and turn regions that differ in size and conformation. In some cases, these differing peripheral regions may comprise half of the structure. Proteins placed together in the same fold category may not have a common evolutionary origin. The structural similarities could arise just from the physics and chemistry of proteins favoring certain packing arrangements and chain topologies.

MACiE

MACiE, which stands for 'Mechanism, Annotation and Classification in Enzymes', contains 260 fully annotated enzyme reaction mechanisms, which comprise 249 EC numbers (166 EC sub-subclasses) and 334 distinct CATH codes.

All entries in MACiE are chosen such that: (a) there is a 3-dimensional crystal structure of the enzyme deposited in the Protein Databank (wwPDB), (b) there is a relatively well understood mechanism known, or at least a chemically meaningful suggestion of a mechanism, which is taken from a variety of sources, including chemical studies, quantum mechanical calculations and structure studies. The mechanism included should always be supported by experimental evidence. It is ideal when there are both computational studies and experimental studies available.

KEGG

KEGG stores more than 8000 enzyme-catalyzed and spontaneous reactions, organized and visualized in about 350 pathway maps. Based on bioinformatics-based genome annotations, the enzymes and reactions are assigned to the more than 1200 genomes so far sequenced. The enzyme nomenclature and classification used is based on the IUBMB EC-number system.

The database is mainly used for the construction of metabolic networks.

KEGG-ENZYME is also derived from the IUBMB Enzyme Nomenclature, but the other data sets like COMPOUND, DRUG, GLYCAN, REACTION, RPAIR are developed and maintained by the Kanehisa Laboratories in the Bioinformatics Center of Kyoto University and the Human Genome Center of the University of Tokyo. In addition to the nomenclature data, the enzyme data comprise substrates, products, reactions, gene names and links to chemical structures of metabolites, reaction diagrams and to metabolic pathways.

Table 10.3-3. Overview on Pathway Data in KEGG as of end of 2010

KEGG PATHWAY	Pathway maps	366
KEGG BRITE	Functional hierarchies	89
KEGG MODULE	Pathway modules	540
KEGG DISEASE	Human diseases	374
KEGG DRUG	Drugs	9,163
KEGG EDRUG	Crude drugs and other natural products	753
KEGG ORTHOLOGY	KEGG Orthology (KO) groups	13,718
KEGG GENOME	Organisms	1,377
KEGG GENES	Genes in high-quality genomes (132 eukarya, 1058 bacteria, 85 archaea)	5,669,906
KEGG DGENES	Genes in draft genomes (16 eukarya)	294,638
KEGG COMPOUND	Metabolites and other small molecules	16,253
KEGG GLYCAN	Glycans	10,969
KEGG REACTION	Biochemical reactions	8,183
KEGG RPAIR	Reactant pair chemical transformations	12,268
KEGG RCLASS	Reaction class	2,304
KEGG ENZYME	Enzyme nomenclature	5,191

MetaCyc, EcoCyc, BioCyc, etc.

MetaCyc (http://metacyc.org/) is a non-redundant reference database of small-molecule metabolism that contains experimentally verified metabolic pathway and enzyme information obtained from the scientific literature. They are displayed within the various pathways or with a graphic reaction diagram and links to the connected pathways.

Presently more than 1900 different organisms are represented. More than 1530 metabolic pathways are stored, with more than 8600 enzymatic reactions and more than 24,000 associated literature citations. MetaCyc stores all enzyme-catalyzed reactions that have been assigned EC numbers by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB). MetaCyc also stores hundreds of additional enzyme-catalyzed reactions that have not yet been assigned an EC number.

The main difference to KEGG is the fact that the literature references are connected to the data, and the metabolic pathways are organized in a different way as compared to the KEGG maps, thus providing smaller entities.

The database started as a reconstruction of the metabolic network of *E. coli*, called EcoCyc. This work is still going on. A special focus is put on the manual construction of the pathways for a small number of organisms. For about 20 organisms (tier 2 in BioCyc) computer-based predictions with tools for gap-filling are frequently done, for almost 500 organisms genome annotation is done by standard procedures.

Reactome

The goal of the Reactome knowledge base is to represent human biological processes, but many of these processes have not been directly studied in humans. Rather, a human event has been inferred from experiments on material from a model organism. In such cases, the model organism reaction is annotated in Reactome, the inferred human reaction is annotated as a separate event, and the inferential

link between the two reactions is explicitly noted. Currently the database holds 1,000 human pathways.

Reactome uses a frame-based knowledge representation. The data model consists of classes (frames) that describe the different concepts (e.g., reaction, simple entity). Knowledge is captured as instances of these classes (e.g., 'glucose transport across the plasma membrane', 'cytosolic ATP'). Classes have attributes (slots) which hold properties of the instances (e.g., the identities of the molecules that participate as inputs and outputs in a reaction).

- PhysicalEntities include individual molecules, multi-molecular complexes and sets of molecules or complexes grouped together on the basis of shared characteristics. Molecules are further classified as genome encoded (DNA, RNA, and proteins) or not (all others). Attributes of a PhysicalEntity instance capture the chemical structure of an entity, including any covalent modifications (in the case of a macromolecule) and its subcellular localization.
- CatalystActivity: PhysicalEntities are paired with molecular functions taken from the Gene Ontology molecular function-controlled vocabulary to describe instances of biological catalysis. An optional ActiveUnit attribute indicates the specific domain of a protein or subunit of a complex that mediates the catalysis. If a PhysicalEntity has multiple catalytic activities, a separate CatalystActivity is created for each. This strategy allows the association of specific activities with specific variant forms of a protein or complex, and also enables easy retrieval of all activities of a protein, or all proteins capable of mediating a specific molecular function.
- Events: The conversion of input entities to output entities in one or more steps - are the building blocks used in Reactome to represent all biological processes. Two subclasses of Event are recognized, ReactionlikeEvent and Pathway. A ReactionlikeEvent is an event that converts inputs into outputs. A Pathway is any grouping of related events. An event may be a member of more than one Pathway.

ChEBI

ChEBI stands for Dictionary of Chemical Compounds of Biological Interest. The ChEBI database is hosted at the European Bioinformatics Institute (EBI). ChEBI includes the following data fields:

- ChEBI Identifer – the unique identifer
- ChEBI Name – the name recommended for use in biological databases, graphical representation of a molecular structure and associated molfile(s)
- IUPAC International Chemical Identifier (InChI) and SMILES strings
- Formula – Molecular formula, Charge, Mass
- ChEBI Ontology, Synonyms – other names together with an indication of their source
- Database Links – manually curated cross-references to other non-proprietary databases
- Registry Number – CAS Registry Number, Beilstein Registry Number, Gmelin Registry Number (if available)
- Citations – Publications which cite the entity along with hyperlinks to their entries.

OMIM-online Mendelian inheritance in man

OMIM is a comprehensive, authoritative, and timely compendium of human genes and genetic phenotypes. The full-text, referenced overviews in OMIM contain information on all known Mendelian disorders and over 12,000 genes. OMIM focuses on the relationship between phenotype and genotype. The entries contain links to other genetics resources.

OMIM is intended for use primarily by physicians and other professionals concerned with genetic disorders, by genetics researchers, and by advanced students in science and medicine. While the OMIM database is open to the public, users seeking information about a personal medical or genetic condition are urged to consult with a qualified physician for diagnosis and for answers to personal questions.

Literature:

Bairoch, A. Nucleic Acids Res. 2000;28:304–305. (ENZYME database)

Barthelmes, J. *et al.* Nucleic Acids Res. 2007;35:D511–514. (BRENDA, AMENDA and FRENDA)

Caspi, R. *et al*. Nucleic Acids Res. 2008;36:D623–D631. (MetaCyc, BioCyc).

Coutinho, P.M., Henrissat, B. In: H.J. Gilbert *et al*. (Eds.) *Recent Advances in Carbohydrate Bioengineering.* Cambridge, UK: Royal Soc. of Chemistry, 1999; p. 3–12. (Carbohydrate-active enzymes)

Degtyarenko, K. *et al*. Nucleic Acids Res. 2008;36:D344–D350. (ChEBI)

Fleischmann, A. *et al*. Nucleic Acids Res. 2004;32:D434–D437. (IntEnz)

Hotelier, T. *et al*. Nucleic Acids Res. 2004;32:D145–147. (ESTHER)

Kanehisa, M. *et al*. Nucleic Acids Res. 2008;36:D480–D484. (KEGG)

Manning, G. e*t al*. Science 2002;298:1912–1934. (Protein Kinase Complement)

Passardi, F. *et al.* Phytochemistry 2007;68:1605–1611. (PeroxiBase)

Rawlings, N.D. *et al.* Nucleic Acids Res. 2008;36:D320–D325. (MEROPS)

Roberts, R.J. *et al*. Nucleic Acids Res. 2007;35:D269–70. (REBASE)

Schomburg, D., Schomburg, I. *Springer Handbook of Enzymes*, 2nd Ed. Heidelberg: Springer, Vols.1–39, S1–S7, 2001–2010.

Index

For a general survey refer to pages 1–2

D

E

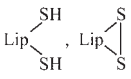
G

H

N

U

Common Abbreviations
(Other abbreviations are defined in the text)

aa	Amino acid
Acc, $AccH_2$	Acceptor, reduced acceptor (unspecified)
ACP	Acyl carrier protein
ATP, ADP, AMP, A	Adenosine tri-, di-, monophosphate, adenosine
bp, kbp	base pair (in DNA), kilobase pairs
cAMP	Cyclic AMP = adenosine 3′,5′-monophosphate
cGMP	Cyclic GMP = guanosine 3′,5′-monophosphate
CoA-SH, CoA-S-	Coenzyme A
CTP, CDP, CMP, C	Cytidine tri-, di-, monophosphate, cytidine
Cyt	Cytochrome
Da, kDa	Dalton, kilodalton (unit of molecular mass)
dATP, dADP, dAMP, dA	Deoxyadenosine tri-, di-, monophosphate, deoxyadenosine
dCTP, dCDP, dCMP, dC	Deoxycytidine tri-, di-, monophosphate, deoxycytidine
dGTP, dGDP, dGMP, dG	Deoxyguanosine tri-, di-, monophosphate, deoxyguanosine
dTTP, dTDP, dTMP, dT	Deoxythymidine tri-, di-, monophosphate, deoxythymidine
DNA	Deoxyribonucleic acid
E	Enzyme
EC number	Enzyme classification according to the IUBMB EC classification
ER	Endoplasmatic reticulum
ETF	Electron transferring flavoprotein
F_{430}	A corrinoid coenzyme (Ni)
FAD, $FADH_2$	Flavin-adenine dinucleotide, reduced flavin-adenine dinucleotide
Fd	Ferredoxin
FMN, $FMNH_2$	Flavin mononucleotide, reduced flavin mononucleotide
Fp	Flavoprotein
ΔG	Change of free energy (see 1.5.1)
G6P	Glucose 6-phosphate
GSH, GSSG	Glutathione, oxidized glutathione
GTP, GDP, GMP, G	Guanosine tri-, di-, monophosphate, guanosine
Ig	Immunoglobulin
ITP, IDP, IMP, I	Inosine tri-, di-, monophosphate, inosine
k	Velocity constant of a reaction (1.5.4)
K	Equilibrium constant of a reaction (see 1.5.1)
K_S, K_I, K_D	Dissociation constants (see 1.5.4, 7.1.2)
K_M	Michaelis constant (see 1.5.4)
kb	Kilobases (10^3 bases)
λ	Wavelength of light
	α-Lipoic acid, oxidized α-lipoic acid
NAD^+, NADH + H^+	Nicotinamide-adenine dinucleotide, reduced nicotinamide-adenine dinucleotide
$NADP^+$, NADPH + H^+	Nicotinamide-adenine dinucleotide phosphate, reduced nicotinamide-adenine dinucleotide phosphate
nt	Nucleotide
NTP, NDP, NMP, N	Any nucleotide tri-, di-, monophosphate or nucleoside
PAP	Adenosine 3,5-diphosphate
PAPS	3-Phosphoadenylylsulfate
PEP	Phosphoenolpyruvate
P_i, PP_i	Inorganic phosphate, inorganic pyrophosphate
pH	Negative decadic logarithm of the H^+ concentration
pK	Negative decadic logarithm of a dissociation constant
PQQ	Pyrroloquinoline quinone
PRPP	α-D-5-Phosphoribosylpyrophosphate
PyrP	Pyridoxal phosphate
RNA	Ribonucleic acid
mRNA, rRNA, tRNA	Messenger-, ribosomal-, transfer ribonucleic acid
R-S-S-R	Disulfide group of amino acids or peptides
S	Svedberg units (sedimentation coefficient)
SAH	S-Adenosylhomocysteine
SAM	S-Adenosylmethionine
THF	5,6,7,8-Tetrahydrofolate
THMPT	5,6,7,8-Tetrahydromethanopterin
ThPP	Thiamin pyrophosphate
UDPG	Uridine diphosphate glucose
UQ, UQH_2	Ubiquinone, reduced ubiquinone
UTP, UDP, UMP, U	Uridine tri-, di-, monophosphate, uridine

Abbreviations for amino acids are listed in Figure 1.3.2, abbreviations for sugars in Figure 4.4.1-1.

Organization of the Chapters

This book was organized in a decimal classification system, which is also used for the index and for the numerous cross-references.

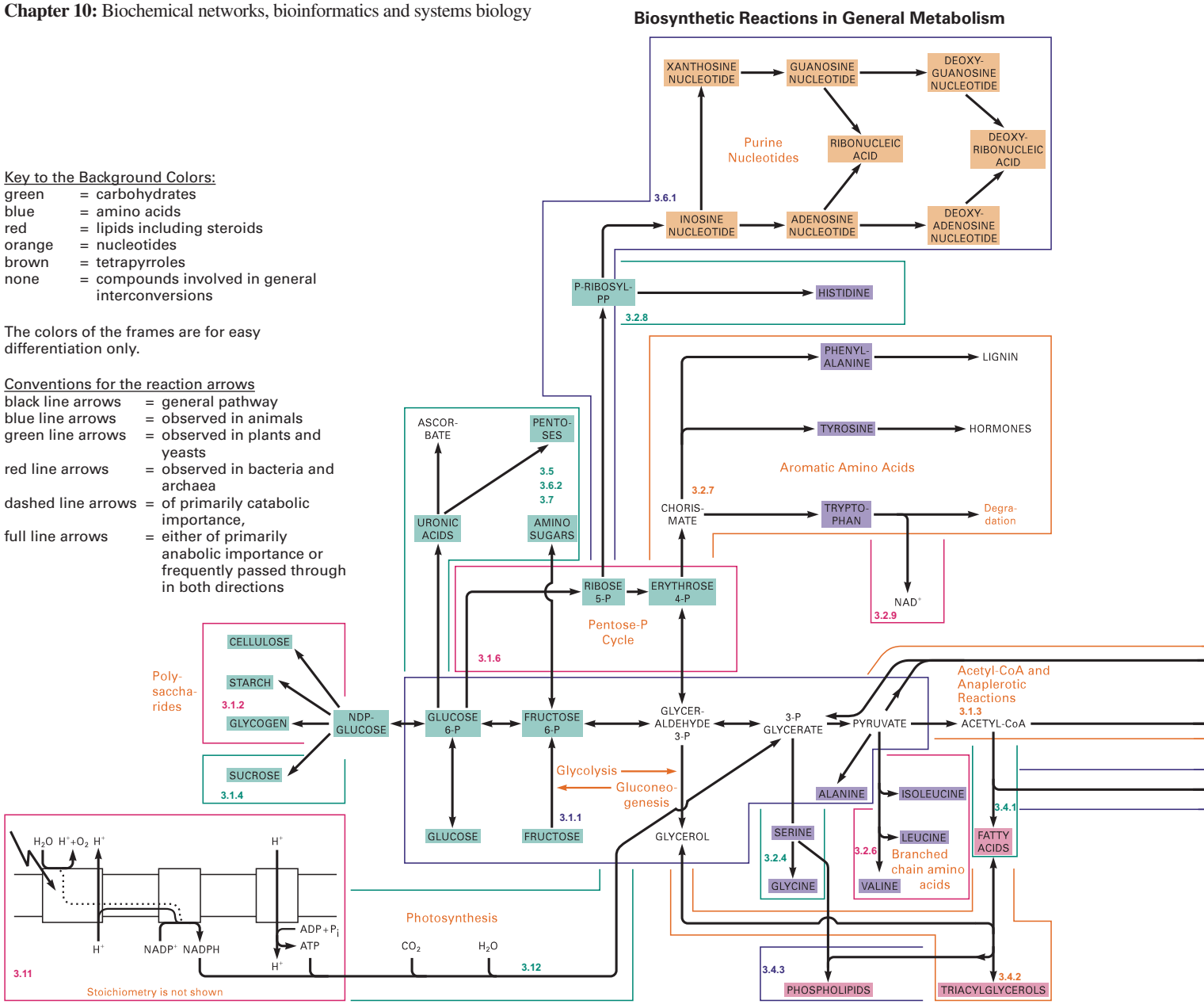
The Figures on this and on the facing page present the majority of the pathways in this book in a schematized way. The chapter and section numbers, which are shown in the drawings, can be used for quick location of the reactions.

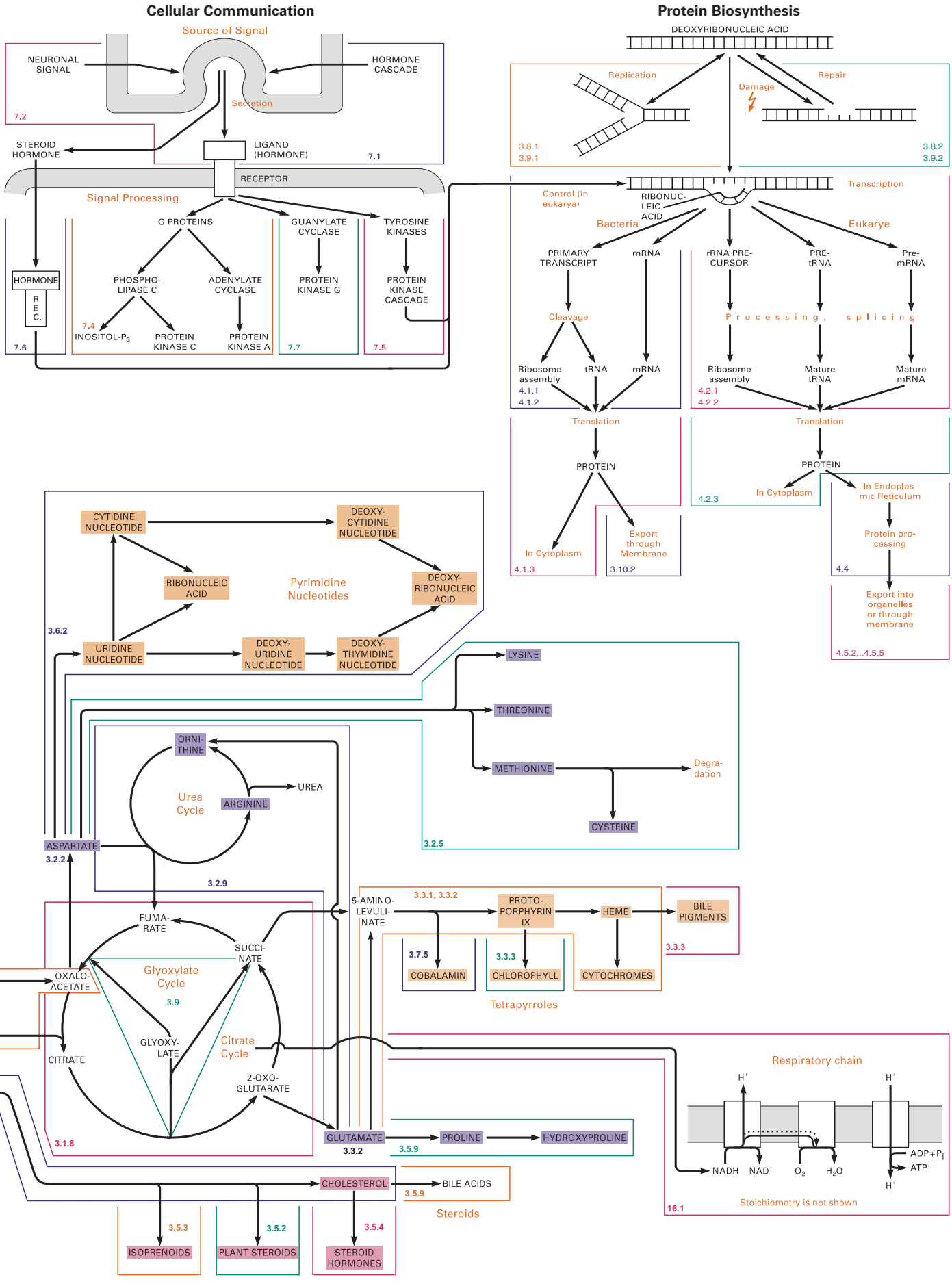
The general setup is as follows:
Chapter 1: Introduction and general aspects, chemistry and physical chemistry
Chapter 2: The cell and its contents: Enzymes, nucleic acids, polymeric carbohydrates and lipids
Chapter 3: General metabolism in animals, plants and bacteria
Chapter 4:.Protein biosynthesis, modifications and degradation
Chapter 5: Viruses
Chapter 6: Transport systems
Chapter 7: Signal transduction and cellular communication
Chapter 8: Immune system
Chapter 9: Blood coagulation and fibrinolysis
Chapter 10: Biochemical networks, bioinformatics and systems biology

Key to the Background Colors:
green = carbohydrates
blue = amino acids
red = lipids including steroids
orange = nucleotides
brown = tetrapyrroles
none = compounds involved in general interconversions

The colors of the frames are for easy differentiation only.

Conventions for the reaction arrows
black line arrows = general pathway
blue line arrows = observed in animals
green line arrows = observed in plants and yeasts
red line arrows = observed in bacteria and archaea
dashed line arrows = of primarily catabolic importance,
full line arrows = either of primarily anabolic importance or frequently passed through in both directions




Cellular Communication
Source of Signal
NEURONAL SIGNAL
HORMONE CASCADE
Secretion
7.2
STEROID HORMONE
LIGAND (HORMONE)
7.1
RECEPTOR
Signal Processing
G PROTEINS
GUANYLATE CYCLASE
TYROSINE KINASES
HORMONE
REC.
PHOSPHO-LIPASE C
ADENYLATE CYCLASE
PROTEIN KINASE G
PROTEIN KINASE CASCADE
7.4
INOSITOL-P3
PROTEIN KINASE C
PROTEIN KINASE A
7.6
7.7
7.5
Protein Biosynthesis
DEOXYRIBONUCLEIC ACID
Replication
Damage
Repair
3.8.1
3.9.1
3.8.2
3.9.2
Control (in eukarya)
RIBONUC-LEIC ACID
Transcription
Bacteria
Eukarye
PRIMARY TRANSCRIPT
mRNA
rRNA PRE-CURSOR
PRE-tRNA
Pre-mRNA
Cleavage
Processing, splicing
Ribosome assembly
tRNA
mRNA
Ribosome assembly
Mature tRNA
Mature mRNA
4.1.1
4.1.2
4.2.1
4.2.2
Translation
Translation
PROTEIN
PROTEIN
In Cytoplasm
In Endoplas-mic Reticulum
4.2.3
In Cytoplasm
Export through Membrane
4.1.3
3.10.2
Protein pro-cessing
4.4
Export into organelles or through membrane
4.5.2...4.5.5
CYTIDINE NUCLEOTIDE
DEOXY-CYTIDINE NUCLEOTIDE
RIBONUCLEIC ACID
Pyrimidine Nucleotides
DEOXY-RIBONUCLEIC ACID
3.6.2
URIDINE NUCLEOTIDE
DEOXY-URIDINE NUCLEOTIDE
DEOXY-THYMIDINE NUCLEOTIDE
LYSINE
THREONINE
METHIONINE
Degra-dation
CYSTEINE
ORNI-THINE
Urea Cycle
UREA
ARGININE
ASPARTATE
3.2.2
3.2.5
3.2.9
5-AMINO-LEVULI-NATE
3.3.1, 3.3.2
PROTO-PORPHYRIN IX
HEME
BILE PIGMENTS
3.3.3
3.7.5
COBALAMIN
3.3.3
CHLOROPHYLL
CYTOCHROMES
Tetrapyrroles
FUMA-RATE
SUCCI-NATE
OXALO-ACETATE
Glyoxylate Cycle
3.9
GLYOXY-LATE
Citrate Cycle
CITRATE
2-OXO-GLUTARATE
3.1.8
GLUTAMATE
3.3.2
PROLINE
3.5.9
HYDROXYPROLINE
Respiratory chain
H+
H+
NADH
NAD+
O2
H2O
ADP+Pi
ATP
H+
Stoichiometry is not shown
16.1
CHOLESTEROL
BILE ACIDS
3.5.9
Steroids
3.5.3
ISOPRENOIDS
3.5.2
PLANT STEROIDS
3.5.4
STEROID HORMONES