The Molecular Biology of Plastids

Cell Culture and Somatic Cell Genetics of Plants, Volume 7A

Edited by

LAWRENCE BOGORAD
Department of Cellular and Developmental Biology
The Biological Laboratories
Harvard University
Cambridge, Massachusetts

INDRA K. VASIL
Laboratory of Plant Cell and Molecular Biology
Department of Vegetable Crops
University of Florida
Gainesville, Florida

ACADEMIC PRESS, INC.
Harcourt Brace Jovanovich, Publishers
San Diego New York Boston London Sydney Tokyo Toronto

Front cover illustration: An adaptation of a map of genes on the chloroplast chromosome of rice. Courtesy of Masahiro Sugiura.

This book is printed on acid-free paper. ♾

Academic Press, Inc.
San Diego, California 92101

United Kingdom Edition published by
Academic Press Limited
24–28 Oval Road, London NW1 7DX

Library of Congress Cataloging-in-Publication Data
(Revised for volume 7)

Cell culture and somatic cell genetics of plants.

Vols. 4-5 edited by Friedrich Constabel, Indra K. Vasil; v. 6 edited by Jeff Schell and Indra K. Vasil; v. 7 edited by Lawrence Bogorad and Indra K. Vasil.
Includes bibliographies and indexes.
Contents: v. 1. Laboratory procedures and their applications -- v. 2. Cell growth, nutrition,cytodifferentiation, and cryopreservation -- [etc.] --
v. 7a. Molecular Biology of Plastids.
v. 7b. The Photosynthetic Appartus: Molecular Biology and Operation.
1. Plant cell culture. 2. Plant cytogenetics. 3. Plant molecular biology. I. Vasil, I. K. II. Constabel, F. III. Schell, Jeff.
QK725.C37 1984 581'.07'24 83-21538
ISBN 0-12-715007-2 (alk. paper) (vol. 7A)
ISBN 0-12-715010-2 (alk. paper) (vol. 7B)

PRINTED IN THE UNITED STATES OF AMERICA
91 92 93 94 9 8 7 6 5 4 3 2 1

Contributors

Numbers in parentheses indicate the pages on which the authors' contributions begin.

S. L. BERRY-LOWE (257), Department of Botany, University of Georgia, Athens, Georgia 30602

LAWRENCE BOGORAD (1, 93), Department of Cellular and Developmental Biology, The Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138

JOHN E. BOYNTON (55), Department of Botany, Duke University, Durham, North Carolina 27706

G. DELP (139), Institut fur Biologie-III der Universität Freiburg, Schanzlestrasse 1, D-7800 Freiburg I. Br., Federal Republic of Germany

ROLAND DOUCE (217), Laboratoire de Physiologie Cellulaire Végétale, URA CNRS n° 576, Département de Biologie Moléculaire et Structurale, Centre d'Etudes Nucléaires de Grenoble et, Université Joseph Fourier, F-38041 Grenoble, France

NICHOLAS W. GILLHAM (55), Department of Zoology, Duke University, Durham, North Carolina 27706

MICHAEL W. GRAY (303), Department of Biochemistry, Faculty of Medicine, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4H7

ELIZABETH H. HARRIS (55), Department of Botany, Duke University, Durham, North Carolina 27706

JACQUES JOYARD (217), Laboratoire de Physiologie Cellulaire Végétale, URA CNRS n° 576, Département de Biologie Moléculaire et Structurale, Centre d'Etudes Nucléaires de Grenoble et, Université Joseph Fourier, F-38041 Grenoble, France

H. KÖSSEL (139), Institut fur Biologie III der Universität Freiburg, Schanzlestrasse 1, D-7800 Freiburg I. Br., Federal Republic of Germany

M. KUNTZ (169), Institut de Biologie Moléculaire des Plantes du CNRS, Université Louis Pasteur, Strasbourg cedex, France

L. MARÉCHAL-DROUARD (169), Institut de Biologie Moléculaire des Plantes du CNRS, Université Louis Pasteur, Strasbourg cedex, France
JEFFREY D. PALMER (5), Department of Biology, Indiana University, Bloomington, Indiana 47405
ANASTASIA PROMBONA (191), Max Planck Institute for Molecular Genetics, D-1000 Berlin 33, Federal Republic of Germany
G. W. SCHMIDT (257), Department of Botany, University of Georgia, Athens, Georgia 30602
DIETMAR STAHL (191), Max Planck Institute for Molecular Genetics, D-1000 Berlin 33, Federal Republic of Germany
ALAP R. SUBRAMANIAN (191), Max Planck Institute for Molecular Genetics, D-1000 Berlin 33, Federal Republic of Germany
MASAHIRO SUGIURA (125), Nagoya University, Chikusa, Nagoya 464, Japan
J. H. WEIL (169), Institut de Biologie Moléculaire des Plantes du CNRS, Université Louis Pasteur, Strasbourg cedex, France

General Preface

Recent advances in the techniques and applications of plant cell culture and plant molecular biology have created unprecedented opportunities for the genetic manipulation of plants. The potential impact of these novel and powerful biotechnologies on the genetic improvement of crop plants has generated considerable interest, enthusiasm, and optimism in the scientific community and is in part responsible for the rapidly expanding biotechnology industry.

The anticipated role of biotechnology in agriculture is based not on the actual production of any genetically superior plants, but on elegant demonstrations in model experimental systems that new hybrids, mutants, and genetically engineered plants can be obtained by these methods and the presumption that the same procedures can be adapted successfully for important crop plants. However, serious problems exist in the transfer of this technology to crop species.

Most of the current strategies for the application of biotechnology to crop improvement envisage the regeneration of whole plants from single, genetically altered cells. In many instances this requires that specific agriculturally important genes be identified and characterized, that they be cloned, that their regulatory and functional controls be understood, and that plants be regenerated from single cells in which such gene material has been introduced and integrated in a stable manner.

Knowledge of the structure, function, and regulation of plant genes is scarce, and basic research in this area is still limited. On the other hand, a considerable body of knowledge has accumulated in the last fifty years on the isolation and culture of plant cells and tissues. For example, it is possible to regenerate plants from tissue cultures of many plant species, including several important agricultural crops. These procedures are now widely used in large-scale rapid clonal propagation of plants. Plant cell culture techniques also allow the isolation of mutant

cell lines and plants, the generation of somatic hybrids by protoplast fusion, and the regeneration of genetically engineered plants from single transformed cells.

Many national and international meetings have been the forums for discussion of the application of plant biotechnology to agriculture. Neither the basic techniques nor the biological principles of plant cell culture are generally included in these discussions or their published proceedings. Following the very enthusiastic reception accorded the two volumes entitled "Perspectives in Plant Cell and Tissue Culture" that were published as supplements to the *International Review of Cytology* in 1980, I was approached by Academic Press to consider the feasibility of publishing a treatise on plant culture. Because of the rapidly expanding interest in the subject both in academia and in industry, I was convinced that such a treatise was needed and would be useful. No comprehensive work of this nature is available or has been attempted previously.

The organization of the treatise is based on extensive discussions with colleagues, the advice of a distinguished editorial advisory board, and suggestions provided by anonymous reviewers to Academic Press. However, the responsibility for the final choice of subject matter included in the different volumes, and of inviting authors for various chapters, is mine. The basic premise on which this treatise is based is that knowledge of the principles of plant cell culture is critical to their potential use in biotechnology. Accordingly, descriptions and discussion of all aspects of modern plant cell culture techniques and research are included in the treatise. The first volume describes every major laboratory procedure used in plant cell culture and somatic cell genetics research, including many variations of a single procedure adapted for important crop plants. The second and third volumes are devoted to the nutrition and growth of plant cell cultures and to the important subject of generating and recovering variability from cell cultures. An entirely new approach is used in the treatment of this subject by including not only spontaneous variability arising during culture, but also variability created by protoplast fusion, genetic transformation, etc. Future volumes are envisioned to cover most other relevant and current areas of research in plant cell culture and its uses in biotechnology.

In addition to the very comprehensive treatment of the subject, the uniqueness of these volumes lies in the fact that all the chapters are prepared by distinguished scientists who have played a major role in the development and/or uses of specific laboratory procedures and in key fundamental as well as applied studies of plant cell and tissue culture.

This allows a deep insight, as well as a broad perspective, based on personal experience. The volumes are designed as key reference works to provide extensive as well as intensive information on all aspects of plant cell and tissue culture not only to those newly entering the field but also to experienced researchers.

Indra K. Vasil

Preface

Volumes 7A and 7B of this series are about plastids—the DNA-containing organelles that set plants apart from other organisms. Volume 7A deals with various aspects of plastid nucleic acid and protein metabolism. The molecular biology and operation of the photosynthetic apparatus are the subjects of the chapters in Volume 7B.

Except for some information on chloroplast gene transmission discussed in Chapter 3 and the composition of chloroplast ribosomes, most of the information in Volume 7A has been acquired during the last 15 years. The first maps of restriction endonuclease sites on chloroplast chromosomes were published in the mid- to late-1970s and now (see Chapter 2) such maps have been made for more than 1000 species. Comparisons of the organization of plastid chromosomes have yielded important information on phylogeny and brought increased understanding of the evolution of these chromosomes. Chapters 4–7 describe our knowledge of the apparatus for nucleic acid and protein metabolism, including our still rudimentary ideas of how some transcripts of chloroplast genes are processed. In a sense, the material in Chapters 1–7 is concerned largely with the molecular biology of what plastids do for themselves. Chapters 9 and 10 describe the organization and operation of the outer plastid membranes and the transport of polypeptides through them. Finally, in a circular way, the discussion of the origin and evolution of the chloroplast genome in Chapter 11 is related to material in Chapters 2 and 3 as well as to other parts of Volume 7A. The acquisition of almost all the information described in Volume 7A has depended on the development of biochemical methods as well as on the tools of molecular biology, which include restriction mapping, cloning, and DNA sequencing.

The first six chapters of Volume 7B contain the authors' views of our current knowledge of the composition, operation, and molecular biology of the apparatus for oxygenic photosynthesis in higher plants and

cyanobacteria and for anoxygenic photosynthesis in other bacteria. Chapters 7, 8, and 9 deal with the composition and synthesis of phycobiliproteins, chlorophylls, and carotenoids—the photosynthetic pigments. Because of space limitations, many interesting and important aspects of plastid biology have been omitted from these two volumes. Most conspicuously absent are extensive discussions of carbon metabolism and its regulation. The only exception in this subject area is Chapter 10, which deals with the molecular biology of ribulose bisphosphate carboxylase–oxygenase. This topic is included mainly because the large subunit of this enzyme is encoded in the chloroplast genome, both its nuclear and plastid genes have been studied in depth, and the problem of assembly is interesting. Beginning in Chapter 11 and continuing through Chapters 12 and 13, a few aspects of the development of plastids and of the photosynthetic apparatus in cyanobacteria are examined. Information on the development of light energy transducing systems of bacteria that carry on anoxygenic photosynthesis is also included in Chapter 3.

One of the most arresting and fascinating features of eukaryotic cell biology is the dispersal in the nuclear and plastid genomes of genes for plastid components. Plastids encode genes for only a modest fraction of the total number of proteins they contain; genes for the remaining proteins are encoded in the nucleus (and perhaps mitochondria?). It is apparent in almost all cases studied to date that multimeric complexes in plastids are made up of products of both plastid and nuclear genes. Possible mechanisms for integrating the expression of plastid and nuclear genomes for plastid components is the subject of Chapter 14; the exposition in this chapter depends on information described and examined in both volumes.

Chapter 15 of Volume 7B also relates to the contents of both Volumes 7A and 7B. Some protein and DNA sequences are presented in various chapters of these two volumes. Authors were asked to include such sequences only when the information had a direct bearing on points being discussed. Chapter 16 is provided for the many readers of these volumes who may wish to study DNA and protein sequences mentioned in the text. At the time of its preparation late in 1990, Chapter 16 was a comprehensive list of sequences of chloroplast genes that had been deposited in computer data bases. It will, no doubt, be out of date by the time these volumes are published, but nevertheless should provide entry into this mass of information for the interested reader.

DNA sequences are available for the total chloroplast genomes of *Marchantia polymorpha*, *Nicotiana tabacum*, and *Oryza sativa*, as well as about 70% of the chloroplast chromosomes of *Zea mays*. However, the number

of proteins encoded by these genomes is not known with certainty, and the functions of proteins that could be encoded in some recognized open reading frames in these genomes are not known. Although, in general, plastid chromosomes are highly conserved with respect to gene content and the arrangements of blocks of genes are constant over large taxonomic distances, there are a few genes that are found in some plastid chromosomes and not in others (Volume 7A, Chapter 2).

This has been an era of rapid acquisition of sequence data. The era has hardly come to an end, but the ability to transform cyanobacteria (Volume 7B, Chapter 2) as well as the plastid genomes of *Chlamydomonas* and higher plants (Volume 7A, Chapters 1 and 4) has marked the beginning of a new era in which information gained from sequencing can be used to study both the molecular biology and the metabolic functioning of plastids as well as the molecular details of photosynthetic mechanisms *in vivo* through genetic manipulation. Many of the chapters in Volume 6 of this series demonstrate how important and valuable it has been to be able to introduce foreign DNA as well as modified normal sequences into plant nuclear genomes.

It is our hope that, like its predecessors, Volumes 7A and 7B will prove to be reliable sources of comprehensive and useful information. We are only beginning to understand the molecular biology of chloroplasts, but we believe that the chapters in these two volumes will serve as milestones on this route.

As usual, in compiling such volumes, we are indebted to the authors who have diligently persevered to present clear statements of sometimes difficult and complicated subjects that are still emerging and are poorly understood. We must recognize our debts to the many scientists who obtained the data dealt with in these chapters. We also wish to thank members of the Editorial Advisory Board for their assistance in the organization of this volume. Finally, we are grateful to the secretarial staffs of the authors and to the technical staff of Academic Press. We would like to particularly recognize Michelle Walker, our editor at Academic Press, for her patience as well as her great contributions to the publication of these volumes.

Lawrence Bogorad
Indra K. Vasil

Cell Culture and Somatic Cell Genetics of Plants
Contents of Previous Volumes

Laboratory Procedures and Their Applications, Volume 1

Cell Growth, Nutrition, Cytodifferentiation and Cryopreservation, Volume 2

Plant Regeneration and Genetic Variability, Volume 3

Part I Regeneration

Cell Culture in Phytochemistry, Volume 4

Part I Introduction

Part II Accumulation of Phytochemicals

Phytochemicals in Plant Cell Culture, Volume 5

Molecular Biology of Plant Nuclear Genes, Volume 6

The Photosynthetic Apparatus: Molecular Biology and Operation, Volume 7B

Part I

Part II

Scale-Up and Automation in Plant Propagation, Volume 8

CHAPTER 1

Introduction

Lawrence Bogorad

Department of Cellular and Developmental Biology
The Biological Laboratories
Harvard University
Cambridge, Massachusetts

In 1682 Nehemiah Grew described green precipitates in leaves. This has been considered the first report of the existence of chloroplasts (Loomis, 1960), but almost another century passed before Joseph Priestley discovered oxygenic photosynthesis in plants in 1771. A few years after that, in 1779, the Dutch physician J. Ingen-Housz found that only the green parts of plants carry on photosynthesis. In 1837 chloroplasts were described in more detail by von Mohl, and in 1883 Arthur Meyer gave the name "grana" to the small black dots he saw in chloroplasts (Loomis, 1960). In simple, elegant experiments, Engelmann showed in 1882 that oxygen-seeking motile bacteria are attracted to the illuminated part of a *Spirogyra* chloroplast *in situ*, and in 1888 Haberlandt showed that oxygen-seeking motile bacteria move to illuminated chloroplasts of *Funaria hygrometrica* Hedw., liberated by cutting a leaf under sucrose on a microscope slide. Thus it was demonstrated that photosynthesis occurs in chloroplasts (Hill, 1975). Anoxygenic photosynthesis by bacteria was recognized much later.

The first indication that plastids have genes came at the beginning of the twentieth-century when Bauer and Correns saw that some plastid traits were not inherited according to Mendel's rules. Because of these and other observations (Rhoades, 1955), botanists began to think not only of the genome, referring then to the collection of nuclear genes, but also of the plastome—implying a collection of plastid genes. DNA was not universally recognized as the preponderant genetic material until the mid-1950s, and unequivocal evidence for the presence of DNA in chloroplasts was not obtained until the 1960s (see Chapters 2A and 3A). Unique chloroplast ribosomes were identified at about the same time, and rRNAs, tRNAs, and ribosomal proteins began to be studied in the late 1960s (see Chapters 6A, 7A, and 8A). Investigations of transcription

by plastid preparations were underway in the mid-1960s, and characterization of chloroplast DNA-dependent RNA polymerase was initiated early in the 1970s (see Chapter 4A). The first endonuclease restriction-site maps of plastid chromosomes were published in the mid-1970s, and molecular hybridization techniques permitted rRNA genes to be located at the same time (see Chapters 2A, 3A, and 6A). The genes *rbcL* and *psbA*, which encode two photosynthetic proteins, were mapped soon afterward, and the first chloroplast protein gene sequence—that for the large subunit of ribulose bisphosphate carboxylase—was published in 1980. Since then the application of techniques of molecular biology has contributed enormously to our understanding of the structure and development of the photosynthetic apparatus in eukaryotes and prokaryotes. Knowledge of the amino acid sequences of plastid proteins gained through sequencing DNA opened the way to relating X-ray diffraction patterns of crystals of ribulose bisphosphate carboxylase–oxygenase and reaction centers of some photosynthetic bacteria to protein structures. In turn, the information on structure has facilitated meaningful site-specific mutagenesis experiments to analyze paths of electron transport, etc. Sequencing of plastid DNA has also revealed the existence of previously unrecognized proteins in the organelle.

These volumes trace the brief history of the molecular biology of plastids and the photosynthetic apparatus against the background of over 200 years of history of photosynthesis research (see reviews by Rabinowitch, 1945; Loomis, 1960; Myers, 1974; Hill, 1975; Arnon, 1977). In concluding a summary of studies of photosynthesis between 1950 and 1975, Arnon (1977) stated that "historians of science may count the last twenty-five years among the golden years of photosynthesis research." The purpose of the present volumes is to examine aspects of photosynthesis and plastid biology about which our knowledge has advanced greatly through research using the tools and points of view of molecular biology. All of this has occurred within the past 15 years or so. Each chapter records and examines one such aspect. The aim in assembling these books has not been to consider every facet of photosynthesis nor to relate all that is known about plastid biology. Rather, the objective has been to present a status report on many features of plastids and photosynthesis that were unknown—indeed unknowable—before the ideas and methods of molecular biology were applied to the analysis of plastid biology.

Volume 7A deals primarily with the molecular biology of plastids and closes with a discussion on the origin and evolution of these organelles. Volume 7B is focused on the molecular biology and operation of the photosynthetic apparatus as well as molecular aspects of plastid differentiation. The final chapter in Volume B enumerates sequenced plastid

genes and their cyanobacterial counterparts; it also provides abbreviated literature citations and information for retrieving the sequences from data banks. This information will be incomplete before the book is published, but it is presented primarily to assist the reader who wishes to consult sequences of genes discussed but not printed here.

Among obvious omissions from this work are detailed discussions of nuclear genes for plastid proteins such as *cab* genes, which encode the light-harvesting chlorophyll *a*/*b*-binding proteins of photosystem II, and *rbcS* genes, which encode the small subunit protein of ribulose bisphosphate–oxygenase. These have been treated in a chapter by Fluhr (1989) in Volume 6 of this series. Another important aspect of photosynthesis that is barely touched on is photosynthetic carbon metabolism; the only segment presented is Chapter 9 in Volume 7B on the enzyme ribulose bisphosphate carboxylase–oxygenase. The very interesting systems for activating enzymes of carbon metabolism and regulating carbon flow are not discussed. Among other subjects that have been omitted are plastid amino acid metabolism, sulphur metabolism, lipid biosynthesis, and nitrate reduction. Another volume on molecular biology of plastids dealing with subjects omitted here will be appropriate in the near future.

The rate of acquisition of knowledge about unexplored aspects of plastid biology will accelerate as a result of the recent successes in genetically transforming and manipulating the chloroplast genome of *Chlamydomonas* (Boynton *et al.*, 1988; Blowers *et al.*, 1989, 1990). The ability to study transient gene expression in tobacco plastids *in situ* (Daniell *et al.*, 1990) and to genetically transform tobacco plastids (Svab *et al.*, 1990) extends the prospect of performing dynamic studies to understand development processes of plastids in flowering plants.

REFERENCES

Arnon, D. I. (1977). Photosynthesis 1950–75: Changing concepts and perspectives. *Encycl. Plant Physiol. New Ser.* **5,** 7–56.

Blowers, A. D., Bogorad, L., Shark, K. B., and Sanford, J. C. (1989). Studies on *Chlamydomonas* chloroplast transformation: Foreign DNA can be stably maintained in the chromosome. *Plant Cell* **1,** 123–132.

Blowers, A. D., Ellmore, G. S., Klein, U., and Bogorad, L. (1990). Transcriptional analysis of endogenous and foreign genes in chloroplast transformants in *Chlamydomonas*. *Plant Cell* **2,** 1059–1070.

Boynton, J. E., Gillham, N. W., Harris, E. H., Hosier, J. P., Johnson, A. M., Jones, A. R., Randolph-Anderson, B. L., Robertson, D., Klein T. M., Shark, K. B., and Sanford, J. C. (1988). Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. *Science* **240,** 1534–1538.

Daniell, H., Vivekananda, J., Nielsen, B. L., Ye, G. N., Tewari, K. K., and Sanford, J. C. (1990). Transient foreign gene expression in chloroplasts of cultured tobacco cells after biolistic delivery of chloroplast vectors. *Proc. Natl. Acad. Sci. USA* **87,** 88–92.

Fluhr, R. (1989). Structure and regulation of light-inducible genes: Genes involved in photosynthesis. *In* "Cell Culture and Somatic Genetics of Plants." Vol 6. "Molecular Biology of Plant Nuclear Genes" (J. Schell and I. K. Vasil, eds.), pp. 134–153. Academic Press, San Diego.

Hill, R. (1975). Days at visual spectroscopy. *Annu. Rev. Plant Physiol.* **26,** 1–11.

Loomis, W. E. (1960). Historical introduction. *Encycl. Plant Physiol.* **5,** 85–114.

Myers, J. (1974). Conceptual developments in photosynthesis, 1924–1974. *Plant Physiol.* **54,** 420–426.

Rabinowitch, E. I. (1945). "*Photosynetheis,* " Vol. 1, 599 pp. Interscience, New York.

Rhoades, M. M. (1955). Interaction of genic and non-genic hereditary units and the physiology of non-genic inheritance. *Encycl. Plant Physiol.* **1,** 19–57.

Svab, Z., Hajdukiewicz, P., and Maliga, P. (1990). Stable transformation of plastids in higher plants. *Proc. Natl. Acad. Sci. USA* **87,** 8526–8530.

CHAPTER 2

Plastid Chromosomes: Structure and Evolution

Jeffrey D. Palmer

Department of Biology
Indiana University
Bloomington, Indiana

I. INTRODUCTION

The three decades since the discovery of chloroplast DNA (cpDNA; reviewed in Kirk, 1986) have been marked by an ever increasing depth and breadth in our understanding of the structure and organization of the molecule. Much of the effort of the few groups working during the 1960s was aimed simply at correctly purifying the "real" cpDNA (Kirk, 1986). The 1970s was a period of discovery of the basic physicochemical properties of the molecule, such as its size and conformation (Kolodner and Tewari, 1975, 1979), and of construction of rudimentary restriction-site and gene maps (e.g., Bedbrook and Bogorad, 1976). The 1980s saw an

explosion of data resulting largely from two parallel and complementary developments: (1) the intensive sequencing and characterization of 10 chloroplast genomes (those of six angiosperms, a bryophyte, and three algae), culminating in the complete sequencing of three land-plant genomes (Shinozaki *et al.*, 1986; Ohyama *et al.*, 1986; Hiratsuka *et al.*, 1989), and (2) the physical and gene mapping of >1,000 chloroplast genomes to address issues of genome and plant evolution (reviewed in Palmer, 1985a,b; Palmer *et al.*, 1988a; Downie and Palmer, 1991).

This storehouse of information forms the basis for the two primary questions this chapter seeks to address. (1) What is the structure and gene organization of cpDNA, and how does this vary in diverse lineages of land plants and algae? (2) What evolutionary processes and molecular mechanisms are responsible for this diversity? Two phylogenetic perspectives are essential to this discussion. First, chloroplasts are undoubtedly of eubacterial, most likely cyanobacterial, origin (Gray, 1989; Chapter 11A). Therefore, the inclusion of bacterial genomes in the comparisons drawn herein highlights both the primitive bacterial features of cpDNA as well as those derived subsequent to its marriage with the eukaryotic cell. Second, our knowledge of cpDNA is quite lopsided phylogenetically: The cpDNAs of a few land plants (mostly angiosperms) are extremely well characterized and those of >1,000 other land plants moderately so, whereas cpDNAs of algae (particularly the great diversity of rhodophytes and chromophytes) are, with few exceptions, only poorly known. Thus, the present generalization that cpDNA evolution is in all respects conservative is based largely on consideration of land plants and may ultimately prove to be, at least in part, invalid for algal lineages.

II. GENOME SIZE EVOLUTION

Compared with the nuclear genome, which varies in size >1,000-fold among angiosperms alone, cpDNA is quite conserved (Table I). The three characterized plastid genomes of nonphotosynthetic plants, which have lost many normal chloroplast genes (Section II.B.2), are 50–73 kb (Table I). Excluding these, the great majority of cpDNAs—representing all of the major lineages of land plants and algae—are between 120 and 160 kb, with only a few genomes approaching 200 kb (Table I). This represents some 30-fold reduction in size compared with the genomes (Bancroft *et al.*, 1989) of the cyanobacterial progenitors of chloroplasts.

TABLE I

Size and Repeat Content of Chloroplast DNAs

Taxa[a]	Genome Size (kb)[b]	Repeat Size (kb)[c]	Reference
Tracheophyta (*a*/*b*)			
Angiosperms (~1,000 species, ~500 genera, ~100 families)	50–217	0–76	Footnote [d]
2 genera, 1 family	50–71[e]	0–24	dePamphilis and Palmer (1989; 1990)
20 genera, 3 families	120–130	none	Footnote [f]
470 genera, 96 families	130–160	10–30	Footnote [d]
4 genera, 4 families	160–180	25–40	Footnote [g]
2 genera, 1 family	200–220	70–80	Footnote [h]
Gymnosperms			
Ginkgo biloba	158	17	Palmer and Stein (1986)
Conifers (7 species, 6 genera, 2 families)	120–130	none	Strauss *et al.* (1988), Lidholm *et al.* (1988)
Ferns (50 species, 13 genera, 6 families)	140–160	10–24	Footnote [i]
Bryophyta (*a*/*b*)			
Marchantia polymorpha	121	10	Ohyama *et al.* (1986)
Physcomitrella patens	122	9	Calie and Hughes (1987)
Sphaerocarpus donnellii	125	nd	Herrmann *et al.* (1980)
Chlorophyta (*a*/*b*)			
Acetabularia (2 species)	(400)[j]	0–10	Footnote [j]
Batophora oerstedii	(400)	nd	Leible *et al.* (1989)
Bryopsis (3 species)	(150)	nd	Misonou *et al.* (1989)
Chara (3 species)	(160–180)	>6	J. Manhart (unpublished)
Chlamydomonas (4 species)	195–292	20–41	Footnote [k]
Chlorella (2 species)	120–174	0–23	Yamada (1983); Schuster *et al.* (1990)
Codium fragile	89	none	Manhart *et al.* (1989)
Coleochaete orbicularis	(100)	none	J. Manhart (unpublished)
Derbesia marina	(100)	nd	Linne von Berg *et al.* (1982)
Eremosphaera virdis	(150)	nd	Linne von Berg *et al.* (1982)
Nitella translucens	(400)	>8	J. Manhart (unpublished)
Pandorina morum	(150–450)	nd	Moore (1990)
Polytoma obtusum	(200)	nd	Siu *et al.* (1975)
Sirogonium melanosporum	(130)	nd	J. Manhart (unpublished)
Spirogyra maxima	130	none	Manhart *et al.* (1990)
Euglenophyta (*a*/*b*)			
Astasia longa	73[e]	6	Siemeister and Hachtel (1989)
Euglena gracilis	130–152	6	Hallick and Buetow (1989)
Rhodophyta (*a*/PB)			
Gracilaria (3 species)	(175)	nd	Goff and Coleman (1988)
Griffithsia pacifica	178	none	Li and Cattolico (1987)
Lomentaria baileyana	(155)	nd	Goff and Coleman (1988)
Plocamium cartilagineum	(170)	nd	Goff and Coleman (1988)
Polyneura latissima	(115)	nd	Goff and Coleman (1988)
Polysiphonia elongata	(120)	nd	Goff and Coleman (1988)
Porphyra yezoensis	200	nd	Cattolico (1986)
Smithora naiadum	(185)	nd	Goff and Coleman (1988)
Glaucophyta (*a*/PB)			
Cyanophora paradoxa	127–136	10–11	Breiteneder *et al.* (1988)

(continues)

TABLE I
Continued

Taxa[a]	Genome Size (kb)[b]	Repeat Size (kb)[c]	Reference
Phaeophyta (*a*/*c*)			
Dictyota dichotoma	123	5	Kuhsel and Kowallik (1987)
Macrocystis integrifolia	(170)	10	Fain *et al.* (1988)
Pylaiella littoralis	191[l]	5	Loiseaux-de Goër *et al.* (1988)
Sphacelaria sp.	(150)[l]	nd	Dalmon *et al.* (1983)
Xanthophyta (*a*/*c*)			
Botrydium granulatum	(130)	nd	Linne von Berg *et al.* (1982)
Tribonema viride	(140)	nd	Linne von Berg *et al.* (1982)
Vaucheria bursata	124	6	Linne von Berg and Kowallik (1988)
Raphidophyta (*a*/*c*)			
Chattonella japonica	(160)	nd	Cattolico (1986)
Bacillariophyta (*a*/*c*)			
Coscinodiscus granii	118	9	Kowallik (1989)
Cyclotella meneghiniana	128	17	C. Bourne and J. Palmer (unpublished)
Odontella sinensis	118	9	Kowallik (1989)
Chrysophyta (*a*/*c*)			
Ochromonas danica	121	15	Li and Cattolico (1988)
Olisthodiscus luteus	150	22	Reith and Cattolico (1986)
Cryptophyta (*a*/*c*/PB)			
Cryptomonas sp.	118	6	Douglas (1988)
Eustigmatophyta (*a*/*c*/PB)			
Monodus sp.	(140)[l]	nd	Li *et al.* (1991)

[a] Photosynthetic pigment composition (in parentheses): *a*, *b*, and *c* are chlorophylls *a*, *b*, and *c*, respectively; PB denotes phycobilins.

[b] Less precise measurements of genome size (by restriction fragment, electron microscopic, or reassociation kinetic analyses) are indicated in parentheses; more precise measurements (by whole genome restriction mapping or DNA sequencing) are not.

[c] nd, not determined. Only those repeats >1 kb are cited. All repeats are inverted duplications except for tandem repeats in *Euglena* and *Acetabularia* and two-copy repeats of undetermined orientation in *Chara*, *Nitella*, and *Macrocystis*.

[d] See Crouse *et al.* (1985) for a listing of genomes mapped until 1985 and Downie and Palmer (1991) for a listing of genomes mapped since then.

[e] These shrunken genomes are from three nonphotosynthetic plants: the angiosperms *Epifagus virginiana* (71-kb genome; dePamphilis and Palmer, 1989; 1990) and *Conopholis americana* (50 kb; S. Downie, C. dePamphilis, and J. Palmer, unpublished) and the euglenophyte *Astasia longa* (73 kb; Siemeister and Hachtel, 1989).

[f] Four groups of angiosperms are known to lack the large inverted repeat: (1) six tribes of legumes and all of their 18 examined genera (Fabaceae; Palmer *et al.*, 1987b; Lavin *et al.*, 1990); (2) two genera of Geraniaceae (*Erodium* and *Sarcocaulon;* P. Calie and J. Palmer, unpublished); (3) one genus of Scrophulariaceae (*Striga;* S. Downie, C. dePamphilis, and J. Palmer, unpublished); (4) *Conopholis* (see footnote [e]).

[g] Includes *Spirodela oligorhiza* (180-kb genome; Lemnaceae; Van Ee *et al.*, 1982), *Nicotiana acuminata* (168 kb; Solanaceae; Shen *et al.*, 1982; R. Olmstead and J. Palmer, unpublished), *Linum* (160–173 kb; Linaceae; Coates and Cullis, 1987), and *Begonia* sp. (170 kb; Begoniaceae; J. Palmer, unpublished).

[h] *Pelargonium* and *Geranium* (Geraniaceae) have duplications of half their genome (Palmer *et al.*, 1987a; P. Calie and J. Palmer, unpublished).

[i] Herrmann *et al.* (1980); Stein *et al.* (1986); Yatskievych *et al.* (1988); Hasebe and Iwatsuki (1990); D. Stein (unpublished).

[j] The early estimates of exceptionally large chloroplast genomes (up to 2,000 kb) in two species of *Acetabularia* (Padmanabhan and Green, 1978; Tymms and Schweiger, 1985) are thought to be erroneously high due to mixed cultures; the genomes are now estimated to be roughly 400 kb (Leible *et al.*, 1989).

[k] Rochaix (1978); Palmer *et al.* (1985b); Turmel *et al.* (1987); Boynton *et al.* (1991).

[l] Equals the sum of both circular molecules for these bicircular genomes.

The one group that is exceptional in this regard are the green algae, whose genome sizes vary at least fivefold, from 89 kb in *Codium* to 400 kb and perhaps larger (see Table I: footnote j) in a few members of each of three major classes: *Acetabularia* and *Bataphora* (Ulvophyceae), *Nitella* (Charophyceae), and *Pandorina* (Chlorophyceae). Chloroplast DNA varies in size by at least 100 kb within the green algal genus *Chlamydomonas*, while variation of from 150 kb to >400 kb has been noted for sexually incompatible, but morphologically indistinguishable, members of a single "species," *Pandorina morum* (Table I).

In the following four sections, I describe the events responsible for size change in cpDNA and discuss the limited evidence available regarding the underlying mechanisms. In one sense, changes in genome size occur in two ways: by changes in the amount of repeated DNA and by changes in sequence complexity. This latter category is divided into three classes: changes in gene content, changes in intron content, and all other types of deletions and insertions.

A. Repeated Sequences

Most chloroplast genomes have very few repeated sequences; however, those they do have often dominate the physical landscape of the genome and are implicated in a variety of processes affecting the evolution of genome size, gene order, and gene sequence. This section focuses on those large repeats (>1 kb in size; Table I), changes in whose size or copy number significantly affect the size of chloroplast genomes. Shorter repeats implicated in processes of genome rearrangement will be considered in Section III.B.2.

The outstanding architectural feature of most cpDNAs is a large inverted duplication, the so-called "inverted repeat" (Fig. 1; Table I). The two copies of this repeat are always found identical within a chromosome, implying the frequent operation of some sort of gene conversion–copy-correction process. The inverted repeat structure will be considered here, within the context of genome size change, in two ways: its phylogenetic distribution and its size. Its roles in other important evolutionary processes will be considered in Sections III.A (as generating molecular heterogeneity), III.B (as limiting opportunities for genome rearrangement), and IV.B (as retarding rates of nucleotide substitution).

Current observations render tenable two alternative phylogenetic scenarios for the origin and subsequent loss of the inverted repeat. First, the inverted repeat was present in the common ancestor of all chloroplast genomes and subsequently lost on multiple occasions in some

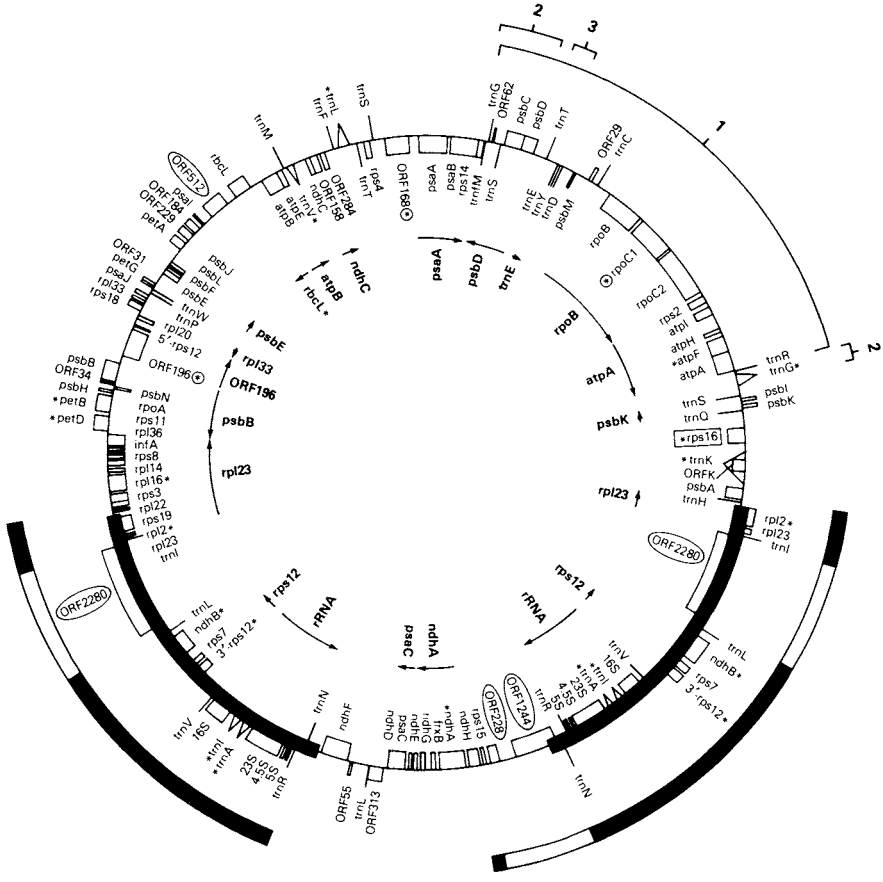


FIG. 1. Gene map of the 156-kb tobacco chloroplast genome, showing selected rearrangements found in rice and *Marchantia*. Genes on the outside of the circle are transcribed counterclockwise; those on the inside are transcribed clockwise. Arrows on the inside of the circle indicate sets of genes thought to constitute operons. In general, the operons are named according to their first gene. The asterisked operon (*rbcL*) consists of a single gene in land plants and chlorophytes but two genes (*rbcL* and *rbcS*) in all other algae (Table IIA). Circled gene names indicate genes present in tobacco and *Marchantia* but absent from rice. The solid boxed name indicates a gene present in tobacco and rice but absent from *Marchantia*. Asterisks denote genes that have the same intron(s) in all three sequenced genomes; circled asterisks denote genes that are split in only two of the three genomes (Table III). The thickened parts of the circle represent the 25.3-kb inverted repeat of tobacco. The thick lines outside the tobacco inverted repeat represent the extent of the inverted repeat in rice (20.8 kb), with regions deleted from rice indicated as open boxes. The numbered brackets outside the circle indicate regions that have been moved by three overlapping inversions in rice relative to tobacco. Gene nomenclature: rRNA genes are indicated by 16S, 23S, 4.5S, and 5S; tRNA genes by *trn*, followed by the one-letter amino acid code; genes for RNA polymerase subunits by *rpo*, followed by a subunit-specific letter;

members of the Tracheophyta, Chlorophyta, Euglenophyta, and Rhodophyta (Table I). Second, the inverted repeat was absent from the ancestral chloroplast genome(s) and instead was gained independently in many different chloroplast lineages, with the only clearly subsequent losses being within land plants. Evidence for the first scenario includes the presence of a complete rRNA operon in all examined inverted repeats and the presence of duplicate, possibly inverted, rRNA operons in the cyanobacterial ancestors of chloroplasts (Bancroft *et al.*, 1989). In all other respects, however, including its overall gene content, size, spacing of the two repeat copies, and relative location and orientation of the rRNA operon relative to the small single-copy region, the inverted repeats of various chloroplast lineages differ substantially. Also noteworthy is the clearly independent and largely specific duplication of rRNA genes in multiple lineages of mitochondria (Palmer, 1985a) and in Euglenophyta chloroplasts (see below). This argues for a strong selective pressure to increase rRNA gene number and, hence, a greater likelihood of parallel duplication of these genes in chloroplasts. Phylogenetic data bearing on these scenarios are quite equivocal, both because it is currently unclear whether chloroplasts themselves originated once or multiple times (Gray, 1989; Chapter 11A) and because inverted repeats, while present in the great majority of cpDNAs, are absent from some or all members of four phyla (Table I).

Land plants allow the only strong conclusions regarding the gain and loss of the inverted repeat. Its near universal presence and strong ho-

---

genes for the 50S and 30S ribosomal proteins by *rpl* and *rps*, respectively, followed by the number of the corresponding *Escherichia coli* protein; genes for components of the thylakoid membrane complexes ATP synthase, photosystem I, photosystem II, and cytochrome b_6/f complex by *atp*, *psa*, *psb*, and *pet*, respectively, followed by a subunit-specific letter; genes for NADH dehydrogenase subunits by *ndh*, followed by the letter of the corresponding mitochondrial subunit; gene for the large subunit of ribulose-1,5-bisphosphate carboxylase by *rbcL*; open reading frames that are conserved between at least two of the three sequenced genomes are indicated by ORF, followed by their lengths in tobacco in codons. General references for the gene maps and structures of the three sequenced genomes are Shinozaki *et al.* (1986), Ohyama *et al.* (1986, 1988), Hiratsuka *et al.* (1989), and Sugiura (1989). The most current source of nomenclature and references for thylakoid membrane genes is Hallick (1989). I have named one other recently identified gene as *ndhH* (Fearnley *et al.*, 1989). References for the various operons are as follows: *rRNA*, *rps12*, *atpB*, and *trnE* (Shinozaki *et al.*, 1986); *ndhA*, (Matsubayashi *et al.*, 1987); *psaC* (Matsubayashi *et al.*, 1987; Schantz and Bogorad, 1988); *rpl23* (Zhou *et al.*, 1989; H. Grüne and P. Westhoff, unpublished); *psbB* (Tanaka *et al.*, 1987; Kohchi *et al.*, 1988a); ORF196 (Kohchi *et al.*, 1988b); *rpl33* (Ohto *et al.*, 1988); *psbE* (Willey and Gray, 1989); *ndhC* (Matsubayashi *et al.*, 1987; Steinmüller *et al.*, 1989) *psaA* (Meng *et al.*, 1988); *psbD* (Yao *et al.*, 1989); *rpoB* (Hudson *et al.*, 1988); *atpA* (Hudson *et al.*, 1987); *psbK* (Murata *et al.*, 1988).

mologies (Fig. 1; Palmer, 1985b) make it clear that the inverted repeat was present in the common ancestor of land plants and that its absence in five groups (Table I) reflects five independent losses. The land-plant lineage also presents the strongest (although still equivocal) case for independent gain of the inverted repeat; thus far the repeat is known to be absent from the few members examined from two of three studied classes of green algae, including the Charyophyceae, from which land plants are thought to derive (Table I).

The inverted repeat of land plants is typically 20–30 kb in size but ranges from 10 to 76 kb (Table I). This accounts for nearly two-thirds of the size variation documented in cpDNA of photosynthetic land plants, whose total range of genome complexities is only 110–150 kb. Most of this inverted repeat size change occurs without change in overall genome complexity, i.e., via spreading of the inverted repeat into previously single-copy regions or via contraction. At one extreme, about 50% of the genome, including many genes present once in all other cpDNAs, lies within the inverted repeat of *Pelargonium hortorum* (Palmer *et al.*, 1987a). A moderate amount of size change also occurs via deletions and insertions internal to the inverted repeat, leading to net changes in genome complexity. These length mutations are usually small, involving segments of no more than 1 kb, although a 6-kb deletion (of the largest gene in the chloroplast genome) has been noted in the inverted repeat of rice (Fig. 1; Hiratsuka *et al.*, 1989) and insertions of 21-kb and 6-kb segments have been reported in the green alga *Chlamydomonas moewusii* (Turmel *et al.*, 1987).

Two types of large tandem repeats have been found in cpDNAs of some algae. Both *Euglena gracilis* and *Astasia longa*, photosynthetic and nonphotosynthetic members, respectively, of the Euglenophyta, have 6-kb tandem repeats containing the entire rRNA operon (Hallick and Buetow, 1989; Siemeister and Hachtel, 1989). The usual form of the genome in both taxa features three entire rRNA repeats and one truncated repeat containing a 16S rRNA pseudogene. Unequal crossing-over between repeats is likely to have given rise to those strains of *E. gracilis* that have one, two, or five complete rRNA operons and one or two partial operons (reviewed in Palmer, 1985a,b; Hallick and Buetow, 1989). An apparently recent case of tandem duplication is found in certain strains of the green alga *Acetabularia mediterranea*, which possess at least five tandem copies of a 10-kb sequence (Tymms and Schweiger, 1985, 1989). This sequence does not encode rRNA; in fact, it is not even present in cpDNA of another species of *Acetabularia* (Tymms and Schweiger, 1985).

B. Gene Content

I. Gene Content Differences

The three completely sequenced cpDNAs, from tobacco, rice, and *Marchantia polymorpha*, each contain about 120 different genes (Fig. 1). The known products of these genes function in two processes: gene expression (4 rRNAs, 30–31 tRNAs, 20 ribosomal proteins, 4 RNA polymerase subunits) and photosynthesis (28 thylakoid proteins, 1 soluble protein). In addition, the land-plant genomes contain about 30 conserved open reading frames (ORFs) of unknown function, some 10 of which are homologous to genes encoding subunits of mitochondrial NADH dehydrogenase.

Comparison of the three sequenced genomes reveals at least 10 differences in gene content (Table IIB; Fig. 1). Several patterns are evident in these differences. (1) None of them involves either known photosynthetic genes or presumptive NADH dehydrogenase genes. (2) Seven of the 10 differences involve ORFs for which no bacterial homologs are known or for which no evidence presently indicates that their products perform the same function in plastids and bacteria (*mbpX*, *mbpY*, *frxC*). (3) The only two protein genes whose functions are reasonably certain and that are lacking in at least one genome are ribosomal protein genes, *rps16* and *rpl21*. (4) The only difference for a gene encoding a stable RNA involves a $tRNA^{Arg}$(CCG) gene that is present in *Marchantia* but absent from the two angiosperms. This difference is easily reconciled with a slightly expanded wobble reading in angiosperms. The few chloroplast genomes (from land plants and *Euglena*) examined in enough detail apparently encode a set of tRNAs sufficient to read all codons (Wakasugi *et al.*, 1986; Ohyama *et al.*, 1988; Hiratsuka *et al.*, 1989; Hallick and Buetow, 1989). Thus, there is no need to invoke import of tRNAs from the cytoplasm, as is known to occur for several mitochondrial lineages, including that of angiosperms (Maréchal-Drouard *et al.*, 1988). (5) The phylogenetic direction of 5 of the 10 differences is reasonably clear. These are the four ORFs that have been lost from the rice genome relative to tobacco and the outlier *Marchantia*, and *frxC*, which has been lost in angiosperms relative to *Marchantia* and the alga *Cyanophora paradoxa* (Table IIB).

A large-scale blot-hybridization survey of angiosperm cpDNAs extends and modifies the above conclusions in two important ways (Downie and Palmer, 1991). First, a clear hierarchy exists with respect to likelihood of gene loss: The rRNA and photosynthetic genes are most

TABLE II
Differences in Gene Content among cpDNAs

A. Genes Present in cpDNA of Some Plants but in Nuclear DNA of Others

Gene[a]	Taxa with Gene in cpDNA	Taxa with Gene in Nuclear DNA	Reference
rpl22	all but legumes	legumes	S. Gantt and J. Palmer (unpublished)
tufA	nongreen algae, most green algae	land plants, some charophytes	Baldauf and Palmer (1990) Baldauf *et al.* (1990)
rbcS	nongreen algae	all chlorophyll *a/b* eukaryotes	Footnote [b]
petF	*Cyanophora*	all chlorophyll *a/b* eukaryotes	Bayer and Schenk (1989), Newman-Spallart *et al.* (1990)
psaF	*Cyanophora*	land plants, *Chlamydomonas*	D. Bryant (unpublished)
atpD	*Cyanophora*	land plants	D. Bryant (unpublished)
atpG	*Cyanophora*	land plants	D. Bryant (unpublished)

B. Genes Present in cpDNA of Some Plants but Known Only as Absences from Others

Gene[a]	Taxa with Gene in cpDNA	Taxa Lacking Gene in cpDNA	Reference
ORF2280	tobacco, *Marchantia*	rice	footnote [c]
ORF1244	tobacco, *Marchantia*	rice	footnote [c]
ORF512	tobacco, *Marchantia*	rice	footnote [c]
ORF228	tobacco, *Marchantia*	rice	footnote [c]
rps16	tobacco, rice	*Marchantia*	footnote [c]
rpl21	*Marchantia*	tobacco, rice	footnote [c]
mbpX	*Marchantia*	tobacco, rice	footnote [c]
mbpY	*Marchantia*	tobacco, rice	footnote [c]
$trnR_{CCG}$	*Marchantia*	tobacco, rice	footnote [c]
frxC	*Marchantia, Cyanophora*	tobacco, rice	footnote [c], D. Bryant (unpublished)
rpl15	*Euglena, Cyanophora*	land plants	Christopher and Hallick (1989) Bryant and Stirewalt (1990)
rps5	*Cyanophora*	land plants	Michalowski *et al.* (1990)
rps10	*Cyanophora*	land plants	D. Bryant (unpublished)
rps17	*Cyanophora*	land plants	Löffelhardt *et al.* (1990)
rpl1	*Cyanophora*	land plants	D. Bryant (unpublished)
rpl3	*Cyanophora*	land plants	Evrard *et al.* (1990)
rpl6	*Cyanophora*	land plants	Bryant and Stirewalt (1990)
rpl18	*Cyanophora*	land plants	Michalowski *et al.* (1990)
rpl19	*Cyanophora*	land plants	D. Bryant (unpublished)
rpl35	*Cyanophora*	land plants	Bryant and Stirewalt (1990)
rvtA	*Scenedesmus*	land plants	Kück (1989)
acpA	*Cyanophora*	land plants	D. Bryant (unpublished)

TABLE II
Continued

B. Genes Present in cpDNA of Some Plants but Known Only as Absences from Others

Gene[a]	Taxa with Gene in cpDNA	Taxa Lacking Gene in cpDNA	Reference
hisH	*Cyanophora*	land plants	D. Bryant (unpublished)
trpD	*Cyanophora*	land plants	D. Bryant (unpublished)
atpI	Everything but *Cyanophora*	*Cyanophora*	D. Bryant (unpublished)
apcA	*Cyanophora*	footnote [d]	Lambert *et al.* (1985)
apcB	*Cyanophora*	footnote [d]	Lambert *et al.* (1985)
apcE	*Cyanophora*	footnote [d]	D. Bryant (unpublished)
cpcA	*Cyanophora*	footnote [d]	Lambert *et al.* (1985)
cpcB	*Cyanophora*	footnote [d]	Lambert *et al.* (1985)
Many	photosynthetic plants	nonphotosynthetic plants	Siemeister and Hachtel (1989) dePamphilis and Palmer (1989; 1990)

[a] Gene nomenclature is described in the legend to Figure 1, footnote d, and in the text. Additional genes listed are *mbpX* and *mbpY* (bacterial membrane permeases), *frxC* (bacterial nitrogenase or photosynthetic protein), *rvtA* (putative reverse transcriptase), *acpA* (acyl carrier protein of fatty acid synthesis), *hisH* (amido transferase of histidine biosynthesis), and *trpD* (anthranilate synthase, component II, of tryptophan biosynthesis).

[b] Starnes *et al.* (1985), Reith and Cattolico (1986), Douglas and Durnford (1989), Hwang and Tabita (1989), Valentin and Zetsche (1989).

[c] From data compiled in Shinozaki *et al.* (1986), Ohyama *et al.* (1986, 1988), Hiratsuka *et al.* (1989), and Sugiura (1989). Not included in this listing are ORFs with no known sequence homology and found in only one of the three sequenced genomes.

[d] The phycobiliprotein genes *apcA, apcB, apcE, cpcA,* and *cpcB* are clearly absent from the three sequenced land-plant cpDNAs and are almost certainly absent from that set of plants and algae (Table I) that simply lacks phycobilisomes altogether.

refractory to loss, with no cases yet apparent; the ribosomal protein genes are next most rarely lost; and the unidentified ORFs are most likely to be lost. Second, gene loss is, in general, quite rare and, from a quantitative standpoint, the situation in rice, where >10% of the coding capacity is absent relative to tobacco, is rather exceptional. Most angiosperm genomes have virtually the same gene content; therefore, gene loss/gain is a minor contributor to overall size change compared with variation in inverted repeat size (Section II.A) and amount of intergenic spacer DNA (Section II.D).

Our view of gene content differences on a deeper evolutionary time scale is biased by the lack of any completely sequenced algal cpDNAs and by the tendency for many of the sequencing efforts in these genomes to have as their starting point genes (usually encoding photosynthetic or ribosomal proteins) known to exist in land-plant cpDNAs. The only two algae for which extensive and fairly random sequencing has

been performed are *Euglena gracilis* and *Cyanophora paradoxa*. Over 100 kb of the 145-kb *Euglena* genome has been sequenced (R. Hallick, personal communication) and the only novel genes found relative to land plants are *tufA* (Montandon and Stutz, 1983; Baldauf and Palmer, 1990) and *rpl5* (Christopher and Hallick, 1989). Given that much of the *Euglena* genome is occupied by introns (see next section) and exceptionally AT-rich DNA of doubtful coding capability (Hallick and Buetow, 1989), it seems likely that it ultimately will be found to contain substantially fewer genes than those of land plants, rather than many new ones (Hallick and Buetow, 1989). This speculation is consistent with the observation that *Euglena* alone among plastid-containing organisms (including both photosynthetic and nonphotosynthetic ones) can survive the apparent loss of its plastid and plastid genome (for a review of this controversial literature, see Hallick and Buetow, 1989). The implication is that the *Euglena* genome lacks genes necessary for metabolic functions other than photosynthesis that might be found in all other plastid genomes (Palmer, 1985a). The recent finding (Siemeister and Hachtel, 1989) that *Astasia longa*, a nonphotosynthetic relative of *Euglena*, has an apparently functional plastid genome of considerably reduced size, and gene content is intriguing in this regard. Perhaps the *Astasia* plastid genome has retained essential nonphotosynthetic genes that *Euglena* has lost.

Southern hybridizations have led to the suggestion that *Chlamydomonas reinhardtii* cpDNA contains several genes involved in DNA metabolism that are lacking from land-plant cpDNA (Oppermann *et al.*, 1989). Experiments with protein synthesis inhibitors suggest that another green alga, *Acetabularia mediterranea*, may encode several chloroplast genes involved in nucleotide metabolism (Schweiger *et al.*, 1986). These reports clearly need validation by gene isolation and sequencing but raise the possibility of considerable divergence of plastid gene content among green algae and their descendants—the land plants.

Cyanophora paradoxa, whose plastid (sometimes termed cyanelle) possesses a rudimentary peptidoglycan cell wall, is the only nonchlorophyll *b*-containing alga whose chloroplast gene content has been examined to any significant extent. Although less than one-third of the 133-kb genome has been sequenced to date (D. Bryant, personal communication), 25 gene content differences relative to land plants are already apparent (Table II). The 24 genes found in *Cyanophora* but not in land-plant cpDNA include 11 translational genes (*tufA* and 10 ribosomal protein genes), 10 photosynthetic genes (five of which encode proteins that are nuclear-encoded in land plants; the other five encode phycobilin proteins, which have no counterpart in the land-plant chloroplast), and 3 genes encoding heretofore unknown classes of products for cpDNA,

namely proteins involved in amino acid (*hisH, trpD*) and fatty acid biosynthesis (*acpA*). Hybridization experiments and sequence analysis strongly suggest that one gene, *atpI*, which is chloroplast-encoded in all other examined cpDNAs, is absent from *Cyanophora* cpDNA (Fig. 2; D. Bryant, unpublished data).

The profound differences in gene content between *Cyanophora* and chlorophytes, as well as the less marked ones among chlorophytes, make clear that overall similarity in genome size can be a very poor indicator of similarity in gene content and of a shared history of gene content reduction. Earlier conclusions drawn by this author (Palmer, 1985a,b) and Cavalier-Smith (1987) to the effect that "most of the transfer of genes from chloroplast to nucleus occurred relatively soon after endosymbiosis . . . and that most plants and algae are now settled into a set compartmentalization of genes between chloroplast and nucleus" (Palmer, 1985b) now seem premature and based on too few data. While it is probably true that the rate of gene transfer/loss has diminished sig-

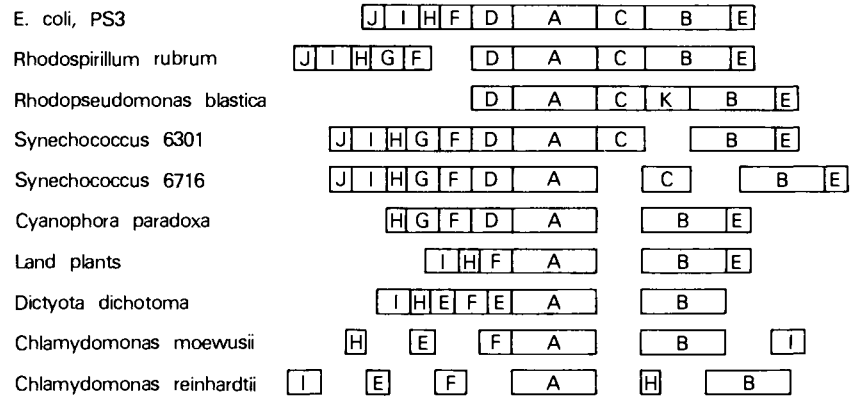


FIG. 2. Organization of ATP synthase genes in chloroplasts and bacteria. Genes drawn without any space between them are found adjacent in the indicated genome and are probably cotranscribed; genes or gene clusters that are shown separated by an open space are physically and transcriptionally unlinked. Genes are drawn proportional to coding length; introns (e.g., in *atpF* in land plants) are not shown. All genes shown have been sequenced except for *atpB* from *Dictyota;* all six genes from *C. moewusii;* and *atpA, atpF, atpH,* and *atpI* from *Chlamydomonas reinhardtii,* all of which have been mapped using heterologous gene probes. *Rhodopseudomonas blastica* has not been examined for several *atp* genes. Data are summarized from Falk and Walker (1988), six bacterial genomes; Breiteneder *et al.* (1988), Lambert *et al.* (1985), and D. Bryant (unpublished), *Cyanophora;* Sugiura (1989), land plants; Hallick and Buetow (1989) and R. Hallick (unpublished), *Euglena;* Kuhsel (1988), *Dictyota;* Turmel *et al.* (1988) and Boynton *et al.* (1991), two *Chlamydomonas* species.

nificantly over the billion years following endosymbiosis, the evidence for continued transfer/loss is mounting rapidly. This view of gene loss as a continuing process that has occurred differentially among plastid lineages encourages the hope that these events will be useful for elucidating some of the major branchings of algal and land-plant evolution. At the same time, it means that one can no longer use the "common" pathway of genome reduction among plastid lineages as an argument in favor of a single primary endosymbiotic origin for plastids.

2. Processes of Genetic Flux

Adherents of the endosymbiotic origin of plastids have often naively assumed that all of the many nuclear genes that encode chloroplast proteins once resided in the plastid and were subsequently transferred from the plastid to the nucleus. Furthermore, it is often assumed that a plant whose chloroplast genome lacks a gene typically found in other cpDNAs will have that gene found in its nucleus instead. In fact, evidence for gene transfer is currently limited to relatively few genes, whereas evidence is mounting that many nuclear genes for chloroplast proteins originated by different pathways. Here, following on the earlier discussion by Bogorad (1975), I outline several ways in which genes for chloroplast proteins might originate, be lost, or be transferred between compartments. I also discuss the relevant evidence, which is essentially of two types: gene distribution patterns and gene sequence phylogenies. It remains to be seen as to which of these processes, or others yet unimagined, will most accurately describe the evolutionary fate of the large number of "missing" genes listed in Table IIB.

a. Gene Transfer Gene transfer, the relocation of an ancestrally bacterial gene from the chloroplast to the nucleus (Bogorad, 1975), is best established for two "primordial" cases of transfer (i.e., occurred soon after symbiosis, before discernable diversification of plastid-containing lineages), two "recent" transfers (i.e., occurred within a monophyletic lineage of plastids), and five ambiguous ones. The two primordial cases involve metabolic enzymes (glyceraldehyde-3-phosphate dehydrogenase [GAPDH] and phosphoglucose isomerase [PGI]), for which phylogenetic analysis establishes that the nuclear gene encoding the chloroplast form of the enzyme is much more closely related to the cognate eubacterial gene than to the nuclear gene encoding the equivalent cytoplasmic isozyme (Martin and Cerff, 1986; Shih *et al.*, 1986; Froman *et al.*, 1989). Weaker evidence, based on overall structural and functional similarities at the enzyme level, supports the notion that several other meta-

bolic enzymes may also represent primordial gene transfers (Weeden, 1983; Knowles *et al.*, 1989). The two clear cases of recent transfer involve the elongation factor Tu gene *tufA*, which was transferred from the chloroplast to the nucleus in the green algal ancestors of land plants (Baldauf and Palmer, 1990), and the ribosomal protein gene *rpl22*, which was transferred to the nucleus during land plant evolution (S. Gantt and J. Palmer, unpublished). The five ambiguous cases all involve photosynthetic genes that are known to be nuclear-encoded in land plants and probably all chlorophytes, but which are chloroplast-encoded in at least one nonchlorophyte, *Cyanophora paradoxa*, and, in the case of *rbcS*, in all five nonchlorophytes examined (Table IIA). These cases will remain ambiguous as to whether transfer was primordial or recent until the question of multiple or single primary origins of chloroplasts is settled, if ever (see Gray, 1989; Chapter 11A).

b. Reverse Gene Transfer Reverse gene transfer, the incorporation of foreign genes into an organellar genome, has been essentially proven for mtDNA (ironically enough, for chloroplast tRNA genes that have invaded plant mtDNA; Joyce and Gray, 1989) but rests on weaker evidence for cpDNA. These potential cases fall into two classes. First, five presumptive genes lie within introns: a reverse transcriptaselike ORF within the *petD* intron of *Scenedesmus obliquus* (Kück, 1989), maturaselike ORFs in the 23S rRNA intron of *Chlamydomonas reinhardtii* (Rochaix *et al.*, 1985) and the *trnK* intron of land plants (Neuhaus and Link, 1987), and ORFs lacking similarity to any characterized genes in two *psbA* introns of *C. moewusii* (Turmel *et al.*, 1989). All five intron ORFs have a restricted phylogenetic distribution within cpDNA suggestive of recent acquisition (as do the introns themselves; see Section II.C). Second, tobacco cpDNA contains several ORFs >100 codons that are absent from cpDNAs of rice and *Marchantia* (Shinozaki *et al.*, 1986; K. Wolfe, unpublished). Assuming that these are functional genes, as opposed to random ORFs occurring simply by chance, these ORFs were either recently acquired by the tobacco genome or lost at least twice independently.

c. Gene Substitution Gene substitution, the loss of a chloroplast gene and the takeover of its function by a primordially nuclear gene (Bogorad, 1975), apparently occurred on several occasions and may turn out to be as important as gene transfer as a mechanism for moving gene function from the plastid to the nucleus. Evidence for this comes from the same phylogenetic logic used to support a cyanobacterial–plastid origin of nuclear genes encoding GAPDH and PGI (see above) but with the reverse conclusion. Here, primary sequence or structure–function

data indicate that a nuclear-encoded chloroplast enzyme is more like a cytoplasmic form than a eubacterial one; examples include glutamine synthetase (Tingey *et al.*, 1988), NADP-dependent malic enzyme (Rothermel and Nelson, 1989), fructose bisphosphate aldolase (Marsh *et al.*, 1989), triosephosphate isomerase (Kurzok and Feierabend, 1984), and, with some ambiguity, phosphoglycerate kinase (Longstaff *et al.*, 1989). The implication is that an ancestrally nuclear gene for each enzyme was duplicated, followed by the adaptation of one member of each gene pair to supplying the chloroplast and then the loss of the corresponding endosymbiont–plastid gene from the cell altogether. Several chloroplast ribosomal proteins encoded by nuclear genes have no bacterial homologs (Gantt, 1988; Zhou and Mache, 1989) and may also have evolved by a substitution pathway.

d. Gene Sharing Gene sharing, a variant form of gene substitution, is now firmly established for yeast mitochondria (Surguchov, 1987) but has not been demonstrated for chloroplasts. In yeast, several examples are known of a single nuclear gene that supplies both cytoplasm and mitochondrion with the same or similar gene product via processes of alternative transcription or splicing.

e. Gene Recruitment Gene recruitment, the evolution of an ancestrally nuclear gene to provide the plastid with a protein of novel function not supplied by the endosymbiont genome, may well have taken place on numerous occasions during the billion years in which the plastid has become genetically and biochemically integrated within the eukaryotic cell. A possible example of this involves the low-molecular weight heat-shock proteins of chloroplasts, which show significant similarity to analogues from the cytoplasm, but for which no corresponding proteins are known from bacteria (Vierling *et al.*, 1989).

f. Gene Loss Gene loss, the disappearance of a gene from the plastid genome without any corresponding functional substitution by some other plastid or nuclear gene, might be expected under two circumstances. First, under normal selective constraints, a plastid gene encoding a protein that plays some nonessential, dispensable role might be lost. This is likely to have occurred in the early stages of endosymbiosis for many bacterial genes that became unnecessary within the context of a photosynthetic eukaryotic cell (e.g., genes for bacterial cell wall synthesis). However, there is no evidence for such loss within modern lineages of photosynthetic plants. Second, a normally essential gene might become dispensable owing to a drastic shift in the plant and plastid's

life-style. The loss of all photosynthetic genes and most other ORFs from the plastid genome of the parasitic, nonphotosynthetic angiosperm *Epifagus virginiana* (dePamphilis and Palmer, 1990) and of several photosynthetic genes from the nonphotosynthetic euglenophyte *Astasia longa* (Siemeister and Hachtel, 1989) clearly falls into this latter category.

3. Mechanisms of Genetic Flux

Gene transfer can be viewed (Bogorad, 1975) as a three-step process involving (1) physical transfer of a duplicate chloroplast gene to the nucleus, (2) functional activation of the transferred gene, and (3) loss of the preexisting chloroplast gene. The middle step is likely to be the most difficult and, therefore, rate-limiting. Evidence that cpDNA sequences are readily transferred to the nucleus comes from the observation that numerous pieces of cpDNA are integrated in nuclear DNA (Timmis and Scott, 1983), many of which appear to result from recent transfer (Cheung and Scott, 1989). How these sequences become transferred, whether by vector-mediated processes involving transposons or viruses or by indirect processes of transformation (i.e., nuclear uptake of DNA from lysed chloroplasts) or even occasional fusion of plastid and nucleus and recombination of their genomes is quite unclear. Whether the intermediate in the physical transfer process is RNA, DNA, or both is also unclear.

Gain of proper function by a transferred gene seems a daunting task, requiring (1) acquisition of sequences, including a promoter, terminator, polyadenylation site, and ribosome-binding site, required for correct expression in the nuclear-cytoplasmic genetic environment; (2) acquisition of regulatory sequences conferring the proper level and timing of expression with respect to tissue type, developmental stage, and environmental influence; (3) acquisition of an N-terminal transit peptide sequence necessary for targeting a cytoplasmically synthesized polypeptide into the chloroplast; (4) removal or alteration of sequences, such as certain types of introns, that might interfere with nuclear-cytoplasmic gene expression; and (5) potential changes in codon usage such that an abundant chloroplast polypeptide would be translated efficiently on cytoplasmic ribosomes. One could imagine these steps occurring slowly and serially, by the gradual recruitment and "improvement" of the necessary expression and transit sequences by a transferred chloroplast gene. Alternatively, a more rapid activation might occur if a chloroplast gene were to become inserted within the N-terminus of a duplicate (hence dispensable) preexisting nuclear gene encoding a chloroplast

polypeptide and then quickly cannibalize its regulatory and transit sequences. While any one step in the activation process may not be difficult, the sum total seems unlikely to be achieved by any but a small fraction of transferred genes. The vast majority seems destined to accumulate mutations that would render them pseudogenes long before they would have any chance of ever becoming functional.

The one step in the activation process for which relevant experimental evidence exists is the gain of transit peptides. Gene fusion experiments have shown that a surprisingly large percentage of short random DNA sequences can serve as transit peptides for import into mitochondria (Lemire *et al.*, 1989, and references therein). For two genes, *rbcS* (Wolter *et al.*, 1988) and *rpl22* (S. Gantt and J. Palmer, unpublished), exon shuffling has been implicated as being involved in the process of transit peptide recruitment.

Once a transferred gene becomes functional in the nucleus, the original chloroplast gene might well be redundant and fall prey to the deletion processes that maintain the compact organization of most plastid genomes. In some cases, as with the *rpl23* and *infA* pseudogenes of certain angiosperms (Zurawski and Clegg, 1987; Wolfe and Sharp, 1988), these deletion events may be fairly subtle and detectable only by careful sequence analysis. The simultaneous existence of gene copies in both plastid and nucleus also allows time for the functional divergence of originally identical genes, in which case both copies might be retained as active genes. The *tufA* gene, present in the chloroplast of green algae but in the nucleus of land plants (Baldauf and Palmer, 1990), might represent just such a case. Certain charophycean green algae, from which land plants are most likely derived, contain in their cpDNAs an intact *tufA* gene that is so exceptionally divergent as to be unlikely to produce a functional elongation factor Tu, while their nuclear genomes contain several uncharacterized *tufA*-like sequences (Baldauf *et al.*, 1990).

Compared with gene transfer, processes of gene substitution and gene recruitment, which require in essence only the elaboration of a transit peptide, and even reverse gene transfer, which at a minimum does not require that particular step, seem simple. Simpler, however, does not always mean easier. Reverse gene transfer may face mechanistic constraints, including (1) more difficult entry of nucleic acids into (intact) chloroplasts compared with their exit (from broken ones), (2) greater likelihood of deletion of foreign sequences in the compact chloroplast genome (see Section II.D), and (3) greater difficulty of recruiting regulatory and control sequences in a genome with a relatively small number of transcriptional units and essentially no separately evolving duplicate genes. Gene substitution, on the face of it a very simple process, may encounter severe functional constraints. Homolo-

gous, primordially chloroplast (bacterial) and nuclear (eukaryotic) genes have been separated for roughly 3 billion yr. Successful functional substitution of a nuclear gene product for a chloroplast one should, in principle, be easier for a monomeric protein than for a polypeptide that interacts with other polypeptides, either within membranes or as part of a multisubunit soluble enzyme. However, even for a monomeric protein, a major barrier to substitution may lie in the adaptation of cytoplasmic and chloroplast proteins to compartmental-specific parameters of pH, ionic concentrations, and metabolite levels. And, of course, gene substitution is not even a possibility for those chloroplast genes, such as photosynthetic ones, for which counterparts in the nucleus simply never existed.

C. Intron Content

Chloroplast genomes of major lineages of eukaryotes contain highly variable numbers of introns (Table III). With the exception of the group III class of 100-bp introns unique to *E. gracilis* (Christopher and Hallick, 1989), most introns are 400–1,000 bp in size. The >100 introns found in *Euglena* (Table III; R. Hallick, unpublished data) amount to >30 kb of the genome (Koller and Delius, 1984), whereas only a single 232-bp intron has been found in >40 genes sequenced to date from *C. paradoxa* (Table III; D. Bryant, unpublished). Thus, variation in intron content accounts for a significant amount of overall variation in genome size among major lineages of algae and between them and land plants (Table III). However, differences in intron content contribute only minimally to genome size variation within land plants (Table III).

No intron is shared by all of the limited number of chloroplast genomes examined thus far, with substantial data on intron content presently available for only four lineages (land plants, *C. reinhardtii*, *E. gracilis*, *C. paradoxa;* Table III). With but one exception (see next paragraph), no introns are shared between even two of these four lineages (e.g., Gingrich and Hallick, 1985), the first three of which have chlorophylls *a* and *b* and are likely to stem from a common primary endosymbiotic event. In all but one case in which introns are found in the same gene in multiple lineages (*rpl16*, *petD*, *rpoC1*, *psbA*, *psaA*, *rrn23;* Table III), the introns are clearly not evolutionarily homologous, being different in location, sequence, and sometimes type. For example, the four introns found in the *C. reinhardtii psbA* gene (Erickson *et al.*, 1984) are of the group I type (Dujon, 1989) and are located in entirely different positions from the four group II (Michel *et al.*, 1989) introns of the *E. gracilis psbA* gene (Karabin *et al.*, 1984).

TABLE III
Distribution of Introns in cpDNA[a]

	Land Plants[b]			Charophytes[c]			*Chlamydomonas*[d]				*Chlorella*[e]					
Gene	T	R	M	C	N	S	R	S	M	E	E	P	So	Eg	Cp	*a/c*
trnL	1	1	1	—	—	—	0	—	—	—	1	—	—	0	1	1
trnI	1	1	1	1	1	0	0	—	—	—	0	—	—	0	0	0
trnA	1	1	1	1	1	1	0	—	—	—	0	—	—	0	0	0
trnV	1	1	1	—	—	—	0	—	—	—	0	—	—	0	—	1
trnG	1	1	1	—	—	—	—	—	—	—	—	—	—	0	—	—
trnK	1	1	1	—	—	—	—	—	—	—	—	—	—	0	—	—
rps12	2[f]	2[f]	2[f]	—	—	—	0	—	—	—	—	—	—	0	—	—
rps16	1	1	x	—	—	—	—	—	—	—	—	—	—	—	—	—
rpl2[g]	1	1	1	—	—	—	0	—	—	—	—	—	—	0	0	—
rpl16[g]	1	1	1	—	—	—	0	—	—	—	—	—	—	3	0	—
atpF	1	1	1	—	—	—	—	—	—	—	—	—	—	3	0	0
petB	1	1	1	—	—	—	—	—	—	—	—	0	—	1	—	0
petD	1	1	1	—	—	—	—	—	—	0	—	0	1	—	—	0
ndhA	1	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—
ndhB	1	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—
rpoC1[g]	1	0	1	—	—	—	—	—	—	—	—	—	—	11	—	—
ORF196[g]	2	0	2	—	—	—	—	—	—	—	—	—	—	—	—	—
ORF168	2	2	1	—	—	—	—	—	—	—	—	—	—	—	0	—
psbA	0	0	0	—	—	—	4	3	2	—	—	—	—	4	0	—
psaA	0	0	0	—	—	—	2[f]	2[f]	2[f]	2[f]	—	—	—	3	—	—
rrn23	0	0	0	—	—	—	1	—	5	6	1	—	—	0	—	—
rrn16	0	0	0	—	—	—	0	—	1	0	—	—	—	0	—	0
rps7	0	0	0	—	—	—	0	—	—	—	—	—	—	0	0	—
rbcL	0	0	0	0	0	0	0	—	—	0	0	—	—	9	—	—
atpI	0	0	0	—	—	—	—	—	—	—	—	—	—	6	—	—
psaB	0	0	0	—	—	—	0	—	—	—	—	—	—	6	0	—
psbB	0	0	0	—	—	—	—	—	—	—	—	—	—	4	0	—
psbC	0	0	0	—	—	—	0	—	—	—	—	—	—	>3	—	—
psbE	0	0	0	—	—	—	—	—	—	—	—	—	—	2	0	—
psbF	0	0	0	—	—	—	—	—	—	—	—	—	—	1	0	—
rps3	0	0	0	—	—	—	0	—	—	—	—	—	—	2	0	—
rps8	0	0	0	—	—	—	—	—	—	—	—	—	—	3	0	—
rps11	0	0	0	—	—	—	0	—	—	—	—	—	—	3	—	—
rps14	0	0	0	—	—	—	—	—	—	—	—	—	—	1	—	—
rps19	0	0	0	—	—	—	—	—	—	—	—	—	—	2	0	—
rpl14	0	0	0	—	—	—	0	—	—	—	—	—	—	1	0	—
rpl23	0	0	0	—	—	—	—	—	—	—	—	—	—	3	—	—
tufA	x	x	x	0	—	x	0	—	—	—	—	—	—	3	0	0
rpoB	0	0	0	—	—	—	0	—	—	—	—	—	—	8	—	—
rpoC2	0	0	0	—	—	—	0	—	—	—	—	—	—	2	—	—
ORF184	0	0	0	—	—	—	—	—	—	—	—	—	—	1	—	—

[a] Numbers indicate numbers of introns found in the indicated gene and taxon; a dash indicates the gene has not been examined; an x means the gene is known to be missing from the indicated genome. Data are from

The one possible case published thus far of much earlier intron gain is that of the *trnL* intron found in land plants and *Cyanophora*. Evidence for homology of the intron in these two groups is strong and includes its identical position in the gene and relatively high degree of sequence similarity (Evrard *et al.*, 1988). This intron is peculiar in being the only group I intron present in the three sequenced land-plant cpDNAs and the only intron of any type known in the >40 sequenced genes from *Cyanophora* cpDNA. Recent studies show that this *trnL* intron is present in most algal chloroplast lineages and, indeed, that it was present in the cyanobacterial ancestors of chloroplasts (Kuhsel *et al.*, 1990; Xu *et al.*, 1990). Moreover, a second chloroplast intron, a group II intron in a *trnV* gene, is also distributed widely in chloroplasts and cyanobacteria (M. Kuhsel and J. Palmer, unpublished). These two tRNA introns are the only introns found thus far in eubacterial genomes.

These observations lead to three conclusions regarding the origin of chloroplast introns. First, most chloroplast introns were probably acquired during chloroplast evolution; only a few introns, perhaps as few as two, were present in the cyanobacterial endosymbiont that gave rise to chloroplasts. Second, most chloroplast genes probably acquired their introns rather late during chloroplast evolution, because these introns are so restricted in phylogenetic distribution. Third, several chloroplast genes (e.g., *psbA*) were probably invaded by introns independently in at least two different lineages. This "introns-late" view of the origin of most chloroplast introns is at odds with the still dominant "introns-early" view propounded by Gilbert (Gilbert *et al.*, 1986) and most textbook writers (e.g., Lewin, 1990) but is in accord with the growing body of evidence indicating that introns have been both acquired and lost

---

several compilations (Ohyama *et al.*, 1988; Dujon, 1989; Hallick and Buetow, 1989; Hiratsuka *et al.*, 1989; Michel *et al.*, 1989; Wolfe, 1989a; Manhart and Palmer, 1990), other reports (Rochaix *et al.*, 1985; Yamada and Shimaji, 1987; Evrard *et al.*, 1988; Kuhsel, 1988; Christopher and Hallick, 1989; Durocher *et al.*, 1989; Keller *et al.*, 1989; Kück, 1989; Liu *et al.*, 1989; Lou *et al.*, 1989; Nickoloff *et al.*, 1989; Reimann and Kück, 1989; Rochaix *et al.*, 1989; Turmel *et al.*, 1989a,b; Bryant and Stirewalt, 1990; Evrard *et al.*, 1990; Kuhsel *et al.*, 1990; Löffelhardt *et al.*, 1990; Michalowski *et al.*, 1990; Robertson *et al.*, 1990), and the unpublished work of H. Bohnert, D. Bryant, S. Fong, R. Hallick, U. Kück, C. Lemieux, and S. Surzycki.

[b] Taxa designations: T, tobacco (*Nicotiana tabacum*); R, rice (*Oryza sativa*); M, *Marchantia polymorpha*.

[c] Taxa designations: C, *Coleochaete orbicularis;* N, *Nitella axillaris;* S, *Spirogyra maxima*.

[d] Taxa designations: R, *C. reinhardtii;* S, *C. smithii;* M, *C. moewusii;* E, *C. eugametos*.

[e] Taxa designations: E, *C. ellipsoidea;* P, *C. protothecoides*.

[f] Fragmented genes known or thought to be *trans*-spliced (*rps12, psaA*) or to produce polypeptide fragments (*rps7*) (see text).

[g] Intron has been lost from this gene in one or more lineages of angiosperms (Zurawski *et al.*, 1984a; Hiratsuka *et al.*, 1989; Downie *et al.*, 1991; S. Downie, J. Logsdon, and J. Palmer, unpublished).

So, *Scenedesmus obliquus;* Eg, *Euglena gracilis;* Cp, *Cyanophora paradoxa;* a/c, phaeophyte (either *Dictyota dichotoma* or *Pylaiella littoralis*).

within recent times in many lineages of organisms and genomes (Rogers, 1989; Perlman and Butow, 1989).

If most of the >100 introns found in chloroplasts are of recent, postendosymbiotic origin (at least with respect to their presence in chloroplast genes) then one should expect the same for introns present in transferred nuclear genes of chloroplast origin. However, claims for common ancestry of introns "across the prokaryote–eukaryote boundary" have rested entirely on the supposed homology (based entirely on intron position) between essentially only two of the introns of the nuclear genes encoding cytosolic and chloroplastic forms of GAPDH (Quigley *et al.*, 1988; Shih *et al.*, 1988). I find these authors' conclusions unconvincing for the following reasons. (1) Essentially, all of the many introns found in other nuclear genes of chloroplast origin appear to be derived. For example, *rbcS* in the nucleus of land plants and green algae contains as many as three introns, whereas cyanobacterial *rbcS* and the chloroplast-localized *rbcS* of all other algal types lack introns entirely (Wolter *et al.*, 1988, and references in Table II). (2) The chance of parallel acquisition of different introns at the same location in a gene seems reasonably high given such factors as the number of different introns already known in GAPDH genes (>30), sequence constraints on intron positioning, and the possibility of intron sliding. (3) Such an introns-early view postulates not only the presence of introns in the common ancestor of eubacteria and eukaryotes, but also their retention throughout the first 2 billion yr of eubacterial diversification, up to the point of chloroplast origin, followed by the massively parallel loss of hundreds, if not thousands, of putative introns in each of the several different eubacterial lineages (Woese, 1987) in existence at that time.

The availability of complete sequences for cpDNAs of two angiosperms and the bryophyte *Marchantia* allows several generalizations about the evolution of intron content in land plants. First, all but 1 of the 20 or so introns found in each genome are shared by at least one angiosperm and *Marchantia* and were, therefore, present in the common ancestor of land plants (Table III). Two of these introns, in *trnI* and *trnA*, were acquired about 450–500 million yr ago in the Charophyceae, the green algal group from which land plants are derived (Manhart and Palmer, 1990; Table III). Second, most of these introns are highly stable components of the land-plant chloroplast genome, being present in all three genomes, with only three cases of intron loss clearly apparent (all of which are precise losses of the entire intron in rice; Hiratsuka *et al.*, 1989). Additional evidence for intron loss comes from hybridization and sequence studies, which reveal that the *rpl2* intron has been lost in parallel in at least six separate lineages of dicots and that the *rpl16* intron has been lost in at least one dicot lineage (Zurawski *et al.*, 1984a; Downie

et al., 1991; S. Downie, J. Logsdon, and J. Palmer, unpublished). Third, there are no known examples of intron gains among the three sequenced genomes (Table III). In summary, the introns found in land plants originated in their green algal progenitors, at least 400 million yr ago, followed by a continuing period of stasis characterized by little or no further gain and only occasional loss. Thus, changes in intron content, like those in gene content, are relatively minor contributors to overall size change during chloroplast genome evolution.

How do chloroplasts gain and lose their introns? The process generally believed to be responsible for the precise loss of introns from genes involves two steps (Dujon, 1989): (1) the reverse transcription of RNAs whose introns have been removed by splicing and (2) homologous recombination between the intronless cDNA and the native gene. Several kinds of evidence suggest that recombination processes are well developed in chloroplasts (Sections III.A, III.B, and IV.B), whereas evidence for reverse transcription has been otherwise lacking. Because land-plant chloroplasts do occasionally lose their introns, but lack any reverse transcriptaselike genes in their three sequenced genomes, one might speculate that they import, at least occasionally, reverse transcriptase from the cytoplasm. On the other hand, given that intron loss is such a rare event even on an evolutionary time scale, one must also consider mechanistically more difficult processes that do not involve reverse transcription. These include both the precise deletion of the intron via normal genomic-based deletion processes (see next section) and gene conversion between an intron-containing gene and its spliced transcript.

The mechanism for intron gain is fairly clear for those group I introns that encode their own site-specific endonucleases that effect the intron's efficient propagation into its cognate sites and, perhaps, its transposition to different sites (Dujon, 1989). Some, but not all, group I introns in *Chlamydomonas* cpDNA appear to fall into this category (Lemieux *et al.*, 1988; Durocher *et al.*, 1989; Boynton *et al.*, 1991), whereas the *trnL* intron of land plants and *Cyanophora* does not. How group II and III introns are spread is less clear; proposed mechanisms include the DNA-level insertion of transposable elements and the DNA- and RNA-level insertion of preexisting group II introns (Wessler, 1989; Rogers, 1989; Mörl and Schmelzer, 1990).

D. Length Mutation

The gain and loss of genes and introns are just two special types of the general class of mutations known as length mutations (insertions and deletions). The great majority of length mutations in cpDNA are much

smaller, only 1–10 bp in size, than these paramount events and occur predominantly in noncoding DNA (including both intergenic spacers and introns themselves). Many of these small-length mutations are flanked by or close to very short direct repeats or occur within homopolymer tracts, suggesting that they take place mainly by slippage and mispairing during DNA replication or repair (Takaiwa and Sugiura, 1982; Zurawski *et al.*, 1984b). Length mutations of 10–1,200 bp in size occur less frequently than the smaller ones and are more likely to occur by recombination than by replication. These recombinational processes include unequal crossing-over between misaligned tandem repeats to produce both deletions and additions (Schlunegger and Stutz, 1984; Blasko *et al.*, 1988; also see Section III.B) and intramolecular recombination between short direct repeats to produce deletions (Ogihara *et al.*, 1988).

The very rare class of length mutations that occurs within coding regions is of two general types. Protein genes generally sustain only very short mutations, often of just 3 bp, that preserve their reading frame. An exception to this rule is the enigmatic ORF2280, the largest gene in most land-plant cpDNAs, yet one whose function is unknown and which can tolerate internal deletions/insertions of as large as 300 bp (Blasko *et al.*, 1988; Zhou *et al.*, 1988). Ribosomal RNA genes tend to sustain a broader range of length mutation, of 1–100 bp in size; these occur mainly in single-stranded loop regions of the rRNA and, thus, do not disrupt its secondary structure (see Chapter 6A).

As noted previously, land-plant cpDNAs vary relatively little in sequence complexity, which ranges from 110 to 150 kb. Thus far, the described nongenic length mutations responsible for this variation are limited in two ways. First, none is larger than ca. 1 kb in size. Second, none is yet identified as an insertion of foreign DNA sequences. Perhaps the best candidates for foreign sequences in cpDNA are its introns (see preceding section); also, subclover cpDNA has been shown to contain several "novel," species-specific sequences of unknown origin (Milligan *et al.*, 1989). This contrasts with the relative abundance of cpDNA sequences in plant mitochondrial and nuclear DNA (Timmis and Scott, 1983; Stern and Palmer, 1984). The compact nature of land-plant cpDNA, together with its relative lack of major size variation and of foreign sequences, implies that the genome is evolving under strong constraint—either mechanistic or selective—to eliminate unnecessary sequences. This constraint is exemplified by nonphotosynthetic flowering plants and euglenophytes, whose plastid genomes have eliminated most or all of their photosynthetic genes (Siemeister and Hachtel, 1989; dePamphilis and Palmer, 1989, 1990).

Green algal cpDNAs show much more variation in total genome size than do land-plant cpDNAs (Table I), appear to have an elevated ratio of 50–1,000-bp long-length mutations to point mutations (Palmer *et al.*, 1985a; Turmel *et al.*, 1987), and have sustained the largest known length mutations (Turmel *et al.*, 1987; Tymms and Schweiger, 1985, 1989). Taken together, these observations suggest that green algal cpDNAs are evolving under more relaxed selective constraints on genome size and/or greater mechanistic pressures to expand and contract.

III. GENOME ORGANIZATION

Three levels of organization of the plastid chromosome can be conveniently recognized: (1) the three-dimensional "nucleoid" arrangement of the multiple genomes present within a plastid (cytological organization); (2) the overall form of plastid chromosomes (molecular conformation); and (3) the fine-scale arrangement and order of genes within a chromosome (gene order). In viewing the entire phylogenetic spectrum of chloroplast-containing eukaryotes, great evolutionary variability is seen in gene order, as compared with only moderate and very little variability in cytological organization and molecular conformation, respectively. The first of these levels of organization is discussed by Gillham *et al.* (Chapter 3A) and will not be dealt with further. The other two levels are discussed in the following two sections.

A. Molecular Conformation

The natural conformation of all characterized chloroplast genomes is circular. The only linear plastid chromosomes described thus far are "unnatural," greatly reduced ones present in albino cereal plants derived from another culture (Day and Ellis, 1985; Ellis and Day, 1986). Evidence gained from both direct (electron microscopic) and indirect (restriction mapping and DNA sequencing of whole genomes) observation of cpDNA shows that the basic structure of almost all chloroplast chromosomes is unicircular; i.e., the entire sequence content of the genome maps as a single circular chromosome. Three examples are now known in which the genome is bicircular, with sequences distributed between two circles of different size. These include two of the three genomes of the Phaeophyta that have been examined in this regard and the one

examined member of the Eustigmatophyta (Table I). In addition to their circular chromosomes, chloroplasts of a few algae contain small circular extrachromosomal plasmids of unknown gene content, some of which also apparently exist as integrants within the chromosome (Ebert *et al.*, 1985, and references therein).

The best characterized bicircular chloroplast genome, that of the brown alga *Pylaiella littoralis*, exists as circles of 133 and 58 kb (Loiseaux-de Goër *et al.*, 1988). The large circle is more AT-rich than the small one and present in about twice the abundance. The two circles are largely dissimilar in sequence content (Loiseaux-de Goër *et al.*, 1988) and clearly do not arise and interconvert by the recombinational process that produces multicircular genomes in angiosperm mitochondria (Palmer, 1990). The 58-kb circle of *P. littoralis* contains several pseudogenes, but as yet no known functional genes have been identified (Loiseaux-de Goër *et al.*, 1988; Markowicz *et al.*, 1988).

A seemingly similar situation to that in *Pylaiella* has been claimed for rice cpDNA, in which it was proposed that individual plants contain a 10:1 ratio of genomes bearing functional and defective copies of *rbcL* and *atpB* (Moon *et al.*, 1987). However, this observation has not been confirmed by other groups studying rice cpDNA, including those who sequenced the entire rice genome (Hiratsuka *et al.*, 1989) and must, therefore, be regarded as unsubstantiated.

While the basic form of the plastid genome is generally unicircular and occasionally bicircular, additional layers of molecular heterogeneity are generated by three types of recombination events, each of which results from or is affected by the presence of specific kinds of repeated elements in the genome. The first type of recombination occurs intramolecularly between the two segments of the large inverted repeats that characterize most cpDNAs (Section II.A; Table I). With but one exception, this recombination results in a 50:50 mixture of two genetically identical but physically distinct inversion isomers that differ only in the relative orientation of their single-copy regions (reviewed in Palmer, 1985a,b). In *Dictyota dichotoma*, one inversion isomer is present in roughly fivefold excess of the other (Kuhsel and Kowallik, 1987; Kuhsel, 1988). The *Dictyota* inverted repeat is also the smallest known (4.7 kb) and if its rRNA operon were to straddle both ends of the repeat then only one genome isomer would contain an entire operon and would likely be selected for. Whether "flip-flop" isomerization occurs by site-specific or general homologous recombination is presently unclear (Palmer *et al.*, 1985a).

A second layer of molecular heterogeneity is generated by intermolecular recombination between cpDNA monomers, which produces circular dimers (Kolodner and Tewari, 1979) and higher-order multimers of

progressively lower abundance (Deng *et al.*, 1989). Genomes lacking a large inverted repeat exist only as head-to-tail circular dimers, whereas genomes having one exist also as head-to-head dimers, implying that recombination occurs between the inverted sequences of two circular monomers (Kolodner and Tewari, 1979). The production of multimers likely occurs by general homologous recombination. The fact that head-to-head dimers are severalfold overrepresented (comprising 70–80% of the dimers in genomes in which the inverted repeat occupies 33% of the chromosome; Kolodner and Tewari, 1979) relative to expectations of random recombination throughout the genome suggests that other factors may also be involved. These might include the inverted repeat's frequent engagement in intramolecular recombination and its possible involvement in genome replication.

The third type of heterogeneity-generating recombination is unequal crossing-over between molecules containing short tandem repeats. Sequence studies have shown that a length-hypervariability evident in single cultures of *Euglena gracilis* results from unequal crossing-over at a 54-bp tandem repeat element (Schlunegger and Stutz, 1984). A similar process has been predicted to account for the size polymorphisms observed in cpDNA preparations from the green alga *Codium fragile* (Manhart *et al.*, 1989).

B. Gene Order

The arrangement of genes on the circular chromosomes of chloroplasts has already been discussed in terms of their presence or absence (Section II.B) and copy number (Section II.A). Here, gene arrangement is considered from the standpoint of the serial order of genes around the chromosome. The diversity of gene orders and patterns of organization found in algae and land plants is described, followed by a discussion of mechanisms that generate gene rearrangement.

1. Patterns of Gene Order and Arrangement

a. Land Plants Several features of the gene arrangement of tobacco cpDNA are evident in Figure 1. First, the large inverted repeat (Section II.A) divides the genome into four segments, the two repeats of 25 kb and the large and small single-copy regions of 87 and 18 kb, respectively. Second, genes are compactly arranged in each segment of the genome, with intergenic spacers generally only a few hundred bp in length. Third, many, although by no means all, of the genes are cotranscribed

as part of operons (Figs. 1 and 2). Many of these operons have been found in identical or similar form in eubacteria, especially cyanobacteria, and, thus, have been retained throughout the ca. 1 billion yr of chloroplast evolution. In some cases, operons have been expanded in a particular chloroplast lineage to include additional genes relative to the operons of cyanobacteria and other chloroplast lineages. The products of some of these additional genes function in completely different enzyme complexes or processes than those of the original genes of the operon. For example, *rps14* is part of the *psaA* operon in land plants but not in *Euglena*, *Cyanophora*, and cyanobacteria (Cantrell and Bryant, 1987; Meng *et al.*, 1988; Cushman *et al.*, 1988; D. Bryant, unpublished). The primary transcripts from these operons also tend to be processed much more extensively in chloroplasts than in such bacteria as *E. coli*, although how chloroplasts compare in this regard to cyanobacteria is poorly understood (see Chapter 5A). Finally, transcription units, be they individual genes or operons, are distributed fairly randomly between strands of the chromosome.

The gene order of tobacco (Fig. 1) is found in most other angiosperms (Palmer, 1985a,b; Palmer *et al.*, 1988a; Downie and Palmer, 1991) and also represents the ancestral gene order of vascular plants (Palmer and Stein, 1986). The completely sequenced genomes of rice (Fig. 1; Hiratsuka *et al.*, 1989) and *Marchantia* (Ohyama *et al.*, 1986, 1988) differ in gene order from that of tobacco by only three and one inversions, despite having been separated from a common ancestor for some 200 and 400 million yr, respectively. In most cases, the variant patterns found in less completely characterized angiosperm genomes also result from one or a few inversions (Palmer, 1985a,b; Downie and Palmer, 1991). In only four groups of land plants—two lineages of legumes (Palmer *et al.*, 1987b, 1988b; Milligan *et al.*, 1989), geranium (Palmer *et al.*, 1987a), and conifers (Strauss *et al.*, 1988)—have more complex cases of rearrangement been described. Most of these appear to involve large numbers of inversions, in some cases accompanied by putative transpositions and insertions of additional "novel" sequences. In three of the four cases (i.e., in legumes and conifers), the gene reshufflings were preceded by the loss of the inverted repeat, whose absence may contribute an element of instability to the genome (Section III.B.2).

With two exceptions, all of the gene rearrangements described in land plants have their boundaries between operons, rather than within them. A 78-kb inversion shared by mung bean and common bean has one endpoint within the *rpl23* operon (between *rps8* and *infA*) (Palmer *et al.*, 1988b). In subclover, the *rpoB* operon is split by a rearrangement—potentially a transposition—that separates *rpoB-rpoC1* from *rpoC2* (Milligan *et al.*, 1989).

b. Green Algae Gene maps of moderate resolution (20–30 protein and rRNA genes; Turmel *et al.*, 1988; Boynton *et al.*, 1991) have been established for two distantly related pairs of closely related, interfertile species of *Chlamydomonas* (Chlorophyceae), while low-resolution maps (13–14 genes; Manhart *et al.*, 1989, 1990) are available for *Spirogyra maxima* (Charophyceae) and *Codium fragile* (Ulvophyceae). Within each pair of interfertile *Chlamydomonas* species, gene order is apparently identical; however, gene order is almost completely different among the two pairs, *Spirogyra* and *Codium*, and land plants. The only similarities involve local conservation of gene order within one or a few sets of two or three genes (*psaAB, psbDC, atpFA, rRNA*) that are known to be cotranscribed in land plants.

Chlamydomonas cpDNAs lack the extensive operon structure of land-plant genomes (Turmel *etal.*, 1988; Boynton *et al.*, 1991). The rRNA operon is virtually the only operon retained, with the dispersion of ancestrally cotranscribed genes best exemplified by the ATP synthase genes (Fig. 2). In all other chloroplast genomes, the *atp* genes bear witness to their bacterial origins, being organized into two operons similar to those of cyanobacteria. In *Chlamydomonas*, however, the six cpDNA-encoded *atp* genes are scattered singly around the genome in a species-specific manner (Fig. 2). Too little is known to say whether operons have been dispensed within other green algal lineages to the extent evident in *Chlamydomonas*. It is notable, however, that the 16S and 23S rRNA genes, which are cotranscribed in virtually all bacteria and other chloroplast genomes, are known to be transcribed separately in *Chlorella ellipsoidea* (Chlorophyceae; Yamada and Shimaji, 1987) and are probably transcribed separately in *Spirogyra* (Manhart *et al.*, 1990) and *Codium* (Manhart *et al.*, 1989).

Gene scrambling is carried to an extreme level in *Chlamydomonas reinhardtii* by the fracturing of two genes into subgenic pieces. The *psaA* gene is split into three widely separated pieces that are transcribed separately and assembled to form an mRNA by *trans*-splicing (Kück *et al.*, 1987). The *rps7* gene appears to be split into two pieces that are thought possibly to give rise to separate polypeptides (Robertson *et al.*, 1990). Neither of these two genes is *trans*-spliced in land plants or other genera of algae, whereas *rps12*, which is uninterrupted in *Chlamydomonas, Euglena*, and *Cyanophora* (Table III), is *trans*-spliced in all examined land plants (Sugiura, 1989).

c. Euglena The *E. gracilis* chloroplast genome has been largely sequenced and, as discussed earlier (Section II.B), likely contains substantially fewer genes than land-plant genomes. Many of the retained genes are grouped into operons of similar organization to those found in land

plants (Figs. 1 and 2; Hallick and Buetow, 1989). No similarities are evident, however, in the global arrangement of operons. There is a pronounced asymmetry of transcription of the *Euglena* genome. Half the circle is transcribed almost entirely off one strand and the other half largely off the other (Koller and Delius, 1984; Hallick and Buetow, 1989). The origin of replication is situated between these two halves in such a manner that each replication fork is oriented in the same direction as transcription (Hallick and Beutow, 1989). The functional significance of this pattern of organization is discussed in Section III.B.2.

d. Cyanophora The only other well-studied chloroplast genome is that of *C. paradoxa*. It contains many of the same operons as those found in land plants and *Euglena* (Figs. 1 and 2) but with a third fundamentally different pattern of global operon arrangement (Lambert *et al.*, 1985; Breiteneder *et al.*, 1988; Löffelhardt *et al.*, 1990; D. Bryant, unpublished). Many of the extra genes present in *Cyanophora* cpDNA relative to land-plant cpDNA (Section II.B; Table II) are internal parts of standard chloroplast operons, with their positions corresponding to those of bacterial operons (e.g., Fig. 2). Similar to land plants, but unlike *Euglena*, transcription switches strands fairly randomly in *Cyanophora*.

e. Chromophytes Low-resolution maps of 13–20 genes have been constructed for six species representing three diverse groups of chromophytes—three diatoms, two brown algae, and a xanthophyte (Kuhsel, 1988; Kowallik, 1989; Loiseaux-de Goër *et al.*, 1988; C. Bourne and J. Palmer, unpublished). These limited data permit two tentative conclusions. First, overall gene order is largely different among the three groups and also different when compared with all other mapped genomes, with only the three diatoms showing limited similarities. Second, as in most other cpDNAs except those of *Chlamydomonas*, several prokaryotic operons appear to be retained (rRNA, *psaA*, *psbD*, *atpA*). The *atpA* operon of the brown alga *Dictyota dichotoma* is notable for containing two highly dissimilar copies of *atpE*, whereas in all other chloroplasts and bacteria a single *atpE* gene is either part of the *atpB* operon or else transcribed singly (Fig. 2).

2. Processes of Gene Rearrangement

The extent of gene order scrambling among examined algal chloroplast genomes is too great to allow inference of the nature of individual rearrangement events. It is only among land plants where gene order differences can, for the most part, be attributed to specific mutations. Of

these, most are inversions, although transposition and even inverted repeat expansion/contraction (Palmer *et al.*, 1988b) have also been invoked as potentially causing known rearrangements. The following discussion will consider these processes in terms of the factors, mechanistic and selective, that are likely either to promote or suppress them.

In general, genetic inversions occur by recombination between inversely oriented repeated sequences. Chloroplasts seem well endowed with active recombination systems (Sections II.A, II.D, III.A, and IV.B) and, in principle, recombination activity should not be rate-limiting (although see below for a possible exception). The relative abundance of appropriately oriented and located (between, rather than within, transcription units) short dispersed repeats seems to be a major factor in determining the prevalence of cpDNA inversions. The completely sequenced cpDNAs of tobacco, whose arrangement is the common one in vascular plants, and of *Marchantia* have no dispersed repeats >50 bp (Ohyama *et al.*, 1988; Shinozaki *et al.*, 1986; Shimada and Sugiura, 1989; K. Wolfe, unpublished) and most plants with a tobaccolike arrangement do not have detectable dispersed repeats as assayed by filter hybridization (Palmer, 1985a,b). Conversely, short repeats appear to be unusually abundant in most of the land-plant genomes that are highly rearranged, including grasses (Howe, 1985; Bowman and Dyer, 1986; Bowman *et al.*, 1988; Shimada and Sugiura, 1989), conifers (Tsai and Strauss, 1989), geranium (Palmer *et al.*, 1987a), and subclover (Milligan *et al.*, 1989). In addition, short dispersed repeats are particularly abundant in the cpDNAs of *Chlamydomonas* species (reviewed in Boynton *et al.*, 1991), whose genomes feature highly divergent gene orders.

Sequencing studies in the grasses and conifers (see above references) have shown that in several instances short repeats are situated close to the endpoints of inversions. However, direct evidence indicating that the repeats were involved in generating the inversions is lacking, and interpretation of these regions is often ambiguous. For instance, a 70-bp repeat is located at or near both endpoints of the largest grass-specific inversion (of about 28 kb; Fig. 1). Howe (1985) interpreted this as evidence that the inversion occurred by intramolecular recombination between the two repeat elements, whereas Hiratsuka *et al.* (1989) proposed a more complicated model involving intermolecular recombination between tRNA genes, followed by deletion of one genome monomer to create the tRNA-fM/G pseudogene found at one inversion endpoint. Howe *et al.* (1988) and Shimada and Sugiura (1989) have noted that the endpoints of the two smaller grass-specific inversions (Fig. 1) are associated with tRNA genes (both functional ones and pseudogenes). Precisely how the tRNA genes, which in general show only limited simi-

larity to each other, might cause recombination is unclear. Howe *et al.* (1988) also discussed the difficulty of interpreting these rather ancient inversions (all three grass inversions are at least 50 million yr old), given the likelihood that subsequent sequence evolution will delete or obscure the sequences originally involved in promoting recombination. As Palmer *et al.* (1988b) have pointed out, a clearer understanding of the role of tRNA genes and short repeats in promoting inversions should come from the sequencing of more recent inversions, such as those occurring within such genera as *Pisum* (Palmer *et al.*, 1985b) and *Oenothera* (Herrmann *et al.*, 1983; K. Sytsma, unpublished). Such analysis should also include the most closely related unrearranged genome in each case (the reference point for the cereal comparisons has been the dicot tobacco, which last shared a common ancestor with cereals some 200 million yr ago).

Assuming short dispersed repeats are important in generating rearrangement, and noting their relative abundance in more rearranged genomes, it becomes important to consider how they might be created and dispersed in the genome. Tsai and Strauss (1989) concluded that repeats of moderate size (633 bp) in conifer cpDNA are unlikely to have been produced as part of a duplicative inversion; instead, they favor the idea that the repeats were created and spread by (duplicative) transposition. Transposition has also been proposed as being involved in repeat-element spread and associated rearrangement in subclover (Milligan *et al.*, 1989) and as having caused an ancient inversion (Fig. 1) that distinguishes the genomes of bryophytes from vascular plants (Zhou *et al.*, 1988). It is important to stress, however, that no classical transposable element has been isolated from these or any other land-plant chloroplast genomes. If these repeats are transposable, then perhaps the described ones are defective elements that depend on undescribed complete elements present elsewhere in the chloroplast or nucleus for their putative transposition.

An alternative means for creating extra copies of a sequence and spreading them about a genome is retroposition, the reverse transcription of a chloroplast transcript and its subsequent integration into the genome. The specialized case whereby a cDNA recombines with its progenitor gene was invoked in Section II.C to explain the observed evolutionary loss of chloroplast introns. The discovery of several short dispersed pseudogenes in cpDNAs of cereals (Bowman *et al.*, 1988; Howe *et al.*, 1988; Shimada and Sugiura, 1989) and subclover (Milligan *et al.*, 1989) supports the idea that some cDNAs might be integrated at random in the genome. Land-plant cpDNAs do not contain discernable reverse transcriptase genes, but chloroplasts have not been assayed for reverse transcriptase activity.

As discussed in the preceding section, rearrangements that disrupt evolutionarily primitive, cotranscribed gene clusters are extremely rare in land-plant cpDNAs, although they appear to be more frequent in green algae, especially *Chlamydomonas*. Regardless of the mechanism involved (e.g., inversion, transposition, retroposition), such dispersive rearrangements, as well as those that fragment uninterrupted or *cis*-spliced genes into *trans*-spliced ones, must in general involve the creation of new transcriptional units and thus the recruitment of new promoter elements. A genome might become preadapted toward gene- and operon-disrupting rearrangements via the amplification and spread of short dispersed repeats that fortuitously contain promoterlike elements or, alternatively, via mutations enabling RNA polymerase to recognize a more generalized promoter motif.

The large inverted repeat characteristic of most land-plant genomes has been discussed as a factor that might retard genome rearrangement, primarily because genomes lacking the repeat tend to be more rearranged than related genomes that retain the repeat (Palmer, 1985a,b; Palmer *et al.*, 1987b; Strauss *et al.*, 1988). Although no evidence forces this correlation into causality, there are at least two plausible explanations for how such a large repeat might confer an element of structural stability upon the chloroplast genome. First, recombination enzymes may be titrated by the lengthy, actively recombining (Sections III.B and IV.B) repeats, thereby decreasing the availability of these enzymes elsewhere in the genome. Second, the repeat may effectively suppress inversions between one single-copy region and either the other single-copy region or the repeat itself, as recombination between the resulting direct repeats would delete a major portion of the genome.

Arguments to explain the surprising resistance of certain bacterial chromosomes to rearrangement have been based on considerations of chromosome structure and function, including gene dosage effects and constraints on chromosome folding and replication (Schmid and Roth, 1983; Palmer, 1985a; Brewer, 1988; Segall and Roth, 1989). Of these, the one that appears most relevant to the available cpDNA evidence is the conjecture of Brewer (1988) that genes are oriented on the *E. coli* chromosome so as to avoid head-on collisions between the machineries of DNA replication and transcription. As discussed above (Section III.B.1), this model also fits well with the available data for *Euglena* cpDNA, but not for other chloroplast chromosomes. If *Euglena* conforms to this model for reasons other than sheer chance, one must wonder why other chloroplast genomes, particularly those of unicellular algae that might be subject to similar selection for rapid growth, do not also. Perhaps as more algal genomes are examined in detail, other cases will be found. One might test this conjecture for *Euglena* by introducing (via transfor-

mation) heavily transcribed regions in opposite orientation of replication forks and examining whether or not such genomes have a reduced fitness, if they survive at all.

IV. GENE SEQUENCES

A. Base Composition and Base Modification

Chloroplast DNA tends to be AT-rich in base composition. Vascular plants have a narrow range of GC content of 36–39%, whereas bryophytes and several diverse lineages of algae exhibit generally lower and more variable compositions of 25–37% GC (reviewed in Palmer, 1985a). The two most exceptional groups in this respect are the nonphotosynthetic alga *Polytoma obtusum* (Siu *et al.*, 1975), whose AT-rich (17% GC) genome might be expected to be lacking in much normal gene function, and green algae in the related genera *Acetabularia* and *Batophora* (Tymms and Schweiger, 1985; Leible *et al.*, 1989), whose genomes (50% GC) are most similar to those of cyanobacteria (52–56% GC) in base composition.

The more AT-rich chloroplast genomes also show a pronounced base-compositional heterogeneity, with noncoding DNA (both intergenic spacers and introns) being AT-rich relative to coding DNA (Ohyama *et al.*, 1988; Hallick and Buetow, 1989). For example, the completely sequenced genome of *Marchantia* has an overall GC content of 29%, with rRNA and tRNA genes being most GC-rich (53% and 52%, respectively), protein genes intermediate (29%), and spacers the least (19%) (Ohyama *et al.*, 1988). Within protein genes, codon usage is highly biased toward AT in the third position (*Marchantia* uses only 12% GC in third positions), with this bias more pronounced for the more AT-rich genomes (e.g., of *Marchantia* and *Euglena*) and less pronounced for angiosperms (Ohyama *et al.*, 1988; Wakasugi *et al.*, 1986; Hallick and Buetow, 1989). A striking feature of *Chlamydomonas* cpDNA is that different classes of genes exhibit quite different codon usage patterns. Photosynthetic genes show an unusually restricted codon usage for several amino acids (e.g., Arg, Gly), for which T is strongly preferred over A in the third position, and fail to use several amino acid codons at all (Woessner *et al.*, 1986). In contrast, a more expanded and conventional codon usage is exhibited by three intron ORFs (Rochaix *et al.*, 1985; Turmel *et al.*, 1989b) and by RNA polymerase and ribosomal protein genes (S. Fong and S. Surzycki, unpublished).

Until 2 yr ago, the only report of cpDNA methylation was that of the developmentally regulated methylation of cytosine in *Chlamydomonas reinhardtii* cpDNA that is still controversially associated with its largely uniparental inheritance (reviewed in Palmer, 1985a; Chapter 3A). However, three recent papers (Ngernprasirtsiri *et al.*, 1988a,b; Gauly and Kössel, 1989) describe tissue-specific methylation of cpDNA in angiosperms. A low level of cytosine methylation was found in nonleaf plastids (chromoplasts, amyloplasts, proplastids), whereas no methylation was detected in leaf plastids (chloroplasts, etioplasts). These authors also suggest that methylation is specifically involved in repressing plastid gene expression in nonleaf plastids. It will be interesting to see whether future studies support these claims for a role of methylation in transcriptional regulation, as well as to investigate possible roles of DNA methylation in such processes as plastid DNA transmission and mismatch repair.

B. Rates and Patterns of Nucleotide Substitutions

Rates of synonymous (silent) substitutions in protein genes are relatively low in cpDNA of angiosperms, about four times lower than in nuclear DNA of plants and mammals and 20–40 times lower than in mammalian mitochondrial DNA (Table IV). Only angiosperm mtDNA evolves more slowly, by a factor of about three. CpDNA also features a

TABLE IV
Nucleotide Substitution Rates in Eukaryotic Genomes

Genome	Synonymous Substitution Rate[a]	Relative Synonymous Rate[b]	Nonsynonymous Substitution Rate[a]
Angiosperm mitochondrial	0.5	1	0.1
Angiosperm chloroplast			
Both single-copy regions	1.5	3	0.2
Large single-copy region	1.4	3	0.2
Small single-copy region	1.8	3	0.3
Inverted repeat	0.3	0.6	0.1
Angiosperm nuclear	5.4	12	0.4
Mammalian nuclear	2–8	4–16	0.5–1.3
Mammalian mitochondrial	20–50	40–100	2–3

[a] Estimated rate of substitutions per site per 10^9 yr. Rates are derived from mean values over all genes and from two estimates of divergence times (60 million yr for maize/wheat and 200 million yr for monocots/dicots). Ranges for mammalian data are over different orders. Data are from Wolfe *et al.* (1987, 1989) and Wolfe (1989b).

[b] Normalized relative to the angiosperm mitochondrial rate.

relatively moderate transition-transversion bias (somewhat less than 2:1 relative to random expectation) compared to the extreme bias (20:1) exhibited by the rapidly evolving mtDNA of mammals (Wolfe *et al.*, 1987; Zurawski and Clegg, 1987). As expected, the rate of nonsynonymous substitutions (which cause amino acid replacements) is substantially lower than the synonymous rate (Table IV).

Temporal (between-lineage) and spatial (within-genome) heterogeneity in substitution rates has been noted, the latter occurring in both a gene-specific and region-specific manner. On average, photosynthetic proteins change significantly more slowly than others (ribosomal proteins, RNA polymerase subunits, NADH dehydrogenase subunits, and ORFs) (Wolfe and Sharp, 1988; Sugiura, 1989; Wolfe, 1989a), implying greater selective constraints on their sequence and structure. Structural and functional constraints on the evolution of specific types of plastid genes and gene products will be considered in the relevant chapters in this volume.

In angiosperm cpDNA, synonymous substitutions occur 4–6 times more rapidly in the two single-copy regions than in the inverted repeat (Table IV). The functional significance, if any, and underlying mechanism responsible for such a profound effect are unclear. Imposition of a strong bias in favor of the wild type sequence upon the copy correction process that maintains identity between the two segments of the inverted repeat should suffice to explain a rate retardation in this region. However, the nature of such a bias-recognition system is completely hypothetical at present time.

Restriction mapping studies in several groups of angiosperms suggest that whole genome mutation rates may vary severalfold, even among closely related lineages (Palmer *et al.*, 1988a; Schilling and Jansen, 1989; Baldwin *et al.*, 1990). These observations need to be extended by comparative sequencing studies in these same groups. The most obvious temporal heterogeneities revealed thus far by gene sequencing are relative decreases in the rate of nonsynonymous substitutions observed over several genes in a lineage of dicots (Wolfe *et al.*, 1987) and in *atpE* among cereals (Palmer, 1985a; Rodermel and Bogorad, 1987).

The whole question of mutation rates in algal chloroplast genomes is largely unexplored. Phylogenetic trees constructed for three chloroplast protein genes suggest that average mutation rates have not varied by more than a factor of two since the origin of plastids (Morden and Golden, 1989; Baldauf and Palmer, 1990; Douglas *et al.*, 1990). However, the data base for these comparisons is much too limited to reveal possible rate heterogeneities within and between algal lineages. Data re-

garding a possible rate retardation in the inverted repeat of algae are too fragmentary and contradictory to be conclusive (Turmel *et al.*, 1987; Fain *et al.*, 1988).

V. CONCLUDING REMARKS

The detailed comparative data base available for land-plant cpDNAs paints a picture of a very conservative and stable genome. With few exceptions, the size, conformation, inverted repeat structure, gene content, gene order, and primary sequence of cpDNA change quite slowly in land plants. Overall, the chloroplast genome of land plants is the most evolutionarily conservative genome known from eukaryotes. The structural conservation of cpDNA contrasts markedly with the fluid nature of plant mitochondrial DNA (Palmer, 1990) and nuclear DNA (Tanksley and Pichersky, 1988).

A major challenge is to understand the biochemical mechanisms and selective constraints that underly the conservative evolution of land-plant cpDNA. As discussed in some detail elsewhere (Palmer, 1990), this will require several lines of study, including (1) experimental analysis of biochemical processes (e.g., DNA replication, repair and recombination, and putative processes of reverse transcription and biased copy-correction) that affect the structure and sequence of cpDNA; (2) examination of the cellular dynamics of cpDNA organization, replication, and transmission; (3) creation and testing of novel forms of cpDNA via laboratory transformation procedures (currently these are available only for the green alga *Chlamydomonas reinhardtii;* Boynton *et al.*, 1991); and (4) continued study of informative evolutionary mutants, such as plants whose genomes are unusually rearranged, have greatly expanded or contracted inverted repeats, or are evolving in the absence of photosynthesis.

Our knowledge of the structure and evolution of algal cpDNAs is so fragmentary that the major immediate challenge is simply to gather more descriptive, comparative data. This cataloging should have several motivations. First, it is presently unclear whether plastids are monophyletic or polyphyletic in origin (see Gray, 1989; Chapter 3A). If the latter is true, then we might expect to find fundamentally different kinds of plastid genomes among the major groups of eukaryotic algae. Second, it is already apparent that the chloroplast genomes of green algae are exceedingly diverse in size and gene arrangement; however, the degree

to which this diversity reflects a few episodes of major genome reorganization, as opposed to continued and intensified processes of deletion, insertion, and rearrangement, is unclear and will require more extensive examination of phylogenetically restricted sets of algal genomes. Finally, study of the complete phylogenetic range of chloroplast genomes should yield a much fuller appreciation of those processes, such as gene transfer and intron gain, that have clearly played a major role in reshaping the genome relative to its endosymbiotic ancestor(s) but are relatively unimportant within the narrow phylogenetic context of land-plant evolution.

ACKNOWLEDGMENTS

I am grateful to Stephen Downie, Maria Kuhsel, John Logsdon, Cliff Morden, Ken Wolfe, and Mimi Zolan for critical reading of the manuscript; to Hans Bohnert, John Boynton, Don Bryant, Rose Ann Cattolico, Annette Coleman, Steve Fong, Rick Hallick, Klaus Kowallik, Ulrich Kück, Maria Kuhsel, Jim Manhart, Linda Moore, Diana Stein, Stefan Surzycki, and Ken Wolfe for supplying unpublished data; and to Deborah Katz-Downie and Jennifer Thomas for assistance in manuscript preparation. Chloroplast DNA research in my laboratory is supported by grants from the NIH (GM-35087) and NSF (DCB-8451660 and BSR-8996262).

REFERENCES

Baldauf, S. L., and Palmer, J. D. (1990). Evolutionary transfer of the chloroplast *tuf*A gene to the nucleus. *Nature* **344,** 262–265.

Baldauf, S. L., Manhart, J. R., and Palmer, J. D. (1990). Different fates of the chloroplast *tuf*A gene following its transfer to the nucleus in green algae. *Proc. Natl. Acad. Sci USA,* **87,** 5317–5321.

Baldwin, B. G., Kyhos, D. W., and Dvorák, J. (1990). Chloroplast DNA evolution and adaptive radiation in the Hawaiian silversword alliance (Asteraceae—Madiinae). *Ann. Mo. Bot. Gard.* **77,** 96–109.

Bancroft, I., Wolk, C. P., and Oren, E. V. (1989). Physical and genetic maps of the genome of the heterocyst forming cyanobacterium *Anabaena* sp. strain PCC7120. *J. Bact.* **171,** 5940–5948.

Bayer, M. G., and Schenk, H. E. A. (1989). Ferredoxin of *Cyanophora paradoxa* Korsch. is encoded on cyanellar DNA. *Curr. Genet.* **16,** 311–313.

Bedbrook, J. R., and Bogorad, L. (1976). Endonuclease recognition sites mapped on *Zea mays* chloroplast DNA. *Proc. Natl. Acad. Sci. USA* **73,** 4309–4313.

Blasko, K., Kaplan, S. A., Higgins, K. G., Wolfson, R., and Sears, B. B. (1988). Variation

in copy number of a 24-base pair tandem repeat in the chloroplast DNA of *Oenothera hookeri* strain Johansen. *Curr. Genet.* **14,** 287–292.

Bogorad, L. (1975). Evolution of organelle and eukaryotic genomes. *Science* **188,** 891–898.

Bowman, C. M., and Dyer, T. A. (1986). The location and possible evolutionary significance of small dispersed repeats in wheat ctDNA. *Curr. Genet.* **10,** 931–941.

Bowman, C. M., Barker, R. F., and Dyer, T. A. (1988). In wheat ctDNA, segments of ribosomal protein genes are dispersed repeats, probably conserved by nonreciprocal recombination. *Curr. Genet.* **14,** 127–136.

Boynton, J. E., Gillham, N. W., Newman, S. M., and Harris, E. H. (1991). Organelle genetics and transformation of *Chlamydomonas*. *In* "Plant Gene Research," Vol. VI. (R. G. Herrmann, ed.). Springer Verlag, New York, in press.

Breiteneder, H. Seiser, C., Löffelhardt, W., Michalowski, C., and Bohnert, H. J. (1988). Physical map and protein gene map of cyanelle DNA from the second known isolate of *Cyanophora paradoxa* (Kies-strain). *Curr. Genet.* **13,** 199–206.

Brewer, B. J. (1988). When polymerases collide: Replication and the transcriptional organization of the *E. coli* chromosome. *Cell* **53,** 679–686.

Bryant, D. A., and Stirewalt, V. L. (1990). The cyanelle genome of *Cyanophora paradoxa* encodes ribosomal proteins not encoded by the chloroplast genomes of higher plants. *FEBS Lett.* **259,** 273–280.

Calie, P. J., and Hughes, K. W. (1987). The consensus land plant chloroplast gene order is present, with two alterations, in the moss *Physcomitrella patens*. *Mol. Gen. Genet.* **208,** 335–341.

Cantrell, A., and Bryant, D. A. (1987). Molecular cloning and nucleotide sequence of the *psa*A and *psa*B genes of the cyanobacterium *Synechococcus* sp. PCC7002. *Plant Mol. Biol.* **9,** 453–468.

Cattolico, R. A. (1986). Chloroplast evolution in algae and land plants. *Trends Ecol. Evol.* **1,** 64–67.

Cavalier-Smith, T. (1987). The simultaneous symbiotic origin of mitochondria, chloroplasts, and microbodies. *Ann. N.Y. Acad. Sci.* **503,** 55–71.

Cheung, W. Y., and Scott, N. S. (1989). A contiguous sequence in spinach nuclear DNA is homologous to three separated sequences in chloroplast DNA. *Theor. Appl. Genet.* **77,** 625–633.

Christopher, D. A., and Hallick, R. B. (1989). *Euglena gracilis* chloroplast ribosomal protein operon: A new chloroplast gene for ribosomal protein L5 and description of a novel organelle intron category designated group III. *Nucleic Acids Res.* **17,** 7591–7608.

Coates, D., and Cullis, C. A. (1987). Chloroplast DNA variability among *Linum* species. *Amer. J. Bot.* **74,** 260–268.

Crouse, E. J., Schmitt, J. M., and Bohnert, H. J. (1985). Chloroplast and cyanobacterial genomes, genes and RNAs: A compilation. *Plant Mol. Biol. Reporter* **3,** 43–89.

Cushman, J. C., Hallick, R. B., and Price, C. A. (1988). The two genes for the P700 chlorophyll *a* apoproteins on the *Euglena gracilis* chloroplast genome contain multiple introns. *Curr. Genet.* **13,** 159–171.

Dalmon, J., Loiseaux, S., and Bazetoux, S. (1983). Heterogeneity of plastid DNA of two species of brown algae. *Plant Sci. Lett.* **29,** 243–253.

Day, A., and Ellis, T. H. N. (1985). Deleted forms of plastid DNA in albino plants from cereal anther culture. *Curr. Genet.* **9,** 671–678.

Deng, X.-W., Wing, R. A., and Gruissem, W. (1980). The chloroplast genome exists in multimeric forms. *Proc. Natl. Acad. Sci. USA* **86,** 4156–4160.

dePamphilis, C. W., and Palmer, J. D. (1989). Evolution and function of plastid DNA: A review with special reference to nonphotosynthetic plants. *In* "Physiology, Biochemis-

try, and Genetics of Nongreen Plastids" (C. D. Boyer, J. C. Shannon, and R. C. Hardison, eds.), pp. 182–202. American Society of Plant Physiology, Rockville, Maryland.

dePamphilis, C. W., and Palmer, J. D. (1990). Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. *Nature (London)* **348,** 337–339.

Douglas, S. E. (1988). Physical mapping of the plastid genome from the chlorophyll *c*-containing alga, *Cryptomonas* sp. *Curr. Genet.* **14,** 591–598.

Douglas, S. E., and Durnford, D. G. (1989). The small subunit of ribulose-1,5-bisphosphate carboxylase is plastid-encoded in the chlorophyll *c*-containing alga *Cryptomonas* sp. *Plant Mol. Biol.* **13,** 13–20.

Douglas, S. E., Durnford, D. G., and Morden, C. W. (1990). Nucleotide sequence of the gene for the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Cryptomonas* sp.: Evidence supporting the polyphyletic origin of plastids. *J. Phycol.*, **26,** 500–508.

Downie, S. R., Olmstead, R. G., Zurawski, G., Soltis, D. E., Soltis, P. S., Watson, J. C., and Palmer, J. D. (1991). Loss of the chloroplast DNA *rpl*2 intron demarcates six lineages of dicotyledons: Molecular and phylogenetic implications. *Evolution,* in press.

Downie, S. R., and Palmer, J. D. (1991). Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. *In* "Plant Molecular Systematics" (P. Soltis, D. Soltis, and J. J. Doyle, eds.). Chapman and Hall, New York, in press.

Dujon, B. (1989). Group I introns as mobile genetic elements: Facts and mechanistic speculations—A review. *Gene* **82,** 91–114.

Durocher, V., Gauthier, A., Bellemare, G., and Lemieux, C. (1989). An optional group I intron between the chloroplast small subunit rRNA genes of *Chlamydomonas moewusii* and *C. eugametos. Curr. Genet.* **15,** 277–282.

Ebert, C., Tymms, M. J., and Schweiger, H.-G. (1985). Homology between 4.3 μm minicircular and plastomic DNA in chloroplasts of *Acetabularia cliftonii. Mol. Gen. Genet.* **200,** 187–192.

Ellis, T. H. N. and Day, A. (1986). A hairpin plastid genome in barley. *EMBO J.* **5,** 2769–2774.

Erickson, J. M., Rahire, M., and Rochaix, J.-D. (1984). *Chlamydomonas reinhardii* gene for the 32,000 mol. wt. protein of photosystem II contains four large introns and is located entirely within the chloroplast inverted repeat. *EMBO J.* **3,** 2753–2762.

Evrard, J. L., Kuntz, M., Straus, N. A., and Weil, J.-H. (1988). A class-I intron in a cyanelle tRNA gene from *Cyanophora paradoxa:* Phylogenetic relationship between cyanelles and plant chloroplasts. *Gene* **71,** 115–122.

Evrard, J. L., Kuntz, M., and Weil, J. H. (1990). The nucleotide sequence of five ribosomal protein genes from the cyanelles of *Cyanophora paradoxa:* Implications concerning the phylogenetic relationship between cyanelles and chloroplasts. *J. Mol. Evol.* **30,** 16–25.

Fain, S. R., Druehl, L. D., and Bailie, D. L. (1988). Repeat and single copy sequences are differentially conserved in the evolution of kelp chloroplast DNA. *J. Phycol.* **24,** 292–302.

Falk, G., and Walker, J. E. (1988). DNA sequence of a gene cluster coding for subunits of the F_o membrane sector of ATP synthase in *Rhodospirillum rubrum. Biochem. J.* **254,** 109–122.

Fearnley, I. M., Runswick, M. J., and Walker, J. E. (1989). A homologue of the nuclear coded 49 kd subunit of bovine mitochondrial NADH-ubiquinone reductase is coded in chloroplast DNA. *EMBO J.* **8,** 665–672.

Froman, B. E., Tait, R. C., and Gottlieb, L. D. (1989). Isolation and characterization of the phosphoglucose isomerase gene from *Escherichia coli. Mol. Gen. Genet.* **217,** 126–131.

Gantt, J. S. (1988). Nucleotide sequences of cDNAs encoding four complete nuclear-encoded plastid ribosomal proteins. *Curr. Genet.* **14,** 519–528.

Gauly, A., and Kössel, H. (1989). Evidence for tissue-specific cytosine-methylation of plastid DNA from *Zea mays*. *Curr. Genet.* **15,** 371–376.

Gilbert, W., Marchionni, M., and McKnight, G. (1986). On the antiquity of introns. *Cell* **46,** 151–154.

Gingrich, J. C., and Hallick, R. B. (1985). The *Euglena gracilis* chloroplast ribulose-1,5-bisphosphate carboxylase gene. I. Complete DNA sequence and analysis of the nine intervening sequences. *J. Biol. Chem.* **260,** 16156–16161.

Goff, L. J., and Coleman, A. W. (1988). The use of plastid DNA restriction endonuclease patterns in delineating red algal species and populations. *J. Phycol.* **24,** 357–368.

Gray, M. W. (1989). The evolutionary origins of organelles. *Trends Genet.* **5,** 294–299.

Hallick, R. B. (1989). Proposals for the naming of chloroplast genes. II. Update to the nomenclature of genes for thylakoid membrane polypeptides. *Plant Mol. Biol. Reporter* **7,** 266–275.

Hallick, R. B., and Buetow, D. E. (1989). Chloroplast DNA. *In* "The Biology of *Euglena*" Vol. IV (D. E. Buetow, ed.), pp. 351–414. Academic Press, New York.

Hasebe, M., and Iwatsuki, K. (1990). Chloroplast DNA from *Adiantum capillusveneris* L., a fern species (Adiantaceae); clone bank, physical map and unusual gene localization in comparison with angiosperm chloroplast DNA. *Curr. Genet.* **17,** 359–364.

Herrmann, R. G., Palta, H. K., and Kowallik, K. V. (1980). Chloroplast DNA from three archegoniates. *Planta* **148,** 319–327.

Herrmann, R. G., Westhoff, P., Alt, J., Winter, P., Tittgen, J., Bisanz, C., Sears, B. B., Nelson, N., Hurt, E., Hauska, G., Viebrock, A., and Sebald, W. (1983). Identification and characterization of genes for polypeptides of the thylakoid membrane. *In* "Structure and Function of Plant Genomes" (O. Ciferri and L. Dure III, eds.), pp. 143–153. Plenum Publishing Corporation, New York.

Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.-R., Meng, B.-Y., Li, Y.-Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., and Sugiura, M. (1989). The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.* **217,** 185–194.

Howe, C. J. (1985). The endpoints of an inversion in wheat chloroplast DNA are associated with short repeated sequences containing homology to *att*-lambda. *Curr. Genet.* **10,** 139–145.

Howe, C. J., Barker, R. F., Bowman, C. M., and Dyer, T. A. (1988). Common features of three inversions in wheat chloroplast DNA. *Curr. Genet.* **13,** 343–349.

Hudson, G. S., Mason, J. G., Holton, T. A., Koller, B., Cox, G. B., Whitfeld, P. R., and Bottomley, W. (1987). A gene cluster in the spinach and pea chloroplast genomes encoding one CF_1 and three CF_0 subunits of the H^+-ATP synthase complex and the ribosomal protein S2. *J. Mol. Biol.* **196,** 283–298.

Hudson, G. S., Holton, T. A., Whitfeld, P. R., and Bottomley, W. (1988). Spinach chloroplast *rpo*BC genes encode three subunits of the chloroplast RNA polymerase. *J. Mol. Biol.* **200,** 639–654.

Hwang, S.-R., and Tabita, R. R. (1989). Cloning and expression of the chloroplast-encoded *rbc*L and *rbs*C genes from the marine diatom *Cylindrotheca* sp. strain N1. *Plant Mol. Biol.* **13,** 69–79.

Joyce, P. B. M., and Gray, M. W. (1989). Chloroplast-like transfer RNA genes expressed in wheat mitochondria. *Nucleic Acids Res.* **17,** 5461–5476.

Karabin, G. D., Farley, M., and Hallick, R. B. (1984). Chloroplast gene for M_r 32,000 polypeptide of photosystem II in *Euglena gracilis* is interrupted by four introns with conserved boundary sequences. *Nucleic Acids Res.* **12,** 5801–5812.

Keller, M., Weil, J. H., and Nair, C. K. K. (1989). Nucleotide sequence of the *psb*B gene of *Euglena gracilis*. *Plant Mol. Biol.* **13,** 723–725.

Kirk, J. T. O. (1986). The discovery of chloroplast DNA. *Bioessays* **4,** 36–38.

Knowles, V. L., Dennis, D. T., and Plaxton, W. C. (1989). Purification of a novel pyruvate kinase from a green alga. *FEBS Lett.* **259,** 130–132.

Kohchi, T., Yoshida, T., Komano, T., and Ohyama, K. (1988a). Divergent mRNA transcription in the chloroplast *psb*B operon. *EMBO J.* **7,** 885–891.

Kohchi, T., Ogura, Y., Umesono, K., Yamada, Y., Komano, T., Ozeki, H., and Ohyama, K. (1988b). Ordered processing and splicing in a polycistronic transcript in liverwort chloroplasts. *Curr. Genet.* **14,** 147–154.

Koller, B., and Delius, H. (1984). Intervening sequences in chloroplast genomes. *Cell* **36,** 613–622.

Kolodner, R., and Tewari, K. K. (1975). The molecular size and conformation of the chloroplast DNA from higher plants. *Biochim. Biophys. Acta* **402,** 372–390.

Kolodner, R., and Tewari, K. K. (1979). Inverted repeats in chloroplast DNA from higher plants. *Proc. Natl. Acad. Sci. USA* **76,** 41–45.

Kowallik, K. V. (1989). Molecular aspects and phylogenetic implications of plastid genomes of certain chromophytes. *In* "The Chromophyte Algae: Problems and Perspectives (J. C. Green, B. S. C. Leadbeater, and W. L. Diver, eds.), pp. 101–124. Clarendon Press, Oxford.

Kück, U. (1989). The intron of a plastid gene from a green alga contains an open reading frame for a reverse transcriptase-like enzyme. *Mol. Gen. Genet.* **218,** 257–265.

Kück, U., Choquet, Y., Schneider, M., Dron, M., and Bennoun, P. (1987). Structural and transcription analysis of two homologous genes for the P700 chlorophyll *a*-apoproteins in *Chlamydomonas reinhardii:* Evidence for *in vivo trans*-splicing. *EMBO J.* **6,** 2185–2195.

Kuhsel, M. (1988). Molecular characterization of the plastid genome of the brown alga *Dictyota dichotoma* (Huds.) Lamour. Ph.D. Thesis, University of Düsseldorf.

Kuhsel, M., and Kowallik, K. V. (1987). The plastome of a brown alga, *Dictyota dichotoma*. II. Location of structural genes coding for ribosomal RNAs, the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase and for polypeptides of photosystems I and II. *Mol. Gen. Genet.* **207,** 361–368.

Kuhsel, M. G., Strickland, R., and Palmer, J. D. (1990). An ancient group I intron is shared by eubacteria and chloroplasts. *Science* **250,** 1570–1573.

Kurzok, H.-G., and Feierabend, J. (1984). Comparison of a cytosolic and a chloroplast triosephosphate isomerase isoenzyme from rye leaves. II. Molecular properties and phylogenetic relationships. *Biochim. Biophys. Acta* **788,** 222–233.

Lambert, D. H., Bryant, D. A., Stirewalt, V. L., Dubbs, J. M., Stevens, S. E., Jr., and Porter, R. D. (1985). Gene map for the *Cyanophora paradoxa* cyanelle genome. *J. Bacteriol.* **164,** 659–664.

Lavin, M., Doyle, J. J., and Palmer, J. D. (1990). Evolutionary significance of the loss of the chloroplast DNA inverted repeat in the Leguminosae subfamily Papilionoideae. *Evolution,* **44,** 390–402.

Leible, M. B., Berger, S., and Schweiger, H.-G. (1989). The plastome of *Acetabularia mediterranea* and *Batophora oerstedii:* Inter- and intraspecific variability and physical properties. *Curr. Genet.* **15,** 355–361.

Lemieux, B., Turmel, M., and Lemieux, C. (1988). Unidirectional gene conversions in the chloroplast of *Chlamydomonas* interspecific hybrids. *Mol. Gen. Genet.* **212,** 48–55.

Lemire, B. D., Frankhauser, C., Baker, A., and Schatz, G. (1989). The mitochondrial tar-

geting function of randomly generated peptide sequences correlates with predicted helical amphiphilicity. *J. Biol. Chem.* **264,** 20206–20215.

Lewin, B. (1990). "Genes IV." Oxford University Press, Oxford.

Li, N., and Cattolico, R. A. (1987). Chloroplast genome characterization in the red alga *Griffithsia pacifica*. *Mol. Gen. Genet*, **209,** 343–351.

Li, N., and Cattolico, R. A. (1988). Study of chloroplast genome organization and gene arrangement in Rhodophyte, *Griffithsia pacifica* and Chromophyte, *Ochromonas danica*. *J. Phycol.* **24** (Suppl.), 8.

Li, N., Hedberg, M., and Cattolico, R. A. (1991). Chloroplast DNA heterogeneity in *Monodus* sp. (Eustigmatophyta). *Curr. Genet.*, in press.

Lidholm, J., Szmidt, A. E., Hällgren, J.-E., and Gustafsson, P. (1988). The chloroplast genomes of conifers lack one of the rRNA-encoding inverted repeats. *Mol. Gen. Genet.* **212,** 6–10.

Linne von Berg, K.-H., and Kowallik, K. V. (1988). Structural organization and evolution of the plastid genome of *Vaucheria sessilis* (Xanthophyceae). *BioSystems* **21,** 239–247.

Linne von Berg, K.-H., Schmidt, M., Linne von Berg, G., Sturm, K., Hennig, A., and Kowallik, K. V. (1982). The chloroplast genome (plastome) from algae of different phylogenetic relationships. *Brit. Phycol. J.* **17,** 235.

Liu, X.-Q., Gillham, N. W., and Boynton, J. E. (1989). Chloroplast ribosomal protein gene *rps*12 of *Chlamydomonas reinhardtii:* wild-type sequence, mutation to streptomycin resistance and dependence, and function in *Escherichia coli*. *J. Biol. Chem.* **264,** 16100–16108.

Löffelhardt, W., Kraus, M., Pfanzagl, B., Götz, M., Brandtner, M., Markmann-Munlish, U., Subramanian, A. R., Michalowski, C., and Bohnert, H. J. (1990). Cyanelle genes for components of the translation apparatus. *In* "Endocytobiology IV" (P. Nardon, V. Gianinazzi, A. M. Grenier, L. Margulis, and D. C. Smith, eds.), pp. 561–564. INRA, Paris.

Loiseaux-de Goër, S. L., Markowicz, Y., Dalmon, J., and Audren, H. (1988). Physical maps of the two circular plastid DNA molecules of the brown alga *Pylaiella littoralis* (L.) Kjellm. *Curr. Genet.* **14,** 155–162.

Longstaff, M., Raines, C. A., McMorrow, E. M., Bradbeer, J. W., and Dyer, T. A. (1989). Wheat phosphoglycerate kinase: Evidence for recombination between the genes for the chloroplastic and cytosolic enzymes. *Nucleic Acids Res.* **17,** 6569–6580.

Lou, J. K., Cruz, F. D., and Wu, M. (1989). Nucleotide sequence of the chloroplast ribosomal protein gene L14 in *Chlamydomonas reinhardtii*. *Nucleic Acids Res.* **17,** 3587.

Manhart, J. R., and Palmer, J. D. (1990). The gain of two chloroplast tRNA introns marks the green algal ancestors of land plants. *Nature* **345,** 268–270.

Manhart, J. R., Kelly, K., Dudock, B. S., and Palmer, J. D. (1989). Unusual characteristics of *Codium vulgare* chloroplast DNA revealed by physical and gene mapping. *Mol. Gen. Genet.* **216,** 417–421.

Manhart, J. R., Hoshaw, R. W., and Palmer, J. D. (1990). Unique chloroplast genome in *Spirogyra maxima* (Chlorophyta) revealed by physical and gene mapping. *J. Phycol.*, **26,** 490–494.

Maréchal-Drouard, L., Weil, J.-H., and Guillemaut, P. (1988). Import of several tRNAs from the cytoplasm into the mitochondria in bean *Phaseolus vulgaris*. *Nucleic Acids Res.* **16,** 4777–4788.

Markowicz, Y., Goër, S. L., and Mache, R. (1988). Presence of a 16S rRNA pseudogene in the bimolecular plastid genome of the primitive brown alga *Pylaiella littoralis*. Evolutionary implications. *Curr. Genet.* **14,** 599–608.

Marsh, J. J., Wilson, K. J., and Lebherz, H. G. (1989). Structural similarities between spinach chloroplast and cytosolic class I fructose-1,6-bisphosphate aldolases. *Plant Physiol.* **91,** 1393–1401.

Martin, W., and Cerff, R. (1986). Prokaryotic features of a nucleus-encoded enzyme. cDNA

sequences for chloroplast and cystolic glyceraldehyde-3-phosphate dehydrogenases from mustard (*Sinapis alba*). *Eur. J. Biochem.* **159,** 323–331.

Matsubayashi, T., Wakasugi, T., Shinozaki, K., Yamaguchi-Shinozaki, K., Zaita, N., Hidaka, T., Meng, B.-Y., Ohto, C., Tanaka, M., Kato, A., Maruyama, T., and Sugiura, M. (1987). Six chloroplast genes (*nadh*A–F) homologous to human mitochondrial genes encoding components of the respiratory chain NADH dehydrogenase are actively expressed: Determination of the splice sites in *nadh*A and *nadh*B pre-mRNAs. *Mol. Gen. Genet.* **210,** 385–393.

Meng, B.-Y., Tanaka, M., Wakasugi, T., Ohme, M., Shinozaki, K., and Sugiura, M. (1988). Cotranscription of the genes encoding two P700 chlorophyll *a* apoproteins with the gene for ribosomal protein CS14: Determination of the transcriptional initiation site by *in vitro* capping. *Curr. Genet.* **14,** 395–400.

Michalowski, C. B., Pfanzage, B. Löffelhardt, W., and Bohnert, H. J. (1990). The cyanelle S10 *spc* ribosomal protein gene operon from *Cyanophora paradoxa*. *Mol. Gen. Genet.*, **224,** 222–231.

Michel, F., Umesono, K., and Ozeki, H. (1989). Comparative and functional anatomy of group II catalytic introns—A review. *Gene* **82,** 5–30.

Milligan, B. G., Hampton, J., and Palmer, J. D. (1989). Dispersed repeats and structural reorganization in subclover chloroplast DNA. *Mol. Biol. Evol.* **6,** 355–368.

Misonou, T., Ishihara, J., Pak, J. Y., and Nitta, T. (1989). Restriction endonuclease analysis of chloroplast and mitochondrial DNAs from *Bryopsis* (Derbesiales, Chlorophyta). *Phycologia* **28,** 422–428.

Montandon, P.-E., and Stutz, E. (1983). Nucleotide sequence of a *Euglena gracilis* chloroplast genome region coding for the elongation factur tu; evidence for a spliced mRNA. *Nucleic Acids Res.* **11,** 5877–5892.

Moon, E., Kao, T.-H., and Wu, R. (1987). Rice chloroplast DNA molecules are heterogenous as revealed by DNA sequences of a cluster of genes. *Nucleic Acids Res.* **15,** 611–630.

Moore, L. J. (1990). The nature and extent of intraspecific variation in chloroplast DNAs of sexually isolated populations of *Pandorina morum* Bory. Ph.D. Thesis, Brown University, Providence, Rhode Island.

Morden, C. W., and Golden, S. S. (1989). *psb*A genes indicate common ancestry of prochlorophytes and chloroplasts. *Nature* **337,** 382–385.

Mörl, M., and Schmelzer, C. (1990). Integration of group II intron bI1 into a foreign RNA by reversal of the self-splicing reaction in vitro. *Cell* **60,** 629–636.

Murata, N., Miyao, M., Hayashida, N., Hidaka, T., and Sugiura, M. (1988). Identification of a new gene in the chloroplast genome encoding a low-molecular-mass polypeptide of photosystem II complex. *FEBS Lett.* **235,** 283–288.

Neuhaus, H., and Link, G. (1987). The chloroplast $tRNA^{LYS}$ (UUU) gene from mustard (*Sinapis alba*) contains a class II intron potentially coding for a maturase-related polypeptide. *Curr. Genet.* **11,** 251–257.

Neumann, C., Brandtner, M., Kraus, M., Jakowitsch, J., Bayer, M. G., Maier, T. L., Schenk, H. E. A., and Löffelhardt, W. (1990). *FEBS Lett.* **268,** 55–58.

Nickoloff, J. A., Christopher, D. A., Drager, R. G., and Hallick, R. B. (1989). Nucleotide sequence of the *Euglena gracilis* chloroplast genes for isoleucine, phenylalanine and cysteine transfer RNAs and ribosomal protein S14. *Nucleic Acids Res.* **17,** 4882.

Ngernprasirtsiri, J., Kobayashi, H., and Akazawa, T. (1988a). DNA methylation as a mechanism of transcriptional regulation in nonphotosynthetic plastids in plant cell. *Proc. Natl. Acad. Sci. USA* **85,** 4750–4754.

Ngernprasirtsiri, J., Kobayashi, H., and Akazawa, T. (1988b). DNA methylation occurred

around lowly expressed genes of plastid DNA during tomato fruit development. *Plant Physiol.* **88,** 16–20.

Ogihara, Y., Terachi, T., and Sasakuma, T. (1988). Intramolecular recombination of chloroplast genome mediated by short direct-repeat sequences in wheat species. *Proc. Natl. Acad. Sci. USA* **85,** 8573–8577.

Ohto, C., Torazawa, K., Tanaka, M., Shinozaki, K., and Sugiura, M. (1988). Transcription of ten ribosomal protein genes from tobacco chloroplasts: A compilation of ribosomal protein genes found in the tobacco chloroplast genome. *Plant Mol. Biol.* **11,** 589–600.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., and Ozeki, H. (1986). Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* **322,** 572–574.

Ohyama, K., Fukuzawa, H., Kohchi, T., Sano, T., Sano, S., Shirai, H., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., and Ozeki, H. (1988). Structure and organization of *Marchantia polymorpha* chloroplast genome I. Cloning and gene identification. *J. Mol. Biol.* **203,** 281–298.

Oppermann, T., Hong, T.-H., and Surzycki, S. J. (1989). Chloroplast and nuclear genomes of *Chlamydomonas reinhardtii* share homology with *Escherichia coli* genes for DNA replication, repair and transcription. *Curr. Genet.* **15,** 39–46.

Padmanabhan, P., and Green, B. (1978). The kinetic complexity of *Acetabularia* chloroplast DNA. *Biochim. Biophys. Acta* **521,** 67–73.

Palmer, J. D. (1985a). Evolution of chloroplast and mitochondrial DNA in plants and algae. *In* "Monographs in Evolutionary Biology: Molecular Evolutionary Genetics" (R. J. MacIntyre, ed.), pp. 131–240. Plenum Publishing Corporation, New York.

Palmer, J. D. (1985b). Comparative organization of chloroplast genomes. *Annu. Rev. Genet.* **19,** 325–354.

Palmer, J. D. (1990). Contrasting modes and tempos of genome evolution in land plant organelles. *Trends Genet.* **6,** 115–120.

Palmer, J. D., and Stein, D. B. (1986). Conservation of chloroplast genome structure among vascular plants. *Curr. Genet.* **10,** 823–833.

Palmer, J. D., Boynton, J. E., Gillham, N. W., and Harris, E. H. (1985a). Evolution and recombination of the large inverted repeat in *Chlamydomonas* chloroplast DNA. *In* "Molecular Biology of the Photosynthetic Apparatus" (C. Arntzen, L. Bogorad, S. Bonitz, and K. Steinback, eds.), pp. 269–278. Cold Spring Harbor, New York.

Palmer, J. D., Jorgensen, R. A., and Thompson, W. F. (1985b). Chloroplast DNA variation and evolution in *Pisum*: Patterns of change and phylogenetic analysis. *Genetics* **109,** 195–213.

Palmer, J. D., Nugent, J. M., and Herbon, L. A. (1987a). Unusual structure of geranium chloroplast DNA: A triple-sized repeat, extensive gene duplications, multiple inversions and new repeat families. *Proc. Natl. Acad. Sci. USA* **84,** 769–773.

Palmer, J. D., Osorio, B., Aldrich, J., and Thompson, W. F. (1987b). Chloroplast DNA evolution among legumes: Loss of a large inverted repeat occurred prior to other sequence rearrangements. *Curr. Genet.* **11,** 275–286.

Palmer, J. D., Jansen, R. K., Michaels, H., Manhart, J., and Chase, M. W. (1988a). Phylogenetic analysis of chloroplast DNA variation. *Ann Mo. Bot. Gard.* **75,** 1180–1218.

Palmer, J. D., Osorio, B., and Thompson, W. F. (1988b). Evolutionary significance of inversions in legume chloroplast DNAs. *Curr. Genet.* **14,** 65–74.

Perlman, P. S., and Butow, R. A. (1989). Mobile introns and intron-encoded proteins. *Science* **246,** 1106–1109.

Quigley, F., Martin, W. F., and Cerff, R. (1988). Intron conservation across the prokaryote–eukaryote boundary: Structure of the nuclear gene for chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize. *Proc. Natl. Acad. Sci. USA* **85,** 2672–2676.

Reimann, A., and Kück, U. (1989). Nucleotide sequence of the plastid genes for apocytochrome b_6 (*pet*B) and subunit IV of the cytochrome b_6–f complex (*pet*D) from the green alge *Chlorella prototheco*ides: Lack of introns. *Plant Mol. Biol.* **13,** 255–256.

Reith, M., and Cattolico, R. A. (1986). Inverted repeat of *Olisthodiscus luteus* chloroplast DNA contains genes for both subunits of ribulose-1,5-bisphosphate carboxylase and the 32,000-dalton Q_B protein: Phylogenetic implications. *Proc. Natl. Acad. Sci. USA* **83,** 8599–8603.

Robertson, D., Boynton, J. E., and Gillham, N. W. (1990). Cotranscription of the wild type chloroplast *atp*E gene encoding the CF_1/CF_0 epsilon subunit with the 3′ half of the *rps7* gene in *Chlamydomonas reinhardtii* and characterization of frameshift mutations in *atp*E. *Mol. Gen. Genet.* **221,** 155–163.

Rochaix, J.-D. (1978). Restriction endonuclease map of the chloroplast DNA of *Chlamydomonas reinhardii*. *J. Mol. Biol.* **126,** 597–617.

Rochaix, J.-D., Rahire, M., and Michel, F. (1985). The chloroplast ribosomal intron of *Chlamydomonas reinhardii* codes for a polypeptide related to mitochondrial maturases. *Nucleic Acids Res.* **13,** 975–984.

Rochaix, J.-D., Kuchka, M., Mayfield, S., Schirmer-Rahire, M., Girard-Bascou, J., and Bennoun, P. (1989). Nuclear and chloroplast mutations affect the synthesis or stability of the chloroplast *psb*C gene product in *Chlamydomonas reinhardtii*. *EMBO J.* **8,** 1013–1021.

Rodermel, S. R., and Bogorad, L. (1987). Molecular evolution and nucleotide sequences of the maize plastid genes for the alpha subunit of CF_1 (*atp*A) and the proteolipid subunit of CF_0 (*atp*H). *Genetics* **116,** 127–139.

Rogers, J. H. (1989). How were introns inserted into nuclear genes? *Trends Genet.* **5,** 213–216.

Rothermel, B. A., and Nelson, T. (1989). Primary structure of the maize NADP-dependent malic enzyme. *J. Biol. Chem.* **264,** 19587–19592.

Schantz, R., and Bogorad, L. (1988). Maize chloroplast genes *ndh*D, *ndh*E and *psa*C. Sequences, transcripts and transcript pools. *Plant Mol. Biol.* **11,** 239–247.

Schilling, E. E., and Jansen, R. K. (1989). Restriction fragment analysis of chloroplast DNA and the systematics of *Viguiera* and related genera (Asteraceae: Heliantheae). *Amer. J. Bot.* **76,** 1769–1778.

Schlunegger, B., and Stutz, E. (1984). The *Euglena gracilis* chloroplast genome: Structural features of a DNA region possibly carrying the single origin of DNA replication. *Curr. Genet.* **8,** 629–634.

Schmid, M. B., and Roth, J. R. (1983). Selection and endpoint distribution of bacterial inversion mutations. *Genetics* **105,** 539–557.

Schuster, A. M., Waddle, J. A., Korth, K., and Meints, R. H. (1990). The chloroplast genome of an exsymbiotic *Chlorella*-like green alga. *Plant Mol. Biol.* **14,** 859–862.

Schweiger, H.-G., de Groot, E. J., Leible, M. B., and Tymms, M. J. (1986). Conservative and variable features of the chloroplast genome of *Acetabularia*. *In* "Regulation of Chloroplast Differentiation," Plant Biology Ser., Vol. 2 (G. Akoyunoglou and H. Senger, eds.), pp. 467–476, Alan R. Liss, New York.

Segall, A. M., and Roth, J. R. (1989). Recombination between homologies in direct and inverse orientation in the chromosome of *Salmonella:* Intervals which are nonpermissive for inversion formation. *Genetics* **122,** 737–747.

Shen, G. F., Chen, K., Wu, M., and Kung, S. D. (1982). *Nicotiana* chloroplast genome IV. *N. acuminata* has larger inverted repeats and genome size. *Mol. Gen. Genet.* **187,** 12–18.

Shih, M.-C., Lazar, G., and Goodman, H. M. (1986). Evidence in favor of the symbiotic origin of chloroplasts: Primary structure and evolution of tobacco glyceraldehyde-3-phosphate dehydrogenase. *Cell* **47,** 73–80.

Shih, M.-C., Heinrich, P., and Goodman, H. M. (1988). Intron existence predated the divergence of eukaryotes and prokaryotes. *Science* **242,** 1164–1166.

Shimada, H., and Sugiura, M. (1989). Pseudogenes and short repeated sequences in the rice chloroplast genome. *Curr. Genet.* **16,** 293–301.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B.-Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., and Sugiura, M. (1986). The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. *EMBO J.* **5,** 2043–2049.

Siemeister, G., and Hachtel, W. (1989). A circular 73 kb DNA from the colourless flagellate *Astasia longa* that resembles the chloroplast DNA of *Euglena:* Restriction and gene map. *Curr. Genet.* **15,** 435–441.

Siu, C.-H., Chiang, K.-S., and Swift, H. (1975). Characterization of cytoplasmic and nuclear genomes in the colorless alga *Polytoma* V. Molecular structure and heterogeneity of leucoplast DNA. *J. Mol. Biol.* **98,** 369–391.

Starnes, S. M., Lambert, D. H., Maxwell, E. S., Stevens, S. E., Porter, R. D., and Shiveley, J. M. (1985). Cotranscription of the large and small subunit genes of ribulose-1,5-bisphosphate carboxylase oxygenase in *Cyanophora paradoxa. FEMS Microbiol. Lett.* **28,** 165–169.

Stein, D. B., Palmer, J. D., and Thompson, W. F. (1986). Structural evolution and flip-flop recombination of chloroplast DNA in the fern genus *Osmunda. Curr. Genet.* **10,** 835–841.

Steinmüller, K., Ley, A. C., Steinmetz, A. A., Sayre, R. T., and Bogorad, L. (1989). Characterization of the *ndh*C-*psb*G-ORF157/159 operon of maize plastid DNA and of the cyanobacterium *Synechocystis* sp. PCC6803. *Mol. Gen. Genet.* **216,** 60–69.

Stern, D. B., and Palmer, J. D. (1984). Extensive and widespread homologies between mitochondrial DNA and chloroplast DNA in plants. *Proc. Natl. Acad. Sci. USA* **81,** 1946–1950.

Strauss, S. H., Palmer, J. D., Howe, G. T., and Doerksen, A. H. (1988). Chloroplast genomes of two conifers lack an inverted repeat and are extensively rearranged. *Proc. Natl. Acad. Sci. USA* **85,** 3898–3902.

Sugiura, M. (1989). The chloroplast chromosomes in land plants. *Annu. Rev. Cell Biol.* **5,** 51–70.

Surguchov, A. P. (1987). Common genes for mitochondrial and cytoplasmic proteins. *Trends Biochem. Sci.* **12,** 335–338.

Takaiwa, F., and Sugiura, M. (1982). Nucleotide sequence of the 16S–23S spacer region in an rRNA gene cluster from tobacco chloroplast DNA. *Nucleic Acids Res.* **10,** 2665–2676.

Tanaka, M., Obokata, J., Chunwongse, J., Shinozaki, K., and Sugiura, M. (1987). Rapid splicing and stepwise processing of a transcript from the *psb*B operon in tobacco chloroplasts: Determination of the intron sites in *pet*B and *pet*D. *Mol. Gen. Genet.* **209,** 427–431.

Tanksley, S. D., and Pichersky, E. (1988). Organization and evolution of sequences in the plant nuclear genome. *In* "Plant Evolutionary Biology" (L. D. Gottlieb and S. K. Jain, eds.), pp. 55–83. Chapman and Hall, New York.

Timmis, J. N., and Scott, N. S. (1983). Sequence homology between spinach nuclear and chloroplast genomes. *Nature* **305,** 65–67.

Tingey, S. V., Tsai, F.-Y., Edwards, J. W., Walker, E. L., and Coruzzi, G. M. (1988). Chlo-

roplast and cytosolic glutamine synthetase are encoded by homologous nuclear genes which are differentially expressed *in vivo*. *J. Biol. Chem.* **263,** 9651–9657.

Tsai, C.-H., and Strauss, S. H. (1989). Dispersed repetitive sequences in the chloroplast genome of Douglas-fir. *Curr. Genet.* **16,** 211–218.

Turmel, M., Bellemare, G., and Lemieux, C. (1987). Physical mapping of differences between the chloroplast DNAs of the interfertile algae *Chlamydomonas eugametos* and *Chlamydomonas moewusii*. *Curr. Genet.* **11,** 543–552.

Turmel, M., Lemieux, B., and Lemiuex, C. (1988). The chloroplast genome of the green alga *Chlamydomonas moewusii:* Localization of protein-coding genes and transcriptionally active regions. *Mol. Gen. Genet.* **214,** 412–419.

Turmel, M., Boulanger, J., and Bergeron, A. (1989a). Nucleotide sequence of the chloroplast *petD* gene of *Chlamydomonas eugametos*. *Nucleic Acid Res.* **17,** 3593.

Turmel, M., Boulanger, J., and Lemieux, C. (1989b). Two group I introns with long internal open reading frames in the chloroplast *psbA* gene of *Chlamydomonas moewusii*. *Nucleic Acids Res.* **17,** 3875–3887.

Tymms, M. J., and Schweiger, H.-G. (1985). Tandemly repeated nonribosomal DNA sequences in the chloroplast genome of an *Acetabularia mediterranea* strain. *Proc. Natl. Acad. Sci. USA* **82,** 1706–1710.

Tymms, M. J., and Schweiger, H.-G. (1989). Significant differences between the chloroplast genomes of two *Acetabularia mediterranea* strains. *Mol. Gen. Genet.* **219,** 199–203.

Valentin, K., and Zetsche, K. (1989). The genes of both subunits of ribulose-1,5-bisphosphate carboxylase constitute an operon on the plastome of a red alga. *Curr. Genet.* **16,** 203–209.

Van Ee, J. H., Vos, Y. J., Bohnert, H. J., and Planta, R. J. (1982). Mapping of genes on the chloroplast DNA of *Spirodella oligorhiza*. *Plant Mol. Biol.* **1,** 117–131.

Vierling, E., Harris, L. M., and Chen, Q. (1989). The major low-molecular-weight heat shock protein in chloroplasts shows antigenic conservation among diverse higher plant species. *Mol. Cell. Biol.* **9,** 461–468.

Wakasugi, T., Ohme, M., Shinozaki, K., and Sugiura, M. (1986). Structures of tobacco chloroplast genes for $tRNA^{Ile}$(CAU), $tRNA^{Leu}$(CAA), $tRNA^{Cys}$(GCA), $tRNA^{Ser}$(UGA) and $tRNA^{Thr}$(GGU): A compilation of tRNA genes from tobacco chloroplasts. *Plant Mol. Biol.* **7,** 385–392.

Weeden, N. F. (1983). Evolutionary affinities of plant phosphoglucose isomerase and fructose-1,6-bisphosphatase isozymes. *Isozymes Curr. Top Biol. Med. Res.* **8,** 53–66.

Wessler, S. R. (1989). The splicing of maize transposable elements from pre-mRNA—A minireview. *Gene* **82,** 127–133.

Willey, D. L., and Gray, J. C. (1989). Two small open reading frames are co-transcribed with the pea chloroplast genes for the polypeptides of cytochrome *b*-559. *Curr. Genet.* **15,** 213–220.

Woese, C. R. (1987). Bacterial evolution. *Microbiol. Rev.* **51,** 221–271.

Woessner, J. P., Gillham, N. W., and Boynton, J. E. (1986). The sequence of the chloroplast *atp*B gene and its flanking regions in *Chlamydomonas reinhardtii*. *Gene* **44,** 17–28.

Wolfe, K. H. (1989a). Compilation of sequences of protein-coding genes in chloroplast DNA including cyanelle and cyanobacterial homologues. *Plant Mol. Biol. Reporter* **7,** 30–48.

Wolfe, K. H. (1989b). Rates of nucleotide substitution in higher plants and mammals. Ph.D. Thesis, University of Dublin.

Wolfe, K. H., and Sharp, P. M. (1988). Identification of functional open reading frames in chloroplast genomes. *Gene* **66,** 215–22.

Wolfe, K. H., Li, W.-H., and Sharp, P. M. (1987). Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad. Sci. USA* **84,** 9054–9058.

Wolfe, K. H., Sharp, P. M., and Li, W.-H. (1989). Rates of synonymous substitution in plant nuclear genes. *J. Mol. Evol.* **29,** 208–211.

Wolter, F. P., Fritz, C. C., Willmitzer, L., Schell, J., and Schreier, P. H. (1988). *rbc*S genes in *Solanum tuberosum:* Conservation of transit peptide and exon shuffling during evolution. *Proc. Natl. Acad. Sci. USA* **85,** 846–850.

Xu, M.-Q., Kathe, S. D., Goodrich-Blair, H., Nierzwicki-Bauer, S. A., and Shub, D. A. (1990). The prokaryotic origin of a chloroplast intron: A self-splicing group I intron in the gene for $tRNA^{Leu}_{UAA}$ of cyanobacteria. *Science* **250,** 1566–1570.

Yamada, T. (1983). Characterization of inverted repeat sequences and ribosomal RNA genes of chloroplast DNA from *Chlorella ellipsoidea. Curr. Genet.* **7,** 481–487.

Yamada, T., and Shimaji, M. (1987). Splitting of the ribosomal RNA operon on chloroplast DNA from *Chlorella ellipsoidea. Mol. Gen. Genet.* **208,** 377–383.

Yao, W. B., Meng, B.-Y., Tanaka, M., and Sugiura, M. (1989). An additional promoter within the protein-coding region of the *psb*D-*psb*C gene cluster in tobacco chloroplast DNA. *Nucleic Acids Res.* **17,** 9583–9591.

Yatskievych, G., Stein, D. B., and Gastony, G. J. (1988). Chloroplast DNA evolution and systematics of *Phanerophlebia* (Dryopteridaceae) and related fern genera. *Proc. Natl. Acad. Sci. USA* **85,** 2589–2593.

Zhou, D.-X., and Mache, R. (1989). Presence in the stroma of chloroplasts of a large pool of a ribosomal protein not structurally related to any *Escherichia coli* ribosomal protein. *Mol. Gen. Genet.* **219,** 204–208.

Zhou, D.-X., Massenet, O., Quigley, F., Marion, M. J., Monéger, F., Huber, P., and Mache, R. (1988). Characterization of a large inversion in the spinach chloroplast genome relative to *Marchantia:* A possible transposon-mediated origin. *Curr. Genet.* **13,** 433–439.

Zhou, D.-X., Quigley, F., Massenet, O., and Mache, R. (1989). Cotranscription of the S10- and *spc*-like operons in spinach chloroplasts and identification of three of their gene products. *Mol. Gen. Genet.* **216,** 439–445.

Zurawski, G., and Clegg, M. T. (1987). Evolution of higher-plant chloroplast DNA-encoded genes: Implications for structure-function and phylogenetic studies. *Annu. Rev. Plant Physiol.* **38,** 391–418.

Zurawski, G., Bottomley, W., and Whitfeld, P. R. (1984a). Junctions of the large single copy region and the inverted repeats in *Spinacia oleracea* and *Nicotiana debneyi* chloroplast DNA: Sequence of the genes for $tRNA^{HIS}$ and the ribosomal proteins S19 and L2. *Nucleic Acids Res.* **12,** 6547–6558.

Zurawski, G., Clegg, M. T., and Brown, A. H. D. (1984b). The nature of nucleotide sequence divergence between barley and maize chloroplast DNA. *Genetics* **106,** 735–749.

CHAPTER 3

Transmission of Plastid Genes

Nicholas W. Gillham

Department of Zoology
Duke University
Durham, North Carolina

John E. Boynton
Elizabeth H. Harris

Department of Botany
Duke University
Durham, North Carolina

I. INTRODUCTION

The discovery that variegated patterns of leaf pigmentation could be inherited in a non-Mendelian fashion was made independently by Correns and Baur in 1909, only 9 years after the rediscovery of Mendel's laws. These two scientists, working respectively with mutants of the four-o'clock (*Mirabilis jalapa*) and the common garden geranium (*Pelargonium zonale*), also defined the two basic patterns of chloroplast gene transmission recognized today in the angiosperms. Correns (1909) found that the inheritance of green, variegated, and white color patterns in seedlings of *M. jalapa* was maternal, as was subsequently found to be the case in most angiosperms. In contrast, Baur (1909) observed that reciprocal crosses of green and variegated plants of *P. zonale* showed a biparental but non-Mendelian pattern of inheritance, yielding green, white, and variegated plants. Both Correns and Baur also observed somatic segregation of the variegated leaf phenotypes, a second characteristic feature of plastid transmission.

The literature on plastid transmission has been summarized in a number of reviews (Connett, 1987; Gillham, 1978; Gillham *et al.*, 1985; Hagemann, 1976, 1979; Kirk and Tilney-Bassett, 1978; Sears, 1980, 1983; Smith, 1988; Tilney-Bassett, 1975; Whatley, 1982) and in the recent volume "Division and Segregation of Organelles," edited by Boffey and Lloyd (1988). The phenotypes of known plastome mutations have been tabulated based on their biochemical and electron microscopic characterization (Börner and Sears, 1986). DNA sequence alterations in some antibiotic-resistant, herbicide-resistant, and photosynthetically deficient plastid mutations have now been determined (e.g., Boynton *et al.*, 1991; Erickson *et al.*, 1985; Fromm *et al.*, 1987; Harris *et al.*, 1989; Winter and Herrmann, 1988).

In this review, we examine the organization, transmission, and segregation of chloroplast genomes in somatic cells and sexual crosses of higher plants and algae. We also review the little that is known about the genetic control of transmission of chloroplast genes and describe the transmission of plastids and plastid genomes following somatic cell fusion in higher plants.

II. CHLOROPLAST GENOME ORGANIZATION AND TRANSMISSION IN SOMATIC CELLS

A. Chloroplast Nucleoids

Chloroplasts are highly polyploid organelles, with DNA molecules of 85–200 kb organized into discrete, membrane-associated bodies reminiscent of bacterial chromosomes. These "nucleoids" appear as fibrous structures that stain with Feulgen or acridine orange (Ris and Plaut, 1962) and with the fluorochrome dye 4',6-diamidino-2-phenylindole (DAPI; Kuroiwa and Suzuki, 1980; Sellden and Leech, 1981). In most higher plants and many green algae, nucleoids are associated with thylakoid membranes scattered through the chloroplast (Hansmann *et al.*, 1985; Kuroiwa *et al.*, 1981; Lindbeck and Rose, 1987; Possingham and Lawrence, 1983); however, peripheral nucleoids associated with the inner envelope membrane are also found, particularly in monocotyledons (Hashimoto, 1985). In the various algal groups, nucleoids can be dispersed, peripheral, central, or ring-shaped (see Kuroiwa *et al.*, 1981; Zachleder and Cepák, 1987), and nucleoid arrangement can also change during the cell cycle in algae (Zachleder *et al.*, 1989) and in the course of plastid development in plants (Hashimoto and Possingham, 1989a; Miyamura *et al.*, 1986). As in bacterial cells (Firshein, 1989), membrane attachment of nucleoids is thought to facilitate their partitioning at cell division and may also play a role in replication of chloroplast DNA (Lindbeck *et al.*, 1987; Rose, 1988).

Transcriptionally active complexes containing complete chloroplast genomes and a specific set of proteins have been isolated from chloroplasts of spinach (Briat *et al.*, 1982) and mustard (Reiss and Link, 1985). Four DNA-binding proteins have been identified in proplastid nucleoids of tobacco (Nemoto *et al.*, 1988, 1989) and chloroplast DNA–protein complexes have been characterized microscopically in *Chlamydomonas* (Chiang *et al.*, 1981).

Chloroplast genome number per cell can vary widely within individual plants and among species (Boffey and Leech, 1982). Proplastids in *Brassica* have been reported to have only a single small nucleoid (Kuroiwa *et al.*, 1981), whereas mature chloroplasts of most plant leaves typically have 10–20 nucleoids, which may contain as few as 2 or 3 (Hansmann *et al.*, 1985; Scott and Possingham, 1980) to as many as 24 (Boffey and Leech, 1982; Lindbeck *et al.*, 1989) chloroplast DNA molecules.

B. Plastid Division and Segregation

The process of plastid division has been reviewed recently (Kuroiwa, 1989; Leech and Pyke, 1988; Oross and Possingham, 1989; Whatley, 1988). In typical plant cells, elongation of the plastid is followed by progressive constriction at right angles to the long axis, to form a dumbbell shape. Thylakoids and nucleoids are partitioned to the two ends shortly before final separation into two daughter plastids (see Hashimoto and Possingham, 1989b). In wheat leaves, actively dividing cells adjacent to the basal meristem have about 7% of their total plastid profiles in the dumbbell configuration, and the half-life of these profiles averages 20 min (Leech and Pyke, 1988).

Nucleoids typically are partitioned more or less equally among daughter chloroplasts, and the plastids themselves also segregate at mitosis, but no mechanism analogous to the mitotic spindle is known that ensures transmission of precisely the same plastid complement to each daughter cell. Thus, both events can be viewed as stochastic processes (Birky, 1983a). The grouping of chloroplast DNA molecules into nucleoids may account for the observation that segregation of chloroplast genomes is often more rapid than would be predicted from the average number of genomes per cell (see Forster *et al.*, 1980; Lee and Haughn, 1980). However, in certain cases, heteroplasmic chloroplasts or heteroplasmic plastid types can persist for prolonged periods (e.g., Bolen *et al.*, 1980; Boynton *et al.*, 1990; Lax *et al.*, 1987; Sager, 1972; Spreitzer and Chastain, 1987). Models for segregation must take these facts into account, as well as the problem of unequal numbers of DNA molecules per nucleoid and variability in numbers of nucleoids per chloroplast (Adams, 1978; Birky, 1983a,b; Birky *et al.*, 1981; VanWinkle-Swift, 1980). The effects of mitotic segregation of organelle genomes, genetic drift, and other factors on organelle gene diversity have been modeled using population genetic theory by Birky *et al.* (1989).

C. Developmental Changes in Plastid Number and DNA Content

Several studies have followed the changes in plastid number and DNA content during leaf development (for a detailed discussion, see Possingham *et al.*, 1988). For example, in young, dividing mesophyll cells of spinach, chloroplast DNA synthesis and chloroplast number appear to keep pace with cell division, maintaining 10–15 chloroplasts per cell, each with an average of about 50 chloroplast genomes (Lawrence and Possingham, 1986). At the onset of leaf expansion, chloroplast DNA

synthesis exceeds the rate of chloroplast division, increasing the genome number to as many as 12,500 copies per cell. Later, chloroplast DNA synthesis declines while chloroplasts continue to divide, thus lowering the DNA content per plastid.

Chloroplast DNA copy number appears to be correlated with nuclear DNA content (Bowman, 1986; Butterfass, 1988; Ellis and Leech, 1985; Ellis *et al.*, 1983; Whiteway and Lee, 1977), but this may be an indirect effect rather than an absolute relationship between genome numbers. In diploid, tetraploid, and hexaploid wheat species, the mean number of chloroplasts per cell increases with ploidy level, but a better correlation is found between chloroplast number and cell size, regardless of ploidy (Pyke and Leech, 1987).

III. TRANSMISSION OF CHLOROPLASTS AND THEIR GENOMES IN CROSSES

A. General Mechanisms

The transmission of plastids and their genomes in crosses has been monitored by light and electron microscopic studies of gametic differentiation and fertilization, by genetic analysis of the inheritance of chloroplast (plastome) mutations, and by examining the transmission of chloroplast DNA restriction fragment length polymorphisms (RFLPs). Several distinct mechanisms result in uniparental-maternal inheritance of chloroplast DNA in algae and higher plants. These include exclusion of plastids from one parent in the gamete, destruction of plastids from one parent in the zygote, and plastid fusion in the zygote followed by destruction of plastid DNA from one parent (see reviews by Dickinson and Li, 1988; Sager and Grabowy, 1985; Sears, 1980; Whatley, 1982). Destruction of otherwise functional plastids should be distinguished from degeneration of mutant plastids incapable of self-replication (see Vaughn, 1981) and from loss of plastids in generative or somatic cells as a result of nuclear mutations, a topic beyond the scope of this review.

B. Specific Examples in Algae and Lower Plants

In the green algae (Chlorophyta), gametes from both parents generally contribute chloroplasts to the zygote, but chloroplast fusion has been

reported in only a few cases. Loss of chloroplast nucleoids from one parent after chloroplast fusion occurs in isogamous species of the unicellular green alga *Chlamydomonas* (see below), in the coenocytic green alga *Dictyosphaeria cavernosa*, and the giant unicellular green alga *Acetabularia calyculus* (Kuroiwa *et al.*, 1985a). However, only in the case of *Chlamydomonas* has the disappearance of nucleoids been correlated with chloroplast DNA destruction. Both chloroplast and mitochondrial nucleoids disappear before mating in male gametes of the green alga *Bryopsis maxima* (Kuroiwa and Hori, 1986; Ogawa, 1988; Saito *et al.*, 1989). In many oogamous algal species, and in some isogamous species in which there is no overt morphological distinction between "male" and "female" gametes, the plastids contributed by the male parent disintegrate in the egg cell following gametic fusion. Prezygotic plastid elimination has been described for the spermatozoids of certain oogamous Xanthophyte and brown algae (see Whatley, 1982).

Mature spermatozoids of many lower land plants have their organelles oriented in a characteristic fashion, with the plastids located at the posterior end. Motile spermatozoids of lower plants typically discard a cytoplasmic vesicle including the plastid(s) before the male gamete reaches the egg (see Whatley, 1982). In spermatogenesis of the fern *Pteris vittata* L., chloroplast volume and chloroplast DNA content were observed to decrease during repeated divisions of the spermatocytes and the chloroplasts then differentiated into amyloplasts and lost their DNA during sperm maturation (Kuroiwa *et al.*, 1988).

C. Specific Examples in Seed Plants

Among the seed plants (Spermatophytina), plastid transmission has been studied in both gymnosperms and angiosperms. Conifers are exceptional among all plants in exhibiting a largely uniparental-paternal pattern of plastid transmission as shown by genetic (Ohba *et al.*, 1971), light and electron microscopic (see Whatley, 1982), and molecular studies (see review by Neale and Sederoff, 1988). In conifer species studied to date, paternal plastids have been shown to enter the egg cell and maternal plastids are often observed to degenerate (Whatley, 1982). Paternal transmission of chloroplast DNA has been demonstrated through the use of RFLPs in a number of conifers (Table I), including Douglas fir (Neale *et al.*, 1986), several pines (Neale and Sederoff, 1989; Wagner *et al.*, 1987), larch (Szmidt *et al.*, 1987), and spruce (Szmidt *et al.*, 1988). Among six hybrids between two larch species, three showed strictly paternal inheritance of RFLP markers, one maternal, and two biparental

(Szmidt *et al.*, 1987). Apparent chloroplast DNA recombinants were also found by analysis of chloroplast DNA from individual trees in a region where *Pinus banksiana* and *Pinus contorta* are known to hybridize naturally (Govindaraju *et al.*, 1988, 1989). In crosses of *Pinus rigida* to *Pinus taeda*, chloroplast DNA is inherited paternally, whereas mitochondrial DNA is inherited maternally (Neale and Sederoff, 1989), but in redwood (*Sequoia sempervirens*), both chloroplast and mitochondrial DNA are transmitted by the paternal parent (Neale *et al.*, 1989). Although data on transmission of chloroplast DNA are lacking for more primitive gymnosperms, light and ultrastructural studies summarized by Whatley (1982) suggest that most of the cytoplasm from the large spermatozoids of the cycad *Dioon* may be discarded just inside the egg cell, whereas in *Gingko* this material is left outside the egg cell.

In angiosperms, the pattern of plastid transmission has been examined in crosses of >50 genera of dicotyledons and monocotyledons (Smith, 1988). Such studies show that plastid transmission is most often maternal, but biparental inheritance is not uncommon and has now been demonstrated in about 20 genera (see Table I; Sears, 1980; Smith, 1988; Tilney-Bassett, 1975). In 11 of these genera, >5% of the progeny in crosses of certain species regularly contain at least some paternally derived plastids (Smith, 1988). Corriveau and Coleman (1988) have identified a number of additional species in which plastid DNA can be found in generative or sperm cells of pollen, suggesting that these have the potential for biparental inheritance (Table I).

Maternal inheritance of plastids in angiosperms normally involves elimination of plastids from the pollen parent at some time between pollen grain (microspore) formation and fertilization of the egg by one of the two sperm nuclei. The haploid, uninucleate microspore divides mitotically to yield a large vegetative cell containing most of the cytoplasm and organelles and a small generative cell whose nucleus then divides mitotically a second time. In the pollen tube, the two resulting sperm cells are surrounded by the cytoplasm of the vegetative cell. At fertilization, one of the sperm nuclei fuses with the egg nucleus, forming a diploid zygote from which the embryo will develop, while the other fuses with the two polar nuclei to yield a triploid cell from which the endosperm of the seed develops.

On the basis of light and electron microscopic observations, Hagemann (1979, 1983) recognized four patterns of plastid transmission in angiosperms (Table II). In the *Lycopersicon* type, exemplified by the tomato (*Lycopersicon esculentum*), all microspore plastids segregate to the vegetative cell, and only chloroplasts from the maternal parent are inherited by the progeny. Nevertheless, transmission of paternal plastids

TABLE I

Genera in which Biparental or Paternal Inheritance of Plastids Has Been Implicated[a]

Species	Evidence for Biparental or Paternal Inheritance			Reference
	Plastome mutant	cpDNA RFLP	DAPI stain of cpDNA	
Chlorophyta				
Chlamydomonas reinhardtii,[b] *C. smithii*[b]	+	+	+	see Harris (1989)
C. eugametos,[b] *C. moewusii*[b]	+	+	+	Coleman and Maguire (1983), Lemieux *et al.* (1980, 1981, 1984)
Pteridophyta				
Scolopendrium vulgare	+			see Kirk and Tilney-Bassett (1978)
Gymnospermae				
Taxodiaceae				
Cryptomeria japonica[b]	+			see Kirk and Tilney-Bassett (1978), Smith (1988)
Sequoia sempervirens[b]		+		Neale *et al.* (1989)
Pinaceae				
Pinus banksiana × *P. contorta*[b]		+		Wagner *et al.* (1987)
Pinus rigida × *P. taeda*[b]		+		Neale and Sederoff (1989)
Pseudotsuga menziesii[b]		+		Neale *et al.* (1986)
Larix decidua × *L. leptolepis*[b]		+		Szmidt *et al.* (1987)
Picea glauca × *P. omorika*[b]		+		Szmidt *et al.* (1988)
Angiospermae				
Aizoaceae				
Fenestraria rhopalophylla			+	Corriveau and Coleman (1988)
Boraginaceae				
Borrago officinalis[b]	+			see Kirk and Tilney-Bassett (1978), Smith (1988)
Cactaceae				
Rhipsalidopsis gaertneri			+	Corriveau and Coleman (1988)
Zygocactus truncatus			+	Corriveau and Coleman (1988)
Campanulaceae				
Campanula alliariafolia			+	Corriveau and Coleman (1988)
C. carpatica			+	Corriveau and Coleman (1988)

C. rapunculoides			+	Corriveau and Coleman (1988)
Lobelia syphilitica			+	Corriveau and Coleman (1988)
Platycodon grandiflorum			+	Corriveau and Coleman (1988)
Caprifoliaceae				
Linnaea borealis			+	Corriveau and Coleman (1988)
Lonicera japonica			+	Corriveau and Coleman (1988)
Caryophyllaceae				
Silene otites × *S. pseudotites*	+			see Kirk and Tilney-Bassett (1978), Smith (1988)
Clusiaceae				
Hypericum perforatum[b]	+		+	Corriveau and Coleman (1988); see Kirk and Tilney-Bassett (1978), Smith (1988)
Convolvulaceae				
Ipomoea nil			+	Corriveau and Coleman (1988)
Droseraceae				
Dionaea muscipula			+	Corriveau and Coleman (1988)
Drosera capillaris			+	Corriveau and Coleman (1988)
Ericaceae				
Rhododendron maximum			+	Corriveau and Coleman (1988)
Rhododendron japonicum × 11 species	+			see Smith (1988)
Fabaceae				
Acacia decurrens × *A. mearnsii*[b]	+			see Smith (1988)
Acacia mearnsii × *A. baileyana*[b]	+			see Smith (1988)
Cicer arietinum			+	Corriveau and Coleman (1988)
Lathyrus japonicus			+	Corriveau and Coleman (1988)
L. odoratus			+	Corriveau and Coleman (1988)
Medicago sativa[b]	+	+	+	Corriveau and Coleman (1988), Lee *et al.*, (1988), see Smith (1988)
M. truncatula[b]	+			see Smith (1988)
Melilotus alba			+	Corriveau and Coleman (1988)
Phaseolus vulgaris	+		−[c]	Corriveau and Coleman (1988)
Pisum sativum			+	Corriveau and Coleman (1988)
Wisteria sinensis			+	Corriveau and Coleman (1988)

(continues)

TABLE I
Continued

Species	Evidence for Biparental or Paternal Inheritance: Plastome mutant	cpDNA RFLP	DAPI stain of cpDNA	Reference
Geraniaceae				
Geranium maculatum[b]	+		+	Corriveau and Coleman (1988)
G. pratense			+	Corriveau and Coleman (1988)
G. sanguineum			+	Corriveau and Coleman (1988)
G. bohemicum × *G. deprehensium*	+			see Smith (1988)
Pelargonium peltatum			+	Corriveau and Coleman (1988)
Pelargonium × *Hortorum (zonale)*[b]	+	+	+	Corriveau and Coleman (1988); Metzlaff *et al.*, (1981); see Kirk and Tilney-Bassett (1978); Smith (1988)
Lamiaceae				
Nepeta cataria	+		−	Corriveau and Coleman (1988)
Liliaceae				
Chlorophytum comosum			+	Corriveau and Coleman (1988)
C. elatum[b]	+		+	Corriveau and Coleman (1988), see Kirk and Tilney-Bassett (1978), Smith (1988)
C. comosum × *C. elatum*	+			see Kirk and Tilney-Bassett (1978), Smith (1988)
Onagraceae				
Epilobium hirsutum	+			see Smith (1988)
E. augustifolium	+		-	Corriveau and Coleman (1988)
E. watsonii × *E. montanum*		+		Schmitz and Kowallik (1987), see Smith (1988)
Oenothera[b] (30 species)	+	+	+	Corriveau and Coleman (1988), see Kirk and Tilney-Bassett (1978), Smith (1988)
Passifloraceae				
Passiflora edulis			+	Corriveau and Coleman (1988)

Plumbaginaceae				
Plumbago auriculata			+	Corriveau and Coleman (1988)
P. capensis			+	Corriveau and Coleman (1988)
P. larpentae			+	Corriveau and Coleman (1988)
P. zeylanica			+	Corriveau and Coleman (1988)
Poaceae				
Pennisetum americanum[b]	+			see Smith (1988)
Secale cereale[b]	+		−	Corriveau and Coleman (1988), see Kirk and Tilney-Bassett (1978), Smith (1988)
Hordeum vulgare × *S. cereale*		+		Soliman *et al.* (1987)
Polygonaceae				
Fagopyrum esculentum[b]	+		−	Corriveau and Coleman (1988), see Kirk and Tilney-Bassett (1978), Smith (1988)
Scrophulariaceae				
Antirrhinum majus	+		−	Corriveau and Coleman (1988), see Kirk and Tilney-Bassett (1978), Smith (1988)
Solanaceae				
Browallia speciosa	+			see Kirk and Tilney-Bassett (1978), Smith (1988)
Solanum tuberosum	+		−	Corriveau and Coleman (1988), see Kirk and Tilney-Bassett (1978), Smith (1988)
S. nigrum	+			see Smith (1988)
Nicotiana plumbaginifolia × *N. tabacum*	+			Medgyesy *et al.* (1986)
Petunia hybrida		+		Cornu and Dulieu (1988)
Strelitziaceae				
Strelitzia reginae			+	Corriveau and Coleman (1988)

[a] Based on genetic data using plastome mutants, inheritance of restriction fragment length polymorphisms (RFLPs) in chloroplast DNA (cpDNA), and 4′,6-diamidino-2-phenylindole (DAPI) staining of chloroplast DNA in pollen generative or sperm cells.

[b] Instances in which biparental or paternal inheritance of plastids is a regular or frequent occurrence.

[c] − indicates no DAPI staining of plastid DNA in pollen was observed, despite genetic evidence for biparental inheritance.

[Adapted from Corriveau and Coleman, 1988, and Smith, 1988.]

TABLE II
Patterns of Plastid Transmission in Male Gametophytes, Gametes, and Crosses in Angiosperms

Transmission Pattern	Typical Species	Transmission of Male Plastids in Crosses
Lycopersicon type (male plastids excluded from generative cell)	*Mirabilis jalapa*	no
	Beta vulgaris	no
	Gossypium hirsutum	no
	Antirrhinum majus	rare
	Petunia hybrida	no[a]
	Lycopersicon esculentum	no
Solanum type (generative cell plastids degenerate)	*Solanum tuberosum*	rare
Triticum type (sperm cell cytoplasm and plastids left outside egg)	*Triticum aestivum*	no
	Hordeum vulgare	no
Pelargonium type (male plastids transmitted to embryo)	*Pelargonium zonale*	yes
	Oenothera spp.	yes

[a] Cornu and Dulieu (1988) report a single mutant line that consistently shows a low frequency of biparental plastid transmission in crosses.

[Data from Hagemann (1979), Hagemann and Schröder (1985), and Sears (1980).]

at low frequency has been documented in both *Nicotiana* (Medgyesy *et al.*, 1986) and *Petunia*, where nuclear genes have been shown to play a role in permitting paternal plastid transmission in up to 2% of the progeny (Cornu and Dulieu, 1988). In the *Solanum* (potato) type, the first partitioning of the cytoplasm during pollen mitosis is more equal, with the generative cell as well as the vegetative cell receiving some plastids. However, in the course of further pollen development, the generative cell plastids are lost (or eliminated), so the sperm cells do not contain plastids. Plants exhibiting the *Solanum* pattern normally transmit their plastids maternally, but occasionally a plant showing biparental plastid transmission is observed (Table I). In the *Triticum* (wheat) type, which is so far restricted to the grasses, plastids are found in the generative and sperm cells. However, when the sperm cell nucleus enters the egg cell, enucleated cytoplasmic bodies containing mitochondria and plastids are left outside (Mogensen, 1988; Mogensen and Rusche, 1985). In the *Pelargonium* pattern of plastid transmission, found in geranium and evening primrose, distribution of plastids to the generative and vegetative cells during the first pollen mitosis results in biparental plastid transmission. Among biparental progeny of *Oenothera*, maternally derived plastids

predominate, whereas in alfalfa (*Medicago sativa*) paternally derived plastids predominate, suggesting the possibility of additional mechanisms influencing plastid inheritance (Lee *et al.*, 1988, 1989; Smith, 1989). Spatial distribution of the plastids in the zygote preceding its asymmetric division into a suspensor and a terminal cell may play a critical role, with paternal plastids in alfalfa being more favorably situated for entry into the terminal cell (Tilney-Bassett and Almouslem, 1989).

Plastid transmission patterns can vary greatly within a single taxon. For example, in the Liliaceae, plastids in the generative cell of *Lilium martagon* appear not to degenerate and may be transmitted to the zygote (Schröder, 1984), whereas young generative cells of *Fritillaria imperialis, Convallaria majalis,* and *Hosta ventricosa* may receive some plastids that subsequently seem to degenerate (Schröder and Hagemann, 1985; Schröder, 1985a, 1986a,b; Vaughn *et al.*, 1980). Plastids are absent from the generative cells of *Gasteria verrucosa* and *Fritillaria meleagris* (Schröder, 1985a,b). In *G. verrucosa,* this may result because plastids cluster at the proximal pole during the first pollen mitosis, assuring that most, if not all, of them remain in the vegetative cell. This same polarization of plastids is seen in *Impatiens* and is probably mediated by the cytoskeleton (Van Went, 1984). In two other Liliaceous species, *Aloë secundiflora* and *Aloë jucunda,* most plastids remain near the center of the microspore close to the vegetative nucleus at the time of the first pollen mitosis but, nevertheless, are normally absent from the mature generative cell (Schröder and Hagemann, 1986).

D. Deletion of Chloroplast DNA Sequences during Pollen Formation

In cultivated cereals such as barley and wheat, Day and Ellis (1984, 1985) have suggested that the evolution of maternal inheritance may be related to the high frequency of defective chloroplast genomes observed in haploid plants derived from cultured anthers. A high percentage of the regenerated plants are albino seedling lethals, containing heterogeneous collections of plastid genomes with substantial fractions of their sequences deleted. Three such deleted molecules of wheat had lost 71, 67, and 42% of their respective chloroplast genomes, but all retained regions assumed to contain origins of replication (Day and Ellis, 1984, 1985). The deleted chloroplast genomes were hypothesized to arise from intramolecular recombination between short, direct repeats (see Bowman and Dyer, 1986). Defective chloroplast genomes with a linear, hairpin structure isolated from barley were proposed to arise by a complex strand

switching during chloroplast DNA replication (Ellis and Day, 1986). In contrast, *Nicotiana*, whose chloroplast genome lacks the short dispersed repeat sequences characteristic of wheat (Howe *et al.*, 1988; Shinozaki *et al.*, 1986), does not produce albino plants in anther cultures unless the anthers are derived from white tissue (Nilsson-Tilgren and von Wettstein-Knowles, 1970). Thus, maternal transmission of the chloroplast genome of *Nicotiana* is not correlated with instability of plastid genomes during pollen development.

IV. GENETIC CONTROL OF PLASTID TRANSMISSION

A. Use of Plastome Mutations to Investigate Plastid Transmission

The genetic control of plastid transmission has been investigated most extensively in species of *Chlamydomonas*, where the plastid genome exhibits a uniparental-maternal pattern of inheritance in >90% of meiotic zygotes, and in two higher plant genera, *Pelargonium* and *Oenothera*, where the plastids themselves are often transmitted biparentally. Most plastome mutations available in higher plants cause alterations in leaf and stem pigmentation that are not very useful for studying rare cases of biparental or paternal chloroplast transmission in species where chloroplast genomes are normally maternally transmitted. In contrast, easily scored and selected chloroplast mutations to antibiotic resistance in *Chlamydomonas reinhardtii*, which have been used extensively to study the transmission of chloroplast genes (reviewed by Gillham, 1978, Harris, 1989; Sager, 1972), facilitate identification of rare exceptional zygotes transmitting the chloroplast genome of the mt^- parent. This approach should now be feasible in higher plants because chloroplast mutations to antibiotic resistance can be isolated in tissue culture and resistant plants regenerated from these calli (e.g., Fluhr *et al.*, 1985). Rare paternal transmission of plastids in *Nicotiana* has been demonstrated by selecting for antibiotic resistance among offspring of crosses between resistant male and sensitive female parents in callus tissue derived from seedlings (Medgyesy *et al.*, 1986). In the future, the use of antibiotic-resistant plastome mutations should make possible investigation of the genetic determinants involved in uniparental-maternal plastid transmission in higher plants.

B. Transmission of Chloroplast DNA in *Chlamydomonas*

Most studies of chloroplast genome transmission in *Chlamydomonas* have involved the heterothallic species *C. reinhardtii,* but important results have also been obtained with the distantly related heterothallic species pair *Chlamydomonas moewusii–Chlamydomonas eugametos* and with the homothallic species *Chlamydomonas monoica.* Many of the experimental results for *C. reinhardtii* have been reviewed elsewhere (see Gillham, 1978; Harris, 1989; Sager, 1972, 1977; Sager and Grabowy, 1985) and will be summarized only briefly here. In *C. reinhardtii,* gametes of opposite mating type fuse pairwise to form diploid zygotes, 95% of which form thick-walled meiotic zygospores and 5% of which divide mitotically as stable vegetative diploids (Ebersold, 1967). Most (90–99%) of the meiotic zygotes transmit chloroplast antibiotic resistance markers only from the mt^+ parent. These are referred to as maternal or, more properly, UP^+ (for uniparental from mating type plus) zygotes (Harris, 1989). The remaining zygotes are exceptional and transmit the chloroplast genome of both parents (biparental zygotes) or rarely of only the mt^- parent (paternal or UP^- zygotes). The frequency of exceptional zygotes can be greatly increased by UV irradiation of mt^+ gametes prior to mating (Gillham *et al.*, 1974; Sager and Ramanis, 1967), by growth of the mt^+ parent for several generations in 5-fluorodeoxyuridine (FdUrd) before mating, which substantially reduces the chloroplast genome number (Wurtz *et al.*, 1977), or by nuclear mating type-linked *mat-3* mutations (Gillham *et al.*, 1987). Somatic segregation and recombination of chloroplast markers from both parents occur among the progeny of biparental zygotes. Recombination of chloroplast genes in *Chlamydomonas* is reviewed by Boynton *et al.* (1991). Pedigree analysis shows that the segregation process is extremely rapid, being 50–70% completed for any given marker by the end of the first postmeiotic mitotic division (Forster *et al.*, 1980; Sager, 1977).

Despite the fact that the uniparental transmission of chloroplast DNA in *Chlamydomonas* zygotes has been unequivocally demonstrated (Boynton *et al.*, 1987; Grant *et al.*, 1980; Lemieux *et al.*, 1980; Mets, 1980), the timing and molecular mechanisms responsible are still in doubt (see Gillham, 1978; Harris, 1989; Sager, 1977). The single chloroplasts of the mt^+ and mt^- parents fuse between 3 and 7 hr after mating, subsequent to the time of nuclear fusion (Cavalier-Smith, 1970). Chloroplast DNA from the mt^+ parent undergoes a shift to a lighter density within the first 6 hr of mating, possibly due to methylation, while that from the mt^- parent disappears (Sager, 1977; Sager and Lane, 1972). However, genetic com-

plementation between nonphotosynthetic mutations in different chloroplast genes is first observed about 24 hr after mating, suggesting that some chloroplast DNA, or stable messenger RNA, from both parents may persist in most zygospores for as long as 2 days (Bennoun *et al.*, 1980).

DAPI staining reveals that chloroplast nucleoids from the mt^- parent vanish within the first 40–50 min after mating, prior to chloroplast fusion (Coleman, 1984; Kuroiwa *et al.*, 1982). After about 10 hr, the 8–10 nucleoids of the mt^+ parent coalesce to form a single, large DAPI-staining structure adjacent to the pyrenoid of the zygotic chloroplast (Kuroiwa *et al.*, 1982). Disappearance of the mt^- nucleoids can be blocked by UV irradiation of mt^+ but not mt^- gametes and by incubation of the young zygotes in various inhibitors including actinomycin D and cycloheximide (Kuroiwa, 1985; Kuroiwa *et al.*, 1983a,b, 1985b). These inhibitors are effective only when added shortly after mating, implying that proteins synthesized in the young zygote may be responsible for the loss of mt^- nucleoids.

To explain the uniparental transmission of chloroplast genes by the mt^+ parent, Sager and her colleagues proposed a restriction-modification model analogous to processes known in *Escherichia coli* and other bacteria (see Sager and Kitchin, 1975). Chloroplast DNA contributed by the mt^+ gamete would be modified to protect it from digestion by a restriction enzyme that becomes activated in the zygote to destroy mt^- chloroplast DNA (Sager and Grabowy, 1985; Sager and Ramanis, 1973; Sager *et al.*, 1984). In support of this model, they reported that cytosine residues in chloroplast DNA of mt^+ but not mt^- gametes became extensively methylated prior to mating, and that a unique methyl transferase activity occurred in mt^+ but not in mt^- gametes.

However, observations of other laboratories have cast doubt on whether or not massive methylation of chloroplast DNA seen in mt^+ gametes is causally related to transmission of mt^+ chloroplast DNA. The nuclear mutation *me-1*, which constitutively methylates >35% of the cytosine residues in chloroplast DNA of either mating type, does not alter the pattern of chloroplast gene transmission (Bolen *et al.*, 1982). Exposure of mt^+ cells undergoing gametogenesis to the methylation inhibitors 5-azacytidine (azaC) or L-ethionine greatly reduces chloroplast DNA methylation, but these hypomethylated mt^+ gametes transmit their chloroplast genomes in the normal, uniparental fashion (Feng and Chiang, 1984).

Sager and her colleagues have argued that these results were inconclusive, because they observed an additional round of methylation in

mt$^+$ but not *mt*$^-$ gametes of the *me-1* mutant (Sager and Grabowy, 1983; Sager *et al.*, 1984) and found that treatment of *mt*$^+$ gametes with higher concentrations of azaC reduces the frequency of maternal zygotes (Sager and Grabowy, 1985). They proposed that azaC as well as other treatments (e.g., UV irradiation and growth in FdUrd) of the *mt*$^+$ parent that increase the frequency of exceptional zygotes prevent activation of the postulated restriction enzyme for chloroplast DNA destruction. This would permit transmission of unmodified chloroplast DNA to the progeny. All the methylation observed so far in chloroplast DNA appears to be nonspecific, however, and no evidence indicates a specific restriction endonuclease involved in elimination of chloroplast DNA from the *mt*$^-$ parent in the zygote. In summary, the restriction-modification hypothesis for uniparental inheritance of chloroplast DNA is not well supported by existing experimental results. The extensive methylation of chloroplast DNA observed in *mt*$^+$ gametes and zygotes quite possibly serves some other function related to the sexual cycle in *C. reinhardtii*.

Based on studies of nucleoid destruction and cellular nucleases in *C. reinhardtii*, Kuroiwa has proposed an alternative hypothesis involving a degradative nuclease for elimination of *mt*$^-$ chloroplast DNA (Kuroiwa, 1985; Kuroiwa *et al.*, 1982, 1983a,b). Alteration of "domains" surrounding *mt*$^+$ chloroplast DNA during gametogenesis is postulated to protect it from zygotic destruction. This model is reminiscent of the idea suggested by Gillham *et al.* (1974) that the zygotic chloroplast could contain a fixed number of "attachment sites" for transmission of chloroplast DNA that are occupied preferentially by chloroplast genomes of the *mt*$^+$ gamete. UV irradiation was hypothesized to increase biparental transmission of chloroplast genes by rendering these sites accessible to *mt*$^-$ genomes (see also Adams, 1978).

Consistent with this model, Kuroiwa *et al.* (1985b) reported that DAPI-stained nucleoids of the *mt*$^-$ parent seemed to be degraded from their periphery within 30 min after zygote formation, becoming smaller in size and disappearing completely within 40–60 min. Gametes and young zygotes were found to possess a Ca^{2+}-requiring enzyme fraction (nuclease C) that degrades rather than restricts DNA (Ogawa and Kuroiwa, 1985a,b, 1986), but evidence associating any of the six polypeptides observed with UP$^+$ inheritance is lacking. Recently, Nakamura *et al.* (1988) have identified a constellation of six polypeptides that label with ^{35}S-methionine in young zygotes. Synthesis of these proteins was inhibited by UV irradiation of *mt*$^+$ but not *mt*$^-$ gametes prior to mating, suggesting that one or more of these polypeptides may be required for uniparental inheritance. Synthesis of these polypeptides was not blocked

by growth and gametogenesis of mt^+ cells on 0.5 m*M* FdUrd, however, although 70% of the mt^+ gametes lacked chloroplast nucleoids and destruction of mt^- nucleoids occurred (Nakamura and Kuroiwa, 1989).

Of the rare (ca. 5%) zygotes of *C. reinhardtii* that form vegetative diploids rather than undergoing meiosis, 50–90% exhibit biparental inheritance of chloroplast genomes, and most of the remainder show UP^+ transmission (Boynton *et al.*, 1987; Gillham, 1963; Matagne, 1981). Because chloroplast genes recombine in vegetative diploids, fusion of the respective organelles must occur (Boynton *et al.*, 1976, 1987; Galloway and Holden, 1985; Gillham, 1969). The frequency of diploids that transmit chloroplast markers uniparentally from the mt^+ parent can be increased by delaying the first mitotic division for 40 hr either by nitrogen starvation or incubation in the dark (VanWinkle-Swift, 1978). This suggests that processes degrading chloroplast DNA from the mt^- parent become activated over time in the undivided diploid cells. The high frequencies of biparental meiotic zygotes that arise following UV irradiation of the mt^+ parent (Sears, 1980) or when a mating type-linked *mat-3* mutation is used in a cross (Gillham *et al.*, 1987) also decline when germination (which can first be induced 2 days after mating) is delayed. In both cases, biparental zygotes are converted largely to UP^+ zygotes. A bias toward UP^+ diploids is not seen in polyethylene glycol (PEG)-mediated fusions between vegetative cells of opposite mating type (Matagne, 1981; Matagne and Hermesse, 1980), but there is disagreement over whether or not biased transmission occurs in PEG-fused gametes (Matagne, 1981; Matagne and Schaus, 1985; Matsuda *et al.*, 1983, 1988).

The models of both Sager and Ramanis (1973) and Kuroiwa (1985) vest genetic control of uniparental transmission of chloroplast genes in *C. reinhardtii* largely in the mt^+ parent. Mating type in *C. reinhardtii* constitutes a region on nuclear linkage group VI, in which crossing-over appears to be suppressed. Several complementing mutations in genes with no obvious relationship to mating (i.e., nicotinamide, acetate, and thiamine auxotrophs) fail to recombine with mating type (Gillham, 1969) and with certain mutations affecting the mating process (for a review, see Goodenough and Ferris, 1987).

Goodenough and her colleagues (Galloway and Goodenough, 1985; Goodenough and Ferris, 1987) have proposed a genetic model that explains the behavior of mating type-linked mutations affecting the mating process and also suggests how genes linked to mating type might control uniparental inheritance. We have modified this model here (Fig. 1) to emphasize the roles played by genes involved in the uniparental transmission of chloroplast and mitochondrial genomes by opposite mating types (Boynton *et al.*, 1987). The model proposes that gametogenesis

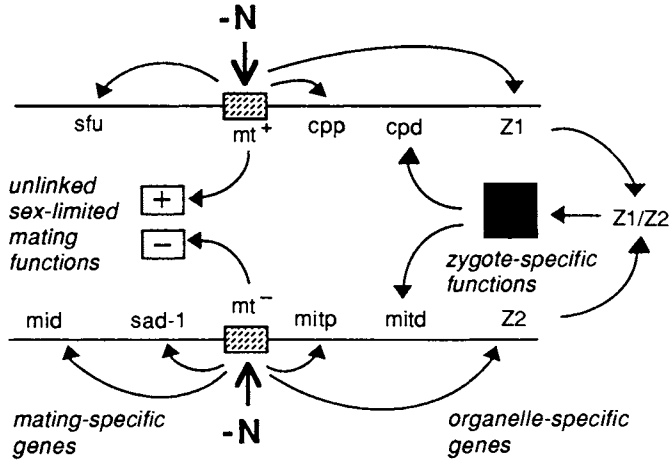


FIG. 1. Model for control of chloroplast and mitochondrial DNA transmission by the mating type locus in *Chlamydomonas reinhardtii*. See text for description. [Modified from Galloway and Goodenough, 1985.]

triggered by nitrogen starvation results in the expression of specific genes whose products are required for protection of mt^+ chloroplast DNA (*cpp*) and mt^- mitochondrial DNA (*mitp*). The *cpp* and *mitp* genes are linked, respectively, to the mt^+ and mt^- alleles. Upon fusion of cells of opposite mating type, products encoded by the postulated genes *Z1* (linked to mt^+) and *Z2* (linked to mt^-) either form a complex or act sequentially to initiate the pathway leading to meiotic zygote formation. This also triggers either activation or synthesis of proteins coded by the *cpd* and *mitd* genes required, respectively, for elimination of mt^- chloroplast DNA and for destruction of mitochondrial DNA from the mt^+ parent. The *cpd* gene is linked to *cpp* and mt^+, whereas the *mitd* gene is linked to *mitp* and mt^-. If either the *Z1* or *Z2* product is made in insufficient quantities in a gamete, a vegetative diploid would be formed, and the observed biparental transmission of both chloroplast and mitochondrial genomes would occur (Boynton *et al.*, 1987), because neither the meiotic pathway nor the *cpd* and *mitd* functions would be activated.

Dominance of mt^- in diploids for mating function, but not for chloroplast gene transmission (Gillham, 1969; Matagne and Mathieu, 1983; Tsubo and Matsuda, 1984), is explained by supposing the existence of a locus designated *mid* (minus dominance) that is linked to the mt^- allele and comprises two functions, *m1* and *m2* (Galloway and Goodenough, 1985; Goodenough and Ferris, 1987). The *m1* gene product is an activator of mt^--specific genes, while the *m2* gene product is a repressor of mt^+-specific genes with the exception of the *cpp* gene. The *mitp* gene

would be expected to be activated by *m1* and this function, like the *cpp* function, should be dominant in diploids. Matagne (1987) has extended Goodenough's model to include an mt^--linked gene (*cge*), encoding the nuclease that destroys mt^- chloroplast DNA, which is activated at the time of gametic fusion by one of a group of genes (*maps*) that may or may not be linked to mt^+.

Obviously, mutations of the appropriate phenotype would allow many of these predictions to be tested, but so far only three mutations affecting chloroplast gene transmission have been described. The *mat-1* mutant, which was reported to be linked to the mt^- allele, to increase dramatically transmission of chloroplast genetic markers from the mt^- parent in crosses (Sager and Ramanis, 1974), and to cause extensive methylation of mt^- chloroplast DNA in gametes (Sager *et al.*, 1981), was shown to be a spontaneous vegetative diploid (Gillham *et al.*, 1987; Sager and Grabowy, 1985). The *mat-3* mutations linked to mt^+ (Gillham *et al.*, 1987) are postulated to affect the *cpd* gene (Fig. 1). The *mat-2* mutation, which was reported to reduce the already low transmission of chloroplast genes by the mt^- parent (Sager and Ramanis, 1974), has been lost.

VanWinkle-Swift and Aubert (1983) have proposed a model that invokes a direct role for mating type in controlling chloroplast gene inheritance in the homothallic species *C. monoica*. In crosses of an erythromycin-resistant mutant (*ery-u1*) of *C. monoica* to wild type, zygotes either segregated all resistant or all sensitive progeny, in contrast to nuclear mutants, which segregated 2:2. Although every cell in this homothallic species presumably has the potential to act either as mt^+ or mt^- in a cross, the mating type of each gamete is postulated to be fixed at the time of fusion and thereby determines the inheritance pattern of its chloroplast DNA. Assuming that the *ery-u1* mutation affects a chloroplast gene, when an *ery-u1* gamete mates as mt^+ all the progeny of that zygote will be erythromycin-resistant, but when the reciprocal mating-type switch occurs that zygote will segregate all erythromycin-sensitive progeny. A nuclear mutant *mtl-1* that is unable to transmit uniparentally inherited markers such as *ery-u1* in crosses to an mtl^+ strain has also been isolated (VanWinkle-Swift and Hahn, 1986). Matings of *mtl-1* cells yielded many nonviable zygotes in which all, instead of half, the chloroplast nucleoids vanished, suggesting that the *mtl-1* strain was defective for the function that normally protects chloroplast DNA of mt^+ origin from destruction, and that only the zygotes in which the mt^+ parent had the wild-type *mtl* allele were viable (VanWinkle-Swift and Salinger, 1988).

In the heterothallic species *C. moewusii*, chloroplast nucleoids from one gamete were seen to disappear following cell fusion, suggesting uni-

parental transmission of chloroplast DNA (Coleman and Maguire, 1983). However, mutations conferring resistance to erythromycin (*er-nM1*) and streptomycin (*sr-nM1*) were found to exhibit biparental non-Mendelian inheritance in this species (Lee and Lemieux, 1986). These mutations were subsequently shown to be single base-pair changes in the chloroplast 16S (*sr-nM1*) and 23S (*er-nM1*) ribosomal DNA genes (Gauthier *et al.*, 1988). This discrepancy regarding the fate of chloroplast DNA was postulated to result from differences in nutrient availability (Lee and Lemieux, 1986), because providing zygotes with a fresh nutrient supplement reduced nucleoid destruction (Coleman and Maguire, 1983). In *C. reinhardtii*, the frequency of exceptional zygotes is known to decline with increasing stringency of nitrogen starvation during gametogenesis of mt^- but not mt^+ cells (Sears *et al.*, 1980).

C. Transmission of Plastids in *Pelargonium* and *Oenothera*

Biparental transmission of chloroplasts in *Pelargonium* has been studied extensively for many years (see Tilney-Bassett, 1988). Crosses of green with variegated plants had long been known to produce complex patterns of plastid inheritance ranging from maternal to biparental to paternal. Tilney-Bassett (1988) has carried out a series of crosses directed at determining the genetic basis for these patterns. From each of two white-margined, variegated leaf chimeric cultivars, two isogenic clones were selected, one producing only green plastids and the other yielding some shoots that contained only white plastids. In reciprocal crosses using these isogenic clones, mutant plastids were less successfully transmitted than normal plastids regardless of whether the variegated plant was the female or male parent. Because neither selection against white embryos nor environmental effects could account for differing frequencies of maternal, biparental, and paternal zygotes from dissimilar cultivars, the variation in inheritance pattern in crosses of dissimilar cultivars was thought to result from genetic differences.

Tilney-Bassett (1988) identified two distinct plastid segregation patterns. Some cultivars (Type I) produced mostly maternal zygotes, some biparental, and few paternal ones and were true-breeding in self-crosses. Others (Type II) produced nearly equal frequencies of maternal and paternal zygotes, with biparental zygotes being rare. On selfing, these gave half Type II and half Type I progeny. These ratios suggested that Type I cultivars were homozygous for a nuclear gene (*Pr*) controlling plastid segregation pattern and that Type II cultivars were hetero-

zygous for the *Pr1* and *Pr2* alleles of this gene. However, in backcrosses, unexpected ratios were obtained that led to the postulate that *Pr1* was a self-compatible recessive allele and *Pr2* was a self-incompatible dominant allele, while *Pr1* and *Pr2* were cross-compatible. Tilney-Bassett also had to assume that sporophytic control existed on the male side (i.e., in a *Pr1 Pr2* plant all pollen grains were phenotypically *Pr2*) and gametophytic control occurred on the female side. Whatever the explanation of these disparate ratios, the results indicate that maternal plastids are selected in *Pr1 Pr1* homozygotes, whereas maternal and paternal plastids are equally successful in *Pr1 Pr2* heterozygotes. Modifier genes that interact with the *Pr* alleles were also hypothesized to account for differences between cultivars having the same transmission pattern.

Biparental plastid transmission in *Oenothera* also has been studied for many years (see reviews by Gillham, 1978; Kirk and Tilney-Bassett, 1978; Kutzelnigg and Stubbe, 1974). The seven pairs of chromosomes in *Oenothera* species are linked together in one or more chains or rings at the first meiotic division as a result of a series of subterminal reciprocal translocations, which prevent independent assortment and restrict recombination to the chromosome ends (Cleland, 1972). Because of this, and the biparental transmission of plastids, crosses between different *Oenothera* species can be made to assess the compatibility of different combinations of plastomes and nuclear genomes.

From a study of plastome–genome interactions among 400 different combinations of the 10 North American and 14 European species in the subgenus *Euoenothera*, Stubbe (1964) concluded that the subgenus contained five different plastid types and three basic haploid nuclear genomes that could be arranged in six diploid combinations. The compatibility responses between these five plastid and six nuclear genome combinations ranged from normal to lethal. Plastome IV was considered the most primitive because it was fully compatible with all nuclear genome combinations, save one, with which it was partially compatible. The three different haploid genomes defined on the basis of genome–plastome compatibility corresponded to the three main species groups of North American *Euoenothera* recognized by Cleland (1972) on the basis of taxonomic criteria.

More recently, the five plastomes postulated by Stubbe have been shown to correspond to five physically distinguishable chloroplast genomes. Gordon *et al.* (1981, 1982) have prepared restriction maps of all five wild-type *Euoenothera* plastomes. Alterations in the positions of restriction sites among the five plastomes result primarily from insertions and deletions in the large single-copy region, although some occur symmetrically within the inverted repeat region. The three plastomes whose

DNAs exhibit the largest differences (I, III, and V) are those regarded by Stubbe as being the newest. Also, plastome II appears on physical grounds to be intermediate between plastomes IV and I, as proposed by Stubbe (1964).

Plastid competition in *Oenothera* has been investigated extensively in crosses between green and variegated cultivars (studies by Schötz, reviewed by Kirk and Tilney-Bassett, 1978). Several conclusions emerged from Schötz's experiments. First, major control of plastid transmission in *Oenothera* was exerted by the maternal parent so nonvariegated plants always accounted for a high percentage of the total. Even among variegated plants, the maternal parent appeared to contribute most of the plastids, an observation confirmed by electron microscopic examination of egg cells and zygotes of *Oenothera lamarckiana* (Meyer and Stubbe, 1974). Second, a paternal effect was revealed by variation in the percentage of white tissue in progeny from different interspecific crosses. This was partly a function of the plastids themselves and partly due to the nuclear background. Third, three groups of green plastids were distinguished, with fast, medium, and slow competitive abilities with respect to white plastids. The fast type corresponds to plastomes I and III of Stubbe, medium to plastome II, and slow to plastomes IV and V (see Tilney-Bassett, 1975). Fourth, the competitive abilities of different plastid types were affected by nuclear genotype. Re-examination of the competitive abilities of different plastome types in a constant nuclear background gave the same results as those obtained by Schötz from comparisons made in various nuclear backgrounds (Chiu *et al.*, 1987). Although the assumption in Schötz's experiments is that differential plastid multiplication underlies the phenomenon of plastid competition, this remains to be proved. Other possibilities such as differential plastid input or differential destruction of plastids in the zygote could also explain these results.

V. CHLOROPLAST TRANSMISSION FOLLOWING SOMATIC CELL FUSION

In many commonly studied crop plants, strictly maternal transmission of plastids in sexual crosses presents a barrier to the creation of new plastome–genome combinations that might yield improved cultivars for agricultural use. Fortunately, this problem has been overcome in a number of species using the techniques of somatic cell hybridization and

regeneration of whole plants (see Galun and Aviv, 1983, 1986). To construct a new plastome–genome combination (1) the nuclear genome of the chloroplast "donor" must be eliminated, (2) a simple method for detection of "cybrids" containing the donor chloroplast is required, (3) "recipient" cells that do not fuse with a donor must be eliminated, and (4) fusion of donor and recipient chloroplasts and recombination of their genomes must be prevented.

The last criterion is easily met as chloroplast fusion appears to occur very rarely in higher plants, effectively blocking recombination of chloroplast genomes from different parents. So far, two single *bona fide* chloroplast DNA recombinants have been reported in higher plants among thousands of progeny screened. One arose in a somatic hybrid between *Nicotiana tabacum* and *Nicotiana plumbaginifolia* (Medgyesy *et al.*, 1985), and the other in somatic hybrids between *N. tabacum* and *Solanum tuberosum* (Thanh and Medgyesy, 1989). In sexual crosses of *Oenothera*, where >7,500 progeny were scored, no chloroplast recombinants were detected (Chiu and Sears, 1985), nor have recombinants been detected among ca. 1,000 *Nicotiana* calli screened by Fluhr *et al.* (1984) or among >400 plants regenerated from *Brassica* protoplast fusions or anther cultures (Kemble *et al.*, 1988). In contrast, mitochondrial genome recombination in cybrids seems to be the norm (Galun and Aviv, 1983; Nagy *et al.*, 1983).

The donor–recipient method has been used successfully to construct cybrids with novel plastome–genome combinations in a variety of interspecific fusions in *Nicotiana* (Table III). In most instances, the donor nuclear genome was eliminated by either X- or gamma-rays, but in one case donor cells were enucleated by centrifugation through a Percoll gradient (Maliga *et al.*, 1982). Two methods have been employed to obtain cybrids from the recipient cell population. The first involves conditions that select against growth of recipient protoplasts and in favor of cybrids or hybrids. The mannitol selection method used to obtain cybrids between *N. tabacum* and *Nicotiana sylvestris* (Galun and Aviv, 1983) is of limited general utility because it relies on the differential sensitivity of protoplasts of these two species to mannitol. A more generally applicable method is iodoacetate treatment of recipient cells to prevent their growth while allowing growth of hybrids between iodoacetate-treated and untreated cells (Medgyesy *et al.*, 1980) and cybrids between irradiated donor cells and iodoacetate-treated recipient cells (Sidorov *et al.*, 1981).

Direct selection for chloroplast markers from the donor such as streptomycin or lincomycin resistance would appear to be the preferred

TABLE III

Cybrid Plants Obtained between Different Plant Species Using Protoplast Fusion and the Donor–Recipient Method[a]

Plastid Donor Species	Recipient Species	Means of Cybrid Selection	Reference
Nicotiana tabacum	*N. sylvestris*	mannitol	Galun and Aviv (1983)
N. tabacum, SR1	*N. sylvestris*	streptomycin	Fluhr *et al.* (1983)
N. tabacum, SR1	*N. plumbaginifolia*	streptomycin	Menczel *et al.* (1982)
N. tabacum, LS1	*N. plumbaginifolia*	iodoacetate	Sidorov *et al.* (1981)
N. tabacum, SR1[b]	*N. plumbaginifolia*	streptomycin	Maliga *et al.* (1982)
N. sylvestris, LR	*N. plumbaginifolia*	lincomycin	Cséplö *et al.* (1984)
N. plumbaginifolia, TBR2	*N. tabacum*, SR1-A15	color	Menczel *et al.* (1986)
N. tabacum, SR1	*Salpiglossis sinuata*	streptomycin	Thanh *et al.* (1988)
S. sinuata	*N. tabacum*, LS1	color	Thanh *et al.* (1988)
Solanum tuberosum[c]	*N. tabacum*, LS1	color	Thanh and Medgyesy (1989)
N. rustica	*N. sylvestris*	mannitol	Aviv *et al.* (1984)

[a] Donor nuclei eliminated with X- or gamma-rays from a ^{60}Co source. Chloroplast gene symbols are as follows: SR1, streptomycin resistance; SR1-A15, streptomycin-resistant and pigment-deficient; LR, lincomycin resistance; LS1, light-sensitive pigmentation; TBR2, terbutryn resistance.

[b] Donor protoplasts enucleated by centrifugation.

[c] A single green plastid recombinant survived.

method for detection of cybrids (Table III). Protoplast-derived clones are screened for their ability to form green, resistant calli on antibiotic-containing media (see Maliga *et al.*, 1982; Cséplö and Maliga, 1984), which can then be used to regenerate resistant plants. Similarly, pigment-deficient plastome mutants can also be used as recipients and green, irradiated donor protoplasts as donors (see Menczel *et al.*, 1986). Green calli containing donor chloroplasts and the recipient nuclear genome are then selected. Irradiation of donor protoplasts with high doses of X-rays does not completely eliminate their ability to form hybrids (Menczel *et al.*, 1982). To destroy the remaining donor chromosomes, the cybrid itself can be irradiated and used as the donor in a backfusion to the original recipient while selecting for the donor plastome (Thanh and Medgyesy, 1989).

Progress with the donor–recipient system in *Nicotiana* has been rapid. New selectable color and resistance markers have been isolated (Cséplö and Maliga, 1984; Cseplö *et al.*, 1984; Fluhr *et al.*, 1985; Svab and Maliga, 1986), and several of the resistance markers have been shown to result in specific base-pair changes in the chloroplast ribosomal RNA genes (Cséplö *et al.*, 1988; Etzold *et al.*, 1987; Fromm *et al.*, 1987). A protoplast fusion method that permits unambiguous determination of whether pig-

ment mutations are of nuclear or plastid origin has been devised (Aviv and Galun, 1985). The sorting out of plastids in *Nicotiana* cybrids has been studied (Fluhr *et al.*, 1983) and the cybrid method used to transfer chloroplasts resistant to triazine herbicides between different *Nicotiana* species (Menczel *et al.*, 1986).

As the donor–recipient technique has been perfected in *Nicotiana*, more distant plastome–genome combinations have been attempted. Successful intertribal chloroplast transfer in reciprocal directions has been reported between *N. tabacum* and *Salpiglossis sinuata* (Thanh *et al.*, 1988; Table III). However, chloroplast transfers involving more distantly related members of the Solanaceae were less successful. Thus, streptomycin-resistant colonies were not detected when the SR1 mutant of *N. tabacum* was used as the chloroplast donor and *Solanum nigrum* as the recipient (Thanh *et al.*, 1988), and a single green chloroplast recombinant was isolated when a green *S. tuberosum* (potato) protoplast culture was used as the donor and a light-sensitive (LS1) plastome mutant of *N. tabacum* as the recipient (Thanh and Medgyesy, 1989).

The donor–recipient method has also been used to construct *Citrus* (Vardi *et al.*, 1987) and *Brassica* cybrids (Menczel *et al.*, 1987; Morgan and Maliga, 1987; Robertson *et al.*, 1987). In the *Citrus* cybrids, no genetic or physical chloroplast markers were available, but mitochondrial DNA restriction patterns similar or identical to the donor were observed in two fusion combinations. In *Brassica*, cybrids were obtained after fusing protoplasts of fertile and cytoplasmic male sterile *Brassica napus* lines carrying the original cytoplasm of *B. napus* and the Ogura cytoplasm of *Raphanus sativus*, respectively (Menczel *et al.*, 1987; Morgan and Maliga, 1987). With two possible exceptions, all cybrids isolated possessed chloroplast DNA derived from the fertile parent although mitochondrial DNA was recombinant. The parental plastid genomes sorted out more swiftly in the cybrids than did the parental and recombinant mitochondrial types.

Use of cybrids and whole-plant regeneration allows the synthesis of an array of new plastome–genome combinations between different species and genera whose construction would normally be precluded by maternal inheritance and by infertility arising because of incompatibility or lack of pairing homology between different parental chromosome sets. The potential of the cybrid system is already being realized as a means for transferring herbicide resistance, a chloroplast trait, between different *Nicotiana* species (Menczel *et al.*, 1986). We may expect similar benefits to accrue from the synthesis of novel plastome–genome combinations.

VI. CONCLUDING REMARKS

Despite 80 years of research, the reason why uniparental inheritance of chloroplast and mitochondrial genomes is so frequent in nature remains a mystery. Speculations include the suggestions that lack of organelle transmission by sperm might eliminate competition in the fertilized egg between organelle genomes of different origin (Grun, 1976); that male gametes, because they are naked at the time of fertilization, may be particularly sensitive to infection by foreign organisms (Coleman, 1982); and that elimination or inactivation of chromosomes or organelle DNA from one parent provides for the selective silencing of the DNA from that parent (Sager and Kitchin, 1975). Whether or not any of these explanations is correct, the phenomenon of uniparental inheritance of organelle genomes must have some selective advantage. Unlike meiosis, which probably arose only once in evolution, uniparental inheritance of chloroplast and mitochondrial genes apparently has arisen independently a number of times and occurs by quite different mechanisms in different taxa.

Literature for this chapter was reviewed through November 1989.

REFERENCES

Adams, G. M. W. (1978). Chloroplast gene transmission in *Chlamydomonas reinhardtii:* A random choice model. *Plasmid* **1,** 522–535.

Aviv, D., and Galun, E. (1985). An in vitro procedure to assign pigment mutations in *Nicotiana* to either the chloroplast or the nucleus. *J. Hered.* **76,** 135–136.

Aviv, D., Bleichman, S., Arzee-Gonen, P., and Galun, E. (1984). Intersectional cytoplasmic hybrids in *Nicotiana*. Identification of plastomes and chondriomes in *N. sylvestris* + *N. rustica* cybrids having *N. sylvestris* nuclear genomes. *Theor. Appl. Genet.* **67,** 499–504.

Baur, E. (1909). Das Wesen und die Erblichkeitsverhältnisse der "varietates albomarginatae hort." von *Pelargonium zonale*. *Z. Vererbungs.* **1,** 330–351.

Bennoun, P., Masson, A., and Delosme, M. (1980). A method for complementation analysis of nuclear and chloroplast mutants of photosynthesis in *Chlamydomonas*. *Genetics* **95,** 39–47.

Birky, C. W., Jr. (1983a). The partitioning of cytoplasmic organelles at cell division. *Int. Rev. Cytol., Suppl.* **15,** 49–89.

Birky, C. W., Jr. (1983b). Relaxed cellular controls and organelle heredity. *Science* **222,** 468–475.

Birky, C. W., Jr., VanWinkle-Swift, K. P., Sears, B. B., Boynton, J. E., and Gillham, N. W. (1981). Frequency distributions for chloroplast genes in *Chlamydomonas* zygote clones: Evidence for random drift. *Plasmid* **6,** 173–192.

Birky, C. W., Jr., Fuerst, P., and Maruyama, T. (1989). Organelle gene diversity under migration, mutation, and drift: Equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics* **121,** 613–627.

Boffey, S. A., and Leech, R. M. (1982). Chloroplast DNA levels and the control of chloroplast division in light-grown wheat leaves. *Plant Physiol.* **69,** 1387–1391.

Boffey, S. A., and Lloyd, D. (eds.) (1988). "The Division and Segregation of Organelles." Cambridge University Press, Cambridge. 252 pp.

Bolen, P. L., Gillham, N. W., and Boynton, J. E. (1980). Evidence for persistence of chloroplast markers in the heteroplasmic state in *Chlamydomonas reinhardtii*. *Curr. Genet.* **2,** 159–167.

Bolen, P. L., Grant, D. M., Swinton, D., Boynton, J. E., and Gillham, N. W. (1982). Extensive methylation of chloroplast DNA by a nuclear gene mutation does not affect chloroplast gene transmission in *Chlamydomonas*. *Cell* **28,** 335–343.

Börner, T., and Sears, B. B. (1986). Plastome mutants. *Plant Mol. Biol. Rep.* **4,** 69–92.

Bowman, C. M. (1986). Copy numbers of chloroplast and nuclear genomes are proportional in mature mesophyll cells of *Triticum* and *Aegilops* species. *Planta* **167,** 264–274.

Bowman, C. M., and Dyer, T. A. (1986). The location and possible evolutionary significance of small dispersed repeats in wheat ctDNA. *Curr. Genet.* **10,** 931–941.

Boynton, J. E., Gillham, N. W., Harris, E. H., Tingle, C. L., VanWinkle-Swift, K., and Adams, G. M. W. (1976). Transmission, segregation, and recombination of chloroplast genes in *Chlamydomonas*. *In* "Genetics and Biogenesis of Chloroplasts and Mitochondria" (Th. Bücher, W. Neupert, W. Sebald, and S. Werner, eds.), pp. 313–322. Elsevier/North-Holland Biomedical Press, Amsterdam.

Boynton, J. E., Harris, E. H., Burkhart, B. D., Lamerson, P. M., and Gillham, N. W. (1987). Transmission of mitochondrial and chloroplast genomes in crosses of *Chlamydomonas*. *Proc. Natl. Acad. Sci. USA* **84,** 2391–2395.

Boynton, J. E., Gillham, N. W., Harris, E. H., Newman, S. M., Randolph-Anderson, B. L., Johnson, A. M., and Jones, A. R. (1990). Manipulating the chloroplast genome of *Chlamydomonas*—Molecular genetics and transformation. *In* "Current Research in Photosynthesis," Vol III, 509–516 (M. Baltscheffsky, ed.) Kluwer Academic Publishers, Netherlands.

Boynton, J. E., Gillham, N. W., and Harris, E. H. (1991). Chloroplast genetics and transformation in *Chlamydomonas*. *In* "Cell Organelles. Advances in Plant Gene Research," Vol. 6 (Th. Hohn, ed.). Springer-Verlag, Vienna, in press.

Briat, J. F., Gigot, C., Laulhère, J. P., and Mache, R. (1982). Visualization of a spinach plastid transcriptionally active DNA–protein complex in a highly condensed structure. *Plant Physiol.* **69,** 1205–1211.

Butterfass, Th. (1988). Nuclear control of plastid division. *In* "The Division and Segregation of Organelles" (S. A. Boffey and D. Lloyd, eds.), pp. 21–38. Cambridge University Press, Cambridge.

Cavalier-Smith, T. (1970). Electron microscopic evidence for chloroplast fusion in zygotes of *Chlamydomonas reinhardii*. *Nature* **228,** 333–335.

Chiang, K.-S., Friedman, E., Malavasic, M. J., Jr., Feng, M.-H. L., Eves, E. M., Feng, T. Y., and Swinton, D. C. (1981). On the folding and organization of chloroplast DNA in *Chlamydomonas reinhardtii*. *Ann. N.Y. Acad. Sci.* **361,** 219–247.

Chiu, W.-L., and Sears, B. B. (1985). Recombination between chloroplast DNAs does not occur in sexual crosses of *Oenothera*. *Mol. Gen. Genet.* **198,** 525–528.

Chiu, W.-L., Stubbe, W., and Sears, B. B. (1987). Plastid inheritance in *Oenothera:* Organelle genome modifies the extent of biparental plastid transmission. *Curr. Genet.* **13,** 181–190.

Cleland, R. E. (1972). "*Oneothera:* Cytogenetics and Evolution." Academic Press, New York. 370 pp.

Coleman, A. W. (1982). Sex is dangerous in a world of potential symbionts or the basis of selection for uniparental inheritance. *J. Theor. Biol.* **97,** 367–369.

Coleman, A. W. (1984). The fate of chloroplast DNA during cell fusion, zygote maturation and zygote germination in *Chlamydomonas reinhardi* as revealed by DAPI staining. *Exp. Cell Res.* **152,** 528–540.

Coleman, A. W., and Maguire, M. J. (1983). Cytological detection of the basis of uniparental inheritance of plastid DNA in *Chlamydomonas moewusii. Curr. Genet.* **7,** 211–218.

Connett, M. B. (1987). Mechanisms of maternal inheritance of plastids and mitochondria: Developmental and ultrastructural evidence. *Plant Mol. Biol. Rep.* **4,** 103–205.

Cornu, A., and Dulieu, H. (1988). Pollen transmission of plastid-DNA under genotypic control in *Petunia hybrida* Hort. *J. Hered.* **79,** 40–44.

Correns, C. (1909). Vererbungsversuche mit blass (gelb) grünen und buntblättrigen Sippen bei *Mirabilis jalapa, Urtica pilulifera* und *Lunaria annua. Z. Vererbungs.* **1,** 291–329.

Corriveau, J. S., and Coleman, A. W. (1988). Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. *Amer. J. Bot.* **75,** 1443–1458.

Cséplö, A., and Maliga, P. (1984). Large scale isolation of maternally inherited lincomycin resistance mutations, in diploid *Nicotiana plumbaginifolia* protoplast cultures. *Mol. Gen. Genet.* **196,** 407–412.

Cséplö, A., Nagy, F., and Maliga, P. (1984). Interspecific protoplast fusion to rescue a cytoplasmic lincomycin resistance mutation into fertile *Nicotiana plumbaginifolia* plants. *Mol. Gen. Genet.* **198,** 7–11.

Cséplö, A., Etzold, T., Schell, J., and Schreier, P. H. (1988). Point mutations in the 23S rRNA genes of four lincomycin resistant *Nicotiana plumbaginifolia* mutants could provide new selectable markers for chloroplast transformation. *Mol. Gen. Genet.* **214,** 295–299.

Day, A., and Ellis, T. H. N. (1984). Chloroplast DNA deletions associated with wheat plants regenerated from pollen: Possible basis for maternal inheritance of chloroplasts. *Cell* **39,** 359–368.

Day, A., and Ellis, T. H. N. (1985). Deleted forms of plastid DNA in albino plants from cereal anther culture. *Curr. Genet.* **9,** 671–678.

Dickinson, H. G., and Li, F. L. (1988). Organelle behaviour during higher plant gametogenesis. *In* "The Division and Segregation of Organelles" (S. A. Boffey and D. Lloyd, eds.), pp. 131–148. Cambridge University Press, Cambridge.

Ebersold, W. T. (1967). *Chlamydomonas reinhardi:* Heterozygous diploid strains. *Science* **157,** 447–449.

Ellis, J. R., and Leech, R. M. (1985). Cell size and chloroplast size in relation to chloroplast replication in light-grown wheat leaves. *Planta* **165,** 120–125.

Ellis, J. R., Jellings, A. J., and Leech, R. M. (1983). Nuclear DNA content and the control of chloroplast replication in wheat leaves. *Planta* **157,** 376–380.

Ellis, T. H. N., and Day, A. (1986). A hairpin plastid genome in barley. *EMBO J.* **5,** 2769–2774.

Erickson, J. M., Rahire, M., Rochaix, J.-D., and Mets, L. (1985). Herbicide resistance and cross-resistance: Changes at three distinct sites in the herbicide-binding protein. *Science* **228,** 204–207.

Etzold, T., Fritz, C. C., Schell, J., and Schreier, P. H. (1987). A point mutation in the chloroplast 16S rRNA gene of a streptomycin resistant *Nicotiana tabacum. FEBS Lett.* **219,** 343–346.

Feng, T.-Y., and Chiang, K.-S. (1984). The persistence of maternal inheritance in *Chlamydomonas* despite hypomethylation of chloroplast DNA induced by inhibitors. *Proc. Natl. Acad. Sci. USA* **81,** 3438–3442.

Firshein, W. (1989). Role of the DNA/membrane complex in prokaryotic DNA replication. *Ann. Rev. Microbiol.* **43,** 89–120.

Fluhr, R., Aviv, D., Edelman, M., and Galun, E. (1983). Cybrids containing mixed and sorted-out chloroplasts following interspecific somatic fusions in *Nicotiana*. *Theor. Appl. Genet.* **65,** 289–294.

Fluhr, R., Aviv, D., Galun, E., and Edelman, M. (1984). Generation of heteroplastidic *Nicotiana* cybrids by protoplast fusion: Analysis for plastid recombinant types. *Theor. Appl. Genet.* **67,** 491–497.

Fluhr, R., Aviv, D., Galun, E., and Edelman, M. (1985). Efficient induction and selection of chloroplast-encoded antibiotic-resistant mutants in *Nicotiana*. *Proc. Natl. Acad. Sci. USA* **82,** 1485–1489.

Forster, J. L., Grabowy, C. T., Harris, E. H., Boynton, J. E., and Gillham, N. W. (1980). Behavior of chloroplast genes during the early zygotic divisions of *Chlamydomonas reinhardtii*. *Curr. Genet.* **1,** 137–153.

Fromm, H., Edelman, M., Aviv, D., and Galun, E. (1987). The molecular basis for rRNA-dependent spectinomycin resistance in *Nicotiana* chloroplasts. *EMBO J.* **6,** 3233–3237.

Galloway, R. E., and Goodenough, U. W. (1985). Genetic analysis of mating locus linked mutations in *Chlamydomonas reinhardii*. *Genetics* **111,** 447–461.

Galloway, R. E., and Holden, L. R. (1985). Transmission and recombination of chloroplast genes in asexual crosses of *Chlamydomonas reinhardii*. II. Comparisons with observations of sexual diploids. *Curr. Genet.* **10,** 221–228.

Galun, E., and Aviv, D. (1983). Cytoplasmic hybridization: Genetic and breeding applications. *In* "Handbook of Plant Cell Culture," Vol. 1, "Techniques for Propagation and Breeding" (D. A. Evans, W. R. Sharp, P. V. Ammirato, and Y. Yamada, eds.), pp. 358–392. Macmillan, New York.

Galun, E., and Aviv, D. (1986). Organelle transfer. *Meth. Enzymol.* **118,** 595–611.

Gauthier, A., Turmel, M., and Lemieux, C. (1988). Mapping of chloroplast mutations conferring resistance to antibiotics in *Chlamydomonas:* Evidence for a novel site of streptomycin resistance in the small subunit rRNA. *Mol. Gen. Genet.* **214,** 192–197.

Gillham, N. W. (1963). Transmission and segregation of a non-chromosomal factor controlling streptomycin resistance in diploid *Chlamydomonas*. *Nature* **200,** 294.

Gillham, N. W. (1969). Uniparental inheritance in *Chlamydomonas reinhardi*. *Amer. Natur.* **103,** 355–388.

Gillham, N. W. (1978). "Organelle Heredity." Raven Press, New York. 602 pp.

Gillham, N. W., Boynton, J. E., and Lee, R. W. (1974). Segregation and recombination of non-Mendelian genes in *Chlamydomonas*. *Genetics* **78,** 439–457.

Gillham, N. W., Boynton, J. E., and Harris, E. H. (1985). Evolution of plastid DNA. *In* "Evolution of Genome Size" (T. Cavalier-Smith, ed.), pp. 299–351. John Wiley and Sons, New York.

Gillham, N. W., Boynton, J. E., Johnson, A. M., and Burkhart, B. D. (1987). Mating type linked mutations which disrupt the uniparental transmission of chloroplast genes in *Chlamydomonas*. *Genetics* **115,** 677–684.

Goodenough, U. W., and Ferris, P. J. (1987). Genetic regulation of development in *Chlamydomonas*. *In* "Genetic Regulation of Development" (W. Loomis, ed.), pp. 171–189. Alan R. Liss, New York.

Gordon, K. H. J., Crouse, E. J., Bohnert, H. J., and Herrmann, R. G. (1981). Restriction

endonuclease cleavage site map of chloroplast DNA from *Oenothera parviflora* (*Euoenothera* Plastome IV). *Theor. Appl. Genet.* **59,** 281–296.

Gordon, K. H. J., Crouse, E. J., Bohnert, H. J., and Herrmann, R. G. (1982). Physical mapping of differences in chloroplast DNA of the five wild-type plastomes in *Oenothera* subsection *Euoenothera*. *Theor. Appl. Genet.* **61,** 373–384.

Govindaraju, D. R., Wagner, D. B., Smith, G. P., and Dancik, B. P. (1988). Chloroplast DNA variation within individual trees of a *Pinus banksiana* and *Pinus contorta* sympatric region. *Can. J. For. Res.* **18,** 1347–1350.

Govindaraju, D. R., Dancik, B. P., and Wagner, D. B. (1989). Novel chloroplast DNA polymorphism in a sympatric region of two pines. *J. Evol. Biol.* **2,** 49–59.

Grant, D. M., Gillham, N. W., and Boynton, J. E. (1980). Inheritance of chloroplast DNA in *Chlamydomonas reinhardtii*. *Proc. Natl. Acad. Sci. USA* **77,** 6067–6071.

Grun, P. (1976). "Cytoplasmic Genetics and Evolution." Columbia University Press, New York. 435 pp.

Hagemann, R. (1976). Plastid distribution and plastid competition in higher plants and the induction of plastom mutations by nitroso-urea-compounds. *In* "Genetics and Biogenesis of Chloroplasts and Mitochondria" (Th. Bücher, W. Neupert, W. Sebald, and S. Werner, eds.), pp. 331–338. Elsevier/North-Holland Biomedical Press, Amsterdam.

Hagemann, R. (1979). Genetics and molecular biology of plastids of higher plants. *Stadler Symp.* **11,** 91–115.

Hagemann, R. (1983). The formation of generative and sperm cells without plastids in angiosperms and the underlying mechanisms. *In* "Fertilization and Embryogenesis in Ovulated Plants" (O. Erdelska, ed.), pp. 97–99. Bratislava Slov. Acad. Sci., Bratislava.

Hagemann, R., and Schröder, M. B. (1985). New results about the presence of plastids in generative and sperm cells of Gramineae. *In* "Reproduction in Seed Plants, Ferns, and Mosses" (M. T. M. Willemse and J. L. Van Went, eds.), pp. 53–55. Pudoc, Wageningen.

Hansmann, P., Falk, H., Ronai, K., and Sitte, P. (1985). Structure, composition, and distribution of plastid nucleoids in *Narcissus pseudonarcissus*. *Planta* **164,** 459–472.

Harris, E. H. (1989). "The Chlamydomonas Sourcebook," Academic Press, San Dieigo. 780 pp.

Harris, E. H., Burkhart, B. D., Gillham, N. W., and Boynton, J. E. (1989). Antibiotic resistance mutations in the chloroplast 16S and 23S rRNA genes of *Chlamydomonas reinhardtii:* Correlation of genetic and physical maps of the chloroplast genome. *Genetics* **123,** 281–282.

Hashimoto, H. (1985). Changes in distribution of nucleoids in developing and dividing chloroplasts and etioplasts of *Avena sativa*. *Protoplasma* **127,** 119–127.

Hashimoto, H., and Possingham, J. V. (1989a). DNA levels in dividing and developing plastids in expanding primary leaves of *Avena sativa*. *J. Exp. Bot.* **40,** 257–262.

Hashimoto, H., and Possingham, J. V. (1989b). Effect of light on the chloroplast division cycle and DNA synthesis in cultured leaf discs of spinach. *Plant Physiol.* **89,** 1178–1183.

Howe, C. J., Barker, R. F., Bowman, C. M., and Dyer, T. A. (1988). Common features of three inversions in wheat chloroplast DNA. *Curr. Genet.* **13,** 343–349.

Kemble, R. J., Yanow, S. A., Wu, S.-C., and Barsby, T. L. (1988). Absence of mitochondrial and chloroplast DNA recombination in *Brassica napus* plants regenerated from protoplasts, protoplast fusions and anther culture. *Theor. Appl. Genet.* **75,** 875–881.

Kirk, J. T. O., and Tilney-Bassett, R. A. E. (1978). "The Plastids: Their Chemistry, Structure, Growth and Inheritance," 2nd Edition. Elsevier/North-Holland, Amsterdam. 960 pp.

Kuroiwa, T. (1985). Mechanisms of maternal inheritance of chloroplast DNA: An active digestion hypothesis. *Microbiol. Sci.* **2,** 267–272.

Kuroiwa, T. (1989). The nuclei of cellular organelles and the formation of daughter organelles by the "plastid-dividing ring." *Bot. Mag. (Tokyo)* **102,** 291–329.

Kuroiwa, T., and Hori, T. (1986). Preferential digestion of male chloroplast nuclei and mitochondrial nuclei during gametogenesis of *Bryopsis maxima* Okamura. *Protoplasma* **133,** 85–87.

Kuroiwa, T., and Suzuki, T. (1980). An improved method for the demonstration of the *in situ* chloroplast nuclei in higher plants. *Cell Struct. Funct.* **5,** 195–197.

Kuroiwa, T., Suzuki, T., Ogawa, K., and Kawano, S. (1981). The chloroplast nucleus: Distribution, number, size, and shape, and a model for the multiplication of the chloroplast genome during chloroplast development. *Plant Cell Physiol.* **22,** 381–396.

Kuroiwa, T., Kawano, S., Nishibayashi, S., and Sato, C. (1982). Epifluorescent microscopic evidence for maternal inheritance of chloroplast DNA. *Nature* **298,** 481–483.

Kuroiwa, T., Kawano, S., and Sato, C. (1983a). Mechanisms of maternal inheritance. I. Protein synthesis involved in preferential destruction of chloroplast DNA of male origin. *Proc. Japan Acad.* **59B,** 177–181.

Kuroiwa, T., Kawano, S., and Sato, C. (1983b). Mechanisms of maternal inheritance. II. RNA synthesis involved in preferential destruction of chloroplast DNA of male origin. *Proc. Japan Acad.* **59B,** 182–185.

Kuroiwa, T., Enomoto, S., and Shihira-Ishikawa, I. (1985a). Preferential destruction of chloroplast nucleoids in zygotes in green alga *Dictyosphaeria cavernosa* and *Acetabularia calyculus*. *Experientia* **41,** 1178–1179.

Kuroiwa, T., Nakamura, S., Sato, C., and Tsubo, Y. (1985b). Epifluorescent microscopic studies on the mechanism of preferential destruction of chloroplast nucleoids of male origin in young zygotes of *Chlamydomonas reinhardtii*. *Protoplasma* **125,** 43–52.

Kuroiwa, H., Sugai, M., and Kuroiwa, T. (1988). Behavior of chloroplasts and chloroplast nuclei during spermatogenesis in the fern, *Pteris vittata* L. *Protoplasma* **146,** 89–100.

Kutzelnigg, H., and Stubbe, W. (1974). Investigations on plastome mutants in Oenothera. 1. General conclusions. *Sub-Cell Biochem.* **3,** 73–89.

Lawrence, M. E., and Possingham, J. V. (1986). Microspectrofluorometric measurement of chloroplast DNA in dividing and expanding leaf cells of *Spinacia oleracea*. *Plant Physiol.* **81,** 708–710.

Lax, A. R., Vaughn, K. C., Duke, S. O., and Endrizzi, J. E. (1987). Structural and physiological studies of a plastome cotton mutant with slow sorting out. *J. Hered.* **78,** 147–152.

Lee, D. J., Blake, T. K., and Smith, S. E. (1988). Biparental inheritance of chloroplast DNA and the existence of heteroplasmic cells in alfalfa. *Theor. Appl. Genet.* **76,** 545–549.

Lee, D. J., Blake, T. K., Smith, S. E., Bingham, E. T., and Carroll, T. W. (1989). Chloroplast genome mapping and plastid ultrastructure analysis of chlorophyll deficient mutants of alfalfa. *Crop Sci.* **29,** 190–196.

Lee, R. W., and Haughn, G. W. (1980). Induction and segregation of chloroplast mutations in vegetative cell cultures of *Chlamydomonas reinhardtii*. *Genetics* **96,** 79–94.

Lee, R. W., and Lemieux, C. (1986). Biparental inheritance of non-Mendelian gene markers in *Chlamydomonas moewusii*. *Genetics* **113,** 589–600.

Leech, R. M., and Pyke, K. A. (1988). Chloroplast division in higher plants with particular reference to wheat. *In* "The Division and Segregation of Organelles" (S. A. Boffey and D. Lloyd, eds.), pp. 39–62. Cambridge University Press, Cambridge.

Lemieux, C., Turmel, M., and Lee, R. W. (1980). Characterization of chloroplast DNA in *Chlamydomonas eugametos* and *C. moewusii* and its inheritance in hybrid progeny. *Curr. Genet.* **2,** 139–147.

Lemieux, C., Turmel, M., and Lee, R. W. (1981). Physical evidence for recombination of chloroplast DNA in hybrid progeny of *Chlamydomonas eugametos* and *C. moewusii*. *Curr. Genet.* **3,** 97–103.

Lemieux, C., Turmel, M., Seligy, V. L., and Lee, R. W. (1984). Chloroplast DNA recombination in interspecific hybrids of *Chlamydomonas:* Linkage between a nonmendelian locus for streptomycin resistance and restriction fragments coding for 16S rRNA. *Proc. Natl. Acad. Sci. USA* **81,** 1164–1168.

Lindbeck, A. G. C., and Rose, R. J. (1987). Is DNA associated with spinach chloroplast vesicles at specific sites? *J. Plant Physiol.* **129,** 425–434.

Lindbeck, A. G. C., Rose, R. J., Lawrence, M. E., and Possingham, J. V. (1987). The role of chloroplast membranes in the location of chloroplast DNA during the greening of *Phaseolus vulgaris* etioplasts. *Protoplasma* **139,** 92–99.

Lindbeck, A. G. C., Rose, R. J., Lawrence, M. J., and Possingham, J. V. (1989). The chloroplast nucleoids of the bundle sheath and mesophyll cells of *Zea mays. Physiol. Plant.* **75,** 7–12.

Maliga, P., Lörz, H., Lázár, G., and Nagy, F. (1982). Cytoplast–protoplast fusion for interspecific chloroplast transfer in *Nicotiana. Mol. Gen. Genet.* **185,** 211–215.

Matagne, R. F. (1981). Transmission of chloroplast alleles in somatic fusion products obtained from vegetative cells and/or 'gametes' of *Chlamydomonas reinhardi. Curr. Genet.* **3,** 31–36.

Matagne, R. F. (1987). Chloroplast gene transmission in *Chlamydomonas reinhardtii.* A model for its control by the mating-type locus. *Curr. Genet.* **12,** 251–256.

Matagne, R. F., and Hermesse, M.-P. (1980). Chloroplast gene inheritance studied by somatic fusion in *Chlamydomonas reinhardtii. Curr. Genet.* **1,** 127–131.

Matagne, R. F., and Mathieu, D. (1983). Transmission of chloroplast genes in triploid and tetraploid zygospores of *Chlamydomonas reinhardtii:* Roles of mating-type gene dosage and gametic chloroplast DNA content. *Proc. Natl. Acad. Sci. USA* **80,** 4780–4783.

Matagne, R. F., and Schaus, M. (1985). About biased and non-biased transmission of chloroplast genes following artificial fusion of gametes in *Chlamydomonas reinhardtii. Curr. Genet.* **10,** 81–85.

Matsuda, Y., Sakamoto, K., and Tsubo, Y. (1983). Biased and non-biased transmission of chloroplast genes in somatic fusion products of *Chlamydomonas reinhardtii. Curr. Genet.* **7,** 339–345.

Matsuda, Y., Saito, T., Umemoto, T., and Tsubo, Y. (1988). Transmission patterns of chloroplast genes after polyethylene glycol-induced fusion of gametes in non-mating mutants of *Chlamydomonas reinhardtii. Curr. Genet.* **14,** 53–58.

Medgyesy, P., Menczel, L., and Maliga, P. (1980). The use of cytoplasmic streptomycin resistance: Chloroplast transfer from *Nicotiana tabacum* into *Nicotiana sylvestris,* and isolation of their somatic hybrids. *Mol. Gen. Genet.* **179,** 693–698.

Medgyesy, P., Fejes, E., and Maliga, P. (1985). Interspecific chloroplast recombination in a *Nicotiana* somatic hybrid. *Proc. Natl. Acad. Sci. USA* **82,** 6960–6964.

Medgyesy, P., Páy, A., and Márton, L. (1986). Transmission of paternal chloroplasts in *Nicotiana. Mol. Gen. Genet.* **204,** 195–198.

Menczel, L., Galiba, G., Nagy, F., and Maliga, P. (1982). Effect of radiation dosage on efficiency of chloroplast transfer by protoplast fusion in Nicotiana. *Genetics* **100,** 487–495.

Menczel, L., Polsby, L. S., Steinback, K. E., and Maliga, P. (1986). Fusion-mediated transfer of triazine-resistant chloroplasts: Characterization of *Nicotiana tabacum* cybrid plants. *Mol. Gen. Genet.* **205,** 201–205.

Menczel, L., Morgan, A., Brown, S., and Maliga, P. (1987). Fusion-mediated combination of Ogura-type cytoplasmic male sterility with *Brassica napus* plastids using X-irradiated CMS protoplasts. *Plant Cell Rep.* **6,** 98–101.

Mets, L. (1980). Uniparental inheritance of chloroplast DNA sequences in interspecific hybrids of *Chlamydomonas. Curr. Genet.* **2,** 131–138.

Metzlaff, M., Börner, T., and Hagemann, R. (1981). Variations of chloroplast DNAs in the genus *Pelargonium* and their biparental inheritance. *Theor. Appl. Genet.* **60,** 37–41.

Meyer, B., and Stubbe, W. W. (1974). Das Zahlenverhältnis von mütterlichen und väterlichen Plastiden in den Zygoten von *Oenothera erythrosepala* Borbas (syn. *Oe. lamarckiana*). *Ber. Deutsch. Bot. Ges.* **87,** 29–38.

Miyamura, S., Nagata, T., and Kuroiwa, T. (1986). Quantitative fluorescence microscopy on dynamic changes of plastid nucleoids during wheat development. *Protoplasma* **133,** 66–72.

Mogensen, H. L. (1988). Exclusion of male mitochondria and plastids during syngamy in barley as a basis for maternal inheritance. *Proc. Natl. Acad. Sci. USA* **85,** 2594–2597.

Mogensen, H. L., and Rusche, M. L. (1985). Quantitative ultrastructural analysis of barley sperm. I. Occurrence and mechanism of cytoplasm and organelle reduction and the question of sperm dimorphism. *Protoplasma* **128,** 1–13.

Morgan, A., and Maliga, P. (1987). Rapid chloroplast segregation and recombination of mitochondrial DNA in *Brassica* cybrids. *Mol. Gen. Genet.* **209,** 240–246.

Nagy, G., Lázár, G., Menczel, L., and Maliga, P. (1983). A heteroplasmic state induced by protoplast fusion is a necessary condition for detecting rearrangements in *Nicotiana* mitochondrial DNA. *Theor. Appl. Genet.* **66,** 203–207.

Nakamura, S., and Kuroiwa, T. (1989). Selective elimination of chloroplast DNA by 5-fluorodeoxyuridine causing no effect on preferential digestion of male chloroplast nucleoids in *Chlamydomonas*. *Eur. J. Cell Biol.* **48,** 165–173.

Nakamura, S., Sato, C., and Kuroiwa, T. (1988). Polypeptides related to preferential digestion of male chloroplast nucleoids in *Chlamydomonas*. *Plant Sci.* **56,** 129–136.

Neale, D. B., and Sederoff, R. R. (1988). Inheritance and evolution of conifer organelle genomes. *In* "Genetic Manipulation of Woody Plants" (J. W. Hanover and E. E. Keathley, eds.), pp. 252–264. Plenum Press, New York.

Neale, D. B., and Sederoff, R. R. (1989). Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. *Theor. Appl. Genet.* **77,** 212–216.

Neale, D. B., Wheeler, N. C., and Allard, R. W. (1986). Paternal inheritance of chloroplast DNA in Douglas-fir. *Can. J. For. Res.* **16,** 1152–1154.

Neale, D. B., Marshall, K. A., and Sederoff, R. R. (1989). Chloroplast and mitochondrial DNA are paternally inherited in *Sequoia sempervirens* D. Don Endl. *Proc. Natl. Acad. Sci. USA*, **86,** 9347–9349.

Nemoto, Y., Kawano, S., Nakamura, S., Mita, T., Nagata, T., and Kuroiwa, T. (1988). Studies on plastid-nuclei (nucleoids) in *Nicotiana tabacum* L. I. Isolation of proplastid-nuclei from cultured cells and identification of proplastid-nuclear proteins. *Plant Cell Physiol.* **29,** 167–177.

Nemoto, Y., Nagata, T., and Kuroiwa, T. (1989). Studies on plastid-nuclei (nucleoids) in *Nicotiana tabacum* L. II. Disassembly and reassembly of proplastid-nuclei isolated from cultured cells. *Plant Cell Physiol.* **30,** 445–454.

Nilsson-Tilgren, T., and von Wettstein-Knowles, P. (1970). When is the male plastome eliminated? *Nature* **227,** 1265–1266.

Ogawa, S. (1988). Disappearance of chloroplast nucleoids during male gamete formation in *Bryopsis plumosa* (Hudson) C. Ag. (Chlorophyceae). *Bot. Gaz.* **149,** 25–29.

Ogawa, K., and Kuroiwa, T. (1985a). Nuclease C. Polymorphism of calcium-dependent nucleases in *Chlamydomonas reinhardtii*. *Plant Cell Physiol.* **26,** 481–491.

Ogawa, K., and Kuroiwa, T. (1985b). Destruction of chloroplast nuclei of the male gamete by calcium and nuclease C in a cell model of *Chlamydomonas reinhardtii*. *Plant Cell Physiol.* **26,** 493–503.

Ogawa, K., and Kuroiwa, T. (1986). Base-specific endo-exonucleolytic activity of *Chlamydomonas* nuclease C1&2. *Plant Cell Physiol.* **27,** 701–710.

Ohba, K., Iwakawa, M., Ohada, Y., and Murai, M. (1971). Paternal transmission of a plastid anomaly in some reciprocal crosses of Suzi, *Cryptomeria japonica* D. Don. *Silvae Genet.* **20,** 101–107.

Oross, J. W., and Possingham, J. V. (1989). Ultrastructural features of the constricted region of dividing plastids. *Protoplasma* **150,** 131–138.

Possingham, J. V., and Lawrence, M. E. (1983). Controls to plastid division. *Int. Rev. Cytol.* **84,** 1–56.

Possingham, J. V., Hashimoto, H., and Oross, J. (1988). Factors that influence plastid division in higher plants. *In* "The Division and Segregation of Organelles" (S. A. Boffey and D. Lloyd, eds.), pp. 1–20. Cambridge University Press, Cambridge.

Pyke, K. A., and Leech, R. M. (1987). The control of chloroplast number in wheat mesophyll cells. *Planta* **170,** 416–420.

Reiss, T., and Link, G. (1985). Characterization of transcriptionally active DNA-protein complexes G. from chloroplasts and etioplasts of mustard (*Sinapis alba* L.). *Eur. J. Biochem.* **148,** 207–212.

Ris, H., and Plaut, W. (1962). Ultrastructure of DNA-containing areas in the chloroplast of *Chlamydomonas. J. Cell Biol.* **13,** 383–391.

Robertson, D., Palmer, J. D., Earle, E. D., and Mutschler, M. A. (1987). Analysis of organelle genomes in a somatic hybrid derived from cytoplasmic male-sterile *Brassica oleracea* and atrazine-resistant *B. campestris. Theor. Appl. Genet.* **74,** 303–309.

Rose, R. J. (1988). The role of membranes in the segregation of plastid DNA. *In* "The Division and Segregation of Organelles" (S. A. Boffey and D. Lloyd, eds.), pp. 171–195. Cambridge University Press, Cambridge.

Sager, R. (1972). "Cytoplasmic Genes and Organelles." Academic Press, New York. 405 pp.

Sager, R. (1977). Genetic analysis of chloroplast DNA in *Chlamydomonas. Adv. Gen.* **19,** 287–340.

Sager, R., and Grabowy, C. (1983). Differential methylation of chloroplast DNA regulates maternal inheritance in a methylated mutant of *Chlamydomonas. Proc. Natl. Acad. Sci. USA* **80,** 3025–3029.

Sager, R., and Grabowy, C. (1985). Sex in *Chlamydomonas:* Sex and the single chloroplast. *In* "The Origin and Evolution of Sex." MBL Lectures in Biology, Vol. 7 (H. O. Halvorson and A. Monroy, eds.), pp. 113–121. Alan R. Liss, New York.

Sager, R., and Kitchin, R. (1975). Selective silencing of eukaryotic DNA. *Science* **189,** 426–433.

Sager, R., and Lane, D. (1972). Molecular basis of maternal inheritance. *Proc. Natl. Acad. Sci. USA* **69,** 2410–2413.

Sager, R., and Ramanis, Z. (1967). Biparental inheritance of nonchromosomal genes induced by ultraviolet irradiation. *Proc. Natl. Acad. Sci. USA* **58,** 931–935.

Sager, R., and Ramanis, Z. (1973). The mechanism of maternal inheritance in *Chlamydomonas:* Biochemical and genetic studies. *Theor. Appl. Genet.* **43,** 101–108.

Sager, R., and Ramanis, Z. (1974). Mutations that alter the transmission of chloroplast genes in *Chlamydomonas. Proc. Natl. Acad. Sci. USA* **71,** 4698–4702.

Sager, R., Grabowy, C., and Sano, H. (1981). The *mat-1* gene in *Chlamydomonas* regulates DNA methylation during gametogenesis. *Cell* **24,** 41–47.

Sager, R., Sano, H., and Grabowy, C. T. (1984). Control of maternal inheritance by DNA methylation in *Chlamydomonas. Curr. Top. Microbiol. Immunol.* **108,** 157–172.

Saito, A., Ogawa, S., and Wada, S. (1989). Variations in morphology and DNA distribu-

tion among chloroplasts in mature germling cell of *Bryopsis plumosa* (Hudson) C. Ag. (Chlorophyceae). *Bot. Gaz.* **150,** 25–29.

Schmitz, U. K., and Kowallik, K. V. (1987). Why are plastids maternally inherited in *Epilobium*? Ultrastructural observations during microgametogenesis. *Plant Sci.* **53,** 139–145.

Schröder, M.-B. (1984). Ultrastructural studies on plastids of generative and vegetative cells in the family Liliaceae. 1. *Lilium martagon* L. *Biol. Zbl.* **103,** 547–555.

Schröder, M.-B. (1985a). Ultrastructural studies on plastids of generative and vegetative cells in the family Liliaceae. 2. *Fritillaria imperialis* and *F. meleagris*. *Biol. Zbl.* **104,** 21–27.

Schröder, M.-B. (1985b). Ultrastructural studies on plastids of generative and vegetative cells in Liliaceae. 3. Plastid distribution during the pollen development in *Gasteria verrucosa* (Mill.) Duval. *Protoplasma* **124,** 123–129.

Schröder, M.-B. (1986a). Ultrastructural studies on plastids of generative and vegetative cells in Liliaceae. 4. Plastid degeneration during generative cell maturation in *Convallaria majalis* L. *Biol. Zbl.* **105,** 427–433.

Schröder, M.-B. (1986b). Ultrastructural studies on plastids of generative and vegetative cells in Liliaceae. 5. The behavior of plastids during pollen development in *Chlorophytum comosum* (Thunb.) Jacques. *Theor. Appl. Genet.* **72,** 840–844.

Schröder, M.-B., and Hagemann, R. (1985). Plastid content of generative cells in Liliaceae. *In* "Reproduction in Seed Plants, Ferns, and Mosses" (M. T. M. Willemse and J. L. Van Went, eds.), p. 52. Pudoc, Wageningen.

Schröder, M.-B., and Hagemann, R. (1986). Ultrastructural studies on plastids of generative and vegetative cells in Liliaceae. 6. Patterns of plastid distribution during generative cell formation in *Aloë secundiflora* and *A. jucunda*. *Acta Bot. Neerl.* **35,** 243–248.

Scott, N. S., and Possingham, J. V. (1980). Chloroplast DNA in expanding spinach leaves. *J. Exp. Bot.* **31,** 1081–1092.

Sears, B. B. (1980). Elimination of plastids during spermatogenesis and fertilization in the plant kingdom. *Plasmid* **4,** 233–255.

Sears, B. B. (1983). Genetics and evolution of the chloroplast. *Stadler Symp.* **15,** 119–139.

Sears, B. B., Boynton, J. E., and Gillham, N. W. (1980). The effect of gametogenesis regimes on the chloroplast genetic system of *Chlamydomonas reinhardtii*. *Genetics* **96,** 95–114.

Sellden, G., and Leech, R. M. (1981). Localization of DNA in mature and young wheat chloroplasts using the fluorescence probe 4′,6-diamidino-2-phenylindole. *Plant Physiol.* **68,** 731–734.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., and Sugiura, M. (1986). The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. *EMBO J.* **5,** 2043–2049.

Sidorov, V. A., Menczel, L., Nagy, F., and Maliga, P. (1981). Chloroplast transfer in *Nicotiana* based on metabolic complementation between irradiated and iodoacetate treated protoplasts. *Planta* **152,** 341–345.

Smith, S. E. (1988). Biparental inheritance of organelles and its implications in crop improvement. *Plant Breed. Rev.* **6,** 361–393.

Smith, S. E. (1989). Influence of parental genotype on plastid inheritance in *Medicago sativa*. *J. Hered.* **80,** 214–217.

Soliman, K., Fedak, G., and Allard, R. W. (1987). Inheritance of organelle DNA in barley and *Hordeum* × *Secale* intergeneric hybrids. *Genome* **29,** 867–872.

Spreitzer, R. J., and Chastain, C. J. (1987). Heteroplasmic suppression of an amber muta-

tion in the *Chlamydomonas* chloroplast gene that encodes the large subunit of ribulose-bisphosphate carboxylase/oxygenase. *Curr. Genet.* **11,** 611–616.

Stubbe, W. (1964). The role of the plastome in evolution of the genus *Oenothera*. *Genetica* **35,** 28–33.

Svab, Z., and Maliga, P. (1986). *Nicotiana tabacum* mutants with chloroplast encoded streptomycin resistance and pigment deficiency. *Theor. Appl. Genet.* **72,** 637–643.

Szmidt, A. E., Aldèn, T., and Hällgren, J.-E. (1987). Paternal inheritance of chloroplast DNA in *Larix*. *Plant Mol. Biol.* **9,** 59–64.

Szmidt, A. E., El-Kassaby, Y. A., Sigurgeirsson, A., Aldèn, T., Lindgren, D., and Hällgren, J.-E. (1988). Classifying seedlots of *Picea sitchensis* and *P. glauca* in zones of introgression using restriction analysis of chloroplast DNA. *Theor. Appl. Genet.* **76,** 841–845.

Thanh, N. D., and Medgyesy, P. (1989). Limited chloroplast gene transfer via recombination overcomes plastome–genome incompatibility between *Nicotiana tabacum* and *Solanum tuberosum*. *Plant Mol. Biol.* **12,** 87–93.

Thanh, N. D., Páy, A., Smith, M. A., Medgyesy, P., and Márton, L. (1988). Intertribal chloroplast transfer by protoplast fusion between *Nicotiana tabacum* and *Salpiglossis sinuata*. *Mol. Gen. Genet.* **213,** 186–190.

Tilney-Bassett, R. A. E. (1975). Genetics of variegated plants. *In* "Genetics and Biogenesis of Mitochondria and Chloroplasts" (C. W. Birky, Jr., P. S. Perlman, and T. J. Byers, eds.), pp. 268–308. Ohio State University Press, Columbus.

Tilney-Bassett, R. A. E. (1988). Inheritance of plastids in *Pelargonium*. *In* "The Division and Segregation of Organelles" (S. A. Boffey and D. Lloyd, eds.), pp. 115–129. Cambridge University Press, Cambridge.

Tilney-Bassett, R. A. E., and Almouslem, A. B. (1989). Variation in plastid inheritance between *Pelargonium* cultivars and their hybrids. *Heredity* **63,** 145–153.

Tsubo, Y., and Matsuda, Y. (1984). Transmission of chloroplast genes in crosses between *Chlamydomonas reinhardtii* diploids: Correlation with chloroplast nucleoid behavior in young zygotes. *Curr. Genet.* **8,** 223–229.

Van Went, J. L. (1984). Unequal distribution of plastids during generative cell formation in *Impatiens*. *Theor. Appl. Genet.* **68,** 305–309.

VanWinkle-Swift, K. P. (1978). Uniparental inheritance is promoted by delayed division of the zygote in *Chlamydomonas*. *Nature* **275,** 749–751.

VanWinkle-Swift, K. P. (1980). A model for the rapid vegetative segregation of multiple chloroplast genomes in *Chlamydomonas:* Assumptions and predictions of the model. *Curr. Genet.* **1,** 113–125.

VanWinkle-Swift, K. P., and Aubert, B. (1983). Uniparental inheritance in a homothallic alga. *Nature* **303,** 167–169.

VanWinkle-Swift, K. P., and Hahn, J.-H. (1986). The search for mating-type-limited genes in the homothallic alga *Chlamydomonas monoica*. *Genetics* **113,** 601–619.

VanWinkle-Swift, K. P., and Salinger, A. P. (1988). Loss of mt^+-derived zygotic chloroplast DNA is associated with a lethal allele in *Chlamydomonas monoica*. *Curr. Genet.* **13,** 331–337.

Vardi, A., Breiman, A., and Galun, E. (1987). *Citrus* cybrids: Production by donor–recipient protoplast-fusion and verification by mitochondrial-DNA restriction profiles. *Theor. Appl. Genet.* **75,** 51–58.

Vaughn, K. C. (1981). Organelle transmission in higher plants: Organelle alteration vs. physical exclusion. *J. Hered.* **72,** 335–337.

Vaughn, K. C., DeBonte, L. R., Wilson, K. G., and Schaeffer, G. W. (1980). Organelle alteration as a mechanism for maternal inheritance. *Science* **208,** 1916–1918.

Wagner, D. B., Furnier, G. R., Saghai-Maroof, M. A., Williams, S. M., Dancik, B. P., and

Allard, R. W. (1987). Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. *Proc. Natl. Acad. Sci. USA* **84,** 2097–2100.

Whatley, J. M. (1982). Ultrastructure of plastid inheritance: Green algae to angiosperms. *Biol. Rev.* **57,** 527–569.

Whatley, J. M. (1988). Mechanisms and morphology of plastid division. *In* "The Division and Segregation of Organelles" (S. A. Boffey and D. Lloyd, eds.), pp. 63–83. Cambridge University Press, Cambridge.

Whiteway, M. S., and Lee, R. W. (1977). Chloroplast DNA content increases with nuclear ploidy in *Chlamydomonas. Mol. Gen. Genet.* **157,** 11–15.

Winter, P., and Herrmann, R. G. (1988). A five-base-pair-deletion in the gene for the large subunit causes the lesion in the ribulose bisphosphate carboxylase/oxygenase-deficient plastome mutant sigma of *Oenothera hookeri. Bot. Acta* **101,** 68–75.

Wurtz, E. A., Boynton, J. E., and Gillham, N. W. (1977). Perturbation of chloroplast DNA amounts and chloroplast gene transmission in *Chlamydomonas reinhardtii* by 5-fluorodeoxyuridine. *Proc. Natl. Acad. Sci. USA* **74,** 4552–4556.

Zachleder, V., and Cepák, V. (1987). Variations in chloroplast nucleoid number during the cell cycle in the alga *Scenedesmus quadricauda* grown under different light conditions. *Protoplasma* **141,** 74–82.

Zachleder, V., Kuptsova, E. S., Los, D. A., Cepák, V., Kubin, S., Shapizugov, J. M., and Semenenko, V. E. (1989). Division of chloroplast nucleoids and replication of chloroplast DNA during the cell cycle of *Dunaliella salina* grown under blue and red light. *Protoplasma* **150,** 160–167.

CHAPTER 4

Replication and Transcription of Plastid DNA

Lawrence Bogorad

Department of Cellular and Developmental Biology
The Biological Laboratories
Harvard University
Cambridge, Massachusetts

I. INTRODUCTION

This chapter is concerned with the replication and transcription of DNA in plastids. It covers steps in nucleic acid metabolism that occur between DNA transmission (Chapter 3A) and RNA processing (Chapter 5A).

II. DNA REPLICATION

Each plastid contains perhaps 10–50 copies of the mainly 120–200 kbp chromosome (Chapter 2A). Chiang and Sueoka (1967) determined that chloroplast DNA replicates twice during the light period in synchro-

nously grown (12 hr light:12 hr dark) *Chlamydomonas reinhardtii* cells, whereas nuclear DNA undergoes two cycles of replication toward the end of the dark period, during which two cell divisions occur.

Replication of maize and pea plastid DNAs (plDNAs) starts with the formation of two displacement loops (D-loops) and continues by both Cairns and rolling circle mechanisms (Kolodner and Tewari, 1975a,b). However, at the time the D-loops were first described, there were no physical maps on which they could be located; the first detailed restriction map of a plDNA accurately showing the positions of two large inverted repeated segments as well as the locations of the genes for ribosomal RNAs (rRNAs) within them were published only in 1976 and 1977 (Bedbrook and Bogorad, 1976; Bedbrook *et al.*, 1977).

D-loops were first mapped in relation to known sequences on *Euglena* plDNA (Koller and Delius, 1982; Schlunegger *et al.*, 1983; Schlunegger and Stutz, 1984; Ravel-Chapuis *et al.*, 1982), which contains three adjacent copies of genes for 16S and 23S rRNAs plus a "supplementary" copy of the gene for the 16S rRNA. DNA replication is estimated to start almost 5–6.8 kb upstream of the 5′ end of this 16 rRNA gene. This site is close to one of the two short inverted repeated sequences that flanks an AT-rich region in a *Bgl*I fragment that contains the "Z region"; this region varies in length from chromosome to chromosome (Jenni and Stutz, 1979; Jenni *et al.*, 1981; Schlunegger and Stutz, 1984). Koller and Delius (1982) concluded that in most cases DNA synthesis proceeds unidirectionally for about "5000 nucleotides before it starts in the other direction—through the Z region and the second palindromic sequence." Wu and her colleagues (Waddel *et al.*, 1984; Wang *et al.*, 1984) identified and mapped two D-loop regions approximately 6.4 kbp apart on the *C. reinhardtii* chloroplast chromosome; one of these regions was subsequently sequenced (Wu *et al.*, 1986). D-loops have also been mapped on pea plDNA (Meeker *et al.*, 1988).

Another approach to identifying possible replication origins on chloroplast chromosomes has been by assaying individual cloned plDNA fragments for their capacities to serve as templates for DNA synthesis *in vitro*. Using a partially purified pea chloroplast DNA polymerase preparation (McKown and Tewari, 1984), analyses of 94% of the maize chloroplast chromosome revealed two regions to be particularly active in this assay (Gold *et al.*, 1987). One is near the *psbA* gene, close to the edge of one of the large inverted repeats that carries rRNA cistrons. The other, stronger, template is located on the opposite side of the chromosome within a 1,368-bp fragment that includes the coding sequence for ribosomal protein L16. The latter region corresponds in sequence to that identified in the D-loop on *C. reinhardtii* plDNA by Wu *et al.* (1986).

Subsequent analyses further delimited this presumptive maize replication origin to within a region of 455 bp (Carrillo and Bogorad, 1988). The latter work also showed that sequences surrounding this region have a strong positive influence on the effectiveness of the template, although replication does not initiate outside of this region.

Finally, Meeker *et al.* (1988) compared the locations of D-loops on pea plDNA with template activity of cloned fragments of most of the chromosome. They identified two D-loops: one within the spacer region between the 16S and 23S rRNA genes and the second downstream of the 23S rRNA gene. Each of these two sequences was highly active as a template for DNA synthesis *in vitro*.

Despite the data described, no primary replication initiation sequence has been defined (if it exists). The closest coincidence is in 145 bp of the 455-bp maize plDNA fragment, which, in part, corresponds in sequence to a *Chlamydomonas* D-loop region and within which DNA synthesis begins *in vitro* (and proceeds bidirectionally); this fragment includes the 3′ terminal sequence encoding chloroplast ribosomal protein L16. Analysis of maize *Eco*RI fragment *x* shows that a bend site exists at or just beyond the 3′ end of the *rpl 16* coding region (Carrillo and Bogorad, 1988). After additional possible replication origins of plDNA have been identified, sequenced, and analyzed for higher orders of structure, judging whether these are or are not specific replication initiation sequences or structures common to plDNAs will be easier. Also, now that foreign DNA can be introduced into chloroplasts of *C. reinhardtii* (Blowers *et al.*, 1989), dissecting the origins of replication functionally *in vivo* may well be possible.

Additional studies of plastid DNA polymerase and the replication process it catalyzes should elucidate other features of plastid DNA replication. McKown and Tewari (1984) reported that a 90-kDa polypeptide is consistently present in pea chloroplast DNA polymerase preparations. Activity staining of polypeptides of DNA polymerase preparations from maize showed the presence of a polypeptide of about the same size upon electrophoresis in polyacrylamide gels (Carrillo and Bogorad, 1988). Nielsen and Tewari (1988) reported that addition of pea chloroplast topoisomerase I greatly stimulates plastid DNA synthesis catalyzed by partially purified pea plastid DNA polymerase.

Studies of albino mutants of wheat and barley that have deletions in plDNA suggest the locations of some possible origins of replication in those species. As much as 80% of the chloroplast genome can be deleted without impairing the replication of mutant chromosomes, suggesting that enzyme(s) of DNA replication may be entirely nuclear. A region in wheat plDNA that must be retained for replication is the large single-

copy portion of the chromosome containing the *psbA*, *psbC*, and *psbD* genes (Day and Ellis, 1984, 1985); this may include one or more replication origins.

At present, we know little about plDNA replication, but the tools are available for greatly advancing our understanding of the process.

III. THE TRANSCRIPTIONAL APPARATUS OF PLASTIDS

A. Plastid RNA Polymerase

1. Isolation, Subunits, Genes

The discovery of DNA and of 70S ribosomes in plastids in the early 1960s led to the expectation that these organelles could synthesize RNA from DNA templates. Indeed, isolated chloroplasts were shown to synthesize RNA (Kirk, 1964), and RNA polymerase activity was found to be associated with thylakoid fragments rather than the stroma (Bogorad, 1968; Tewari and Wildman, 1969). DNA strands appear to stretch from thylakoid to thylakoid in chloroplast nucleoids (e.g., Kislev *et al.*, 1965); the RNA polymerase is probably associated with the thylakoids by being on the DNA strands.

The maize enzyme can be solubilized by incubating thylakoids at 37°C (Bottomley *et al.*, 1971) with or without 0.5 *M* KCl (Smith and Bogorad, 1974). Endogenous DNase probably digests the DNA during the incubation, thus liberating DNA fragments with associated RNA polymerase molecules inasmuch as solubilization of the polymerase can be aided by adding DNase to chloroplast preparations (e.g., spinach: Briat and Mache, 1980). On the other hand, Polya and Jagendorf (1971a,b) solubilized the polymerase from wheat plastid fragments with 1 *M* KCl. One of the above procedures or combinations of them have also been used to make crude chloroplast RNA polymerase extracts (sometimes designated as an *in vitro* chloroplast transcription system) from chloroplasts of spinach (Briat and Mache, 1980; Lerbs *et al.*, 1983; Greenberg *et al.*, 1984; Gruissem *et al.*, 1983a,b; Orozco *et al.*, 1985), mustard (Link, 1984), *Phaseolus* (Ness and Woolhouse, 1980), *Euglena* (Narita *et al.*, 1985), and peas (Tewari and Goel, 1983; Boyer and Mullet, 1986). Extracts can be readily enriched for the enzyme and freed of DNA by diethylaminoethanol (DEAE) chromatography. Polymerase activity has generally been enriched further by combinations of additional protein

purification methods including ammonium sulfate and/or polyethylene glycol precipitation and glycerol gradient centrifugation. The maize plastid enzyme has been estimated to have a molecular mass of about 500 kDa from density gradient centrifugation, and Smith and Bogorad (1974) reported that the most highly purified enzyme preparations contain polypeptides with apparent molecular masses of 180, 140, 100, 95, 85, 78, and 40 kDa. The major polypeptides in a more recent highly purified preparations of RNA polymerase from maize chloroplasts have apparent masses of 180, 120, 85, 78, 64, 61, 55 and 38 kDa (Hu and Bogorad, 1990 and J. Hu, R. F. Troxler and L. Bogorad, unpublished). Tewari and Goel (1983) concluded that "polypeptides of 180, 140, 110, 95, 65, 47 and 27 kDa probably constitute the RNA polymerase" of pea chloroplasts. In contrast, Briat *et al.* (1987) reported that a highly purified spinach chloroplast RNA polymerase preparation contained polypeptides of 150, 110, 102, 80, 75, and 38 kDa. This array corresponds, in general, to the components of preparations previously described by Lerbs *et al.* (1983, 1985). Apparent size differences of polypeptides can easily result from variations in the conditions used for gel electrophoresis and the molecular weight standards that are used in the analyses, but, apparently, based on the sequences of genes discussed below, the differences in molecular masses of the largest polypeptides of maize and spinach, 180 and 150 kDa, respectively, is real.

The *Escherichia coli* RNA polymerase core enzyme is a complex of two α subunits (37 kDa each), one β subunit (151 kDa), and one β' subunit (155 kDa). The *E. coli* genes encoding the α, β, and β' subunits are designated *rpoA*, *rpoB*, and *rpoC*, respectively. Genes with some homology to these have been mapped on chloroplast chromosomes of tobacco (Shinozaki *et al.*, 1986), *Marchantia* (Ohyama *et al.*, 1986), spinach (Hudson *et al.*, 1988), maize (Ruf and Kossel, 1988; Hu and Bogorad, 1990; Igloi *et al.*, 1990), and rice (Hiratsuka *et al.*, 1989) and have been sequenced. The plastid genes were named to reflect homologies with the *E. coli* genes, although their protein products had not been detected at the time.

Purton and Gray (1989) reported that an antibody against a pea chloroplast *rpoA*–Lac Z fusion protein recognizes a 43-kDa polypeptide in transcriptionally active pea chloroplast extracts. Hu and Bogorad (1990) used a different approach to investigate relationships between chloroplast *rpo* genes and components of the transcriptional apparatus. They found that the N-terminal amino acid sequence of the 38-kDa polypeptide in a maize plastid RNA polymerase preparation corresponds to that predicted by the sequence of the maize *rpoA* gene (Ruf and Kassel, 1988); the N-terminal amino acid sequence of a 180-kDa polypeptide in the preparation was identical to that deduced from the DNA sequence

of maize *rpo*C_2; and the N-terminal amino acid sequence of the 120-kDa polypeptide in the preparation corresponded to the sequence deduced from *rpoB*. Furthermore, amino acid residues two through nine of a 78-kDa polypeptide were found to correspond to that expected from the DNA sequence of maize *rpo*C_1 (J. Hu, R. F. Troxler, and L. Bogorad, unpublished). Thus, on one hand, the products of the maize plastid *rpo*C_1, *rpoB*, *rpo*C_2, and *rpoA* genes have been identified as the 180-, 120, 78-, and 38-kDa polypeptides in maize RNA polymerase preparations, respectively; on the other hand, the probability that these polypeptides are indeed components of the plastid RNA polymerase is strongly supported by the partial homologies of their genes to sequences of *E. coli rpo* genes. In the highly purified maize polymerase preparations, the polypeptides in the 61 kDa band are the α and β subunits of the ribulose bisphosphate carboxylase binding protein; the 55 kDa band contains the large subunit of ribulose bisphosphate carboxylase (J. Hu, R. F. Troxler, and L. Bogorad, unpublished).

Aside from various deletions and insertions, the plastid gene designated *rpoA* would encode a polypeptide about 26–28% homologous to some regions of the *E. coli* α subunit (spinach: Sijben-Muller *et al.*, 1986; Hudson *et al.*, 1988; maize: Ruf and Kossel, 1988). There are also parts of the *E. coli* and plastid *rpoA* genes—insertions and deletions—that are very dissimilar. Comparing plastid genes with one another, the deduced amino acid sequences of the products of the maize and liverwort *rpoA* genes, for example, are 51.8% homologous to one another, whereas the maize, spinach, and tobacco genes are about 68% homologous (Ruf and Kossel, 1988).

The *rpoB* gene of the maize plDNA would encode a basic protein of 122 kDa that would be about 37.5% homologous to the 151-kDa acidic *E. coli* β subunit (Igloi *et al.*, 1990). The largest single deletion in the maize and other plastid *rpoB* genes is a block of 103–105 amino acids near the carboxy terminus. This domain has been deleted without effect on *E. coli* polymerase function (Glass *et al.*, 1986). Igloi *et al.* (1990) identified a nucleotide-binding domain in the protein predicted by the *rpoB* sequence and a zinc finger-type of sequence in the protein expected from the maize *rpo*C_1.

Two separate chloroplast genes contain sequences that would encode polypeptides homologous to parts of the 155-kDa *E. coli* β' subunit. The *rpo*C_1 genes of maize, rice, spinach, tobacco, and *Marchantia* plastids encode polypeptides of 677–684 amino acids at the N-terminal fragment of the *E. coli* β' gene (*Marchantia:* Ohyama *et al.*, 1986; tobacco: Shinozaki *et al.*, 1986; spinach: Hudson *et al.*, 1988; maize: Igloi, 1990; rice: Hiratsuka *et al.*, 1989). The chloroplast gene designated *rpo*C_2, which con-

tains some sequences that encode regions homologous to parts of the 829 C-terminal amino acids of the *E. coli* β' subunit, is especially interesting. More than 500 of the amino acids encoded in plastid *rpoC*$_2$ genes are absent from the corresponding C-terminal portion of the bacterial gene. These can be accounted for by the presence in the plant genes of an additional single block of 626 amino acids and the absence from these genes of a block of 40 amino acids present in the *E. coli* gene. Overall, although the regions of homology between the *E. coli* and plastid genes are unmistakable, differences are substantial in the predicted protein subunits; therefore, differences in some enzymatic activities would not be surprising.

Superimposed on the features common to known plastid *rpoC*$_2$ genes that distinguish them from the 5′ portion of the corresponding *E. coli* gene are differences among the plastid genes themselves. Inserted into the *rpoC*$_2$ genes of rice and maize, but absent from those of tobacco, spinach, and *Marchantia*, is a block of approximately 450 bp of DNA-encoding acidic heptameric repeat units containing stretches of glutamyl, tyrosyl, and leucyl residues with regular spacing (Igloi *et al.*, 1990). Another major difference among *rpoC*$_2$ genes is their size. The *rpoC*$_2$ of maize encodes a protein of 1,527 amino acids and that of rice one of 1,513 amino acids in contrast to the 1,391 to 1,361 and 1,386 amino acids of the tobacco, spinach, and liverwort genes, respectively. Analyses of additional *rpoC*$_2$ genes should establish whether this large insertion is a specific feature of grasses or of some smaller or larger taxonomic group. The molecular masses of the protein products of the maize and rice *rpoC*$_2$ genes deduced from the DNA sequences would be about 175 kDa, and the products of the spinach, tobacco, and liverwort *rpoC*$_2$ genes would be 156 kDa. Thus, the observed difference in the sizes of the largest polypeptides in maize versus spinach plastid RNA polymerase preparations—180 versus 150 kDa—is not an artifact of analyses, and these polypeptides are both encoded by *rpoC*$_2$ genes. Except for the large insertion common to rice and maize, all the *rpoC*$_2$ products are quite homologous.

One of the most intriguing differences between the protein products of tobacco–liverwort–spinach versus maize–rice *rpoC*$_2$ is the acidic nature of the heptameric repeats within the 150-amino acid insertion in the maize–rice proteins. Yeast and other eukaryotic activator proteins also have acidic regions that are necessary for activity (Johnson and McKnight, 1989). The presence of such regions in *rpoC*$_2$ genes raises questions about possible functional differences between the polymerases of maize and rice versus those of tobacco, spinach, and *Marchantia*. The origin and evolution of the approximately 450-bp insert is yet another

subject of great interest. If the large differences in *rpoC*$_2$ genes between those of rice and maize versus those of *Marchantia*, tobacco, and spinach are indeed limited to monocotyledonous plants or some subset of them, this would be the first plastid genome feature identified as being introduced at this taxonomic level.

2. Multiple Plastid RNA Polymerases?

It has long been suggested that plastids might contain more than one kind of DNA-dependent RNA polymerase. Dark-grown maize seedlings incorporate very little ^{32}P-phosphate into plastid rRNAs, but upon illumination they do so rapidly (Bogorad, 1968; Bogorad and Woodcock, 1970). The synthesis of plastid rRNA in maize is blocked by the antibiotic rifampicin (Bogorad and Woodcock, 1970), and chloroplast rRNA synthesis has also been reported to be inhibited by rifampicin in *Euglena* (Brown *et al.*, 1979), *Acetabularia* (Brandle and Zetsche, 1971), the filamentous brown alga *Pylaiella littoralis* (Loiseaux *et al.*, 1975), *Marchantia polymorpha* (Loiseaux *et al.*, 1975), and segments of *Nicotina rustica* leaves (Munsche and Wollgiehn, 1973). However, transcription of calf thymus DNA added to fragments of maize chloroplasts capable of transcription *in vitro* (Bogorad and Woodcock, 1970) was inhibited only slightly by rifampicin (a maximum of about 30% at very high concentrations of the antibiotic under special conditions), and all subsequent analyses of solubilized and bound plastid RNA polymerases from higher plants and *Euglena* (e.g., Polya and Jagendorf, 1971a,b; Bottomley *et al.*, 1971; Smith and Bogorad, 1974; Briat and Mache, 1980; Hallick *et al.*, 1976; Briat *et al.*, 1979; Kidd and Bogorad, 1980; Gruissem *et al.*, 1983a,b; Tewari and Goel, 1983; Lerbs *et al.*, 1983; Sarhan and Chevrier, 1984; Orozco *et al.*, 1985) have failed to reveal a species of plastid RNA polymerase sensitive to rifampicin *in vitro*. The discrepancy between the effect of rifampicin on maize seedling plastid rRNA synthesis and the failure of the antibiotic to inhibit plastid RNA polymerase activity of maize plastid fragments *in vitro* led to the suggestion (Bogorad and Woodcock, 1970) that maize plastids may contain at least two types of RNA polymerase activity: rifampicin-sensitive and -resistant forms. *Chlamydomonas reinhardtii* may be exceptional; Surzycki (1969; also Ratcliff and Surzycki, 1974) has identified a rifampicin-sensitive RNA polymerase fraction in *C. reinhardtii* extracts.

To explain differential solubilization of RNA polymerase activities and, in a few cases, differences in template preference of fractions, multiple RNA polymerases in plastids have been suggested. Bottomley *et al.* (1971) found that their "crude soluble RNA polymerase" fraction of

maize plastids contained only about 60% of the total activity. Purification and identification of proteins in the soluble fraction have been described above (Jolly and Bogorad, 1980; Hu and Bogorad, 1990) and will also be addressed later. Joussaume (1973) proposed that two forms of DNA-dependent RNA polymerase may exist in chloroplasts of pea leaves: RNA polymerase activity that is liberated first by simply breaking plastids and an activity that is solubilized from thylakoids by Triton X-100. The two activities were found to differ in their pH optima and some other properties. Hallick *et al.* (1976) designated their *Euglena* chloroplast RNA polymerase complexed with proteins and chloroplast DNA the transcriptionally active complex (TAC). It is reported to transcribe almost exclusively rRNA genes (Rushlow *et al.*, 1980). Hallick and his coworkers (Gruissem *et al.*, 1983a,b; Greenberg *et al.*, 1984; Narita *et al.*, 1985) subsequently solubilized chloroplast DNA-dependent RNA polymerase activity from the *Euglena* complex, but some TAC remained. Based on differences in properties, they concluded that two distinct DNA-dependent RNA polymerases exist in *Euglena* chloroplasts: one that remains in the TAC and transcribes only rDNA and an enzyme that can be solubilized that transcribes transfer RNA (tRNA) genes. The solubilized enzyme was found to be sensitive to heparin, but TAC activity was not. In addition, the two fractions differed in their salt optima. TACs can also be prepared from chloroplasts of other plants by the same procedures, but these reportedly have broader transcription ranges than the *Euglena* TAC (Briat *et al.*, 1979; Blanc *et al.*, 1981; Reiss and Link, 1985; LeBrun *et al.*, 1986) and, in contrast to the soluble *Euglena* polymerase, Sun *et al.* (1986) and Briat *et al.* (1987) have shown that purified soluble pea and spinach chloroplast RNA polymerases can transcribe cloned homologous 16S rRNA genes very actively.

Narita *et al.* (1985) report that the major polypeptides in *Euglena* TAC have molecular masses of 118 and 85 kDa. Bulow *et al.* (1987) found that the most prominent proteins in a mustard TAC preparation were 81, 57, 19, and 17 kDa. Briat and Mache (1980) found about 30 polypeptides in a spinach chloroplast TAC and concluded that the RNA polymerase activity seemed to be due to five polypeptide subunits: 69, 60, 55, 34, and 15 kDa. The only strong claim for specificity of a TAC is for transcription of *Euglena* rDNAs. There is, as yet, no agreement as to the sizes or number of major polypeptides that might be associated with TAC polymerase activity. Further physical characterization of the active part of this complex is necessary to be sure that it is a different enzyme.

What constitutes "a different RNA polymerase"? The best characterized plastid RNA polymerase is the soluble enzyme discussed in detail above, four of whose subunits are encoded by the plastid genes *rpoA*,

rpoB, *rpoC*$_1$, and *rpoC*$_2$. A "different RNA polymerase" could be one with an entirely different set of subunits, or a polymerase with the same basic set of polypeptides plus one or two additional polypeptides (e.g., a basic set of polypeptides plus the equivalent of bacterial σ factors), or even the same set of polypeptides with one of them modified. Whether a TAC is fundamentally the soluble enzyme that exhibits some different properties because of the way it is associated with the template, or TAC RNA polymerases are entirely different enzymes, or even TACs prepared from different plant species are the same remain to be worked out.

Evidence for a second RNA polymerase activity in maize plastid extracts that transcribes from relaxed templates *in vitro* (Zaitlin *et al.*, 1989) is discussed below.

3. Template Conformation Affects Transcription

Jolly and Bogorad (1980) found that the major peak of maize plastid RNA polymerase activity eluting from a DEAE-cellulose column with KCl (designated the peak fraction (PF) activity) was at the same position when the template was either calf thymus DNA or a negatively supercoiled chimeric plasmid containing a cloned maize plastid DNA sequence. However, the maximum activity was much lower with the latter. A column fraction immediately following (and in some preparations overlapping) the major peak of activity contained a polypeptide of 27.5 kDa, the S factor, which stimulated transcription of the chimeric plasmid 5- to 15-fold but did not change the rate of transcription of calf thymus DNA. Furthermore, the enzyme in the main peak of activity transcribed the plasmid vector sequences and cloned plastid DNA sequences about equally, whereas when the 27.5-kDa S polypeptide was added, the cloned plastid DNA sequences were transcribed about ninefold more actively than were vector sequences; however, almost all the preferential transcription of plastid DNA sequences was lost when a relaxed rather than a negatively supercoiled template was provided, thus demonstrating the requirement for a supercoiled template. Along the same line, Sun *et al.* (1986) found that pea plastid RNA polymerase specifically transcribes cloned pea 16S rDNA from negatively supercoiled but not from linear templates. When a maize RNA polymerase preparation with the S polypeptide present was provided with a supercoiled template carrying the divergently transcribed *rbcL* and *atpBE* genes of the maize plastid chromosome, the transcript ratio for the two genes was 3 *rbcL*:1 *atpBE*, showing that the enzyme also discriminates between these two plastid genes *in vitro* (Jolly *et al.*, 1981). Highly purified maize plastid enzyme initiates transcription of the *rbcL* gene from

the supercoiled template at the same site as occurs *in vivo* (Crossland *et al.*, 1984; Hu and Bogorad, 1990; J. Hu and L. Bogorad, unpublished), but it has been reported that for spinach only crude extracts of chloroplasts do so (Gruissem, 1989); however, no evidence indicates that this property is lost upon purification of the spinach enzyme. In fact, a number of experiments in which transcription initiation is assumed to be accurate have been conducted with spinach RNA polymerase further purified by DEAE chromotography (Orozco *et al.*, 1985; Gruissem and Zurawski, 1985a,b).

In the maize plastid RNA polymerase experiments described above, transcription rates and specificity of transcription from completely relaxed templates were compared with those from negatively supercoiled templates at σ values (a measure of superhelical density) of approximately -0.06, i.e., the form in which they were isolated from *E. coli*. A much more detailed analysis of transcription as a function of template topology was carried out by Stirdivant *et al.* (1985). For use in these experiments, maize chloroplast RNA polymerase was purified by DEAE chromatography and gel filtration to obtain enzyme preparations that did not alter (i.e., relax or nick) supercoiled templates. Enzyme activity was measured with relaxed templates and with a series of negatively supercoiled topoisomers of an *E. coli* plasmid carrying a cloned segment of the maize chloroplast containing the divergently transcribed *rbcL* and *atpBE* genes (Stirdivant *et al.*, 1985). These experiments demonstrated the following. (1) Transcription of the adjacent, divergently transcribed maize plastid genes *rbcL* or *atpBE* from a relaxed template carrying both genes cloned into pBR322 is, at most, about 1/50 of the maximum transcription rate from an appropriately negatively supercoiled template (transcription of pBR322 sequences by this enzyme is very low [Jolly and Bogorad, 1980]). (2) The negative superhelical density of the template required for the maximum rate of transcription is different for the two genes (it is also different for maize *psbA* (unpublished) and maize *rps4* [Russell and Bogorad, 1987] but in general falls between σ values of about -0.05 to -0.08). (3) The transcription/$-\sigma$ profiles for *rbcL* and *atpB* are strikingly different, and relatively small differences in conformation of the template carrying the two divergently transcribed genes strongly affect the relative transcription of the two genes. At σ -0.022, the ratio of transcription of *rbcL*/*atpBE* is about 5:1; at -0.160, it is about 1:2.4.

Kolodner *et al.* (1975b, 1976) observed supercoiled DNA molecules in preparations from chloroplasts. They found that chloroplast DNAs from leaves of peas, spinach, and lettuce had σ values of -0.087 to -0.089. Their observations implied that plastids contain gyrase and that tran-

scription *in vivo* is most likely to be from negatively supercoiled templates.

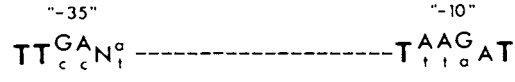
The observation that chloroplast RNA polymerase transcribes cloned plDNA genes well from negatively supercoiled templates DNA and poorly from relaxed templates (Jolly and Bogorad, 1980; Crossland *et al.*, 1984; Stirdivant *et al.*, 1985; Russell and Bogorad, 1987) was used by Lam and Chua (1987) to demonstrate, indirectly, that chloroplast extracts active in transcription contain gyrase, as was to be expected from the observations of Kolodner and Tewari (1976) that indicated plastid DNAs are negatively supercoiled. Novobiocin, an inhibitor of gyrase, was found to reduce transcription from supercoiled templates in chloroplast extracts (Lam and Chua, 1987). Thompson and Mosig (1985) showed that extracts of *Chlamydomonas* contain gyrase and that novobiocin and naldixic acid affect transcription of some segments of plastid DNA *in vivo*. Pyke *et al.* (1989) reported that wheat chloroplast extracts contain proteins of 96–101 kDa that react with antibody to yeast topoisomerase II. Chloroplasts contain topoisomerase I (Siedlecki *et al.*, 1983; Nielsen and Tewari, 1988). Except for the observations of Thompson and Mosig (1985, 1987, 1988), no direct evidence indicates that changes in the conformation of the chloroplast chromosome affects transcription of plastid genes *in vivo*, but it would not be surprising to find that it does in view of the properties of the enzyme *in vitro*.

The σ values of three chloroplast DNAs determined by Kolodner *et al.* (1976) were for chromosomes stripped of all proteins. The torsional stress *in vivo* is very likely to be lower because of the proteins—including RNA polymerase—that are associated with the chromosome *in situ*. For example, the *E. coli* chromosome with associated proteins has a σ value of approximately -0.025 *in vivo* (Bliska and Cozzarelli, 1987) in contrast to the σ value of -0.065 observed for Form I pASA21 plasmid DNA isolated from *E. coli* cells and freed of proteins. Proteins associated with DNA *in vivo* reduce the negative superhelical density. On the other hand, when Borowiec and Gralla (1987) studied the rate of open complex formation with *E. coli* RNA polymerase using four variants of the *LAC* p^s promoter, the rate was highest with DNA at a σ value of approximate -0.05. However, inasmuch as the *E. coli* chromosome has a σ value of approximately -0.025 *in vivo*, if other *E. coli* promoters also function poorly at such a value, how does the polymerase work well in the cell? Perhaps some proteins bind to the DNA and alter the conformation locally. The σ value of protein-free DNA versus the optima for transcription *in vitro* is somewhat different for plastids. The superhelicity of plDNA *in vivo* is not known but the sigma values of -0.087 to -0.089 of the three naked plastid chromosomes studied by Kolodner

et al. (1976) show them to be more negatively supercoiled than isolated *E. coli* DNA, which has a σ value of about -0.065. Judging from the situation in *E. coli*, the plDNA is most likely at much less negative σ values than about -0.09 *in vivo*. On the other hand, the four maize plastid genes studied—*rbcL*, *atpBE*, *rps4*, and *psbA*—are most actively transcribed *in vitro* by maize plastid RNA polymerase from templates of σ values of about -0.05 to -0.08 (Stirdivant *et al.*, 1985; Russell and Bogorad, 1987); the negative superhelicity of all or parts of the chromosomes may be in these ranges *in vivo*. The conformation of the plastid chromosome is set by the ratio of topoisomerase I to gyrase activities together with the amounts and natures of DNA-binding proteins (including RNA polymerase). DNA-binding proteins could affect conformation either generally or specifically and locally, which, in turn, could influence transcription of individual genes or groups of genes. To understand the transcriptional apparatus and processes in plastids, information is needed about DNA-associating proteins as well as about RNA polymerase. Current knowledge of such proteins is discussed in a later section of this chapter.

4. Transcription Initiation

A proposed canonical promoter sequence on one DNA strand for *E. coli* genes transcribed by the RNA polymerase holoenzyme with σ 70 is 5′-TTGACA . . . 17 nt . . . TATAAT-3′ (Reznikoff *et al.*, 1985). The 3′ end of the -10 hexanucleotide TATAAT lies seven nucleotides upstream of the transcription initiation site. The other conserved hexanucleotide, TTGACA, is centered at about 35 bp upstream of the start point. The spacing between the -10 and -35 regions is somewhat variable, but a separation by 17 bp seems to be favored. Based simply on comparisons of sequences 5′ to the probable—but undemonstrated—transcription start sites of the small number of plastid genes whose sequences were known at that time, Steinmetz *et al.* (1983) identified a pattern of conserved sequences resembling those of *E. coli* promoters associated with these plastid genes. The "consensus sequences" for the "-35" and "-10" elements deduced from this analysis were:



Bold uppercase letters represent nucleotides that are present in >90% of the sequences; uppercase letters are nucleotides that occur at this position in 50–90% of the cases, and lowercase letters represent nucleotides

present in 20–50% of the cases. N denotes a position at which none of the four nucleotides is present at a frequency >40%. Kung and Lin (1985) carried out a similar analysis.

Link (1984) first studied the effects of deletions in a presumed promoter region on transcription *in vitro*. He used the mustard chloroplast gene *psbA* and a mustard chloroplast extract. Link and Langridge (1984) had determined that the conserved sequences upstream of the transcription initiation site of the mustard *psbA* gene is 5′-TTGACA . . . 18 nt . . . TATACT-3′. A clone of the *psbA* gene, starting 30 nucleotides upstream of the first T in the "−35" region, i.e., at about −65, was transcribed actively *in vitro*. Transcription was not as strong from a deletion in which the first two T's of the "−35" sequence had been removed and transcription from a construct from which the entire "−35" element had been removed was low but clearly detectable. Link (1984) noted that the region between the −35 and −10 elements of the mustard *psbA* promoter includes nucleotide sequences resembling the TATA box of eukaryotic cells (and the prokaryotic −10) and suggested that some features of eukaryotic transcription initiation sequences might be recognized by plastid RNA polymerase. In the experiments described above, Link (1984) assayed transcription from linear templates. The gene *psbA* is strongly transcribed *in vitro* by chloroplast RNA polymerase preparations. Small amounts of transcription can be seen even from linear templates containing this gene, but Link (1984) did not report the relative rates of transcription from linear versus relaxed versus negatively supercoiled templates. The effects of DNA conformation on the relative effectiveness of two maize plastid gene promoters is discussed elsewhere in this chapter.

Bradley and Gatenby (1985) studied the transcription of the maize gene *atpB in vitro* from cloned supercoiled templates by a maize chloroplast extract. The conserved sequences in the promoter region are 5′-TTGACA . . . 18 nt . . . TAGTAT (Krebbers *et al.*, 1982). They found that deletions 5′ to the "−35" region had no effect on the transcription rate but that transcription could not be detected from a deleted template retaining only 24 nucleotides upstream of the presumed transcription initiation site—i.e., lacking a −35 element. Each of several single mutations they made in the "−35" sequence reduced transcription *in vitro* about 10-fold. They concluded that no sequences upstream of the "−35" were necessary for strong active transcription *in vitro* but that the "−35" sequence was essential because removing it completely eliminated transcription and that the introduction of individual mutations into the 5′-TTGACA-3′ sequence sharply reduced transcription.

Upstream of the coding sequence of the maize and spinach plastid

genes for transfer RNA_2Met (*trn*M2) is the sequence 5′-TTGCTT . . . 17 nt . . . TATAAT-3′ (Steinmetz *et al.*, 1983; Gruissem and Zurawski, 1985b). Gruissem and Zurawski (1985b) analyzed the effects of various deletions and mutations in this region on transcription by spinach chloroplast RNA polymerase partially purified by DEAE chromatography. The templates were provided as "predominantly form I" plasmids. A template containing all of the "−35"-like sequence 5′-TTGCTT-3′ plus an immediately upstream A was transcribed strongly, but even with the first TT pair deleted, about 40% of wild type "transcription efficiency" was retained; the deletion of the next G reduced it to about 20% of wild type. Thus, it cannot be said from these experiments that a complete −35 sequence is essential for transportation of spinach *trn*M_2 *in vitro* by the partially purified spinach polymerase *in vitro*. However, deletion of the entire "−35" sequence reduced transcription efficiency to about 4% of wild type. This, plus results of experiments in which the "−35" sequences of some plastid genes were fused to "−10" regions of others, led Gruissem and Zurawski (1985a) to conclude that "−35" sequences strongly influence the rate of transcription of spinach plastid genes *in vitro*. They proposed, from the outcome of mix-and-match experiments in which putative −35 elements were interchanged, that the order of effectiveness of −35 sequences from spinach plDNA is TTGACA (in *psbA* and *atpB*) > TTGCTT (*trn*M_2) > TTGCGC (*rbcL*). However, the effectiveness of −35 promoter elements may not vary independently of other parts of the promoter complex.

The capacity to introduce foreign DNA stably into the *Chlamydomonas* chloroplast chromosome (Boynton *et al.*, 1988; Blowers *et al.*, 1989) has opened the way for studying features of *Chlamydomonas* chloroplast promoters *in vivo*. This is important because various possible σ factors, repressor proteins, or activator proteins may be lacking from transcriptionally active preparations and because the conformation of the template can affect the relative effectiveness of a promoter (Stirdivant *et al.*, 1985). Deleting all sequences 5′ to position −24 (relative to the transcription initiation site) of *Chlamydomonas atpB*—thus eliminating any possible "−35" type of sequence—has no effect on the relative rate of transcription or on the transcription initiation site from this promoter (Blowers *et al.*, 1990). This stands in striking contrast to the results of the experiments of Bradley and Gatenby (1985) in which a maize *atpB* construct deleted to about the same point was not transcribed *in vitro*. Additional analyses of the *atpB* promoter in transgenic *Chlamydomonas* (U. Klein, J. DeCamp, and L. Bogorad, unpublished) show that deletions that do not disturb the "−10" sequence but remove all sequences normally upstream do not eliminate transcription, although the tran-

scription rate from some deletions may be somewhat lower. In the *Chlamydomonas* experiments, candidate *atpB* promoter sequences were fused to the coding region for the GUS (β-glucuronidase) gene and the transcription rates were measured in pulse experiments *in vivo*. Transcription of GUS from the unaltered *atpB* promoter was the same as the rate of transcription of the endogenous *atpB*, which served as an internal control. On the other hand, comparable analyses of the promoter region of the 165 rRNA gene *in vivo* in *Chlamydomonas* have revealed that all of the TTGACA sequence is essential for transcription, thus, there are two strikingly different types of promoters in *C. reinhardtii* chloroplasts (J. DeCamp, U. Klein, and L. Bogorad, unpublished data). It remains to be seen whether this also true in plDNA of other species.

Finally, Gruissem *et al.* (1986) have observed that cloned spinach plastid genes *trnR1* and *trnS1*, containing no recognizable sequences with homology to other "−10" and "−35" sequences, can be transcribed *in vitro;* deletions of 5′ sequences almost down to the coding regions do not eliminate transcription of these genes *in vitro*. It has been suggested that these genes may have internal promoters.

The insensitivity of isolated higher plant chloroplast RNA polymerases to rifampicin has been discussed above, but tagetitoxin, a chlorosis-inducing phytotoxin produced by *Psuedomonas syringae* pv. *tagetis, blocks the accumulation of rRNAs as well as transcripts of rbcL, psbA*, and perhaps of all plastid RNAs in greening tissues (Lukens *et al.*, 1987). This toxin inhibits pea chloroplast RNA polymerase activity but has only a slight effect on transcription by RNA polymerase II from wheat germ nuclei (Mathews and Durbin, 1990). Which step in transcription is blocked by this compound is not known.

To summarize, transcription has been studied experimentally from only a few chloroplast promoters of a small number of species. An intact "−35" sequence is reported to be absolutely required in the maize *atpB* experiments of Bradley and Gatenby (1985) and for the 165 rRNA genes of *C. reinhardtii*. In the case of mustard *psbA* and spinach *trn*M_2, all or parts of the "−35" sequence can be removed without completely eliminating transcription but the rates are reduced, sometimes drastically, from such deleted templates. In the cases of the *Chlamydomonas atpB* gene, sequences upstream of the "−10" element—including any sequence that might function as a "−35" block—do not influence transcription *in vivo*. In the cases of the spinach genes *trnR1* and *trnS1*, sequences within the coding region may serve as promoters. Thus, except for homologies to sequences of *E. coli* σ 70 promoters (Reznikoff *et al.*, 1985), most plastid promoters remain relatively poorly defined. Even the most detailed studies of one or a few plastid gene promoters from a

single species may fail to reveal the extent of divergence of promoters of a single plastid chromosome. Also, analyses of plastid gene promoters from a few species—mostly angiosperms—cannot reveal how much diversity may exist among all chloroplast promoters. It is worth noting that Reznikoff *et al.* (1985) deduced some generalizations regarding *E. coli* σ 70 promoters from knowledge of >100 *E. coli* promoter sequences (!) but even then could go on to enumerate exceptional findings, including the facts that some promoter mutations do not involve alterations in the canonical "−35" and "−10" sequences and that some promoters lack one of these canonical sequences "without a commensurate effect on promoter activity." Superimposed on the observations of these 100 genes is the conclusion that sequences of promoters of 31 positively regulated *E. coli* genes deviate significantly from the consensus sequence of 94 "typical promoters," i.e., genes transcribed by *E. coli* σ 70 RNA polymerase (Raibaud and Schwartz, 1984).

The conclusions regarding our present knowledge of plastid promoters are as follows. (1) If sequences resembling *E. coli* σ 70 promoters are found at the expected location 5′ to the experimentally defined transcription initiation site of a plastid gene, they may indeed function as a promoter. (2) If no such sequences are found, a promoter may well be there nonetheless, but we may not recognize it yet.

5. Transcription Termination

How do 3′ ends of plastid transcripts form? This problem is intertwined with the question of how transcript stability may be influenced by resistance to 3′ exonuclease action. Platt (1986) has pointed out that 3′ ends of transcripts in prokaryotes can be generated directly by termination (e.g., at some sequence 3′ to a stem-loop encoding sequence) or by 3′ exonuclease processing back from a termination sequence. He states that "Usually there is no reason to doubt the assumed purpose of hairpins as termination sites [hairpins plus a short downstream uridine sequence is enough to create a 3′ terminus *in vitro* for most bacterial genes], but in the larger context of controlling gene expression, the ability to function as a barrier to exonuclease degradation may be . . . dominant . . . the dual function may prove to be a common feature of genetic control elements."

Short inverted repeated sequences that could encode stems of small stem-loop structures in transcripts occur after the ends of the protein-coding regions of many plastid genes. The similarity between such stem-loop structures and associated sequences to transcription termination regions of bacteria was noted (e.g., Zurawski *et al.*, 1981), and

this led to the suggestion that they play some role in transcription termination in plastids. Stern and Gruissem (1987) have argued that stem-loop regions play no role in termination. They used a partially purified spinach chloroplast RNA polymerase preparation (Gruissem *et al.*, 1983a) that can also process tRNA gene transcripts to transcribe templates with or without a stem-loop-coding sequence located between a chloroplast promoter and the coding regions for the genes *trnH* or *trnM*. Because this preparation could process tRNA gene transcripts, they assayed for termination by looking for $tRNA^{His}$ and $tRNA^{Met}$ molecules. These tRNAs accumulated even when the stem-loop and surrounding sequences were present; therefore, they concluded that the stem-loop sequences have no effect on termination, although Stern and Gruissem (1987) pointed out that their spinach preparation might well have lost an essential termination factor(s). They determined that the spinach RNA polymerase–tRNA gene transcript-processing preparation they used contained 3′ exonuclease activity that trimmed several spinach plastid transcripts (these were made from linearized templates using SP6 or T7 RNA polymerases and were then added to the spinach RNA polymerase preparation) to beyond the inverted repeated sequences at the ends of these genes. After incubation, the 3′ ends of *rbcL* and *psbA* mRNAs corresponded precisely to those of these mRNAs isolated from plastid preparations, although *petD* and *rpoA* transcripts were a few nucleotides longer or shorter than the *in vivo* forms. Finally, the effect of the presence of inverted repeated sequences on RNA stability was assessed by comparing the lifetime in the RNA polymerase preparation of transcripts of *petD*—an RNA containing the inverted repeated sequences—with the lifetimes of several transcripts lacking such structures (an internal sequence of *petB* and 5′ sequences from *psbA* and *rbcL*). Transcripts of *petD* mRNA made and processed *in vivo* were isolated and incubated in the polymerase preparation. These transcripts survived much longer than any of the other RNAs tested but, curiously, a longer *petD* transcript made *in vitro* had a very short half-life in the spinach chloroplast transcription mixture. The mechanisms for termination of transcription in plastids are not known but the possibility that stem-loops plus adjacent sequences play some roles in these processes is not excluded by the data now available. Structures such as stem-loops on RNA may well affect the stability of transcripts, as discussed by Platt (1986) and by Brawerman (1989), but, as yet, no strong experimental evidence indicates this is the case of plastid transcripts. Experiments with transgenic plastids *in situ* are likely to provide clues that can be exploited to study the problem *in vitro*.

B. Proteins that Affect Plastid DNA Transcription or Bind to Plastid DNA

Transcription systems of prokaryotes (e.g., *E. coli*) are comprised of a multimeric core RNA polymerase plus a variety of σ factor proteins that affect promoter selectivity, and, in addition, negative and positive regulatory proteins are involved in the transcription of some genes. Polypeptides of maize chloroplast RNA polymerase corresponding to those of the *E. coli* core enzyme have been identified (see above); however, although some additional polypeptides are present in highly purified plastid polymerase preparation, none have been shown unequivocally to be the functional equivalents of *E. coli* σ proteins.

Transcriptional regulation contributes to changes in RNA levels for some plastid genes in some species in some developmental conditions (e.g., Klein and Mullet, 1990; Troxler *et al.*, 1989). Kobayashi (Chapter 12B) discusses evidence indicating that DNA methylation may influence transcription of plastid genes in the development of chromoplasts in fruit, but it would not be surprising to also find proteins that are involved in the negative and/or positive regulation of transcription of plastid genes. A few protein fractions and proteins that alter the activity or binding of an RNA polymerase to plDNA sequences have been identified.

There are two reports of plant proteins that influence behavior of *E. coli* core RNA polymerase in a manner that might be expected of transcriptional σ factors. Surzycki and Shellenbarger (1976) described a 51-kDa polypeptide from *C. reinhardtii* that stimulates the formation of open binary complexes with *C. reinhardtii* chloroplast DNA by *E. coli* core RNA polymerase in the presence of rifampicin. On the basis of such behavior the polypeptide was designated "σ 2." Surzycki and Shellenbarger (1976) also reported that the 51-kDa polypeptide affected the "core *Chlamydomonas* plastid RNA polymerase"; however, unfortunately, too little is known about that enzyme at present to interpret the results. Along a similar line, Bulow and Link (1988) reported that a mustard plastid protein fraction, which itself lacks DNA-binding activity, enhances the formation of complexes between *E. coli* core RNA polymerase and mustard *psbA* promoter sequences. They describe their protein fraction as having "σ-like activity."

As discussed above, Jolly and Bogorad (1980) identified a 27.5-kDa polypeptide that is required by appropriately prepared maize plastid RNA polymerase to preferentially transcribe cloned chloroplast genes from negatively supercoiled bacterial plasmids. Briat *et al.* (1984) isolated

a 17-kDa double-strand DNA-binding protein from the spinach TAC; this small polypeptide cross-reacts with antibodies against *E. coli* and cyanobacterial HU—a DNA-binding protein. A 115-kDa protein from peas has been reported (Lam *et al.*, 1988) to bind and protect from exonuclease III digestion a region around the transcription start site of the plastid *rbcL*. Crevel *et al.* (1989) have isolated from cyanobacteria, as well as from spinach chloroplasts, a 10-kDa single-strand DNA-binding protein. And, finally, the work of Thompson and Mosig (1987) implies that when cloned in *E. coli* all or part of the *E. coli* integration host factor (IHF) protein (e.g., Prentki and Galas, 1987) interacts with the *Chlamydomonas* chloroplast P_A promoter (the novobiocin responsive promoter) that they studied.

In only a few of the cases described above has a protein been shown to influence the activity of a homologous RNA polymerase. Only scraps of information on these proteins have been acquired so far, and none has been shown to affect the specificity of plastid RNA polymerase transcription. The S protein (Jolly and Bogorad, 1980) probably acts generally rather on specific genes, although not enough data are available to support such an assertion.

In the course of searching for chloroplast promoter-binding proteins in plastid extracts, a fraction with such activity was identified eluting from a DEAE column just ahead of the PF RNA polymerase fraction. The binding fraction (BF) that preceded the PF fraction from the DEAE column was found to have RNA polymerase activity. Unlike the PF RNA polymerase, which strongly transcribes from supercoiled but barely from relaxed DNA, the BF polymerase is active with relaxed templates as well as negatively supercoiled ones (it was tested with two maize chloroplast genes, *psbA* and *rbcL*). Whether the BF activity is of the same multimeric enzyme present in the PF fraction plus one or more additional polypeptides or if the RNA polymerase in the BF fraction is entirely different from that in the PF fraction is not known (Zaitlin *et al.*, 1989).

Thompson and Mosig (1984, 1985, 1987) have found that novobiocin, an inhibitor of gyrase, promotes the transcription of at least one segment of *C. reinhardtii* chloroplast DNA *in vivo*. Transcripts of the novobiocin responsive segment, designated P_A, are "overaccumulated" when the segment is cloned in an *E. coli* strain carrying the *himA42* mutation in which the α subunit of *E. coli* IHF is altered. They interpret this observation "to mean that . . . repression of P_A by IHF minimally requires both binding of IHF to a site overlapping P_A and binding of one or more additional proteins, perhaps IHF itself, to sequences upstream of P_A." Surette and Chaconas (1989) have demonstrated that transfer of mu DNA *in vitro* occurs with optimal efficiency from donor plasmid DNA of

σ value -0.06. Inasmuch as the *E. coli* chromosome is at a σ of about -0.025 *in vivo*, the reaction would not be expected to occur. Localized or general alterations of plastid chromosome topology, through the action of topoisomerases or other conformation modifying proteins, could contribute to the regulation of plastid gene expression. The IHF protein facilitates the process.

The development of the photosynthetic apparatus in the photosynthetic bacterium *Rhodobacter capsulatus* apparently shows that gene expression is affected by alterations in DNA topology. This photosynthetic purple nonsulfur bacterium grows on an exogenous carbon source when atmospheric oxygen is available but can also grow anaerobically and photosynthetically at lower oxygen tensions. The photosynthetic apparatus is synthesized and assembled under reduced oxygen tension. Zhu and Hearst (1988) followed the effects of gyrase inhibitors on mRNA levels of more than a dozen genes in *R. capulatus* and concluded that DNA supercoiling is very likely involved in the differential expression of photosynthetic genes in response to the level of oxygen. However, Cook *et al.* (1989) could not show that the DNA changed conformation using trimethylpsoralen to "freeze" the DNA in its conformation *in vivo*. They suggested that novobiocin might be blocking transcription by preventing the removal of supercoils.

It is also interesting in considering plastid development that Blake and Gralla (quoted in Borowiec and Gralla, 1987) observed "changes in unrestrained superhelicity dependent upon growth conditions" in *E. coli*.

C. Transcription *in Vivo*

Plastid genes are most often quite tightly packed on the chromosome. Numerous examples show divergently transcribed genes whose promoters are very close to one another (Link and Bogorad, 1980) and overlapping transcripts that arise from opposite strands (e.g., Schwarz *et al.*, 1981). Furthermore, a single region of the chromosome is often represented by transcripts of many different sizes. Some of the multiple transcripts are products of processing of a single large transcript (Chapter 5A), but there are two reports of alternative promoters that lie within clusters of plastid genes that are also transcribed as long polycistronic RNAs. An alternative promoter has been found in the protein-coding region of the *psbD-psbC* gene cluster in tobacco plDNA (Yao *et al.*, 1989), and Haley and Bogorad (1990) have determined that transcription initiation at alternative promoters contributes to the generation of overlapping transcripts for segments of the maize gene clusters *psbE-psbF-psbL-*

ORF40 and ORF31-*petE*-ORF42. In almost all of the cases cited, one can imagine RNA polymerase molecules colliding—i.e., of transcription from one promoter interfering with transcription from another. However, there are many copies of the chromosome in each plastid. Whether all active genes are being transcribed from the same chromosome or, in cases of promoters that could interfere, only one of the two promoters is being used on a single chromosome is not known.

In the course of light-induced development of chloroplasts from etioplasts in dark-grown maize seedlings, the activity of plastid RNA polymerase, measured on a calf thymus DNA template, rises about fourfold in 16 hr (Apel and Bogorad, 1976). Amounts of the two largest maize RNA polymerase polypeptides increase little, if at all, in that time interval compared with the change in activity. The production of a new polypeptide that stimulates the activity of the polymerase or changes in conditions that affect polymerase activity in the plastid are obvious possibilities, but no evidence supports such proposals at present. Also, it has not been shown that RNA polymerase activity is limiting before illumination. However, the greatly accelerated rate of incorporation of ^{32}P-phosphate into plastid rRNAs, which occurs soon after illumination of etiolated maize leaves (Bogorad, 1968), may reflect the light-induced increase in RNA polymerase activity.

The large amount of phosphate in tissues of higher plants dilutes out added ^{32}P-phosphate and, thus, except for the rRNAs that are so abundant, this severely limits the ability to look at rates of transcription of plastid genes *in vivo*. However, this can be accomplished with the single-celled flagellate *C. reinhardtii*. Relative transcription rates have been determined for the genes *rrn* (1.000), *rbcL* (0.312), *psbA* (0.132), *rp116* (0.049), *atpB* (0.043), *tufA* (0.030), and *psaB* (0.013) by ^{32}P-phosphate pulse-labeling (Blowers *et al.*, 1990). Defining segments of the promoter regions and *trans*-acting elements responsible for these 80-fold different rates by using the *Chlamydomonas* chloroplast transcription system should be possible.

There is a transitory increase in the sizes of the pools of transcripts of a number of genes during the light-induced development of etioplasts into chloroplasts in maize (Bedbrook *et al.*, 1978; Rodermel and Bogorad, 1985; Sheen and Bogorad, 1988); however, pools of transcripts of *psbA* (Rodermel and Bogorad, 1985) and *petE* (J. Haley and L. Bogorad, unpublished) continue to increase in the illuminated dark-grown seedlings well after those of other photoresponsive genes drop. Klein and Mullet (1990) observed similar transitory changes for *rbcL* transcripts in 8-day-old barley seedlings when dark-grown plants are illuminated. They also reported that the rate of *psbA* transcription in 4.5-day-old dark-grown

barley seedlings increased about fourfold after 4 hr of illumination but began to decline by 14 hr of illumination. (Photoregulated plastid development is discussed in Chapter 12B.)

Mullet and Klein (1987) designed a plastid run-on assay to determine the relative rates of transcription of plastid genes *in vivo:* Labeled RNAs are generated by isolated plastids after lysis in a transcription mixture. Mullet and Klein (1987) concluded that in 4.5-day-old dark-grown barley seedlings the relative transcription of most of the plastid genes they studied by this method was affected little, if at all, by illumination. Deng and Gruissem (1987, 1988; Deng *et al.*, 1987, 1989) used the same assay to compare transcription of a number of plastid genes in dark-grown and light-grown spinach seedlings and advanced more far-reaching conclusions. Under their assay conditions and at the times that they took samples, they could see no differences in the transcription of any one of 10 plastid genes relative to one another (except perhaps *psaA*), by the run-on transcription assay of lysed plastids. They went on to argue that all plastid gene transcription is constitutive, i.e., proportional to promoter strength (which was assumed to be constant), and that this is the case in all plastids—etioplasts, chloroplasts, and chromoplasts—at all times. They have proposed that it is only the relative stability of certain transcripts and not the relative rate of transcription, that changes during light-induced and other types of plastid differentiation. (This question is also addressed by Kobayashi in Chapter 12B.) Although changes in transcript stability during plastid development have not been demonstrated to date, they may occur for some transcripts in some plants and may well occur for spinach seedlings at all times. However, strong evidence was obtained subsequently (Klein and Mullet, 1990) indicating that the generalization that plastid transcription is entirely constitutive in all plants in all forms of plastids and that all pool size changes are solely the results of posttranscriptional events is incorrect. Klein and Mullet (1990) found that the run-on assay they (Mullet and Klein, 1987) and Deng and Gruissem (1987, 1988) used earlier was flawed in that some of the transcription products could have been destroyed by plastid ribonuclease action. Klein and Mullet (1990) optimized the assay and then found that overall plastid transcription activity is near maximum and largely light-independent in apical sections of 4.5-day-old dark-grown barley seedlings, except that, even under these conditions, *psbA* transcription was enhanced when dark-grown seedlings were illuminated. Illumination of 8-day-old dark-grown barley seedlings, on the other hand, increased the transcription of *psbA*, *rbcL*, and 16S rDNA differentially, i.e., not all to the same relative extent. Furthermore, they found that transcription of the maize plastid genes *rbcL* and *psbA* in-

creased to different extents (in the range of 3.5- to 4.5-fold) upon illumination of 9-day-old dark-grown maize seedlings and that the increase in rate of transcription of 16S rDNA during illumination was much smaller. Although not demonstrated to date, changes in RNA stability may well occur also. Overall, it is most likely that the sizes of some transcript pools are affected largely by transcriptional changes and others are altered by posttranscriptional changes, and that this varies from species to species.

What is it in the plastid that changes during greening that alters the transcription rate? In the single-celled rhodophyte *Cyanidium caldarium,* heme acts directly or indirectly to specifically induce transcription of genes for apophycocyanin and apoallophycocyanin (Troxler *et al.*, 1989; R. Troxler, S. Rodermel, and L. Bogorad, unpublished). Do other small molecules act with or as effectors or derepressors for other chloroplast genes? Are there σ-like factors that influence transcription from specific chloroplast genes? Are there topological changes in all or parts of plastid chromosomes during differentiation of plastids into any form? And could such changes affect the transcription of some plastid genes differently than others?

IV. FUTURE RESEARCH DIRECTIONS

A substantial number of questions remain regarding transcription in plastids: What are plastid promoter structures? How diverse are they? What are the characteristics of terminator sequences or structures? Is there more than one type of DNA-dependent RNA polymerase in plastids? If there are two or more, how do they differ from one another in polypeptide composition and specificity? Are genes encoding some components of the plastid transcriptional apparatus located in the nuclear genome? Are there repressor and effector proteins or σ-like factors that are involved in differential transcription of various plastid genes? Are they encoded in nuclear or plastid genes? How is differential transcription of plDNA affected *in vivo*? Do localized or general changes in template topology play roles in differential expression during plastid development? What affects transcript stability, and how does stability contribute to the size of the transcript pool of each gene?

ACKNOWLEDGMENTS

The preparation of this chapter and the research from the author's laboratory described herein have been supported by the National Science Foundation and the National Institute of General Medical Sciences. I am also grateful to Jim Hu for reading an earlier version of this chapter and for providing helpful suggestions.

REFERENCES

Apel, K., and Bogorad, L. (1976). Light-induced increase in the activity of maize plastid DNA-dependent RNA polymerase. *Eur. J. Biochem.* **67,** 615–620.

Bedbrook, J., and Bogorad, L. (1976). Endonuclease recognition sites mapped on *Zea mays* chloroplast DNA. *Proc. Natl. Acad. Sci. USA* **73,** 4309–4313.

Bedbrook, J. R., Kolodner, R., and Bogorad, L. (1977). *Zea mays* chloroplast ribosomal RNA genes are part of a 22,000 base pair inverted repeat. *Cell* **11,** 739–749.

Bedbrook, J. R., Link G., Coen, D. M., Bogorad, L., and Rich, A. (1978). Maize plastid gene expressed during photoregulated development. *Proc. Natl. Acad. Sci. USA* **75,** 3060–3064.

Blanc, M., Briat, J.-F., and Laulhere, J.-P. (1981). Influence of the ionic environment on the *in vitro* transcription of the spinach plastid DNA by a selectively bound RNA-polymerase DNA complex. *Biochim. Biophys. Acta* **655,** 374–382.

Bliska, J. B., and Cozzarelli, N. R. (1987). Use of site-specific recombination as a probe of DNA structure and metabolism in vivo. *J. Mol. Biol.* **194,** 205–218.

Blowers, A. D., Bogorad, L., Shark, K. B., and Sanford, J. C. (1989). Studies on *Chlamydomonas* chloroplast transformation: Foreign DNA can be stably maintained in the chromosome. *Plant Cell* **1,** 123–132.

Blowers, A. D., Ellmore, G. S., Klein, U., and Bogorad, L. (1990). Transcriptional analysis of endogenous and foreign genes in chloroplast transformants of *Chlamydomonas reinhardtii*. *Plant Cell,* **2,** 1059–1070.

Bogorad, L. (1968). Control mechanisms in plastid development. *In* "Control Mechanisms in Developmental Processes," *Devel. Biol. Supplement* **1,** 1–31. Academic Press, New York.

Bogorad, L., and Woodcock, C. L. F. (1970). Rifamycins: The inhibition of plastid RNA synthesis *in vivo* and variable effects on chlorophyll formation in maize leaves. *In* "Autonomy and Biogenesis of Mitochondria and Chloroplasts," pp. 92–97. North-Holland Publ. Co., Amsterdam.

Borowiec, J. A., and Gralla, J. D. (1987). All three elements of the lac ps promoter mediate its transcriptional response to DNA supercoiling. *J. Mol. Biol.* **195,** 89–97.

Bottomley, W., Smith, H. J., and Bogorad, L. (1971). RNA polymerases of maize. I. The partial purification and properties of the chloroplast enzyme. *Proc. Natl. Acad. Sci. USA* **68,** 2412–2416.

Boyer, S. K., and Mullet, J. E. (1986). Characterization of *P. sativum* chloroplast psbA transcripts produced *in vivo, in vitro* and in *E. coli*. *Plant Mol. Biol.* **6,** 229–243.

Boynton, J. E., Gillham, N. W., Harris, E. H., Hosier, J. P., Johnson, A. M., Jones, A. R., Randolph-Anderson, B. L., Robertson, D., Klein, T. M., Shark, K. B., and Sanford, J. C. (1988). Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. *Science* **240,** 1534–1538.

Bradley, D., and Gatenby, A. A. (1985). Mutational analysis of the maize chloroplast ATPase-beta subunit gene promoter: The isolation of promoter mutants in *E. coli* and their characterization in a chloroplast *in vitro* transcription system. *EMBO J.* **4,** 3641–3648.

Brandle, E., and Zetsche, K. (1971). Die Wirkung von Rifampicin auf die RNA- und Proteinsynthese sowie die Morphogenese und den Chlorophyllgehalt kernhaltiger und kernloser Acetabularia-zellen. *Planta* **99,** 46–55.

Brawerman, G. (1989). mRNA decay: Finding the right targets. *Cell* **57,** 9–10.

Briat, J.-F., and Mache, R. (1980). Properties and characterization of a spinach chloroplast RNA polymerase isolated from a transcriptionally active DNA–protein complex. *Eur. J. Biochem.* **111,** 503–509.

Briat, J.-F., Laulhere, J.-P., and Mache, R. (1979). Transcription activity of a DNA–protein complex isolated from spinach plastids. *Eur. J. Biochem.* **98,** 285–292.

Briat, J.-F., Letoffe, S., Mache, R., and Rouviere-Yaniv, J. (1984). Similarity between the bacterial histone-like protein HU and a protein from spinach chloroplasts. *FEBS Lett.* **172,** 75–79.

Briat, J.-F., Bisanz-Seyer, C., and Lescure, A.-M. (1987). *In vitro* transcription initiation of the rDNA operon of spinach chloroplast by a highly purified soluble homologous RNA polymerase. *Curr. Genet.* **11,** 259–263.

Brown, R. D., Bastia, R., and Haselkorn, R. (1970). *In* "RNA Polymerase and Transcription" (L. Silvestri, ed.), pp. 309–328. North-Holland Publ. Co., Amsterdam.

Bulow, S., and Link, G. (1988). Sigma-like activity from mustard (*Sinapis alba* L.) chloroplasts conferring DNA-binding and transcription specificity to *E. coli* core RNA polymerase. *Plant Mol. Biol.* **10,** 349–357.

Bulow, S., Reiss, T., and Link, G. (1987). DNA-binding proteins of the transcriptionally active chromosome from mustard (*Sinapis alba* L.) chloroplasts. *Curr. Genet.* **12,** 157–159.

Carrillo, N., and Bogorad, L. (1988). Chloroplast DNA replication *in vitro:* Site-specific initiation from preferred templates. *Nucleic Acids Res.* **16,** 5603–5620.

Chiang, K.-S., and Sueoka, N. (1967). Replication of chloroplast DNA in *Chlamydomonas reinhardi* during vegetative cell cycle: Its mode and regulation. *Biochemistry* **57,** 1506–1513.

Cook, D. N., Armstrong, G. A., and Hearst, J. E. (1989). Induction of anaerobic gene expression in *Rhodobacter capsulatus* is not accompanied by a local change in chromosomal supercoiling as measured by a novel assay. *J. Bacteriol.* **171,** 4836–4843.

Crevel, G., Laine, B., Sautiere, P., and Galleron, C. (1989). Isolation and characterization of DNA-binding proteins from the cyanobacterium *Synechococcus* sp. PCC 7002 (*Agmenellum quadruplicatum*) and from spinach chloroplasts. *Biochim. Biophys. Acta* **1007,** 36–43.

Crossland, L. D., Rodermel, S. R., and Bogorad, L. (1984). Single gene for the large subunit of ribulosbisphosphate carboxylase in maize yields two differentially regulated mRNAs. *Proc. Natl. Acad. Sci. USA* **81,** 4060–4064.

Day, A., and Ellis, T. H. N. (1984). Chloroplast DNA deletions associated with wheat

plants regenerated from pollen: Possible basis for maternal inheritance of chloroplasts. *Cell* **39,** 359–368.

Day, A., and Ellis, T. H. N. (1985). Deleted forms of plastid DNA in albino plants from cereal anther culture. *Curr. Genet.* **9,** 671–678.

Deng, X.-W., and Gruissem, W. (1987). Control of plastid gene expression during development: The limited role of transcriptional regulation. *Cell* **49,** 379–387.

Deng, X.-W., and Gruissem, W. (1988). Constitutive transcription and regulation of gene expression in non-photosynthetic plastids of higher plants. *EMBO J.* **7,** 3301–3308.

Deng, X.-W., Stern, D. B., Tonkyn, J. C., and Gruissem, W. (1987). Plastid run-on transcription. Application to determine the transcriptional regulation of spinach plastid genes. *J. Biol. Chem.* **262,** 9641–9648.

Deng, X.-W., Tonkyn, J. C., Peter, G. F., Thornber, J. P., and Gruissem, W. (1989). Post-transcriptional control of plastid mRNA accumulation during adaptation of chloroplasts to different light quality environments. *Plant Cell* **1,** 645–654.

Glass, R. E., Jones, S. T., Nene, V., Nomura, T., Fujita, N., and Ishihama, A. (1986). Genetic studies on the subunit of *Escherichia coli* RNA polymerase VIII. Localisation of a region involved in promoter selectivity. *Mol. Gen. Genet.* **203,** 487–491.

Gold, B., Carrillo, N., Tewari, K. K., and Bogorad, L. (1987). Nucleotide sequence of a preferred maize chloroplast genome template for *in vitro* DNA synthesis. *Proc. Natl. Acad. Sci. USA* **84,** 194–198.

Greenberg, B. M., Narita, J. O., DeLuca-Flaherty, C., Gruissem, W., Rushlow, K. A., and Hallick, R. B. (1984). Evidence for two RNA polymerase activities in *Euglena gracilis* chloroplasts. *J. Biol. Chem.* **259,** 14880–14887.

Gruissem, W. (1989). Chloroplast gene expression: How plants turn their plastids on. *Cell* **56,** 161–170.

Gruissem, W., and Zurawski, G. (1985a). Analysis of promoter regions for the spinach chloroplast *rbcL, atpB* and *psbA* genes. *EMBO J.* **4,** 3375–3383.

Gruissem, W., and Zurawski, G. (1985b). Identification and mutational analysis of the promoter for a spinach chloroplast transfer RNA gene. *EMBO J.* **4,** 1637–1644.

Gruissem, W., Greenberg, B. M., Zurawski, G., Prescott, D. M., and Hallick, R. B. (1983a). Biosynthesis of chloroplast transfer RNA in a spinach chloroplast transcription system. *Cell* **35,** 815–828.

Gruissem, W., Narita, J. O., Greenberg, B. M., Prescott, D. M., and Hallick, R. B. (1983b). Selective *in vitro* transcription of chloroplast genes. *J. Cell. Biochem.* **22,** 31–46.

Gruissem, W., Elsner-Menzel, C., Latshaw, S., Narita, J. O., Schaffer, M. A., and Zurawski, G. (1986). A subpopulation of spinach chloroplast tRNA genes does not require upstream promoter elements for transcription. *Nucleic Acids Res.* **14,** 7541–7556.

Haley, J., and Bogorad, L. (1990). Alternative promoters are used for genes within maize chloroplast polycistronic transcription units. *Plant Cell* **2,** 323–333.

Hallick, R. B., Lipper, C., Richards, O. C., and Rutter, W. J. (1976). Isolation of a transcriptionally active chromosome from chloroplasts of *Euglena gracilis*. *Biochemistry* **15,** 3039–3045.

Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.-R., Meng, B.-Y., Li, Y.-Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., and Sugiura, M. (1989). The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.* **217,** 185–194.

Hu, J., and Bogorad, L. (1990). Maize chloroplast RNA polymerase: The 180-, 120-, and

38-kilodalton polypeptides are encoded in chloroplast genes. *Proc. Natl. Acad. Sci. USA* **87,** 1531–1535.

Hudson, G. S., Holton, T. A., Whitfield, P. R., and Bottomley, W. (1988). Spinach chloroplast *rpoBC* genes encode three subunits of the chloroplast RNA polymerase. *J. Mol. Biol.* **200,** 639–654.

Igloi, G. L., Meinke, A., Dory, I., and Kossel, H. (1990). Nucleotide sequence of the maize chloroplast rpo B/C1/C2 operon: Comparison between the derived protein primary structures from various organisms with respect to functional domains. *Mol. Gen. Genet.* **221,** 379–394.

Jenni, B., and Stutz, E. (1979). Analysis of *Euglena gracilis* chloroplast DNA mapping of a DNA sequence complementary to 16S rRNA outside of the three rRNA gene sets. *FEBS Lett.* **102,** 95–99.

Jenni, B., Fasnacht, M., and Stutz, E. (1981). The multiple copies of the *Euglena gracilis* chloroplast genome are not uniform in size. *FEBS Lett.* **125,** 175–179.

Johnson, L. B., and McKnight, S. L. (1989). Eukaryotic transcriptional regulatory proteins. *Annu. Rev. Biochem.* **58,** 799–839.

Jolly, S. O., and Bogorad, L. (1980). Preferential transcription of cloned maize chloroplast DNA sequences by maize chloroplast RNA polymerase. *Proc. Natl. Acad. Sci. USA* **77,** 822–826.

Jolly, S. O., McIntosh, L., Link, G., and Bogorad, L. (1981). Differential transcription *in vivo* and *in vitro* of two adjacent maize chloroplast genes: The large subunit of ribulose bisphosphate carboxylase and the 2.2-kilobase gene. *Proc. Natl. Acad. Sci. USA* **78,** 6821–6825.

Joussaume, M. (1973). Mise en evidence de deux formes de RNA polymerase dependante du DNA dans les chloroplastes isoles de feuilles de Pois. *Physiol. Veg.* **11,** 69–82.

Kidd, G. H., and Bogorad, L. (1980). A facile procedure for purifying maize chloroplast RNA polymerase from whole cell homogenates. *Biochim. Biophys. Acta* **609,** 14–30.

Kirk, J. T. O. (1964). DNA-dependent RNA synthesis in chloroplast preparations. *Biochem. Biophys. Res. Comm.* **14,** 393–397.

Kislev, N., Swift, H., and Bogorad, L. (1965). Nucleic acids of chloroplasts and mitochondria in swiss chard. *J. Cell Biol.* **25,** 327–344.

Klein, R. R., and Mullet, J. E. (1990). Light-induced transcription of chloroplast genes. *psbA* transcription is differentially enhanced in illuminated barley. *J. Biol. Chem.* **265,** 1895–1902.

Koller, B., and Delius, H. (1982). Origin of replication in chloroplast DNA of *Euglena gracilis* located close to the region of variable size. *EMBO J.* **1(8),** 995–998.

Kolodner, R. D., and Tewari, K. K. (1975a). Chloroplast DNA from higher plants replicates by both the Cairns and the rolling circle mechanism. *Nature* **256,** 708–711.

Kolodner, R., and Tewari, K. K. (1975b). Presence of displacement loops in the covalently closed circular chloroplast deoxyribonucleic acid from higher plants. *J. Biol. Chem.* **250,** 8840–8847.

Kolodner, R., Tewari, K. K., and Warner, R. C. (1976). Physical studies on the size and structure of the covalently closed circular chloroplast DNA from higher plants. *Biochim. Biophys. Acta* **447,** 144–155.

Krebbers, E., Larrinua, I. M., McIntosh, L., and Bogorad, L. (1982). The maize chloroplast genes for β and ϵ subunits of the photosynthetic coupling factor CF_1, are fused. *Nucleic Acids Res.* **10,** 4985–5502.

Kung, S. D., and Lin, C. M. (1985). Chloroplast promoters from higher plants. *Nucleic Acids Res.* **13,** 7543–7549.

Lam, E., and Chua, N.-H. (1987). Chloroplast DNA gyrase and *in vitro* regulation of transcription by template topology and novobiocin. *Plant Mol. Biol.* **8,** 415–424.

Lam, E., Hanley-Bowdoin, L., and Chua, N.-H. (1988). Characterization of a chloroplast sequence-specific DNA binding factor. *J. Biol. Chem.* **263,** 8288–8293.

LeBrun, M., Briat, J.-F., and Laulhere, J.-P. (1986). Characterization and properties of the spinach chloroplast transcriptionally active chromosome isolated at high ionic strength. *Planta* **169,** 505–512.

Lerbs, S., Briat, J.-F., and Mache, R. (1983). Chloroplast RNA polymerase from spinach: Purification and DNA-binding proteins. *Plant Mol. Biol.* **2,** 67–74.

Lerbs, S., Brautigam, E., and Parthier, B. (1985). Polypeptides of DNA-dependent RNA polymerase of spinach chloroplasts: Characterization by antibody-linked polymerase assay and determination of sites of synthesis. *EMBO J.* **4,** 1661–1666.

Link, G. (1984). DNA sequence requirements for the accurate transcription of a protein-coding plastid gene in a plastid *in vitro* system from mustard (*Sinapis alba* L.). *EMBO J.* **3,** 1697–1704.

Link, G., and Bogorad, L. (1980). The sizes, locations and directions of transcription of two genes on a cloned maize DNA sequence. *Proc. Natl. Acad. Sci. USA* **77,** 1832–1836.

Link, G., and Langridge, U. (1984). Structure of the chloroplast gene for the precursor of the M_r 32,000 photosystem II protein from mustard (*Sinapis alba* L.). *Nucleic Acids Res.* **12,** 945–958.

Loiseaux, S., Mache, R., and Rozier, C. (1975). Rifampicin inhibition of the plastid rRNA synthesis of *Marchantia polymorpha*. *J. Cell Sci.* **17,** 327–335.

Lukens, J. H., Mathews, D. E., and Durbin, R. D. (1987). Effect of tagetitoxin on the levels of ribulose 1,5-bisphosphate carboxylase, ribosomes, and RNA in plastids of wheat leaves. *Plant Physiol.* **84,** 808–813.

Mathews, D. E., and Durbin, R. D. (1990). Tagetitoxin inhibits RNA synthesis directed by RNA polymerases from chloroplasts and *Escherichia coli*. *J. Biol. Chem.* **265,** 493–498.

Matsuda, Y., and Surzycki, S. J. (1980). Chloroplast gene expression in *Chlamydomonas reinhardi*. *Mol. Gen. Genet.* **180,** 463–474.

McKown, R. L., and Tewari, K. K. (1984). Purification and properties of a pea chloroplast DNA polymerase. *Proc. Natl. Acad. Sci. USA* **81,** 2354–2358.

Meeker, R., Nielsen, B., and Tewari, K. K. (1988). Localization of replication origins in pea chloroplast DNA. *Mol. Cell. Biol.* **8,** 1216–1223.

Mullet, J. E., and Klein, R. R. (1987). Transcription and RNA stability are important determinants of higher plant chloroplast RNA levels. *EMBO J.* **6,** 1571–1579.

Munsche, D., and Wollgiehn, R. (1973). Die Synthese von ribosomaler RNA in Chloroplasten von *Nicotiana Rustica*. *Biochim. Biophys. Acta* **294,** 106–117.

Narita, J. O., Rushlow, K. E., and Hallick, R. B. (1985). Characterization of a *Euglena gracilis* chloroplast RNA polymerase specific for ribosomal RNA genes. *J. Biol. Chem.* **260,** 11194–11199.

Ness, P. J., and Woolhouse, H. W. (1980). RNA synthesis in *Phaseolus* chloroplasts. *J. Exp. Bot.* **31,** 223–233.

Nielsen, B. L., and Tewari, K. K. (1988). Pea chloroplast topoisomerase I: Purification, characterization, and role in replication. *Plant Mol. Biol.* **11,** 3–14.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., and Ozeki, H. (1986). Chloroplast gene sequence deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* **322,** 572–574.

Orozco, E. M., Jr., Mullet, J. E., and Chua, N.-H. (1985). An *in vitro* system for accurate transcription initiation of chloroplast protein genes. *Nucleic Acids Res.* **13,** 1283–1302.

Platt, T. (1986). Transcription termination and the regulation of gene expression. *Annu. Rev. Biochem.* **55,** 339–372.

Polya, G. M., and Jagendorf, A. T. (1971a). Wheat leaf RNA polymerases I. Partial purifi-

cation and characterization of nuclear, chloroplast and soluble DNA-dependent enzymes. *Arch. Biochem. Biophys.* **146,** 635–648.

Polya, G. M., and Jagendorf, A. T. (1971b). Wheat leaf RNA polymerases II. Kinetic characterization and template specificities of nuclear, chloroplast, and soluble enzymes. *Arch. Biochem. Biophys.* **146,** 649–657.

Prentki, P., and Galas, D. J. (1987). *Escherichia coli* integration host factor binds specifically to the ends of the insertion sequence IS1 and to its major insertion hot-spot in pBR322. *J. Mol. Biol.* **195,** 261–272.

Purton, S., and Gray, J. C. (1989). The plastid *rpoA* gene encoding a protein homologous to the bacterial RNA polymerase alpha subunit is expressed in pea chloroplasts. *Mol. Gen. Genet.* **217,** 77–84.

Pyke, K. A., Marrison, J., and Leech, R. M. (1989). Evidence for a type II topoisomerase in wheat chloroplasts. *FEBS Lett.* **242(2),** 305–408.

Raibaud, O., and Schwartz, M. (1984). Positive control of transcription initiation in bacteria. *Annu. Rev. Genet.* **18,** 173–206.

Ratcliff, S., and Surzycki, S. (1974). DNA-dependent RNA polymerases from *Chlamydomonas reinhardtii*. *J. Cell. Biol.* **63,** 281a.

Ravel, Chapuis, P., Heizmann, P., and Nigon, V. (1982). Electron microscopic localization of the replication origin of *Euglena gracilis* chloroplast DNA. *Nature* **300,** 78–81.

Reiss, T., and Link, G. (1985). Characterization of transcriptionally active DNA–protein complexes from chloroplasts and etioplasts of mustard (*Sinapis alba* L.). *Eur. J. Biochem.* **148,** 207–212.

Reznikoff, W. S., Siegele, D. A., Cowing, D. W., and Gross, C. A. (1985). The regulation of transcription initiation in bacteria. *Annu. Rev. Genet.* **19,** 355–387.

Rodermel, S. R., and Bogorad, L. (1985). Maize plastid photogenes: Mapping and photoregulation of transcript levels during light-induced development. *J. Cell Biol.* **100,** 463–476.

Ruf, M., and Kossel, H. (1988). Structure and expression of the gene coding for the alpha-subunit of DNA-dependent RNA polymerase from the chloroplast genome of *Zea mays*. *Nucleic Acids Res.* **16,** 5741–5754.

Rushlow, K. E., Orozco, E. M., Jr., Lipper, C., and Hallick, R. B. (1980). Selective *in vitro* transcription of *Euglena* chloroplast ribosomal RNA genes by a transcriptionally active chromosome. *J. Biol. Chem.* **255,** 3786–3792.

Russell, D., and Bogorad, L. (1987). Transcription analysis of the maize chloroplast gene for the ribosomal protein S4. *Nucleic Acids Res.* **15,** 1853–1867.

Sarhan, F., and Chevrier, N. (1984). A comparative study of nuclear and chloroplast DNA dependent RNA polymerases from wheat leaves. *Int. J. Biochem.* **16,** 707–711.

Schlunegger, B., and Stutz, E. (1984). The *Euglena gracilis* chloroplast genome: Structural features of a DNA region possibly carrying the single origin of DNA replication. *Curr. Genet.* **8,** 629–634.

Schlunegger, B., Fasnacht, M., Stutz, E., Koller, B., and Delius, H. (1983). Analysis of a polymorphic region of the *Euglena gracilis* chloroplast genome. *Biochim. Biophys. Acta* **739,** 114–121.

Schwarz, Z., Jolly, S. O., Steinmetz, A. A., and Bogorad, L. (1981). Overlapping divergent genes in the maize chloroplast chromosome and *in vitro* transcription of the gene for tRNAHis. *Proc. Natl. Acad. Sci. USA* **78,** 3423–3427.

Sheen, J.-Y., and Bogorad, L. (1988). Differential expression in bundle sheath and mesophyll cells of maize genes for photosystem II components encoded by the plastid genome. *Plant Physiol.* **86,** 1020–1026.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., and Sugiura, M. (1986). The complete nucleotide sequence of the tobacco chloroplast genome; its gene organization and expression. *EMBO J.* **5,** 2043–2049.

Siedlecki, J., Zimmermann, W., and Weissbach, A. (1983). Characterization of a prokaryotic topoisomerase I activity in chloroplast extracts from spinach. *Nucleic Acids Res.* **11,** 1523–1536.

Sijben-Muller, G., Hallick, R. B., Alt, J., Westhoff, P., and Herrmann, R. G. (1986). Spinach plastid genes coding for initiation factor IF-1, ribosomal protein S11 and RNA polymerase α-subunit. *Nucleic Acids Res.* **14,** 1029–1044.

Smith, H. J., and Bogorad, L. (1974). The polymerase of *Zea mays* chloroplasts. *Proc. Natl. Acad. Sci. USA* **71,** 4839–4842.

Steinmetz, A., Krebbers, E. T., Schwarz, Z., Gubbins, E. J., and Bogorad, L. (1983). Nucleotide sequences of five maize chloroplast transfer RNA genes and their flanking regions. *J. Biol. Chem.* **9,** 5503–5511.

Stenzel, T. T., Patel, P., and Bastia, D. (1987). The integration host factor of *Escherichia coli* binds to bent DNA at the origin of replication of the plasmid pSC101. *Cell* **49,** 709–717.

Stern, D. B., and Gruissem, W. (1987). Control of plastid gene expression: 3′ Inverted repeats act as mRNA processing and stabilizing elements, but do not terminate transcription. *Cell* **51,** 1145–1157.

Stirdivant, S. M., Crossland, L. D., and Bogorad, L. (1985). DNA supercoiling affects *in vitro* transcription of two maize chloroplast genes differently. *Proc. Natl. Acad. Sci. USA* **82,** 4886–4890.

Sun, E., Shapiro, D. R., Wu, B. W., and Tewari, K. K. (1986). Specific *in vitro* transcription of 16S rRNA gene by pea chloroplast RNA polymerase. *Plant Mol. Biol.* **6,** 429–439.

Surette, M. G., and Chaconas, G. (1989). A protein factor which reduces the negative supercoiling requirement in the Mu DNA strand transfer reaction is *Escherichia coli* integration host factor. *J. Biol. Chem.* **264,** 3028–3034.

Surzycki, S. J. (1969). Genetic functions of the chloroplast of *Chlamydomonas reinhardi:* Effect of rifampim on chloroplast DNA-dependent RNA polymerase. *Proc. Natl. Acad. Sci. USA* **63,** 1327–1334.

Surzycki, S. J., and Shellenbarger, D. L. (1976). Purification and characterization of a putative sigma factor from *Chlamydomonas reinhardi. Proc. Natl. Acad. Sci. USA* **73,** 3961–3965.

Tewari, K. K., and Goel, A. (1983). Solubilization and partial purification of RNA polymerase from pea chloroplasts. *Biochemistry* **22,** 2142–2148.

Tewari, K. K., and Wildman, S. G. (1969). Function of chloroplast DNA II. Studies on DNA-dependent RNA polymerase activity of tobacco chloroplasts. *Biochim. Biophys. Acta* **186,** 358–372.

Thompson, R. J., and Mosig, G. (1984). Light and genetic determinants in the control of specific chloroplast transcripts in *Chlamydomonas reinhardtii. Plant Physiol.* **76,** 1–6.

Thompson, R. J., and Mosig, G. (1985). An ATP-dependent supercoiling topoisomerase of *Chlamydomonas reinhardtii* affects accumulation of specific chloroplast transcripts. *Nucleic Acids Res.* **13,** 873–891.

Thompson, R. J., and Mosig, G. (1987). Stimulation of a *Chlamydomonas* chloroplast promoter by novobiocin *in situ* and *E. coli* implies regulation by torsional stress in the chloroplast DNA. *Cell* **48,** 281–287.

Thompson, R. J., and Mosig, G. (1988). Integration host factor (IHF) represses a *Chlamydomonas* chloroplast promoter in *E. coli*. *Nucleic Acids Res.* **16,** 3313–3326.

Troxler, R. F., Lin, S., and Offner, G. D. (1989). Heme regulates expression of phycobiliprotein photogenes in the unicellular rhodophyte, *Cyanidium caldarium*. *J. Biol. Chem.* **264,** 20596–20601.

Waddell, J., Wang, X.-M., and Wu, M. (1984). Electron microscopic localization of the chloroplast DNA replicative origins in *Chlamydomonas reinhardii*. *Nucleic Acids Res.* **12,** 3843–3856.

Wang, X.-M., Chang, C. H., Waddell, J., and Wu, M. (1984). Cloning and delimiting one chloroplast DNA replicative origin of *Chlamydomonas*. *Nucleic Acids Res.* **12,** 3857–3872.

Wu, M., Lou, J. K., Chang, D. Y., Chang, C. H., and Nie, Z. Q. (1986). Structure and function of a chloroplast DNA replication origin of *Chlamydomonas reinhardtii*. *Proc. Natl. Acad. Sci. USA* **83,** 6761–6765.

Yao, W. B., Meng, B. Y., Tanaka, M., and Sugiura M. (1989). An additional promoter within the protein-coding region of the *psbD-psbC* gene cluster in tobacco chloroplast DNA. *Nucleic Acids Res.* **17,** 9583–9591.

Zaitlin, D., Hu, J., and Bogorad, L. (1989). Binding and transcription of relaxed DNA templates by fractions of maize chloroplast extracts. *Proc. Natl. Acad. Sci. USA* **86,** 876–880.

Zhu, Y. S., and Hearst, J. E. (1988). Transcription of oxygen-regulated photosynthetic genes requires DNA gyrase in *Rhodobacter capsulatus*. *Proc. Natl. Acad. Sci. USA* **85,** 4209–4213.

Zurawski, G., Perrot, B., Bottomley, W., and Whitfeld, P. R. (1981). The structure of the gene for the large subunit of ribulose 1,5-bisphosphate carboxylase from spinach chloroplast DNA. *Nucleic Acids Res.* **9,** 3251–3270.

CHAPTER 5

Transcript Processing in Plastids: Trimming, Cutting, Splicing

Masahiro Sugiura

Nagoya University
Chikusa, Nagoya 464, Japan

I. INTRODUCTION

The chloroplast genome contains about 130 genes (reviewed by Sugiura, 1989), and most of the chloroplast genes are transcribed polycistronically. Multiple transcripts are observed for most chloroplast gene clusters and these are mainly the results of multiple RNA processing of the primary transcripts. Processing of rRNA and tRNA precursors and of precursors from split genes are required to form functional RNA molecules. Some chloroplast genes are known to be constitutively transcribed, suggesting that posttranscriptional processing of primary transcripts represents an important step in the control of chloroplast gene expression (reviewed by Gruissem, 1989a). This chapter discusses our current knowledge of the processing of transcripts from chloroplast protein-coding genes. Transcript-processing activities in chloroplasts include the trimming of 3′ ends of transcripts, the cutting of 5′-untranslated sequences and intercistronic sequences, and the removal of intron sequences from mRNA precursors. The processing of rRNA

and tRNA precursors in chloroplasts is presented in Chapters 6A and 7A and a recent review (Gruissem, 1989b).

II. 3′-TRIMMING OF mRNA PRECURSORS

Most chloroplast transcription units contain short inverted repeats at their 3′ ends. These inverted repeats, which can potentially form stem-loop structures, were originally thought to function as transcriptional terminators. The role of spinach chloroplast inverted repeats found at the 3′ end of chloroplast transcription units has been examined using an *in vitro* transcription system (Stern and Gruissem, 1987). It was found that the inverted repeat is ineffective as transcription terminators *in vitro* but serves as accurate and efficient RNA-processing signals. Precursor mRNAs are processed in a 3′–5′ direction by a nuclease activity present in the transcription system, generating nearly homogeneous 3′ ends distal to the inverted repeat. In some cases, the 3′ ends generated *in vitro* coincide with those found for chloroplast mRNAs *in vivo*. RNAs containing inverted repeats at their 3′ ends exhibit a greatly enhanced stability.

The stability and protein interaction of the 3′-inverted repeat of chloroplast mRNAs have been investigated in detail (Stern *et al.*, 1989). The mutational analysis of the *petD* inverted repeat revealed that the formation of a stem-loop is necessary, but not always sufficient, for RNA stabilization *in vitro*. Precursor and processed *petD* 3′-inverted repeats have been shown to bind different sets of proteins. Comparison of the bound *petD* proteins with proteins that bind to *rbcL* and *psbA* reveals that binding of certain proteins is gene-specific. A spinach chloroplast-processing activity that catalyzes the conversion of the *petD* precursor mRNA to the mature mRNA possessing a 3′-inverted repeat was further characterized (Stern and Gruissem, 1989). The activity requires Mg^{++} or Mn^{++}, but not K^{+}, and is biochemically distinct from *Escherichia coli* mRNA processing.

The nuclear *NAC2* locus of *Chlamydomonas* has been shown to encode a *trans*-acting factor that acts in a gene-specific manner to control the stability of *petD* transcript (Kuchka *et al.*, 1989). The second exon of *Chlamydomonas psaA* is cotranscribed with *psbD*, but the mutation in the *NAC2* gene has no effect on *psaA* message maturation or stability. In barley, the accumulation of *psbD* transcripts induced by blue light has been suggested to require *de novo* synthesis of a nuclear-encoded gene

product (Gamble and Mullet, 1989). These observations suggest that nuclear-encoded proteins function in chloroplast mRNA maturation and differential mRNA stability as a major control step in chloroplast gene expression.

III. CUTTING OF mRNA PRECURSORS

Most chloroplast genes are cotranscribed and produce polycistronic primary transcripts. Extensive processing (cutting) events then result in the accumulation of complicated sets of overlapping transcription products (e.g., Westhoff, 1985; Barkan *et al.*, 1986; Berends *et al.*, 1987; Tanaka *et al.*, 1987; Ohto *et al.*, 1988; Woodbury *et al.*, 1988). Cutting of the chloroplast mRNA precursor generally occurs in its 5′-untranslated and intercistronic regions.

The analysis of chloroplast mRNAs has revealed that several transcripts contain multiple 5′ ends. Multiple 5′ ends were found for *rbcL* (Crossland *et al.*, 1984; Mullet *et al.*, 1985; Erion, 1985), *atpB* (Mullet *et al.*, 1985), *psbA* (Boyer and Mullet, 1986), the *psbB* operon (Westhoff, 1985; Tanaka *et al.*, 1987), the *rpl23* cluster (Thomas *et al.*, 1988); *rps16* (Neuhaus *et al.*, 1989), and the *psbDC* cluster (Berends *et al.*, 1987; Yao *et al.*, 1989). *In vitro* capping experiments revealed that the longest RNAs contain 5′-triphosphates representing the primary transcripts and that shorter RNAs must be processed forms of the primary transcripts. The longest maize *rbcL* mRNA has been demonstrated to be specifically cut to shorter RNA species *in vitro* by an activity present in maize chloroplast lysate (Hanley-Bowdoin *et al.*, 1985). Multiple transcriptional initiation occurs, however, in the *rpl23* cluster (Thomas *et al.*, 1988) and the *psbDC* cluster (Yao *et al.*, 1989).

Chloroplast polycistronic transcripts are generally processed into many overlapping shorter RNA species. Some of the shorter RNAs are monocistronic but others are not. Detailed analysis of polycistronic transcripts have been done for the *psbB* operon (Barkan *et al.*, 1986; Rock *et al.*, 1987; Tanaka *et al.*, 1987; Westhoff and Herrmann, 1988) and the *psbDC* operon (Gamble *et al.*, 1988; Gamble and Mullet, 1989; Yao *et al.*, 1989).

The *psbB* operon contains the genes for the 47-kDa chlorophyll *a* apoprotein (*psbB*) and the 10 kDa phosphoprotein (*psbH*) of photosystem II and for cytochrome b_6 (*petB*) and subunit IV (*petD*) of the cytochrome *b*/*f* complex in this order. In spinach, the RNA pattern is complex and

resolves into 18 major RNA species (Westhoff and Herrmann, 1988). All RNA species derive from one DNA strand and hybridize in an overlapping fashion, and they arise from the putative 5.6-kb primary transcript by processing rather than multiple transcription initiation and termination. Processing results ultimately in the formation of monocistronic mRNA for each of the two photosystem II polypeptides and a dicistronic mRNA for both subunits of the cytochrome *b/f* complex. These mono- and dicistronic mRNAs are likely to be major translatable mRNAs. However, almost all of the transcripts from the maize *psbB* operon cosediment with polysomes, suggesting that they are translated and intercistronic processing is not always required for translation of these RNAs (Barkan, 1988). The primary transcript that undergoes stepwise processing to yield discrete products is shown in Fig. 1. The formation of monocistronic *psbB* and *psbH* RNAs requires cleavage of the precursor before and after the *psbH* cistron. Three such cleavage sites were identified: two in close proximity in front of *psbH* and one between *psbH* and *petB*

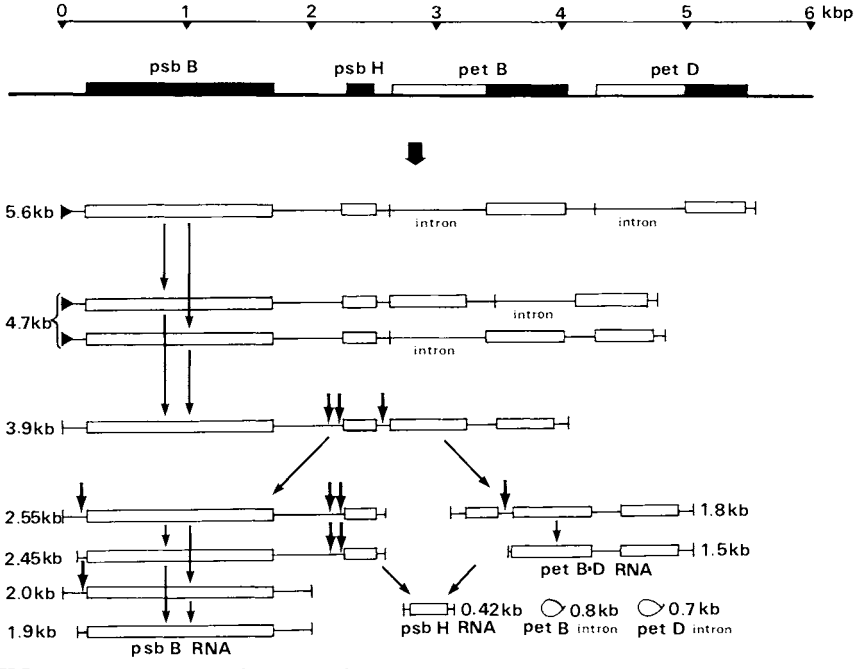


FIG. 1. RNA cutting and splicing diagram for the *psbB* operon. Triangles indicate transcription initiation sites, and bold arrows show cutting sites. [From Westhoff and Herrmann (1988) and Tanaka *et al.* (1987), with extensive modifications.]

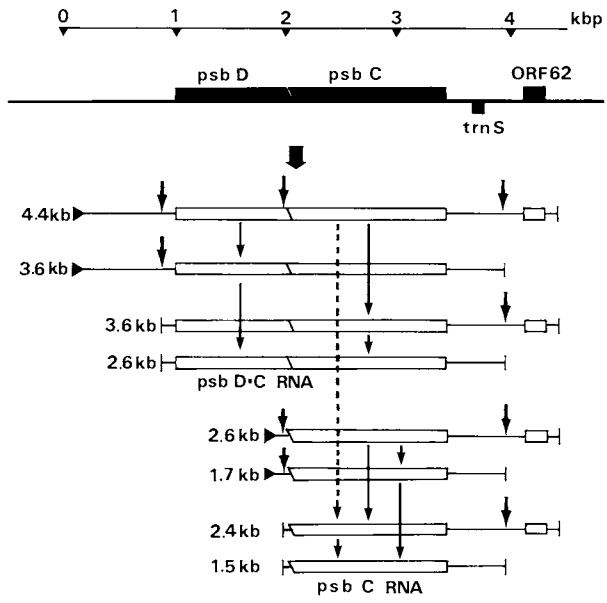


FIG. 2. RNA cutting diagram for the *psbD/C* operon. Triangles indicate transcription initiation sites, and bold arrows show cutting sites. [From Yao *et al.* (1989), with extensive modifications.]

(Westhoff and Herrmann, 1988). Comparison of DNA sequences in the vicinity of these sites reveal two conserved regions, a hexanucleotide motif (YGGAA/TY), occurring in tandem. The cleaved sites are located in the first (before *petB*) or in both motifs (before *psbH*). The conservation of these sequence elements in other plants suggests the involvement of a sequence-specific endonuclease in cutting of intercistronic regions.

Transcripts from the *psbDC* operon have been analyzed in detail in tobacco (Yao *et al.*, 1989). This operon contains in linear sequence the overlapping genes encoding the D2 (*psbD*) and 43-kDa (*psbC*) proteins of photosystem II and ORF62 (Fig. 2). Eight major RNA species were detected, four RNAs containing both the *psbD* and *psbC* sequences and the other four containing the *psbC* sequence. Transcription starts at two different sites, before *psbD* and within *psbD* and produces 4.4-kb and 2.6-kb primary transcripts. Six other RNA species are likely to arise from the precursors by cutting. Unlike the *psbB* processing, no conserved sequences were detected near or at the cleavage sites. This observation suggests that cutting enzymes are gene-specific.

IV. SPLICING OF mRNA PRECURSORS

Introns in chloroplast genes were first reported for the 23S rDNA of *Chlamydomonas reinhardii* (Rochaix and Malnoe, 1978). Most genes possessing introns in higher plant chloroplast genomes contain single introns, while *Euglena* and *Chlamydomonas* genes for polypeptides contain multiple introns (i.e., Hallick *et al.*, 1985; Shinozaki *et al.*, 1986a; Sugiura *et al.*, 1987; Plant and Gray, 1988). Six chloroplast tRNA genes in higher plants have introns but none are known in algae. Chloroplast genes containing introns so far found are listed in Table I.

Introns found in chloroplast genes can be classified into four groups on the basis of the intron boundary sequences and of the possible secondary structures (Shinozaki *et al.*, 1986a; Christopher and Hallick, 1989). Group I introns can be folded with a secondary structure typical of group I introns of fungal mitochondrial genes (Michel and Dujon, 1983). The introns of *trnL* and 23S rDNA belong to this group. Introns of *trnI* and *trnA* can be folded into a secondary structure, which is similar to the postulated structure of group II introns in fungal mitochondrial genes. Group III introns include most of the chloroplast introns and have conserved boundary sequences GTGYGRY at the 5′ ends and RYCNAYY(Y)YNAY at the 3′ ends. Their postulated secondary structures are similar to those of group II introns. A fourth intron group has recently been described for *Euglena rpl23*, *rps19*, *rps3*, *rpl14*, *rps8*, *rps14*, *rpl16*, and *tufA* (Christopher and Hallick, 1989; Montandon *et al.*, 1987). These introns are uniform in size (95–109 bp), share common features with each other, and are distinct from group I–III introns.

The presence of introns can be predicted by sequence homologies with the corresponding *E. coli* genes (e.g., tRNA genes, ATP synthase genes, ribosomal protein genes) and by the conserved intron boundary sequences. It is not easy to deduce the 3′ ends of potential introns because of their low conservation. To confirm the existence of introns and to determine the splice sites of pre-mRNAs, reverse transcription (cDNA) analysis is carried out. A primer that is complementary to a putative second exon is hybridized to total chloroplast RNA, and cDNA is synthesized from the primer using reverse transcriptase in the presence of dideoxyribonucleoside triphosphates. The sequence ladder obtained in this way indicates an exact splice site of the pre-mRNA. This method was successfully applied for *petB* and *petD* (Tanaka *et al.*, 1987; Fukuzawa *et al.*, 1987), *ndhA* and *ndhB* (Matsubayashi *et al.*, 1987), *rpl16* (Posno *et al.*, 1986), *atpF* (Hudson *et al.*, 1987), *tufA* (Montandon *et al.*, 1987), *psbF* (Cushman *et al.*, 1988), *rpl16* and *rps8* (Christopher and Hal-

TABLE I
Chloroplast Genes Containing Introns

Gene	Number of Introns: Higher Plants	*Euglena*	*Chlamydomonas*	Comments
23SrDNA	0	0	1	ORF[a], an intron in *Chlorella*
trnL-UAA	1	0		
trnI-GAU	1	0	0	
trnA-UGC	1	0	0	
trnV-UAC	1	0		
trnG-UCC	1	0		
trnK-UUU	1	0		ORF
rps3	0	2		
rsp8	0	3		
rps12	3 exons	0		*trans*-splicing
rps14	0	1		
rps16	1			
rps19	0	2		
rp12	1	0		no intron in spinach
rp114	0	1	0	
rp116	1	3	0	
rp123	0	3		pseudogene in spinach
rpoC1	1			no intron in rice and maize
tufA	no gene	3		
rbcL	0	9	0	
psaA	0	3	3 exons	*trans*-splicing
psaB	0	6	0	
psbA	0	4	4	
psbC	0	1		ORF
psbE	0	2		
psbF	0	1		
petB	1			no intron in *Chlorella*
petD	1		0	no intron in *Chlorella*
atpF	1			
ndhA	1			
ndhB	1			

[a] ORF means the presence of an ORF in an intron.

lick, 1989). Alternatively, S1 nuclease analysis was used to identify splice sites of pre-mRNAs from *atpF* (Bird *et al.*, 1985), *petB* and *petD* (Rock *et al.*, 1987; Westhoff and Herrmann, 1988), *rpoC1* (Hudson *et al.*, 1988), and *rpl16* (Zhou *et al.*, 1989). The monocots rice and maize have no intron in *rpoCl* and, thus, comparison with their sequences supports the proposed splice sites in the dicot and other plant *rpoC1* (Hiratsuka *et al.*, 1989; Shimada *et al.*, 1990; Igloi *et al.*, 1990).

The tobacco gene for ribosomal protein *rps12* is divided into one copy of 5′ *rps12* and two copies of 3′ *rps12* (Torazawa *et al.*, 1986). 5′ *rps12* contains exon 1 of 38 codons and 3′ *rps12* consists of exon 2 of 78 codons, a 536-bp intron, and exon 3 of 7 codons. This gene structure was designated as a "divided" gene (Shinozaki *et al.*, 1986b). The 5′ and 3′ *rps12*s are separated by 28 kbp and are transcribed independently. These two transcripts are spliced in *trans* to produce a mature mRNA for S12 (Zaita *et al.*, 1987; Koller *et al.*, 1987; Hildebrand *et al.*, 1988). The 3′-flanking sequence of exon 1 and the 5′-flanking sequence of exon 2 fit the conserved boundary sequences of chloroplast group III introns. It is noteworthy that the tobacco *rps12* requires both *cis* and *trans* splicing for its mature mRNA. Interestingly, liverwort *rsp12* is divided into two parts (Fukuzawa *et al.*, 1986), and *Euglena rps12* is an uninterrupted gene (Montandon and Stutz, 1984).

The *Chlamydomonas psaA* gene is also divided into three parts (Kück *et al.*, 1987). The first exon of 30 codons is 50 kbp away from the second exon (60 codons), which is itself 90 kbp from exon 3 (661 codons). All exons are flanked by the consensus intron boundary sequences. The three exons are transcribed independently as precursors, and the synthesis of mature *psaA* mRNA involves the assembly in *trans* of three separate transcripts (Choquet *et al.*, 1988). Interestingly, exon 2 is cotranscribed with the upstream *psbD* gene, and *psaB* is uninterrupted as are *psaA* and *psaB* of higher plants. The tobacco *psaA* and *psaB* genes are cotranscribed together with the downstream *rps14* gene (Meng *et al.*, 1988).

V. CONCLUSIONS

Accumulating evidence indicates that expression of some chloroplast genes is effectively controlled at the posttranscriptional level. The relative transcriptional activities for several chloroplast genes remain nearly constant during chloroplast development, although their steady-state mRNA level changes dramatically (Deng and Gruissem, 1987; Mullet and Klein, 1987). Some chloroplast genes are constitutively transcribed, and the differential accumulation of their mRNAs in different tissues is controlled at the posttranscriptional level (Deng and Gruissem, 1987). The ratio of spliced to unspliced mRNAs from several chloroplast genes varies in different tissues (Barkan, 1989). Rapid splicing occurs in mature leaf tissue (e.g., Tanaka *et al.*, 1987); however, only a low proportion of

transcripts is spliced in roots and meristem-proximal leaf tissue, which is indicative of a regulatory role of splicing activities in chloroplast gene expression (Barkan, 1989).

Characterization of the molecular components involved in the post-transcriptional regulation has just been initiated. Several specific proteins that bind the 3′-inverted repeat of chloroplast mRNA have been identified in spinach (Stern *et al.*, 1989), and three proteins having typical RNA-binding domains have been isolated from tobacco chloroplasts (Li and Sugiura, 1990), although the function of most of these proteins is still unclear. The availability of *in vitro* processing systems from chloroplasts facilitates the study of the processing reaction in more detail.

Chloroplast introns can be classified into three to four groups, suggesting the presence of complex splicing pathways. Self-splicing *in vitro* of pre-mRNAs from chloroplast split genes has not been reported. The group III introns have the conserved intron boundary sequences. Interestingly, three tRNA gene introns from higher plants are included in group III. The intron boundary sequence is similar to that found in nuclear protein-coding genes, suggesting that at least one group of chloroplast intron sequences is removed by a mechanism similar to that operating in the nucleus. Splicing of nuclear pre-mRNAs is generally catalyzed by protein–RNA complexes. RNA molecules are thought not to be imported into chloroplasts from the cytoplasm, which suggests that the RNA components, if any, should be encoded in the chloroplast genome. Recently a chloroplast-encoded RNA molecule possibly involved in *trans*-splicing of *psaA* pre-mRNAs was detected in *Chlamydomonas* (Goldschmidt-Clermont *et al.*, 1990), and the tobacco chloroplast genome has been found to encode small RNA species other than tRNAs and rRNA (Matsubayashi *et al.*, unpublished). No *in vitro* splicing systems from chloroplasts are currently available. This makes it difficult to analyze individual steps in splicing and to detect factors involved in the splicing reaction in chloroplasts.

REFERENCES

Barkan, A. (1988). Proteins encoded by a complex chloroplast transcription unit are each translated from both monocistronic and polycistronic mRNAs. *EMBO J.* **7**, 2637–2644.

Barkan, A. (1989). Tissue-dependent plastid RNA splicing in maize: Transcripts from four plastid genes are predominantly unspliced in leaf meristerms and roots. *Plant Cell* **1**, 437–445.

Barkan, A., Miles, D., and Taylor, W. C. (1986). Chloroplast gene expression in nuclear, photosynthetic mutants of maize. *EMBO J.* **5,** 1421–1427.

Berends, T., Gamble, P. E., and Mullet, J. E. (1987). Characterization of the barley chloroplast transcription units containing psaA-psaB and psbD-psbC. *Nucleic Acids Res.*, **15,** 5217–5240.

Bird, C. R., Koller, B., Auffret, A. D., Huttly, A. K., Howe, C. J., Dyer, T. A., and Gray, J. C. (1985). The wheat chloroplast gene for CF_0 subunit of ATP synthase contains a long intron. *EMBO J.* **4,** 1381–1388.

Boyer, S. K., and Mullet, J. E. (1986). Characterization of *P. sativum* chloroplast psbA transcripts produced *in vivo, in vitro* and in *E. coli*. *Plant Mol. Biol.* **6,** 229–243.

Choquet, Y., Goldschmidt-Clermont, M., Girard-Bascou, J., Kück, U., Bennoun, P., and Rochaix, J.-D. (1988). Mutant phenotypes support a *trans*-splicing mechanism for the expression of the tripartite *psaA* gene in the *C. reinhardtii* chloroplast. *Cell* **52,** 903–913.

Christopher, D. A., and Hallick, R. B. (1989). *Euglena gracilis* chloroplast ribosomal protein operon: A new chloroplast gene for ribosomal protein L5 and description of a novel organelle intron category designated group III. *Nucleic Acids Res.* **17,** 7591–7607.

Crossland, L. D., Rodermel, S. R., and Bogorad, L. (1984). Single gene for the large subunit of ribulosebisphosphate carboxylase in maize yields two differentially regulated mRNAs. *Proc. Natl. Acad. Sci. USA* **81,** 4060–4064.

Cushman, J. C., Christopher, D. A., Little, M. C., Hallick, R. B., and Price, C. A. (1988). Organization of the *psbE, psbF, orf38,* and *orf42* gene loci on the *Euglena gracilis* chloroplast genome. *Curr. Genet.* **13,** 173–180.

Deng, X.-W., and Gruissem, W. (1987). Control of plastid gene expression during development: The limited role of transcriptional regulation. *Cell* **49,** 379–387.

Erion, J. L. (1985). Characterization of the mRNA transcripts of the maize, ribulose-1,5-bisphosphate carboxylase, large subunit gene. *Plant Mol. Biol.* **4,** 169–179.

Fukuzawa, H., Kohchi, T., Shirai, H., Ohyama, K., Umesono, K., Inokuchi, H., and Ozeki, H. (1986). Coding sequences for chloroplast ribosomal protein S12 from liverwort, *Marchantia polymorpha*, are separated far apart on the different DNA strands. *FEBS Lett.* **198,** 11–15.

Fukuzawa, H., Yoshida, T., Kohchi, T., Okumura, T., Sawano, Y., and Ohyama, K. (1987). Splicing of group II introns in mRNAs coding for cytochrome b_6 and subunit IV in the liverwort *Marchantia polymorpha* chloroplast genome. *FEBS Lett.* **220,** 61–66.

Gamble, P.E., and Mullet, J. E. (1989). Blue light regulates the accumulation of two *psbD-psbC* transcripts in barley chloroplasts. *EMBO J.* **8,** 2785–2794.

Gamble, P. E., Berends Sexton, T., and Mullet, J. (1988). Light-dependent changes in *psbD* and *psbC* transcripts of barley chloroplasts: Accumulation of two transcripts maintains *psbD* and *psbC* translation capability in mature chloroplasts. *EMBO J.* **7,** 1289–1297.

Goldschmidt-Clermont, M., Girard-Bascou, J., Choquet, Y., and Rochaix, J.-D. (1990). Trans-splicing mutants of *Chlamydomonas reinhardtii*. *Mol. Gen. Genet.* **223,** 417–425.

Gruissem, W. (1989a). Chloroplast gene expression: How plants turn their plastids on. *Cell* **56,** 161–170.

Gruissem, W. (1989b). Chloroplast RNA: Transcription and processing. *In* "Biochemistry of Plants," Vol. 15 (A. Marcus, ed.), pp. 151–191. Academic Press, San Diego.

Hallick, R. B., Gingrich, J. C., Johanningmeier, U., and Passavant, C. W. (1985). Introns in *Euglena* and *Nicotiana* chloroplast protein genes. *In* "Molecular Form and Function of the Plant Genome" (L. van Vloten-Doting, G. S. P. Groot, and T. C. Hall, eds.), pp. 211–231. Plenum Press, New York.

Hanley-Bowdoin, L., Orozco, E. M., and Chua N.-H. (1985). In vitro synthesis and processing of the maize chloroplast transcript encoded by a ribulose-1,5-bisphosphate carboxylase large subunit gene. *Mol. Cell Biol.* **5,** 2733–2745.

Hildebrand, M., Hallick, R. B., Passavant, C. W., and Bourque, D. P. (1988). Trans-splicing in chloroplasts: The *rps*12 loci of *Nicotiana tabacum*. *Proc. Natl. Acad. Sci. USA* **85,** 372–376.

Hiratsuka, J., Shimada, H., Whittier, R. F., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, M., Honji, Y., Sun, C. R., Meng, B. Y., Li, Y. Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., and Sugiura, M. (1989). The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.* **217,** 185–194.

Hudson, G. S., Mason, J. G., Holton, T. A., Koller, B., Cox, G. B., Whitfield, P. R., and Bottomley, W. (1987). A gene cluster in the spinach and pea chloroplast genomes encoding one CF_1 and three CF_0 subunits of the H^+-ATP synthase complex and the ribosomal protein S2. *J. Mol. Biol.* **196,** 283–298.

Hudson, G. S., Holton, T. A., Whitfeld, P. R., and Bottomley, W. (1988). Spinach chloroplast *rpoBC* genes encode three subunits of the chloroplast RNA polymerase. *J. Mol. Biol.* **200,** 639–654.

Igloi, G. L., Meinke, A., Döry, I., and Kössel, H. (1990). Nucleotide sequence of the maize chloroplast rpoB/C_1/C_2 operon: Comparison between the derived protein primary structures from various organisms with respect to functional domains. *Mol. Gen. Genet.*, **221,** 379–394.

Koller, B., Fromm, H., Galun, E., and Edelman, M. (1987). Evidence for in vivo *trans* splicing of pre-mRNAs in tobacco chloroplasts. *Cell* **48,** 111–119.

Kuchka, M. R., Goldschmidt-Clermont, M., van Dillewijn, J., and Rochaix, J.-D. (1989). Mutation at the Chlamydomonas nuclear *NAC2* locus specifically affects stability of the chloroplast *psbD* transcript encoding polypeptide D2 of PS II. *Cell* **58,** 869–876.

Kück, U., Choquet, Y., Schneider, M., Dron, M., and Bennoun, P. (1987). Structural and transcriptional analysis of two homologous genes for the P700 chlorophyll *a*-apoproteins in *Chlamydomonas reinhardtii:* Evidence for *in vivo trans*-splicing. *EMBO J.* **6,** 2185–2195.

Li, Y., and Sugiura, M. (1990). Three distinct ribonucleoproteins from tobacco chloroplasts: Each contains a unique amino terminal acidic domain and two ribonucleoprotein concensus motifs. *EMBO J.* **9,** 3059–3066.

Matsubayashi, T., Wakasugi, T., Shinozaki, K., Yamaguchi-Shinozaki, K., Zaita, N., Hidaka, T., Meng, B. Y., Ohto, C., Tanaka, M., Kato, A., Maruyama, T., and Sugiura, M. (1987). Six chloroplast genes (*ndh*A to F) homologous to human mitochondrial genes encoding components of the respiratory-chain NADH dehydrogenase are actively expressed: Determination of the splice-sites in *ndh*A and *ndh*B pre-mRNAs. *Mol. Gen. Genet.* **210,** 385–393.

Meng, B. Y., Tanaka, M., Wakasugi, T., Ohme, M., Shinozaki, K., and Sugiura, M. (1988). Cotranscription of the genes encoding two P700 chlorophyll *a* apoproteins with the gene for ribosomal protein CS14: Determination of the transcriptional initiation site by in vitro capping. *Curr. Genet.* **14,** 395–400.

Michel, F., and Dujon, B. (1983). Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast- and nuclear-encoded members. *EMBO J.* **2,** 33–38.

Montandon, P. E., and Stutz, E. (1984). The genes for the ribosomal proteins S12 and S7 are clustered with the gene for the EF-Tu protein on the chloroplast genome of *Euglena gracilis*. *Nucleic Acids Res.* **12,** 2851–2859.

Montandon, P. E., Knuchel-Aegerter, C., and Stutz, E. (1987). *Euglena gracilis* chloroplast DNA: The untranslated leader of *tuf*A-ORF206 gene contains an intron. *Nucleic Acids Res.* **15,** 7809–7822.

Mullet, J. E., and Klein, R. R. (1987). Transcription and RNA stability are important determinants of higher plant chloroplast RNA levels. *EMBO J.* **6,** 1571–1579.

Mullet, J. E., Orozco, E. M., Jr., and Chua, N.-H. (1985). Multiple transcripts for higher plant *rbc*L and *apt*B genes and localization of the transcription initiation site of the *rbc*L gene. *Plant Mol. Biol.* **4,** 39–54.

Neuhaus, H., Scholz, A., and Link, G. (1989). Structure and expression of a split chloroplast gene from mustard (*Sinapis alba*): Ribosomal protein gene *rps16* reveals unusual transcriptional features and complex RNA maturation. *Curr. Genet.* **15,** 63–70.

Ohto, C., Torazawa, K., Tanaka, M., Shinozaki, K., and Sugiura, M. (1988). Transcription of ten ribosomal protein genes from tobacco chloroplasts: A compilation of ribosomal protein genes found in the tobacco chloroplast genome. *Plant Mol. Biol.* **11,** 589–600.

Plant, A. L., and Gray, J. C. (1988). Introns in chloroplast protein-coding genes of land plants. *Photosyn. Res.* **16,** 23–39.

Posno, M., van Vliet, A., and Groot, G. S. P. (1986). The gene for *Spirodela oligorhiza* chloroplast ribosomal protein homologous to *E. coli* ribosomal protein L16 is split by a large intron near its 5′ end: Structure and expression. *Nucleic Acids Res.* **14,** 3181–3195.

Rochaix, J.-D., and Malnoe, P. (1978). Anatomy of the chloroplast ribosomal DNA of *Chlamydomonas reinhardtii*. *Cell* **15,** 661–670.

Rock, C. D., Barkan, A., and Taylor, W. C. (1987). The maize plastid *psb*B-*psb*F-*pet*B-*pet*D gene cluster: Spliced and unspliced *pet*B and *pet*D RNAs encode alternative products. *Curr. Genet.* **12,** 69–77.

Shimada, H., Fukuda, M., Ishikawa, M., and Sugiura, M. (1990). Rice chloroplast RNA polymerase genes: The absence of an intron in *rpoC1* and the presence of an extra sequence in *rpoC2*. *Mol. Gen. Genet.*, **221,** 395–402.

Shinozaki, K., Deno, H., Sugita, M., Kuramitsu, S., and Sugiura, M. (1986a). Intron in the gene for the ribosomal proteins S16 of tobacco chloroplast and its conserved boundary sequences. *Mol. Gen. Genet.* **202,** 1–5.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamda, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., and Sugiura, M. (1986b). The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. *EMBO J.* **5,** 2043–2049.

Stern, D. B., and Gruissem, W. (1987). Control of plastid gene expression: 3′ Inverted repeats act as mRNA processing and stabilizing elements, but do not terminate transcription. *Cell* **51,** 1145–1157.

Stern, D. B., and Gruissem, W. (1989). Chloroplast mRNA 3′ end maturation is biochemically distinct from prokaryotic mRNA processing. *Plant Mol. Biol.* **13,** 615–625.

Stern, D. B., Jones, H., and Gruissem, W. (1989). Function of plastid mRNA 3′ inverted repeats: RNA stabilization and gene-specific protein binding. *J. Biol. Chem.* **264,** 18742–18750.

Sugiura, M. (1989). The chloroplast chromosomes in land plants. *Annu. Rev. Cell. Biol.* **5,** 51–70.

Sugiura, M., Shinozaki, K., Tanaka, M., Hayashida, N., Wakasugi, T., Matsubayashi, T., Ohto, C., Torazawa, K., Meng, B. Y., Hidaka, T., and Zaita, N. (1987). Split genes and *cis/trans* splicing in tobacco chloroplasts. *In* "Plant Molecular Biology" (D. von Wettstein and N.-H. Chua, eds.), pp. 65–76, Plenum, New York.

Tanaka, M., Obokata, J., Chunwongse, J., Shinozaki, K., and Sugiura, M. (1987). Rapid splicing and stepwise processing of a transcript from the *psbB* operon in tobacco chlo-

roplast: Determination of the intron sites in *petB* and *petD*. *Mol. Gen. Genet.* **209,** 427–431.

Thomas, F., Massenet, O., Dorne, A. M., Briat, J. F., and Mache, R. (1988). Expression of the *rpl*23, *rpl*2 and *rps*19 genes in spinach chloroplasts. *Nucleic Acids Res.* **16,** 2461–2472.

Torazawa, K., Hayashida, N., Obokata, J., Shinozaki, K., and Sugiura, M. (1986). The 5′ part of the gene for ribosomal protein S12 is located 30 kbp downstream from its 3′ part in tobacco chloroplast genome. *Nucleic Acids Res.* **14,** 3143.

Westhoff, P. (1985). Transcription of the gene encoding the 51 kd chlorophyll a-apoprotein of the photosystem II reaction centre from spinach. *Mol. Gen. Genet.* **201,** 115–123.

Westhoff, P., and Herrmann, R. G. (1988). Complex RNA maturation in chloroplasts: The *psb*B operon from spinach. *Eur. J. Biochem.* **171,** 551–564.

Woodbury, N. W., Roberts, L. L., Palmer, J. D., and Thompson, W. F. (1988). A transcription map of the pea chloroplast genome. *Curr. Genet.* **14,** 75–89.

Yao, W. B., Meng, B. Y., Tanaka, M., and Sugiura, M. (1989). An additional promoter within the protein-coding region of the *psb*D-*psb*C gene cluster in tobacco chloroplast DNA. *Nucleic Acids Res.* **17,** 9583–9591.

Zaita, N., Torazawa, K., Shinozaki, K., and Sugiura, M. (1987). *Trans* splicing in vivo: Joining of transcripts from the 'divided' gene for ribosomal protein S12 in the chloroplasts of tobacco. *FEBS Lett.* **210,** 153–156.

Zhou, D.-X., Quigley, F., Massenet, O., and Mache, R. (1989). Cotranscription of the S10- and *spc*-like operons in spinach chloroplasts and identification of three of their gene products. *Mol. Gen. Genet.* **216,** 439–445.

CHAPTER 6

rRNAs and rRNA Genes of Plastids

G. Delp
H. Kössel

Institut für Biologie III der Universität Freiburg
Schanzlestrasse 1
D-7800 Freiburg I. Br.
Freiburg, Federal Republic of Germany

I. INTRODUCTION

The genetic system of chloroplasts depends on ribosomes, tRNAs, and translation factors, which are located within the organellar compartment and which mediate the translation of chloroplast DNA-encoded mRNAs. Historically, the detection of chloroplast ribosomes (Lyttelton, 1960) was reported even before the discovery of chloroplast DNA (Ris and Plaut, 1962; Chun *et al.*, 1963). The characterization of chloroplast ribosomal subunits as 30S and 50S subunits and their interaction with antibiotics such as chloramphenicol (Anderson and Smillie, 1966) and

erythromycin (Mets and Bogorad, 1971) first indicated a prokaryotic nature of chloroplast ribosomes. With the development of modern gene technology, cloning and mapping chloroplast rRNA genes became feasible in the mid-1970s (Bedbrook *et al.*, 1977), which finally permitted the sequence analysis of the complete maize chloroplast rRNA operon (Schwarz, and Kössel, 1980; Koch *et al.*, 1981; Edwards and Kössel, 1981). From this work, a close phylogenetic relationship between the maize chloroplast and bacterial rRNA operon and its rRNA transcripts became clearly evident, which, in addition to earlier morphological observations, also has lent strong support to an endosymbiotic origin of chloroplasts on the molecular level.

Subsequently, many more chloroplast rRNA genes have been analyzed, not only with respect to their fine structures but also with respect to the mode of their expression and the function of their rRNA products. In this chapter, we will review the present knowledge on the rRNAs and rRNA genes from chloroplasts with emphasis on more recent observations and developments concerning the expression of rRNA genes and the maturation and function of their rRNA products.

II. STRUCTURE OF CHLOROPLAST rRNA GENES

A. Copy Numbers and Orientation

As depicted in Figure 1, chloroplast DNAs from most higher plants, and also from algae such as *Chlamydomonas* and *Chlorella*, contain a duplicate set of rRNA genes within each copy of a pair of inverted repeat regions. However, several notable exceptions from the inverted repeat situation have been found. Plastid DNA from certain members of the legume and conifer family lack an inverted repeat region (Chapter 2A) and, therefore, contain only one set of rRNA genes. A second exception is the alga *Euglena gracilis*, in which rRNA operons are arranged in tandem repeat units, the number of which fluctuates between one and five in different strains of the alga. The situation for the *E. gracilis* strain Z, with three complete rRNA operons preceded by a supplementary 16S rRNA gene (the latter being linked to an open reading frame [ORF], see below), is depicted in Fig. 1. Other *Euglena* strains contain supplementary 5S rRNA and partial 23S rRNA genes (Roux *et al.*, 1983; for review, see Koller *et al.*, 1988).

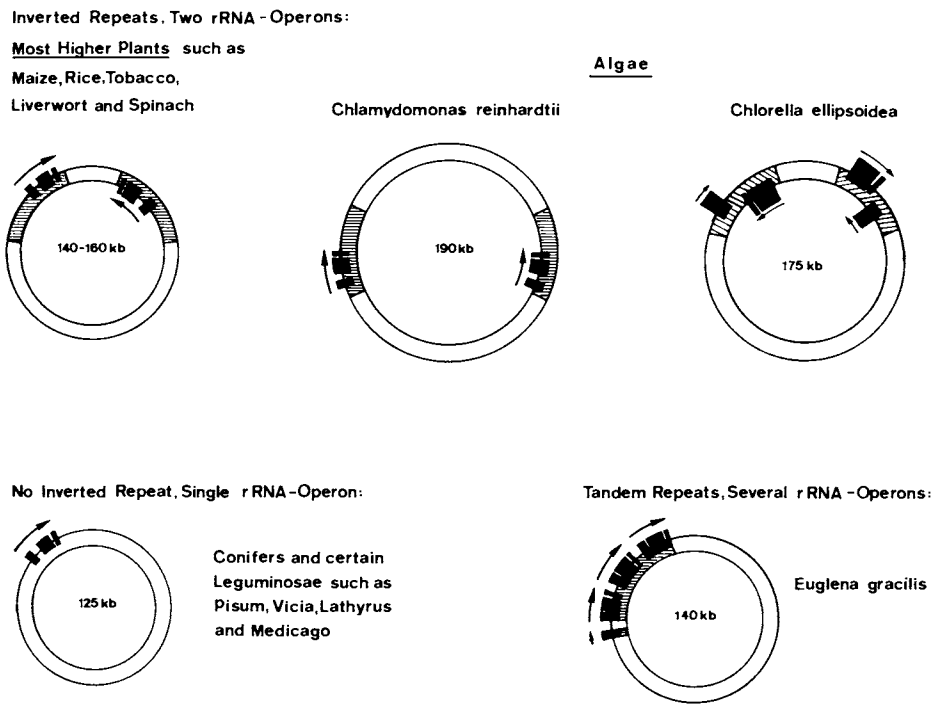


FIG. 1. Repeat regions and rRNA operons on chloroplast genomes. Repeat regions are marked by shading. Ribosomal RNA genes are symbolized by the black segments corresponding to 16S rDNA (middle-sized segments), 23S rDNA (large segments), and 5S rDNA (small segments), with the arrows indicating the polarity of transcription. Note that the segment sizes are only approximately drawn to scale.

B. Operon Structures and Flanking tRNA Genes

Chloroplast rRNA genes are arranged in the form of operons, which show the typically prokaryotic gene order of 16S, 23S, and 5S rDNAs with transcribed spacers between the genes. As depicted in Fig. 2, the existence within the 16S/23S spacer of two tRNA genes coding for $tRNA^{Ile}$ and $tRNA^{Ala}$ appears to be a universal trait for all chloroplast rRNA operons (Graf *et al.*, 1980; Koch *et al.*, 1981), and this again resembles the situation of bacterial rRNA operons (for review of the latter, see Fournier and Ozeki, 1985). However, superimposed on this basic skeleton, small rRNA and tRNA genes appear, which together with intervening sequences at certain positions are characteristic for higher plants or algal species. As indicated in Fig. 2, in higher plants the region distal to

the 5′ end of the 16S rRNA gene contains a $tRNA^{Val}$ gene and the region proximal to the 5S rRNA gene harbors a $tRNA^{Arg}$ gene (Schwarz *et al.*, 1981; Przybyl *et al.*, 1984). Also, the presence of long intervening sequences within the spacer tRNA genes (Koch *et al.*, 1981) and of a gene located between the 23S and 5S rRNA genes, which codes for a separate 4.5S rRNA, is characteristic of chloroplast rRNA operons of higher plants. The 4.5S rRNA gene is structurally equivalent to the 3′ terminus of the bacterial, cyanobacterial, and algal 23S rRNA genes (Edwards *et al.*, 1981; Strittmatter and Kössel, 1984); therefore, it demonstrates the tendency of chloroplast 23S rRNA genes to be fragmented at their ends. In the *Chlamydomonas* chloroplast rRNA operon, other small rRNAs, a 3S and a 7S rRNA, are encoded. Again, positional and sequence homology of the corresponding genes show that they represent equivalents of 5′-terminal regions of other plastid and bacterial 23S rRNAs (Rochaix and Darlix, 1982). In addition to these terminal fragmentations, internal fragmentation of plastid 23S rRNA by hidden breaks (see below) is commonly observed in higher plants. It should, however, be emphasized that all the small rRNAs are contained within the primary transcripts of rRNA operons, in which they constitute a full-length 23S rRNA, and that fragmentation occurs only posttranscriptionally, apparently signalled by the sequences of the respective intergenic regions introduced during evolution.

Intervening sequences are present in the small subunit rRNA genes of *Chlamydomonas moewusii* (Durocher *et al.*, 1989) and in the large subunit rRNA genes of *Chlamydomonas reinhardtii* (Rochaix *et al.*, 1985) and *Chlorella ellipsoidea* (Yamada and Shimaji, 1987a). The intron of the *C. reinhardtii* 23S rRNA gene contains an ORF that codes for an RNA maturaselike protein; therefore, this rRNA operon is a chimeric operon, which includes genes for rRNAs and tRNAs as well as for one potential mRNA, although evidence for translation of the ORF has not been presented so far. A chimeric rRNA operon, which contains a single 16S rRNA gene linked to an ORF of 406 codons, has also been observed in the plastome of certain *E. gracilis* strains (Roux and Stutz, 1985). Although the existence of a 3.6-kb transcript containing sequences from both the genes shows that this chimeric operon is transcribed, evidence for the translation of the ORF is still lacking (Koller *et al.*, 1988).

Euglena 16S rDNA leader sequences contain a pseudo $tRNA^{Ile}$ gene (Orozco *et al.*, 1980; Miyata *et al.*, 1982; Roux and Stutz, 1985), which apparently originates from a transposition of the spacer $tRNA^{Ile}$ gene to the leader region and subsequent divergence to the pseudogene structures. However, sequences homologous to other tRNA genes have also been identified (McGarvey *et al.*, 1988).

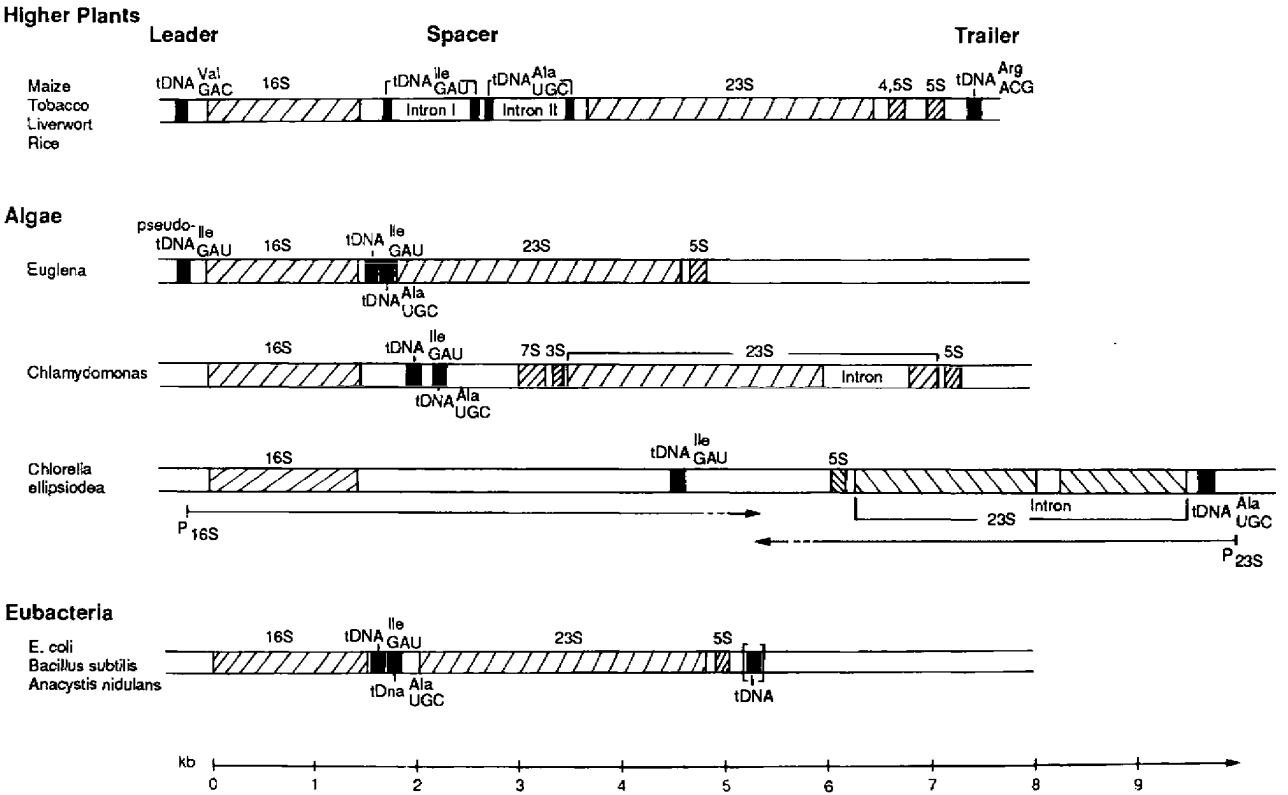


FIG. 2. Alignment of rRNA operons from chloroplasts of higher plants and algae and from eubacteria. Genes coding for rRNAs and tRNAs are marked by shading and black segments, respectively. In the case of *Chlorella ellipsoidea*, the two operons and their polarities are indicated by the long arrows starting at the respective promoter sites P_{16S} and P_{23S}.

The rRNA genes of the *Chlorella ellipsoidea* plastome deviate from the canonical order because they are contained in two closely linked smaller operons of opposite orientation (Yamada and Shimaji, 1987b). As depicted in Fig. 2, the first operon contains the genes for 16S rRNA and $tRNA^{Ile}$, whereas the second operon is composed of the genes coding for $tRNA^{Ala}$, 23S rRNA, and 5S rRNA. This situation is obviously due to the inversion of a 5-kb region by which the original single rRNA operon has been split between the two spacer tRNA genes.

C. Sequence Comparison and Numerical Phylogenetic Relationships

With the exception of a limited number of 4.5S and 5S rRNA sequences, which have been analyzed by direct RNA sequencing, all rRNA sequences are derived from the respective DNA sequences. A compilation of published rDNA and rRNA sequences is given in Table I.

Homologies in terms of percentage of identical nucleotide positions have been obtained from the comparison of the various rRNA operon components (Kössel *et al.*, 1983). As expected, structural genes (rDNAs and tDNAs) show maximum homology (>90%) among higher plant species, somewhat less homology (80–90%) between higher plant and algal species, and even less, but still strong, homology (65–77%) with the bacterial counterparts. On the other hand, leader, intergenic, and trailer regions as well as intervening sequences show considerably less sequence conservation among higher plant species and no significant homologies are detectable between such regions of higher plants and algal and bacterial rRNA operons.

In contrast to the strong homology observed between chloroplast and bacterial rRNA genes, homology with nuclear and mitochondrial rRNA genes is much smaller. In maize, the small subunit rDNA sequences from the three genetic compartments that code for a cytoplasmic 17S rRNA (nuclear encoded), an 18S rRNA (mitochondria), and the plastid 16S rRNA are known (Messing *et al.*, 1984; Chao *et al.*, 1984; Schwarz and Kössel, 1980). An alignment of the three sequences (G. Delp and H. Kössel, unpublished observation) still shows 47% homology between the two organellar rRNAs, which probably reflects their common origin from eubacterial 16S rRNA, whereas only 36 and 35% homology is observed between the nuclear encoded cytoplasmic 17S rRNA and the chloroplast and mitochondrial small subunit rRNAs, respectively. A dot matrix comparison of the three sequences using different criterions of

TABLE I

Sequenced rDNAs/rRNAs and tDNAs/tRNAs from Plastid rRNA Operons

rDNA Leader Sequences Containing a $tRNA^{Val}_{GAC}$ Gene[a]

Species	Reference
Glycine max	von Allmen and Stutz (1988)
Marchantia polymorpha	Ohyama *et al.* (1988)
Nicotiana tabacum	Tohdoh *et al.* (1981)
Oryza sativa	Hiratsuka *et al.* (1989)
Pisum sativum	Shapiro and Tewari (1986)
Sinapis alba	Przybyl *et al.* (1984)
Spinacea oleracea	Briat *et al.* (1982)
Zea mays	Schwarz *et al.* (1981)

16S rDNAs

Species	Length of Mature RNAs (nt[b])	Reference
Chlamydomonas eugametos	1,448	Durocher *et al.* (1989)
Chlamydomonas moewusii	1,448 (intron of 402 nt)	Durocher *et al.* (1989)
Chlamydomonas reinhardtii	1,475	Dron *et al.* (1982)
Chlorella ellipsoidea	1,583	Yamada (1988)
Euglena gracilis strain Z supplementary 16S rDNA	1,491	Graf *et al.* (1982)
	1,486	Roux *et al.* (1983)
Glycine max	1,499	von Allmen and Stutz (1988)
Marchantia polymorpha	1,496	Kohchi *et al.* (1988)
Nicotiana tabacum	1,489	Tohdoh and Sugiura (1982)
Oryza sativa	1,490	Hiratsuka *et al.* (1989)
Pisum sativum	1,488	Stummann *et al.* (1988)
Pylaiella littoralis[c]	1,505	Markowicz *et al.* (1988a)
Zea mays	1,491	Schwarz and Kössel (1980)

Spacer rDNAs Including Genes for $tRNA^{Ile}_{GAU}$ and $tRNA^{Ala}_{UGC}$

Species	Introns in Spacer tRNA Genes	DNA	RNA	Reference
Chlamydomonas reinhardtii	no introns	X		Schneider and Rochaix (1986)
Chlorella ellipsoidea[d]	no introns	X		Yamada and Shimaji (1987a)
Euglena gracilis strain z	no introns	X		Graf *et al.* (1980)
Glycine max[e]	with introns	X		De Lanversin *et al.* (1987)
Marchantia polymorpha	with introns	X		Ohyama *et al.* (1988)
Nicotiana tabacum	with introns	X		Sugiura *et al.* (1985)

(continues)

TABLE I
Continued

Species	Introns in Spacer tRNA Genes	DNA	RNA	Reference
Olisthodiscus luteus	no introns	X		Delaney and Cattolico (1989)
Oryza sativa	with introns	X		Hiratsuka *et al.* (1989)
Pisum sativum	with introns	X		Stummann *et al.* (1988)
Pylaiella littoralis	no introns	X		Markowicz *et al.* (1988b)
Spinacea oleracea	with introns	X		Massenet *et al.* (1987)
			$tRNA^{Ile}$	Guillemaut and Weil (1982)
Zea mays	with introns	X		Koch *et al.* (1981)
			$tRNA^{Ile}$	Guillemaut and Weil (1982)

23S rDNAs

Species	Length of Mature RNAs (nt)	Reference
Chlorella ellipsoidea	2,965 (intron of 243 nt)	Yamada and Shimaji (1987a)
Euglena gracilis strain Z	2,890	Yepiz-Plascencia *et al.* (1988)
Marchantia polymorpha	2,811	Kohchi *et al.* (1988)
Nicotiana tabacum	2,810	Takaiwa and Sugiura (1982b)
Oryza sativa	2,883	Hiratsuka *et al.* (1989)
Pisum sativum	2,813	Stummann *et al.* (1988)
Zea mays	2,890	Edwards and Kössel (1981)

7S and 3S rDNAs

Species	Length of Mature RNAs (nt)	Reference
Chlamydomonas reinhardtii	7S: 282 3S: 47	Rochaix and Darlix (1982)

4. 5S rDNAs/rRNAs

Species	DNA	RNA	Length of Mature RNAs (nt)	Reference
Acorus calamus		X	105	Bobrova *et al.* (1987)
Allium tuberosum		X	103	Zhen-Qi *et al.* (1986a)

TABLE I
Continued

Species	DNA	RNA	Length of Mature RNAs (nt)	Reference
Apium graveolens		X	103	Zhen-Qi *et al.* (1986a)
Commelina communis		X	103	Zhen-Qi *et al.* (1986a)
Dryopteris accuminata		X	103	Takaiwa *et al.* (1982)
Hordeum vulgare		X	95	Zhen-Qi *et al.* (1986a)
Jungermannia subulata		X	100	Yamano *et al.* (1985)
Ligularia calthifolia		X	105	Bobrova *et al.* (1987)
Lycopersicon esculentum		X	103	Zhen-Qi *et al.* (1986b)
Marchantia polymorpha		X	103	Yamano *et al.* (1985)
	X		102	Kohchi *et al.* (1988)
Mnium rugicum		X	103	Troitsky *et al.* (1984)
Nicotiana tabacum	X		103	Takaiwa and Sugiura (1980a)
		X	103	Takaiwa and Sugiura (1980b)
Oryza sativa	X		94	Hiratsuka *et al.* (1989)
Pisum sativum	X		105	Stummann *et al.* (1988)
Triticum aestivum		X	96	Wildeman and Nazar (1980)
Spinacea oleracea		X	106	Kumagai *et al.* (1982)
Spirodela oligorhiza	X		102	Keus *et al.* (1983a)
Zea mays	X		95	Edwards *et al.* (1981)

5S rDNAs/rRNAs

Species	DNA	RNA	Length of Mature RNAs (nt)	Reference
Chlamydomonas reinhardtii	X	X	121	Schneider *et al.* (1985)
Chlorella ellipsoidea	X		120	Yamada and Shimaji (1986)
Cyanophora paradoxa		X	118	Maxwell *et al.* (1986)
Dryopteris accuminata		X	122	Takaiwa and Sugiura (1982a)
Euglena gracilis strain B	X		121	El-Gewely *et al.* (1984)
Euglena gracilis strain Z	X	X	115	Karabin *et al.* (1983)
Jungermannia subulata		X	119	Yamano *et al.* (1984)
Juniperus media		X	121	Van den Eynde *et al.* (1988)
Lemna minor		X	120	Dyer and Bowman (1979)
Marchantia polymorpha	X		122	Kohchi *et al.* (1988)
		X	119	Yamano *et al.* (1984)
Nicotiana tabacum	X	X	121	Takaiwa and Sugiura (1981)
Oryza sativa	X		120	Hiratsuka *et al.* (1989)
Pelargonium zonale	X		121	Eck *et al.* (1987)
Pisum sativum	X		121	Stummann *et al.* (1988)
Porphyra umbilicalis		X	121	Van den Eynde *et al.* (1988)
Spinacea oleracea	X		121	Audren *et al.* (1987)
		X	122	Delihas *et al.* (1981)
Spirodela oligorhiza	X		120	Keus *et al.* (1983a)
Vicia faba		X	121	Van den Eynde *et al.* (1988)
Zea mays	X		121	Strittmatter and Kössel (1984)

(continues)

TABLE I
Continued

Trailer rDNAs Containing a $tRNA^{Arg}_{ACG}$ Gene[f]	
Species	References
Marchantia polymorpha	Ohyama *et al.* (1988)
Nicotiana tabacum	Kato *et al.* (1985)
Oryza sativa	Hiratsuka *et al.* (1989)
Pelargonium zonale	Hellmund *et al.* (1984)
Pisum sativum	Shapiro and Tewari (1986)
Spirodela oligorhiza	Keus *et al.* (1984)
Zea mays	Dormann-Przybyl *et al.* (1986)

[a] Various 16S rDNA leader sequences, particularly from algae not containing tRNA genes, may be found as part of the 16S rDNA sequences listed in the following section of this table.

[b] nt, nucleotide.

[c] The bimolecular plastid genome of this alga also contains sequences encoding a pseudo-16S rRNA.

[d] Due to the inversion and separation of the second half of the rRNA operon of *C. ellipsoidea* (see Fig. 2), the rDNA spacer is not a true spacer.

[e] From the rDNA spacer of *Glycine max* only the sequences of the tRNA exons have been published.

[f] Various trailer sequences, particularly from algae not containing tRNA genes, may be found as part of 23S and/or 5S rDNA sequences listed in the previous sections of this table.

similarity leads to 62% (mitochondrial 18S–chloroplast 16S) and 21% (mitochondrial 18S–cytoplasmic 17S) of alignment length (Chao *et al.*, 1984).

Compilations of 16S and 23S rRNA sequences in the form of secondary structures also contain many of the plastid rRNA species (Gutell *et al.*, 1985; Gutell and Fox, 1988). On the level of secondary structures, several stems, loops, and bulges characteristic of plastid rRNAs are recognizable. They appear suitable as gross taxonomic markers in analogy to stem, bud, and leaf structures of plants at the macroscopic level.

Chloroplast rDNA sequences are commonly included for establishing quantitative phylogenetic relationships among species (for review, see Chapter 11A). For this purpose, comparisons of small ribosomal subunit RNA sequences have been used most extensively (Cedergren *et al.*, 1988; Giovannoni *et al.*, 1988). But in addition to this, phylogenetic trees have also been derived from 23S rRNAs (Leffers *et al.*, 1987; Cedergren *et al.*, 1988) and from 4.5S and 5S rRNAs (Hori *et al.*, 1985; Bobrova *et al.*, 1987; Van den Eynde *et al.*, 1988). All rRNA-based phylogenetic trees, not unexpectedly, place the plastids within the eubacterial kingdom as a coherent group with cyanobacteria.

III. EXPRESSION OF CHLOROPLAST rRNA GENES

A. RNA Polymerases Involved in rRNA Transcription

The transcription of chloroplast genes is mediated by a chloroplast-specific DNA-dependent RNA polymerase, whose origin and diversity is a matter of dispute (for reviews see Briat *et al.*, 1986; Weil, 1987; Ruf and Kössel, 1988; Little and Hallick, 1988). Concerning the diversity of plastid RNA polymerases, different enzymes have been postulated for the transcription of mRNA, tRNA, and rRNA coding genes (Gruissem *et al.*, 1983a,b) and for different developmental stages of plastids (Reiss and Link, 1985). In particular, the complex termed transcriptionally active chromosome (TAC), first described by Hallick *et al.* (1976) and Schiemann *et al.* (1977), consisting of plastid DNA-bound RNA polymerase, was isolated from many chloroplast species including algae and higher plants (for review, see Briat *et al.*, 1987). In *Euglena,* this complex exclusively produces rRNA run-off transcripts, whereas a soluble RNA polymerase activity from the stroma fraction specifically transcribes genes coding for tRNA (Greenberg *et al.*, 1984; Gruissem *et al.*, 1986) and mRNA (Orozco *et al.*, 1985; Little and Hallick, 1988). In higher plants, however, the situation appears to be different. Here also the soluble RNA polymerase can correctly initiate and transcribe rRNA genes as was demonstrated for pea (Sun *et al.*, 1986) and spinach (Briat *et al.*, 1987) chloroplasts. On the other hand, the TAC from spinach chloroplasts apparently can also transcribe protein-coding genes (Briat *et al.*, 1987). These observations and the existence of consensus promoter sequences, which are identical for rRNA and protein-coding genes (see Table II), lend support to the concept of a single plastid core RNA polymerase, whose specificity is modulated by σ-like or other yet unknown cofactors. However, further refinement of *in vitro* transcription systems from chloroplast will be necessary to obtain a final picture with regard to the diversity and specificity of chloroplast RNA polymerases and in particular of rRNA transcription. (Plastid transcription is discussed in more detail in Chapter 4A.)

B. rRNA Promoters and Regulation of Transcript Levels

All plastid rRNA genes are arranged within operons, which contain complete sets of rRNA genes and which are transcribed in the form of

TABLE II

Promoter Sequences of Plastid rRNA Operons[a]

Species	Sequences[b]	Methods of Identification[c]	Reference
Chlamydomonas moewusii	GAATTGACAAAAAAATTAAATAAGTAAATTAAGAATAGATAA (−200)	1	Durocher *et al.* (1989)
Chlorella ellipsoidea[d]	P_1TTGACGTAAATATAAATTGTAGGTTATCGTATAAAAGTATGA (↓ ↓ −200)	1, 2, 3	Yamada and Shimaji (1987b)
Chlorella ellipsoidea	P_2TTGACGTAAATAAGCTCTGTAGGTTATTCTTGAAGGAAGATA (↓)		
Euglena gracilis st. B[e]	AAAGCGCGCTTTTAGTGTACGCATTATAGTAAATGTGCCCCT (↓ −50)	1, 2, 4	McGarvey *et al.* (1988)
Marchantia polymorpha	TTGACATAATAGGGTAGGTATGGGTATACTAGAAATGAGCTT (−110)	1	Kohchi *et al.* (1988)
Nicotiana tabacum	TTGACGTGAGGGGGCAGGGATGGCTATATTTCTGGGAGCGAA (−110)	1	Tohdoh *et al.* (1981)
Pisum sativum	TTGACACAAGGGGGGGTAAGGCCATATAATATTTATGGGAGG (↓ −150)	1	Stummann *et al.* (1988)
		2, 3, 4	Sun *et al.* (1989)
Pylaiella littoralis	TTGACAATTTCTATATATGGCATGTACTCTGATAGGTAAGGT (−90)	1	Markowicz *et al.* (1988a)

	−110		
Sinapis alba	T T G A C G T G A G G G G G T A G G G G T A G C T A T A T T T C T G G G A G C G A A	1	Przybyl *et al.* (1984)
	−170 ↓		
Spinacea oleracea[f]	P_1 T T T T G A C T T G C T C C C C C G C C G T G A T T G A A T G A G A A T G A A T A A	1, 2, 3	Lescure *et al.* (1985)
	↓ −110		
Spinacea oleracea	P_2 T T G A C G T G A G G G G G T A G G G A T G G C T A T A T T T C T G G G A G C G A A		
	−120 ↓		
Spirodela oligorhiza	T T G A C G T G A T A G G G T A G G G A T G G C T A T A T T G C T G G G A G C C G A	1, 2	Keus *et al.* (1983b)
	−120 ↓		
Zea mays	T T G A C G T G A T A G G G T A G G G T T G G C T A T A C T G C T G G T G G C G A A	1, 3	Schwarz *et al.* (1981)
		2, 4	Strittmatter *et al.* (1985)
		5	Delp *et al.* (1987)
Consensus sequences[g]	−35 −10		
Plastid rRNA promoters	T T G A C R T ——— 14–17 bp ——— T A T a Y T		
Plastid promoters	T T G a N a N ——— 12–23 bp ——— T A T a a T		Kung and Lin (1985)
E. coli promoters	T T G A C a t ——— 15–21 bp ——— T A t A a T		Hawley and McClure (1983)

[a] Various 16S rDNA leader sequences, which may contain unidentified promoter sequences, may be found in the 16S rDNA sequences referred to in Table I.

[b] Sequences of the −10 and −35 regions are underlined. Positions of transcriptional start sites are marked by an arrow. The positions marked by $-n$ indicate the distances to the start of the 16S rRNA genes.

[c] Methods applied for identification are as follows: 1, sequence homology; 2, S1 mapping; 3, *in vitro* transcription or RNA polymerase binding; 4, *in vitro* capping of transcript; 5, mutagenesis.

[d] Promoters P_1 and P_2 refer to the 16S/$tRNA^{Ile}$ and $tRNA^{Ala}$/23S/5S transcripts, respectively (see Fig. 2).

[e] Sequences homologous to the −35 consensus sequences are not recognizable in this promoter.

[f] Promoters P_1 and P_2 refer to the two tandem promoters of the $tDNA^{Val}$/16S rDNA intergenic region.

[g] N, any of the four nucleotides; R, purine nucleotides; Y, pyrimidine nucleotides.

long primary transcripts (Hartley and Head, 1979; Hartley, 1979; Kössel *et al.*, 1982) ranging in size from about 5 (*Euglena*) to 8 kb (*Chlamydomonas* and higher plants; see Fig. 2). This implies positioning of a single transcriptional start site in the region located upstream of the 16S rRNA genes. An exception to this situation is the split rRNA operon from the *Chlorella* plastome (see Fig. 2), where a second promoter is necessary for the transcript containing $tRNA^{Ala}$, 23S rRNA, and 5S rRNA (Yamada and Shimaji, 1987b).

Identification of transcription start sites has mainly been achieved by S_1 mapping, binding of RNA polymerases, and homology of the identified regions with prokaryotic promoter consensus sequences (Table II). It should, however, be noted that S_1 mapping does not allow discrimination between the primary 5′ ends and 5′ termini produced by processing. In view of this general uncertainty, the 5′ terminus of the primary rRNA transcript from maize chloroplasts has been identified by the more specific *in vitro* capping reaction, which is only possible with RNAs carrying a 5′-triphosphate group characteristic of the primary 5′ end (Strittmatter *et al.*, 1985). The start site identified by this method at position 117 upstream of the 16S rRNA gene has confirmed earlier conclusions based on *Escherichia coli* RNA polymerase binding of a promoter within the $tDNA^{Val}$/16S rDNA intergenic region (Schwarz *et al.*, 1981). That the sequences upstream of this start site function as active promoter also *in vivo* could be demonstrated by testing the wild type and a mutated sequence in an *E. coli* promoter test vector (Delp *et al.*, 1987).

In the spinach rRNA operon, the situation appears to be more complicated, because two promoters, P_1 and P_2, within the $tDNA^{Val}$/16S rDNA region apparently exist and precede transcriptional start sites at positions 164 (P_1) and 114 (P_2) upstream of the 16S rRNA gene (Briat *et al.*, 1987). Both promoters are active in the heterologous *E. coli* system *in vivo* and *in vitro*, although P_1 is more active than P_2; however, in the homologous system, promoter P_2 is apparently not active, because the corresponding *in vivo* transcript could not be detected by S_1 mapping and because purified RNA polymerase from spinach chloroplasts did not initiate transcription *in vitro* at P_2. On the other hand, the spinach chloroplast promoter P_2 is homologous with respect to position and sequence to the maize chloroplast promoter, and sequences similar to P_2 are also well conserved in other plant species (see Table II). The inactivity of P_2 in spinach is, therefore, unexpected and it is tempting to speculate that at least in spinach, and perhaps in the other plant species, P_2 is a regulated promoter that is only active under certain physiological conditions.

No coherent picture has so far emerged for the regulation of plastid rRNA operons *in vivo* or *in vitro*. Piechulla *et al.* (1985) have observed a 3-fold decrease of plastid rRNA levels during ripening of tomato fruits, which, however, is relatively small as compared with the more than 10-fold reduction of various mRNA levels. Aguettaz *et al.* (1987) have found a close correlation of 16S rRNA levels with the decrease and increase of plastid DNA levels during the growth cycle of mixotrophic spinach cell suspensions, in which amyloplasts change to chloroplasts. This apparent absence of direct regulation of rRNA transcription has led to the proposal that rRNA transcription is regulated by gene dosage and that the varying multiplicity of plastid DNA causes synthesis of various amounts of rRNAs and ribosomes (Bendich, 1987). Evidence for transcriptional regulation *in vivo* has, however, been presented by Klein and Mullet for barley and maize plastids (Mullet and Klein, 1987; Klein and Mullet, 1990). They have measured steady-state levels as well as transcriptional-rate levels of 16S rRNA in dependence of light. In 4.5-day-old dark-grown barley seedlings a rapid 10-fold decrease of rRNA transcription is observed after both continuation of dark and exposure to light which is suggestive for a direct downward transcriptional regulation. Under identical conditions, however, the steady state levels of 16S rRNA remain constant in the dark and decrease only threefold under light conditions. Therefore, the steady-state levels of rRNA appear to be regulated largely by the balance between stability and breakdown and only in part by transcriptional regulation. A minimum of rRNA transcription is reached after growth in the dark for 8 days in barley and 9 days in maize (Klein and Mullet, 1990). Subsequent illumination for 16 (barley) or 14 (maize) hours leads to a tenfold (barley) and fivefold (maize) increase of 16S rRNA transcription rate without concomitant increase of the steady state levels of 16S rRNA. This again is strong evidence for direct transcriptional regulation (but contrary to the above mentioned situation in 4.5-day-old dark-grown seedlings for an *upward* regulation) and clearly shows that for interpretation of steady-state levels, both, new synthesis and degradation rates have to be correlated. As an example where the rRNA loss appears to be caused by an increase of degradation rather than decrease of synthesis, studies on the senescence of detached wheat leaves (Lamattina et al., 1988) have to be mentioned.

From the existence of several conserved hairpin structures and of short reading frames coding for peptides of 8 (spinach, tobacco, and mustard; for the latter, see Przybyl *et al.*, 1984) and 19 (maize) amino acids, regulation of the plastid rRNA operons by translation-mediated attenuation has been suggested (Briat *et al.*, 1983). Data obtained with

an *E. coli* heterologous test system, in which the spinach rRNA leader sequences were linked to the *E. coli* Gal K gene, are consistent with this proposal (Laboure *et al.*, 1988). Final proof of this regulatory mechanism would, however, require a homologous test system.

So far no RNA species connecting the upstream $tRNA^{Val}$ sequence with the 16S rRNA leader region have been detected (Kössel *et al.*, 1982; Strittmatter *et al.*, 1985). In view of the rRNA promoter(s) identified distal to the $tRNA^{Val}$ gene, the simplest interpretation would be a completely independent transcription of the $tRNA^{Val}$ gene by the promoters tentatively identified upstream of the $tRNA^{Val}$ genes (Schwarz *et al.*, 1981; Tohdoh *et al.*, 1981; Strittmatter *et al.*, 1985). A transcriptional terminator distal to this tRNA gene has, however, not been identified. A palindromic structure consistently observed immediately distal to the $tRNA^{Val}$ genes of various species is not followed by an oligoA/U track, as is typical for prokaryotic terminators. Therefore, in spite of the intergenic promoter(s), $tRNA^{Val}$ is still possibly cotranscribed with the rRNA operon but is processed so rapidly that a common precursor RNA cannot be detected. In this case, the conserved palindromic structures distal to the $tRNA^{Val}$ sequences might function as processing sites. S_1 data obtained with *in vitro E. coli* RNA polymerase transcripts from the tobacco rRNA leader region lend support to a cotranscription of the $tRNA^{Val}$ gene (Tohdoh *et al.*, 1981). This transcription would imply that at least two promoters are functioning simultaneously for the rRNA operon: the promoter(s) identified in the $tDNA^{Val}$/16S rDNA intergenic region and the promoter(s) proximal to the $tRNA^{Val}$ genes. Further investigations, particularly for characterization of and the discrimination between transcriptional terminators and processing sites, are necessary to clarify these questions.

C. Cotranscription of 5S rRNA

The apparent absence of 5S rRNA sequences in the large primary rRNA transcript (Hartley, 1979; Kössel *et al.*, 1982) and the presence of promoterlike structures in the 4.5S/5S intergenic region of duckweed, spinach, and *Pelargonium* (Keus *et al.*, 1983a, 1984; Audren *et al.*, 1987; Eck *et al.*, 1987) have been used as arguments for support of an independent transcription of 5S rRNA genes; however, several other observations contradict this possibility. First, the putative intergenic promoter sequences are not conserved in several other species. Second, the putative promoter sequences of spinach and *Pelargonium* are not active in the

homologous *in vitro* transcription system of spinach (Audren *et al.*, 1987) and in the heterologous *E. coli* minicell system (Eck *et al.*, 1987), respectively. Third, the existence of full-length intergenic RNA connecting the 4.5 and 5S rRNA sequences could be demonstrated by S_1 and primer extension techniques in maize (Strittmatter and Kössel, 1984) and spinach (Audren *et al.*, 1987). Therefore, an independent 5S rRNA transcription appears unlikely, and the failure to detect 5S rRNA sequences in the primary transcript may be explained by rapid cleavage of the 5S rRNA sequences from the precursor rRNA.

D. Termination of rRNA Transcription: Is $tRNA^{Arg}$ a Trailer tRNA?

The question of whether the $tRNA^{Arg}$ gene, which is present in the distal regions of 5S rDNA in higher plants, is transcribed separately or as a trailer tRNA contained in the primary rRNA transcript has been addressed by several authors. In the case of *Spirodela* (Keus *et al.*, 1984) and tobacco (Kato *et al.*, 1985), promoterlike sequences have been detected in the 5S rDNA/$tDNA^{Arg}$ intergenic regions. However, functional testing of these putative promoters or identification of the respective 5′ RNA termini by S_1 mapping and *in vitro* capping has not been reported. Dormann-Przybyl *et al.* (1986) have demonstrated by S_1 and primer extension mapping that in maize a nearly full-length intergenic RNA is connected with the 3′ end of 5S rRNA or the 5′ end of $tRNA^{Arg}$ precursors and that the 5′ termini of the latter are not preceded by recognizable promoterlike sequences. On the other hand, an RNA species containing the complete intergenic region and both the 5S rRNA and $tRNA^{Arg}$ sequences could not be detected. Therefore, the question whether the $tRNA^{Arg}$ is a genuine trailer tRNA (which would also imply that termination of the rRNA operon transcription occurs distal to the $tRNA^{Arg}$ gene) or the $tRNA^{Arg}$ gene is a separate transcriptional unit remains unanswered. Development of functional test systems for transcriptional termination of plastid genes will be essential to resolve this question.

In view of the existence of most of the intergenic RNA and for the reason of functional homology with the trailer tRNA genes of bacterial rRNA operons (for review, see Fournier and Ozeki, 1985), which are cotranscribed, we favor the possibility of a cotranscription of the $tRNA^{Arg}$ and termination of the primary rRNA transcript distal to this trailer tRNA gene. It should be noted that the alternative modes of $tRNA^{Arg}$ transcription are not mutually exclusive; therefore (analogously

to the situation discussed for the $tRNA^{Val}$ gene; see above), cotranscription and separate transcription of the $tRNA^{Arg}$ gene might occur simultaneously or as alternative modes of different developmental stages.

E. Processing of Primary rRNA Transcripts

The pathway leading from the primary rRNA transcript to mature rRNAs and tRNAs requires several processing activities.

1. As judged from bacterial rRNA precursors, several nucleases are likely involved in endonucleolytic cleavage and exonucleolytic trimming of rRNA-flanking sequences.
2. The action of several tRNA-specific processing enzymes is required, such as RNaseP for producing the 5′ termini and other nucleases and CCA nucleotidyltransferase for maturation of the 3′ termini of the spacer tRNAs and perhaps also for leader and trailer tRNAs.
3. Splicing activities are necessary for removal of introns that occur in spacer tRNA or 23S rRNA precursors.

With the exception of tRNA-specific enzymes (for review, see Chapter 7A), no processing enzymes specifically involved in rRNA processing and maturation have been characterized from plastids. However, at least certain putative processing sites of the primary rRNA transcript have been identified by S_1 and reverse transcriptase mapping. For instance, in maize the internucleotide bond between positions 30 and 31 upstream of the mature 16S rRNA sequences probably represents a cleavage site for an RNaseIII-like activity, because this position is embedded in a single-stranded bulge, which is flanked by double-stranded stems (Strittmatter *et al.*, 1985). The folding of 5′- and 3′-terminal precursor sequences of 16S rRNA to double-stranded stem structures similar to the *E. coli* 16S precursors is, per se, a strong argument for the involvement of an RNaseIII-like activity in the endonucleolytic cleavage of plastid 16S rRNA (Schwarz *et al.*, 1981; Tohdoh and Sugiura, 1982). Similar stem structures are found between 5′-flanking sequences of 23S and 3′-flanking sequences of 4.5S rRNAs (the equivalent of the 3′-terminal region of bacterial 23S rRNAs), from which an RNaseIII-like activity has been postulated also for the plastid 23S–4.5S rRNA processing (Edwards *et al.*, 1981).

Several putative processing sites have been mapped in the RNAs encoded by the intergenic regions between 23S and 4.5S rDNA, 4.5S and 5S rDNA, and 5S rDNA and $tDNA^{Arg}$ of maize (Strittmatter and Kössel, 1984; Dormann-Przybyl *et al.*, 1986) and spinach (Audren *et al.*, 1987).

As evident from the identification of several cleavage sites obtained with 5′- and 3′-terminally labeled S_1 probes or by primer extension mapping, the cleavage reactions apparently do not follow a certain order but rather occur in a random sequence.

In only a few cases the terminal positions of 16S and 23S rRNAs have been identified by S_1 mapping, whereas in the majority of rRNA sequences derived from rDNA sequences (see Table I) terminal positions are deduced only from a comparison with other species. In view of the limited accuracy of the S_1 technique, not to speak of the uncertainty of homology deductions, the identifications of the termini of both 16S and 23S rRNAs remain to be confirmed by more direct methods such as end-labeling for virtually all the 16S and 23S rRNA (rDNA) sequences listed in Table I. As an exception, the 3′-terminal sequences of *Euglena* 16S rRNA have been analyzed by direct RNA sequencing, which not only has allowed identification of the 3′-terminal nucleotide but also of one N^6-dimethyladenosine residue within the 3′-terminal loop region and one 2′-O-methylutridine residue in the 5′-flanking, single-stranded region of the last stem (Steege *et al.*, 1982). Also in this work, the susceptibility of plastid 16S rRNA from *Euglena* to cleavage with colicine E3 has been demonstrated. N^6-dimethyladenosine residues, presumably as two successive residues near the 3′ end, and a 7-methylguanosine residue, presumably located at the position homologous to position 527 of *E. coli* 16S rRNA, have been detected in chloroplast 16S rRNA from pea by antibody reactions with ribosomal 30S subunits (Trempe and Glitz, 1981; Trempe *et al.*, 1982). No other examples of modified bases have been reported for plastid rRNAs. In view of the conservation of sequences that surround the positions of modified bases in the *E. coli* 16S rRNA, base modifications of plastid rRNAs probably also occur in the other positions known to be modified in *E. coli* (Schwarz and Kössel, 1980).

F. Hidden Breaks of 23S rRNAs

When large ribosomal subunit RNAs from higher plant chloroplasts are isolated under denaturing conditions, only a small portion of intact 23S rRNA is obtained, while the major portion appears as several distinct species of shorter chain lengths (Leaver, 1973). Pulse-chase experiments have shown that this fragmentation is due to postmaturation processes (Mache *et al.*, 1978). The existence of a relatively larger fraction of unbroken 23S rRNA in plastids of older leaves as compared with younger leaves indicates higher stability of 23S rRNA in matured plastids as compared with newly developing plastids (Lerbs and Wollgiehn, 1982). In

view of these observations, this fragmentation is probably not a mere isolation artifact. It is, however, reminiscent of the 5′-terminal fragmentation of *Chlamydomonas* 23S rRNA, which leads to 7S and 3S rRNAs, and of the 3′-terminal fragmentation of higher plant chloroplast 23S rRNAs, which results in 4.5S rRNAs (see Fig. 2). In analogy to this, hidden breaks of plastid 23S rRNA may reflect processing steps necessary for certain unknown structural requirements and/or functions of plastid ribosomes. Ribosome inactivation and/or initiation of ribosome degradation may be considered as alternative functions of hidden breaks.

The positions of two major hidden breaks have been determined in maize chloroplast 23S rRNA by S_1 mapping (Kössel *et al.*, 1985). The position of the two hidden breaks coincide with stem-loop structures, which in comparison with bacterial 23S rRNA are either specifically inserted in plastid 23S rRNA, as in the case of the stem-loop structure at positions 953–965 (which is completely missing in *E. coli* 23S rRNA), or specifically reduced, as in the case of the stem-loop structure at positions 1822–1841. Extra stem-loop structures per se, which are also present in other positions of the 23S rRNA secondary structure, or altered stem-loop structures or the loss of stem-loop structures, as caused for 16S rRNA by a 23-bp deletion in 16S rDNA, apparently do not lead to hidden breaks. Therefore, which additional structural features or nucleases are required to mediate the specific cleavage reactions leading to hidden breaks at the two altered stem-loop structures exclusively are not clear. An attractive hypothesis that remains to be tested is self-cleavage of 23S rRNA by ribozymelike structures residing in 23S rRNA.

IV. ANTIBIOTIC RESISTANCES RESIDING IN PLASTID rRNAs

Plastome-encoded mutations conferring drug or herbicide resistance are likely to be useful tools as selectable markers in establishing stable chloroplast transformations. For this reason, a search for such mutants started soon after the detection of chloroplast DNA in the 1960s.

Maternally inherited mutants conferring resistance against antibiotics that interact with prokaryoticlike ribosomes were isolated from *Chlamydomonas* as early as 1965 by Gillham (for reviews, see Montandon *et al.*, 1985; Fromm *et al.*, 1987). Originally, such mutants were considered to produce altered ribosomal proteins, as was observed for most of the

corresponding bacterial mutants. However, Montandon *et al.* (1985) first reported a streptomycin-resistant *Euglena* mutant, which showed a base substitution in the plastid 16S rRNA gene. The corresponding change in the 16S rRNA is a C–U transition in position 876, which is equivalent to the C residue at position 912 of *E. coli* 16S rRNA. This position is conserved in all eubacterial and plastid small subunit rRNAs and together with the conserved G residue at position 888 forms a base pair at the base of helix 29 (Dams *et al.*, 1988). Thus the C–U transition conferring streptomycin resistance leads to a conversion of this CG to a UG base pair, thereby reducing the stability of helix 29. On the other hand, all archaebacterial and eukaryotic small subunit rRNAs already contain a U residue (and therefore a helix 29 of reduced stability) in this position, which is in accordance with the streptomycin resistance of eukaryotic ribosomes. That this C–U transition confers streptomycin resistance also to *E. coli* ribosomes has been demonstrated by site-directed mutagenesis (Montandon *et al.*, 1986). More recently, streptomycin mutants residing in plastid 16S rRNA genes have also been reported for tobacco (Etzold *et al.*, 1987; Fromm *et al.*, 1989) and *Chlamydomonas* (Gauthier *et al.*, 1988). This work not only extends the observation of rRNA-based antibiotic resistance to higher plant and *Chlamydomonas* chloroplasts but at the same time demonstrates that besides position 912 at least two other positions, which correspond to A-523 and C-525 of *E. coli* 16S rRNA, can confer streptomycin resistance by mutation. Complementary to the streptomycin resistance residing in mutated 16S rRNA structures, a plastome mutant coding for an altered ribosomal protein S12 of tobacco chloroplasts has been described recently (Galili *et al.*, 1989).

Several mutants leading to maternally inherited spectinomycin resistance from *Chlamydomonas* (Harris *et al.*, 1987) and tobacco (Fromm *et al.*, 1987) have been characterized as mutations of the respective 16S rRNA coding genes. These mutations reduce the stability of a conserved stem structure, which is composed of the complementary pentanucleotides corresponding to positions 1063–1067 and 1189–1193 of helix 36 of the *E. coli* 16S rRNA secondary structure. Interestingly, the nucleotide sequences of the wild-type 16S rRNA genes reported for maize (Schwarz and Kössel, 1980), *Euglena* (Graf *et al.*, 1982), and rice (Hiratsuka *et al.*, 1989) chloroplasts predict weakened structures in the same region. This suggests a natural spectinomycin resistance of these species, which has been confirmed experimentally, at least with maize and *Euglena* (Fromm *et al.*, 1987).

Point mutations conferring antibiotic resistance have also been characterized for 23S rRNA genes of chloroplasts. In *Chlamydomonas*, an

erythromycin- and spiramycin-resistant mutant results from a C–G transversion at the position of chloroplast 23S rRNA, which corresponds to position 2611 of the peptidyltransferase loop region of *E. coli* 23S rRNA (Gauthier *et al.*, 1988). In tobacco, lincomycin-resistant mutants have been mapped at the positions corresponding to 2032, 2055, and 2059 of *E. coli* 23S rRNA, which again alter the peptidyltransferase region (Cseplö *et al.*, 1988).

ACKNOWLEDGMENTS

We wish to thank Dr. G. Igloi for critical reading of the manuscript and Mrs. E. Schiefermayr for technical help. Our own contributions cited in this article have been supported by grants from the Deutsche Forschungsgemeinschaft (SFB 46 and 206) and from the Fonds der Chemischen Industrie (FRG).

REFERENCES

Aguettaz, P., Seyer, P., Pesey, H., and Lescure, A. M. (1987). Relations between the plastid gene dosage and the levels of 16S RNA and RBC-L gene transcripts during amyloplast to chloroplast change in mixotrophic spinach cell suspensions. *Plant Mol. Biol.* **8,** 169–177.

Anderson, L. A., and Smillie, R. M. (1966). Binding of chloramphenicol by ribosomes from chloroplasts. *Biochem. Biophys. Res. Commun.* **23,** 535–539.

Audren, H., Bisanz-Seyer, C., Briat, J.-F., and Mache, R. (1987). Structure and transcription of the 5S rRNA gene from spinach chloroplasts. *Curr. Genet.* **12,** 263–270.

Bedbrook, J. R., Kolodner, R., and Bogorad, L. (1977). *Zea mays* chloroplast ribosomal RNA genes are part of a 22000 base pair inverted repeat. *Cell* **11,** 739–749.

Bendich, A. J. (1987). Why do chloroplast and mitochondria contain so many copies of their genome? *BioEssays* **6,** 279–282.

Bobrova, V. K., Troitsky, A. V., Ponomarev, A. G., and Antonov, A. S. (1987). Low-molecular-weight rRNAs sequences and plant phylogeny reconstruction: Nucleotide sequences of chloroplast 4.5S rRNAs from *Acorus calamus* (Araceae) and *Ligularia calthifolia* (Asteraceae). *Pl. Syst. Evol.* **156,** 13–27.

Briat, J. F., Dron, M., Loiseaux, S., and Mache, R. (1982). Structure and transcription of the spinach chloroplast rDNA leader region. *Nucleic Acids Res.* **10,** 6865.

Briat, J. F., Dron, M., and Mache, R. (1983). Is transcription of higher plant chloroplast ribosomal operons regulated by premature termination? *FEBS Lett.* **163,** 1–5.

Briat, J. F., Lescure, A. M., and Mache, R. (1986). Transcription of the chloroplast DNA: A review. *Biochimie* **68,** 981–990.

Briat, J.-F., Bisanz-Seyer, C., and Lescure, A.-M. (1987). In vitro transcription initiation of the rDNA operon of spinach chloroplast by a highly purified soluble homologous RNA polymerase. *Curr. Genet.* **11,** 259–263.

Cedergren, R., Gray, M. W., Abel, Y., and Sankoff, D. (1988). The evolutionary relationships among known life forms. *J. Mol. Evol.* **28,** 98–112.

Chao, S., Sederoff, R., and Levings III, C. S. (1984). Nucleotide sequence and evolution of the 18S rRNA gene in maize mitochondria. *Nucleic Acids Res.* **12,** 6629–6644.

Chun, E. H. L., Vaughan, M. H., and Rich, A. (1963). The isolation and characterization of DNA associated with chloroplast preparations. *J. Mol. Biol.* **7,** 130–141.

Cseplö, A., Etzold, T., Schell, J., and Schreier, P. H. (1988). Point mutations in the 23S rRNA genes of four lincomycin resistant *Nicotiana plumbaginifolia* mutants could provide new selectable markers for chloroplast transformation. *Mol. Gen. Genet.* **214,** 295–299.

Dams, E., Hendriks, L., Van de Peer, Y., Neefs, J.-M., Smits, G., Vandenbempt, I., and De Wachter, R. (1988). Compilation of small ribosomal subunit RNA sequences. *Nucleic Acids Res.* **16,** r87–r173.

Delaney, T. P., and Cattolico, R. A. (1989). Chloroplast ribosomal DNA organization in the chromophytic alga *Olisthodiscus luteus*. *Curr. Genet.* **15,** 221–229.

De Lanversin, G., Pillay, D. T. N., and Jacq, B. (1987). Sequence studies on the soybean chloroplast 16S–23S rDNA spacer region. Comparison with other angiosperm sequences and proposal of a generalized RNA secondary structure model for the intergenic regions. *Plant Mol. Biol.* **10,** 65–82.

Delihas, N., Andersen, J., Sprouse, H. M., and Dudock, B. (1981). The nucleotide sequence of the chloroplast 5S ribosomal RNA from spinach. *Nucleic Acids Res.* **9,** 2801–2805.

Delp, G., Igloi, G. L., Beck, C. F., and Kössel, H. (1987). Functional in vivo verification in *Escherichia coli* of promoter activities from the rDNA/tDNA^{Val}(GAC) leader region of *Zea mays* chloroplasts. *Curr. Genet.* **12,** 241–246.

Dormann-Przybyl, D., Strittmatter, G., and Kössel, H. (1986). The region distal to the ribosomal RNA operon from chloroplasts of maize contains genes coding for tRNA^{Arg}(ACG) and tRNA^{Asp}(GUU). *Plant. Mol. Biol.* **7,** 419–431.

Dron, M., Rahire, M., and Rochaix, J.-D. (1982). Sequence of the chloroplast 16S rRNA gene and its surrounding regions of *Chlamydomonas reinhardtii*. *Nucleic Acids Res.* **10,** 7609–7620.

Durocher, V., Gauthier, A., Bellemare, G., and Lemieux, C. (1989). An optional group I intron between the chloroplast small subunit rRNA genes of *Chlamydomonas moewusii* and *C. eugametos*. *Curr. Genet.* **15,** 277–282.

Dyer, T. A., and Bowman, C. M. (1979). Nucleotide sequences of chloroplast 5S ribosomal ribonucleic acid in flowering plants. *Biochem. J.* **183,** 595–604.

Eck, R., Lazarus, C. M., Baldauf, F., Metzlaff, M., and Hagemann, R. (1987). Sequence analysis and *Escherichia coli* minicell transcription test of *Pelargonium* plastid 5S rDNA. *Mol. Gen. Genet.* **207,** 514–516.

Edwards, K., and Kössel, H. (1981). The rRNA operon from *Zea mays* chloroplasts: Nucleotide sequence of 23S rDNA and its homology with *E. coli* 23S rDNA. *Nucleic Acids Res.* **9,** 2853.

Edwards, K., Bedbrook, J., Dyer, T., and Kössel, H. (1981). 4.5S rRNA from *Zea mays* chloroplast shows structural homology with the 3′ end of prokaryotic 23S rRNA. *Biochem. Internat.* **2,** 533.

El-Gewely, M. R., Helling, R. B., and Dibbits, J. G. T. (1984). Sequence and evolution of the regions between the RRN operons in the chloroplast genome of *Euglena gracilis bacillaris*. *Mol. Gen. Genet.* **194,** 432–443.

Etzold, T., Fritz, C. C., Schell, J., and Schreier, P. H. (1987). A point mutation in the chloroplast 16S rRNA gene of a streptomycin resistant *Nicotiana tabacum*. *FEBS Lett.* **219,** 343–346.

Fournier, M. J., and Ozeki, H. (1985). Structure and organization of the transfer ribonucleic acid genes of *Escherichia coli* K-12. *Microbiol. Rev.* **49,** 379–397.

Fromm, H., Edelman, M., Aviv, D., and Galun, E. (1987). The molecular basis for rRNA-dependent spectinomycin resistance in *Nicotiana* chloroplasts. *EMBO J.* **6,** 3233–3237.

Fromm, H., Galun, E., and Edelman, M. (1989). A novel site for streptomycin resistance in the "530 loop" of chloroplast 16S ribosomal RNA. *Plant Mol. Biol.* **12,** 499–505.

Galili, S., Fromm, H., Aviv, D., Edelman, M., and Galun, E. (1989). Ribosomal protein S12 as a site for streptomycin resistance in *Nicotiana* chloroplasts. *Mol. Gen. Genet.* **218,** 289–292.

Gauthier, A., Turmel, M., and Lemieux, C. (1988). Mapping of chloroplast mutations conferring resistance to antibiotics in *Chlamydomonas*: Evidence for a novel site of streptomycin resistance in the small subunit rRNA. *Mol. Gen. Genet.* **214,** 192–197.

Giovannoni, S. J., Turner, S., Olsen, G. J., Barns, S., Lane, D. J., and Pace, N. R. (1988). Evolutionary relationships among cyanobacteria and green chloroplasts. *J. Bacteriol.* **170,** 3584–3592.

Graf, L., Kössel, H., and Stutz, E. (1980). Sequencing of the 16S-23S spacer in a ribosomal RNA operon of *Euglena gracilis* chloroplast DNA reveals two tRNA genes. *Nature* **286,** 908–910.

Graf, L., Roux, E., Stutz, E., and Kössel, H. (1982). Nucleotide sequence of a *Euglena gracilis* chloroplast gene coding for the 16S rRNA: Homologies to *E. coli* and *Zea mays* chloroplast 16S rRNA. *Nucleic Acids Res.* **10,** 6369.

Greenberg, B. M., Gruissem, W., and Hallick, R. B. (1984). Accurate processing and pseudouridylation of chloroplast transfer RNA in a chloroplast transcription system. *Plant Mol. Biol.* **3,** 97–109.

Gruissem, W., Greenberg, B. M., Zurawski, G., Prescott, D. M., and Hallick, R. B. (1983a). Biosynthesis of chloroplast transfer RNA in a spinach chloroplast transcription system. *Cell* **35,** 815–828.

Gruissem, W., Narita, J. B., Greenberg, B. M., Prescott, D. M., and Hallick, R. B. (1983b). Selective *in vitro* transcription of chloroplast genes. *J. Cell Biochem.* **22,** 31–46.

Gruissem, W., Greenberg, B. M., Zurawski, G., and Hallick, R. B. (1986). Chloroplast gene expression and promoter identification in chloroplast extracts. *Methods Enzymol.* **118,** 253–270.

Guillemaut, P., and Weil, J. H. (1982). The nucleotide sequence of the maize and spinach chloroplast isoleucine transfer RNA encoded in the 16S to 23S rDNA spacer. *Nucleic Acids Res.* **10,** 1653–1659.

Gutell, R. R., and Fox, G. E. (1988). A compilation of large subunit RNA sequences presented in a structural format. *Nucleic Acids Res.* **16,** r175–r269.

Gutell, R. R., Weiser, B., Woese, C. R., and Noller, H. F. (1985). Comparative anatomy of 16S-like ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **32,** 155–216.

Hallick, R. B., Lipper, C., Richards, O. C., and Rutter, W. (1976). Isolation of a transcriptionally active chromosome from chloroplasts of *Euglena gracilis*. *Biochemistry* **15,** 3039–3045.

Harris, E. H., Boynton, J. E., and Gillham, N. W. (1987). Interaction of nuclear and chloroplast mutations in biogenesis of chloroplast ribosomes in *Chlamydomonas*. *In* "Molecular and Cellular Aspects of Algal Development" (W. Weissner and D. G. Robinson, eds.), pp. 142–149. Springer Verlag, Heidelberg.

Hartley, M. R. (1979). The synthesis and origin of chloroplast low molecular weight rRNA in spinach. *Eur. J. Biochem.* **96,** 311–320.

Hartley, M. R., and Head, C. (1979). The synthesis of chloroplast high-molecular-weight ribosomal ribonucleic acid in spinach. *Eur. J. Biochem.* **96,** 301–309.

Hawley, D. K., and McClure, W. R. (1983). Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* **11,** 2237–2255.

Hellmund, D., Metzlaff, M., and Serfling, E. (1984). A transfer RNA^{Arg} gene of *Pelargonium* chloroplasts, but not a 5S RNA gene, is efficiently transcribed after injection into *Xenopus* oocyte nuclei. *Nucleic Acids Res.* **12,** 8253–8268.

Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.-R., Meng, B.-Y., Li, Y.-Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., and Sugiura, M. (1989). The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of cereals. *Mol. Gen. Genet.* **217,** 185–194.

Hori, H., Lim, B.-L., and Osawa, S. (1985). Evolution of green plants as deduced from 5S rRNA sequences. *Proc. Natl. Acad. Sci. USA* **82,** 820–823.

Karabin, G. D., Narita, J. O., Dodd, J. R., and Hallick, R. B. (1983). *Euglena gracilis* chloroplast ribosomal RNA transcription units: Nucleotide sequence polymorphism in 5S rRNA genes and 5S rRNAs. *J. Biol. Chem.* **258,** 14790–14796.

Kato, A., Takaiwa, F., Shinozaki, K., and Sugiura, M. (1985). Location and nucleotide sequence of the genes for tobacco chloroplast $tRNA^{Arg}$ (ACG) and $tRNA^{Leu}$ (UAG). *Curr. Genet.* **9,** 405–409.

Keus, R. J. A., Roovers, D. J., Dekker, A. F., and Groot, G. S. P. (1983a). The nucleotide sequence of the 4.5S and 5S rRNA genes and flanking regions from *Spirodela oligorhiza* chloroplasts. *Nucleic Acids Res.* **11,** 3405–3410.

Keus, R. J. A., Dekker, A. F., van Roon, M. A., and Groot, G. S. P. (1983b). The nucleotide sequences of the regions flanking the genes coding for 23S, 16S and 4.5S ribosomal RNA on chloroplast DNA from *Spirodela oligorhiza*. *Nucleic Acids Res.* **11,** 6465–6474.

Keus, R. J. A., Dekker, A. F., Kreuk, K. C. J., and Groot, G. S. P. (1984). Transcription of ribosomal DNA in chloroplasts of *Spirodela oligorhiza*. *Curr. Genet.* **9,** 91–98.

Klein, R. R., and Mullet, J. E. (1990). Light-induced transcription of chloroplast genes. *J. Biol. Chem.* **265,** 1895–1902.

Koch, W., Edwards, K., and Kössel, H. (1981). Sequencing of the 16S-23S spacer in a ribosomal RNA operon of *Zea mays* chloroplast DNA reveals two split tRNA genes. *Cell* **25,** 203–213.

Kohchi, T., Shirai, H., Fukuzawa, H., Sano, T., Komano, T., Umesono, K., Inokuchi, H., Ozeki, H., and Ohyama, K. (1988). Structure and organization of *Marchantia polymorpha* chloroplast genome. IV. Inverted repeat and small single copy regions. *J. Mol. Biol.* **203,** 353–372.

Koller, B., Roux, E., Montandon, P.-E., and Stutz, E. (1988). A chimeric transcript containing a 16S rRNA and a potential mRNA in chloroplasts of *Euglena gracilis*. *Plant Mol. Biol.* **10,** 339–347.

Kössel, H., Edwards, K., Koch, W., Langridge, P., Schiefermayr, E., Schwarz, Zs., Strittmatter, G., and Zenke, G. (1982). Structural and functional analysis of an rRNA operon and its flanking tRNA genes from *Zea mays* chloroplasts. *Nucleic Acids Res. Symp. Series* **11,** 117.

Kössel, H., Edwards, K., Fritzsche, E., Koch, W., and Schwarz, Zs. (1983). Phylogenetic significance of nucleotide sequence analysis. *In* "Protein and Nucleic Acids in Plant Systematics" (U. Jensen and D. E. Fairbrothers, eds.). Springer-Verlag, Berlin and Heidelberg.

Kössel, H., Natt, E., Strittmatter, G., Fritzsche, E., Gozdzicka-Jozefiak, A., and Przybyl, D. (1985). Structure and expression of rRNA operons from plastids of higher plants.

In "Molecular form and function of the plant genome" (L. van Vloten-Doting, G. S. P. Groot, and T. C. Hall, eds.), pp. 183–199. Plenum Publishing Corp., New York.

Kumagai, I., Pieler, T., Subramanian, A. R., and Erdmann, V. A. (1982). Nucleotide sequence and secondary structure analysis of spinach chloroplast 4.5S RNA. *J. Biol. Chem.* **257,** 12924–12928.

Kung, S. D., and Lin, C. M. (1985). Chloroplast promoters from higher plants. *Nucleic Acids Res.* **13,** 7543–7549.

Laboure, A.-M., Lescure, A.-M., and Briat, J.-F. (1988). Evidence for a translation-mediated attenuation of a spinach chloroplast rDNA operon. *Biochimie* **70,** 1343–1352.

Lamattina, L., Pinedo, M., Yudi, V. P., Pont Lezica, R. F., and Conde, R. D. (1988). Estimation of rRNA synthesis and degradation rates in senescing wheat leaves. *Arch. Biochem. Biophys.* **260,** 285–292.

Leaver, C. J. (1973). Molecular integrity of chloroplast rRNA. *Biochem. J.* **135,** 237–240.

Leffers, H., Kjems, J., Ostergaard, L., Larsen, N., and Garrett, R. A. (1987). Evolutionary relationships amongst archaebacteria. A comparative study of 23S ribosomal RNAs of a sulfur-dependent extreme thermophile, an extreme halophile and a thermophilic methanogen. *J. Mol. Biol.* **195,** 43–61.

Lerbs, S., and Wollgiehn, R. (1982). The occurrence of 2 nicks in the heavy chloroplast ribosomal RNA of *Nicotiana rustica*. *Biochem. Physiol. Pflanz.* **177,** 431–439.

Lescure, A.-M., Bisanz-Seyer, C., Pesey, H., and Mache, R. (1985). *In vitro* transcription initiation of the spinach chloroplast 16S rRNA gene at two tandem promoters. *Nucleic Acids Res.* **13,** 8787–8796.

Little, M. C., and Hallick, R. B. (1988). Chloroplast *rpo*A, *rpo*B, and *rpo*C genes specify at least three components of a chloroplast DNA-dependent RNA polymerase active in tRNA and mRNA transcription. *J. Biol. Chem.* **263,** 14302–14307.

Lyttelton, J. W. (1960). Nucleoproteins of white clover. *Biochem. J.* **74,** 82–90.

Mache, R., Jalliffier-Verne, M., Rozier, C., and Loiseaux, S. (1978). Molecular weight determination of precursor, mature and post-mature plastid ribosomal RNA from spinach using fully denaturing conditions. *Biochim. Biophys. Acta* **517,** 390–399.

Markowicz, Y., Loiseaux-De Goër, S., and Mache, R. (1988a). Presence of a 16S rRNA pseudogene in the bi-molecular plastid genome of the primitive brown alga *Pylaiella littoralis*. Evolutionary implications. *Curr. Genet.* **14,** 599–608.

Markowicz, Y., Mache, R., and Loiseaux-De Goër, S. (1988b). Sequence of the plastid rDNA spacer region of the brown alga *Pylaiella littoralis* (L.) Kjellm. Evolutionary significance. *Plant Mol. Biol.* **10,** 465–469.

Massenet, O., Martinez, P., Seyer, P., and Briat, J.-F. (1987). Sequence organization of the chloroplast ribosomal spacer of *Spinacea oleracea* including the 3′ end of the 16S rRNA and the 5′ end of the 23S rRNA. *Plant Mol. Biol.* **10,** 53–63.

Maxwell, E. S., Liu, J., and Shively, J. M. (1986). Nucleotide sequences of *Cyanophora paradoxa* cellular and cyanelle-associated 5S ribosomal RNAs: The cyanelle as a potential intermediate in plastid evolution. *J. Mol. Evol.* **23,** 300–304.

McGarvey, P., Helling, R. B., Lee, J.-Y., Engelke, D. R., and El-Gewely, R. M. (1988). Initiation of rrn transcription in chloroplasts of *Euglena gracilis bacillaris*. *Curr. Genet.* **14,** 493–500.

Messing, J., Carlson, J., Hagen, G., Rubenstein, I., and Oleson, A. (1984). Cloning and sequencing of the ribosomal RNA genes in maize: The 17S region. *DNA* **3,** 31.

Mets, L. J., and Bogorad, L. (1971). Mendelian and uniparental alteration in erythromycin binding by plastid ribosomes. *Science* **174,** 707–709.

Miyata, T., Kikuno, R., and Ohshima, Y. (1982). A pseudogene cluster in the leader region of the *Euglena* chloroplast 16S-23S rRNA genes. *Nucleic Acids Res.* **10,** 1771–1780.

Montandon, P. E., Nicolas, P., Schürmann, P., and Stutz, E. (1985). Streptomycin-resistance of *E. gracilis* chloroplasts: Identification of a point mutation in the 16S rRNA gene in an invariant position. *Nucleic Acids Res.* **13,** 4299–4310.

Montandon, P. E., Wagner, R., and Stutz, E. (1986). *E. coli* ribosomes with a C912 to U base change in the 16S rRNA are streptomycin resistant. *EMBO J.* **5,** 3705–3708.

Mullet, J. E., and Klein, R. R. (1987). Transcription and RNA stability are important determinants of higher plant chloroplast RNA levels. *EMBO J.* **6,** 1571–1579.

Ohyama, K., Fukuzawa, H., Kohchi, T., Sano, T., Sano, S., Shirai, H., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., and Ozeki, H. (1988). Structure and organization of *Marchantia polymorpha* chloroplast genome. I. Cloning and gene identification. *J. Mol. Biol.* **203,** 281–298.

Orozco, E. M., Rushlow, K. E., Dodd, J. R., and Hallick, R. B. (1980). *Euglena gracilis* chloroplast rRNA transcription units: Nucleotide sequence homology between the 16S-23S ribosomal spacer and the 16S rRNA leader regions. *J. Biol. Chem.* **255,** 10997–11003.

Orozco, E. M., Mullet, J. E., and Chua, N.-H. (1985). An *in vitro* system for accurate transcription initiation of chloroplast protein genes. *Nucleic Acids Res.* **13,** 1283–1302.

Piechulla, B., Chonoles Imlay, K. R., and Gruissem, W. (1985). Plastid gene expression during fruit ripening in tomato. *Plant Mol. Biol.* **5,** 373–384.

Przybyl, D., Fritzsche, E., Edwards, K., Kössel, H., Falk, H., Thompson, J. A., and Link, G. (1984). The ribosomal RNA genes from chloroplasts of mustard (*Sinapis alba* L.): Mapping and sequencing of the leader region. *Plant Mol. Biol.* **3,** 147–158.

Reiss, T., and Link, G. (1985). Characterization of transcriptionally active DNA–protein complexes from chloroplasts and etioplasts of mustard (*Sinapis alba* L.). *Eur. J. Biochem.* **148,** 207–212.

Ris, H., and Plaut, W. (1962). Ultrastructure of DNA containing areas in the chloroplast of *Chlamydomonas*. *J. Cell. Biol.* **13,** 383–391.

Rochaix, J.-D., and Darlix, J.-L. (1982). Composite structure of the chloroplast 23S rRNA genes of *Chlamydomonas reinhardtii*. *J. Mol. Biol.* **159,** 383–395.

Rochaix, J.-D., Rahire, M., and Michel, F. (1985). The chloroplast ribosomal intron of *Chlamydomonas reinhardtii* codes for a polypeptide related to mitochondrial maturases. *Nucleic Acids Res.* **13,** 975–984.

Roux, E., and Stutz, E. (1985). The chloroplast genome of *Euglena gracilis:* The mosaic structure of a DNA segment linking the extra 16S rRNA gene with *rrn* operon A. *Curr. Genet.* **9,** 221–227.

Roux, E., Graf, L., and Stutz, E. (1983). Nucleotide sequence of a truncated rRNA operon of the *Euglena gracilis* chloroplast genome. *Nucleic Acids Res.* **11,** 1957–1968.

Ruf, M., and Kössel, H. (1988). Structure and expression of the gene coding for the α-subunit of DNA-dependent RNA polymerase from the chloroplast genome of *Zea mays*. *Nucleic Acids Res.* **16,** 5741–5754.

Schiemann, J., Wollgiehn, R., and Parthier, B. (1977). Isolation of a transcription active RNA polymerase DNA complex from *Euglena* chloroplasts. *Biochem. Physiol. Pflanzen.* **171,** 474–478.

Schneider, M., and Rochaix, J.-D. (1986). Sequence organization of the chloroplast ribosomal spacer of *Chlamydomonas reinhardtii:* Uninterrupted rRNAile and tRNAala genes and extensive secondary structure. *Plant Mol. Biol.* **6,** 265–270.

Schneider, M., Darlix, J.-L., Erickson, J., and Rochaix, J.-D. (1985). Sequence organization of repetitive elements in the flanking regions of the chloroplast ribosomal unit of *Chlamydomonas reinhardtii*. *Nucleic Acids Res.* **13,** 8531–8541.

Schwarz, Zs., and Kössel, H. (1980). The primary structure of 16S rDNA from *Zea mays* chloroplast is homologous to *E. coli* 16S rRNA. *Nature* **283,** 739–742.

Schwarz, Zs., Kössel, H., Schwarz, E., and Bogorad, L. (1981). A gene coding for $tRNA^{Val}$ is located near 5' terminus of 16S rRNA gene in *Zea mays* chloroplast genome. *Proc. Natl. Acad. Sci. USA* **78,** 4748–4752.

Shapiro, D. R., and Tewari, K. K. (1986). Nucleotide sequences of transfer RNA genes in the *Pisum sativum* chloroplast DNA. *Plant Mol. Biol.* **6,** 1–12.

Steege, D. A., Graves, M. C., and Spremulli, L. L. (1982). *Euglena gracilis* chloroplast small subunit rRNA. Sequence and base-pairing potential of the 3' terminus, cleavage by colicin E3. *J. Biol. Chem.* **257,** 10430–10439.

Strittmatter, G., and Kössel, H. (1984). Cotranscription and processing of 23S, 4.5S and 5S rRNA in chloroplasts from *Zea mays*. *Nucleic Acids Res.* **12,** 7633–7647.

Strittmatter, G., Gozdzicka-Jozefiak, A., and Kössel, H. (1985). Identification of an rRNA operon promoter from *Zea mays* chloroplasts which excludes the proximal tRNAVal (GAC) from the primary transcript. *EMBO J.* **4,** 599–604.

Stummann, B. M., Lehmbeck, J., Bookjans, G., and Henningsen, K. W. (1988). Nucleotide sequence of the single ribosomal RNA operon of pea chloroplast DNA. *Physiol. Plant* **72,** 139–146.

Sugiura, M., Shinozaki, K., and Ohme, M. (1985). Tobacco chloroplast genes for transfer RNAs. *In* "Molecular Form and Function of the Plant Genome" (L. van Vloten-Doting, G. S. P. Groot, and T. C. Hall, eds.), pp. 325–334. Plenum Publishing Corp., New York.

Sun, E., Shapiro, D. R., Wu, B. W., and Tewari, K. K. (1986). Specific *in-vitro* transcription of 16S rRNA gene by pea chloroplast RNA polymerase. *Plant Mol. Biol.* **6,** 429–439.

Sun, E., Wu, B.-W., and Tewari, K. K. (1989). *In vitro* analysis of the pea chloroplast 16S rRNA gene promoter. *Mol. Cell. Biol.* **9,** 5650–5659.

Takaiwa, F., and Sugiura, M. (1980a). Nucleotide sequence of the 4.5S and 5S ribosomal RNA genes from tobacco chloroplasts. *Mol. Gen. Genet.* **180,** 1–4.

Takaiwa, F., and Sugiura, M. (1980b). The nucleotide sequence of 4.5S ribosomal RNA from tobacco chloroplasts. *Nucleic Acids Res.* **8,** 4125–4129.

Takaiwa, F., and Sugiura, M. (1981). Heterogeneity of 5S RNA species in tobacco chloroplasts. *Mol. Gen. Genet.* **182,** 385–389.

Takaiwa, F., and Sugiura, M. (1982a). The nucleotide sequence of chloroplast 5S ribosomal RNA from a fern, *Dryopteris acuminata*. *Nucleic Acids Res.* **10,** 5369–5373.

Takaiwa, F., and Sugiura, M. (1982b). The complete sequence of a 23S rRNA gene from tobacco chloroplasts. *Eur. J. Biochem.* **124,** 13–19.

Takaiwa, F., Kusuda, M., and Sugiura, M. (1982). The nucleotide sequence of chloroplast 4.5S rRNA from a fern, *Dryopteris acuminata*. *Nucleic Acids Res.* **10,** 2257–2260.

Tohdoh, N., and Sugiura, M. (1982). The complete nucleotide sequence of a 16S ribosomal RNA gene from tobacco chloroplasts. *Gene* **17,** 213–218.

Tohdoh, N., Shinozaki, K., and Sugiura, M. (1981). Sequence of a putative promoter region for the rRNA genes of tobacco chloroplasts. *Nucleic Acids Res.* **9,** 5399–5406.

Trempe, M. R., and Glitz, D. G. (1981). Chloroplast ribosome structure. Electron microscopy of ribosomal subunits and localization of N^6,N^6-dimethyladenosine by immunoelectronmicroscopy. *J. Biol. Chem.* **256,** 11873–11879.

Trempe, M. R., Ohgi, K., and Glitz, D. G. (1982). Ribosome structure. Localization of 7-methylguanosine in the small subunits of *Escherichia coli* and chloroplast ribosomes by immunoelectron microscopy. *J. Biol. Chem.* **257,** 9822–9829.

Troitsky, A. V., Bobrova, V. K., Ponomarev, A. G., and Antonov, A. S. (1984). The nucleotide sequence of chloroplast 4.5S rRNA from *Mnium rugicum* (*Bryophyta*): Mosses also possess this type of RNA. *FEBS Lett.* **176,** 105–109.

Van den Eynde, H., De Baere, R., De Roeck, E., Van de Peer, Y., Vandenberghe, A., Willekens, P., and De Wachter, R. (1988). The 5S ribosomal RNA sequences of a red

algal rhodoplast and a gymnosperm chloroplast. Implications for the evolution of plastids and cyanobacteria. *J. Mol. Evol.* **27,** 126–132.

von Allmen, J. M., and Stutz, E. (1988). The soybean chloroplast genome: Nucleotide sequence of a region containing tRNAVal (GAC) and 16S rRNA gene. *Nucleic Acids Res.* **16,** 1200.

Weil, J. H. (1987). Organization and expression of the chloroplast genome. *Plant Sci.* **49,** 149–157.

Wildeman, A. G., and Nazar, R. W. (1980). Nucleotide sequence of wheat chloroplastid 4.5S ribonucleic acid. Sequence homology in 4.5S RNA species. *J. Biol. Chem.* **255,** 11896–11900.

Yamada, T. (1988). Nucleotide sequence of the chloroplast 16S ribosomal RNA gene from the unicellular green alga *Chlorella ellipsoidea*. *Nucleic Acids Res.* **16,** 9865.

Yamada, T., and Shimaji, M. (1986). Nucleotide sequence of the 5S rRNA gene from the unicellular green alga *Chlorella ellipsoidea*. *Nucleic Acids Res.* **14,** 9529.

Yamada, T., and Shimaji, M. (1987a). An intron in the 23S rRNA gene of the *Chlorella* chloroplasts: Complete nucleotide sequence of the 23S rRNA gene. *Curr. Genet.* **11,** 347–352.

Yamada, T., and Shimaji, M. (1987b). Splitting of the ribosomal RNA operon on chloroplast DNA from *Chlorella ellipsoidea*. *Mol. Gen. Genet.* **208,** 377–383.

Yamano, Y., Ohyama, K., and Komano, T. (1984). Nucleotide sequences of chloroplast 5S ribosomal RNA from cell suspension cultures of the liverworts *Marchantia polymorpha* and *Jungermannia subulata*. *Nucleic Acids Res.* **12,** 4621–4624.

Yamano, Y., Kohchi, T., Fukuzawa, H., Ohyama, K., and Komano, T. (1985). Nucleotide sequences of chloroplast 4.5S ribosomal RNA from a leafy liverwort, *Jungermannia subulata*, and a thalloid liverwort, *Marchantia polymorpha*. *FEBS Lett.* **185,** 203–207.

Yepiz-Plascencia, G. M., Jenkins, M. E., and Hallick, R. B. (1988). Nucleotide sequence of the *Euglena gracilis* chloroplast 23S rRNA gene of the *rrnC* operon. *Nucleic Acids Res.* **16,** 9340.

Zhen-Qi, C., Hong, Z., Guo-Ya, L., and Xiao-Yang, L. (1986a). The nucleotide sequences of chloroplast 4.5S rRNA from four species of plants, celery (*Apium graveolens*), barley (*Hordeum vulgare*), Chinese chive (*Allium tuberosum*) and dayflower (*Commelina communis*). *FEBS Lett.* **200,** 193–196.

Zhen-Qi, C., Xiao, X., and E-Sheng, W. (1986b). The nucleotide sequence of 4.5S rRNA from tomato chloroplasts. *Biochim. Biophys. Acta* **866,** 89–91.

CHAPTER 7

tRNAs and tRNA Genes of Plastids

L. Maréchal-Drouard
M. Kuntz
J. H. Weil

Institut de Biologie Moléculaire des Plantes du CNRS
Université Louis Pasteur
Strasbourg cedex, France

I. ORGANIZATION AND NUMBER OF CHLOROPLAST tRNA GENES

The circular chloroplast (cp) genomes of land plants typically possess a large single-copy (LSC) and a small single-copy (SSC) region separated by two large inverted repeated (IR) sequences (which contain the rRNA genes), with the exception of some leguminous plants that do not contain the large inverted repeats (for a review, see Chapter 2A). In tobacco (Fig. 1), the cp genome codes for 30 tRNA species (Shinozaki *et al.*, 1986;

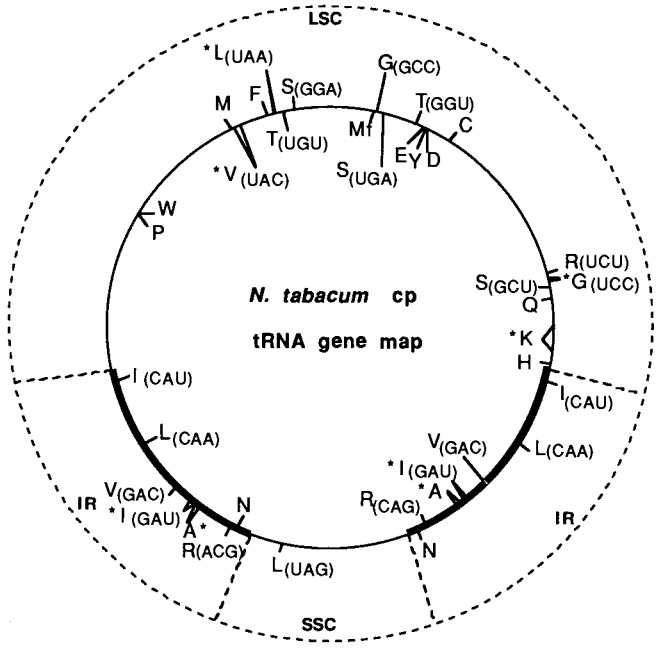


FIG. 1. Transfer RNA gene organization of the chloroplast (cp) genome of *Nicotiana tabacum*. The inverted repeats (IR) are shown with bold lines. LSC and SSC indicate the large and small single-copy regions, respectively. The tRNA genes are shown by the one-letter codes for amino acids with their anticodons in parentheses (when several isoaccepting species are present). Asterisks indicate split genes. Genes shown on the outside of the circle are transcribed counterclockwise; those inside are transcribed clockwise.

Sugiura and Wakasugi, 1989); 23 encoded by single genes (located in the LSC or SSC region) and 7 encoded by genes present twice (in the IR sequences). In rice, the same set of tRNA genes was found (Hiratsuka *et al.*, 1989), but in this case the $tRNA^{His}$ gene, which is at the edge of the LSC region in tobacco, is now included in the IR sequences. In the case of the liverwort *Marchantia polymorpha* (Ohyama *et al.*, 1986), the IR sequences are shorter and code only for five tRNA species (the other tRNA genes are located in the single-copy regions). However, a total of 31 tRNA species are encoded by the liverwort cp genome; the additional tRNA gene (with respect to tobacco and rice) is a $tRNA^{Arg}$(CCG) gene. A single $tRNA^{Glu}$ gene, which encodes the tRNA involved in δ-aminolevulinate (DALA) synthesis (see Section XII), and a single $tRNA^{Gln}$ gene, which encodes a tRNA mischarged by glutamate (see Section XIII), have been identified in these cp genomes.

Euglena gracilis is a unicellular green alga that has an unusual cp genome organization in that it does not contain the usual two large IR

sequences but, instead, has three sequences that are repeated in tandem and each code for the 23S, 16S, and 5S rRNAs and two tRNAs, namely $tRNA^{Ala}$ and $tRNA^{Ile}$ located in the ribosomal spacer between the 16S and 23S rRNA genes (Orozco *et al.*, 1980). A total of 25 tRNA genes (for a review, see Sprinzl *et al.*, 1989) have been sequenced on the cp genome of *E. gracilis* strain Z (23 are located in the single-copy region of the genome and present as single genes, while there are three copies of the $tRNA^{Ala}$ and $tRNA^{Ile}$ genes), but it should be noted that this genome has not yet been completely sequenced.

Very few tRNA gene sequences from plastids of other unicellular organisms have been determined. In the case of *Chlamydomonas reinhardtii* (a green alga) and *Pylaeilla littoralis* (a brown alga), the spacer separating the 16S and 23S rRNA genes and containing the $tRNA^{Ile}$(GAU) and $tRNA^{Ala}$ genes (Schneider and Rochaix, 1986; Markowicz *et al.*, 1988) has been sequenced. It should be noted that these two tRNA genes have been found in the ribosomal spacer of all plastid genomes studied except in *Chlorella ellipsoidea* (a green alga), where the 16S rRNA gene is followed only by a $tRNA^{Ile}$ gene, whereas the 23S rRNA and $tRNA^{Ala}$ genes are located on the opposite strand (Yamada and Shimaji, 1987). The nucleotide sequences of the plastid $tRNA^{Trp}$, $tRNA^{Gly}$(UCC), $tRNA^{Glu}$ and $tRNA^{Cys}$ genes from *Chlamydomonas reinhardtii* have also been reported (Zhang and Spreitzer, 1989; O'Neill *et al.*, 1990). In the case of *Cyanophora paradoxa* cyanelles (Turner *et al.*, 1989; Evrard *et al.*, 1990), six tRNA genes are known to date, namely the $tRNA^{Ile}$(GAU) and $tRNA^{Ala}$ genes located in the ribosomal spacer (Janssen *et al.*, 1987) and the $tRNA^{Glu}$, $tRNA^{Leu}$(UAA) (Evrard *et al.*, 1988), $tRNA^{Gly}$(GCC), and $tRNA^{Ser}$(GGA) (Kuntz *et al.*, 1988) genes.

II. tRNA PSEUDOGENES

Pseudogenes have been found in a number of cp genomes. For instance, a pseudo-$tRNA^{Ile}$ gene and a pseudo-$tRNA^{Trp}$ gene (El-Gewely *et al.*, 1984) have been found in *E. gracilis* strain *bacillaris*. A pseudo-$tRNA^{Pro}$ gene with a GGG anticodon has been sequenced in liverwort (Ohyama *et al.*, 1986). A $tRNA^{fMet/Gly}$ chimeric structure has been described in wheat (Howe, 1985) and rice (Hiratsuka *et al.*, 1989). A pseudo-$tRNA^{Thr}$ gene has been reported in barley (Oliver and Poulsen, 1984), wheat (Quigley and Weil, 1985), and rice (Hiratsuka *et al.*, 1989).

The determination of the complete nucleotide sequence of rice cp DNA has revealed the presence of four tRNA pseudogenes (namely

pseudo-tRNA^{Gly}, pseudo-tRNA^{Ile}, pseudo-tRNA^{Thr}, and pseudo-tRNA^{Glu}) and of one tRNA$^{fMet/Gly}$ chimeric pseudogene (Hiratsuka *et al.*, 1989). These pseudogenes are present in the LSC region, whereas none are found in the SSC region and IR sequences. Similar pseudogenes are found in wheat (for a review, see Shimada and Sugiura, 1989), but, in contrast, only one pseudogene, the pseudo-tRNA^{Arg} gene, is present in the tobacco cp genome (Kato *et al.*, 1985).

It has been proposed that some of these pseudogenes arose by homologous recombination between two tRNA genes (along conserved regions of the tRNA genes) and that such recombinations are responsible for some of the rearrangements that occurred in the cp genomes (Hiratsuka *et al.*, 1989). The fact that tRNA pseudogenes have been found at the ends of inversions in the cp genome (e.g., rice) supports this theory. It is tempting to speculate that large inversions found in the cp genome of other plants and flanked by tRNA genes (e.g., in the Compositae family and particularly in sunflower [Heyraud *et al.*, 1987]) could also have arisen by recombination between tRNA genes.

III. INTERVENING SEQUENCES IN cp tRNA GENES

No intron has been found in the *E. gracilis* cp tRNA genes, whereas six tRNA genes are interrupted in the case of higher plant (Fig. 1) and liverwort cp genomes (Ohyama *et al.*, 1986; Shinozaki *et al.*, 1986; Hiratsuka *et al.*, 1989). These introns, which are very long as compared with the size of the mature tRNA (325–886 bp in liverwort; 503–2526 bp in tobacco), must be spliced efficiently because the corresponding mature tRNAs are found in relatively high amounts in cps. The introns found in the tRNA^{Ala}, tRNA^{Gly}(UCC), tRNA^{Ile}(GAU), tRNA^{Lys}, and tRNA^{Val}(UAC) genes belong to class II of structured introns, whereas the intron found in the tRNA^{Leu}(UAA) gene belongs to class I of structured introns (Michel and Dujon, 1983). Interestingly, the latter intron is the only class I intron found in land-plant cp genomes up to now, and a similar intron was also found in the corresponding gene in *C. paradoxa* cyanelles (Evrard *et al.*, 1988), but not in *E. gracilis* cps. The class II introns found in the tRNA^{Ala} and tRNA^{Ile}(GAU) genes (located in the spacer of the ribosomal operons) are present in all land plants studied so far but not in other organisms (Markowicz *et al.*, 1988). All the cp tRNA gene introns are located in the anticodon loop, except that of the tRNA^{Gly}(UCC) gene, where the intron is located in the D stem of the

tRNA. Recently, Manhart and Palmer (1990) have determined the nucleotide sequence of the cp $tRNA^{Ala}$ and $tRNA^{Ile}$ genes from three green algae, namely *Coleochaete*, *Nitella*, and *Spirogyra*, and concluded from the distribution of class II introns in these genes that these taxa are part of the lineage that gave rise to land plants.

In cps, splicing of these introns by an autocatalytic process (for a review, see Sharp, 1987) has not been observed, suggesting that specific cp factors are involved in the splicing of these introns. It is interesting to note that the interrupted gene coding for $tRNA^{Lys}$ might also encode a protein that has structural features related to those of maturases (Neuhaus and Link, 1987), but it remains to be demonstrated that such a protein is actually synthesized and involved in splicing of the cp class II introns. Recently, G. Delp, G. L. Igloi, and H. Kössel (unpublished results) have obtained the splicing (and 5′ processing) of a maize cp $tRNA^{Ala}$ precursor *in vitro* using a cp soluble extract.

IV. *IN VITRO* TRANSCRIPTION OF cp tRNA GENES

Transcription of cp tRNA genes has been obtained in nonchloroplastic *in vitro* systems, using purified *Escherichia coli* DNA-dependent RNA-polymerase (e.g., in the case of the tobacco $tRNA^{Val}$(GAC) [Tohdoh *et al.*, 1981]) or a HeLa cell extract, which has been shown to transcribe a *E. gracilis* tRNA gene cluster (Gruissem *et al.*, 1982).

A *E. gracilis* cp soluble crude extract has been shown to transcribe homologous tRNA genes *in vitro* (Greenberg and Hallick, 1986), except the $tRNA^{Ala}$ and $tRNA^{Ile}$(GAU) genes that are located in the ribosomal spacer. However, cotranscription of these two tRNA genes with the rRNA genes was obtained, using a cp DNA-bound RNA polymerase activity, termed transcriptionally active chromosome, which was shown to transcribe preferentially rRNAs (Greenberg *et al.*, 1984b; see also Chapter 6A). Greenberg and Hallick (1986) also stated that the *E. gracilis* cp soluble transcription system did not transcribe the spinach cp $tRNA^{Met_2}$ gene; however, this gene and other spinach cp tRNA genes could be transcribed *in vitro* by a homologous cp soluble extract (Gruissem *et al.*, 1983). Interestingly, this spinach cp transcription system was also found to transcribe a *E. gracilis* tRNA gene cluster.

The use of such a crude transcription system provided valuable information on the cp tRNA gene promoters and on the maturation of tRNA precursors (see Sections V and VI); however, how many different RNA

polymerase activities are present in cps is unclear, and further work on purified cp RNA polymerase is still needed. In maize, transcription of a $tRNA^{His}$ by a purified homologous RNA polymerase has been reported (Schwartz *et al.*, 1981).

Transcription of cp tRNA genes has also been obtained *in organello* in run-on transcription experiments using isolated plastids (Deng *et al.*, 1987; Jayabaskaran *et al.*, 1990).

V. CHLOROPLAST tRNA GENE PROMOTERS

Comparison of the 5'-flanking sequences of land-plant cp tRNA genes as well as of protein genes has revealed the presence of prokaryoticlike promoters upstream of these genes. As illustrated in Figure 2 in the case of *trnE*, two conserved motifs similar to the bacterial "−35" and "−10" boxes can be identified upstream of this gene in land-plant cps and in *C. paradoxa* cyanelles. That these motifs represent the actual promoter elements of this gene is also suggested by S_1 mapping experiments, which demonstrated that a precursor transcript starts 8 bp downstream of the "−10" element (Ohme *et al.*, 1985). However, proof of the involvement of such motifs in the transcription of cp tRNA genes have, up to now, only been obtained in the case of spinach $tRNA^{Met_2}$ by expression of the corresponding gene *in vitro* in a homologous, soluble transcription system (Gruissem *et al.*, 1983): In this case, mutations in both the "−35" and "−10" boxes cause a decrease of transcription (Gruissem and Zurawski, 1985b). However, the mutational analysis revealed that the upstream regions can tolerate considerable sequence alteration and yet retain significant transcriptional activity.

Surprisingly, transcription of some spinach cp tRNA genes was not abolished in this *in vitro* system when the 5' region of these genes was deleted (Gruissem *et al.*, 1986), suggesting that other types of promoters, possibly internal to the coding region, are also involved in the transcription, at least for a subset of cp tRNA genes. This hypothesis is strengthened by the observation that many cp tRNA genes do not exhibit the typical "−35" and "−10" motifs found upstream of *trnE* and of other cp tRNA genes.

More information on cp tRNA gene transcription will probably be obtained in the near future with further studies on the cp RNA polymerase(s) and further mutational studies. The availability of cp transfor-

```
                    "-35"                "-10"
N.t.  CATTCGATTATATTGACAATTTCAAAAAACTGATCATACTATGATCATAGTATGATGGCG..13
V.f.  CATTCCATTATATTGACAATTTCAAAAAACTGTTCATACTATGAACATAGTAGAATGGAA..13
P.s.  CGAATCATTATATTGACAATTTCAAAAAACTGTTCATACTATGAACATAGTAGAATGGAG..13
S.o.  TATTCGATTATATTGACAATTTCAAAAAACTGTTCATACTATGAACATAGTATGATGGCA..12
T.a.  TTTTCGAGTATATTGACAATTCCAAAAAACTGCTCATACTATGATTATAGTATAATCACG..32
O.s.  TCTTCGAGTATATTGACAATTCCAAAAAACTGCTCATACTATCATTATAGTATAATGAGG..32
M.p.  CTAAAAAAACTATTGACAGGAAATCATATAATTGAGTAAGATAAGGTTAAGGGATAAA.....0
C.p.  ATTCTAAAGTAGTTGACATATTTATTTTCATAGTA.TAATATTAAATATATAAATATATC...4
Cons  ------A-tataTTGACAatt-caaaaaactg-tcaTActAT-A--atagta--at----
                    "-35"                "-10"
```

FIG. 2. Alignment of the 5'-flanking region of the chloroplast *trnE* genes from various organisms. "−35" and "−10" indicate the conserved elements resembling the consensus *E. coli* promoter elements. The numbers at the right indicate the number of bases up to the beginning of the coding region. "Cons" indicates the consensus sequence: bases that are conserved in all cases are in capitals; those conserved in all examined angiosperms are in lowercase. The first seven lines show the sequence flanking the cp *trnE* gene from *Nicotiana tabacum* (*N.t.*), *Vicia faba* (*V.f.*), *Pisum sativum* (*P.s.*), *Spinacia oleracea* (*S.o.*), *Triticum aestivum* (*T.a.*), *Oryza sativa* (*O.s.*), and *Marchantia polymorpha* (*M.p.*), respectively. The corresponding region of the cyanelle genome from *Cyanophora paradoxa* (*C.p.*) is also shown.

mation techniques will also provide *in vivo* transcription data. The construction of complete cp DNA transcription maps (Woodbury *et al.*, 1988) could also be helpful. However, such studies are difficult in the case of tRNA genes because some tRNA genes are cotranscribed with other genes and because tRNA precursors are then processed (see Section VI). Unambiguous data have only been obtained in a few cases of cotranscribed genes. For instance, the cp $tRNA^{Glu}$, $tRNA^{Tyr}$, and $tRNA^{Asp}$ genes, which are clustered in all land plants studied, have been shown to be cotranscribed in tobacco (Ohme *et al.*, 1985). Cotranscription of $tRNA^{Ile}$(GAU) and $tRNA^{Ala}$ with the 16S and 23S rRNA genes has also been demonstrated (see Chapter 6A).

In *E. gracilis*, many cp tRNA genes are tightly clustered and most likely form transcriptional units (for a review, see Hallick and Buetow, 1989). It is also noteworthy that no prokaryoticlike promoter sequence has been found upstream of these tRNA genes. Instead, all *E. gracilis* cp tRNA gene clusters contain an upstream conserved sequence that reads $GT^{A}_{T}NTA^{AA}_{TT}A$ and could function as a promoter. Alternatively, because at least some of these genes are transcribed in a HeLa cell extract, these

genes (as well as those spinach cp tRNA genes that do not need any upstream sequence) could be transcribed in a manner similar to the mechanism used by eukaryotic RNA polymerase III.

VI. PROCESSING OF cp tRNAs

Chloroplast tRNAs, like other RNAs, are synthesized as immature mono- or polycistronic precursors. In prokaryotes and eukaryotes, the 5′-leader sequence is removed by a site-specific endonuclease (RNase P) that cleaves precisely at the 5′ end of the mature tRNA. An RNase P-like activity has also been identified in the cp soluble transcription extract of *E. gracilis* and spinach cps (Greenberg *et al.*, 1984a; Greenberg and Hallick, 1986) and in S30 (30,000g supernatant) lysates of spinach and tobacco cps (Yamaguchi-Shinozaki *et al.*, 1987). In *E. coli* and *Bacillus subtilis*, RNase P has been well characterized and shown to contain an RNA moiety. In contrast, cp RNase P apparently does not have an RNA subunit (Wang *et al.*, 1988).

Maturation at the 3′ end of tRNAs is quite different in prokaryotes and eukaryotes. In *E. coli*, the 3′ extension is cleaved by an endonuclease and subsequently trimmed by an exonuclease (RNase D) up to the 3′-CCA end of the mature tRNA. In eukaryotes, an endonuclease cleaves precisely at the 3′ end of the tRNA domain. An RNase activity responsible for the 3′-terminus processing has been detected in the soluble transcription extracts (Greenberg *et al.*, 1984a; Greenberg and Hallick, 1986) of spinach and *E. gracilis* cps and in S30 lysates of spinach and tobacco cps (Hamaguchi-Shinozaki *et al.*, 1987). In spinach, 3′ maturation of cp tRNAs was shown to be produced by a single endonucleolytic cut (Wang *et al.*, 1988).

The cp tRNA genes generally do not encode the 3′-terminal CCA sequence, which must therefore be added posttranscriptionally. However, in many cases the tRNA gene sequence includes the first C of the CCA sequence. Furthermore, a few genes such as $tRNA^{Ala}$(UGC) from *Chlorella ellipsoidea*, $tRNA^{Ile}$(GAU) from *Glycine max*, and $tRNA^{Lys}$(UUU) from *Pisum sativum* have reportedly an encoded 3′-CCA sequence (see Sprinzl *et al.*, 1989). There is also no 3′-CCA sequence in nuclear tRNA genes. A 3′-CCA sequence is present in some bacterial tRNA genes, for instance in *B. subtilis* and in the case of the very few known tRNA genes from cyanobacteria; however, in other bacteria such as *E. coli*, although a CTP,ATP tRNA nucleotidyltransferase activity is present, the 3′-

terminal CCA sequence is always encoded in the tRNA genes. A tRNA nucleotidyltransferase activity has been identified in cps of spinach (Greenberg *et al.*, 1984a) and in the case of the $tRNA^{Phe}$ gene (which encodes the first C of the CCA), this enzyme, after cleavage of the 3′ extension after the first C, adds the second C and the A (Wang *et al.*, 1988).

Thus, despite the prokaryotic origin of the cps, cp tRNA precursors are processed by a pathway similar to that utilized for the maturation of the tRNA precursors coded for by nuclear genes.

Contradictory results were obtained when attempts were made to determine whether the 5′ and 3′ maturation of tRNAs is ordered (Wang *et al.*, 1988; Yamaguchi-Shinozaki *et al.*, 1987) or not (Marion-Poll *et al.*,1988).

Maturation of tRNAs also includes ribonucleoside modifications. Little is known about the enzymes catalyzing these modifications; however, a pseudouridylate synthetase activity has been detected in the spinach cp soluble transcription extract (Greenberg *et al.*, 1984a).

VII. CONCENTRATIONS OF ISOACCEPTING tRNAs AND CODON USAGE IN cps

It has been shown in *E. coli* and in *Saccharomyces cerevisiae* that differences in the population of isoaccepting tRNAs are related to differences in the usage of synonymous codons and that a strong correlation exists between codon usage and tRNA content. Also, in *E. coli*, the rate of translation is limited at some codons by the availability of the matching tRNAs; therefore, codon usage determines translation rate (Sorensen *et al.*, 1989).

In cps, Pfitzinger *et al.* (1987) have shown that a correlation exists between the amounts of tRNAs specific for a given amino acid and the frequency of the codons specifying this amino acid. Furthermore, the population of isoaccepting tRNAs is adjusted to the frequency of synonymous codons (codons coding for the same amino acid) used in cp protein genes (Table I). Therefore, this adjustment probably reflects the selection for an optimum translation rate in cps.

It should also be noted that the tRNA species encoded by genes present in duplicate on the cp genomes (in the IR sequences) are not found in higher amounts than the tRNAs encoded by single-copy genes (Pfit-

TABLE I

Synonymous Codon Usage and Relative Content of Isoaccepting Species in Chloroplasts

	tRNA Corresponding to:								
	Arg	Arg	Gly	Gly	Ile	Ile	Leu	Leu	Leu[a]
Anticodons[b]	UCU	ACG	GCC	UCC	CAU[c]	GAU	UAA	CAA	UAG
Codons	AGA AGG	CGU CGC CGA CGG	GGU GGC	GGA GGG	AUA	AUU AUC	UUA	UUG	CUU CUC CUA CUG
Codon usage in chloroplasts (%)[d]	31	69	49	51	26	74	36	21	43
Percentage of each isoacceptor in its tRNA family (in bean chloroplasts)	26	74	50	50	32	68	36	22	42

[a] Codon reading is according to Sugiura *et al.* (1986).

[b] Possible nucleotide modifications have not been considered.

[c] Modification of C allows recognition of the isoleucine codon AUA and not of the methionine codon AUG (Francis and Dudock, 1982).

[d] Codon usage in the chloroplasts of tobacco, maize, and spinach was taken into account. The values represent the percentage of each codon in the family of synonymous codons. [Modified from Pfitzinger *et al.*, 1987.]

zinger *et al.*, 1987; Pillay *et al.*, 1984), indicating that in cps the relative concentrations of the various tRNAs are not controlled by gene dosage.

VIII. REGULATION OF PLASTID tRNA LEVELS

The strength of their respective promoters probably influences the relative levels of tRNAs in cps, as was demonstrated *in vitro* for the spinach *trnM*2 gene and derivatives of it, in which the genuine *trnM*2 promoter elements have been replaced by promoter elements from other cp genes (Gruissem and Zurawski, 1985a). These experiments have confirmed the concept of promoter strength for cp genes. In addition, because cp mRNA levels are also strongly influenced by RNA stability (Gruissem *et al.*, 1988), RNA stability may also be an important determinant of cp tRNA levels; however, no data demonstrating this possibility have been reported up to now.

Differences in the ratio of spliced and unspliced cp mRNAs have been observed in different tissues (Barkan, 1988), but no data are available in the case of tRNAs. However, in the case of the $tRNA^{Lys}$ gene, the level of a transcript corresponding to an unspliced precursor has been studied during mustard seedling development, and it has been shown that the concentration of this transcript is light-independent and shows a transient peak 48 hr after sowing (Hughes *et al.*, 1987). Chloroplast tRNA levels also may depend on the rate of tRNA maturation at the 5′ and 3′ ends. However, it should be noted that *in vitro* tRNA maturation by a soluble cp extract seems very fast and, therefore, may not play a regulatory role, at least in green leaves. But it should be mentioned that specific binding of cp proteins to the 3′ end of the mustard $tRNA^{Lys}$ precursor was detected (Nickelsen and Link, 1989) and could be involved in the rate of $tRNA^{Lys}$ maturation.

Concerning $tRNA^{Phe}$, we have recently shown that the two isoacceptors species present in bean cps differ only by posttranscriptional modifications and that their relative amounts vary during leaf development: In etiolated leaves, the undermodified $tRNA^{Phe}$ only represents 15% of total cp $tRNA^{Phe}$, during development and greening it increases to reach 60% in 8-day-old leaves, and then it drops down to 9% in senescing leaves (Pfitzinger *et al.*, 1990).

Thus, although limited information is available, both transcriptional and posttranscriptional mechanisms of regulation probably contribute to determine the relative levels of tRNAs in cps.

The mechanisms of regulation that control the overall levels of cp tRNAs during development also remain largely speculative, although it is known that cp tRNAs are present not only in photosynthetically active tissues but also (in lower amounts) in nonphotosynthetic tissues such as roots and young embryogenic cotyledons and in dark-grown plants (Burkard *et al.*, 1972; Merrick and Dure, 1972). More recently, the presence of cp tRNAs was observed in yellow senescent leaves, although the corresponding genes apparently were not transcribed any longer in these organs (Jayabaskaran *et al.*, 1990), whereas transcription of cp tRNA genes was detected in dark-grown maize plants (C. Jayabaskaran *et al.*, unpublished results). In *Euglena*, like in higher plants, cp tRNAs apparently are quantitatively regulated by light (Goins *et al.*, 1973).

IX. CODON RECOGNITION IN cps

As mentioned above, there are only 30 different tRNA species encoded by the higher plant cp genomes, whereas all 61 sense codons are found in cp protein genes, and, according to the "wobble" hypothesis, a minimum of 32 tRNA species would be required to decode the 61 code words. Furthermore, because $tRNA^{Leu}$(UAA) and $tRNA^{Leu}$(CAA) that are present in cps should be able to decode the same leucine codon UGG, an apparent deficit of three tRNA species appears in higher plant cps. In fact, a comparison of the anticodons of the higher plant cp tRNAs with all sense codons shows an apparent deficit of tRNAs that can normally recognize the following codons: GCU/C (Ala), CGC/A/G (Arg), CUU/C (Leu), and CCU/C (Pro) (Sugiura and Wakasugi, 1989). Because it is generally assumed that there is no import of cp tRNAs (unlike some mitochondrial tRNAs that are imported from the cytosol), these observations suggest that a special mechanism of codon recognition might operate in cps, at least in some cases.

A "two-out-of-three" mechanism (only two bases of the anticodon get paired with two bases of the codon) has been shown to operate in an *in vitro* protein-synthesizing system from *E. coli* (Samuelson *et al.*, 1980). According to this mechanism, the single cp $tRNA^{Ala}$(UGC), $tRNA^{Arg}$(ACG), and $tRNA^{Pro}$(UGG) could read all four Ala, Arg, and Pro codons, respectively (by G-C base-pairing of only the second and third bases of the anticodon with the first and the second bases of the codon). In addition, the cp $tRNA^{Leu}$(UAG) could read all four CUN leucine co-

dons due to a U–N wobble, if the U is unmodified (which apparently is the case). In agreement with this hypothesis, H. Pfitzinger and P. Guillemaut (unpublished data) have recently shown that, in fact, the m^7G present at position 36 of the anticodon facilitates the U–U and U–C wobble interaction. Thus, using these mechanisms, the cp-encoded tRNAs should be able to read all possible sense codons.

It is also interesting to note that a comparison of total and polysome-bound tRNAs has shown that all higher plant cp tRNA species are apparently used in protein biosynthesis (Pfitzinger *et al.*, 1987).

X. GENERAL FEATURES OF cp tRNAs

With few exceptions, cp tRNAs possess the invariant and semi-invariant nucleotides characteristic of prokaryotic and eukaryotic cytoplasmic tRNAs. Their primary structure is very similar to prokaryotic tRNAs (the sequence homology is about 70%), but they show little homology to eukaryotic cytoplasmic tRNAs (Sprinzl *et al.*, 1989). The 3'-terminal CCA is not coded in the cp tRNA genes but, rather, is added post-transcriptionally.

Chloroplast tRNAs can generally be aminoacylated *in vitro* using cp or prokaryotic enzymes, whereas they are not charged by the enzymes from the cytoplasm of the same cell (Steinmetz and Weil 1989).

An examination of the modified nucleosides present in the cp tRNAs (Fig. 3) listed in the Sprinzl compilation (Sprinzl *et al.*, 1989) leads to the following general conclusions.

1. Only 32 sites are modified in the cp tRNA molecules, whereas 55 sites are modified in cytosolic tRNAs (for a review, see Björk and Kohli, 1989). However, cp tRNAs are more modified than eubacterial tRNAs (21 modified sites; Björk and Kohli, 1989).
2. Among the special features found in almost all cp tRNAs are the following modified nucleosides: Gm at position 18, D at position 20, m^5U at position 54, and ψ at position 55. In contrast to plant cytosolic tRNAs, the abundance of acp^3U at position 47 in cp tRNAs should be noted.
3. At position 35, ψ has been found in wheat cp $tRNA^{Glu}$ (Sprinzl *et al.*, 1989).
4. At position 36 of cp $tRNA^{Leu}$(UAG), the modified nucleoside m^7G has been found (Sprinzl *et al.*, 1989). Nucleosides at this position are

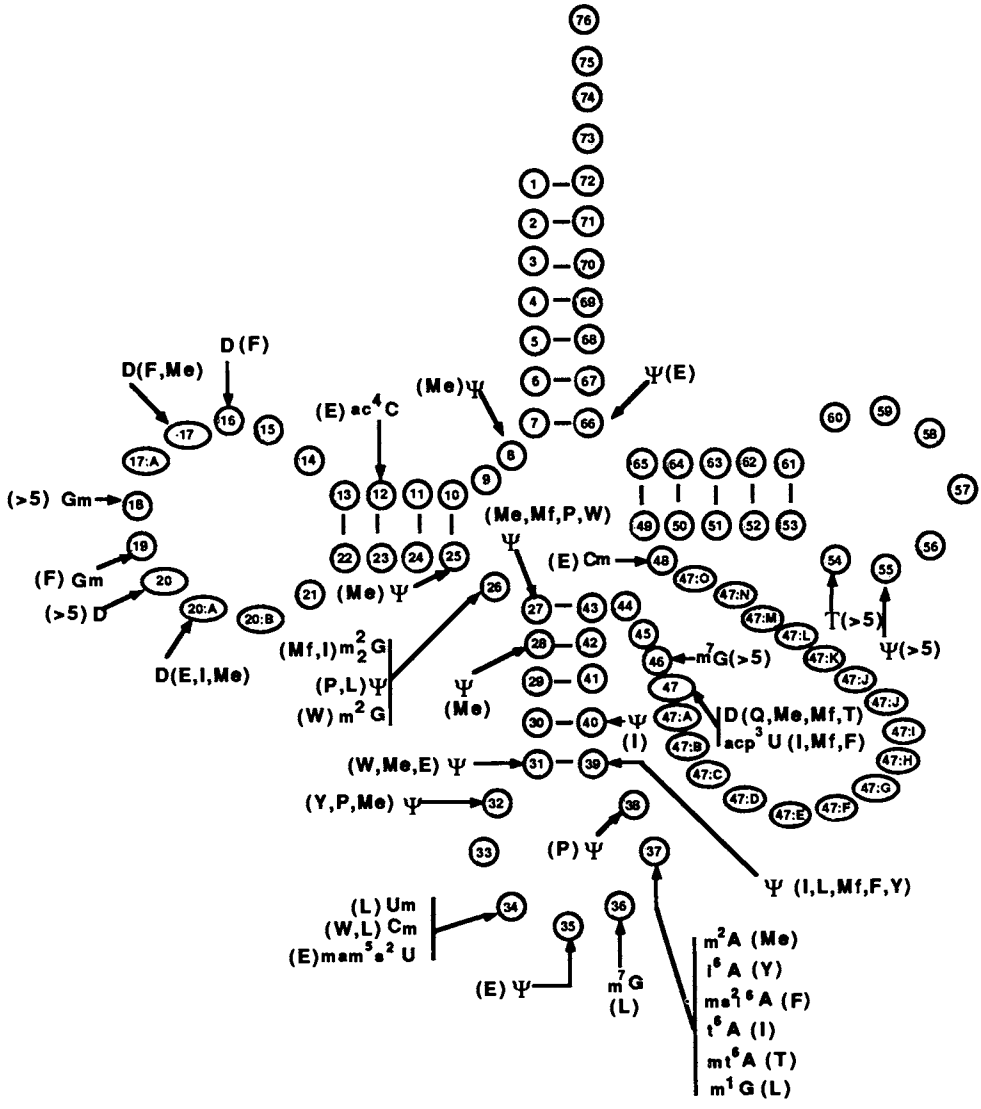


FIG. 3. Modified nucleosides found in cp tRNAs. The structures of the modified nucleosides as well as the abbreviations can be found in Sprinzl *et al.*, 1989. The cp tRNAs containing these modified nucleosides are indicated in parentheses by the one-letter code for amino acids. Me, methionine elongator tRNA; Mf, methionine initiator tRNA. >5 indicates that more than five tRNAs contain the modified nucleoside at this position. [Data compiled from Sprinz *et al.*, 1989.]

rarely modified, and the role of m^7G in the anticodon is unknown. Previously, m^7G had only been found at position 46, regardless of the origin of the tRNA.

5. The nucleotide adjacent to the anticodon at position 37 is frequently modified in cp tRNAs as in other tRNAs. However, it can be pointed out that ms^2i^6A has been found in eubacterial and organellar tRNAs but is usually not present in cytosolic tRNAs of eukaryotes.

The presence of modified nucleosides in cp tRNAs raises questions concerning the biosynthesis of minor nucleosides in cps. The complete sequence of three cp genomes is now determined (Shinozaki *et al.*, 1989; Hiratsuka *et al.*, 1989; Ohyama *et al.*, 1986), and no gene coding for a tRNA-modifying enzyme has been found yet. Because at least 40 open reading frames, or unidentified reading frames, are present in these genomes, the possibility exists that some cp-specific tRNA modifications (m^7G at position 36, ms^2i^6A at position 37, and acp^3U at position 47) are catalyzed by enzymes that are coded for by the cp genome. However, most of the tRNA-modifying enzymes probably are coded for by nuclear genes, synthesized in the cytoplasm, and transported into the cps, as, for instance, in the case of tRNA-modifying enzymes catalyzing the formation of modified nucleosides in yeast mitochondrial tRNAs (for a review, see Björk and Kohli, 1989).

XI. PECULIAR FEATURES OF cp $tRNA^{Ile_1}$

In plant cps, two $tRNAs^{Ile}$ are found: the major $tRNA^{Ile}$ (or $tRNA^{Ile_2}$) recognizes the codons AUU and AUC, while the minor $tRNA^{Ile}$ (or $tRNA^{Ile_1}$) probably recognizes the codon AUA. Surprisingly, the gene coding for spinach cp $tRNA^{Ile_1}$ was found to contain a CAU anticodon, which is normally a methionine anticodon (Kashdan and Dudock, 1982). This CAU anticodon was also found in the corresponding cp tRNA gene in other higher plants and in liverwort. In the case of spinach, the C residue is modified posttranscriptionally into C*, and Francis and Dudock (1982) have demonstrated that this $tRNA^{Ile}$(C*AU) is aminoacylated with isoleucine but not with methionine. Similar observations have been made in the case of bacteriophage T4 (Guthrie and McClain, 1973, 1979) and *E. coli* $tRNA^{Ile}$. In *E. coli*, the anticodon of the minor $tRNA^{Ile}$ gene is CAU, but this $tRNA^{Ile}$ has a lysidine (L) at the first position of its anticodon, which allows it to recognize the AUA codon (Muramatsu *et al.*,

1988a). Lysidine is a novel type of modified cytidine with a lysine moiety and with a positive charge. Probably because of this unique feature (L at the first position of the anticodon), it recognizes A but not G at the third position of the codon. In addition, it has been shown that the replacement of L by C changes both the aminoacylation and the codon recognition properties of this tRNA (Muramatsu *et al.*, 1988b). In the case of plant cp $tRNA^{Ile_1}$, it can be postulated that this tRNA probably recognizes the isoleucine codon AUA due to the modification of a C at the first position of the anticodon (possibly also into a L residue). However, it should be mentioned that spinach cp $tRNA^{Ile_1}$ contains an extra unpaired nucleotide within the double-stranded anticodon stem and that this nucleotide is m_2^2G, which is characteristically found in eukaryotic but not in prokaryotic tRNAs. Furthermore, this tRNA shows little similarity to other isoleucine tRNAs and no marked similarity to either prokaryotic, eukaryotic, or cp isoleucine tRNAs. This tRNA is as similar to the methionine elongator tRNA of *E. coli* or spinach cps as it is to other isoleucine tRNAs.

XII. A cp $tRNA^{Glu}$ IS INVOLVED IN CHLOROPHYLL BIOSYNTHESIS

It has been demonstrated in barley (Schön *et al.*, 1986) that a cp $tRNA^{Glu}$(UUC) participates in the metabolic conversion of glutamic acid into DALA, which is a precursor of chlorophyll (the tetrapyrrole ring of chlorophyll is synthesized from eight DALA molecules). The glutamate acceptor activities can be fractionated into three fractions on high-pressure liquid chromatography columns, but only the first fraction can support DALA formation in a reconstitution assay.

The formation of DALA from glutamate is specific for plant cps and blue-green algae (cyanobacteria), whereas in animals, yeast, and some bacteria, DALA synthesis occurs via condensation of glycine and succinyl-CoA. The barley cp $tRNA^{Glu}$, which participates in DALA biosynthesis, has a high similarity to the cp $tRNA^{Glu}$ from other plants, and, furthermore, the requirement for a tRNA was also demonstrated for DALA synthesis in *C. reinhardtii* (Huang *et al.*, 1984), *E. gracilis* (Weinstein *et al.*, 1986), and *C. vulgaris* (Weinstein and Beale, 1985) as well as in *Methanobacterium thermoautrophicum* (Friedmann and Thauer, 1986), suggesting that tRNA-dependent synthesis of DALA occurs in all plant cps and also in other organisms. The cp $tRNA^{Glu}$ has a 5-methylamino-

methyl-2-thiouridine (mam^5s^2U) at the first position of the anticodon, and this modified nucleoside has only been found up to now in tRNAs of prokaryotic organisms.

XIII. PROTEIN SYNTHESIS IN ORGANELLES REQUIRES MISAMINOACYLATION OF $tRNA^{Gln}$

Three glutamate-accepting tRNA species could be identified in barley cps (Schön *et al.*, 1988). The first $tRNA^{Glu}$ is involved in the biosynthesis of DALA (see above) and probably in the incorporation of glutamate into proteins. The two other glutamate-accepting tRNA species differ only by a ribose methylation of G_{18} and both have a U*UG anticodon (specific for glutamine). These mischarged Glu-$tRNA^{Gln}$ species can be converted into Gln-$tRNA^{Gln}$ by a crude cp extract. This reaction requires a specific amidotransferase and glutamine or asparagine as amide donors.

Chloroplasts, cyanobacteria, and plant and animal mitochondria do not have any detectable Gln-tRNA synthetase activity, and the glutamine residues required for protein biosynthesis are provided by conversion of glutamate attached to an "incorrectly" charged $tRNA^{Gln}$. This phenomenon of mischarging catalyzed by Glu-tRNA synthetase is widespread (Schön *et al.*, 1988). It has been especially described in several species of gram-positive bacteria (Wilcox and Nirenberg, 1968) and therefore appears to be conserved during evolution. However, the direct aminoacylation of a $tRNA^{Gln}$ by glutamine occurs in the cytoplasm of eukaryotic cells and in gram-negative bacteria. These observations raise questions about the origin of organelles and the evolution of the mechanisms involved in protein biosynthesis.

REFERENCES

Barkan, A. (1988). Proteins encoded by a complex chloroplast transcription unit are each translated from multiple mono- and polycistronic mRNAs. *EMBO J.* **7**, 2637–2644.

Björk, G. R., and Kohli, J. (1990). Synthesis and function of modified nucleosides in tRNA. *In* "Chromatography and Modification of Nucleosides" (C. N. Gehrke and K. C. Kuo, eds.), pp. 813–867. Elsevier, Amsterdam.

Burkard, G., Vaultier, J. P., and Weil, J. H. (1972). Differences in the level of plastid-specific tRNAs in chloroplasts and etioplasts of *Phaseolus vulgaris*. *Phytochemistry* **11**, 1351–1353.

Deng, X. W., Stern, D. B., Tonkyn, J. C., and Gruissem, W. (1987). Plastid run-on transcription: application to determine the transcriptional regulation of spinach plastid genes. *J. Biol. Chem.* **262,** 9641–9648.

El-Gewely, M. R., Helling, R. B., and Dibbits, J. G. T. (1984). Sequence and evolution of the regions between the *rrn* operons in the chloroplast genome of *Euglena gracilis bacillaris*. *Mol. Gen. Genet.* **194,** 432–443.

Evrard, J. L., Kuntz, M., Straus, N. A., and Weil, J. H. (1988). A class-I intron in a cyanelle tRNA gene from *Cyanophora paradoxa:* Phylogenetic relationship between cyanelles and plant chloroplasts. *Gene* **71,** 115–122.

Evrard, J. L., Kuntz, M., and Weil, J. H. (1990). The nucleotide sequence of five ribosomal protein genes from the cyanelles of *Cyanophora paradoxa:* Implications concerning the phylogenetic relationship between cyanelles and chloroplasts. *J. Mol. Evol.* **30,** 16–25.

Francis, M., and Dudock, B. (1982). Nucleotide sequence of a spinach chloroplast isoleucine tRNA. *J. Biol. Chem.* **257,** 11195–11198.

Friedmann, H. C., and Thauer, R. K. (1986). Ribonuclease-sensitive δ-aminolevulinic acid formation from glutamate in cell extract of *Methanobacterium thermoautrophicum*. *FEBS Lett.* **207,** 84–88.

Goins, D. J., Reynolds, R. J., Schiff, J. A., and Barnett, W. E. (1973). A cytoplasmic regulatory mutant of *Euglena:* Constitutivity for the light-inducible chloroplast transfer RNAs. *Proc. Natl. Acad. Sci. USA* **70,** 1749–1752.

Greenberg, B. M., and Hallick, R. B. (1986). Accurate transcription and processing of 19 *Euglena* chloroplast tRNAs in a *Euglena* soluble extract. *Plant Mol. Biol.* **6,** 89–100.

Greenberg, B. M., Gruissem, W., and Hallick, R. B. (1984a). Accurate processing and pseudouridylation of chloroplast transfer RNA in a chloroplast transcription system. *Plant Mol. Biol.* **3,** 97–109.

Greenberg, B. M., Narita, J. O., DeLuca-Flaherty, C., Gruissem, W., Rushlow, K. A., and Hallick, R. B. (1984b). Evidence for two RNA polymerase activities in *Euglena gracilis* chloroplasts. *J. Biol. Chem.* **259,** 14880–14887.

Gruissem, W., and Zurawski, G. (1985a). Analysis of promoter regions for the spinach chloroplast *rbcL, atpB* and *psbA* genes. *EMBO J.* **4,** 3375–3383.

Gruissem, W., and Zurawski, G. (1985b). Identification and mutational analysis of the promoter for a spinach chloroplast transfer RNA gene. *EMBO J.* **4,** 1637–1644.

Gruissem, W., Prescott, D. M., Greenberg, B. M., and Hallick, R. B. (1982). Transcription of *E. coli* and *Euglena* chloroplast tRNA gene clusters and processing of polycistronic transcripts in a HeLa cell-free system. *Cell* **30,** 81–92.

Gruissem, W., Greenberg, B. M., Zurawski, G., Prescott, D. M., and Hallick, R. B. (1983). Biosynthesis of chloroplast transfer RNA in a spinach chloroplast transcription system. *Cell* **35,** 815–828.

Gruissem, W., Elsner-Menzel, C., Latshaw, S., Narita, J. O., Schaffer, M. A., and Zurawski, G. (1986). A subpopulation of spinach chloroplast tRNA genes does not require upstream promoter elements for transcription. *Nucleic Acids Res.* **14,** 7541–7556.

Gruissem, W., Barkan, A., Deng, X. W., and Stern, D. (1988). Transcriptional and post-transcriptional control of plastid mRNA levels in higher plants. *Trends Genet.* **4,** 758–763.

Guthrie, C., and McClain, W. H. (1973). Conditionally lethal mutants of bacteriophage T4 defective in production of a transfer RNA. *J. Mol. Biol.* **81,** 137–155.

Guthrie, C., and McClain, W. H. (1979). Rare transfer ribonucleic acid essential for phage growth. Nucleotide sequence comparison of normal and mutant T4 isoleucine-accepting transfer ribonucleic acid. *Biochemistry* **18,** 3786–3795.

Hallick, R. B., and Buetow, D. E. (1989). Chloroplast DNA. *In* "The Biology of *Euglena*," (D. E. Buetow, ed.) Vol. IV, pp. 351–414. Academic Press, San Diego.

Heyraud, F., Serror, P., Kuntz, M., Steinmetz, A., and Heizmann, P. (1987). Physical map and gene localization on sunflower (*Helianthus annuus*) chloroplast DNA: Evidence for an inversion of a 23.5-kbp segment in the large single copy region. *Plant Mol. Biol.* **9,** 485–496.

Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C. R., Meng, B. Y., Li, Y. Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., and Sugiura, M. (1989). The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.* **217,** 185–194.

Howe, C. J. (1985). The endpoints of an inversion in wheat chloroplast DNA are associated with short repeated sequences containing homology to att-lambda. *Curr. Genet.* **10,** 139–145.

Huang, D. D., Wang, W. Y., Gough, S. P., and Kannangara, L. G. (1984). δ-aminolevulinic acid-synthesizing enzymes need an RNA moiety for activity. *Science* **225,** 1482–1484.

Hughes, J. E., Neuhaus, H., and Link, G. (1987). Transcript levels of two adjacent chloroplast genes during mustard (*Sinapis alba* L.) seedling development are under differential temporal and light control. *Plant Mol. Biol.* **9,** 355–363.

Janssen, I., Mucke, H., Löffelhardt, W., and Bohnert, H. J. (1987). The central part of the cyanelle rDNA unit of *Cyanophora paradoxa:* Sequence comparison with chloroplasts and cyanobacteria. *Plant Mol. Biol.* **9,** 479–484.

Jayabaskaran, C., Kuntz, M., Guillemaut, P., and Weil, J. H. (1990). Variations in the levels of chloroplast tRNAs and aminoacyl-tRNA synthetases in senescing leaves of *Phaseolus vulgaris*. *Plant Physiol.*, **92,** 136–140.

Kashdan, M. A., and Dudock, B. S. (1982). The gene for a spinach chloroplast isoleucine tRNA has a methionine anticodon. *J. Biol. Chem.* **257,** 11191–11194.

Kato, A., Takaiwa, F., Shinozaki, K., and Sugiura, M. (1985). Location and nucleotide sequence of the genes for tobacco chloroplast $tRNA^{Arg}$(ACG) and $tRNA^{Leu}$(UAG). *Curr. Genet.* **9,** 405–409.

Kuntz, M., Evrard, J. L., and Weil, J. H. (1988). Nucleotide sequence of the $tRNA^{Ser}$GGA and $tRNA^{Gly}$GCC genes from cyanelles of *Cyanophora paradoxa*. *Nucleic Acids Res.* **16,** 8733.

Manhart, J. R., and Palmer, J. (1990). The gain of two chloroplast tRNA introns marks the green algal ancestors of land plant. *Nature (London)* **345,** 268–270.

Marion-Poll, A., Hibbert, C. S., Radebaugh, C. A., and Hallick, R. B. (1988). Processing of mono-, di- and tricistronic transfer RNAs precursors in a spinach or pea chloroplast soluble extract. *Plant Mol. Biol.* **11,** 45–56.

Markowicz, Y., Mache, R., and Loiseaux-de Goër, S. (1988). Sequence of the plastid rDNA spacer of the brown alga *Pylaiella littoralis* (L.) Kjellm. Evolutionary significance. *Plant Mol. Biol.* **10,** 465–469.

Merrick, W. C., and Dure, L. S. (1972). The developmental biochemistry of cotton seed embryogenesis and germination. IV. Levels of cytoplasmic and chloroplastic transfer ribonucleic acid species. *J. Biol. Chem.* **247,** 7988–7999.

Michel, F., and Dujon, B. (1983). Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast-, and nuclear-encoded members. *EMBO J.* **2,** 33–38.

Muramatsu, T., Nishihawa, K., Nemoto, F., Kuchino, Y., Nishinura, S., Miyazawa, T.,

and Yokoyama, S. (1988a). Codon and amino acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. *Nature* **336,** 179–181.

Muramatsu, T., Yokoyama, S., Horie, N., Matsuda, A., Veda, T., Yamaizumi, Z., Kuchino, Y., Nishinura, S., and Miyazawa, T. (1988b). A novel lysine-substituted nucleoside in the first position of the anticodon of minor isoleucine tRNA from *Escherichia coli*. *J. Biol. Chem.* **263,** 9261–9267.

Neuhaus, H., and Link, G. (1987). The chloroplast $tRNA^{Lys}$(UUU) gene from mustard (*Sinapis alba*) contains a class-II intron potentially coding for a maturase-related polypeptide. *Curr. Genet.* **11,** 251–257.

Nickelsen, J., and Link, G. (1989). Interaction of a 3′ RNA region of the mustard trnK gene with chloroplast proteins. *Nucl. Acids Res.* **17,** 9637–9648.

Ohme, M., Kamogashira, T., Shinozaki, K., and Sugiura, M. (1985). Structure and cotranscription of tobacco chloroplast genes for $tRNA^{Glu}$(UCC), $tRNA^{Tyr}$(GUA) and $tRNA^{Asp}$(GUC). *Nucleic Acids Res.* **13,** 1045–1056.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., and Ozeki, H. (1986). Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* **322,** 572–574.

Oliver, R. P., and Poulsen, C. (1984). Structure of a heavily transcribed region of barley chloroplast DNA. Transfer RNA genes for serine (UGA), glycine (GCC,UCC), formylmethionine and threonine (GGU). *Carlberg Res. Commun.* **49,** 647–673.

O'Neill, G. P., Schön, A., Chow, H., Chen, M. W., Kim, Y. C., and Söll, D. (1990). Sequence of $tRNA^{Glu}$ and its genes from the chloroplast genome of *Chlamydomonas reinhardtii*. *Nucl. Acids Res.* **18,** 5893.

Orozco, E. M., Rushlow, K. E., Dodd, J. R., and Hallick, R. B. (1980). *Euglena gracilis* chloroplast ribosomal RNA transcription units: II. Nucleotide sequence homology between the 16 S–23 S ribosomal RNA spacer and the 16 S ribosomal RNA leader regions. *J. Biol. Chem.* **255,** 10997–11003.

Pfitzinger, H., Guillemaut, P., Weil, J. H., and Pillay, D. T. N. (1987). Adjustment of the tRNA population to the codon usage in chloroplasts. *Nucleic Acids Res.* **15,** 1377–1386.

Pfitzinger, H., Maréchal-Drouard, L., Pillay, D. T. N., Weil, J. H., and Guillemaut, P. (1990). Variations during leaf development of the relative amounts of two bean (*Phaseolus vulgaris*) chloroplast $tRNA^{Phe}$ which differ in their minor nucleotide content. *Plant Mol. Biol.* **14,** 969–975.

Pillay, D. T. N., Singh, G. P., and Guillemaut, P. (1984). Localization of four genes for three $tRNA^{Leu}$s on the physical map of the soybean (*Glycine max* L.) chloroplast genome. *Gene* **29,** 351–354.

Quigley, F., and Weil, J. H. (1985). Organization and sequence of the five tRNA genes and of an unidentified reading frame in the wheat chloroplast genome: Evidence for gene rearrangement during the evolution of chloroplast genomes. *Curr. Genet.* **9,** 495–503.

Samuelson, T., Elias, P., Lusting, F., Axberg, T., Fölsch, G., Akeson, B., and Lagerkvist, U. (1980). Aberrations of the classic codon reading scheme during protein synthesis in vitro. *J. Biol. Chem.* **255,** 4583–4588.

Schneider, M., and Rochaix, J. D. (1986). Sequence organization of the chloroplast ribosomal spacer of *Chlamydomonas reinhardtii:* Uninterrupted $tRNA^{Ile}$ and $tRNA^{Ala}$ genes and extensive secondary structure. *Plant Mol. Biol.* **6,** 265–270.

Schön, A., Krupp, G., Gough, S., Berry-Lowe, S., Kannangara, C. G., and Söll, D. (1986). The RNA required in the first step of chlorophyll biosynthesis is a chloroplast glutamate tRNA. *Nature* **322,** 281–284.

Schön, A., Kannangara, C. G., Gough, S., and Söll, D. (1988). Protein biosynthesis in organelles requires misaminoacylation of tRNA. *Nature* **331,** 187–190.

Schwartz, Z., Jolly, S. O., Steinmetz, A. A., and Bogorad, L. (1981). Overlapping divergent genes in the maize chloroplast chromosome and in vitro transcription of the gene for $tRNA^{His}$. *Proc. Natl. Acad. Sci. USA* **78,** 3423–3427.

Sharp, P. A. (1987). Splicing of messenger RNA precursors. *Science* **235,** 766–771.

Shimada, H., and Sugiura, M. (1989). Pseudogenes and short repeated sequences in the rice chloroplast genome. *Curr. Genet.* **16,** 293–301.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimida, H., and Sugiura, M. (1986). The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. *EMBO J.* **5,** 2043–2049.

Sorensen, M. A., Kurland, C. G., and Pedersen, S. (1989). Codon usage determines translation rate in *Escherichia coli*. *J. Mol. Biol.* **207,** 365–377.

Sprinzl, M., Hartmann, T., Weber, J., Blank, J., and Zeidler, R. (1989). Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* (Suppl.) **17,** 1–172.

Steinmetz, A., and Weil, J. H. (1989). Protein synthesis in chloroplasts. *In* "The Biochemistry of Plants," (A. Marcus, ed.) Vol. 15, pp. 193–227. Academic Press, New York.

Sugiura, M., and Wakasugi, T. (1989). Compilation and comparison of transfer RNA genes from tobacco chloroplasts. *Crit. Rev. Plant Sci.* **8,** 89–101.

Tohdoh, N., Shinozaki, K., and Sugiura, M. (1981). Sequence of a putative promotor region for the rRNA genes of tobacco chloroplast DNA. *Nucleic Acids Res.* **9,** 5399–5406.

Turner, S., Burger-Wiersma, T., Giovannoni, S. J., Mur, L. R., and Pace, N. R. (1989). The relationship of a prochlorophyte *Prochlorothrix hollandica* to green chloroplasts. *Nature* **337,** 380–382.

Wang, M. J., Davies, N. W., and Gegenheimer, P. (1988). Novel mechanisms for maturation of chloroplast transfer RNA precursor. *EMBO J.* **7,** 1567–1574.

Weinstein, J. D., and Beale, S. I. (1985). RNA is required for enzymatic conversion of glutamate to δ-aminolevulinate by extracts of *Chlorella vulgaris*. *Arch. Biochem. Biophys.* **239,** 87–93.

Weinstein, J. D., Mayer, S. M., and Beale, S. I. (1986). Stimulation of δ-aminolevulinic acid formation in algae by heterologous RNA. *Plant Physiol.* **82,** 1096–1101.

Wilcox, M., and Nirenberg, M. (1968). Transfer RNA as a cofactor coupling amino acid synthesis with that of protein. *Proc. Natl. Acad. Sci. USA* **61,** 229–236.

Woodbury, N. W., Roberts, L. L., Palmer, J. D., and Thompson, W. F. (1988). A transcription map of the pea chloroplast genome. *Curr. Genet.* **14,** 75–89.

Yamada, T., and Shimaji, M. (1987). Splitting of the ribosomal RNA operon on chloroplast DNA from *Chlorella ellipsoidea*. *Mol. Gen. Genet.* **208,** 377–383.

Yamaguchi-Shinozaki, K., Shinozaki, K., and Sugiura, M. (1987). Processing of precursor tRNAs in a chloroplast lysate. *FEBS Lett.* **215,** 132–136.

Zhang, D., and Spreitzer, R. J. (1989). Nucleotide sequence of the *Chlamydomonas reinhardtii* chloroplast genes for tryptophan and glycine transfer RNAs. *Nucleic Acids Res.* **17,** 8873.

CHAPTER 8

Ribosomal Proteins, Ribosomes, and Translation in Plastids

Alap R. Subramanian
Dietmar Stahl
Anastasia Prombona

Max Planck Institute for Molecular Genetics
D-100 Berlin 33, Federal Republic of Germany

I. INTRODUCTION

Plant cells are distinguished by the presence of three separate classes of ribosomes, signifying the three distinct translational systems they maintain. These three classes are designated by the sedimentation values and the subcellular compartment in which they function, i.e., cytoplasmic

80S, mitochondrial 75S, and plastid 70S ribosomes. Plant cytoplasmic 80S ribosomes (reviewed last by Davies and Larkin, 1980) are structurally and functionally similar to the cytoplasmic ribosomes of animal cells. On the other hand, plant mitochondrial ribosomes possess a distinctly different structural motif from the ribosomes of animal mitochondria (review in the next volume of this series; Grohmann *et al.*, 1990). Plastid ribosomes, which are an integral part of the differentiable unique organelle that only photosynthetic eukaryotes possess, naturally have no counterparts in the animal or fungal systems.

In this chapter, we shall summarize recent work on the plastid translation system in higher plants. Emphasis will be on the ribosomal proteins (r-proteins), their genes, and aspects of gene expression and evolution. Particular attention will be given to correlating the protein nomenclature from different laboratories so that a common terminology will develop. The r-protein genes located in the plastid DNA of three angiosperms and a lower nonflowering plant are now known. The clustering pattern of these prokaryotic type genes and their operon structure, with similarities to and differences from the corresponding operons in eubacteria, will be discussed. Over 40 chloroplast r-protein genes are assumed to be in the nuclear genome. Thirteen of these have been cloned and characterized at the cDNA level. If the endosymbiosis hypothesis (Bogorad, 1975; Chapter 11A) is true, these nuclear genes are presumably of eubacterial origin. They are particularly interesting because they would have been incorporated into the eukaryotic chromosomal structure at a rather late stage in nuclear evolution. The features of the known cDNA clones and the diverse chloroplast-targeting sequences they carry are discussed. This chapter is restricted mainly to the work on higher plants and, for illustration, concentrated on the results from the authors' laboratory. For work on the algal chloroplast system of *Chlamydomonas reinhardtii*, the reader is referred to an overview article by Bogorad *et al.* (1977) and to recent papers from the laboratory of Gillham, Boynton, and Harris (see Chapter 3A).

II. PLASTID TRANSLATION SYSTEM

In their stroma, plastids contain a complete system for transcription and translation. The translational system includes a full set of >30 tRNAs (discussed in Chapter 7A), ribosomes, and translational factors. Differ-

ent maturation–modification enzymes, which act on tRNAs, rRNA, mRNA, and r proteins, are also present in the stroma. Genes for the majority of these components are located in the nucleus (Table I).

Plastid ribosomes exclusively synthesize those proteins that are encoded in the plastid genome. The list includes about 50 proteins that have been identified as occurring in chloroplasts and about 40 open reading frames whose protein products have not yet been identified (Umesono and Ozeki, 1987). So far, no evidence indicates that any of the plastid-coded proteins has a site of action in the cell outside of the plastid; however, this possibility has not been experimentally ruled out.

The relative contribution of plastid-synthesized proteins to the total cellular protein mass varies with the organ and the tissue. It is typically considerably greater than that of the mitochondrial ribosomes (a possible exception could be the pollen tube rich in mitochondria). In leaf, about 50% of the protein mass is generally quoted (Ellis, 1984) as the product of chloroplast ribosomes. Over 25% of the total leaf ribosomes (e.g., calculated from the data of Feierabend and Schrader-Reichhardt [1976] for young rye plants) are chloroplast ribosomes.

Because plastid ribosomes are responsible for the synthesis of a huge amount of biomass, any increase in their synthetic efficiency could have a considerable impact. In principle, genetic engineering of r-proteins would allow increase in such efficiency. For example, work on *Escherchia coli* ribosomes has demonstrated that mutations in specific r-proteins can

TABLE I
Gene Compartmentation of the Plastid Translation System

Component	Gene Compartment	
	Plastid	Nucleus
rRNA	+	–
r-Proteins	+	++
Initiation factors	+	++
Elongation factors	–	+
Termination factors	–	+
tRNA	+	–
rRNA processing and base-modifying enzymes	–	+
r-Protein-modifying enzymes	–	+
tRNA base-modifying and terminal CCA-adding enzymes	–	+
Aminoacyl-tRNA synthetases	–	+
mRNA maturation enzymes	–	+

increase or decrease the ribosome travel rate on mRNA. Genetic engineering of plants along these lines is, at present, only a distant dream.

III. PLASTID r-PROTEINS

Plastid ribosomes belong to the eubacterial 70S type. The rRNA of plastid ribosomes is discussed in Chapter 6A. The protein composition of higher plant plastid ribosomes has so far not been reviewed; therefore, in this section we have collected and critically evaluated the data from the literature on this subject.

Two-dimensional gel electrophoresis has been used since 1972 for separating chloroplast r-proteins to determine their total number. A summary of the results from these experiments is given in Table II. These data (covering six higher plants and two algae) indicate that the

TABLE II
Number of r-Proteins in Chloroplast Ribosomes as Estimated by Two-Dimensional Gel Electrophoresis

Organism	r-Proteins			Reference
	Ribosome	Small Subunit	Large Subunit	
Chlamydomonas reinhardtii	48	22	26	Hanson *et al.* (1974)
Chlamydomonas reinhardtii	59	25	34	Brügger and Boschetti (1975)
Chlamydomonas reinhardtii	64	31	33	Schmidt *et al.* (1983)
Euglena gracilis	56–60	22–24	30–34	Freyssinet (1977)
Triticum aestium	~75	n.d.	n.d.	Jones *et al.* (1972)
Spinacia oleracea	~40	n.d.	n.d.	Gualerzi *et al.* (1974)
Spinacia oleracea	57	24	33	Dorne *et al.* (1984)
Spinacia oleracea	59–61	23–25	36	Posno *et al.* (1984b)
Spinacia oleracea	58–65	23–26	35–39	Subramanian (1985; this review)
Nicotiana tabacum	55–58	22–23	33–35	Capel and Bourque (1982)
Pisum sativum	56	24	32	Eneas-Filho *et al.* (1981)
Pisum sativum	59	25	34	Udalova *et al.* (1984)
Vicia faba	54	22	32	Hachtel (1985)
Zea mays	~61	~25	~36	Subramanian (unpublished)
Average[a]	59	24	35	
Escherichia coli	54	21	33	Wittmann (1982)

[a] Data before 1980 excluded.
n.d., not determined.

same number of r-proteins is probably present in all chloroplast ribosomes, independently of species and taxa. The data also suggest a slightly larger number of protein components (five to six more) in the chloroplast ribosome than those present in the eubacterial ribosome of *E. coli*. These "extra" proteins are called Plastid-specific r-proteins (PSrps) and are discussed later.

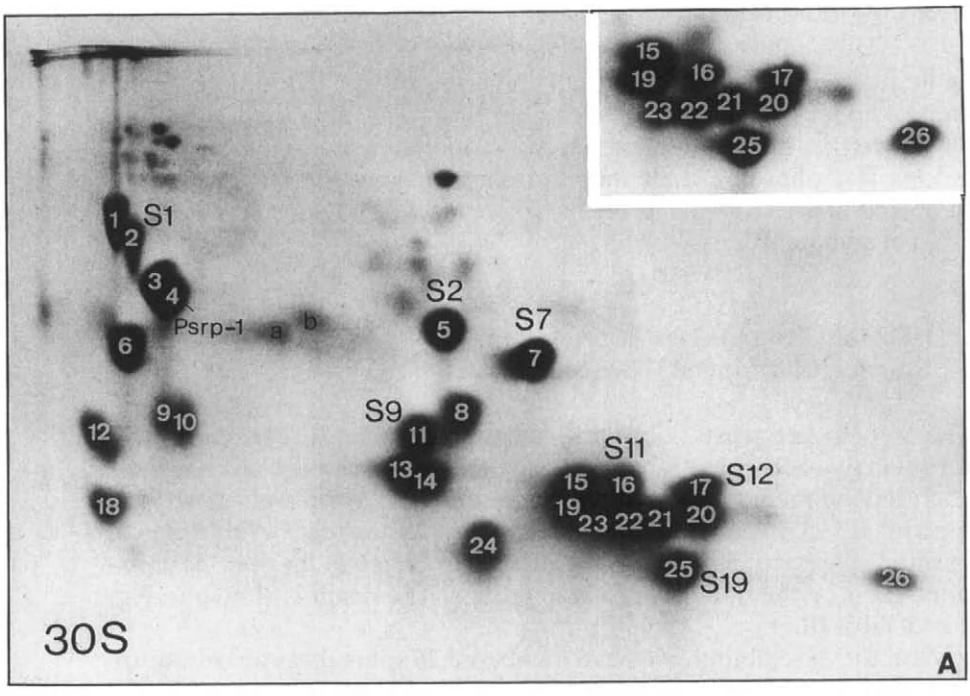
A. r-Protein Composition–Correlation: Spinach Chloroplast Ribosome

The best-studied plastid ribosome is that of the spinach chloroplast. Unfortunately, each of the laboratories working on this problem has used an independent nomenclature (Dorne *et al.*, 1984; Posno *et al.*, 1984; Subramanian, 1985); a common nomenclature is now needed. With this goal in mind, we have made a careful attempt to correlate the spot designations used by the different research groups. The result is shown in Fig. 1 and Table III.

With the 30S subunit, we have numbered 26 spots that apparently are present in near-stoichiometric amounts. Several unnumbered, high M_r spots of relatively low stoichiometry (e.g., spot a, b in Fig. 1) are also present. Based on our experience with *E. coli* r-proteins (Subramanian, 1974), some of the latter are probably translational factors and/or minor dimer forms of certain r-proteins. With the 50S subunit, there are 38 numbered spots. Again, several apparent minor spots have not been numbered for reasons mentioned above.

Table III gives the M_r values for the proteins of these spots and a correlation of the nomenclatures used for the same (or equivalent) spots in the literature. At the resolution in Fig. 1, some spots have not been previously separated and numbered (e.g., A: 30S-12, 30S-23, 30S-26; B: 50S-2, 50S-6, 50S-19, 50S-21). On the other hand, a few spots have been numbered by one or the other group, but we did not observe them (Table III).

In several cases, a spot can be identified as the chloroplast homolog of a eubacterial r-protein because of positive immunological cross-reaction (e.g., S1 [Hahn *et al.*, 1988], L1, L2, L3, L6, L12, L13, L17, S7, S9, S11, S12 and S19 [Bartsch *et al.*, 1982; Bartsch, 1985]). In other cases, identifications have been made from amino acid sequence data of the isolated protein (e.g., L2 [Kamp *et al.*, 1987], L12 [Bartsch *et al.*, 1982], L14, L16, and L22 [Zhou *et al.*, 1989], L21 [Smooker *et al.*, 1990], PSrp-1 [Johnson *et al.*, 1990], S2 and S7 [D. Stahl and A. R. Subramanian, unpublished]). In all, 19 spots (Table III) have been identified in this


15
19
16
17
23 22
21
20
25
26
1
S1
2
3
4
S2
Psrp-1
a
b
5
S7
6
7
S9
8
9
10
11
12
S11
13
14
15
16
17
S12
18
19
23 22 21
20
24
25
S19
26
30S
A

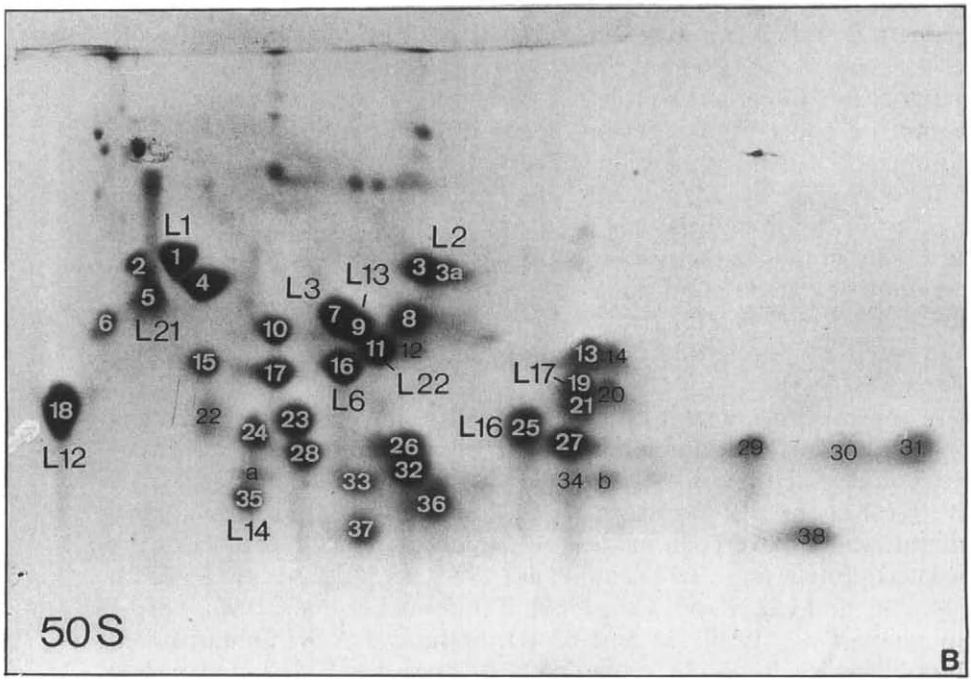

L1
L2
1
2
3
3a
4
L13
L3
5
7
9
8
6
L21
10
11
12
13
14
15
17
16
L22
L17
19
L6
21
20
18
22
24
23
L16
25
27
L12
26
28
29
30
31
a
32
33
34 b
35
36
L14
37
38
50S
B

FIG. 1. Ribosomal proteins of the small (A)(30S) and large (B)(50S) subunits of spinach chloroplast ribosome separated by two-dimensional gel electrophoresis. Correlation with the numbering systems used by other laboratories are given in Table III. The designation S or L and a number (e.g., L12) is given in those cases where the protein spot has been identified as corresponding to an *E. coli* r-protein by amino acid sequence and/or immunological cross-reaction data (see text for references). The inset (30S) shows a better resolution of the S11–S19 region.

TABLE III

Correlation of Spinach Chloroplast r-Protein Data from Different Laboratories

Spot Numbering[a]	M_r by SDS Gel after HPLC Purification[b]	Corresponding Spot Designation[a]	
Berlin		Grenoble	Amsterdam
	30S Subunit		
—	—	1 (LSU)	1 (LSU)
1[c] = S1	43,000	2	2
2[c] = S1	36,750	3	3
3	32,000	4	4
4[c] = PSrp-1	31,000	5	5
5[c] = S2	27,000	6	6
6	25,000	7	7
7[c] = S7	18,500	11	11
8	15,500	10	13
9	18,200	9	9
10	18,200	—	10
11[c] = S9	16,500	12	12
12	16,500	—	—
13	17,800	13	15
14	18,200	19	14
15	15,000	14	16
16[c] = S11	15,500	15	17
17[c] = S12	15,500	17	18
18	14,000	18	—
19	15,500	16	19
20	14,000	22	20
21	14,000	21	22
22	12,500	—	23
23	14,000	—	—
24	12,500	24	24
25[c] = S19	12,000	23	25
26	11,200	—	—
—		20 (SSU)	21 (SSU)
			—
a, b	29,200	—	8
	50S Subunit		
—		1	1
—		2	2
1[c] = L1	31,500	3	3
2	31,000	—	—

(continues)

TABLE III
Continued

Spot Numbering[a]	M_r by SDS Gel after HPLC Purification[b]	Corresponding Spot Designation[a]	
Berlin		Grenoble	Amsterdam
3[c] = L2	30,500	5	4
3a[c] = L2	30,500	4	5
4	30,000	8	6
5[c] = L21	30,000	7	7
6	25,300	—	—
7[c] = L3	25,300	9	8
8	24,500	12	11
9[c] = L13	24,500	11	9
10	24,500	10	10
11[c] = L22	22,000	13	13
12	21,500	6	12
13	16,500	22	17
14	20,000	—	17
15	21,500	16	15
16[c] = L6	21,500	15	14
17	20,000	14	16
18[c] = L12	17,500	20	—
19[c] = L17	16,500	—	—
20	16,500	—	22
21	16,000	—	—
22	16,500	21	19, 20
23	16,000	17	18
24	16,500	23	21
25[c] = L16	15,000	24	24
26	16,500	25	25
27	16,000	26	26
28	16,000	18	23
29	13,000	32	32
30	11,500	35	33
31	12,500	36	34
32	15,000	—	28
33	12,000	27	27
34	12,000	31	—
35 = L14	13,500	29	30
36	11,500	28	29
37	11,000	30	31
38	11,000	33	35
39		—	36
a		—	—
b		—	—
—		19	—
—		37	—

[a] Berlin numbering is as given in Fig. 1; Grenoble and Amsterdam numberings are as given in Dorne *et al.* (1984) and Posno *et al.* (1984), respectively.

[b] Determined in the authors' laboratory (unpublished).

[c] Identification of a spot as the homolog of an *E. coli* r protein is based on immunological cross-reaction and/or amino acid sequence data (see text).

LSU, large subunit; SSU, small subunit of Pubisco.

manner. Two spots that often appear in preparations of chloroplast ribosomes (e.g., Posno *et al.*, 1984) are shown to be the subunits of contaminating ribulose bisphosphate carboxylase (Table III).

B. Modifications of Plastid r-Proteins

One of the chloroplast r-proteins (L2) is methylated at the NH_2 terminus (Kamp *et al.*, 1987), and evidence for the phosphorylation of one or two chloroplast r-proteins has been obtained (Posno *et al.*, 1984; Guitton *et al.*, 1984). Protein modifications in the plastid ribosome apparently follow a different pattern from that in eubacterial ribosomes. For example, r-protein L2 is not methylated in *E. coli*. Another such example is r-protein L12, which occurs in the NH_2-terminal acetylated form in *E. coli* but is unmodified in the chloroplast (Bartsch *et al.*, 1982). Physiologically regulated r-protein phosphorylation is a phenomenon in eukaryotic 80S ribosomes, but it does not occur in *E. coli* or in other tested eubacteria; therefore, plastid r-protein phosphorylation apparently deserves more serious attention.

C. Homology Relationships of Chloroplast r-Proteins

The primary structures of over 25 chloroplast r-proteins are known, one of them (L12) from amino acid sequencing (Bartsch *et al.*, 1982) and the rest from nucleotide sequencing of the genes or the cDNA clones. In each of these cases (Table IV), the chloroplast protein is an apparent homolog of a corresponding eubacterial r-protein. The identities in amino acid sequence between homologous chloroplast and *E. coli* r-proteins vary over a surprisingly large range (Table IV). Some r-proteins are highly conserved (e.g., S12, L36), whereas others (e.g., L22, L23) are at a value barely sufficient to establish homology. The average percent of identity is 44%. Because many of the amino acid replacements are often conservative (e.g., serine–threonine, leucine–isoleucine–valine), the effective average similarity is significantly higher than this value.

The five nuclear-encoded r-proteins in this list (Table IV) have the average sequence identity of 42% to their *E. coli* homolog. Thus, the location of the r-protein gene, whether in the plastid or in the nucleus, has not apparently influenced the amino acid sequence divergence. The precursors of these five nuclear-coded chloroplast r-proteins carry long targeting sequences (discussed in a later section), which are presumably not of eubacterial origin.

TABLE IV
Homology Relationship between the r-Proteins of Plastids and *E. coli*

r-Protein	Protein Chain Length			Identity in Sequence (%)
	Plastid	*E. coli*	Difference	
S2	236	240	−4	39
S3	224	232	−8	38
S4	201	203	−2	39
S7	156	153	+3	43
S8	136	129	+7	38
S11	143	128	+15	48
S12	124	123	+1	66
S14	103	98	+5	39
S15	90	88	+2	34
S16	62	82	−20	32
S18	163	74	+89	32
S19	93	91	+2	42
L2	272	272	0	43
L9[a]	160	148	+12	30
L12[a]	133	120	+13	48
L13[a]	190	142	+48	54
L14	123	123	0	52
L16	136	136	0	53
L20	119	117	+2	41
L22	126	110	+16	27
L23	93	100	−7	25
L24[a]	149	103	+46	35
L33	66	54	+12	35
L35[a]	73	65	+8	41
L36	37	38	−1	62

[a] Encoded in the nucleus (Gantt, 1988; Giese and Subramanian, 1989; Phua *et al.*, 1989; Smooker *et al.*, 1990). All others are encoded in the plastid genome; data for maize (Subramanian *et al.*, 1990), except S15, S16, L2, L20, and L33, which are for rice (Hiratsuka *et al.*, 1989).

D. PSrps

Based on electrophoretic analysis, chloroplast ribosomes contain a significantly larger number of r-proteins than the 54 known to comprise a typical eubacterial ribosome (Wittmann, 1982). Although counting spots in a two-dimensional gel is not the same as counting proteins, further evidence indicates the necessity of this conclusion. Chloroplast r-proteins have been purified by high-performance liquid chromatogra-

phy in our laboratory, and many of the purified proteins have been sequenced. In two cases, the sequence data showed no homology to any of the known *E. coli* r-proteins. More definitive evidence in this direction has come from the isolation of cDNA clones that encode a chloroplast r-protein with no amino acid sequence homology to any of the known r-proteins (Johnson *et al.*, 1990; see also Zhou and Mache, 1989). We have named this protein PSrp-1. Two pea chloroplast cDNA clones reported earlier by Gantt (1988) also apparently belong to this category.

From an evolutionary point of view, PSrps will form an extremely interesting group. They are likely to be proteins that the presumed chloroplast progenitor did not possess and were acquired during the period when endosymbiosis was being established; therefore, they are proteins that will have no essential role in the translational process per se but may be important in some aspects of the cellular integration of plastid protein synthesis. In this context, we should recall that eukaryotic cytoplasmic ribosomes contain about 90 r-proteins, 50% more than in the eubacteria. The amino acid sequence of over half of them are not yet known. We have proposed the hypothesis that PSrps are homologs of certain cytoplasmic r-proteins, which were recruited by the plastid ribosome from the endosymbiont host (Johnson *et al.*, 1990).

IV. GENES FOR PLASTID r-PROTEINS

The most outstanding character of plastid r-protein genes *vis-à-vis* gene regulation is their distribution (in 2:1 ratio) between the nuclear chromosome and the plastid nucleoid. This entails the synthesis of about 40 r-proteins in one common compartment (the cytoplasm) and the equivalent level of 20 r-proteins in the hundreds of cellular sites (plastids) of the second compartment. However, unlike the similar situation with respect to the genes of other plastid assemblies, here a basis for understanding gene regulation may already exist, i.e., from the detailed work on *E. coli* r-protein biosynthesis. The bacterial ribosome assembly is a precisely ordered process. It begins with the binding of two r-proteins, called initiator proteins (S4 and S7 for the small subunit; L3 and L20/L24 for the large subunit), to each rRNA, followed by the site-specific binding of other r-proteins (primary binding proteins). A conformational change of the rRNA then occurs and is followed by the final binding of the full complement of r-proteins; therefore, it is important that all the nuclear- and plastid-coded r-proteins are present stoichiometrically and

that the synthesis is turned on and off coordinately. The mechanism of that regulation remains to be elucidated, but the current work on the cloning and characterization of the plastid r-protein genes is a step in that direction.

A. Genes Located in the Plastid Genome

The complete nucleotide sequences of the plastid genomes of two angiosperms are known (rice and tobacco) and they both encode the same 20 r-proteins (see Table IV and Chapters 5A and 15B). The third completely sequenced plastid genome is that of a nonflowering, nonvascular lower plant (*Marchantia polymorpha*). It also encodes 20 r-proteins but with an exchange of *rps*16 with *rpl*21 (Ohyama *et al.*, 1986). The only other plastid genome where all the r-protein genes have been physically mapped and nearly completely sequenced is that of maize. A map of the r-protein loci in this genome (Subramanian *et al.*, 1990) is shown in Fig. 2.

Ten of the r-protein genes in the plastid genome occur in a large cluster and there are two smaller clusters of three and two r-protein genes. The remaining five r-protein genes occur separately. The large cluster also includes the genes for a component of an *E. coli*-like RNA polymerase (α subunit) and an initiation factor (IF-1).

The distribution of the r-protein gene loci in maize (Fig. 2) is identical to that in rice (Hiratsuka *et al.*, 1989); the same probably also is true for rye (Prombona and Subramanian, 1989, and unpublished results). The r-protein gene distribution in the dicot group (as typified in tobacco) differs from this pattern on three points. (1) The *rps*15 gene in maize, rice, and rye is a double-copy gene located in the inverted repeat (IR) sequence, whereas in tobacco it is a single-copy gene located in the small single-copy region. (2) In cereals (rye, rice, wheat, and maize), *rps*2 is located closer to *rps*4 than to *rps*16 (Fig. 1). In tobacco, *rps*2 is closer to *rps*16 due to the inversion of a 28-kbp segment in the long single-copy region. (3) Two *rps*19 genes occur in the two IRs of maize and rice, but only one complete *rps*19 gene occurs in tobacco (in IR-II).

Because of the typical IR-containing organization of the plastid DNA, six r-protein genes in maize and four in tobacco are double-copy genes. *Marchantia*, which has a relatively short IR, contains no double-copy r-protein genes. The r-protein genes in those higher plants and algae, which lack the IR structure (e.g., certain legumes, gymnosperms, *Euglenophyta*) would all be single-copy genes, as is the case in *E. coli*, where all r-protein genes are single-copy genes. Plastids that contain IR struc-

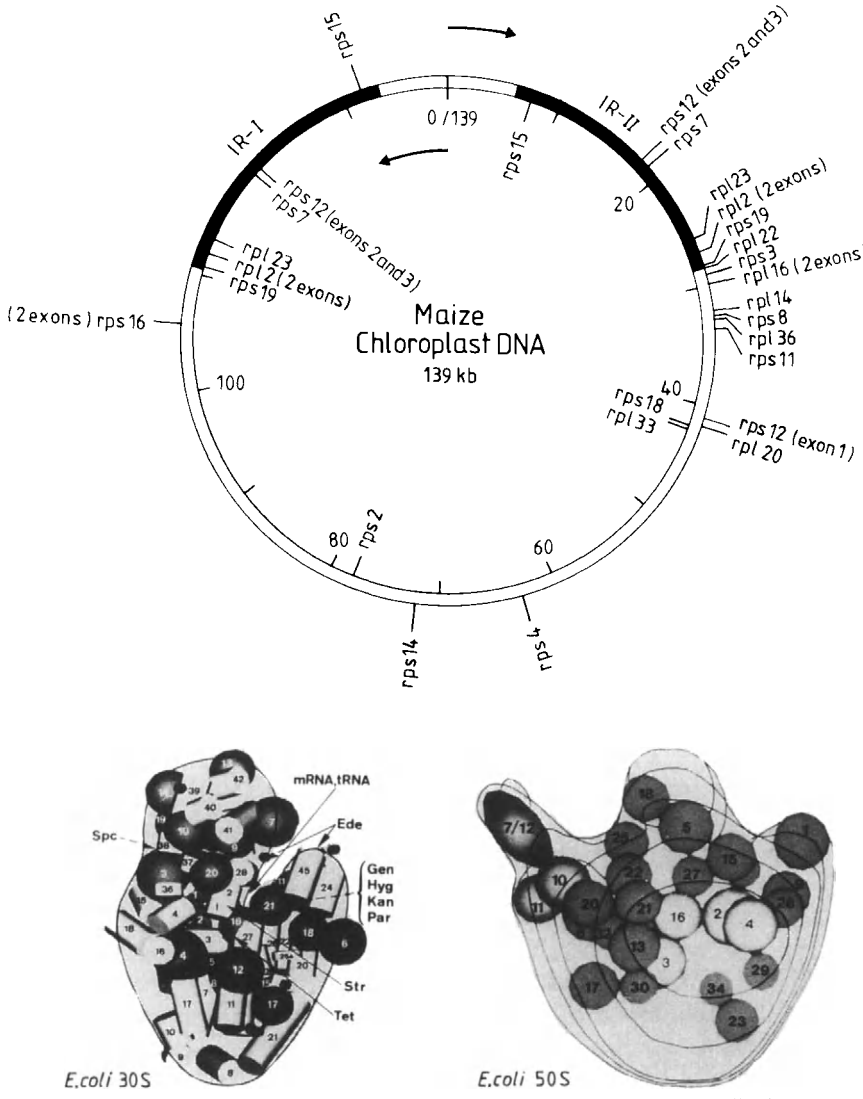


FIG. 2. Physical map of the maize plastid genome showing the experimentally determined loci of ribosomal protein genes (top). The current three-dimensional models of the two subunits of the *E. coli* ribosome (bottom). Arrows show the direction of transcription (genes written outside the circle are transcribed clockwise; those inside counterclockwise). Positions of many of the plastid-encoded proteins (e.g., S12, L2, L16) can be discerned in the three-dimensional models. [The 3-D models taken with permission from Wittmann, 1989.]

tures must presumably be equipped with some additional yet unknown mechanisms to equibalance the transcription–translation of their r-protein genes.

Four of the r-protein genes in the plant plastid genome are intron-containing, i.e, those for L2, L16, S12, and S16. Especially interesting is the case of the single intron in *rpl*2, which is precisely excised and lost in spinach and other members of *Caryophyllaceae* (Zurawski *et al.*, 1984; G. Zurawski, personal communication). As a rule, the r-protein introns belong to the group II type and are self-splicing. The most remarkable of any known r-protein genes is *rps*12, which has three exons, two of which occur in one transcription unit and the third (encoding the NH_2-terminal part) in a separate transcription unit at a great distance (Fig. 2). The maturation of *rps*12 mRNA involves the difficult process of *trans*-splicing (for details, see Chapter 5A).

The patterns of plastid-coded r-protein gene organization in *Euglena gracilis* and in the cyanelle of *Cyanophora paradoxa* are considerably different from those discussed here. For example, both of these contain additional plastid-coded genes (e.g., L3, L5, S17, EF-Tu) and a different kind of intron distribution. Because of space limitations, this subtopic is not discussed in this chapter.

B. r-Protein Genes Located in the Nuclear Genome

A beginning has been made in isolating and characterizing the chloroplast r-protein genes in the nuclear genome. So far 13 of them have been characterized at the cDNA level (Gantt, 1988; Giese and Subramanian, 1989; Phua *et al.*, 1989; Zhou and Mache, 1989; Gantt and Thompson, 1990; Johnson *et al.*, 1990; Smooker *et al.*, 1990; P. M. Smooker and A. R. Subramanian, unpublished results; R. Mache, J. S. Gantt, and D. P. Bourque, personal communication). A schematic diagram of the coding and noncoding segments of five of the characterized cDNA clones are shown in Fig. 3. They all contain a relatively short 5′ upstream and a longer 3′ downstream region of noncoding sequences. The latter presumably contains a polyadenylation signal (apparently two in the case of L13), but the consensus sequence of this signal is presently unclear. The coding sequences include that of a targeting peptide and of the mature r-protein. The latter (Table IV) has a similar degree of amino acid identity to the corresponding *E. coli* r-proteins as those encoded in the plastid genome.

Researchers working on chloroplast r-proteins (see references in Table II) have noted that the latter are on the average larger than *E. coli* r-pro-

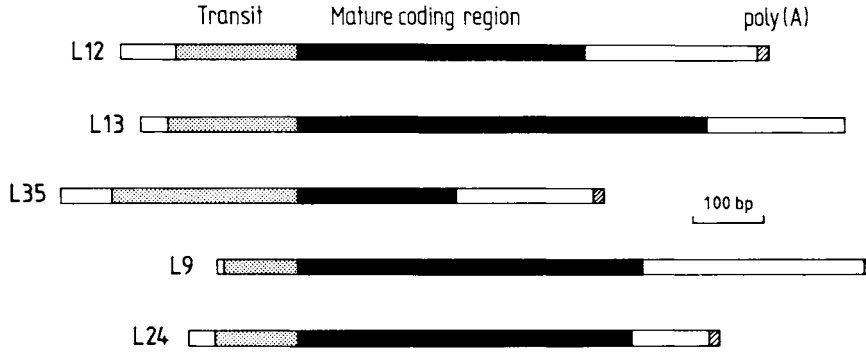


FIG. 3. Schematic representation of the cDNA clones for five chloroplast r proteins. The 5′ and 3′ noncoding regions are unshaded. Evidence indicates that two species of L13 and mRNA each polyadenylated at a separate site may exist (Phua *et al.*, 1989). L9 and L24 are from pea (Gantt, 1988); the others are from spinach (authors' laboratory). Taken from Subramanian *et al.* 1990.

teins. Because the chloroplast r-proteins encoded in the plastid genome are typically of the same size as their *E. coli* homologs (Table IV), the nuclear-coded chloroplast r-proteins are expected to be larger. The work with cDNA clones has confirmed this surmise (Phua *et al.*, 1989; Smooker *et al.*, 1990). The increase in size has mainly come from added extensions at the NH_2 and COOH termini and not due to insertions. A mechanistic origin for such extensions at the NH_2 terminus can be proposed. The nuclear-coded plastid r-protein genes must contain a sequence stretch to target the r-protein. This sequence stretch should also include a cleavage site for the envelope endopeptidase that finally removes the targeting sequence. This additional nucleotide sequence appears fused exactly at the initiating AUG triplet of the original eubacterial gene, abolishing the latter. Because targeting peptides do not have an exact length requirement (see Chapter 10A), the fused sequence stretch could be of variable size, with an extra sequence between the cleavage and fusion sites. According to this hypothesis, the NH_2 extensions would have had no functional importance.

The precise lengths of the transit peptides have been deduced in four cases (L12, L21, L35, and PSrp-1), where the N-terminal sequences of the mature proteins have been determined by amino acid sequencing. The transit peptides are 56 (L12), 55 (L21), 86 (L35), and 66 (PSrp-1) amino acid residues long (Giese and Subramanian, 1989; Smooker *et al.*, 1990; Johnson *et al.*, 1990, and unpublished results). The L35 targeting peptide is longer than the r-protein, whose transport it directs. It is also the longest plastid transit peptide known so far. The transit peptides

show only insignificant sequence identity to each other. They exhibit a characteristic amino acid composition: high content of serine, threonine and proline, lack of tryptophan and tyrosine, and virtual absence of aspartic- and glutamic acids.

The next major step in this field would be the characterization of the nuclear genes of these chloroplast r-proteins. Do these genes occur in single, a few, or many copies? What kind of regulatory regions do they have in their 5′ upstream and 3′ downstream regions to turn them on and off simultaneously? Finally, what is the pattern of distribution of these genes on the plant chromosome (e.g., are they segregated on a few chromosomes or are they randomly dispersed on all)? The answer to the last question has special significance *vis-à-vis* the plasticity of the nuclear chromosomal organization at the time ($>10^9$ yr ago) when endosymbiosis was occurring. Early transmission genetics experiments on the erythromycin loci (i.e., nuclear genes encoding a chloroplast r-protein) in *Chlamydomonas* have shown a distribution in two linkage groups (see Bogorad *et al.*, 1977).

C. Nuclear–Plastid Gene Allocation

Does a single principle govern the known distribution of certain r-protein genes in the plastid genome and others in the nuclear genome? It is too early to answer this question, but it can be discussed.

Twelve of the plastid-encoded r-proteins are constituents of the small ribosomal subunit and eight are components of the large subunit. The ribosome of *E. coli* contains 21 and 33 r-proteins in the small and large subunit, respectively (Table II). Thus, a considerably greater proportion of the small subunit r-proteins are encoded in the plastid (57% vs. 24%). This correlates with the key function of the small ribosomal subunit, i.e., capturing mRNA and positioning it correctly on the decoding site (e.g., see review, Subramanian, 1985). The mRNAs translated by the plastid ribosome are all plastid-encoded.

A second correlation is seen with the ribosome assembly pathway. The bacterial homologs of nearly all the organelle-encoded 50S r-proteins are components of the RI_{50} (assembly intermediate) particle and includes an assembly initiator protein L20. Similarly, the plastid-encoded S proteins include nine such assembly intermediate proteins and the two assembly initiator proteins S4 and S7 of the 30S subunit. The rRNA genes of the plastid ribosome are organelle-encoded. Thus, the ribosome assembly pathway probably has played a key role in the plastid–nuclear gene allocation.

Mutants of *E. coli* that lack one or another of 15 of the *E. coli* r-proteins have been isolated. Those lacking r-proteins are inferred to be not essential for mRNA translation per se. It is a striking fact that only one protein of this group (L33) is maintained in the plastid genome (Fig. 2); the remainder are all presumably nuclear-coded. Therefore, the plastid-encoded r-proteins can be said to belong to an essential category.

V. ORGANIZATION OF PLASTID r-PROTEIN GENES

The 20 r-protein genes of the plastid genome in higher plants are dispersed at 12 loci (Fig. 2). Northern blot analyses have shown (e.g., Ohto *et al.*, 1988) that all the r-protein genes are transcribed. In those cases where the genes are duplicated, whether or not both gene loci are equally transcribed is not known. Multiple transcript bands are the rule indicating polycistronic primary transcripts and subsequent processing. Detailed transcript analysis, including the determination of the 5′ terminus, has been done for some r-protein genes (e.g., *rps*2, *rps*4, *rps*14, *rps*16).

A. Operon Structure

Three examples of the operon organization of plastid r-protein genes are shown in Fig. 4. In the first example, (i.e., that of the largest r-protein operon in the plastid genome), selected genes that belong to three operons in *E. coli* or *B. subtilis* (S10, *spc*, alpha) are fused to form a single transcription unit. The IF-1 gene occurs in the *spc* operon of *B. subtilis* but not of *E. coli* (Boylan *et al.*, 1989).

In the second example (i.e., S12 operon), an extraordinary rearrangement is found: The 5′ portion of the S12 gene is attached to a separate promoter and is located >22 kbp away. Here the cyanobacterial gene organization (e.g., *A.* nidulans; Meng *et al.*, 1989) is similar to that in other eubacteria, but this is not always the case. For example, the gene cluster for L10-L12-β-β' (RNA polymerase) in cyanobacteria has a different organization from that in *E. coli* (Sibold and Subramanian, 1990).

The third example is probably the most interesting because it involves introgression between two functionally distinct gene clusters. Here the *rps*2 gene occurs in the same transcription unit with the genes for several subunits of the thylakoid enzyme complex ATP-synthase. Another

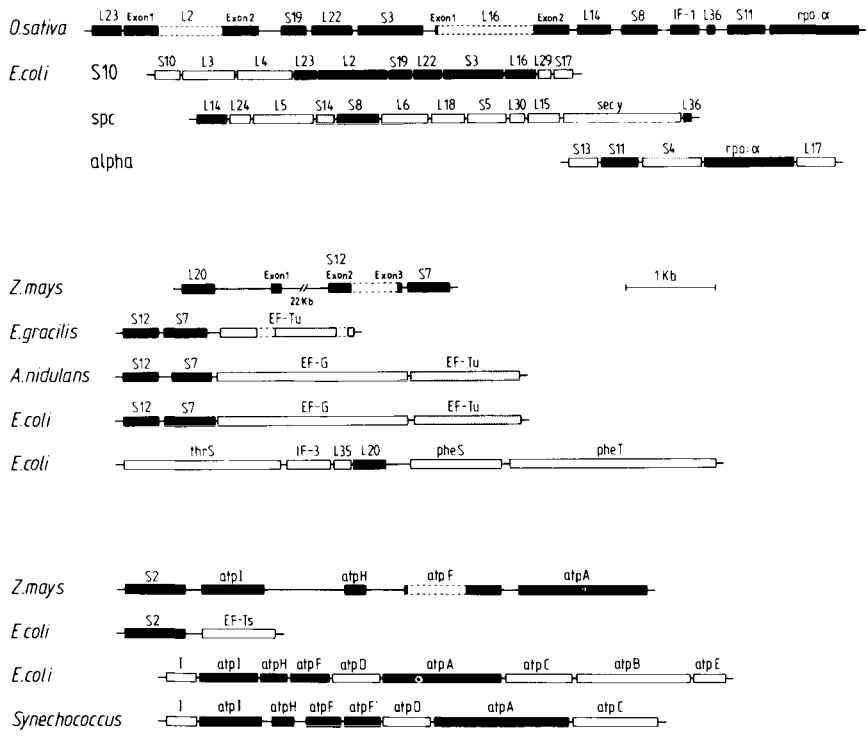


FIG. 4. Three types of gene clusters found for plastid r-proteins. The operons of the homologous eubacterial genes are also shown. Unfilled blocks represent eubacterial genes absent in the plastid genome. See text for details. The organization of S12 operon in Euglena (*E. gracilis*) is from Montandon and Stutz (1984).

similar example is the operon *psaA-psaB-rps*14, where the gene for r-protein S14 occurs together with the genes for two proteins of the photosystem I. Because translational and photosynthetic systems differ considerably in the quantitative requirement of their protein components, especially during plastid development, complex regulatory features are to be expected in these operons.

In *E. coli*, r-protein synthesis is autogenously regulated; i.e., one of the r-proteins encoded by the operon is a repressor that acts by binding to the mRNA (Nomura *et al.*, 1984). The sequence of the mRNA region where the repressor r-protein binds is similar to the sequence of the binding site of this r-protein on its rRNA. However, evidence indicating the existence of such a mechanism in plastids has not been found (e.g., nucleotide sequence homologous to the rRNA-binding site of r-protein

S7 [a repressor in *E. coli*] does not occur in the known sequence of the maize S12 operon [Giese *et al.*, 1987]). It has recently been found that the autogenous mechanism may also not be operative in *B. subtilis* (Henkin *et al.*, 1989). Interestingly, in these aspects plastids apparently resemble the gram-positive branch of the eubacteria more than the gram-negative branch.

B. Transcription–Translation Signals

Transcription of plastid r-protein genes is initiated from promoters that are similar to those found in *E. coli* (reviewed in Subramanian *et al.*, 1990), e.g., TTGTAC(−35) . . . 17 n . . . TTTAAT(−10). Mapping the transcript termini of maize *rps*2 and *rps*4 have revealed that the 5′ end nucleotides begin at expected positions (D. Stahl and A. R. Subramanian, unpublished results; Russel and Bogorad, 1987). No sharp differences between the mapped r-protein promoters and those of other more highly expressed plastid genes (e.g., *rbcL*) have been found. This indicates that regions outside the −10/−35 elements of the promoter as well as mRNA stability could play an important role in plastid transcription–mRNA accumulation (for further discussion, see Chapters 4A and 14B).

The initiation codon of plastid r-protein genes is overwhelmingly an AUG triplet. Two exceptions are *rps*19 and *rpl*2, which are initiated from GUG and ACG codons. The termination codon mostly used in plastid genes is UAA (70%); the UAG and UGA codons are used to the extent of 19 and 11%, respectively (Bonham-Smith and Bourque, 1989). The universal genetic code, without modification, is used in the plastid genome.

Several plastid r-protein genes contain an *E. coli*-like Shine-Dalgarno sequence (Ruf and Kössel, 1988), i.e., a binding site at the 5′ proximal part of the initiating AUG that is complementary to the 3′ end of 16S rRNA, but there are several exceptions as well (e.g., the *rps*3 and *rpl*16 genes of maize contain no discernible Shine-Dalgarno sequences). The essentiality of an *E. coli*-like Shine-Dalgarno sequence in plastid translation is, at present, doubtful. Whether or not some of the PSrps of the plastid ribosome discussed earlier have perhaps taken over this function is still unknown. Shine-Dalgarno-directed initiation is not found in the cytoplasmic 80S ribosomes, where instead a sliding mechanism operates. An aspect of the sliding mechanism was recently revealed by experiments using the mRNA of nuclear-coded chloroplast L12 (Giese and Subramanian, 1990).

VI. RIBOSOME-FREE PLASTIDS IN HIGHER PLANTS

Bleached or chlorotic leaves are often produced in higher plants by nuclear mutations (Börner *et al.*, 1976) or nonpermissive growth temperature (Feierabend and Schrader-Reichhardt, 1976). In the case of rye (*Secale cereale*), this condition is reached at 32°C. It has been shown that heat-bleached rye leaves contain normal amounts of cytoplasmic 80S ribosomes and most other cytoplasmic and plastid components including the plastid DNA, but they lack plastid ribosomes (e.g., Feierabend *et al.*, 1988, a recent reference). The absence of plastid ribosomes means that the proteins encoded in the plastid DNA (including the r-proteins) are no longer synthesized; therefore, new plastid ribosomes could no longer be assembled. The young seedlings with bleached leaves, however, contain plastid ribosomes in their embryonic leaf primordia. These embryonic plastid ribosomes are divided between daughter cells during each cell division, ultimately producing cells that contain only ribosome-free plastids. Those cells that retain embryonic plastid ribosomes are located at the distal end of the leaf, and they develop normal chloroplasts and color the tips green.

We have cloned rye plastid DNA and characterized and sequenced several of its plastid r-protein genes (Prombona and Subramanian, 1989; Prombona, 1989). These sequenced genes as well as the rRNA and some tRNA genes were used as probes to analyze the transcription of plastid DNA in heat-bleached rye.

The plastid genome of higher plants encodes four polypeptides, which, as derived from the amino acid sequence data, are homologs of the α, β, and β' subunits of eubacterial RNA polymerase. If plastid DNA transcription completely depends on an RNA polymerase holoenzyme incorporating these three subunits, no transcripts are expected in ribosome-free plastids. The experiment with heat-bleached rye leaves showed (Fig. 5) an abundant level of plastid transcription. Thus, this experiment revealed that either the essential polypeptide subunits of a functionally active RNA polymerase in plastids is nuclear-coded or the bleached leaves do contain a certain amount of plastid ribosomes.

The patterns of rRNA and tRNA transcripts in heat-bleached rye, shown in Fig. 5, are of particular interest because they appear to reveal that the processing of the precursors of these molecules is severely inhibited. Bands corresponding to mature tRNA (~90 nt) are not seen in the 32°C grown (heat-bleached) plastids, and mature 16S rRNA species (~1,500 nt) is seen only in small amounts. An accumulation of the precursors of higher molecular sizes are seen (Prombona, 1989). Further

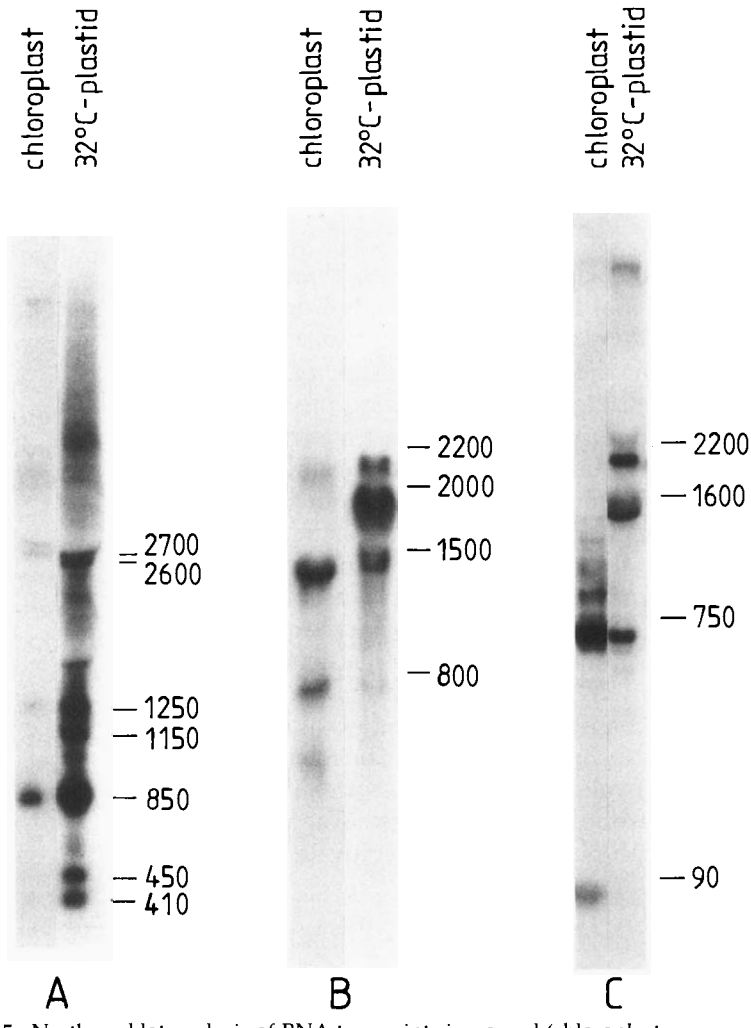


FIG. 5. Northern blot analysis of RNA transcripts in normal (chloroplast, grown at 22°C) and heat-bleached (32°C) plastids of rye. A. Transcripts of r-protein S15 gene. B. Transcripts of 16S rRNA gene. C. Transcripts of the $tRNA^{M}$-$tRNA^{G}$ gene cluster. The mature RNA species (850, 1,500, and 90 nt in A, B, and C, respectively) are found in the chloroplast lanes. Abundant levels of transcripts are found in heat-bleached plastid lanes together with large accumulations of precursor forms.

experiments on this system are necessary to substantiate the inferences from these preliminary experiments.

VII. CONCLUDING REMARKS

The system of plastid r-proteins and their genes has many attractive features as a tool for understanding the origin, functional details, and regulation of synthesis of the translational apparatus. It contains >60 genes, possibly including the homologs of all eubacterial r-protein genes and several additional genes encoding proteins that are specific for the plastid ribosome. According to a recent hypothesis, the latter group of proteins was acquired from the cytoplasmic ribosomes of the endosymbiont host (Johnson *et al.*, 1990). Elucidating the functional roles of these proteins could perhaps bridge the existing gap (e.g., in the initiation step) between the cytoplasmic 80S and eubacterial 70S ribosomes.

Plants are unique in having three translational systems and three categories of ribosomes. Early preliminary experiments (Vasconcelos and Bogorad, 1971) show that few or possibly no ribosomal components (r-protein or rRNA) are used in common among these categories of ribosomes. Thus, plant nuclei are expected to contain >200 genes encoding the r-proteins of its three kinds of ribosomes. This aspect will probably be of practical value when the genes are all cloned, for it will then provide a large resource for testing possible agronomically useful restriction fragment length polymorphism in cultivated plants.

REFERENCES

Bartsch, M. (1985). Correlation of chloroplast and bacterial ribosomal proteins by cross-reactions of antibodies specific to purified *Eschericia coli* ribosomal proteins. *J. Biol. Chem.* **260,** 237–241.

Bartsch, M., Kimura, M., and Subramanian, A.-R. (1982). Purification, primary structure, and homology relationships of a chloroplast ribosomal protein. *Proc. Natl. Acad. Sci. USA* **79,** 6871–6875.

Bogorad, L. (1975). Evolution of organelles and eukaryotic genomes. *Science* **188,** 891–898.

Bogorad, L., Davidson, J. N., and Hanson, M. R. (1977). The genetics of the chloroplast ribosome in *Chlamydomonas reinhardtii*. *In* "Nucleic Acids and Protein Synthesis in Plants" (L. Bogorad and J. H. Weil, eds.), pp. 135–154. Plenum Publishing Corporation, New York.

Bonham-Smith, P. C., and Bourque, D. P. (1989). Translation of chloroplast-encoded mRNA: Potential initiation and termination signals. *Nucleic Acids Res.* **17,** 2057–2080.

Börner, T., Schumann, B., and Hagemann, R. (1976). Biochemical studies on a plastid ribosome-deficient mutant of *Hordeum vulgare*. *In* "Genetics and Biogenesis of Chloroplasts and Mitochondria" (T. Bücher, W. Neupert, W. Sebald, and W. Werner, eds.), pp. 41–48. Elsevier, Amsterdam.

Boylan, S. A., Suh, I. W., Thomas, S. M., and Price, C. W. (1989). Genes encoding the alpha core subunit of *Bacillus subtilis* RNA polymerase is cotranscribed with the genes for initiation factor 1 and ribosomal protein B (L36), S13, S11 and L17. *J. Bacteriol.* **171,** 2553–2562.

Brügger, M., and Boschetti, A. (1975). Two-dimensional gel electrophoresis of ribosomal proteins from streptomycin-sensitive and streptomycin-resistant mutants of *Chlamydomonas reinhardtii*. *Eur. J. Biochem.* **58,** 603–610.

Capel, M. S., and Bourque, D. P. (1982). Characterization of *Nicotiana tabacum* chloroplast and cytoplasmic ribosomal proteins. *J. Biol. Chem.* **257,** 7746–7755.

Davies, E., and Larkins, B. A. (1980). Ribosomes. *In* "The Biochemistry of Plants," Vol. I, "The Plant Cell" (N. E. Tolbert, ed.), pp. 413–437. Academic Press, New York.

Dorne, A.-M., Eneas-Filho, J., Heizmann, P., and Mache, R. (1984). Comparison of ribosomal proteins of chloroplast from spinach and of *E. coli*. *Mol. Gen. Genet.* **193,** 129–134.

Ellis, R. J. (1984). Principles of chloroplast biogenesis. *In* "Chloroplast Biogenesis" (R. J. Ellis, ed.), pp. 1–9. Cambridge University Press, Cambridge.

Eneas-Filho, J., Hartley, M. R., and Mache, R. (1981). Pea chloroplast ribosomal proteins: Characterization and site of synthesis. *Mol. Gen. Genet.* **184,** 484–488.

Feierabend, J., and Schrader-Reichhardt, U. (1976). Biochemical differentiation of plastids and other organelles in rye leaves with a high-temperature-induced deficiency of plastid ribosomes. *Planta* **129,** 133–145.

Feierabend, J., Schlüter, W., and Tebartz, K. (1988). Unassembled polypeptides of the plastidic ribosomes in heat-treated 70S-ribosome-deficient rye leaves. *Planta* **174,** 542–550.

Freyssinet, G. (1977). Characterization of cytoplasmic and chloroplast ribosomal proteins of *Euglena gracilis*. *Biochimie* **59,** 597–610.

Gantt, J. S. (1988). Nucleotide sequences of cDNAs encoding four complete nuclear-encoded plastid ribosomal proteins. *Curr. Genet.* **14,** 519–528.

Gantt, J. S. and Thompson M. D. (1990). Plant cytosolic ribosomal protein S11 and chloroplast ribosomal protein CS17. *J. Biol. Chem.* **265,** 2763–2767.

Giese, K., and Subramanian, A. R. (1989). Chloroplast ribosomal protein L12 is encoded in the nucleus: Construction and identification of its cDNA clones and nucleotide sequence including the transit peptide. *Biochemistry* **28,** 3525–3529.

Giese, K., and Subramanian, A. R. (1990). Enhanced translational utilization of chloroplast ribosomal protein mRNAs from two AUG codons shown by site-directed mutation. *Biochemistry* **29,** 10562–10566.

Giese, K., Subramanian, A. R., Larrinua, I. M., and Bogorad, L. (1987). Nucleotide sequence, promoter analysis, and linkage mapping of the unusually organized operon encoding ribosomal proteins S7 and S12 in maize chloroplast. *J. Biol. Chem.* **262,** 15251–15255.

Grohmann, L., Graack, H. R., and Subramanian, A. R. (1992). Mitochondrial ribosomal rRNAs and proteins and their genes. *In* "The Molecular Biology of Mitochondria," Vol. 8. Academic Press, New York.

Gualerzi, C., Janda, H. G., Passow, H., and Stöffler, G. (1974). Studies on the protein moiety of plant ribosomes. *J. Biol. Chem.* **249,** 3347–3355.

Guitton, C., Dorne, A.-M., and Mache, R. (1984). *In organello* and *in vitro* phosphorylation of chloroplast ribosomal proteins. *Biochem. Biophys. Res. Comm.* **121,** 297–303.

Hachtel, W. (1985). Biosynthesis and assembly of chloroplast ribosomal proteins in isolated chloroplasts from *Vicia faba* L. *Biochem. Physiol. Pflanzen* **180,** 115–124.

Hahn, V., Dorne, A.-M., Mache, R., Ebel, J.-P., and Stiegler, P. (1988). Identification of an *Escherichia coli* S1-like protein in the spinach chloroplast ribosome. *Plant Mol. Biol.* **10,** 459–464.

Hanson, M. R., Davidson, J. N., Mets, L. J., and Bogorad, L. (1974). Characterization of chloroplast and cytoplasmic ribosomal proteins of *Chlamydomonas reinhardtii* by two-dimensional gel electrophoresis. *Mol. Gen. Genet.* **132,** 105–118.

Henkin, T. M., Moon, S. H., Mattheakis, L. C., and Nomura, M. (1989). Cloning and analysis of the *spc* ribosomal protein operon of *Bacillus subtilis:* Comparison with the *spc* operon of *Escherichia coli. Nucleic Acids Res.* **17,** 7469–7486.

Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.-R., Meng, B.-Y., Li, Y.-Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., and Sugiura, M. (1989). The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.* **217,** 185–194.

Johnson, C. H., Kruft, V., and Subramanian, A. R. (1990). Identification of a plastid-specific ribosomal protein in the 30S subunit of chloroplast ribosomes and isolation of the cDNA clone encoding its cytoplasmic precursor. *J. Biol. Chem.* **265,** 12790–12795.

Jones, B. L., Nagabhushan, N., Gulyas, A., and Zalik, S. (1972). Two dimensional acrylamide gel electrophoresis of wheat leaf cytoplasmic and chloroplast ribosomal proteins. *FEBS Lett.* **23,** 167–170.

Kamp, R. M., Srinivasa, B. R., von Knoblauch, K., and Subramanian, A. R. (1987). Occurrence of a methylated protein in chloroplast ribosomes. *Biochemistry* **26,** 5866–5870.

Meng, B.-Y., Shinozaki, K., and Sugiura, M. (1989). Genes for the ribosomal proteins S12 and S7 and elongation factors EF-G and EF-Tu of the cyanobacterium, *Anacystis nidulans:* Structural homology between 16S rRNA and S7 mRNA. *Mol. Gen. Genet.* **216,** 25–30.

Montandon, P. E., and Stutz, E. (1984). The genes for the ribosomal proteins S12 and S7 are clustered with the gene for the EF-Tu protein on the chloroplast genome of *Euglena gracilis. Nucleic Acids Res.* **12,** 2851–2859.

Nomura, M., Gourse, R., and Baughman, G. (1984). Regulation of the synthesis of ribosomes and ribosomal components. *Annu. Rev. Biochem.* **53,** 75–117.

Ohto, C., Torazawa, K., Tanaka, M., Shinozaki, K., and Sugiura, M. (1988). Transcription of ten ribosomal protein genes from tobacco chloroplasts: A compilation of ribosomal protein genes found in the tobacco chloroplast genome. *Plant Mol. Biol.* **11,** 589–600.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., and Ozeki, H. (1986). Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* **322,** 572–574.

Phua, S. H., Srinivasa, B. R., and Subramanian, A. R. (1989). Chloroplast ribosomal protein L13 is encoded in the nucleus and is considerably larger than its bacterial homologue. *J. Biol. Chem.* **264,** 1968–1971.

Posno, M., van Noort, M., Debise, R., and Groot, G. (1984). Isolation, characterization, phosphorylation and site of synthesis of *Spinacia* chloroplast ribosomal proteins. *Curr. Genet.* **8,** 147–154.

Prombona, A. (1989). Molekulare Untersuchungen an Plastiden von *Secale cereale* (Roggen). Ph.D. Dissertation, Free University of Berlin, Berlin.

Prombona, A., and Subramanian, A. R. (1989). A new rearrangement of angiosperm chloroplast DNA in rye (*Secale cereale*) involving translocation and duplication of ribosomal *rpS*15 gene. *J. Biol. Chem.* **264,** 19060–19065.

Ruf, M., and Kössel, H. (1988). Occurrence and spacing of ribosome recognition sites in mRNAs of chloroplasts from higher plants. *FEBS Lett.* **240,** 41–44.

Russel, D., and Bogorad, L. (1987). Transcription analysis of the maize chloroplast gene for the ribosomal protein S4. *Nucleic Acids Res.* **15,** 1853–1867.

Schmidt, R. J., Richardson, C. B., Gillham, N. W., and Boynton, J. E. (1983). Site of synthesis of chloroplast ribosomal proteins in *Chlamydomonas*. *J. Cell Biol.* **96,** 1451–1463.

Sibold, C., and Subramanian, A. R. (1990). Cloning and characterization of the genes for ribosomal proteins L10 and L12 from *Synechocystis* Sp. PCC 6803. *Biochim. Biophys. Acta,* **1050,** 61–68.

Smooker, P. M., Kruft, V., and Subramanian, A. R. (1990). A ribosomal protein is encoded in the chloroplast DNA in a lower plant but in the nucleus in angiosperms. *J. Biol. Chem.*, **265,** 16699–16703.

Subramanian, A. R. (1974). Sensitive separation procedure for *Escherichia coli* ribosomal proteins and the resolution of high-molecular weight components. *Eur. J. Biochem.* **45,** 541–546.

Subramanian, A. R. (1985). The ribosome: Its evolutionary diversity and the functional role of one of its components. *Essays Biochem.* **21,** 45–85.

Subramanian, A. R., Smooker, P. M., and Giese, K. (1990). Chloroplast ribosomal proteins and their genes. *In* "The Ribosome. Structure, Function, and Evolution" (W. E. Hill, A. Dahlberg, R. A. Garrett, P. B. Moore, D. Schlessinger and J. R. Warner, eds.), pp. 655–663. American Society for Microbiology, Washington, D.C.

Udalova, G. V., Yurina, N. P., and Odintsova, M. S. (1984). Ribosomes of the chloroplasts: Stoichiometry of proteins. *Dokl. Akad. Nauk SSR* **278,** 1258–1262.

Umesono, K., and Ozeki, H. (1987). Chloroplast gene organization in plants. *Trends Genet.* **3,** 281–287.

Vasconcelos, A. C. L., and Bogorad, L. (1971). Proteins of cytoplasmic, chloroplast, and mitochondrial ribosomes of some plants. *Biochim. Biophys. Acta* **228,** 492–502.

Wittmann, H. G. (1982). Components of bacterial ribosomes. *Annu. Rev. Biochem.* **51,** 155–183.

Wittmann, H. G. (1989). Ribosomen und Proteinbiosynthese. *Biol. Chem. Hoppe-Seyler* **370,** 87–99.

Zhou, D.-X., and Mache, R. (1989). Presence in the stroma of chloroplasts of a large pool of a ribosomal protein not structurally related to any *Escherichia coli* ribosomal protein. *Mol. Gen. Genet.* **219,** 204–208. Erratum: *Mol. Gen. Genet.* **223,** 167, 1990.

Zhou, D.-X., Quigley, F., Massenet, O., and Mache, R. (1989). Cotranscription of the S10- and *spc*-like operons in spinach chloroplasts and identification of three of their gene products. *Mol. Gen. Genet.* **216,** 439–445.

Zurawski, G., Bottomley, W., and Whitfeld, P. R. (1984). Junctions of the large single copy region and the inverted repeats in *Spinacia oleracea* and *Nicotiana debneyi* chloroplast DNA: Sequence of the genes for $tRNA^{His}$ and the ribosomal proteins S19 and L2. *Nucleic Acids Res.* **12,** 6547–6558.

CHAPTER 9

Structure, Organization, and Properties of Plastid Envelope Membranes

Roland Douce
Jacques Joyard

Laboratoire de Physiologie Cellulaire Végétale
URA CNRS nº 576
Département de Biologie Moléculaire et Structurale
Centre d'Etudes Nucléaires de Grenoble et
Université Joseph Fourier
F-38041 Grenoble, France

I. INTRODUCTION

A large variety of plastid types is found in plants, and they are all, to some extent, interconvertible. The best-known plastid is the chloroplast, which occurs in leaves and other green tissues. Other plastids include proplastids, which are undifferentiated plastids present in meristematic cells; etioplasts, which are plastids formed in leaves from plants grown in the dark; amyloplasts, which contain large starch grains and are found in stems, roots, and tubers; and chromoplasts, which are enriched in carotenoids and are present in petals, fruits, and even roots. The concept of autonomy and continuity of plastids originated from mircoscopy studies (e.g., see Guillermond, 1941) and was further supported by numerous biochemical analyses (Kirk and Tilney-Bassett, 1978). The only common membrane structure among all plastid types is their two limiting membranes. This membrane system, known as the plastid envelope, remains present during plastid division and is involved in plastid continuity, because it is transmitted to the daughter cells during cell division. Therefore, it is not surprising that envelope membranes are involved in plastid biogenesis (Douce and Joyard, 1979, 1984; Douce *et al.*, 1984; Douce and Joyard, 1990).

Plastids are crucial for plant cell metabolism: They are the site for CO_2 reduction and its assimilation into carbohydrates, amino acids, fatty acids, and various terpenoid compounds (Kirk and Tilney-Bassett, 1978). They are also the site for NO_2 and SO_4 reduction and their assimilation into amino acids (Kirk and Tilney-Bassett, 1978). Envelope membranes, which are at the border between plastids and their surrounding cytosol, are therefore essential for the control of uptake of raw material, for all synthesis occurring in the plastids, and for the regulation of the export to the cytosol of newly synthesized molecules necessary for cell metabolism.

In this review, we focus on the organization of the envelope membranes from a structural, biochemical, and physiological point of view.

II. STRUCTURE OF THE PLASTID ENVELOPE

A. Electron Microscopic Analyses of the Plastid Envelope

In chloroplasts from spinach mesophyll cells, the apparent thickness of both envelope membranes (when observed under the electron microscope) is about 5.5 nm (Fig. 1), like that of thylakoids (Carde *et al.*, 1982). Freeze-fracture studies of chloroplast and etioplast envelopes show that the two envelope membranes clearly differ from each other with respect to particle distribution (Sprey and Laetsch, 1976; Simpson, 1978a,b; Cline *et al.*, 1985). Values for total bilayer particle densities are probably in the range of 1,500–3,000 particles/μm^2, and 7,500–10,000 particles/μm^2, for the outer and inner envelope membranes, respectively, from spinach, barley (Simpson, 1978b), and pea chloroplasts (Cline *et al.*, 1985). These differences probably reflect the high lipid : protein ratio (2.5–3 mg lipid/mg protein) of the outer envelope membrane compared with that of the inner envelope membrane (about 1–1.2 mg lipid/mg protein) (Block *et al.*, 1983a), which are responsible for the major difference in membrane densities: 1.08 g/cm^3 for the outer envelope membrane and 1.13 g/cm^3 for the inner envelope membrane (Cline *et al.*, 1981; Block *et al.*, 1983a).

B. Occurrence, Organization, and Putative Role of the Contact Points between Outer and Inner Envelope Membranes

There are no continuities between the outer and inner envelope membranes. However, the intermembrane space is not completely electron translucent. For example, Carde *et al.* (1982) showed that in many places electron dense areas seem to cement both membranes (contact sites). The density of such structures is greater in developing plastids than in mature chloroplasts, suggesting that *in vivo* occurrence of the contact points between the two envelope membranes is probably restricted in both time and space. The density of the contact points strongly decreases when isolated intact chloroplasts are incubated in a hypertonic medium (0.6 *M* sorbitol), which separates the two envelope membranes by a wide and irregular gap due to the differences in their permeability properties, thus suggesting that the contact points are fragile structures

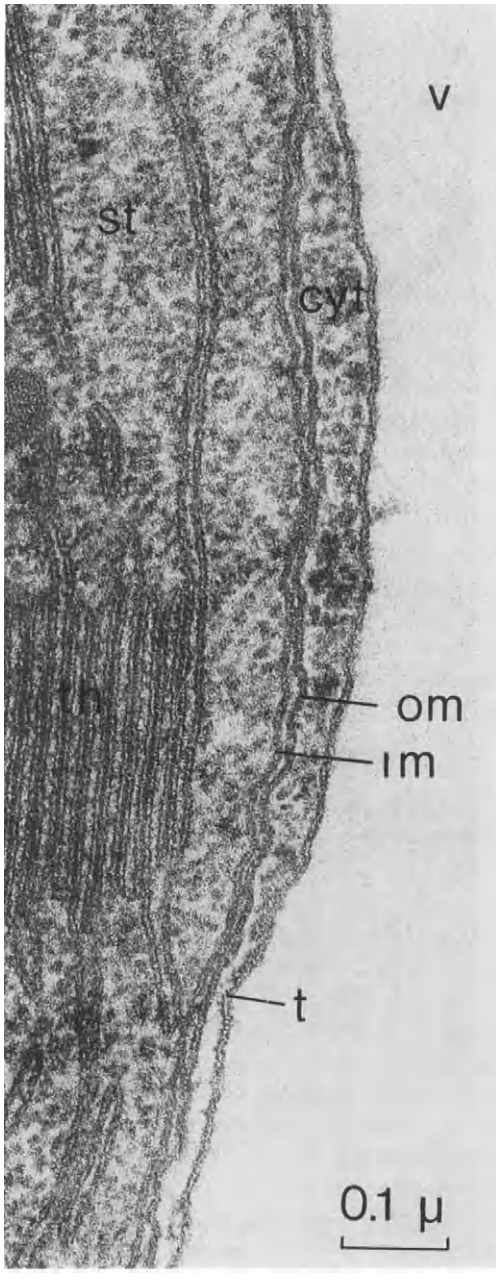


FIG. 1. Thin section of part of spinach (*Spinacia oleracea* L.) leaf cell. The chloroplast envelope consists of two morphologically and topologically distinct membranes. cyt, cytosol; im, inner envelope membrane; om, outer envelope membrane; st, stroma; t, tonoplast; th, thylakoids; v, vacuole. [Micrograph adapted from Carde *et al.* (1982), with permission.]

(Carde *et al.*, 1982). Intact chloroplasts maintained in a high osmolarity medium were also used by Cline *et al.* (1985) and Cremers *et al.* (1988) to analyze, by freeze-fracture, the structure of contact sites between outer and inner envelope membranes. For instance, using ultrarapidly frozen pea chloroplasts, Cline *et al.* (1985) observed structures that may be related to the putative contact sites. These structures were studied, using freeze-fracture and freeze substitution, by Cremers *et al.* (1988) and were clearly recognized in freeze-fracture replicas of envelope membranes by ridges of a high density of intramembrane particles. However, the chemical nature of the intermembrane contact points is completely unknown. One possibility is the nonbilayer forming monogalactosyldiacylglycerol (MGDG), the major glycerolipid in envelope membranes, which could favor membrane fusion (see Section II.B).

Chua and Schmidt (1979) first proposed that these contact points could be the site where plastid proteins that are synthesized on cytoribosomes enter the chloroplast. In support of this hypothesis, Carde *et al.* (1982) observed that some thin threads (uncoiled proteins?) apparently pass through both envelope membranes joining the cytoplasmic space to the stroma space. Although further work is needed to establish whether or not the suggestive pictures presented by Carde *et al.* (1982) actually reflect protein transport through both envelope membranes, the existence of envelope structure(s) involved in the recognition and transport of proteins is highly likely. In fact, some evidence demonstrates that contact sites between the inner and outer mitochondrial membranes are involved in the transfer of proteins from the cytosol to the mitochondrial matrix (see Schwaiger *et al.*, 1987). It is therefore very tempting to suggest that the same is true for plastids. Unfortunately, the evidence is still lacking. Pain *et al.* (1988), using anti-idiotypic antibodies, suggested that contact sites between outer and inner envelope membranes contain binding sites for precursor proteins. However, characterization of this receptor is still controversial (for reviews, see Schmidt and Mishkind, 1986; Mishkind and Scioli, 1988; Keegstra *et al.*, 1989; Douce and Joyard, 1990) and hampers unambiguous demonstration that contact points are indeed the site for protein transport across the envelope membranes.

C. Relationship between the Inner Envelope Membrane and the Internal Membranes of Plastids

In mature spinach chloroplasts, no free connections appear between the intermembrane space of the envelope and the thylakoid space (Carde *et*

al., 1982). However, observations of numerous small vesicles evaginating from the inner envelope membranes are quite common in developing plastids (e.g., see Carde *et al.*, 1982). Such figures strongly suggest that thylakoids could be formed by continuous invagination of the inner envelope membrane (e.g., see Wellburn, 1982). The observation that purified inner envelope membranes and thylakoids share the same polar lipid composition (Cline *et al.*, 1981; Block *et al.*, 1983b) strongly favors the idea that fusion of vesicles produced by the inner envelope membrane with growing thylakoids could occur. In addition, MGDG, the major lipid in plastid membranes, forms hexagonal type II (H_{II}) structures are of major interest as intermediates in membrane fusion (Ellens *et al.*, 1989). Therefore, MGDG (which represents almost half of the lipid content of inner envelope vesicles and thylakoids) probably favors fusions between plastid membranes. On the other hand, both envelope membranes are completely different from thylakoids with respect to their protein composition and their functions (Douce *et al.*, 1984). If the vesicles deriving from the inner envelope represent rudimentary thylakoids, considerable modifications must take place following the initial step of invagination: Vesicles deriving from the inner envelope membrane should be modified by both degradation of inner envelope proteins (unless they consist only of a lipid matrix) and addition of new polypeptides to acquire unique thylakoid characteristics. For instance, Wrischer (1989) studied the localization of photosynthetic activity during development in young plastids (with a single prothylakoid) and mature chloroplasts from maize. The activity of photosystem I (detected by photo-oxidation of 3,3′-diaminobenzidine) and of photosystem II (detected by photoreduction of thiocarbamyl nitrotetrazolium blue) could be detected *in situ*, by electron microscopic cytochemical methods, only in thylakoids (even at early stages of development) but never in envelope membranes (Wrischer, 1989).

However, some major problems are related to thylakoid formation from envelope vesicles. It is well established that membrane synthesis never occurs *de novo*, a membrane always derives from a preexisting membrane, and thylakoids are probably not an exception to this rule. Therefore, vesicles deriving from the inner envelope should be added to a preexisting membrane, whose nature is totally unknown. Interestingly, cauliflower proplastids contain only a rudimentary internal membrane system (prethylakoids), connected with the inner envelope membrane (Journet and Douce, 1985). We have demonstrated that these internal membranes contain chlorophyll and numerous specific polypeptides, which are absent from envelope membranes, but both mem-

brane systems also contain some polypeptides (e.g., a major 31,000-daltons polypeptide) and some enzymatic activities (acyl–acyl-carrier protein [ACP] : 1-acyl-*sn*-glycerol-3-phosphate acyltransferase) in common (Alban *et al.*, 1988, 1989). Finally, another major problem is understanding how vesicle formation from the inner envelope membrane is triggered under local control.

D. Relationship between the Outer Envelope Membrane and Extraplastidial Membranes

Close associations of the outer envelope membrane and extraplastidial membranes has been reported (for a review, see Douce and Joyard, 1979). Such structural relationships have been taken as an argument for continuity of the outer envelope membrane with the cell endomembrane system. As discussed by Douce and Joyard (1979) and Douce *et al.* (1984), appropriate biochemical investigations do not support the generalization of the endomembrane concept to the outer envelope membrane. However, direct continuities of the outer envelope membrane with cytoplasmic membranes, and especially with the endoplasmic reticulum, may occur in some specialized tissues. For instance, in secretory cells such as the resin canal from pine needles, multilobate leucoplasts are sheathed by a layer of fenestrated endoplasmic reticulum associated with the envelope membrane (Carde and Bernard-Dagan, 1982). Terpenes synthesized in envelope membranes are discharged into the endoplasmic reticulum, where they accumulate before being released into the periplasmic space of the cell (Carde and Bernard-Dagan, 1982). Under these conditions, the endoplasmic reticulum behaves simply as a passive corridor that transfers these harmful terpenes from their biosynthetic site (envelope membranes) to the accumulation site without contact with the cytoplasm (Douce *et al.*, 1984).

III. CHEMICAL COMPOSITION OF PLASTID ENVELOPE MEMBRANES

The key step in the preparation of plastid envelope membranes is the purification of large amounts of intact organelles devoid of extraplastidial contaminants (Douce and Joyard, 1982). This can now be achieved with almost all kinds of plastids (chloroplasts, proplastids, etioplasts,

amyloplasts, and chromoplasts). Procedures to separate total envelope membranes from chloroplasts (Douce *et al.*, 1973) and to fractionate outer and inner envelope membranes (Cline *et al.*, 1981; Block *et al.*, 1983a,b) are available. Separating envelope membranes from cauliflower bud proplastids (Alban *et al.*, 1988) and sycamore cell amyloplasts (Ngernprasirtsiri *et al.*, 1988; Alban *et al.*, 1988) is also possible.

A. Polypeptides

Envelope membranes contain a very large number of polypeptides with M_r ranging from <10,000 to >120,000 daltons (Douce *et al.*, 1984). Analyses of spinach chloroplast envelope polypeptides by two-dimensional gel electrophoresis (Fig. 2) have revealed >75 distinct polypeptides after staining with Coomassie Blue (Joyard *et al.*, 1982). But this number is widely underestimated: During the different chromatographic steps used for purification of integral envelope proteins (e.g., see Covès *et al.*, 1986, 1988b), numerous additional polypeptides present in very low amounts can be detected, especially after silver nitrate staining. In fact, envelope membranes probably contain several hundred different polypeptides.

The envelope polypeptides fall roughly into two groups on the basis of their staining intensity. There are several major polypeptides (E54, E37, E30, E14, and E12), and a multitude of faint bands spread all over the gels. Such a distribution probably reflects the dual function of the plastid envelope: Most of the major polypeptides are probably involved in the transport of metabolites across the envelope, whereas the minor polypeptides could be involved in the synthesis of plastid components (glycerolipids, prenylquinones, chlorophyllide, etc.). For instance, one of the major envelope polypeptides has a M_r of 30,000 (E30) and was identified by Flügge and Heldt (1976) as the phosphate–triose phosphate translocator. Because the inner envelope membrane regulates the flow of metabolites between the cytosol and the chloroplast stroma, the envelope polypeptide E30 is a good marker for the inner envelope membrane. In contrast, protochlorophyllide oxidoreductase is only a minor envelope component and has a M_r of 36,000–37,000 (Joyard *et al.*, 1990). Unfortunately, except for these two examples, almost no other envelope proteins have been identified on the polypeptide profile obtained from chloroplasts. Interestingly, envelope membranes isolated from chloroplasts contain both the large and the small subunits of ribulose-1,5-bisphosphate carboxylase (Pineau and Douce, 1974; Joyard *et al.*, 1982). This probably reflects contamination by the major stromal protein (about 80% of the stromal proteins).

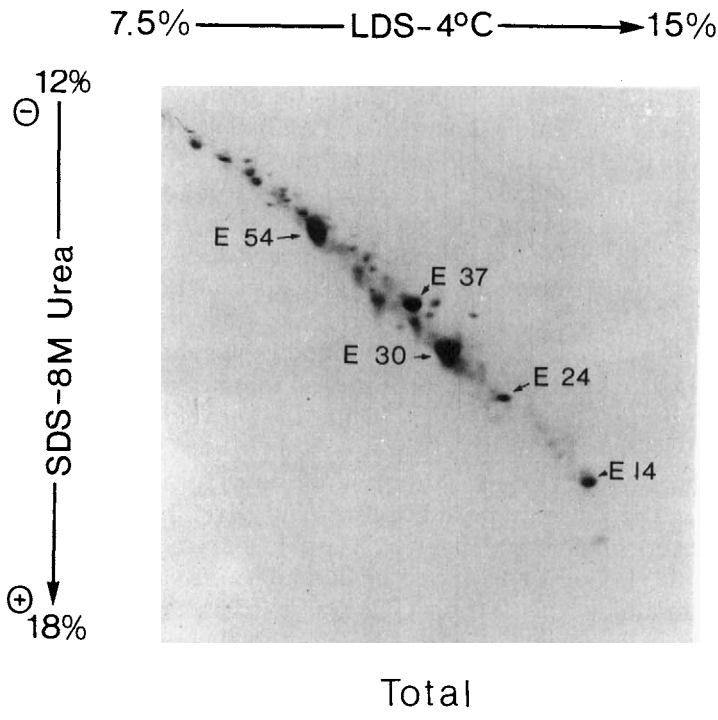


FIG. 2. Two-dimensional electrophoretic analyses of envelope membranes. In the first dimension, an LDS-polyacrylamide gel with a 7.5–15% acrylamide gradient (at 4°C) was used. In the second dimension, an SDS-polyacrylamide gel with a 12–18% acrylamide gradient, in presence of 8 *M* urea (at room temperature), was used. [Adapted from Joyard *et al.*, 1982, with permission.]

As a first step toward characterization of envelope polypeptides, proteolytic digestion and immunological techniques were used to identify those outer membrane polypeptides that were accessible from the cytosolic side.

Of all the nonpenetrating proteolytic enzymes that have been used on intact chloroplasts, thermolysin was demonstrated to be the most suitable protease to probe the cytosolic surface of the outer envelope membrane (Joyard *et al.*, 1983, 1987; Cline *et al.*, 1984). Up to 20 different envelope polypeptides are susceptible to mild digestion of isolated intact chloroplasts by thermolysin. Most of the thermolysin-sensitive envelope polypeptides are not extracted by a mixture of chloroform–methanol (2:1, v/v). A clear exception was a 10,000-daltons polypeptide (E10), which is hydrophobic and is an integral membrane polypeptide. All these thermolysin-sensitive polypeptides are probably located on the outer envelope membrane because E30, an inner envelope polypeptide,

was not affected by the treatment. This hypothesis was confirmed by using rabbit antibodies raised against polypeptides sensitive (E10 and E24) and insensitive (E30 and E37) to thermolysin (Joyard *et al.*, 1983). Using these antibodies, Joyard *et al.* (1983) demonstrated that only antibodies to E10 and E24 induced agglutination of intact chloroplasts or gave a green fluorescence at the outer surface of intact chloroplasts (after addition of fluorescent goat antibodies raised against rabbit antibodies). From these experiments, it is possible to conclude that E10 and E24, and probably all thermolysin-sensitive envelope polypeptides, are accessible from the cytosolic side of the outer membrane of the chloroplast envelope (Joyard *et al.*, 1983). These antibodies were used for the characterization of membrane fractions enriched in outer (antibodies to E10 and E24) and inner (antibodies to E30 and E37) envelope membranes (Block *et al.*, 1983a).

When isolated intact chloroplasts or envelope membranes are incubated in the presence of radioactive ATP, several phosphoproteins, located in envelope membranes (Soll and Buchanan, 1983; Laing and Chriseller, 1984; Soll, 1985), can be identified after autoradiography of polyacrylamide gels. When GTP was used instead of ATP, only two outer envelope proteins were labeled: 23,000- and 32,500-daltons polypeptides (Soll *et al.*, 1988). A 64,000-daltons phosphoprotein was very rapidly labeled at very low ATP concentration (30 n*M*) (Soll and Bennett, 1988). This polypeptide is a soluble protein and is probably present in the intermembrane space of the envelope. All these results suggest that protein kinases are present in envelope membranes. However, at present data are insufficient to conclude that phosphoprotein and protein kinases have a major physiological role (e.g., as enzymes for regulation of cellular functions).

B. Glycerolipids

Colorless lipids are, in terms of quantity, the major constituent of both envelope membranes. In contrast to extraplastidial membranes, plastid membranes (envelope and thylakoids) are characterized by a low phospholipid content and by the presence of glycolipids, i.e., galactolipids and sulfolipid (Douce *et al.*, 1973). The major plastid glycerolipids are galactolipids, which contain one or two galactose molecule(s) attached to the *sn*-3 position of the glycerol backbone, i.e., MGDG and digalactosyldiacylglycerol (DGDG). Interestingly, these glycerolipids are also the major components in cyanobacteria and oxychlorobacteria, which

are supposed to have a common ancestor with chloroplasts.

Siebertz *et al.* (1979) have analyzed envelope and thylakoid glycerolipids and have demonstrated that their structure is identical in both membrane systems (Fig. 3). Galactolipids in thylakoid and envelope membranes contain a high amount of polyunsaturated fatty acids: Up to 95% (in some species) of the total fatty acid is linolenic acid (18:3). In nongreen plastids, 18:3 is still a major component, although appreciable amounts of 18:1 and 18:2 are present. Therefore, the most abundant molecular species of MGDG and DGDG have 18:3 at both *sn*-1 and *sn*-2 positions of the glycerol backbone (Fig. 3). Some plants, such as pea, having almost only 18:3 in MGDG are called "18:3 plants." Other plants, such as spinach, contain significant amounts of 16:3 in MGDG and are called "16:3 plants" (Heinz, 1977). These two types are also found when nongreen plastids are analyzed: For instance, sycamore is an 18:3 plant, whereas cauliflower is a typical 16:3 plant. The positional distribution of 16:3 in MGDG is highly specific: This fatty acid is only present at the *sn*-2 position of glycerol, and is almost excluded from the *sn*-1 position. Therefore, two major structures are found in galactolipids: one with C18 fatty acids at both *sn* positions and one with C18 and C16 fatty acids, respectively, at the *sn*-1 and *sn*-2 positions (Fig. 3). The first one is typical of "eukaryotic" lipids (such as phosphatidylcholine [PC]) and the second one corresponds to a "prokaryotic" structure. These differences are probably due to galactolipid biosynthetic pathways (for reviews see Roughan and Slack, 1982; Joyard and Douce, 1987).

Plastid membranes sometimes contain galactolipids with three (trigalactosyldiacylglycerol [tri-GDG]) and four (tetra-GDG) galactoses, which are not naturally present in plants. They are formed *in vitro*, during the course of envelope membrane purification, owing to a galactolipid–galactolipid galactosyltransferase (Van Besouw and Wintermans, 1978; Heemskerk and Wintermans, 1987), which catalyzes an enzymatic galactose exchange between MGDG and DGDG according to the following reactions:

$$2\text{MGDG} \rightarrow \text{DGDG} + \text{diacylglycerol}$$
$$\text{MGDG} + \text{DGDG} \rightarrow \text{tri-GDG} + \text{diacylglycerol}$$
$$2\ \text{DGDG} \rightarrow \text{tetra-GDG} + \text{diacylglycerol}$$

This enzyme is susceptible to proteolytic digestion by the nonpenetrating protease thermolysin and is therefore located on the cytosolic side of the outer envelope membrane (Dorne *et al.*, 1982). In addition, the galactolipid–galactolipid galactosyltransferase is present in all plastids analyzed so far, chloroplasts and nongreen plastids (Alban *et al.*,

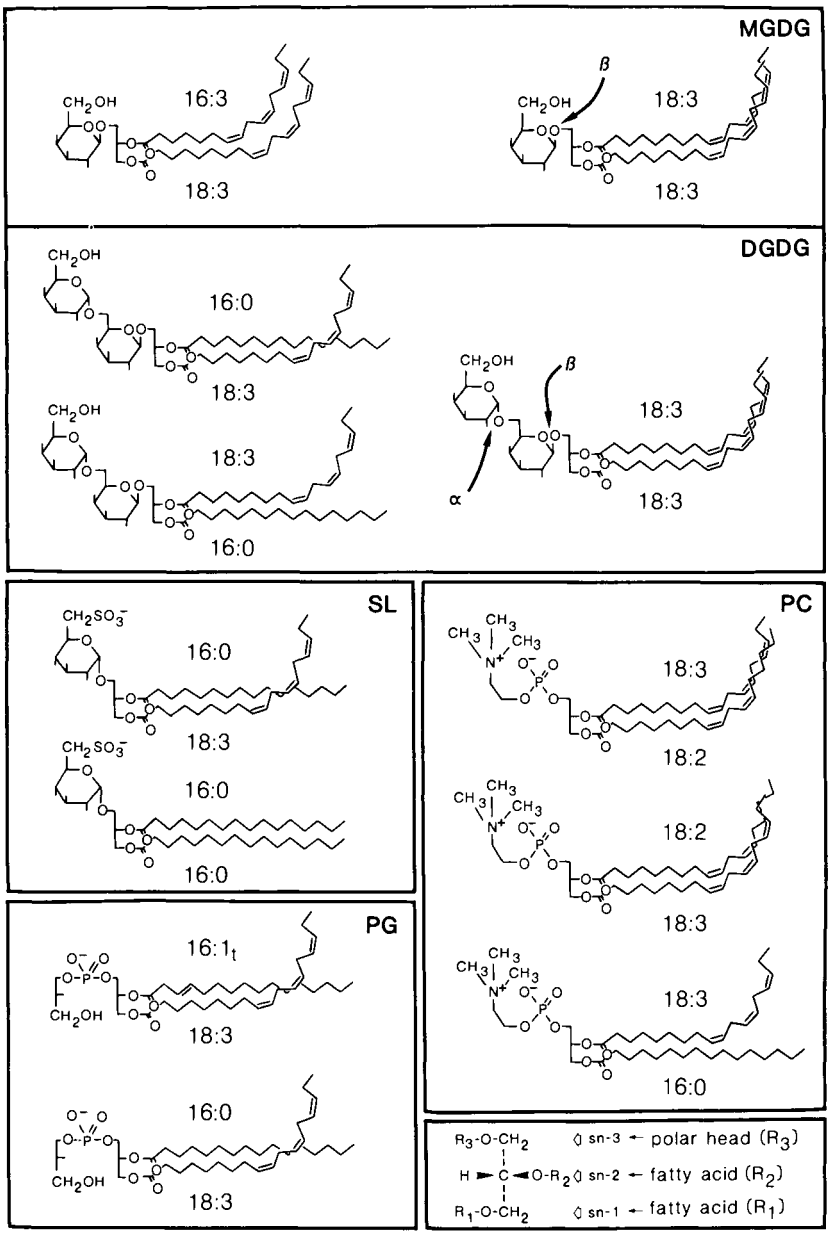


FIG. 3. Structure of the major envelope membrane glycerolipids. DGDG, diagalactosyldiacylglycerol; MGDG, monogalactosyldiacylglycerol; PC, phosphatidylcholine; PG, phosphatidylglycerol; SL, sulfolipid.

1988). Consequently, to obtain a glycerolipid composition, which could represent the *in vivo* situation within plastid membranes, the galactosyltransferase should be destroyed prior to fractionation of plastids (Dorne *et al.*, 1982; Douce *et al.*, 1984).

The glycerolipid composition of envelope membranes from chloroplasts and nongreen plastids is given in Table I. As expected after thermolysin treatment, no diacylglycerol can be detected in envelope membranes (whereas it represents about 10% of the total envelope glycerolipids from nontreated plastids). The most striking feature is that the glycerolipid pattern is almost identical in envelope membranes from chloroplasts, etioplasts, or other nongreen plastids (Douce and Joyard, 1979; Douce *et al.*, 1987).

In the outer envelope membranes, the two galactolipids MGDG and

TABLE I
Glycerolipid Composition of Plant Cell Membranes[a]

Organelle	MGDG	DGDG	SL	PC	PG	PI	PE	DPG
Mitochondria								
Pea leaves								
Total membranes	0	0	0	40	2	6	46	14
Cauliflower buds								
Total membranes	tr	tr	0	37	2	8	38	13
Sycamore cells								
Total membranes	0	0	0	43	3	6	35	13
Inner membrane	*0*	*0*	*0*	*41*	*2.5*	*5*	*37*	*14.5*
Outer membrane	*0*	*0*	*0*	*54*	*4.5*	*11*	*30*	*0*
Peroxisomes								
Potato tubers	0	0	0	52	0	0	48	0
Plastids								
Spinach chloroplasts								
Thylakoids	52	26	6.5	4.5	9.5	1.5	0	0
Envelope								
Inner membrane	*49*	*30*	*5*	*6*	*8*	*1*	*0*	*0*
Outer membrane	*17*	*29*	*6*	*32*	*10*	*5*	*0*	*0*
Total envelope	32	30	6	20	9	4	0	0
Pea etioplasts								
Total envelope	34	31	6	17	5	4	0	0
Cauliflower proplastids								
Total envelope	31.5	27.5	6	20	9	4.5	1	0

[a] This table summarizes some of the data obtained in the authors' laboratory on the glycerolipid composition of plant cell membranes. Composition of plastid membranes was obtained after thermolysin treatment of intact organelles.

DPG, diphosphatidylglycerol or cardiolipin; DGDG, digalactosyldiacylglycerol; MGDG, monodigalactosyldiacylglycerol; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PG, phosphatidylglycerol; PI, phosphatidylinositol; SL, sulfolipid.

DGDG are present in a ratio of 0.6:1, whereas in the inner envelope membrane this ratio is 1.6:1 and is highly comparable with that determined for thylakoids (Block *et al.*, 1983b). Sulfolipid is present in all plastid membranes and is a sulfonic acid derivative of 6-deoxyglucose, as demonstrated by Benson *et al.* (1959). Billecocq *et al.* (1972) and Billecocq (1974, 1975) have shown, by means of specific antibodies, that galactolipids and sulfolipid are present in the cytosolic leaflet of the outer envelope membrane. All these observations clearly demonstrate that the outer envelope membrane does not derive from extraplastidial membranes but, rather, is a typical plastidial membrane.

PC is the major envelope phospholipid and represents about 30–35% of the outer envelope glycerolipid (Cline *et al.*, 1981; Block *et al.*, 1983b). Using phospholipase *c* digestion of intact chloroplasts, Dorne *et al.* (1985, 1989) demonstrated that chloroplast PC is concentrated in the outer leaflet of the outer membrane. PC present only in this membrane leaflet facing the cytosol derives from the endoplasmic reticulum via a phospholipid transfer protein (Miquel *et al.*, 1987). It is interesting to note that the only major difference between cyanobacterial and chloroplast glycerolipids is due to PC, which is generally absent from cyanobacteria and is considered to be a typical eukaryotic lipid (Fig. 3).

The major phospholipid in the inner envelope membrane and in thylakoids is phosphatidylglycerol (PG), which is unique because of a $16{:}1_{trans}$ fatty acid at the *sn*-2 position of the glycerol backbone (Fig. 3) (Haverkate and van Deenen, 1965). Therefore, this phospholipid is different from that found in extraplastidial membranes. The $16{:}1_{trans}$ fatty acid is present in PG from all chloroplast membranes, envelope as well as thylakoids, but is absent from plastid envelope membranes purified from nonphotosynthetic tissues (Alban *et al.*, 1988).

Finally, Table I also indicates that plastid membranes are devoid of phosphatidylethanolamine (PE), which is a major component (together with PC) of mitochondrial, endoplasmic reticulum, or other extraplastidial membranes. The presence of very small amounts of PE in envelope membranes from cauliflower bud proplastids (Alban *et al.*, 1988) or from *Chlamydomonas* chloroplasts (Mendiola-Morgenthaler *et al.*, 1985) reflects the presence of a small amount of contaminating extraplastidial membranes. Preparing such plastids in a totally pure state is extremely difficult.

C. Pigments

In contrast to thylakoids, envelope membranes from chloroplasts or nongreen plastids are yellow, due to the presence of carotenoids and the

TABLE II
Carotenoid Composition of Plastid Membranes[a]

Plastid	Anthera-xanthin	Lutein + Zea-xanthin	Viola-xanthin	Neoxanthin	β-carotene
Spinach chloroplasts					
Thylakoids	–	37	22	15.5	25
Envelope	7	21	56	4	12
Sycamore amyloplasts					
Envelope	–	20	65	5	10
Cauliflower proplastids					
Envelope	–	15	58	14.5	12.5

[a] Values expressed as percentage of total amount of carotenoids. –, not detected.

absence of chlorophyll. Interestingly, envelope membranes from chloroplasts and from most nongreen plastids present the same absorption spectrum, thus demonstrating that they have a very close pigment composition (Table II). In envelope membranes from all plastid types, violaxanthin is the major carotenoid, whereas thylakoids are richer in β-carotene (Douce *et al.*, 1984). For instance, the xanthophyll : β-carotene ratio is about 6 in chloroplast envelope membranes, but only 3 in thylakoids. Both chloroplast envelope membranes contain carotenoids (Block *et al.*, 1983b), but whether or not these molecules are present as protein-carotenoid complexes is not yet known because the procedures used to separate envelope proteins are not mild enough to preserve the integrity of such complexes, if they do occur. The possible role(s) of carotenoids in envelope membranes is still a matter of conjecture (Douce *et al.*, 1984; Douce and Joyard, 1990).

Although devoid of chlorophyll, the plastid envelope contains chlorophyll precursors such as protochlorophyllide (Pineau *et al.*, 1986). Interestingly, Hinterstroisser *et al.* (1988) and Peschek *et al.* (1989) have demonstrated that chlorophyll-free cytoplasmic membranes from *Anacystis nidulans*, in contrast to thylakoids, contain significant amounts of protochlorophyllide and chlorophyllide.

D. Prenylquinones

Lipid extracts of spinach chloroplast envelopes absorb light in the UV range from 230 to 290 nm. This observation demonstrates that in addition to carotenoids envelope membranes from spinach chloroplasts contain prenylquinones as genuine components. As shown in Table III, both envelope membranes contain plastoquinone-9, α-tocopherol, and

TABLE III
Prenylquinone Composition of Spinach Chloroplast Membranes[a]

Prenylquinone	Thylakoids	Envelope	
		Inner Membrane	Outer Membrane
α-tocopherol	24	67	81
Phylloquinone	6	0.7	0.4
Plastoquinone-9	70	32	18

[a] Results expressed as percentage of total prenylquinones.

phylloquinone (Soll *et al.*, 1985). Although envelope membranes contain qualitatively the same prenylquinones as thylakoids, these compounds are present in different amounts. For instance, in contrast to thylakoids, the major prenylquinone in envelope membranes is α-tocopherol, whereas it is plastoquinone-9 in thylakoids (Soll *et al.*, 1985): The α-tocopherol : plastoquinone-9 ratio is therefore much higher in the envelope fraction (about 2.3) than in the thylakoids (about 0.3).

Unfortunately, although the participation of prenylquinones in electron transfer is well established for thylakoids, nothing is known of their function in envelope membranes. α-tocopherol (vitamin E) is regarded as the most important lipid-soluble antioxidant (Machlin, 1984); therefore, it possibly plays such a role in envelope membranes. The exact function of envelope plastoquinone-9 is obscure; however, the importance of this prenylquinone for electron and proton transfer suggests that it could play a role in fatty acid and carotenoid desaturation.

IV. ENVELOPE MEMBRANES AS A SITE OF MEMBRANE BIOGENESIS

Most of the compounds found in envelope membranes are also present in thylakoids (galactolipids, sulfolipid, pigments, etc.). In addition, we saw above that, except for proteins, all envelope membranes share the same chemical components present in similar proportions. Therefore, it was reasonable to suggest that the envelope membranes, which is the only membrane system to be present in all plastids, could play a role in plastid biogenesis (Douce and Joyard, 1979).

A. Glycerolipid Biosynthesis

Glycerolipid biosynthesis requires the assembly of three parts: fatty acids, glycerol, and a polar head group. Twenty years ago, the endoplasmic reticulum was considered the only site for glycerolipid synthesis within the plant cell. Fortunately, this view is no longer held, because it has been clearly demonstrated that plastids contain a complete set of enzymes for glycerolipid biosynthesis (Table IV). First, Douce and Guillot-Salomon (1970) demonstrated that chloroplasts and etioplasts can incorporate *sn*-glycerol-3-phosphate into lysophosphatidic acid, phosphatidic acid, and diacylglycerol (the enzymes of the Kornberg-Pricer pathway) and then into MGDG (after addition of UDP-galactose [UDP-gal]). Then, Douce (1974) demonstrated that the envelope membranes were the site of MGDG synthesis. The pivotal role of envelope membranes in glycerolipid formation was extended when Joyard and Douce (1977) demonstrated that all the enzymes of the chloroplast Kornberg-Pricer pathway are associated with or localized on envelope membranes. Mudd and his coworkers (e.g., Mudd *et al.*, 1987) demonstrated that PG, the major plastid phospholipid, is also synthesized on the inner envelope membrane. Finally, sulfolipid is also synthesized in envelope membranes (Heinz *et al.*, 1989).

The first enzyme of the Kornberg-Pricer pathway, the acyl-ACP : *sn*-glycerol-3-phosphate acyltransferase (EC 2.3.1.?), is a "soluble" enzyme that produces lysophosphatidic acid (lyso-PA) according to the following reaction:

sn-glycerol-3-phosphate + acyl-ACP →
1-acyl-*sn*-glycerol-3-phosphate + ACP

This enzyme, which is recovered in the stroma after chloroplast fractionation, is closely associated with the inner envelope membane, and lyso-PA is released directly into the membrane (Joyard and Douce, 1977). This enzyme is specific for acyl-ACP thioesters (Frentzen *et al.*, 1983). Regardless of the thioester used, a striking specificity for the *sn*-1 position of the glycerol is observed (Joyard *et al.*, 1979; Frentzen *et al.*, 1983). Furthermore, when a mixture of 16:0 and 18:1 thioesters is offered, 18:1 is preferentially used (Frentzen *et al.*, 1983). In nongreen plastids from cauliflower buds and amyloplasts from sycamore cells, this enzyme was also soluble and showed the same specificities and selectivities as those in mature chloroplasts (Alban *et al.*, 1989). Frentzen *et al.* (1983) demonstrated that the selectivity of acylation for the *sn*-1 position by 18:1 or 16:0 strongly depends on the *sn*-glycerol-3-phosphate concentration. For example, incorporation of 18:1 into lyso-PA was mark-

TABLE IV
Enzymatic Activities and Translocators Associated with the Plastid Envelope Membranes

Enzymatic Activities	Localization
Enzymes involved in lipid metabolism	
Acyl-CoA synthase	OM
Acyl-CoA thioesterase	IM
Acyl-ACP : *sn*-gylcerol-3-phosphate acyltransferase	stroma + IM
Acyl-ACP : 1-acyl-*sn*-glycerol-3-phosphate acyltransferase	OM + IM
Phosphatidate phosphatase	IM
Monogalactosyldiacylglycerol (MGDG) synthase	IM
Sulfolipid synthase	IM(?)
MGDG acyltransferase	OM
Digalactosyldiacylglycerol acyltransferase	OM
Galactolipid : galactolipid galactosyltransferase	OM
Phosphatidate cytidyltransferase	IM
CDP-diacylglycerol : *sn*-glycerol-3-phosphate phosphatidyltransferase	IM
Phosphatidylglycerophosphatase	IM
Steryl glucoside acyltransferase	OM(?)
Enzymes involved in pigment, prenylquinones, and flavonoid biosynthesis	
Synthesis of geranylgeraniol derivatives (i.e., phytyl pyrophosphate, etc.)	stroma + IM
Homogentisate decarboxylase	IM
SAM : methyl-6-geranylgeranylquinol methyltransferase	IM
SAM : methyl-6-phytylquinol methyltransferase	IM
SAM : methyl-6-solanesylquinol methyltransferase	IM
SAM : α-tocopherol methyltransferase	IM
SAM : α-tocotrienol methyltransferase	IM
Zeaxanthine epoxidase	IM(?)
Phytoene synthase	IM(?)
Protochlorophyllide oxidoreductase	OM
SAM : flavonol-O-methyltransferase	IM(?)
Others	
Mg^{2+}-dependent ATPase	IM
Hexokinase	OM
Protein kinase(s)	OM + IM
Pore protein	OM
Protein receptor(s)	OM
Translocators (phosphate–triose-phosphate, sulfate, nucleotides, malate–α-ketoglutarate, malate–glutamate, glycolate–glycerate, glucose, oxaloacetate, etc.)	IM

IM, inner membrane; OM, outer membrane; SAM, S-adenosylmethionine. Putative localization when (?) is added.

edly increased by decreasing the *sn*-glycerol-3-phosphate concentration. The acyl-ACP : *sn*-glycerol-3-phosphate acyltransferase was also shown to be more specific for fatty acids esterified to the ACP II isoform, because it possesses a fivefold lower K_m for oleoyl-ACP II than for oleoy-ACP I (Guerra *et al.*, 1986), thus channeling specific fatty acids linked to ACP II toward the envelope Kornberg-Pricer pathway (for a discussion, see Joyard and Douce, 1987).

Bertrams and Heinz (1981) first provided biochemical data for the purified chloroplast acyl-ACP : *sn*-glycerol-3-phosphate acyltransferase from pea and spinach. According to the plant species analyzed, between one and three isomeric forms of acyltransferase activity were recovered (Bertrams and Heinz, 1981; Nishida *et al.*, 1987; Douady and Dubacq, 1987). For instance, in squash chloroplasts, three isoforms (AT1–AT3) were purified (Nishida *et al.*, 1987). These isoforms are very similar to each other in physicochemical characteristics, such as M_r (about 40,000) and p*I* (5.5–5.6). The sequences of the cDNA corresponding to two isoforms of the squash chloroplast protein were recently determined (Ishizaki *et al.*, 1988). Acyl-ACP : *sn*-glycerol-3-phosphate acyltransferase is synthesized as a precursor protein of 396 amino acids, which is processed to become a mature protein of 368 amino acids, losing a transit peptide of 28 amino acids (Ishizaki *et al.*, 1988).

The acyl-ACP : 1-acyl-*sn*-glycerol-3-phosphate acyltransferase (EC 2.3.1.?) catalyzes the acylation of lyso-PA to form phosphatidic acid according to the following reaction:

$$\text{1-acyl-}sn\text{-glycerol-3-phosphate} + \text{acyl-ACP} \rightarrow \text{1,2-diacyl-}sn\text{-glycerol-3-phosphate} + \text{ACP}$$

This enzyme is firmly bound to the inner envelope membrane of chloroplasts (Joyard and Douce, 1977) and nongreen plastids (Alban *et al.*, 1989). The optimum pH for phosphatidic acid synthesis is alkaline (7.5–8). In 18:3 plants such as pea, this enzyme is probably a rate-limiting step in phosphatidic acid synthesis (Gardiner *et al.*, 1984; Andrews *et al.*, 1985). This is not the case in 16:3 plants such as spinach.

The enzyme is highly specific for palmitic acid (16:0). Because lyso-PA used for this reaction is already esterified at the *sn*-1 position, the enzyme will direct 16:0 fatty acid to the available *sn*-2 position (Joyard *et al.*, 1979; Frentzen *et al.*, 1983). Interestingly, the same results were obtained with nongreen plastids purified from cauliflower buds and with amyloplasts from sycamore cells (Alban *et al.*, 1989).

Therefore, the two plastidial acyltransferases have distinct specificities and selectivities for acylation of *sn*-glycerol-3-phosphate. Together, they led to the formation of phosphatidic acid having 18:1 and 16:0 fatty

acids, respectively, at the *sn*-1 and *sn*-2 positions of the glycerol backbone. This structure is typical for the so-called prokaryotic glycerolipids, which are found in 16:3 plants (see above). In contrast, in 18:3 plants, plastid glycerolipids (MGDG) do not contain C16 fatty acids at the *sn*-2 position and, therefore, probably do not derive directly from the envelope Kornberg-Pricer pathway. The origin of the most abundant molecular species of MGDG and DGDG having 18:3 at both the *sn*-1 and *sn*-2 positions of the glycerol backbone is still under investigation.

Diacylglycerol biosynthesis occurs in the inner envelope membrane owing to a phosphatidate phosphatase (EC 3.1.2.2) (Joyard and Douce, 1977), which catalyzes the following reaction:

1,2-diacyl-*sn*-glycerol-3-phosphate →
1,2-diacyl-*sn*-glycerol + inorganic phosphate (P_i)

This enzyme is membrane-bound, active at alkaline pH (9.0), and is highly sensitive to cations (Joyard and Douce, 1979). In addition, 18:3 plants have a rather low phosphatidate phosphatase activity, in contrast to 16:3 plants (Heinz and Roughan, 1983). Therefore, the level of phosphatidate phosphatase activity in envelope membranes may be responsible for the difference observed between 18:3 and 16:3 plants with respect to MGDG synthesis. Alban *et al.* (1989) recently demonstrated that this was also true for nongreen plastids. Plastid envelope membranes isolated from both 16:3 and 18:3 plants have the same capacity to form MGDG when supplied with diacylglycerol and UDP-gal. However, very little MGDG is formed from *sn*-glycerol-3-phosphate in plastids (or plastid envelope membranes) from 18:3 plants, although the biosynthesis of phosphatidic acid is very active, thus suggesting that the limiting step is indeed the formation of diacylglycerol (Heinz and Roughan, 1983).

The inner envelope membrane of spinach chloroplasts is characterized by the presence of a UDP-gal–diacylglycerol galactosyltransferase (EC 2.4.1.46) (or MGDG synthase), which transfers a galactose from a water-soluble donor, UDP-gal, to the hydrophobic acceptor molecule diacylglycerol for synthesizing MGDG (Douce, 1974):

UDP gal + 1,2-diacyl-*sn*-glycerol →
1,2-diacyl-3-O-β-D-galactopyranosyl-*sn*-glycerol + UDP

This enzyme is present in all plastids (chloroplasts and nongreen plastids) analyzed so far (Douce *et al.*, 1984). UDP and *N*-ethylmaleimide are inhibitors of the MGDG synthase (Heemskerk and Wintermans, 1987; Covès *et al.*, 1986, 1988a). Covès *et al.* (1986) achieved a partial purification of this enzyme from spinach chloroplast envelope membranes. (3 - [(3 - cholamidopropyl) - dimethylammonia] - 1 - propane sulfonate) (CHAPS) was used to solubilize envelope membranes and was rather

efficient for MGDG synthase. After solubilization, hydroxyapatite chromatography was used to fractionate envelope proteins, and a fraction containing most of the MGDG synthase activity, but <5% of the envelope proteins, was purified (Covès *et al.*, 1986). Because this fraction was almost entirely devoid of lipids, Covès *et al.* (1988a) were able to demonstrate a strong lipid requirement for the MGDG synthase. For example, addition of total envelope lipids stimulated the MGDG synthase activity up to 2 μmol galactose incorporated into MGDG/hr/mg protein, thus corresponding to a 25–30-fold increase in the specific activity, compared with the value in the solubilized envelope. Acidic glycerolipids, and especially PG, were shown to be the best activators of the enzyme: In the presence of saturating amounts of PG, the maximum activity obtained was as high as 10 μmol galactose incorporated into MGDG/hr/mg protein (Covès *et al.*, 1988a). The preparation of the delipidated MGDG synthase-enriched fraction, together with the development of optimal assay conditions, allowed the determination of kinetic parameters for both UDP-gal and the hydrophobic substrate diacylglycerol: The binding of one substrate to the enzyme was shown to be independent of the binding of the other (Covès *et al.*, 1988a). Further purification of MGDG synthase, up to several thousandfolds, has now been achieved, using cholate and CHAPS (Block *et al.*, 1991).

The biosynthesis of DGDG is still a problem. Van Besouw and Wintermans (1978) and Heemskerk and Wintermans (1987) have proposed that this galactolipid could be synthesized by the galactolipid–galactolipid galactosyltransferase (see reaction p. 227). In fact, this enzyme is the only envelope enzyme characterized so far that can form DGDG (Heemskerk and Wintermans, 1987). However, its position, on the outer surface of the outer envelope membrane (Dorne *et al.*, 1982), raises several questions. For instance, nothing is known about the transfer of DGDG from the cytosolic side of the outer envelope membrane to the other plastid membranes. Furthermore, because isolated intact plastids contain little diacylglycerol, how and where diacylglycerol produced by the galactolipid–galactolipid galactosyltransferase is further metabolized is not known.

B. Formation of Other Glycerolipids

Kleppinger-Sparace *et al.* (1985) demonstrated that intact chloroplasts can incorporate SO_4^{2-} into sulfolipid. Joyard *et al.* (1986) demonstrated that sulfolipid was synthesized in isolated intact chloroplasts from acetate and *sn*-glycerol-3-phosphate. These authors observed a competition

between MGDG and sulfolipid synthesis for diacylglycerol, which suggests that envelope membranes are involved in sulfolipid biosynthesis (Joyard *et al.*, 1986). After extensive studies to identify the natural sulfoquinovosyl donor (Shibuya *et al.*, 1963; Hoppe and Schwenn, 1981; Kleppinger-Sparace *et al.*, 1985; Saidha and Schiff, 1989; Gupta and Sastry, 1988), different nucleoside 5′-diphospho-sulfoquinovoses were recently synthesized by Heinz *et al.* (1989). They also demonstrated that both UDP- and GDP-sulfoquinovoses significantly increased sulfolipid synthesis by spinach chloroplasts and by isolated envelope membranes, UDP-sulfoquinovose being twice as active as the GDP derivative. Therefore, sulfolipid synthesis in envelope membranes is probably due to a UDP-sulfoquinovose–diacylglycerol sulfoquinovosyltransferase, which catalyzes the following reaction:

1,2-diacyl-*sn*-glycerol + UDP-sulfoquinovose →
1′,2′-di-*O*-acyl-3′-*O*-(6-deoxy-6-sulfo-α-D-glucopyranosyl)-*sn*-glycerol + UDP

As presented above, chloroplast PG is unique because it contains a $16{:}1_{trans}$ fatty acid at the *sn*-2 position (Mudd *et al.*, 1987). PG synthesis in isolated chloroplasts has been unsuccessful for a long time because the optimal experimental conditions for assay in the microsomal fraction were actually inhibitory for chloroplast PG synthesis. Mudd and de Zacks (1981) first demonstrated the ability of intact chloroplasts to synthesize PG. Starting from phosphatidic acid, all the enzymes involved in PG formation are localized in the inner envelope membrane (Andrews and Mudd, 1985). PG synthesis requires the coordinated functioning of the following three enzymes:

1. Phosphatidate cytidyltransferase (EC 2.7.7.41)

phosphatidic acid + CTP →
CDP-diacylglycerol + inorganic pyrophosphate

2. CDP-diacylglycerol : *sn*-glycerol-3-phosphate 3-phosphatidyltransferase (EC 2.7.8.5)

CDP-diacylglycerol + *sn*-glycerol-3-phosphate →
3-*sn*-phosphatidyl-1′-*sn*-glycerol-3-phosphate + CMP

3. Phosphatidylglycerolphosphatase (EC 3.1.3.27)

3-*sn*-phosphatidyl-1′-*sn*-glycerol-3-phosphate →
PG + P_i

As discussed by Mudd *et al.* (1987), the complete synthesis of PG can be demonstrated in envelope membranes by the judicious use of specific substrates and inhibitors: If only 16:0-ACP and lyso-PA are added to

envelope membranes, only phosphatidic acid is synthesized; then, CDP-diacylglycerol can be synthesized if CTP and Mg^{2+} are added to the incubation mixture; the subsequent addition of *sn*-glycerol-3-phosphate allows the formation of PG, or phosphatidylglycerophosphate (if mercury ions are added to inhibit phosphatidylglycerophosphatase activity). Additional proof for the demonstration of PG formation from phosphatidic acid synthesized via the envelope Kornberg-Pricer pathway was provided by analysis of the nature and the position of the fatty acids in PG (Sparace and Mudd, 1982).

C. Prenylquinone Biosynthesis

Plastid prenylquinones, especially α-tocopherol and plastoquinone-9, are essential compounds for photosynthesis, mostly as antioxidants or electron carriers. Chloroplasts can synthesize their own prenylquinones (Schulze-Siebert and Schultz, 1987), and the inner envelope membranes play a major role in this synthesis (Soll *et al.*, 1985).

The synthesis of α-tocopherol and plastoquinone-9 in envelope membranes (Fig. 4; Table IV) involves a series of membrane-bound enzymes (Soll, 1987). Tocopherols are synthesized from homogentisic acid and a C_{20}-prenyl unit to form 2-methyl-6-prenylquinol. By a series of methylations and cyclization, the 2-methyl-6-prenylquinol gives rise successively to 2,3-dimethyl-6-prenylquinol, γ-tocopherol (or γ-tocotrienol), and finally α-tocopherol (α-tocotrienol). Plastoquinone-9 is synthesized from homogentisic acid and solanesyl-pyrophosphate to form 2-methyl-6-solanesylquinol, which is methylated and oxidized to form successively plastoquinol-9 and plastoquinone-9. All the enzymes involved (Fig. 4) have been localized in the inner envelope membrane of spinach chloroplasts (Soll *et al.*, 1985) and pea (Soll, 1987). In pepper (*Capsicum annuum*) chromoplasts, tocopherol synthesis also occurs according to the same pathway on the limiting membranes (Camara, 1985). From these chloroplasts, enzymes involved in tocopherol synthesis were solubilized and purified (Camara, 1985).

D. Chlorophyll Synthesis

As discussed above, Pineau *et al.* (1986) demonstrated that purified envelope membranes from mature spinach chloroplasts contain small but consistent pools of protochlorophyllide and chlorophyllide. Incubation of envelope membrane under weak light in the presence of NADPH

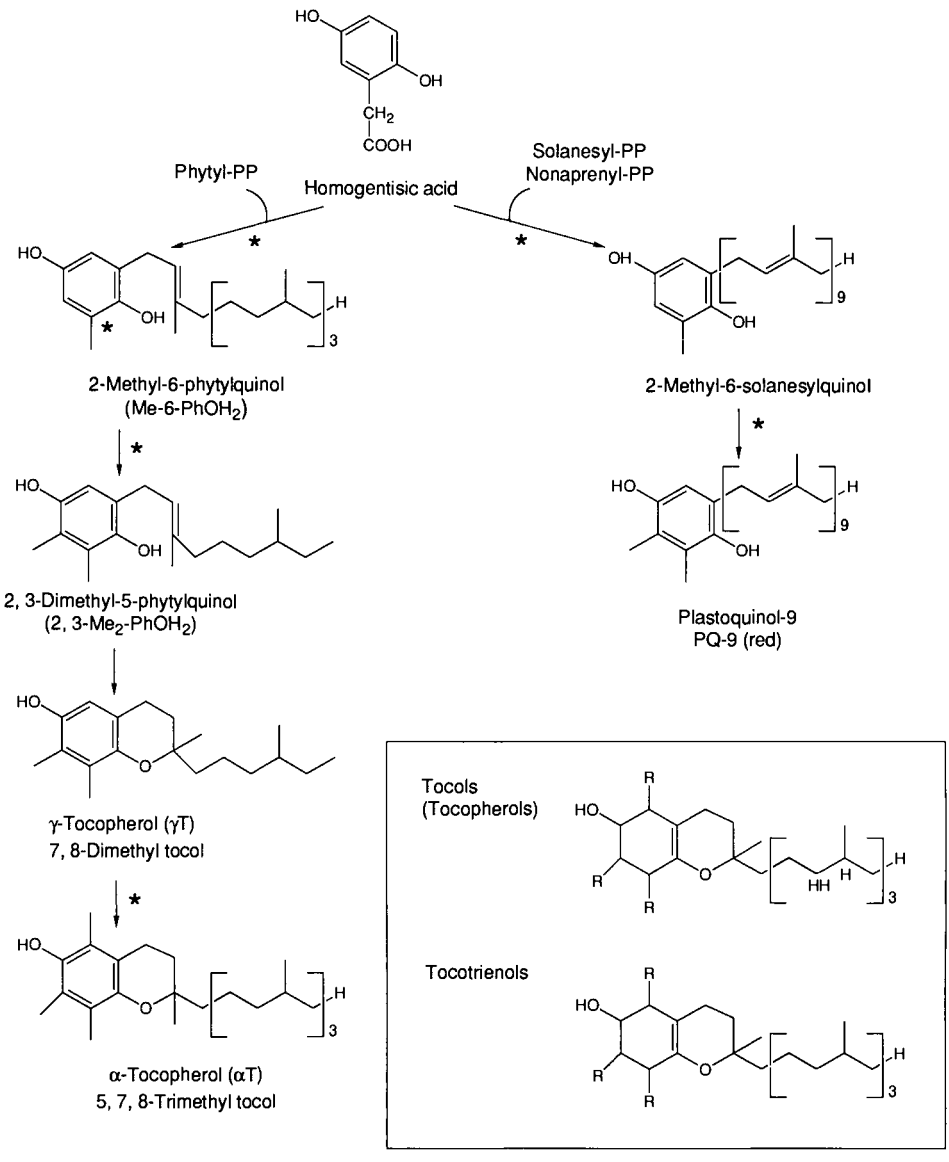


FIG. 4. α-tocopherol and plastoquinone-9 biosynthesis on spinach chloroplast envelope. Reactions marked by a star are localized on the inner envelope membrane. (Adapted with permission from Soll, 1987.)

induced a progressive decrease in the level of fluorescence at 636 nm (attributed to protochlorophyllide), together with a parallel increase in fluorescence level at 680 nm (attributed to chlorophyllide) (Pineau *et al.*, 1986). This reaction was light-dependent and the substitution of NADH and NADPH was ineffective, thus suggesting a possible phototransformation of protochlorophyllide to chlorophyllide due to a protochlorophyllide oxidoreductase (Pineau *et al.*, 1986). Using specific antibodies raised against oat protochlorophyllide oxidoreductase, we have demonstrated that, in mature spinach chloroplasts, this enzyme was only present in envelope membranes and not in the stroma or thylakoids (Joyard *et al.*, 1990). The envelope polypeptide identified using the antibody has a M_r of 37,000, a value very close to protochlorophyllide oxidoreductase from oat etioplasts. In contrast, the etioplast enzyme is mostly concentrated in the prolamellar body and is barely detectable in envelope membranes (Shaw *et al.*, 1985).

V. ENVELOPE MEMBRANES AND THE TRANSPORT OF METABOLITES

Studies with isolated intact chloroplasts have demonstrated the different permeability properties of each envelope membrane. Only the most important points will be summarized here. The reader is referred to the excellent reviews or books recently published on this topic (Heber and Heldt, 1981; Edwards and Walker, 1983; Heldt and Flügge, 1987; Douce and Joyard, 1990).

Small neutral molecules such as O_2, CO_2, or H_2O, but not larger molecules such as sucrose, readily pass through envelope membranes, probably by simple diffusion mechanisms. The same is true for most small anions, which cross the envelope membranes as neutral protonated anions or most amino acids (however, dicarboxylate translocators catalyze rapid transfer of glutamate, aspartate, glutamine, and asparagine).

The outer envelope membrane is freely permeable to small molecules, up to a molecular weight of about 10,000 due to a pore-forming protein (Flügge and Benz, 1984). Therefore, the inner envelope membrane is the true functional border between the plastid stroma and the cytosol. Specific translocators localized in the inner envelope membrane (Table IV) regulate the flow of metabolites necessary for the major plant metabolic pathways such as photosynthesis and photorespiration (Heber and Heldt, 1981). For instance, photorespiration in C_3 plants involves the participation of three plant cell organelles (chloroplasts, peroxisomes,

and mitochondria); therefore, the transport of the different photorespiratory metabolites (glycolate, glycerate, 2-oxoglutarate, glutamate, glutamine, etc.) across the inner envelope membrane is essential for the regulation of the pathway. Nonetheless, not all plastids are alike in their transport functions, and distinctions based on species, tissues, developmental stage, and environment are frequently evident. Unfortunately, critical comparisons are not often made.

A. The Outer Envelope Membrane Contains a Pore-Forming Protein

In contrast to the chloroplast inner envelope membrane, the outer membrane is freely permeable to small molecules, both charged and uncharged. In fact, the transfer of all these small molecules does not occur directly through the lipid bilayer but, rather, is due to a specific membrane structure that allows the passage of water and of a selected set of molecules through the membrane. Using silicone oil filtering centrifugation techniques, Flügge and Benz (1984) examined the penetration of some radioactive substances of various sizes, ranging from M_r 182 (sorbitol) to M_r 66,000 (bovine serum albumine), into the intermembrane space of the chloroplast envelope. The cutoff curve obtained demonstrates an exclusion limit for penetration of between 7,000 and 13,000 daltons, probably around 9,000–10,000 (Flügge and Benz, 1984), a value much higher than that determined for mitochondria or gram-negative bacteria. The size of the pore for the chloroplast outer envelope membrane was around 3 nm. The activity was successfully reconstituted in artificial lipid bilayers, and the strong voltage dependence of the chloroplast pore protein suggests regulation similar to that demonstrated for mitochondria (Freitag *et al.*, 1982). The pore has a single-channel conductance of 0.72 nsec (at 0.1 *M* KCl in the aqueous phase), which is a linear function of the concentration of the salt solution, and demonstrates the absence of any specific site for ions (Flügge and Benz, 1984). Unfortunately, the identity of the outer envelope polypeptide(s) involved in the formation of the pore protein is still unknown.

B. The Phosphate–Triose-Phosphate Translocator

This translocator is the best-characterized envelope protein and one of the major envelope membrane proteins (15–20% of the total envelope protein). The phosphate translocator is located on the inner envelope membrane and catalyzes a strict counter-exchange of P_i (HPO_4^{2-}), triose phosphate, and 3-phosphoglycerate (3-PGA) (Heldt and Rapley, 1970).

This transport system, at the border between the stroma and the cytosol, is a pivotal system for regulation of photosynthesis (Heber and Heldt, 1981; Flügge and Heldt, 1986; Heldt and Flügge, 1987).

As mentioned above, the phosphate–triose-phosphate translocator operates in a strict counter-exchange mode: For each molecule entering the chloroplast, one is exported. The translocator is specific for P_i and triose-phosphate (transported as divalent anions), with K_m values ranging from 0.1 to 0.4 m*M* (Fliege *et al.*, 1978). During photosynthesis, the phosphate–triose-phosphate translocator mostly exchanges P_i for dihydroxyacetone phosphate to maintain stromal P_i concentration at a constant level (Heber and Heldt, 1981). The phosphate–triose-phosphate translocator is essential to the control of the carbon fluxes to the cytosol: The level of exchangeable P_i in the cytosol promoting either sucrose synthesis in the cytosol or starch synthesis within the plastid.

Because the substrates transported by the phosphate translocator are bound as divalent anions, positively charged groups such as lysine and arginine residues are probably present at the binding site. Lysine- and arginine-specific reagents such as pyridoxal phosphate and 2,4,6-trinitrobenzene sulfonate are strong inhibitors of P_i transport and were found suitable for identification of the translocator because, for instance, the reversible binding of pyridoxal phosphate to the lysine residue can be rendered irreversible by reduction of the Schiff base formed by $NaBH_4$. Using this procedure, Flügge and Heldt (1976) identified with SDS-PAGE a 29,000–30,000-daltons polypeptide (E30) as the phosphate–triose-phosphate translocator; further identification was provided using sulfhydryl reagents (*N*-ethylmaleimide or *p*-chloromercuribenzene sulfonate). More specific labeling of the phosphate–triose-phosphate translocator can be achieved, using 1,2-ditritio-1,2-[2,2′-disulfo-4,4′-diisothiocyano]-diphenylethane ($[^3H]_2$-DIDS), a powerful inhibitor of 3-phosphoglycerate-dependent O_2-evolution by isolated intact chloroplasts (Rumpho *et al.*, 1988). DIDS first binds reversibly to the protein, and then a covalent reaction of the isothiocyano group of DIDS with an ε-amino group of lysine takes place, which results in a dramatic conformational change of the protein. All the different methods used labeled the same 29,000–30,000-daltons polypeptide (E30) in chloroplast envelope membranes from C_3 plants. Flügge and Heldt (1981) isolated the protein involved in phosphate transport and determined the hydrodynamic properties of translocator–triton X-100 micelles (Flügge and Heldt, 1987). From these data, it was concluded that the phosphate–triose-phosphate translocator exists in its functional state as a dimer (M_r 61,000) of two identical subunits (M_r 29,000). Activity of the purified phosphate translocator has been reconstituted in liposomes: Uptake of radioactively labeled substrates ($[^{32}P]$phosphate or 3-$[^{14}C]$-PGA) into the

liposomes was only observed when the liposomes were preloaded with P_i or 3-PGA but not when preloaded with KCl or 2-PGA, thus demonstrating that reconstituted transport involves an obligatory and specific counter-exchange of anions (Flügge and Heldt, 1986), as previously demonstrated with isolated intact chloroplasts.

The phosphate–triose-phosphate translocator is present in almost all envelope membranes analyzed so far. For instance, Thompson *et al.* (1987) and Rumpho *et al.* (1988) have identified an envelope polypeptide from C_4 mesophyll chloroplasts corresponding to the phosphate–triose-phosphate translocator. In nongreen plastids, the existence of such a translocator was not obvious. However, Western blotting experiments with antibodies to spinach E30, using purified envelope membranes from cauliflower proplastids (Alban *et al.*, 1988) clearly demonstrate that the major 28,000-dalton polypeptide of the cauliflower proplastid envelope fraction and the phosphate–triose-phosphate translocator from spinach chloroplasts have closely related antigenic sites. In envelope membranes from sycamore amyloplasts, a 31,000-dalton polypeptide has been identified using the same procedure (Alban *et al.*, 1988; Ngernprasirtsiri *et al.*, 1988).

As with almost all envelope proteins, the phosphate translocator is coded for by nuclear DNA and synthesized on cytoribosomes. Flügge and Wessel (1984) have determined an apparent M_r of 40,000 for the precursor protein. Then, during posttranslational import into chloroplasts, the precursor is processed to its mature size. Using oligonucleotide probes modelled from the amino acid sequence of tryptic digests prepared from isolated translocator, Flügge *et al.* (1989) selected cDNA clones coding for the entire 404 amino acid residues of the precursor (corresponding to a M_r of 44,234). The mature part of the translocator contains at least seven hydrophobic regions (and, therefore, probably membrane-spanning regions) and a highly polar domain at the C-terminus (Flügge *et al.*, 1989).

C. Dicarboxylate Translocators

Chloroplasts from C_3 and C_4 plants contain specific translocators for L-malate, L-aspartate, L-glutamate, and 2-oxoglutarate in the inner envelope membrane (Heber and Heldt, 1981; Heldt and Flügge, 1987). Dicarboxylate transport involves a counter-exchange process. Initially, dicarboxylate transport into chloroplasts was attributed to a single translocator, but further studies have demonstrated that at least two distinct translocators might be involved (Somerville and Ogren, 1983; Werner-Washburne and Keegstra, 1985; Woo *et al.*, 1987; Flügge *et al.*,

1988): For instance, the functioning of two translocators involved in 2-oxoglutarate–malate and glutamate–malate exchanges results in a net 2-oxoglutarate–glutamate exchange with no net malate transport (Woo *et al.*, 1987). In other words, 2-oxoglutarate is transported inward on a 2-oxoglutarate translocator and glutamate outward on a dicarboxylate translocator with malate as the counter-ion on both translocators in a cascadelike manner. Further work by Yu and Woo (1988) demonstrated that a third dicarboxylate translocator was present in envelope membranes from oat and spinach: a glutamine translocator, which was unable to transport dicarboxylates other than L-glutamate. The three dicarboxylate translocators present in the chloroplast inner envelope membrane from C_3 plants are essential for photorespiratory NH_3 assimilation in the chloroplast.

Chloroplast stroma contains an NADP-malate dehydrogenase, which is involved in the photoreduction of oxaloacetate to malate. In C_3 plants, this process may be part of the malate–oxaloacetate shuttle involved in the transfer of reducing equivalents across the chloroplast envelope. In C_4 plants such as maize (NADP-malic enzyme type), reduction of oxaloacetate is critical for carbon assimilation. Oxaloacetate transport across envelope membranes involves a specific translocator, demonstrated in envelope membranes from maize and spinach chloroplasts by Hatch *et al.* (1984). At present, the details of oxaloacetate transport in plastids remain a mystery, and more work is needed to confirm the mechanism. Malate efflux and oxaloacetate influx probably occur on separate carriers because the uptake of oxaloacetate was not inhibited by a large excess of malate (Hatch *et al.*, 1984).

The dicarboxylate translocators have not yet been identified unambiguously. However, Somerville and Somerville (1985) have analyzed the polypeptide pattern of chloroplast envelope membranes from a mutant of *Arabidopsis* deficient in dicarboxylate transport: A major envelope polypeptide of M_r 42,000 was present at a very low level in the mutant, compared with the wild type. Therefore, it is possible that this polypeptide could be a component of the dicarboxylate translocator, but the definitive proof for unambiguous identification remains to be given (Somerville and Somerville, 1985).

D. The Glycerate–Glycolate Translocator

During photorespiration, glycolate is formed within chloroplasts by a two-step reaction involving oxygenation of ribulose-1,5-bisphosphate and dephosphorylation of phosphoglycolate. Glycolate is then excreted from the chloroplast to be oxidized in the peroxisomes, and glycerate is

taken up into the stroma to be phosphorylated and recycled within the Benson, Bassham, and Calvin cycle (Tolbert, 1980). Howitz and McCarty (1983) provided evidence for the existence of a glycolate translocator localized in the inner envelope of chloroplasts (however, glycolate uptake was measured, whereas glycolate is excreted from chloroplasts under physiological conditions). Evidence for a glycerate translocator was provided by Robinson (1982). In fact, glycolate possibly can exchange directly with glycerate and hydroxyl ions. This would allow a flexible stoichiometry for the exchange reaction that would accommodate the 2:1 ratio needed for the reaction by which glycolate is metabolized and results in a stoichiometry of two glycolate molecules leaving the chloroplast stroma in exchange for one glycerate entering (Howitz and McCarty, 1985a,b). Interestingly, the maximum transport rates of the glycerate–glycolate translocator are high enough to sustain the photorespiratory carbon fluxes (Howitz and McCarty, 1985a,b).

E. Other Translocators

ATP and ADP are transported into chloroplasts by counter-exchange, mediated by a translocator that is highly specific for ATP import (ADP import is about 10 times lower). This is in contrast to the situation in well-coupled mitochondria, which exhibit a preferential and rapid uptake of external ADP rather than ATP. In C_4 plants such as *Digitaria sanguinalis*, the ATP–ADP translocator was found to be more active (Huber and Edwards, 1976). In addition, the activity of the ATP–ADP translocator probably depends on the developmental stage of the plastid because ATP is necessary during the early stages of plastid development (Douce and Joyard, 1984). Surprisingly, Ngernprasirtsiri *et al.* (1989) showed that a 32,000-daltons polypeptide in amyloplast inner envelope membrane from sycamore cells cross-reacts with the ATP–ADP translocator from *Neurospora crassa* mitochondria. This polypeptide was not immunologically detectable in chloroplast envelope membranes from green sycamore cells. However, the possibility of cross-contamination of amyloplast envelope preparations by mitochondria cannot be entirely ruled out.

Low rates of carrier-mediated uptake of sulfate into spinach chloroplasts have been demonstrated by Mourioux and Douce (1979). Upon addition of sulfate to the incubation medium, one molecule of stromal P_i is exchanged for one molecule of sulfate by a strict counter-exchange process (Mourioux and Douce, 1979). Sulfate uptake into chloroplasts is essential for synthesis of sulfur-containing amino acids (mostly cysteine) and sulfolipid within chloroplasts.

F. Proton Permeability of Envelope Membranes

In a CO_2-free medium, illuminated intact chloroplasts slowly excrete protons against the existing gradient. Under these conditions, a proton gradient is maintained between the stroma and the external medium, with the stroma being more alkaline than the medium (see Heber and Heldt, 1981). A possible mechanism for exporting H^+ out of chloroplasts could involve an electrogenic pump driven by ATP. Douce et al. (1973) characterized on isolated envelope membranes a Mg^{2+}-ATPase, and they demonstrated that it was distinct from the Mg^{2+}-dependent ATPase activity of the thylakoid coupling factor. This ATPase was further characterized by Joyard and Douce (1975), Maury *et al.* (1981), McCarty *et al.* (1984), McCarty and Selman (1986), and Nguyen and Siegenthaler (1983, 1985). Unfortunately, the purification of the enzyme has led to contradictory results about the M_r of the enzyme and its subunits (Nguyen and Siegenthaler, 1985; McCarty and Selman, 1986). In addition, the involvement of the envelope ATPase in H^+ excretion is still unclear. For instance, an alkalinization of the stroma is observed when isolated intact chloroplasts are incubated in the dark in the presence of ATP (Champigny and Joyard, 1978; Robinson, 1985), which suggests the operation of an ATPase proton pump on the envelope. Huber and Maury (1980) and Maury *et al.* (1981) have proposed that this was indeed the case: The envelope ATPase being involved in an ATP driven, Mg^{2+}-dependent, reversible H^+/K^+ exchange across the envelope membrane. In contrast, Robinson (1985) concluded that the pH gradient was unlikely to be maintained by an electrogenic proton pump driven by ATP. Therefore, the exact function of the envelope ATPase remains to be demonstrated; for instance, a role in protein transport across the envelope has been proposed (for reviews, see Schmidt and Mishkind, 1986; Mishkind and Scioli, 1988; Keegstra *et al.*, 1989; Douce and Joyard, 1990). In addition, alternative mechanisms for proton efflux from chloroplasts need to be explored.

VI. CONCLUDING REMARKS

Although we have learned a great deal about plastid envelope membranes and their importance in plant cell metabolism over the past 15 years, many important questions remain unanswered.

Among the numerous envelope polypeptides separated by two-dimensional electrophoresis, only a few of them have been unambiguously identified: The large and small subunits of ribulose-1,5-bisphos-

phate carboxylase (Pineau and Douce, 1974; Joyard *et al.*, 1982); the phosphate–triose-phosphate translocator (Flügge and Heldt, 1976) and the protochlorophylide oxidoreductase (Joyard *et al.*, 1990). In contrast, characterization of the envelope ATPase (Nguyen and Siegenthaler, 1985; McCarty and Selman, 1986) and of the putative protein receptor of the outer envelope membrane (Cornwell and Keegstra, 1987; Pain *et al.*, 1988) have led to somewhat contradictory results. The major limitation in identification of envelope proteins is the extreme difficulty encountered in purifying them. Envelope membranes are only a minor membrane structure within the plant cell: They do not represent >1–2% of the total chloroplast protein. In addition, only a few polypeptides are major envelope components, most of them (e.g., the MGDG synthase) are only present in very limited amounts. Finally, envelope membranes have one of the highest lipid : protein ratios; therefore, their complete solubilization requires large amounts of detergents that might be deleterious to enzymatic activities.

Numerous functions of plastid envelope membranes are still unexplored. For instance, the identity and functions of the majority of the envelope protein kinases and phosphoproteins remain unknown. Obviously, the pleiotropic responses of protein kinase activation in the plant cell requires identification of the target substrates present in the envelope membranes. Some progress has been made in this respect, but much more remains to be done. Another major function of envelope membranes is still poorly understood: Fatty acids, which are synthesized only in plastids, should be exported to the cytosol. The postulated mechanisms could involve two envelope proteins: an acyl-CoA thioesterase and an acyl-CoA synthetase (Joyard and Douce, 1977; Joyard and Stumpf, 1980, 1981), which are localized on the inner and the outer envelope membranes, respectively (Block *et al.*, 1983c; Andrews and Keegstra, 1983). In fact, neither the mechanisms involved nor the nature of the fatty acid thioesters exported are established, despite the major importance of fatty acid availability in the cytosol for phospholipid and triacylglycerol biosynthesis.

The assumption that envelope membranes are very similar in all plastids is based on analyses of the chemical composition (glycerolipids, pigments, prenylquinones, etc.). However, differentiation of all plastid types and their developmental transitions are associated with, or dependent on, marked changes in the specific enzymatic activities found in the plastid envelope membranes. Tissue-specific, plant-specific, or light-regulated expression of genes encoding envelope proteins are major events in plastid biogenesis, but the identification of specific differences between envelope membranes from different plastid types is still in its infancy.

ACKNOWLEDGMENTS

The work described in this review is largely a summary of the work performed in our laboratory during the last few years. We wish to thank deeply Dr. M. A. Block, Dr. A. J. Dorne, Dr. C. Alban, and Dr. J. Covès for their involvement in this research. Contínuous financial support by the CNRS to URA 576 and by the Commissariat à l'Energie Atomique (CEA) to our laboratory is also gratefully acknowledged.

REFERENCES

Alban, C., Joyard, J., and Douce, R. (1988). Preparation and characterization of envelope membranes from non-green plastids. *Plant Physiol.* **88,** 709–717.

Alban, C., Joyard, J., and Douce, R. (1989). Comparison of glycerolipid biosynthesis in non-green plastids from sycamore (*Acer pseudoplatanus*) cells and cauliflower (*Brassica oleracea*) buds. *Biochem. J.* **259,** 775–783.

Andrews, J., and Keegstra, K. (1983). Acyl-CoA synthetase is located in the outer membrane and acyl-CoA thioesterase in the inner membrane of pea chloroplast envelopes. *Plant Physiol.* **72,** 735–740.

Andrews, J., and Mudd, J. B. (1985). Phosphatidylglycerol synthesis in pea chloroplasts. Pathway and localization. *Plant Physiol.* **79,** 259–265.

Andrews, J., Ohlrogge, J. B., and Keegstra, K. (1985). Final step of phosphatidic acid synthesis in pea chloroplasts occurs in the inner envelope membrane. *Plant Physiol.* **78,** 459–465.

Benson, A. A., Daniel, H., and Wiser, R. (1959). A sulfolipid in plants. *Proc. Natl. Acad. Sci. USA* **45,** 1582–1587.

Bertrams, M., and Heinz, E. (1981). Experiments on enzymatic acylation of sn-glycerol 3-phosphate acyltransferases from chloroplasts. *Plant Physiol.* **68,** 653–657.

Billecocq, A. (1974). Structure des membranes biologiques: Localisation des galactosyldiglycérides dans les chloroplasts au moyen des anticorps spécifiques. II—Etude en microscopie électronique à l'aide d'un marquage à la peroxydase. *Biochim. Biophys. Acta* **352,** 245–251.

Billecocq, A. (1975). Structure des membranes biologiques: Localisation du sulfoquinovosyldiglycéride dans les diverses membranes des chloroplastes au moyen des anticorps spécifiques. *Ann. Immunol. (Inst. Pasteur)* **126C,** 337–352.

Billecocq, A., Douce, R., and Faure, M. (1972). Structure des membranes biologiques: Localisation des galactosyldiglycérides dans les chloroplastes au moyen des anticorps spécifiques. *C.R. Acad. Sci.* **275,** 1135–1137.

Block, M. A., Dorne, A.-J., Joyard, J., and Douce, R. (1983a). Preparation and characterization of membrane fractions enriched in outer and inner envelope membranes from spinach chloroplasts. I—Electrophoretic and immunochemical analyses. *J. Biol. Chem.* **258,** 13273–13280.

Block, M. A., Dorne, A.-J., Joyard, J., and Douce, R. (1983b). Preparation and characterization of membrane fractions enriched in outer and inner envelope membranes from spinach chloroplasts. II—Biochemical characterization. *J. Biol. Chem.* **258,** 13281–13286.

Block, M. A., Dorne, A.-J., Joyard, J., and Douce, R. (1983c). The acyl-CoA synthetase and

the acyl-CoA thioesterase are located respectively on the outer and on the inner membrane of the chloroplast envelope. *FEBS Lett.* **153,** 377–381.

Block, M. A., Marechal, E., Joyard, J., and Douce, R. (1991). in preparation.

Camara, B. (1985). S-adenosylmethionine : γ-tocopherol methyltransferase (*Capsicum* chloroplasts). *Methods Enzymol.* **111,** 544–553.

Carde, J. P., and Bernard-Dagan, C. (1982). Compartimentation de la synthèse terpénique chez le Pin maritime. *Bull. Soc. Bot. Fr., Actual. Bot.* **129,** 53–70.

Carde, J. P., Joyard, J., and Douce, R. (1982). Electron microscopic studies of envelope membranes from spinach plastids. *Biol. Cell.* **44,** 315—324.

Champigny, M. L., and Joyard, J. (1978). Analyse du mécanisme de stimulation, par l'ATP de la pénétration du carbone inorganique dans les chloroplastes intacts, isolés de l'épinard. *C.R. Acad. Sci. Paris* **286,** 1791–1794.

Chua, N. H., and Schmidt, G. W. (1979). Transport of proteins into chloroplasts and mitochondria. *J. Cell Biol.* **81,** 461–483.

Cline, K., Andrews, J., Mersey, B., Newcomb, E. H., and Keegstra, K. (1981). Separation and characterization of inner and outer envelope membranes of pea chloroplasts. *Proc. Natl. Acad. Sci. USA* **78,** 3595–3599.

Cline, K., Werner-Washburne, M., Andrews, J., and Keegstra, K. (1984). Thermolysin is a suitable protease for probing the surface of intact pea chloroplasts. *Plant Physiol.* **75,** 675–678.

Cline, K., Keegstra, K., and Staehelin, L. A. (1985). Freeze-fracture electron microscopic analysis of ultrarapidly frozen envelope membranes on intact chloroplasts and after purification. *Protoplasma* **125,** 111–123.

Cornwell, K. L., and Keegstra, K. (1987). Evidence that a chloroplast surface protein is associated with a specific binding site for the precursor to the small subunit of ribulose-1,5-bisphosphate carboxylase. *Plant Physiol.* **85,** 780–785.

Covès, J., Block, M. A., Joyard, J., and Douce, R. (1986). Solubilization and partial purification of UDP-galactose : diacylglycerol galactosyltransferase activity from spinach chloroplast envelope. *FEBS Lett.* **208,** 401–406.

Covès, J., Joyard, J., and Douce, R. (1988a). Lipid requirement and kinetic studies of solubilized UDP-galactose : diacylglycerol galactosyltransferase activity from spinach chloroplast envelope membranes. *Proc. Natl. Acad. Sci. USA* **85,** 4966–4970.

Covès, J., Pineau, B., Joyard, J., and Douce, R. (1988b). Solubilization of chloroplast envelope membranes by detergents and its effects on activity of envelope marker enzymes. *Plant Physiol. Biochem.* **26,** 151–163.

Cremers, F. F. M., Voorhout, W. F., Van der Krift, T. P., Leunissen-Bijvelt, J. J. M., and Verkleij, A. J. (1988). Visualization of contact sites between outer and inner envelope membranes in isolated chloroplasts. *Biochim. Biophys. Acta* **933,** 334–340.

Dorne, A. J., Block, M. A., Joyard, J., and Douce, R. (1982). The galactolipid : galactolipid galactosyltransferase is located on the outer membrne of the chloroplast envelope. *FEBS Lett.* **145,** 30–34.

Dorne, A.-J., Joyard, J., Block, M. A., and Douce, R. (1985). Localization of phosphatidylcholine in outer envelope membrane of spinach chloroplasts. *J. Cell Biol.* **100,** 1690–1697.

Dorne, A.-J., Joyard, J., and Douce, R. (1989). Do thylakoids really contain phosphatidylcholine? *Proc. Natl. Acad. Sci. USA* **87,** 71–74.

Douady, D., and Dubacq, J. P. (1987). Purification of acyl-CoA : glycerol-3-phosphate acyltransferase from pea leaves. *Biochim. Biophys. Acta* **921,** 615–619.

Douce, R. (1974). Site of biosynthesis of galactolipids in spinach chloroplasts. *Science* **183,** 852–853.

Douce, R., and Guillot-Salomon, T. (1970). Sur l'incorporation de la radioactivité du *sn-*

glycérol 3-phosphate-^{14}C dans le monogalactosyldiglycéride des plastes isolés. *FEBS Lett.* **11,** 121–126.

Douce, R., and Joyard, J. (1979). Structure and function of the plastid envelope. *Adv. Bot. Res.* **7,** 1–116.

Douce, R., and Joyard, J. (1984). The regulatory role of the plastid envelope during development. "Topics in Photosynthesis," Vol. 5, "Chloroplast Biogenesis" (N. R. Baker and J. Barber, eds.), pp. 71–132. Elsevier, Amsterdam.

Douce, R., and Joyard, J. (1990). Biochemistry and Function of the plastid envelope. *Annu. Rev. Cell. Biol.* **6,** 173–216.

Douce, R., Holtz, R. B., and Benson, A. A. (1973). Isolation and properties of the envelope of spinach chloroplasts. *J. Biol. Chem.* **248,** 7215–7222.

Douce, R., Block, M. A., Dorne, A. J., and Joyard, J. (1984). The plastid envelope membranes: Their structure, composition, and role in chloroplast biogenesis. *In* "Subcellular Biochemistry," Vol. 10 (D. B. Roodyn, ed.), pp. 1–86. Plenum Press, New York.

Douce, R., Alban, C., Bligny, R., Block, M. A., Covès, J., Dorne, A.-J., Journet, E. J., Joyard, J., Neuburger, M., and Rebeillé, F. (1987). Lipid distribution and synthesis within the plant cell. *In* "The Metabolism, Structure, and Function of Plant Lipids" (P. K. Stumpf, J. B. Mudd, and W. D. Nes, eds.), pp. 255–263. Plenum Publishing Corporation, New York.

Edwards, G., and Walker, D. A. (1983). "C_3, C_4: Mechanisms, and Cellular and Environmental Regulation, of Photosynthesis." Backwell Science Publications, Oxford.

Ellens, H., Siegel, D. P., Alford, D., Yeagle, P. L., Boni, L., Lis, L. J., Quinn, P. J., and Bentz, J. (1989). Membrane fusion and inverted phases. *Biochemistry* **28,** 3692–3703.

Fliege, R., Flügge, U. I., Werdan, K., and Heldt, H. W. (1978). Specific transport of inorganic phosphate, 3-phosphoglycerate and triosephosphates across the inner membrane of the envelope in spinach chloroplasts. *Biochim. Biophys. Acta* **502,** 232–247.

Flügge, U. I., and Benz, R. (1984). Pore-forming activity in the outer membrane of the chloroplast envelope. *FEBS Lett.* **68,** 85–89.

Flügge, U. I., and Heldt, H. W. (1976). Identification of a protein involved in phosphate transport of chloroplasts. *FEBS Lett.* **82,** 29–33.

Flügge, U. I., and Heldt, H. W. (1981). The phosphate translocator of the chloroplast envelope. Isolation of the carrier protein and reconstitution of transport. *Biochim. Biophys. Acta* **638,** 296–304.

Flügge, U. I., and Heldt, H. W. (1981). Chloroplast phosphate- triosephophate-phosophoglycerate translocator : its identification, isolation, and reconstitution. *Meth. Enzymol.* **125,** 716–730.

Flügge, U. I., and Wessel, D. (1984). Cell-free synthesis of putative precursors for envelope membrane polypeptides of spinach chloroplasts. *FEBS Lett.* **168,** 255–259.

Flügge, U. I., Woo, K. C., and Heldt, H. W. (1988). Characteristics of 2-oxoglutarate and glutamate transport in spinach chloroplasts. *Planta* **174,** 534–541.

Flügge, U. I., Fischer, K., Gross, A., Sebald, W., Lottspeich, F., and Eckerskorn, C. (1989). The triose phosphate-3-phosphoglycerate-phosphate translocator from spinach chloroplasts: Nucleotide sequence of a full-length cDNA clone and import of the *in vitro* synthesized precursor protein into chloroplasts. *EMBO J.* **8,** 39–46.

Freitag, H., Neupert, W., and Benz, R. (1982). Purification and characterisation of a pore protein of the outer mitochondrial membrane from *Neurospora crassa*. *Eur. J. Biochem.* **123,** 629–636.

Frentzen, M., Heinz, E., McKeon, T., and Stumpf, P. K. (1983). Specificities and selectivities of glycerol-3-phosphate acyltransferase and monoacylglycerol-3-phosphate acyltransferase from pea and spinach chloroplasts. *Eur. J. Biochem.* **129,** 629–639.

Gardiner, S. E., Heinz, E., and Roughan, P. G. (1984). Rates and products of long-chain

fatty acid synthesis from [1−^{14}C]-acetate in chloroplasts isolated from leaves of 16:3 and 18:3 plants. *Plant Physiol.* **74,** 890–896.

Gounaris, K., and Barber, J. (1983). Monogalactosyldiancylglycerol: The most abundant polar lipid in nature. *Trends Biochem. Sci.* **8,** 378–381.

Guerra, D. J., Ohlrogge, J. B., and Frentzen, M. (1986). Activity of acyl carrier protein isoforms from plant fatty acid metabolism. *Plant Physiol.* **82,** 448–453.

Guillermond, L. (1941). "The Cytoplasm of the Plant Cell," Chronica Botanica ed. Waltham, Massachusetts.

Gupta, S. D., and Sastry, P. S. (1988). Development and use of chlorotetracycline fluorescence as a measurement assay of chloroplast envelope-bound Mg^{2+}. *Arch. Biochem. Biophys.* **260,** 125–133.

Hatch, M. D., Dröscher, L., Flügge, U. I., and Heldt, H. W. (1984). A specific translocator for oxaloacetate transport in chloroplasts. *FEBS Lett.* **178,** 15–19.

Haverkate, F., and van Deenen, L. L. M. (1965). Isolation and chemical characterization of phosphatidylglycerol from spinach leaves. *Biochim. Biophys. Acta* **106,** 78–92.

Heber, U., and Heldt, H. W. (1981). The chloroplast envelope: Structure, function, and role in leaf metabolism. *Annu. Rev. Plant Physiol.* **32,** 139–168.

Heemskerk, J. W. M., and Wintermans, J. F. G. M. (1987). Role of the chloroplast in the leaf acyl-lipid synthesis. *Physiologia Plantarum* **70,** 558–568.

Heinz, E. (1977). Enzymatic reactions in galactolipid biosynthesis. *In* "Lipids and Lipid Polymers" (M. Tevini and H. K. Lichtenthaler, eds.), pp. 102–120. Springer, Berlin.

Heinz, E., and Roughan, P. G. (1983). Similarities and differences in lipid metabolism of chloroplasts isolated from 18:3 and 16:3 plants. *Plant Physiol.* **72,** 273–279.

Heinz, E., Schmidt, H., Hoch, M., Jung, K. H., Binder, H., and Schmidt, R. R. (1989). Synthesis of different nucleoside 5′-diphospho-sulfoquinovoses and their use for studies on sulfolipid biosynthesis in chloroplasts. *Eur. J. Biochem.* **184,** 445–453.

Heldt, H. W., and Flügge, U. I. (1987). Subcellular transport of metabolites in plant cells. *In* "The Biochemistry of Plants," Vol. 12, "Physiology of Metabolism" (D. D. Davies, ed.), pp. 50–85. Academic Press, New York.

Heldt, H. W., and Rapley, L. (1970). Specific transport of inorganic phosphate, 3-phosphoglycerate and dihydroxyacetone phosphate and of dicarboxylates across the inner membrane of spinach chloroplasts. *FEBS Lett.* **10,** 143–147.

Hinterstroisser, B., Missbichler, A., Pineau, B., and Peschek, G. A. (1988). Detection of protochlorophyllide in chlorophyll-free plasma membrane preparations from *Anacystis nidulans*. *Biochem. Biophys. Res. Commun.* **154,** 839–846.

Hoppe, W., and Schwenn, J. D. (1981). *In vitro* biosynthesis of the plant sulfolipid: On the origin of the sulfonate group. *Z. Naturforsch.* **36C,** 820–826.

Howitz, K. T., and McCarty, R. E. (1983). Evidence for a glycolate transporter in the envelope of pea chloroplasts. *FEBS Lett.* **154,** 339–342.

Howitz, K. T., and McCarty, R E. (1985a). Kinetic characteristics of the chloroplast envelope glycolate transporter. *Biochemistry* **24,** 2645–2652.

Howitz, K. T., and McCarty, R. E. (1985b). Substrate specificity of the pea chloroplast glycolate transporter. *Biochemistry* **24,** 3645–3650.

Huber, S. C., and Edwards, G. E. (1976). A high-affinity ATP translocator in mesophyll chloroplasts of *Digitaria sanguinalis*, a plant having the C-4 dicarboxylic acid pathway of photosynthesis. *Biochim. Biophys. Acta* **440,** 675–687.

Huber, S. C., and Maury, W. (1980). Effect of magnesium on intact chloroplasts. 1. Evidence for activation of (sodium) potassium/proton exchange across the chloroplast envelope. *Plant Physiol.* **65,** 350–354.

Ishizaki, O., Nishida, I., Agata, K., Eguchi, G., and Murata, N. (1988). Cloning and nu-

cleotide sequence of cDNA for the plastid glycerol-3-phosphate acyltransferase from squash. *FEBS Lett.* **238,** 424–430.

Journet, E. P., and Douce, R. (1985). Enzymic capacities of purified cauliflower bud plastids for lipid synthesis and carbohydrate metabolism. *Plant Physiol.* **79,** 458–467.

Joyard, J., and Douce, R. (1975). Mn^{2+}-dependent ATPase of the envelope of spinach chloroplasts. *FEBS Lett.* **51,** 335–340.

Joyard, J., and Douce, R. (1977). Site of synthesis of phosphatidic acid and diacylglycerol in spinach chloroplasts. *Biochim. Biophys. Acta* **486,** 273–285.

Joyard, J., and Douce, R. (1979). Characterization of phosphatidate phosphohydrolase activity associated with chloroplast envelope membranes. *FEBS Lett.* **102,** 147–150.

Joyard, J., and Douce, R. (1987). Galactolipid biosynthesis. *In* "The Biochemistry of Plants," Vol. 9, "Lipids" (P. K. Stumpf, ed.), pp. 215–274. Academic Press, New York.

Joyard, J., and Stumpf, P. K. (1980). Characterization of an acyl-CoA thioesterase associated with the envelope of spinach chloroplasts. *Plant Physiol.* **65,** 1039–1043.

Joyard, J., and Stumpf, P. K. (1981). Synthesis of long-chain acyl-CoA in chloroplasts envelope membranes. *Plant Physiol.* **67,** 250–256.

Joyard, J., Chuzel, M., and Douce, R. (1979). Is the chloroplast envelope a site of galactolipid synthesis? Yes! *In* "Advances in the Biochemistry and Physiology of Plant Lipids" (L. A. Appelqvist and C. Liljenberg, eds.), pp. 181–186. Elsevier, Amsterdam.

Joyard, J., Grossman, A. R., Bartlett, S. G., Douce, R., and Chua, N. H. (1982). Characterization of envelope membrane polypeptides from spinach chloroplasts. *J. Biol. Chem.* **257,** 1095–1101.

Joyard, J., Billecocq, A., Bartlett, S. G., Block, M. A., Chua, N. H., and Douce, R. (1983). Localization of polypeptides to the cytosolic side of the outer envelope membrane of spinach chloroplasts. *J. Biol. Chem.* **258,** 10000–10006.

Joyard, J., Blée, E., and Douce, R. (1986). Sulfolipid synthesis from $^{35}SO_4^{2-}$ and [1-^{14}C]-acetate in isolated intact spinach chloroplasts. *Biochim. Biophys. Acta* **879,** 78–87.

Joyard, J., Dorne, A. J., and Douce, R. (1987). Use of thermolysin to probe the cytosolic surface of the outer envelope membrane from plastids. *Methods Enzymol.* **148,** 195–206.

Joyard, J., Pineau, B., Block, M. A., Albrieux, C., and Douce, R. (1990). Envelope membranes from mature spinach chloroplasts contain a NADPH : protochlorophyllide reductase on the cytosolic side of the outer membrane. *J. Biol. Chem.*, **265,** 21820–21827.

Keegstra, K., Olsen, L. J., and Theg, S. M. (1989). Chloroplastic precursors and their transport across the envelope membranes. *Annu. Rev. Plant Physiol.* **40,** 471–501.

Kirk, J. T. O., and Tilney-Bassett, R. A. E. (1978). "The Plastids, their Chemistry, Structure, Growth and Inheritance." Elsevier, Amsterdam.

Kleppinger-Sparace, K. F., Mudd, J. B., and Bishop, D. G. (1985). Biosynthesis of sulfoquinovosyldiacylglycerol in higher plants: The incorporation of $^{35}SO_4$ by intact chloroplasts. *Arch. Biochem. Biophys.* **240,** 859–865.

Laing, W. A., and Christeller, J. T. (1984). Chloroplast phosphoproteins: Distribution of phosphoproteins within spinach chloroplasts. *Plant Sci. Lett.* **36,** 99–104.

Maury, W. J., Huber, S. C., and Moreland, D. E. (1981). Effects of magnesium on intact chloroplasts. II. Cation specificity and involvement of the envelope ATPase in (sodium) potassium/proton exchange across the envelope. *Plant Physiol.* **68,** 1257–1263.

McCarty, D. R., and Selman, B. R. (1986). Properties of a partially purified nucleoside triphosphatase (NTPase) from the chloroplast envelope of pea. *Plant Physiol.* **80,** 908–912.

McCarty, D. R., Keegstra, K., and Selman, B. R. (1984). Characterization and localization of the ATPase associated with pea chloroplast envelope membranes. *Plant Physiol.* **76,** 584–588.

Mendiola-Morgenthaler, L. R., Eichenberger, W., and Boschetti, A. (1985). Isolation of chloroplast envelopes from *Chlamydomonas*. Lipid and polypeptide composition. *Plant Sci. Lett.* **41,** 97–104.

Miquel, M., Block, M. A., Joyard, J., Dorne, A.-J., Dubacq, J.-P., Kader, J.-C., and Douce, R. (1987). Protein-mediated transfer of phosphatidylcholine from liposomes to spinach chloroplast envelope membranes. *Biochim. Biophys. Acta* **937,** 219–228.

Mishkind, M. L., and Scioli, S. E. (1988). Recent developments in chloroplast protein transport. *Photosyn. Res.* **19,** 153–184.

Mourioux, G., and Douce, R. (1979). Transport du sulfate à travers la double membrane limitante, ou envelope, des chloroplastes d'épinard. *Biochimie* **61,** 1283–1292.

Mudd, J. B., and de Zachs, R. (1981). Synthesis of phosphatidylglycerol by chloroplasts from leaves of *Spinacia oleracea* L. (spinach). *Arch. Biochem. Biophys.* **209,** 584–591.

Mudd, J. B., Andrews, J., and Sparace, S. A. (1987). Phosphatidylglycerol synthesis in chloroplast membranes. *Methods Enzymol.* **148,** 338–345.

Ngernprasirtsiri, J., Harinasut, P., Macherel, D., Strzalka, K., Takabe, T., Akazawa, T., and Kojima, K. (1988). Isolation and characterization of the amyloplast envelope-membrane from cultured white-wild cells of sycamore (*Acer pseudoplatanus* L.). *Plant Physiol.* **87,** 371–378.

Ngernprasirtsiri, J., Takabe, T., and Akazawa, T. (1989). Immunochemical analysis shows that an ATP/ADP-translocator is associated with the inner-envelope membranes of amyloplasts from *Acer pseudoplatanus* L. *Plant Physiol.* **89,** 1024–1027.

Nguyen, T. D., and Siegenthaler, P. A. (1983). Proteins and polypeptides of envelope membranes from spinach chloroplasts. Properties of a membrane-bound ATPase. *FEBS Lett.* **164,** 67–70.

Nguyen, T. D., and Siegenthaler, P. A. (1985). Purification and some properties of an Mg^{2+}, Ca^{2+}- and calmodulin-stimulated ATPase from spinach chloroplast envelope membranes. *Biochim. Biophys. Acta* **840,** 99–106.

Nishida, I., Frentzen, M., Ishizaki, O., and Murata, N. (1987). Purification of isomeric forms of acyl-[acyl-carrier-protein] : glycerol-3-phosphate acyltransferase from greening squash cotyledons. *Plant Cell Physiol.* **28,** 1071–1079.

Pain, D., Kanwar, Y. S., and Blobel, G. (1988). Identification of a receptor for protein import into chloroplasts and its localization to envelope contact sites. *Nature (London)* **313,** 232–237.

Peschek, G., Hinterstroisser, B., Wastyn, M., Kunter, O., Pineau, B., Missbichler, A., Lang, J., (1989). Characterization of protochlorophyllide and chlorophyllide by spectrophotometry, spectrofluorimetry, solvent partition and high performance chromatography. *J. Biol. Chem.*, **264,** 11827–11832.

Pineau, B., and Douce, R. (1974). Analyse électrophorétique des protéines de l'enveloppe des chloroplastes d'épinard. *FEBS Lett.* **47,** 255–259.

Pineau, B., Dubertret, G., Joyard, J., and Douce, R. (1986). Fluorescence properties of the envelope membranes from spinach chloroplasts. Detection of protochlorophyllide. *J. Biol. Chem.* **261,** 9210–9215.

Robinson, S. P. (1982). Transport of glycerate across the envelope membrane of isolated chloroplasts. *Plant Physiol.* **70,** 1032–1038.

Robinson, S. P. (1985). The involvement of stromal ATP in maintaining the pH gradient across the chloroplast envelope in the light. *Biochim. Biophys. Acta* **806,** 187–194.

Roughan, P. G., and Slack, C. R. (1982). Cellular organization of glycerolipid metabolism. *Annu. Rev. Plant Physiol.* **33,** 97–132.

Rumpho, M. E., Edwards, G. E., Yousif, A. E., and Keegstra, K. (1988). Specific labeling

of the phosphate translocator in C3 and C4 mesophyll chloroplasts by tritiated dihydro-DIDS (1,2-ditritio-1,2-[2,2'-disulfo-4,4'-diisothiocyano] diphenylethane). *Plant Physiol.* **86,** 1193–1198.

Saidha, T., and Schiff, J. A. (1989). The role of mitochondria in sulfolipid biosynthesis by *Euglena* chloroplasts. *Biochim. Biophys. Acta* **1001,** 268–273.

Schmidt, G. W., and Mishkind, M. L. (1986). The transport of proteins into chloroplasts. *Annu. Rev. Biochem.* **55,** 879–912.

Schulze-Siebert, D., and Schultz, G. (1987). Full autonomy is isoprenoid synthesis in spinach chloroplasts. *Plant Physiol. Biochem.* **25,** 145–153.

Schwaiger, M., Herzog, V., and Neupert, W. (1987). Characterization of translocation contact sites involved in the import of mitochondrial proteins. *J. Cell Biol.* **105,** 235–246.

Shaw, P., Henwood, J., Oliver, R., and Griffiths, T. (1985). Immunogold localisation of protochlorophyllide oxidoreductase in barley etioplasts. *Eur. J. Cell Biol.* **39,** 50–55.

Shibuya, I., Yagi, T., and Benson, A. A. (1963). *In* "Microalgae and Photosynthetic Bacteria" (Jap. Soc. Plant Physiol., ed.), pp. 627–636. University of Tokyo Press, Tokyo.

Siebertz, H. P., Heinz, E., Linscheid, M., Joyard, J., and Douce, R. (1979). Characterization of lipids from chloroplast envelopes. *Eur. J. Biochem.* **101,** 177–185.

Simpson, D. J. (1978a). Freeze-fracture studies on barley plastid membranes. I. Wild-type etioplasts. *Carlsberg Res. Commun.* **43,** 145–170.

Simpson, D. J. (1978b). Freeze-fracture studies on barley plastid membranes. II. Wild-type chloroplasts. *Carlsberg Res. Commun.* **43,** 365–389.

Soll, J. (1985). Phosphoproteins and protein kinase activity in isolated envelopes of pea (*Pisum sativum* L.) chloroplasts. *Planta* **166,** 394–400.

Soll, J. (1987). α-tocopherol and plastoquinone synthesis in chloroplast membranes. *Methods Enzymol.* **148,** 383–392.

Soll, J., and Bennett, J. (1988). Localization of a 64-kDa phosphoprotein in the lumen between the outer and inner envelopes of pea chloroplasts. *Eur. J. Biochem.* **175,** 301–307.

Soll, J., and Buchanan, B. B. (1983). Phosphorylation of chloroplast ribulose bisphosphate carboxylase/oxygenase small subunit by an envelope-bound protein kinase *in situ*. *J. Biol Chem.* **258,** 6686–6689.

Soll, J., Schultz, G., Joyard, J., Douce, R., and Block, M. A. (1985). Localization and synthesis of prenylquinones in isolated outer and inner envelope membrane from spinach chloroplasts. *Arch. Biochem. Biophys.* **238,** 290–299.

Soll, J., Fisher, I., and Keegstra, K. (1988). A guanosine 5'-triphosphate-dependent protein kinase is located in the outer envelope membrane of pea chloroplasts. *Planta* **176,** 488–496.

Somerville, S. C., and Ogren, W. L. (1983). An *Arabidopsis thaliana* mutant defective in chloroplast dicarboxylate transport. *Proc. Natl. Acad. Sci. USA* **80,** 1290–1294.

Somerville, S. C., and Somerville, C. R. (1985). A mutant of *Arabidopsis* deficient in chloroplast dicarboxylate transport is missing an envelope protein. *Plant Sci. Lett.* **37,** 217–220.

Sparace, S. A., and Mudd, J. B. (1982). Phosphatidylglycerol synthesis in spinach chloroplasts: Characterization of the newly synthesized molecule. *Plant Physiol.* **70,** 1260–1264.

Sprey, B., and Laetsch, W. M. (1976). Chloroplast envelopes of *Spinacia oleracea* L. III. Freeze-fracturing of chloroplast envelopes. *Z. Pflanzenphysiol.* **78,** 360–371.

Thompson, A. G., Brailsford, M. A., and Beechey, R. B. (1987). Identification of the phosphate-translocator from maize mesophyll chloroplasts. *Biochem. Biophys. Res. Commun.* **143,** 164–169.

Tolbert, N. E. (1980). Photorespiration. *In* "The Biochemistry of Plants," Vol. 2, "Metabolism and Respiration" (D. D. Davies, ed.), pp. 487–523. Academic Press, New York.

Van Besouw, A., and Wintermans, J. F. G. M. (1978). Galactolipid formation in chloroplast envelopes. I. Evidence for two mechanisms in galactosylation. *Biochim. Biophys. Acta* **529,** 44–53.

Wellburn, A. R. (1982). Bioenergetic and ultrastructural changes associated with chloroplast development. *Int. Rev. Cytol.* **80,** 133–191.

Werner-Washburne, M., and Keegstra, K. (1985). L-aspartate transport into pea chloroplasts. Kinetic and inhibitor evidence for multiple transport systems. *Plant Physiol.* **78,** 221–227.

Woo, K. C., Flügge, U. I., and Heldt, H. W. (1987). A two-translocator model for the transport of 2-oxoglutarate and glutamate during ammonia assimilation in the light. *Plant Physiol.* **84,** 624–632.

Wrischer, M. (1989). Ultrastructural localization of photosynthetic activity in thylakoids during chloroplast development in maize. *Planta* **177,** 18–23.

Yu, J., and Woo, K. C. (1988). Glutamine transport and the role of the glutamine translocator in chloroplasts. *Plant Physiol.* **88,** 1048–1054.

CHAPTER 10

Chloroplast Protein Transport

S. L. Berry-Lowe
G. W. Schmidt

Department of Botany
University of Georgia
Athens, Georgia

I. INTRODUCTION

Plastids (e.g., etioplasts, chloroplasts, leucoplasts) are responsible for much of the metabolism unique to vascular plants and algae, including

photosynthesis and chlorophyll, lipid, and starch biosynthesis. Two-dimensional gel analysis of proteins, calculation of the chloroplast genome-coding capacity, and genetic analysis suggested that most chloroplast proteins are encoded by the nuclear genome (reviewed by Ellis, 1981, 1983). Sequencing has confirmed that the chloroplast genome encodes relatively few proteins (Shinozaki *et al.*, 1986; Ohyama *et al.*, 1986). Consequently, many plastid proteins must be transported into the chloroplast after transcription in the nucleus and translation on free ribosomes in the cytoplasm.

In eukaryotes, secreted proteins have an amino-terminal "signal sequence" of 16–29 amino acids, which directs the precursor protein into the endoplasmic reticulum. The signal peptide is removed concomitant with translation to form the mature protein (Blobel and Dobberstein, 1975; for a review, see Reithmeier, 1988). In vascular plants and green algae, the small subunit (SS) of ribulose bisphosphate carboxylase oxygenase was also found to be synthesized as a precursor (Dobberstein *et al.*, 1977; Chua and Schmidt, 1978; Highfield and Ellis, 1978; Cashmore *et al.*, 1978). Studies showed that the precursor (pSS) is transported and processed after translation by a mechanism requiring ATP but not plastid protein synthesis (Chua and Schmidt, 1978; Highfield and Ellis, 1978; Smith and Ellis, 1979; Grossman *et al.*, 1980). The amino-terminal sequence, or transit sequence, is much longer than a typical signal sequence. Proteins that are transported into mitochondria also are synthesized as precursors with an amino-terminal targeting sequence. Transport and processing of these precursors is similar to that in chloroplasts because it also occurs posttranslationally (for a review, see Ellis and Robinson, 1985).

There are many recent reviews on chloroplast transport (Schmidt and Mishkind, 1986; Keegstra *et al.*, 1987, 1989; Keegstra and Bauerle, 1988; Lubben *et al.*, 1988; Mishkind and Scioli, 1988; Keegstra, 1989). This review highlights recent studies on the structure of the transit peptide and the receptors and enzymes involved in binding and processing.

II. TRANSIT PEPTIDES

The ribulose bisphosphate carboxylase SS has been used as the primary model to determine which domains are responsible for directing the precursor across the chloroplast membranes. Also, a number of studies in-

volve the transit peptides of the precursors of the light-harvesting chlorophyll *a/b*-binding protein (pLHCP), plastocyanin, and ferredoxin. Recently, many other genes have been characterized that encode precursors that are transported into the chloroplast; however, studies utilizing the transit sequences of these proteins have been limited. This discussion is divided into two sections describing (A) the sequence and structure of the different transit peptides and (B) the functional studies utilizing mutant and chimeric fusion proteins.

A. Structure of the Transit Peptides

1. Transit Sequence of the SS of Ribulose Bisphosphate Carboxylase

The first pSS transit sequence was determined by microsequencing the *in vitro* translated precursor of the green alga *Chlamydomonas* (Schmidt *et al.*, 1979). This transit peptide is 44 amino acids in length, hydrophobic, basic, and removed in the chloroplast stroma after transport. Subsequent analysis of partial cDNA and full-length genomic clones isolated from pea (Bedbrook *et al.*, 1980) and soybean (Berry-Lowe *et al.*, 1982) suggested that the transit peptides of the two dicotyledonous plants are similar; however, only low amino acid identity exists between the transit peptides of the two dicotyledonous plants and *Chlamydomonas*. Clones encoding SS have now been isolated from >20 genera (for listings, see Keegstra *et al.*, 1989; Meagher *et al.*, 1989). The lengths of the transit peptides vary from 44 or 45 amino acids in *Chlamydomonas* (Schmidt *et al.*, 1979; Goldschmidt-Clermont and Rahire, 1986) to 59 amino acids in cucumber (Greenland *et al.*, 1987).

A transit peptide consensus sequence for the SS of vascular plants is published by Keegstra *et al.* (1989) and is used in all comparisons in this paper (see Table I). Although comparison of the consensus sequence with the *Chlamydomonas* sequence is difficult, because the latter is much shorter, some regions are similar. From studies utilizing *in vitro* translation of the *Chlamydomonas* SS and subsequent import of the product into pea or spinach chloroplasts (Mishkind *et al.*, 1985b), it was concluded that there are three conserved domains in the algal and corresponding regions of known vascular plants' transit peptides. Although the concept of functional domains is not established, domain delineation still helps to describe transit sequence structure and function. Domain I extends from amino acids 1–20 (numbering corresponds to the longer vascular

TABLE I
Amino Acid Sequences of Transit Peptides

TRANSIT SEQUENCE

STROMA

MASS MLSVATRTNP AQASMVAPFT GLKSAASFPV
MAAVI AKSSVSAAVA RPARSSVRPM
M AVCTVYTIPT TTHLGSSFNQ NNKQVFFNYK
M AASVSRAICV QKPGSKCTRD REATSFARRS VAAPRPPHAK AAGVIRSDSG
MAAATTTTSR PLLLSRQQAA ASSLQCRLPR
MA TTPALYGTAV STSFLRTQPM PMSVTTTKAF
MASTLSTL SVSASLLPKQ QPMVASSLPT
MAA HSIFTTTSTT NSFLYPISSS SSSPNINSSF LGVSLNVNAK
MASQTLVS PSPLSSHSLL RTSFSGVSVK
MASSML SATTVPLQQG GGLSEFSGLR SSASLPMRRN ATSDDFMSAV
SLQV SNKGFSEFSG LRTSSAIPFG RKTNDDLLSV
<CLS KKFEVAEFAG LRSSGCVTFS NKESSFFDVV
MA SLPVNKIIPS
MAQILAPST QWQMRITKTS PCATPITSKM
MAQILA PSTQWQMRFT KSSRHASPIT SNTWSSLLMK
MAAATTTTTT SSSISFSTKP SPSSSKSPLP ISRFSLPFSL NPNKSSSSSR RRGIKSSSPS SISAVLNTTT NVTTTPSPTK
MAAAAPSPSS SAFSKTLSPS SSTSSTLLPR STFPFPHHPH KTTPPPLHLT HTHIHIHSQR RRFTISNVIS TNQKVSQTEK
MA AATSSSPISL TAKPSSKSPL PISRFSLPFS LTPQKDSSRL HRPLAISAVL
MAS FSFFGTIPSS PTKASVFSLP VSVTTLPSFP RRRATRVSVS
MATAVSTV GAATRAPLNL NCSSAGASVP TSGFLGSSLK
MAQ SVSLSTIASP ILSQKPGSSV KSTPPCMASF
MAQVSR ICNGVQNPSL ISNLSKSSQR KSPLSVSLKT QQHPRAYPIS SSWGLKKSGM
MA QINNMAQGIQ TLNPNSNFHK PQVPKSSSFL VFGSKKLKNS ANSMLVLKKD
MAQISS MAQGIQTLSL NSSNLSKTQK GPLVSNSLFF GSKKLTQISA KSLGVFKKDK
MASLSA TTTVRVQPSS SSLHKLSQGN GRCSSIVCLD
MAELIQDK
M ATTFSASVSM QATSLATTTR ISFQKPVLVS
MAHCLAAVS SFSPSAVRRR LSSQVANVVS SRSSVSFHSR
M STTFCSSVSM QATSLAATTR ISFQKPALVS
MATQ FSASVSLQTS CLATTRISFQ KPALISNHGK
MATMACASSL TFPSAQTQKS FFGTNVKQTP VLSFPRPTVA
MASST LSSLSSTPLQ
MVAMAMASLQ SSMSSLSLSS
MALLCFNSFT TTPVTSSSSL FPHPTANPIS
M ASVSSIFGCG
<RA SPPSESRAPL
MATANQAHLF NHYSSNSRFI HFTSRNTSSK
MA ALATSQLVAT RAGLGVPDAS TFRRGAAQGL RGARASAAAD TLSMRTSARA
MA ALATSQLATS GTVLGVTDRF RRPGFQGLRP RNPADAALGM RTIGASAAPK

THYLAKOID

MAAS TMALSSPSFA
MATCAI QQSAFVGQAV
SN TLMSCGIPAV CPSFLSSTKS
M ASACASSTIA AVAFSSPSSR
MA AAVTAAVSFP STKSTPLSTR TSSVITHEKI NFNKVPLYYR NVSVGGKVGT
MA AAVTAAVSLP YSNSTSLPIR TSIVAPERLV
MTTAV TAAVSFPSTK TTSLSARSSS VISPDKISYK
MAMATQAT LFSPSSLSSA KPIDTRLTTS FKQPSAVTFA
MAMATQASLF TPPLSVPKST TAPWKQSLVS
M ATSVMSSLSL KPSSFGVDTK
MASTVMSSL SLKPTFTLEK
MQALS
MAL TMRNPAVKAS SRVAPSSRRA LRVACQAQKN ETASKVGTAL
MIISIFNQ LHLTENSSLM ASFTLSSATP SQLCSSKNGM FAPSLALAKA
MAALQNPVAL QSRTTTAVAA LSTSSTTSPP KPFSLSFSSS TATFNPLRLK
MAAML ASKQGAFMGR

	PROTEIN	SPECIES	REFERENCE
SRKQNLDITS IASNGGRVQC	RuBPC SSU	Consensus	Keegstra et al, 1989
AALKPAVKAA PVAAPAQANQ	RuBPC SSU	Chlamydomonas	Goldschmidt-Clermont and Rahire, 1986
RSSSSNNTLF TTRPSYVITC	Ribulokinase	Spinach	Roesler and Ogren, 1988
AGRGQHCSPL RAVVDAAPIQ	PPDikinase	Maize	Matsuoka et al, 1988
RPGSSLFAGQ GQASTPNVRC	FBPase	Wheat	Raines et al., 1988
SNGFLGLKTS LKRGDLAVAM	Ferredoxin	Pea	Elliott et al., 1989
NMGQALFGLK AGSRGRVTAM	Ferredoxin	Silene	Smeekens et al., 1985
FGQSLTLYAV TTPKPLTVFA	SODismutase	Tomato	Perl-Treves et al., 1988
LAPQFSTLAT SNFKPLTVVA	SODismutase	Pea	Scioli and Zilinskas, 1988
SFRTHAVGTS GGPRRAPTEA	GAPDH	Maize	Quigley et al., 1988
VAFQTSVIGG GNSKRGVVEA	GAPDH	Tobacco	Shih et al., 1986
SAQLTPKTTR STPVKGERVA	GAPDH	Tobacco	Shih et al., 1986
STTLLSSSNN NRRRNNSSIR	Nitrate Rdtase	Spinach	Back et al., 1988
WSSLVMKQTK KVAHSAKFRV	Gln Synthase	Pea	Tingey et al., 1988
QNKKTSSAKF RVLAVKSDGS?	Gln Synthase	Bean	Lightfoot et al., 1988
PTKPETFISR FAPDQPRKGA?	ACLS I	Tobacco	Mazur et al., 1987
TETFVSRFAP DEPRKGSDVL?	ACLS II	Tobacco	Mazur et al., 1987
NSPVNVAPPS PEKTDKNKTF	ACLS I	Rape	Rutledge (personal communication)
ANSKKDQDRT ASRRENPSTF	ACLS II	Rape	Rutledge (personal communication)
KHTNVRFPSS SRTTSMTVKA	RuBPC Activase	Spinach	Werneke et al., 1988
PLRRQLPRLG LRNVRAQAGG?	Heat Shock Prot	Pea	Vierling et al., 1988
TLIGSELRPL KVMSSVSTAE	EPSPS	Arabidopsis	Klee et al., 1987
SIFMQKFCSF RISASVATAQ	EPSPS	Petunia	Gasser et al., 1988
VLRVVRKSSF RISASVATAE	EPSPS	Tomato	Gasser et al., 1988
WGKSSFPTLR TSRRRSFISA	AcylCP	Spinach	Scherer and Knauf, 1987
ESAQSAATAA AASSGYERRN	AcylCP	Squash	Ishizaki et al., 1988
NHGRTNLSFN LSRTRLSISC	AcylCP	Rape	Safford et al., 1988
QMSFVSISSR PSSLRFKICC	AcylCP	Barley	Hansen, 1987
TTNLSFNLRR SIPTRFSISC	AcylCP	Turnip	Rose et al., 1987
TNLSFNLRRS IPSRRLSVSC?	AcylCP	Arabidopsis	Post-Beittenmiller et al., 1989
AAVAVSARKS TSASTKCTEE?	L13 Rib Prot	Maize	Phua et al., 1989
HSFAANLKTC SQFPNKSSGF?	L9 Rib Prot	Pea	Gantt, 1988
NSFLGQPLSP ITLSPFLQGK?	L24 Rib Prot	Pea	Gantt, 1988
RVRIGLPTNC LKGFRILTPI?	L18 Rib Prot	Pea	Gantt, 1988
VSMAPNSSLR NKAIRTERRS?	L25 Rib Prot	Pea	Gantt, 1988
RAPQRSATRQ HQARQGPRRM	ADPGP	Wheat	Olive et al., 1989
LFLTKTSHFR RPKRCFHVNN	StarchPhos	Potato	Brisson et al., 1989
APRHQQQARR GGRFPSLVVC	Waxy	Maize	Klosgen et al., 1986
QSRKAHRGSR RCLSVVVSAT?	Waxy	Barley	Rohde et al., 1988
GKAVKLSPSS SEITGNGRVT	LHCPII(I)	Consensus	Keegstra et al., 1989
GKSQNEFIRK VGNFGEGRIT	LHCPII(II)	Tomato	Pichersky et al., 1987b
KFAAAMPVSV GATNSMSRFS?	LHCPI(I)	Tomato	Pichersky et al., 1987a
RNGSIVGTTK ASFLGGRRLR?	LHCPI(II)	Tomato	Pichersky et al., 1988
	Fd Reductase	Ice Plant	Michalowski et al., 1989
FKKVSLNNVS ISGRVGTIRA	Fd Reductase	Pea	Newman and Gray, 1988
KVPLYYRNVS ATGKMGPIRA	Fd Reductase	Spinach	Jansen et al., 1988
SKPASRHHSI RAAAAAEGKR?	PSI Subunit II	Spinach	Lagoutte, 1988
FSTPKQLKST VSVTRPIRAM	PSI Subunit II	Tomato	Hoffman et al.,1988
SAVKGLPSLS RSSASFTVRA	PSII 10kD	Spinach	Lautner et al., 1988
TSVKGLPSLA RSSSSFKVVA	PSII 10kD	Potato	Eckes et al., 1986
SRVNIAAKPQ RAQRLVVRAE	PSI 11 kDa	Chlamydomonas	Franzen et al., 1989
AASALAAAVS LSAPSAAMAD	PSI 18 kDa	Chlamydomonas	Franzen et al., 1989
GRVNVLISKE RIRGMKLTCQ	Rieske Fe-S	Spinach	Steppuhn et al., 1987
ILTASKLTAK PRGGALGTRM	D-ATP Synthase	Spinach	Hermans et al., 1988
SSFAPAPKGV ASRGSLQVVA	G-ATP Synthase	Chlamydomonas	Yu and Selman, 1988

(continues)

TABLE I
Continued

	TRANSIT SEQUENCE
LUMEN	
	MATVTS SAAVAIPSFA GLKASSTTRA ATVKVAVATP RMSIKASLKD
	MA AITSATVTIP SFTGLKLAVS SKPKTLSTIS RSSSATRAPP KLALKSSLKD
	MATVASSAA VAVPSFTGLK ASGSIKPTTA KIIPTTTAVP RLSVKASLKN
	MRFRPSIV
	M ALRAAQSAKA GVRAARPNRA TAVVCKAQKV
	MAASL QSTATFLQSA KIATAPSRGS SHLRSTQAVG KSFGLETSSA RLTCSFQSDF KDFTGKCSDA
	M AASLQAAATL MQPTKLRSNT LQLKSNQSVS KAFGLEHYGA KVTCSLQSDF KELAHKCVEA
	MAAS LQASTTFLQP TKVASRNTLQ LRSTQNVCKA FGVESASSGG RLSLSLQSDL KELANKCVDA
	MATALCNK AFAAAPVARP ASRRSAVVVR ASGSDVSRRA
	M ASTACFLHHH AAISSPAAGR GSAAQRYQAV SIKPNQIVCK AQKQDDNEAN VLNSGVSRRL
	M ALASKVATRP AVASRRGAVV VRASGESRRA
	MAQ AMASMAGLRG ASQAVLEGSL QISGSNRLSG PTTSRVAVPK MGLNIRAQQV SAEAETSRRA
	MLQLANRS VRAKAARASQ SARSVSCAAA KRGADVAPLT
ENVELOPE	
	MESRV LSRTTAIAAL PKLFRPSREA ASFGFATGVK TPVGLVKDGG SLTWGRQLRP VLLLEPVQTG

[a] Transit peptide sequences are aligned at the presumptive processing site at the carboxyl terminus, in single-letter amino acid designation. ? indicates a presumptive processing site. < indicates unknown amino-terminal sequences. Sequences are grouped according to targeting compartment.

plant sequence) and is present with only one exception (wheat) in the genera studied. Domain II extends from amino acids 25–33 and seems to be required for import in *Chlamydomonas* (Mishkind *et al.*, 1985). Domain III extends from amino acids 41–59 and is highly conserved in the higher-plant transit sequences but less conserved in *Chlamydomonas*. Karlin-Neumann and Tobin (1986) also concluded, on the basis of amino acid similarities, that three conserved domains in transit sequences form a "common framework" in the known transit peptides (several pSS and pLHCP sequences and a single ferredoxin sequence). Domain I corresponds to amino acids 1–8, domain II to amino acids 27–32, and domain III to amino acids 54–57.

2. Chlorophyll *a*/*b*-Binding Protein Transit Sequences

Cashmore (1984) isolated the first LHCP cDNA containing the transit sequence from pea. *cab* genes encoding these proteins have now been isolated from at least 10 higher-plant genera (for a list, see Keegstra *et al.*, 1989). These sequences have been classified into types I and II of light-harvesting complex I (for photosystem I [PS I]-associated LHCP I) and types I and II of light-harvesting complex II (LHC II) on the basis of the mature protein sequence and function as well as nucleotide sequence comparisons (Pichersky *et al.*, 1987a). There is considerable amino acid sequence identity between transit peptides within a category

	PROTEIN	SPECIES	REFERENCE
VGVVVAATAA AGILAGNAMA	Plastocyanin	Silene	Smeekens et al., 1985
FGVIAVATAA SIVLAGNAMA	Plastocyanin	Arabidopsis	Vorst et al., 1988
VGAAVVATAA AGLLAGNAMA	Plastocyanin	Spinach	Rother et al., 1986
ALLSVCFGLL TFLYSGSAFA	OEE1	Synechocystis	Philbrick and Zilinskas, 1988
GQAAAAAALA TAMVAGSANA	OEE1	Chlamydomonas	Mayfield et al., 1989
VKIAGFALAT SALVVSGASA	OEE1	Arabidopsis	Ko et al., 1990
SKIAGFALAT SALVVSGASA	OEE1	Pea	Wales et al., 1989
TKLAGLALAT SALIASGANA	OEE1	Spinach	Tyagi et al., 1987
ALAGFAGAAA LVSSSPANAA	OEE2	Chlamydomonas	Mayfield et al., 1987
ALTVLIGAAA VGSKBSPADA	OEE2	Spinach	Jansen et al., 1987
VLGGLLASAV AAVAPKAALA?	OEE3	Chlamydomonas	Mayfield et al., 1989
MLGFVAAGLA SGSFVKAVLA	OEE3	Spinach	Jansen et al., 1987
SALAVTASIL LTTGAASASA	Cyt c 552	Chlamydomonas	Merchant and Bogorad, 1987
PVCSRREKTA VQPCRAASGS?	Pi Transporter	Spinach	Flügge et al., 1989

(particularly LHCP II type I, because many of these genes have been sequenced) but poorer identity when the different classes are compared with each other (Table I). All of the pLHCP transit peptides begin with the amino acids methionine-alanine. The three domains defined by Karlin-Neumann and Tobin (1986) can be identified, particularly in the LHCP II type I class peptides. The pLHCP transit peptides of both wheat and pea contain an abundance of serine and threonine residues and a central region containing the basic amino acids arginine and lysine (Lamppa *et al.*, 1985). Lamppa *et al.* (1985) postulate that the positions of the hydroxylated amino acids and their ability to be phosphorylated may be important in transport. Phosphorylation–dephosphorylation as a motive force for protein import is discussed below.

One striking feature of the tomato *cab* 7 gene (Pichersky *et al.*, 1988) is that it encodes an intron in the transit sequence. Other genes known to contain introns in the nucleotide sequence encoding the transit sequence include the maize glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene (Quigley *et al.*, 1988) and the *Arabidopsis* acyl carrier protein gene (Post-Beittenmiller *et al.*, 1989). Introns have been suggested to separate exons, each containing a functional domain (Gilbert, 1979). Quigley *et al.* (1988) show that the two introns in the GAPDH transit sequence do separate the three common domains (described by Karlin-Neumann and Tobin, 1986), which can be identified in this transit sequence. One intron in the GAPDH transit sequence occurs between

domains I and II and the other between domains II and III, very near domain III. The intron in the tomato *cab* 7 gene occurs between domains I and II, although alignment of this PS I type 1I sequence with the more common photosystem II (PS II) type I consensus sequence is somewhat difficult. The intron in the *Arabidopsis* acyl carrier protein transit sequence also occurs in the amino portion of the peptide after domain I. Further support for the domain model comes from the observation that most SS genes have an intron placed just two amino acids past the cleavage site, within the mature sequence (Berry-Lowe *et al.*, 1982). The second intron of the *Arabidopsis* acyl carrier protein gene also occurs just past the processing site, inside the mature sequence. It will be of interest to isolate and compare more genes containing transit sequences with introns to determine whether or not they occur in similar positions and delineate functional domains (i.e., do introns occur between the domains hypothesized to be important for transport, processing, or targeting?). We may expect genes encoding introns in the transit sequence to be isolated from ancient gene lineages if introns did originally have this function but were lost through evolution.

3. Transit Sequences of Other Nuclear-Encoded Genes

Many other nuclear-encoded genes with presumptive transit sequences have recently been isolated. These are summarized by Keegstra *et al.* (1989) and in Table I. Some of the products these genes encode are directed to the stroma or thylakoid membrane and others to the thylakoid lumen. Many of these genes encode proteins that function in basic metabolic processes unique to the chloroplast such as photosynthesis. Amino acid sequences of five chloroplast ribosomal protein genes, one envelope phosphotranslocator gene, and several starch enzyme genes (which function in the chloroplast or amyloplast) are also included in Table I. Frequently, strong conservation is observed when the amino acid sequences of transit peptides encoded by the same gene from different plants are compared. For example, the amino acid sequences of the transit peptides of the barley and maize *waxy* genes are very similar. On the other hand, this is not always the case. In part due to the length discrepancy, the amino acid sequences of the transit peptides of the acyl carrier proteins from six plants differ substantially. Although the three-domain framework (Karlin-Neumann and Tobin, 1986) may be applicable in some cases, it is by no means universal, and for many of these other transit sequences there is no obvious conservation of the domains (Table I).

B. Functional Studies

1. Studies with Ribulose Bisphosphate Carboxylase SS Peptides

Comparisons of amino acids from the transit peptides, particularly that of pSS, have provided a basis for functional studies to determine the role of conserved domains in targeting and transport. Most studies have been carried out by one of two procedures. One method has been to mutate the transit or mature nucleotide sequence, transcribe it *in vitro* with the SP6 system, translate it *in vitro*, and examine import of the product in isolated chloroplasts. The other method has been to transform these altered gene constructions into plants and examine the transport and localization of a fusion or reporter protein.

Initial experiments suggested that the transit peptide alone is required for import (see Fig. 1). No import of *Chlamydomonas* or pea SS was detected in isolated spinach and pea chloroplasts when the transit peptide was removed by transit peptidase prior to import assays (Mishkind *et al.*, 1985b). The essential role of the transit peptide for import was also demonstrated by Anderson and Smith (1986) using an *in vitro* transcription system to generate mRNA for a mature SS; the translation product corresponding to the mature SS of pea was not imported into isolated chloroplasts.

After transformation into tobacco, the product of a construction of the SS transit sequence fused to the neomycin phosphotransferase (NPT II) gene was transported and processed in the chloroplast (Schreier *et al.*, 1985; Van den Broeck *et al.*, 1985; Kuntz *et al.*, 1986). In contrast to these results, when the same gene fusion was made and tested in an *in vitro* system, rather than by transformation, transport occurred but at very low efficiency (Wasmann *et al.*, 1988). The reason for this discrepancy is unclear. Fusions of the SS transit sequence with the coding regions for the ribulose bisphosphate carboxylase large subunit (LS) from *Anacystis* (Gatenby *et al.*, 1988), the β subunit of the maize chloroplast ATPase (Gatenby *et al.*, 1988), or the brome mosaic virus (BMV) coat protein (Lubben *et al.*, 1989) resulted in products that could be imported by isolated chloroplasts after *in vitro* transcription and translation. Processing of the BMV fusion protein was incomplete and resulted in an improper size of mature product. The chimeric SS-LS gene product was processed properly however, and a portion could be detected in holoenzyme.

Other studies suggest that the mature sequence is also a factor in import. As reported above, transport of the *in vitro* transcribed and translated product was inefficient when the SS transit sequence was fused

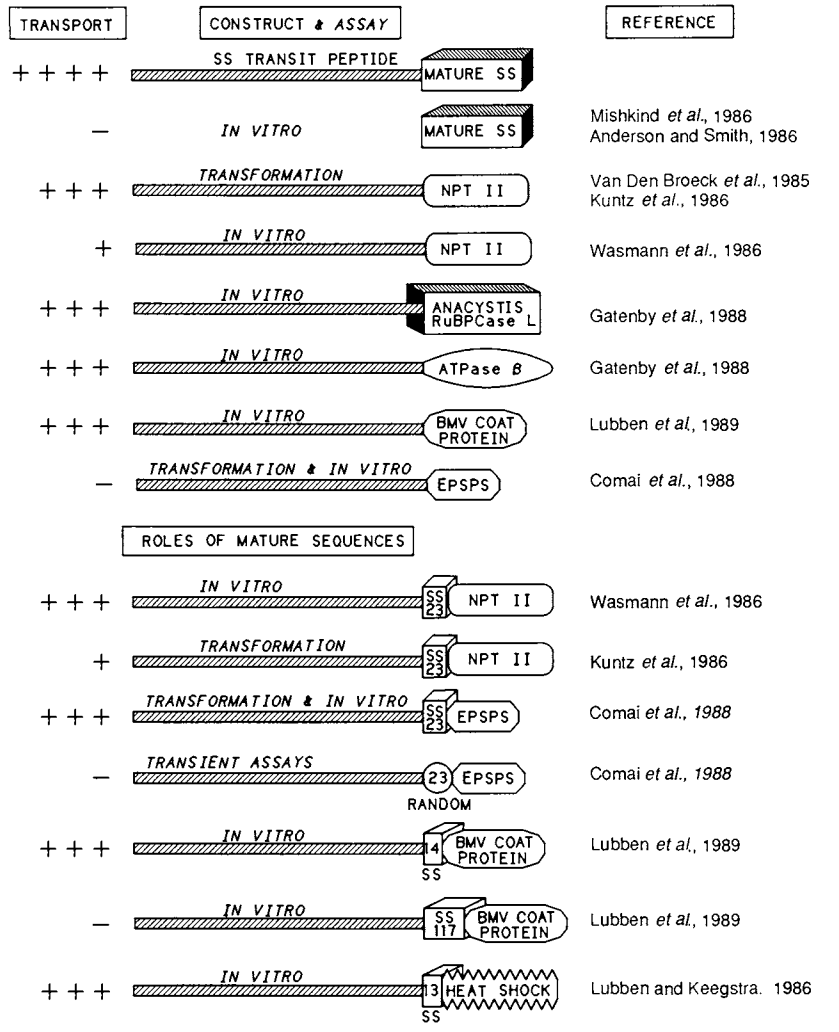


FIG. 1. Import studies with RuBPCase small subunit precursor (pSS) transit sequences. Upper portion: Transport when chimeras were constructed utilizing the pSS transit peptide and sequences of various passenger proteins. Lower portion: Transport of constructions composed of the pSS transit sequence plus amino-terminal sequences of mature small subunit (SS). Levels of transport are indicated as minimal (+) to maximal (+ + + +). Abolished or undetectable transport is denoted as (–). Assays are designated as either *in vitro* (*in vitro* transcription–translation products were imported into isolated chloroplasts) or *in vivo* (gene constructions were transformed into plants and levels of fusion proteins assayed). References are given on the right. BMV, brome mosaic virus coat protein; EPSP, 5-enolpyruvylshikimate-3-phosphate synthase gene; NPT II, neomycin phosphotransferase; RuBPCase, ribulose bisphosphate carboxylase.

directly to NPT II. Import could be improved to wild-type efficiency if the nucleotide sequence encoding the first 23 amino acids of the mature SS were inserted between the transit sequence and the NPT II gene (Wasmann *et al.*, 1986). When the same construct was tested in transformed plants, transport of the product occurred but was less than if the NPT II gene was fused directly to the transit sequence (Kuntz *et al.*, 1986). Again, the *in vitro* and *in vivo* results are contradictory. Similarly, transport occurred *in vitro* or *in vivo* only when the 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) gene was fused to the gene encoding the SS transit sequence and 23 amino acids from the mature SS (Comai *et al.*, 1988). When Comai *et al.* (1988) replaced the nucleotide sequence encoding the 23 amino acids from the mature SS with a sequence encoding 24 random, but hydrophobic and negatively charged, amino acids, transport of EPSPS was abolished, as assayed by transient assays.

When the gene encoding the SS transit sequence and first 13 amino acids of the mature sequence was fused to a cytosolic heat-shock protein gene, the *in vitro* transcribed and translated protein was imported into isolated chloroplasts (Lubben and Keegstra, 1986). Transport of the product was also reported when the SS transit sequence was fused to amino acids 1–14 of the SS mature sequence and to the BMV coat protein, but not when the entire SS mature sequence (portion encoding amino acids 1–117) was fused to the BMV coat protein (Lubben *et al.*, 1989). In another case, the nucleotide sequence encoding the chloroplast gene *psbA* and its upstream open reading frame (encoding 16 amino acids) was fused to the pSS transit sequence and to the nucleotides encoding the first 23 amino acids of the SS mature sequence. This construct was transformed into tobacco plants (Cheung *et al.*, 1988) and the fusion protein was localized to the thylakoid membrane. Therefore, the composition of part or all of the mature sequence also influences transport. An implication of the Cheung *et al.* (1988) study is that the 32-kDa PS II reaction center polypeptide (D1) can be posttranslationally inserted into thylakoids. This is surprising because the mRNA is translated by thylakoid-bound ribosomes (Herrin and Michaels, 1985).

Deletions of various portions of the gene encoding the pSS transit sequence have been tested for their subsequent effect on transport and processing. Figure 2 summarizes the experiments and findings of these attempts to define the functional domains for chloroplast import. When nucleotides encoding the three carboxyl-terminal amino acids of the transit peptide were deleted and the truncated sequence fused to the gene encoding amino acids 1–22 of the SS mature peptide and to the NPT II gene, the fusion protein was transported into the chloroplast

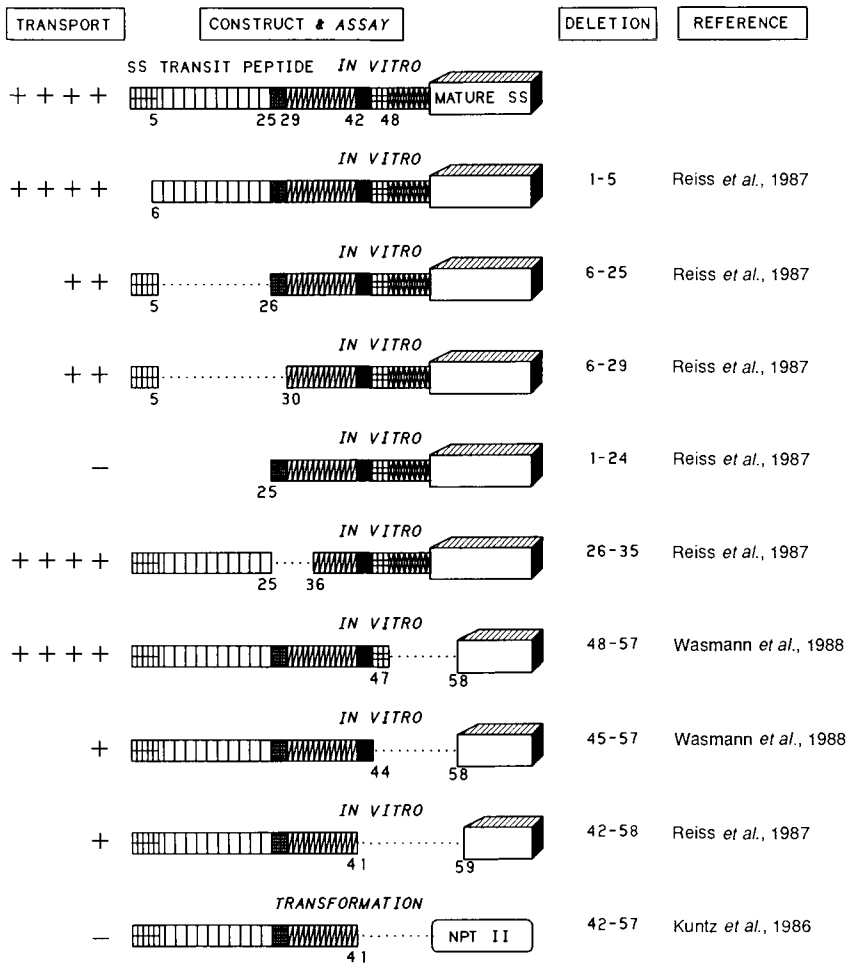


FIG. 2. Effects of pSS transit sequence deletions on chloroplast import. Import levels are indicated as in Figure 1. Deletions in the pSS transit sequence are shown as missing transit peptide segments and/or a dotted line. The wild-type construction is shown at the top for comparison.

in transformed plants (Kuntz *et al.*, 1986). Deletion of nucleotides encoding the 10 most carboxyl-terminal amino acids of the transit peptide did not affect transport of the *in vitro* protein product, while deletion of nuceleotides encoding the 13 most carboxyl-terminal amino acids reduced transport of the protein by 75%, and deletion of the nucleotides encoding the 17 most carboxyl-terminal amino acids nearly abolished transport entirely as compared with the wild type (Wasmann *et al.*, 1988;

Reiss *et al.*, 1987, respectively). These data are in agreement with another study: no transport occurred when the nucleotides encoding the 16 most carboxyl-terminal amino acids of the transit peptide were deleted and the truncated gene was fused to an NPT II gene and transformed into tobacco (Kuntz *et al.*, 1986). In fact, the fusion protein was detected, but in the chloroplast envelope rather than in the stroma. Reiss *et al.* (1987) and Kuntz *et al.* (1986) also altered the site of the junction of the transit and mature peptides, either by deleting and/or replacing the nucleotides encoding one or several amino acids, including the first amino acid, methionine, of the mature SS peptide. No reduction in transport (or processing) of the protein was detected. The authors (Kuntz *et al.*, 1986; Wasmann *et al.*, 1988; Ostrem *et al.*, 1989; Reiss *et al.*, 1987, 1989) suggest that the conserved region at the carboxyl terminus of the pSS transit peptide (domain III) is primarily important for processing rather than transport.

In-frame deletions or duplications of the nucleotides encoding 20 amino acids in the central region of the SS transit sequence have no effect on transport or processing of the protein (Reiss *et al.*, 1987). Deletions of nucleotides encoding the amino terminus of the transit peptide were also examined by Reiss *et al.* (1987) (Fig. 2). There was no effect on transport of the *in vitro* protein product if nucleotides encoding only the first five amino acids were deleted. If regions encoding amino acids 6–25 or 6–29 were deleted, transport of the *in vitro* protein occurred but was reduced; however, if the nucleotides encoding amino acids 1–24 were deleted, transport of the protein into isolated chloroplasts was nearly abolished. Subsequently, Reiss *et al.* (1989) analyzed 10 mutants for binding of protein to the chloroplast and transport of protein into chloroplasts but could not specify a specific domain of the transit peptide important only for transport.

Most recently, De Almeida *et al.* (1989) determined that the presence of the SS transit sequence in constructions that are transformed into tobacco affects gene expression. The constructions consist of the SS promoter region fused to *neo* (NPT II) or *bar* (phosphinothricin acetyltransferase) reporter genes with and without the SS transit sequence. Expression at both mRNA and protein levels is higher when the transit sequence is included in the construct. Import into the chloroplast was detected, but further characterization, including processing, was not examined.

Naturally occurring mutations that may affect the structure of transit sequences of stromal proteins have been observed in only one case. Schmidt *et al.* (1984) calculated, on the basis of electrophoretic mobility, that the erythromycin resistance conferred by a mutation in the L-6 ri-

bosomal protein of chloroplasts might include a deletion of half of the sequences of the transit peptide. In wild-type cells, the L-6 precursor behaves as a 30.1-kDa polypeptide that is matured to a size of 24.4kDa. In contrast, the corresponding translation product from *ery-M1b* cells is 22.4 kDa and can be matured to a 19.9-kDa product; thus, the transit peptide of the mutant would be 2.5 kDa, while that of the wild type is 5.7 kDa. Unfortunately, the dispensible sequences of the L-6 transit peptide have not been determined.

2. Studies with LHCP, Ferredoxin, and Plastocyanin Transit Peptides

In the experiments described above, the pSS transit peptide was used to direct the product to the chloroplast stroma. The localization of the products of genes fused to pLHCP transit sequence was of interest because the *cab* gene products are directed to the thylakoid membrane. The suggestion that the mature peptide contains a signal for transport and association with the thylakoid membrane was supported by the observation that, when the SS transit sequence was fused to the mature wheat LHCP sequence and transcribed and translated *in vitro*, transport and targeting of the protein to the thylakoid membrane occurred (Lamppa, 1988). When the pLHCP transit sequence was fused to partial LHCP mature sequences (encoding various numbers of amino acids) and the *gus* (β-glucuronidase) reporter gene and was transformed into tobacco, optimum transport into the chloroplast of the fusion protein occurred when the nucleotides encoding amino acids 1–24 or 1–53 from the mature LHCP peptide were included (Kavanagh *et al.*, 1988). Although transport occurred, only 10% of the fusion protein activity was associated with thylakoid membranes. Similarly, Kohorn *et al.* (1986) made six deletions in the mature coding region of a *cab* gene. The deleted regions corresponded to amino acids that were thought to cross or extend into the lipid bilayer, stroma, or lumen. The proteins from four of the constructs were transported and processed by isolated chloroplasts and were found in association with the thylakoid membrane, although they did not form LHC II complexes. The proteins from the other two mutant constructions were not detected in the chloroplast at all and either may not have been transported or may have been rapidly degraded. These studies have utilized gene constructions with deletions in portions of the gene encoding the mature sequence. Also, Buvinger *et al.* (1989) showed that a synthesized peptide corresponding to the 20 amino-terminal amino acids of the pLHCP transit peptide competes with native transit peptide and interferes with the transport of pLHCP into isolated pea chloroplasts.

Experiments have also been carried out using the transit sequences of plastocyanin and ferredoxin, proteins targeted to the thylakoid lumen and chloroplast stroma, respectively. Smeekens *et al.* (1986, 1987) suggested that there are two functional domains of the plastocyanin transit peptide: one directing the peptide into the chloroplast stroma, and the other directing the protein across the thylakoid membrane. When the sequence encoding the transit peptide of ferredoxin was fused to sequences encoding either plastocyanin or yeast superoxide dismutase, the *in vitro* transcription–translation fusion protein was imported and localized in the chloroplast stroma. Conversely, when the sequence encoding the transit peptide of plastocyanin was fused to the sequence encoding either the ferredoxin or yeast superoxide dismutase gene, the fusion protein was not detected in the thylakoid lumen but, rather, as an intermediate in the chloroplast stroma with the transit peptide still present (Smeekens *et al.* 1986, 1987). This shows that proper compartment targeting is not entirely controlled by the transit peptide.

Deletions of the nucleotide sequences encoding the ferredoxin transit or mature peptides have also been analyzed. Deletions in the gene encoding the mature peptide of ferredoxin resulted in reduced binding to the chloroplast and reduced transport if the first four amino-terminal amino acids were removed, or low binding and undetectable transport if amino acids 1–47 of the mature peptide were removed (Smeekens *et al.*, 1989). Wild-type binding to the chloroplast and transport was found if the nucleotide sequence encoding the two most carboxyl-terminal amino acids of the ferredoxin transit peptide were deleted. No binding or transport was detected when the sequence encoding the seven most carboxyl amino acids of the transit peptide were deleted (Smeekens *et al.*, 1989). These results agree with the SS studies and suggest that the carboxyl terminus, particularly the glycine-arginine-valine residues of domain III, is important for transport. If a nucleotide linker was inserted between the genes encoding the ferredoxin transit and mature sequences, a fusion protein that is extremely hydrophobic and basic could be produced *in vitro*. Binding of this fusion protein to the chloroplast was very efficient, but transport was not detected, perhaps because the mature peptide of ferredoxin is normally very acidic (Smeekens *et al.*, 1989).

3. Transit Sequences as Determinants of Plastid Localization

Gene constructions of the plastocyanin or ferredoxin transit sequences fused to the mature plastocyanin sequence were placed behind a cauli-

flower mosaic virus 35S promoter and transformed into tomato with the Ti plasmid (De Boer *et al.*, 1988). The fusion protein derived from the plastocyanin transit and mature sequence was detected in the chloroplast and was processed correctly. Although significant levels of plastocyanin mRNA were produced, very little of the fusion protein derived from the ferredoxin transit sequence fused to the plastocyanin sequence was detected in the chloroplast. Plastocyanin protein was also detected in nonphotosynthetic tissue in chromoplasts and leucoplasts, suggesting that all plastids have similar transport machinery (De Boer *et al.*, 1988). This is supported by recent experiments using the transit sequences of the *waxy* gene, which normally targets the protein to the amyloplast. Klosgen *et al.* (1989) found that the intact waxy precursor could be transported and processed in isolated chloroplasts from maize, pea, or potato. The waxy transit sequence also was fused to the nucleotide sequence encoding part of the waxy mature sequence and to the *gus* reporter gene. Transport and processing of the protein produced by *in vitro* translation was detected within the isolated chloroplasts. It is interesting that the efficiency of chloroplast binding and transport was better with the *gus* fusion protein than the waxy protein alone. However, the GUS activity measurements can easily be overestimated. Transport into isolated amyloplasts was attempted and, although some association of protein with the amyloplasts was detected, it could not be concluded that actual transport had occurred because intact amyloplast preparations are difficult to prepare.

Import into chloroplasts and mitochondria is similar because, in both cases, proteins are synthesized as precursors with a transit sequence on cytoplasmic ribosomes and must cross a double membrane into the organelle. As expected, fusion proteins consisting of translated nucleotides encoding chloroplast transit sequences of ferredoxin or plastocyanin fused to the sequence encoding yeast superoxide dismutase (see above) were not detected in the mitochondria. Hurt *et al.* (1986), however, showed that when part of the mitochondrial cytochrome *c* oxidase subunit IV gene was fused to sequences encoding either the LS or SS, the fusion protein was transported into the yeast mitochondria.

"Stop transfer" sequences arrest transport of signal proteins across the endoplasmic reticulum membrane. When these sequences were fused to sequences encoding the pSS transit and mature sequences and assayed *in vitro* by uptake of the fusion protein into isolated chloroplasts, no effect on import was detected (Lubben *et al.*, 1987). In agreement with studies for precursors of integral membrane proteins such as pLHCP (Schmidt *et al.*, 1981), hydrophobic stretches obviously are ignored by the chloroplast import apparatus.

III. STRUCTURAL CONSIDERATIONS

More chimeric gene constructions need to be tested to generalize about which transit peptide domains are required for transport. Additionally, most of the gene constructions have utilized only the SS transit peptide; therefore, the question of what features or domains are important for targeting to different regions of the chloroplast remains unanswered. No specific amino acids or regions of the transit peptides have been shown to be essential for transport. The three conserved domains of amino acids are detectable in some, but not all, peptides. In general, the transit sequences begin with methionine-alanine. Some transit peptides contain a glycine-glycine-X sequence near the carboxyl terminus, which has been implicated in processing (Lopez-Otin *et al.*, 1989). The mature peptide, or portions of it, may be important for transport. This inconsistency suggests that the secondary and tertiary structures of the transit peptides (or the precursor) may be important. We have studied the transit peptides for common patterns of hydrophobicity and structure. Initially, these characteristics from all the available transit peptide sequences were examined. The analysis of only six transit peptide sequences (two from stroma, two from thylakoid, and two from thylakoid lumen localized peptides) are presented (Figs. 3 and 4). The six, arbitrarily chosen, are the transit peptides of the genes encoding the pSS (consensus), a ferredoxin, the pLHCP (consensus), a Fe-S protein of the cytochrome b_6/f complex (FeS), a plastocyanin, and an oxygen-evolving complex 1 (OEE1) 33-kDa polypeptide.

A. Hydrophobicity of Selected Transit Sequences

Figure 3 displays the hydrophobicity profile of the six selected transit peptides. The six sequences were examined according to Kyte and Doolittle (1982) with the Intelligenetics Hydrophobicity program and a window of six residues. Of the >60 transit peptide sequences initially examined, only the starch phosphorylase, ADP-glucose (ADPG) pyrophosphorylase (a partial sequence, truncated on the amino-terminal end) and, to some extent, waxy and the envelope phosphotranslocator were primarily neutral to extremely hydrophilic across the entire transit peptide. All the others are hydrophobic at the amino terminus. This is the one generalization that can be made about these peptides. The central region varies in having differing lengths of both hydrophobic and hydrophilic regions. The carboxyl terminus is either hydrophobic or hy-

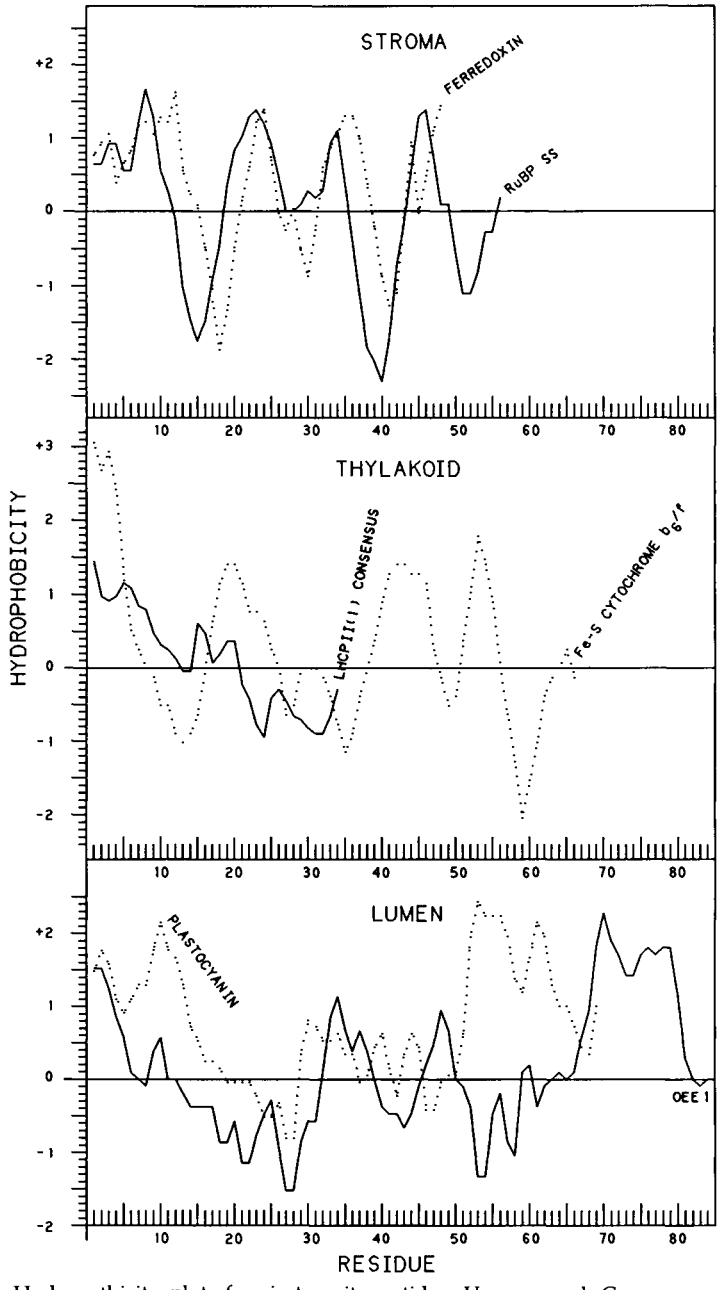


FIG. 3. Hydropathicity plots for six transit peptides. Upper panel: Compares plots obtained for the transit peptides of the pSS consensus sequence (Keegstra *et al.*, 1989) and

drophilic. The SS sequence ends with a hydrophilic region and ferredoxin with a hydrophobic region. Except for the hydrophilic region at the carboxyl terminus of pSS that is absent in that of ferredoxin, the two transit peptides have similar patterns of hydrophobicity (Fig. 3, upper panel). Examination of these plots suggests that other chimeric genes be constructed and tested. The SS transit peptide has an extended hydrophilic tail when compared with ferredoxin. It would be interesting to test the effect of an added hydrophilic tail on ferredoxin transport or, conversely, the effect of deleting the hydrophilic tail on pSS transport.

Both pLHCP II (type I consensus) and FeS, proteins that are targeted to the thylakoid, are their most hydrophobic at the amino terminus of the peptide (Fig. 3, middle panel). Both sequences also end with a hydrophilic region. Other comparisons are difficult because the pLHCP sequence is so much shorter, although it would be of interest to determine whether transport into the thylakoid is retained if the largely hydrophobic region of FeS between residues 35 and 59 are deleted. The FeS transit peptide profile actually resembles a slightly extended stroma-targeting peptide profile more than it resembles the pLHCP transit peptide profile. There are no reports using the FeS transit peptide in chimeric constructions.

The two lumen-targeting transit peptides, plastocyanin and OEE1, have hydrophobic amino-terminal regions (Fig. 3, lower panel). Although the lengths of OEE1 and plastocyanin differ, both have extremely large hydrophobic regions at the carboxyl end. A hydrophobic domain, whose presumed function is as a signal peptide for cotranslational membrane transport, is also found in precursors of lumen proteins synthesized in cyanobacteria (Kuwabara *et al.*, 1987). If the variable central region (amino acids 30–40) of OEE1 is compressed, or the hydrophilic region (amino acids 50–60) of plastocyanin deleted, these two hydrophobicity plots would be quite similar. The first 35 residues of transit sequences for OEE1 and plastocyanin resemble extended pLHCP plots,

---

ferredoxin (Smeekens *et al.*, 1985b), both of which are targeted to the stroma. Middle panel: Plots obtained for transit peptides of two thylakoid proteins, LHCP II (type I consensus sequence) (Keegstra *et al.*, 1989) and the Fe-S protein of the cytochrome b_6/f complex (Steppuhn *et al.*, 1987). Lower panel: Plots obtained for transit peptides for proteins directed to the thylakoid lumen, plastocyanin (Rother *et al.*, 1986), and *oee1* gene product, a 33-kDa polypeptide of the photosystem II water-oxidative complex (Tyagi *et al.*, 1987). The Intelligenetics Hydrophobicity program, utilizing the parameters of Kyte and Doolittle (1982) with a window of six residues, was used for the analysis. The amino acid residue numbers are plotted on the abscissa and the amount of hydrophobicity on the ordinate. Hydrophobic regions are plotted above and hydrophilic regions below the midpoint.

suggesting they could be truncated approximately at amino acid 27 and still support transport.

The transit peptide hydrophobicity plots shown in Figure 3 are calculated from the transit peptide only and not in the presence of the mature peptide. One might expect the hydrophobicity profile to vary at the carboxyl terminus in the presence of the mature peptide, because hydrophobicity is calculated over a window of several residues. We did not detect any large differences in hydrophobicity profiles of the transit sequences of the six genes when analyzed alone or in the presence of approximately 20 amino acids of the mature peptides (data not shown). The hydrophobicity plots do not show a consistent pattern within any class of transit peptides (i.e., all that direct peptides to the stroma, thylakoid, or thylakoid lumen), although the three classes do all resemble each other, particularly in that the amino-terminal region is hydrophobic. The overall pattern of hydrophobicity does suggest other chimeric and mutated genes be constructed to determine the importance of hydrophobic and hydrophilic domains. Proteins that are truncated at the carboxyl terminus and in the central region also are of interest, although problems may be encountered by altering the size of the peptides. Neutral, or filler, regions may need to be inserted in these constructs.

B. Secondary Predictions for Selected Transit Peptides

We examined the transit peptides for common structural features such as α-helices, β-sheets, β-turns, and coils. In all cases, the parameters used were those of Chou and Fasman (1978). This was done for all available transit sequences initially with the Intelligenetics Structure program. This program, however, did not resolve each residue to a single structure; therefore, the six transit peptides were also analyzed using the Wisconsin GCG Peptide Structure program. Each residue is analyzed and plotted according to its ability to form α-helices, β-sheets, or β-turns. From these plots, each amino acid residue for the six transit peptides was assigned a single structure. The structures for the six transit peptides are shown in Figure 4. There is no obvious consensus structure. We expected all six to have a conserved structure at the amino terminus because all are hydrophobic in this region. Five of the peptides begin with an α-helix, while the amino-terminal region of the FeS transit peptide, although being very hydrophobic, forms a β-sheet. We have not examined all of the mature peptide structures in

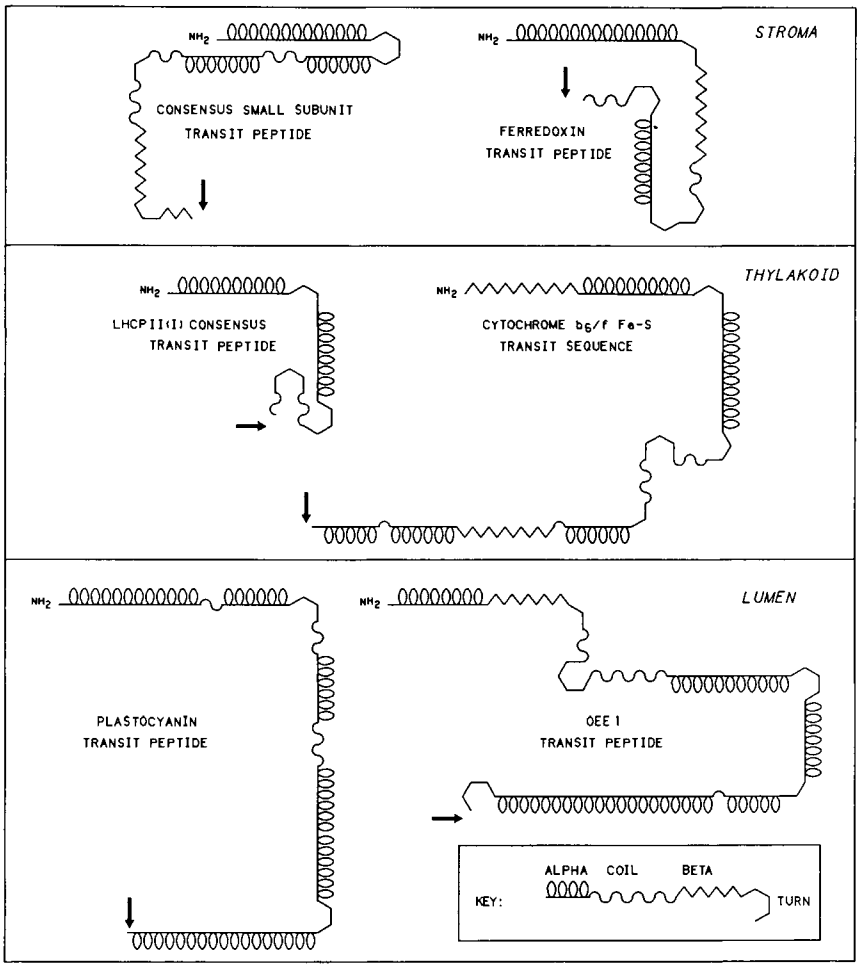


FIG. 4. Predicted secondary structures of six transit peptides. Transit peptides are those described in Figure 3. Structures were determined with the Wisconsin GCG Peptide Structure program utilizing the parameters of Chou and Fasman (1978) and plotted according to the key shown at the bottom. The number of amino acids that participate in generating each structural entity correspond to the number of unit symbols for the α-helices, β-sheets, coils, and β-turns.

conjunction with the transit peptides. The presence of the mature sequence may alter the secondary structure predictions of the transit peptide, although we found that the predicted structures of the six transit peptides in Figure 4 were not affected by the inclusion of the mature peptide.

IV. TRANSIT PEPTIDASES

The enzymes responsible for maturation of the precursors of imported proteins in chloroplasts are presumably essential for establishing the functional integrity of most plastid complexes. Only one enzyme has been reported that possesses activity in its precursor form (EPSPS) (della-Cioppa *et al.*, 1986). However, in our laboratory Cammarata (unpublished) determined that precursors of the apoproteins of the light-harvesting complexes can assemble pigments *in vitro*. Therefore, transit sequences may not cause large changes in the tertiary structure of the mature portions of chloroplast precursors.

Based on analyses conducted so far and conjectures concerning biogenetic pathways, we suppose that there are at least three to four classes of transit peptidases. These include one or two species of stromal enzyme that participate in maturation steps for stromal as well as certain thylakoid proteins and, through the formation of maturation intermediates, proteins that are imported into the thylakoid lumen. The second class are those that reside in the thylakoid membrane and function in the maturation of lumenal proteins and, perhaps, thylakoid proteins such as D1, the 42-kDa chlorophyll *a*-binding protein of PS II, and cytochrome *f*, each of which is synthesized as a precursor within chloroplasts (Grebanier *et al.*, 1978; Michel *et al.*, 1988; Mattoo *et al.*, 1989). The final transit peptidase class would be responsible for maturation of precursors to proteins of the chloroplast envelope and its intermembrane space. This is a subset of plastid proteases that have been incompletely characterized even though these enzymes are immensely important in regulating the subunit stoichiometry for multisubunit complexes and/or provide housekeeping functions in chloroplasts (Mishkind *et al.*, 1985a).

V. STROMAL TRANSIT PEPTIDASES

The most thorough work is on the soluble SS transit peptidase localized in the chloroplast stroma. Its activity was first analyzed in whole cell extracts of *Chlamydomonas reinhardtii* by Dobberstein *et al.* (1977) using *in vitro* synthesized ribulose bisphosphate carboxylase SS as a substrate. In addition to demonstrating that it processes pSS to its mature size via endoproteolysis, Dobberstein *et al.* (1977) also determined that it was sensitive to reagents that modify sulfhydryl groups but not metal che-

lators. The ability of the crude preparation to achieve precise maturation was demonstrated through microsequencing of the processed precursor by Schmidt *et al.* (1979). Schmidt *et al.* (1984) showed that the transit peptidase in *Chlamydomonas* cell extracts also matures precursor of nuclear-encoded ribosomal proteins *in vitro*, but these workers also detected potential intermediates in the maturation of L-21 and L-18.

It is interesting that the algal enzyme differs considerably from the counterpart of vascular plants. As shown by Robinson and Ellis (1984a), the counterpart from vascular plants is insensitive to sulfhydryl modifiers and, instead, has a requirement for heavy metal ions. Thus, in the evolution of transit peptides and transit peptidases, simultaneous changes in the enzyme and the processing site that it acts upon apparently have occurred. Also, it is both interesting and perplexing that maturation of SS precursors can be accomplished in one step when substrates produced *in vitro* are used in transit peptidase assays with preparations of soluble proteins. In contrast, import of algal precursors into vascular plant chloroplasts and maturation of precursors with total lysates of chloroplasts have enabled detection of maturation intermediates (Robinson and Ellis, 1984b; Mishkind *et al.*, 1985b). These observations have led to a premise that maturation of stromal proteins occurs in two steps. Microsequencing data have indicated that the first maturation site as well as the initial transit peptidase has been evolutionarily conserved (Mishkind *et al.*, 1985b). However, Ostrem *et al.* (1989) recently argued that the processing intermediates arise from a subordinate processing step that becomes the recourse if normal, one-step processing is impaired.

The stromal transit peptidase of pea chloroplasts has a native size of 180 kDa and exhibits a pH optimum of 8.5–9 (Robinson and Ellis, 1984a). It acts on pSS and the precursors of ferredoxin, ferredoxin–NADP-reductase, and plastocyanin, processing the last to a maturation intermediate that undergoes transport through thylakoid membranes where, as described below, a second maturation step occurs (Hageman *et al.*, 1986). Recently, Musgrove *et al.* (1989) characterized three other soluble endopeptidases, resolved by DEAE-Sephacel and Phenyl-Sepharose chromatography, that act on preplastocyanin. The function of these enzymes is unclear.

Lamppa and Abad (1987) examined processing of the 31-kDa pLHCP of wheat by lysed chloroplasts from pea and wheat seedlings. Processing by lysed chloroplasts results in the formation of a 25-kDa product, whereas processing upon import into intact chloroplasts results in 25- and 26-kDa LHCP polypeptides. They detected no processing activity by the membrane fractions of lysed chloroplasts. Time-course assays did

not reveal the appearance of processing intermediates, although the maturation of pLHCP is relatively slow and never was completely accomplished. Subsequent analyses (Abad *et al.*, 1989) showed that the processing activity is directed at the level of the transit peptide. Like pSS transit peptidase, the pLHCP transit peptidasse is insensitive to iodoacetate and EGTA but is inhibited by EDTA and 1,10-phenanthroline. From gel filtration analyses, it was estimated that the enzyme has a molecular size of approximately 240-kDa. Abad *et al.* (1989) also characterized processing of pSS by wheat and pea chloroplasts. Pea chloroplast extracts apparently do not mature the wheat pSS, although they efficiently mature wheat pLHCP. This result indicates that the enzymes for pSS and pLHCP are different; however, the activities for maturation of pSS and pLHCP coelute from gel filtration columns. Moreover, the products of gene constructs possessing pea transit peptides on wheat mature SS were matured by pea chloroplast extracts, indicating the transit peptide is of major importance in the recognition of precursors by transit peptidases.

A. One or Two Steps in the Maturation of Precursors of Stromal Proteins?

The possibility that maturation of the pSS is matured in two steps, culminating in the formation of the mature 14-kDa polypeptide, was first evident from time-course studies of precursor processing by pea chloroplast extracts (Robinson and Ellis, 1984b). They also pretreated 20-kDa pSS with iodoacetate to achieve carboxymethylation of the cysteine at the junction of the transit peptide and the mature SS sequence. After incubation with purified transit peptidase, an 18-kDa intermediate was the only proteolytic product. The main conclusions of this work are that maturation of pSS occurs in two steps and, because no separation of activities could be achieved, that the same protease is responsible for both proteolytic events. Anderson and Smith (1986) showed that stromal extracts do not convert transit peptides, generated by *in vitro* transcription–translation, to intermediates. Thus, domains that include portions of mature SS are also necessary for transit peptidase recognition in the process of formation of maturation intermediates.

Two steps, however, are not obligatory for maturation of pSS *in vitro*. Dobberstein *et al.* (1977) showed that a 14-kDa fragment, representing the intact transit peptide, was removed by incubation of antibody-bound pSS with *Chlamydomonas* cell extracts. With a different SS antibody preparation, a different result was obtained by Marks *et al.* (1985):

Algal cell extracts generated an 18.5-kDa intermediate between the 22-kDa precursor and the 16.5-kDa mature subunit. Moreover, the maturation intermediate is the exclusive product when iodoacetate is included in the processing mixtures. The reagent apparently modifies the cysteine adjacent to the maturation site of the transit peptide, resulting in an inability of the final transit peptidase to effect cleavage. Gupta and Beevers (1987) immunoprecipitated a 65-kDa intermediate of the 62-kDa nitrite reductase of pea following incubation of the 67-kDa precursor with pea chloroplast extracts. In this instance, complete precursor maturation was difficult to achieve.

Schmidt *et al.* (1984) detected intermediates upon *in vitro* maturation of *Chlamydomonas* ribosomal proteins L-21 and L-18 with algal cell extracts. Subsequently, Schmidt *et al.* (1985) studied chloroplast ribosome biogenesis by pulse-labeling cells undergoing nitrogen-dependent differentiation from gametes to vegetative cells. In control vegetative cells, a polypeptide of 17 kDa could be immunoprecipitated from the soluble proteins labeled with $^{35}SO_4^-$ for 5 min with L-18 antibody; during the subsequent chase, this disappeared with a coincidental increase in the mature L-18 of 15.5 kDa. However, the prevalence of the putative intermediate produced by partial maturation of the 18.5-kDa L-18 precursor was enhanced in cells differentiating from gametes into vegetative cells. Moreover, its conversion to mature L-18 was much slower than usual, indicating gametes possess reduced amounts of the protease that generates mature L-18. Another striking finding in the study of Schmidt *et al.* (1985) is that inhibition of chloroplast protein synthesis for 3 hr by lincomycin eliminates maturation of the L-18 intermediate. From the earlier studies, it was established that pre-L-18 can be completely matured in the absence of chloroplast ribosome assembly. Hence, Schmidt *et al.* (1985) concluded that, of the two transit peptidases required for L-18 maturation, the second is synthesized within the chloroplast. None of the other chloroplast ribosomal proteins were impaired in maturation after prolonged inhibition of chloroplast protein synthesis. Unfortunately, it was not determined if extracts from lincomycin-treated cells actually are deficient in an L-18 transit peptidase.

Hayden *et al.* (1986) discovered that maturation of LHCP precursors may be temperature-sensitive in some plant species. They observed that pLHCP accumulates in thylakoids when maize leaves are exposed to high intensities of light at 5°C for 6 hr. Thus, a 31-kDa polypeptide was detected in stained profiles and immunoblots of thylakoids as well as purified LHCP complexes. The exposure to dual stress led to modification of antenna properties without changes in pigment content. They concluded that a transit peptidase is inactivated by photoinhibition at

low temperatures. However, the precursor (or a maturation intermediate) can insert into thylakoids and assemble with other mature light-harvesting pigment–protein complexes. Cooper and Ort (1987) found that insertion of LHCP precursors also may be sensitive to chilling in tomato. Plants that were subjected to 2–4°C for 16 hr in darkness and then rewarmed in the light synthesize a 35-kDa polypeptide that accumulates in the stroma during the first 2 hr. Peptide mapping reportedly demonstrated primary structure homology between the 27-kDa LHCP apoproteins and the 35-kDa stromal protein. Both studies indicate that the transit peptidase may be cold- and light- (and/or dark-?) sensitive. Because no *in vitro* assays for transit peptidase activity were made, it is not clear whether the effects are exerted on processing enzymes or are due to irreversible changes in the precursor substrates.

B. Thylakoid Transit Peptidases

Characterization of transit peptidases of the thylakoid lumen was initiated by Hageman *et al.* (1986). Using preplastocyanin, produced by *in vitro* transcription–translation, a protease that completes maturation of this precursor was solubilized from thylakoids by Triton X-100. However, the effectiveness of the membrane protease was significantly enhanced by the provision of a stromal transit peptidase to generate processing intermediates. Kirwin *et al.* (1987) partially purified the thylakoid enzyme and studied its specificity and sensitivity to inhibitors. The size of the enzyme could not be determined and of several inhibitors tested, including phenylmethylsulfonyl fluoride (PMSF) leupeptin, sulfhydryl modifying reagents, and benzamidine, none were effective in blocking processing. Also, EDTA and EGTA stimulated rather than inhibited plastocyanin maturation. Musgrove *et al.* (1989) mention that 1,10-*o*-phenanthroline also stimulates this protease. A subsequent study (Kirwin *et al.*, 1989) reports that the precursor of the PS II oxygen-evolving complex is also processed by the purified thylakoid protein.

C. Transit Peptidase Recognition Sites

Based on sequence analysis of a limited number of chloroplast precursors, final maturation in vascular plants apparently occurs following a small, basic, hydrophobic amino acid series (Schmidt and Mishkind, 1986; Smeekens *et al.*, 1985b). As can be determined from Table I, the idea of a consensus primary sequence no longer seems valid. If one supposes that secondary structure determines maturation sites, no univer-

sal rule can be invoked (Fig. 4). Hence, the specificity as well as the number of transit peptidases is a void in our knowledge about protein import. Moreover, accumulating evidence indicates that sequences in the mature proteins are important in directing precursor maturation. For example, Abad *et al.* (1989) found that deletion of 91 amino acids from the C-terminus of pLHCP from wheat causes loss of the ability of the *in vitro* transcription–translation product to undergo maturation by wheat chloroplast extracts, and, reportedly, this construct could not be imported into chloroplasts. This effect was ascribed to major changes in precursor conformation. On the other hand, transit sequences generally are removed from imported chimeric proteins produced by fusing transit sequences to mature sequences of nonchloroplast proteins. Unless there are fortuitous features of the "passenger" proteins, these findings indicate that transit sequences are direct targets of the maturases.

D. Degradation of Transit Peptides

Analyses of chloroplast extracts following *in vitro* import of radiolabeled precursors have not enabled detection of the transit peptide or its fragments after maturation of the precursors. Anderson and Smith (1986) found that incubation of the translation product synthesized from an *in vitro* transcript of a gene fragment corresponding to the transit peptide plus six amino acids of the mature SS is rapidly degraded by stromal extracts of pea chloroplasts. The only detectable intermediate of the degradation process had a slightly increased electrophoretic mobility relative to that of the transit peptide itself; this was argued to be the result of removal of the mature SS sequences by transit peptidase. Time-course measurements showed that both the transit peptide and the initial proteolysis product disappeared rather rapidly with apparent half-lives of 5 min. Degradation was found to be inhibited by EDTA and 1,10-*O*-phenanthroline. Although iodoacetate blocked appearance of the intermediate, presumably by modifying the cysteine adjoining the precursors maturation site, the reagent had no effect on transit peptide degradation. These workers could not determine whether or not transit peptides themselves are imported by isolated chloroplasts.

E. Tissue Specificity of Transit Peptidases

All available evidence indicates that the genes encoding transit peptidases are expressed in all plant cell types. Thus, leucoplasts from the endosperm of developing castor oil seeds (Boyle *et al.*, 1986), organelles

that are specialized for fatty acid synthesis, import and mature pSS. Recently, Kirwin *et al.* (1989) reported low amounts of a soluble transit peptidase in wheat germ extracts; it is extremely effective in processing the precursor of the 33-kDa polypeptide of the PS II oxygen-evolving complex to a maturation intermediate. They suggest the enzyme is of proplastid origin.

VI. INSERTION OF THYLAKOID PROTEINS

A. Biogenesis of LHCPs

The first studies of the import of LHCP precursors into isolated chloroplasts indicated that the process of maturation and pigment binding are very efficient (Schmidt *et al.*, 1981). Pigment binding, at least partially, is facilitated by the occurrence of a pool of accessible chlorophylls in thylakoids because import, insertion, and formation of pigment-protein complexes can be achieved with chloroplasts kept in darkness (Schmidt *et al.*, 1981). However, with chloroplasts from *Lemna* and maize but not barley, LHCP integration with pigments in membranes requires supplements of δ-aminolevulinic acid and *S*-adenosylmethionine as chlorophyll biosynthesis precursor and cofactor, respectively (Kohorn *et al.*, 1986; Chitnis *et al.*, 1986). LHCP apoproteins characteristically are heterogeneous in size due to the occurrence of multigene families (Dunsmuir, 1985) but also are a result of the existence of alternative maturation sites (Kohorn *et al.*, 1986). Thus, pLHCP produced by *in vitro* transcription–translation was observed to be converted into as many as three distinct apoproteins upon incubation with isolated chloroplasts (Kohorn *et al.*, 1986; Kohorn and Tobin, 1986).

Viitanen *et al.* (1988) and Lamppa (1988) established that the transit sequence of pLHCP is not essential for thylakoid integration. They constructed a gene encoding the mature LHCP apoprotein sequence but possessing a transit sequence from wheat pSS. Upon import of the *in vitro* transcription–translation product into pea thylakoids, the mature LHCP was recovered exclusively in thylakoids.

The sequences required for insertion of LHCP apoproteins into thylakoids have been investigated by deletion analysis of one *cab* gene from *Lemna* (Kohorn *et al.*, 1986). The analysis is somewhat complex because, until recently, the assay also entailed envelope transport. The deletions of the mature apoprotein that are the most deleterious to membrane

insertion are those toward the carboxyl terminus. One of the essential regions required for insertion (and/or protection against degradation by plastid proteases) is the last of the three (or four) membrane-spanning domains and the other is a hydrophilic sequence that is immediately proximal on the amino-terminal side. Site-directed mutagenesis of a *Lemna cab* gene has led to identification of a lysine in the last membrane-spanning domain and a glutamate in the first membrane-spanning domain that is essential for assembly of LHCP apoproteins into a pigment-protein complex. Moreover, an arginine in the first membrane span was found to be essential for accumulation of LHCP in isolated chloroplasts, although the mutant precursor binds to chloroplast envelopes. Interestingly, changing a glutamate to arginine (first membrane domain) and arginine to glutamate (last membrane domain) enables assembly of pigments in the *in vitro* system (Kohorn and Tobin, 1987). Cline *et al.* (1989) also demonstrated that the carboxyl domains of LHCP are necessary for integration in chloroplast lysates. Thus, the absence of integration by C-terminal deletion derivatives of pLHCP observed by Kohorn *et al.* (1986) were not due necessarily to a block at the level of transport through the envelope membranes.

B. Thylakoid Insertion Factors

In a study of the import of *Lemna* and barley pLHCP into developing barley chloroplasts, Chitnis *et al.* (1986, 1988) found that maturation, but not membrane integration or association with oligomeric LHC pigment-protein complex (a trimer of LHCP proteins), depends on the stage of plastid development. Only after 5 hr of greening of 6-day-old barley seedlings are plastids capable of processing pLHCP. These data indicate that maturation of pLHCP can occur after thylakoid insertion and assembly into an LHC complex, and this process depends on light-induced factors; however, it was not possible to determine if pLHCP itself is capable of pigment binding. Surprisingly, fully green seedlings do not support stable accumulation of *Lemna* LHCP in either the precursor or mature forms as determined by *in vitro* transport assays.

Cline (1986) attempted to determine if pLHCP could be inserted directly into thylakoids. Hybrid-selected mRNA was translated *in vitro* and incubated with chloroplast lysates. Thylakoid insertion of pLHCP is maximal when osmoticum is low (0.11 *M* sorbitol), ATP is provided, and the mixture is buffered to pH 8.0, conditions that diminish thylakoid stacking. Subsequently, Cline (1988) found that incubation conditions favoring retention of grana stacks influence pLHCP integration and

maturation in chloroplast lysates. In chloroplasts prepared with a high osmoticum (0.33 *M* sorbitol) and 10 m*M* $MgCl_2$, enhanced maturation of the precursor occurred and light could substitute for ATP in the insertion steps. Digitonin fractionation of thylakoids also revealed that residual pLHCP was capable of translocation into appressed membranes. Hence, the removal of the transit sequence may not be a requirement for protein movement to the thylakoid stacks. It was not shown that the precursors are within the stacks but the effects of exogenous proteases seem to indicate the precursor is on the stroma-exposed regions of grana. More importantly, Cline (1988) demonstrated that a stromal factor is required for membrane integration of pLHCP and many other thylakoid proteins produced upon *in vitro* translation of total leaf poly(A) mRNA. As mentioned above, Hayden *et al.* (1986) obtained results from *in vivo* studies that support the inherent ability of pLHCP to be inserted into thylakoids. They discovered that chilling maize leaves at 5°C at high light intensities for 6 hr leads to accumulation of a higher-molecular weight thylakoid polypeptide that is recognized by LHCP antibodies. Cooper and Ort (1987) reported similar effects of chilling on the biogenesis of LHCP in tomato but, in this case, the precursor is not integrated into thylakoids.

Chitnis *et al.* (1987) extended the analysis of LHCP thylakoid assembly during plastid development. Upon import into barley plastids at the early stages of development, most of the *Lemna* pLHCP is associated with thylakoids of the recipient plastids but it is removed by NaOH washing. They also studied chloroplast lysates and confirmed the observation of Cline (1986) that ATP and a stromal factor are required for thylakoid insertion. Chitnis *et al.* (1987) suggested that the stromal factor, which is maximally abundant at 12 hr of greening of barley seedlings, might be required for palmitoylation of pLHCP (Mattoo and Edelman, 1987) or for changing the precursor's conformation, either of which might alter the hydrophobicity to facilitate membrane insertion. Partial characterization of the stromal factor has indicated it separates as a 65-kDa on a gel filtration column and does not contain an RNA moiety (Fulson and Cline, 1988).

The most recent work on pLHCP import and thylakoid integration has shown that a membrane potential is necessary for the integration step. Uncoupler treatments lead to accumulation of mature LHCP in the stroma (Cline *et al.*, 1989), indicating also that the stroma is the intermediate location of the mature apoproteins en route to thylakoids and maturation can occur in this compartment prior to membrane insertion. On the other hand, Chitnis *et al.* (1988) conducted pulse-chase experi-

ments with chloroplasts from barley seedlings that had been exposed to light for 8 hr and found that pLHCP maturation can also occur after thylakoid integration.

The consensus from the LHCP import studies is that precursor import is followed by transport from the envelope through the stroma and then thylakoid integration occurs. A stromal factor is required for LHCP or pLHCP integration, and both forms of the protein can be assembled with other LHCP components in the membrane. The site of precursor maturation is not absolute; the soluble pLHCP transit peptidase apparently can act on either soluble or membrane-bound forms of the precursors.

C. Thylakoid Membrane Transport

Plastocyanin and polypeptides of the PS II water-oxidation complex are processed in the stroma to maturation intermediates that possess the hydrophobic domain at the carboxyl termini of their transit peptides (see above). Subsequently, the intermediates are translocated through the thylakoid lumen where their maturation is completed, as discussed above. Recently, Kirwin *et al.* (1989) succeeded in the *in vitro* reconstitution of thylakoid membrane transport using the precursor and processing intermediate of the 33-kDa water-oxidation polypeptide, the product of the *oee1* nuclear gene, as a substrate. Previous attempts to obtain thylakoid preparations that support transport into the lumen apparently failed because preplastocyanin and its intermediates are translocated poorly. However, the 33-kDa protein, probably as a maturation intermediate, is specifically imported into the lumen by a transport system that is stimulated by ATP.

VII. RECEPTORS

The receptors of the envelope membranes for precursor import have been studied by (1) analysis of the saturation kinetics of binding, (2) assessment of the effects of inhibitors and proteases, (3) solubilization of binding protein, (4) cross-linking of precursors to envelope proteins, and (5) studies with anti-idiotypic antibodies to portions of transit sequences. A second receptor class, functioning in import into the thy-

lakoid lumen of processing intermediates of plastocyanin and polypeptides of the oxygen-evolving complex, should exist in thylakoids, but none of its characteristics are known. Perhaps even more receptors of thylakoids facilitate insertion of proteins such as LHCPs into the lipid bilayer, but no evidence indicates that they exist.

Although transit sequences usually possess a net positive charge that might confer an affinity of precursor for the phospholipid components of the outer envelope, early studies noted that protease treatment of isolated chloroplasts destroys import capability, presumably by protein receptor degradation (Chua and Schmidt, 1978; Cline *et al.*, 1985). More recently, it was demonstrated that the receptors are also inactivated by *N*-ethylmaleimide (Friedman and Keegstra, 1989). For reasons of parsimony in the membrane transport mechanism, it was suggested that import receptors occur at contact sites of the outer and inner membranes (Dobberstein *et al.*, 1977; Chua and Schmidt, 1978). Subsequently, many strategies have been devised for identification of a proteinaceous envelope receptor. Localization of receptors at contact sites is still attractive, and supporting evidence for this has been obtained for chloroplasts (Pain *et al.*, 1988) as well as mitochondria (Schleyer and Neupert, 1985). Bitsch and Kloppstech (1986) used increasing concentrations of nonionic and zwitterionic detergents to progressively solubilize binding sites and then restored precursor binding to the residual membranes after removal of the detergents by dialysis. Their data indicate that binding sites are very resistant to solubilization, an expected characteristic of components at contact sites.

With isolated chloroplast envelopes used in precursor binding at 4°C, Pfister *et al.* (1982) calculated that there are 625–3,000 binding sites per chloroplast. This is close to the 1,500 receptors per plastid as determined by Cline *et al.* (1985), who employed nigericin to block binding of pSS and pLHCP produced by *in vitro* translation of hybrid-selected mRNA. Cline *et al.* (1985) also examined the effects of pretreating chloroplasts with thermolysin. An intriguing aspect of their data is that binding and import of LHCP precursors is much more sensitive to protease pretreatment than is that of pSS. As suggested previously (Schmidt and Mishkind, 1986), this could indicate the existence of separate receptors or transport complexes for the stromal versus thylakoid membrane proteins. In support of this idea, Bitsch and Kloppstech (1986) mention that the pH optima for envelope binding is different for pSS and pLHCP. In the most recent study of precursor binding (Friedman and Keegstra, 1989), binding assays at 0°C led to an estimate of 1,500–3,000 receptors for pSS in a pea chloroplast envelope. They also calculated that pSS binding exhibits high affinity (dissociation constant of 6–10 n*M*), 30

times higher than that of *Neurospora* mitochondrial binding measurements for apocytochrome *c* (Henning *et al.*, 1983). Assuming the receptor has a size of 60 kDa (see below), Friedmann and Keegstra (1989) calculated that the receptor constitutes 0.3% of the total protein of the inner and outer envelope membranes. Correspondingly, an average of 25 receptors per thin cross-section should be presented during electron microscopy.

Photoactivatable cross-linking reagents have been employed in two studies of chloroplast receptors. The strategy of Cornwell and Keegstra (1987) was to modify pSS, produced by *in vitro* translation of hybrid-selected mRNA, with ^{3}H-*p*-azidophenylacylbromide in darkness. Following incubation with chloroplasts, washing to remove unbound precursor, and UV-irradiation to effect cross-linking to proteins associated with pSS, radioactivity was most intense in an 86-kDa conjugate. Accounting for the size of pSS, it was concluded that a 66-kDa polypeptide (or a conjugate of more than one cross-linked polypeptide) represents a major component of the envelope receptor. Cornwell and Keegstra (1987) note that there is no outer membrane protein of this size, but this could be due to the receptor's low abundance. Kaderbhai *et al.* (1988) used a similar approach but esterified a photoactivatable cross-linking reagent to a synthetic peptide corresponding to the first 24 amino acids of the wheat pSS. A bewildering aspect of this paper is that the ^{125}I-iodination of the peptide was apparently accomplished by the chloramine T method, but the sequence contains no tyrosines, histidines, cysteines, or tryptophans as labeling substrates. Nevertheless, upon presentation to envelopes from pea chloroplasts, UV cross linking led to recovery of the transit peptide with a 52-kDa protein of the outer envelope and a 30-kDa polypeptide of the inner membrane. The interpretation of two-dimensional gel electrophoresis results was that the inner membrane protein must be equivalent to the phosphate translocator, the most abundant protein of chloroplast envelopes. Although the 30-kDa region of the two-dimensional gels reveals a heterogeneous smear of polypeptide(s), Kaderbhai *et al.* (1988) elected to hypothesize that the major inner envelope protein has dual functions: transport of phosphorylated three-carbon carbohydrates and import of cytosolic precursors.

An elegant but controversial study of chloroplast receptors is the work of Pain *et al.* (1988). Using a synthetic peptide corresponding to the 30 C-terminal residues of the transit sequence of pea pSS, a second set of antibodies to the anti-transit sequence immunoglobulins were generated. As the second, anti-idiotypic antibodies might recognize transit peptide-binding sites on envelopes, these were used in immunolocali-

zation and immunoblot analyses and also as potent inhibitors of pSS import into isolated chloroplasts. Immunolocalization at the light microscopy level indicated that presumed receptors are clustered in envelopes and, at the electron microscopy level, these are found at inner–outer membrane contact sites. The antibody recognized two chloroplast polypeptides: the major 30-kDa protein of envelopes and, in the stroma, a protein assumed to be the LS of ribulose bisphosphate carboxylase. The latter finding is disconcerting because binding sites for transit sequences by LS is surprising. Presumably, prior to holoenzyme assembly, the SS is mature. However, pSS might associate with a "chaperonin" before interacting with the ribulose bisphosphate carboxylase LS, which also interacts with chaperonin (Landry and Bartlett, 1989). If this is true, chaperonin possibly interacts with the precursor forms of imported proteins at the level of transit sequences. Therefore, the 52-kDa protein that was detected with anti-idiotypic antibody by Pain *et al.* (1988) actually might be chaperonin of 60 kDA (Roy *et al.*, 1988; see also Chapter 10B) rather than LS.

VIII. ENERGY REQUIREMENTS

Import of pSS into spinach chloroplasts was observed to occur in darkness without energy supplements (Chua and Schmidt, 1978), but Grossman *et al.* (1980) documented that chloroplasts from pea seedlings required the addition of ATP to support protein translocation. It is now firmly established that ATP, and not an envelope membrane potential, is required for protein import (Grossman *et al.*, 1980; Cline *et al.*, 1985). However, the site and number of transport components for hydrolysis of ATP has been difficult to resolve. Schindler *et al.* (1986), Pain and Blobel (1987), and Theg *et al.* (1989) have employed inhibitors, uncouplers, phosphorylated substrates, and exogenous enzymes and nucleotides to block envelope ATPase, to consume external ATP, or to raise the internal ATP concentration. The effects of such manipulations on pSS import indicate that ATP is consumed inside. Experiments of similar design by Flügge and Hinz (1986) lead to the opposite conclusion and suggest that external ATP actually is utilized in the space between the inner and envelope membranes. Subsequently, Hinz and Flügge (1988) provided evidence that the ATP–inorganic phosphate translocator is not involved in protein import: Low concentrations of pyridoxal phosphate completely block import but barely affect activity of the translocator. On

the other hand, a direct correlation was found for the concentration dependence of pyridoxal phosphate inhibition of import and inhibition of dephosphorylation of a 51-kDa envelope. Schindler and Soll (1986), Schindler *et al.* (1986), and Flügge and Hinz (1986) also found that fluoride and molybdate, inhibitors of protein phosphatases, were potent import inhibitors. Moreover, Hinz and Flügge (1988) determined that there must be two ATP-dependent steps in transport because the K_M for phosphorylation of the 51-kDa polypeptide, assuming this event is essential for protein import, was determined to be 5 μM, whereas the apparent K_M for ATP for protein import was estimated to be 900 μM. Recently, Olsen *et al.* (1989) also identified two ATP-dependent steps through studies involving thorough depletion of ATP from cell-free translation mixtures used in pSS synthesis. A high-affinity ATP site, requiring 50–100 μM ATP, promotes precursor binding to chloroplast envelopes, whereas protein translocation itself requires millimolar concentrations of ATP.

Schindler and Soll (1986), Hinz and Flügge (1988), and Schindler *et al.* (1986) proposed that protein phosphorylation–dephosphorylation participates in protein import, perhaps by promoting precursor conformational changes. Imported proteins probably are unfolded during envelope transport, as evidenced by the demonstration of della-Cioppa and Kishore (1988) that glyphosate binding to the precursor of EPSP inhibits chloroplast import. Soll (1988) has characterized an outer envelope protein kinase from pea chloroplasts; it is 70 kDa and, with histone as a substrate, displays a K_M of 450 μM for ATP. If this is an essential component for protein import, antibodies against the kinase should inhibit protein transport.

IX. PROSPECTS

Primary and secondary structure analyses have failed to identify the functional domains of transit sequences for import. Equally disappointing, estimations of hydropathicity of transit sequences reveal little uniformity for transit peptides except for the C-terminal hydrophobic domain characteristic of proteins that are imported into the thylakoid lumen. Surprisingly, no consensus sequences can be identified that serve as recognition signals for transit peptidases. Although deletion analyses of precursors undoubtedly will continue, it is justifiable to doubt whether or not transit peptide functions will be defined by *in vitro*

mutational analyses. Clearly, the future of understanding of precursor import hinges on crystal structure analysis of several classes of precursors. As techniques for production proteins in transformants of *Escherichia coli* are progressively improving, this is not unfeasible.

The energy requirements and receptors for import still are unclear, especially how ATP hydrolysis is coupled to the process of protein transport in the case of the envelope and how a membrane potential facilitates insertion of proteins into thylakoids. As suggested above, analyses with antibodies against ATPases and putative receptor components might behave as specific inhibitors of protein import but also might be useful in immunopurification of import components. Finally, it was suggested previously that mutants of plants and algae with conditional defects in protein import and maturation should be sought (Schmidt and Mishkind, 1986). The validity of the suggestion has not been diminished by research efforts in the past few years. Analyses with such mutants undoubtedly will elucidate facets of the import process that are escaping detection by conventional *in vitro* transport studies.

ACKNOWLEDGMENTS

We are grateful to Valery Maples and Dr. Brigitte Bruns and Dr. Kirk Cammarata for their editorial help. This review was supported in part by funds from the McKnight Foundation and National Science Foundation.

REFERENCES

Abad, M. S., Clark, S. E., and Lamppa, G. K. (1989). Properties of a chloroplast enzyme that cleaves the chlorophyll a/b binding protein precursor. Optimization of an organelle-free reaction. *Plant Physiol.* **90,** 117–124.

Anderson, S., and Smith, S. M. (1986). Synthesis of the small subunit of ribulose 1,5-bisphosphate carboxylase from genes cloned into plasmids containing the SP6 promoter. *Biochem. J.* **240,** 709–715.

Back, E., Burkhart, W., Moyer, M., Privalle, L., and Rothstein, S. (1988). Isolation of cDNA clones coding for spinach nitrite reductase: Complete sequence and nitrate induction. *Mol. Gen. Genet.* **212,** 20–26.

Bedbrook, J. R., Smith, S. M., and Ellis, R. J. (1980). Molecular cloning and sequencing of cDNA encoding the precursor to the small subunit of chloroplast ribulose-1,5-bisphosphate carboxylase. *Nature (London)* **287,** 692–697.

Berry-Lowe, S. L., McKnight, T. D., Shah, D. M., and Meagher, R. B. (1982). The nucleo-

tide sequence, expression, and evolution of one member of a multigene family encoding the small subunit of ribulose-1,5-bisphosphate carboxylase in soybean. *J. Mol. Appl. Gen.* **1,** 483–498.

Bitsch, A., and Kloppstech, K. (1986). Transport of proteins into chloroplasts—Reconstitution of the binding capacity for nuclear-coded precursor proteins after solubilization of envelopes with detergents. *Eur. J. Cell Biol.* **40,** 160–166.

Blobel, G., and Dobberstein, B. (1975). Transfer of proteins across membranes. *J. Cell Biol.* **67,** 835–851.

Boyle, S. A., Hemmingsen, S. M., and Dennis, D. T. (1986). Uptake and processing of the precursor to the small subunit of ribulose 1,5-bisphosphate carboxylase by leucoplasts from the endosperm of developing castor oil seeds. *Plant Physiol.* **81,** 817–822.

Brisson, N., Giroux, H., Zollinger, M., Camirand, A., and Simard, C. (1989). Maturation and subcellular compartmentation of potato starch phosphorylase. *Plant Cell* **1,** 559–566.

Buvinger, W. E., Michel, H., and Bennett, J. (1989). A truncated analog of a pre-light harvesting chlorophyll a/b protein II transit peptide inhibits protein import into chloroplasts. *J. Biol. Chem.* **264,** 1195–1202.

Cashmore, A. R. (1984). Structure and expression of a pea nuclear gene encoding a chlorophyll a/b-binding polypeptide. *Proc. Natl. Acad. Sci. USA* **81,** 2960–2964.

Cashmore, A. R., Broadhurst, M. K., and Gray, R. E. (1978). Cell-free synthesis of leaf protein: Identification of an apparent precursor of the small subunit of ribulose-1,5-bisphosphate carboxylase. *Proc. Natl. Acad. Sci. USA* **75,** 655–659.

Cheung, A. Y., Bogorad, L., Van Montagu, M., and Schell, J. (1988). Relocating a gene for herbicide tolerance: A chloroplast gene is converted into a nuclear gene. *Proc. Natl. Acad. Sci. USA* **85,** 391–395.

Chitnis, P. R., Harel, E., Kohorn, B. D., Tobin, E. M., and Thornber, J. P. (1986). Assembly of the precursor and processed light-harvesting chlorophyll a/b protein of *Lemna* into the light-harvesting complex of barley etiochloroplasts. *J. Cell Biol.* **102,** 982–988.

Chitnis, P. R., Nechushtai, R., and Thornber, J. P. (1987). Insertion of the precursor of the light-harvesting chlorophyll a/b protein into the thylakoids requires the presence of a developmentally regulated stromal factor. *Plant Mol. Biol.* **10,** 3–12.

Chitnis, P. R., Morishige, D. T., Nechushtai, R., and Thornber, J. P. (1988). Assembly of the barley light-harvesting chlorophyll a/b proteins in barley etiochloroplasts involves processing of the precursor on thylakoids. *Plant Mol. Biol.* **11,** 95–107.

Chou, P. Y., and Fasman, G. D. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol.* **47,** 45–148.

Chua, N.-H., and Schmidt, G. W. (1978). Post-translational transport into intact chloroplasts of a precursor to the small subunit of ribulose-1,5-bisphosphate carboxylase. *Proc. Natl. Acad. Sci. USA* **75,** 6110–6114.

Cline, K. (1986). Import of proteins into chloroplasts: Membrane integration of a thylakoid precursor protein reconstituted in chloroplast lysates. *J. Biol. Chem.* **261,** 14804–14810.

Cline, K. (1988). Light-harvesting chlorophyll a/b protein. Membrane insertion, proteolytic processing, assembly into LHC II, and localization to appressed membranes occurs in chloroplast lysates. *Plant Physiol.* **86,** 1120–1126.

Cline, K. J., Werner-Washburne, M., Lubben, T. H., and Keegstra, K. (1985). Precursors to two nuclear-encoded chloroplast proteins bind to the outer envelope membrane before being imported into chloroplasts. *J. Biol. Chem.* **260,** 3691–3697.

Cline, K., Fulsom, D. R., and Viitanen, P. V. (1989). An imported thylakoid protein accumulates in the stroma when insertion into thylakoids is inhibited. *J. Biol. Chem.* **264,** 14225–14232.

Comai, L., Larson-Kelly, N., Kiser, J., Mau, C. J. D., Pokalsky, A. R., Shewmaker, C. K., McBride, K., Jones, A., and Stalker, D. M. (1988). Chloroplast transport of a ribulose

bisphosphate carboxylase small subunit-5-enolpyruvyl 3-phosphoshikimate synthase chimeric protein requires part of the mature small subunit in addition to the transit peptide. *J. Biol. Chem.* **263,** 15104–15109.

Cooper, P., and Ort, D. R. (1987). The effect of chilling on the production of photosystem II light harvesting chlorophyll a/b binding protein in tomato. *In* "Plant Gene Systems and their Biology." *UCLA Symp.* **62,** 353–362.

Cornwell, K. L., and Keegstra, K. (1987). Evidence that a chloroplast surface protein is associated with a specific binding site for the precursor to the small subunit of ribulose-1,5-bisphosphate carboxylase. *Plant Physiol.* **85,** 780–785.

De Almeida, E. R. P., Gossele, V., Muller, C. G., Dockx, J. Reynaerts, A., Botterman, J., Krebbers, E., and Timko, M. P. (1989). Transgenic expression of two marker genes under the control of an *Arabidopsis* rbcS promoter: Sequences encoding the Rubisco transit peptide increase expression levels. *Mol. Gen Genet.* **218,** 78–86.

De Boer, D., Cremers, F., Teertstra, R., Smits, L., Hille, J., Smeekens, S., and Weisbeek, P. (1988). In vivo import of plastocyanin and a fusion protein into developmentally different plastids of transgenic plants. *EMBO J.* **7,** 2631–2635.

della-Cioppa, G., and Kishore, G. M. (1988). Import of a precursor protein into chloroplasts is inhibited by the herbicide glyphosate. *EMBO J.* **7,** 1299–1305.

della-Cioppa, G., Bauer, S. C., Klein, B. K., Shah, D. M., Fraley, R. T., and Kishore, G. M. (1986). Translocation of the precursor of 5-enolpyruvylshikimate-3-phosphate synthase into chloroplast of higher plants in vitro. *Proc. Natl. Acad. Sci. USA* **83,** 6873–6877.

Dobberstein, B., Blobel, G., and Chua, N.-H. (1977). In vitro synthesis and processing of a putative precursor for the small subunit of ribulose 1,5-bisphosphate carboxylase of *Chlamydomonas reinhardtii Proc. Natl. Acad. Sci. USA* **74,** 1082–1085.

Dunsmuir, P. (1985). The petunia chlorophyll a/b binding protein genes: A comparison of Cab genes from different gene families. *Nucleic Acids Res.* **13,** 2503–2518.

Eckes, P., Rosahl, S., Schell, J., and Willmitzer, L. (1986). Isolation and characterization of a light-inducible, organ-specific gene from potato and analysis of its expression after tagging and transfer into tobacco and potato shoots. *Mol. Gen. Genet.* **205,** 14–22.

Elliott, R. C., Pedersen, T. J., Fristensky, B., White, M. J., Dickey, L. F., and Thompson, W. F. (1989). Characterization of a single copy gene encoding ferredoxin I from pea. *Plant Cell* **1,** 681–690.

Ellis, R. J. (1981). Chloroplast proteins: Synthesis, transport, and assembly. *Annu. Rev. Plant Phys.* **32,** 111–137.

Ellis, R. J. (1983). Chloroplast protein synthesis: Principles and problems. *In* "Subcellular Biochemistry," Vol. 9 (D. B. Roodyn, ed.), pp. 237–261. Plenum Publishing Corporation, New York.

Ellis, R. J., and Robinson, C. (1985). Post-translational transport and processing of cytoplasmically-synthesized precursors of organellar proteins. *In* "The Enzymology of Post-Translational Modification of Proteins," Vol. 2, pp. 25–39. Academic Press Inc., London.

Flügge, U. I., and Hinz, G. (1986). Energy dependence of protein translocation into chloroplasts. *Eur. J. Biochem.* **160,** 563–570.

Flügge, U. I., Fischer, K., Gross, A., Sebald, W., Lottspeich, F., and Eckerskorn, C. (1989). The triose phosphate-3-phosphoglycerate-phosphate translocator from spinach chloroplasts: Nucleotide sequence of a full-length cDNA clone and import of the *in vitro* synthesized precursor protein into chloroplasts. *EMBO J.* **8,** 39–46.

Franzen, L.-G., Frank, G., Zuber, H., and Rochaix, J.-D. (1989). Isolation and characterization of cDNA clones encoding the 17.9 and 8.1 kDa subunits of photosystem I from *Chlamydomonas reinhardtii. Plant Mol. Biol.* **12,** 463–474.

Friedman, A. L., and Keegstra, K. (1989). Chloroplast protein import. Quantitative analysis of precursor binding. *Plant Physiol.* **89,** 993–999.

Fulson, D. R., and Cline, K. (1988). A soluble protein factor is required in vitro for membrane insertion of the thylakoid precursor protein, pLHCP. *Plant Physiol.* **88,** 1146–1153.

Gantt, J. S. (1988). Nucleotide sequences of cDNAs encoding four complete nuclear-encoded plastid ribosomal proteins. *Curr. Genet.* **14,** 519–528.

Gasser, C. S., Winter, J. A., Hironaka, C. M., and Shah, D. M. (1988). Structure, expression, and evolution of the 5-enolpyruvylshikimate-3-phosphate synthase genes of petunia and tobacco. *J. Biol. Chem.* **263,** 4280–4289.

Gatenby, A. A., Lubben, T. H., Ahlquist, P., and Keegstra, K. (1988). Imported large subunits of ribulose bisphosphate carboxylase/oxygenase, but not imported β-ATP synthase subunits, are assembled into holoenzyme in isolated chloroplasts. *EMBO J* **7,** 1307–1314.

Gilbert, W. (1979). Introns and exons: Playgrounds of evolution. *In* "Eucaryotic Gene Regulation," Symposium on Molecular and Cellular Biology, Vol. 14 (R. Axel, T. Maniatis, and C. F. Fox, eds.), pp. 1–14. Academic Press, New York.

Goldschmidt-Clermont, M., and Rahire, M. (1986). Sequence, evolution and differential expression of the two genes encoding variant small subunits of ribulose bisphosphate carboxylase/oxygenase in *Chlamydomonas reinhardtii*. *J. Mol. Biol.* **191,** 421–432.

Grebanier, A. E., Coen, D. M., Rich, A., and Bogorad L. (1978). Membrane proteins synthesized but not processed by isolated maize chloroplasts. *J. Cell Biol.* **78,** 734–746.

Greenland, A. J., Thomas, M. V., and Walden, R. M. (1987). Expression of two nuclear genes encoding chloroplast proteins during early development of cucumber seedlings. *Planta* **170,** 99–110.

Grossman, A., Bartlett, S., and Chua, N.-H. (1980). Energy-dependent uptake of cytoplasmically synthesized polypeptides by chloroplasts. *Nature (London)* **285,** 625–628.

Gupta, S. C., and Beevers, L. (1987). Regulation of nitrite reductase. Cell-free translation and processing. *Plant Physiol.* **83,** 750–754.

Hageman, J., Robinson, C., Smeekens, S., and Weisbeek, P. (1986). A thylakoid processing protease is required for complete maturation of the lumen protein plastocyanin. *Nature (London)* **324,** 567–569.

Hansen, L. (1987). Three cDNA clones for barley leaf acyl carrier proteins I and III. *Carlsberg Res. Commun.* **52,** 381–392.

Hayden, D. B., Baker, N. R., Percival, M. P., and Beckwith, P. B. (1986). Modification of the photosystem II light harvesting chlorophyll a/b protein complex in maize during chill-induced photoinhibition. *Biochim. Biophys. Acta* **851,** 86–92.

Henning, B., Koehler, H., and Neupert, W. (1983). Receptor sites involved in posttranslational transport of apocytochrome c into mitochondria: Specificity, affinity, and number of sites. *Proc. Natl. Acad. Sci. USA* **80,** 4963–4967.

Hermans, J. Rother, Ch., Bichler, J., Steppuhn, J., and Hermann, R. G. (1988). Nucleotide sequence of cDNA clones encoding the complete precursor for subunit delta of thylakoid-located ATP synthase from spinach. *Plant Mol. Biol.* **10,** 323–330.

Herrin, D., and Michaels, A. (1985). The chloroplast 32 kDa protein is synthesized on thylakoid-bound ribosomes in *Chlamydomonas reinhardtii*. *FEBS Lett.* **184,** 90–95.

Highfield, P. E., and Ellis, R. J. (1978). Synthesis and transport of the small subunit of chloroplast ribulose bisphosphate carboxylase. *Nature (London)* **271,** 420–424.

Hinz, G., and Flügge, U. I. (1988). Phosphorylation of a 51-kDa envelope membrane polypeptide involved in protein translocation into chloroplasts. *Eur. J. Biochem.* **175,** 649–659.

Hoffman, N. E., Pichersky, E., Malik, V. S., Ko, K., and Cashmore, A. R. (1988). Isolation

and sequence of a tomato cDNA clone encoding subunit II of the photosystem I reaction center. *Plant Mol. Biol.* **10,** 435–445.

Hurt, E. C., Goldschmidt-Clermont, M., Pesold-Hurt, B., Rochaix, J.-D., and Schatz, G. (1986). A mitochondrial presequence can transport a chloroplast-encoded protein into yeast mitochondria. *J. Biol. Chem.* **261,** 11440–11443.

Ishizaki, O., Nishida, I., Agata, K., Eguchi, G., and Murata, N. (1988). Cloning and nucleotide sequence of cDNA for the plastid glycerol-3-phosphate acyltransferase from squash. *FEBS Lett.* **238,** 424–430.

Jansen, T., Rother, C., Steppuhn, J., Reinke, H., Beyreuther, K., Jansson, C., Andersson, B., and Herrmann, R. G. (1987). Nucleotide sequence of cDNA clones encoding the complete '23 kDa' and '16 kDa' precursor proteins associated with the photosynthetic oxygen-evolving complex from spinach. *FEBS Lett.* **216,** 234–240.

Jansen, T., Reilander, H., Steppuhn, J., and Herrmann, R. G. (1988). Analysis of cDNA clones encoding the entire precursor-polypeptide for ferredoxin: $NADP^+$ oxidoreductase from spinach. *Curr. Genet.* **13,** 517–522.

Kaderbhai, M. A., Pickering, T., Austen, M. B., and Kaderbhai, N. (1988). A photoactivatable synthetic peptide labels 30 kDa and 52 kDa polypeptides of the chloroplast inner envelope membrane. *FEBS Lett.* **232,** 313–316.

Karlin-Neumann, G. A., and Tobin, E. M. (1986). Transit peptides of nuclear-encoded chloroplast proteins share a common amino acid framework. *EMBO J.* **5,** 9–13.

Kavanagh, T. A., Jefferson, R. A., and Bevan, M. W. (1988). Targeting a foreign protein to chloroplasts using fusions to the transit peptide of a chlorophyll a/b protein. *Mol. Gen. Genet.* **215,** 38–45.

Keegstra, K. (1989). Transport and routing of proteins into chloroplasts. *Cell* **56,** 247–253.

Keegstra, K., and Bauerle, C. (1988). Targeting of proteins into chloroplasts. *BioEssays* **9,** 15–19.

Keegstra, K., Bauerle, C., Lubben, T., Smeekens, S., and Weisbeek, P. (1987). Targeting of proteins into chloroplasts. *In* "Plant Gene Systems and their Biology," pp. 363–370. Alan R. Liss, New York.

Keegstra, K., Olsen, L. J., and Theg, S. M. (1989). Chloroplastic precursors and their transport across the envelope membranes. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40,** 471–501.

Kirwin, P. M., Elderfield, P. D., and Robinson, C. (1987). Transport of proteins into chloroplasts. Partial purification of a thylakoidal processing peptidase involved in plastocyanin biogenesis. *J. Biol. Chem.* **262,** 16386–16390.

Kirwin, P. M., Meadows, J. W., Shackleton, J. B., Musgrove, J. E., Elderfield, P. D., Mould, R., Hay, N. A., and Robinson, C. (1989). ATP-dependent import of a lumenal protein by isolated thylakoid vesicles. *EMBO J.* **8,** 2251–2255.

Klee, H. J., Muskopf, Y. M., and Gasser, C. S. (1987). Cloning of an *Arabidopsis thaliana* gene encoding 5-enolpyruvyl-shikimate-3-phosphate synthase: Sequence analysis and manipulation to obtain glyphosate-tolerant plants. *Mol. Gen. Genet.* **210,** 437–442.

Kosgen, R. B., Gierl, A., Schwarz-Sommer, Z., and Saedler, H. (1986). Molecular analysis of the waxy locus of *Zea mays*. *Mol. Gen. Genet.* **203,** 237–244.

Klosgen, R. B., Saedler, H., and Weil, J.-H. (1989). The amyloplast-targeting transit peptide of the waxy protein of maize also mediates protein transport in vitro into chloroplasts. *Mol. Gen. Genet.* **217,** 155–161.

Ko, K., Granell, A., Bennett, J., and Cashmore, A. R. (1990). Isolation and characterization of cDNAs from *Lycopersicon esculentum* and *Arabidopsis thaliana* encoding the 33 kd protein of the photosystem II-associated oxygen-evolving complex. *Plant Mol. Biol.*, **14,** 217–227.

Kohorn, B. D., and Tobin, E. M. (1986). Chloroplast import of light-harvesting chlorophyll

a/b-proteins with different amino termini and transit peptides. *Plant Physiol.* **82,** 1172–1174.

Kohorn, B. D., and Tobin, E. M. (1987). Amino acid charge distribution influences the assembly of apoprotein into light-harvesting complex II. *J. Biol. Chem.* **262,** 12897–12899.

Kohorn, B. D., Harel, E., Chitnis, P. R., Thornber, J. P., and Tobin, E. M. (1986). Functional and mutational analysis of the light-harvesting chlorophyll a/b protein of thylakoid membranes. *J. Cell Biol.* **102,** 972–981.

Kuntz, M., Simons, A., Schell, J., and Schreier, P. H. (1986). Targeting of protein to chloroplasts in transgenic tobacco by fusion to mutated transit peptide. *Mol. Gen. Genet.* **205,** 454–460.

Kuwabara, T., Reddy, K. J., and Sherman, L. A. (1987). Nucleotide sequence of the gene from the cyanobacterium *Anacycstis nidulans* R2 encoding the MN-stabilizing protein involved in photosystem II water oxidation. *Proc. Natl. Acad. Sci. USA* **84,** 8230–8234.

Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157,** 105–132.

Lagoutte, B. (1988). Cloning and sequencing of spinach cDNA clones encoding the 20 kDa PS I polypeptide. *FEBS Lett.* **232,** 275–280.

Lamppa, G. K. (1988). The chlorophyll a/b-binding protein inserts into the thylakoids independent of its cognate transit peptide. *J. Biol. Chem.* **263,** 14996–14999.

Lamppa, G. K., and Abad, M. S. (1987). Processing of a wheat light-harvesting chlorophyll a/b protein precursor by a soluble enzyme from higher plant chloroplasts. *J. Cell Biol.* **105,** 2641–2648.

Lamppa, G. K., Morelli, G., and Chua, N.-H. (1985). Structure and developmental regulation of a wheat gene encoding the major chlorophyll a/b-binding polypeptide. *Mol. Cell. Biol.* **5,** 1370–1378.

Landry, S. J., and Bartlett, S. G. (1989). The small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase and its precursor expressed in *Escherichia coli* are associated with groEL protein. *J. Biol. Chem.* **264,** 9090–9093.

Lautner, A., Klein, R., Ljungberg, U., Reilander, H., Bartling, D., Andersson, B., Reinke, H., Beyreuther, K., and Hermann, R. G. (1988). Nucleotide sequence of cDNA clones encoding the complete precursor for the "10-kDa" polypeptide of photosystem II from spinach. *J. Biol. Chem.* **263,** 10077–10081.

Lightfoot, D. A., Green, N. K., and Cullimore, J. V. (1988). The chloroplast-located glutamine synthetase of *Phaseolus vulgaris* L: Nucleotide sequence, expression in different organs and uptake into isolated chloroplasts. *Plant Mol. Biol.* **11,** 191–202.

Lopez-Otin, C., Simon-Mateo, C., Martinez, L., and Vinuela, E. (1989). Gly-Gly-X, a novel consensus sequence for the protoeolytic processing of viral and cellular proteins. *J. Biol. Chem.* **264,** 9107–9110.

Lubben, T. H., and Keegstra, K. (1986). Efficient in vitro import of a cytosolic heat shock protein into pea chloroplasts. *Proc. Natl. Acad. Sci. USA* **83,** 5502–5505.

Lubben, T. H., Bansberg, J., and Keegstra, K. (1987). Stop-transfer regions do not halt translocation of proteins into chloroplasts. *Science* **238,** 1112–1114.

Lubben, T. H., Theg, S. M., and Keegstra, K. (1988). Transport of proteins into chloroplasts. *Photosyn. Res.* **17,** 173–194.

Lubben, T. H., Gatenby, A. A., Ahlquist, P., and Keegstra, K. (1989). Chloroplast import characteristics of chimeric proteins. *Plant Mol. Biol.* **12,** 13–18.

Marks, D. B., Keller, B. J., and Hoober, J. K. (1985). *In vitro* processing of precursors of thylakoid membrane proteins of *Chlamydomonas reinhardtii* y-1. *Plant Physiol.* **79,** 108–113.

Matsuoka, M., Ozeki, Y., Yamamoto, N., Hirano, H., Kano-Murakami, Y., and Tanaka, Y.

(1988). Primary structure of maize pyruvate, orthophosphate dikinase as deduced from cDNA sequence. *J. Biol. Chem.* **263,** 11080–11083.

Mattoo, A. K., and Edelman, M. (1987). Intramembrane translocation and posttranslational palmitoylation of the chloroplast 32-kDa herbicide binding protein. *Proc. Natl. Acad. Sci. USA* **84,** 1497–1501.

Mattoo, A. K., Marder, J. B., and Edelman, M. (1989). Dynamics of the photosystem II reaction center. *Cell* **56,** 241–246.

Mayfield, S. P., Rahire, M., Frank, G., Zuber, H., and Rochaix, J.-D. (1987). Expression of the nuclear gene encoding oxygen-evolving enhancer protein 2 is required for high levels of photosynthetic oxygen evolution in *Chlamydomonas reinhardtii*. *Proc. Natl. Acad. Sci. USA* **84,** 749–753.

Mayfield, S. P., Schirmer-Rahire, M., Frank, G., Zuber, H., and Rochaix, J.-D. (1989). Analysis of the genes of the OEE1 and OEE3 proteins of the photosystem III complex from *Chlamydomonas reinhardtii*. *Plant Mol. Biol.* **12,** 683–693.

Mazur, B. J., Chui, C.-F., and Smith, J. K. (1987). Isolation and characterization of plant genes coding for acetolactate synthase, the target enzyme for two classes of herbicides. *Plant Physiol.* **85,** 1110–1117.

Meagher, R. B., Berry-Lowe, S., and Rice, K. (1989). Molecular evolution of the small subunit of ribulose bisphosphate carboxylase: Nucleotide substitution and gene conversion. *Genetics,* **123,** 845–863.

Merchant, S., and Bogorad, L. (1987). The Cu(II)-repressible plastidic cytochrome c. Cloning and sequence of a complementary DNA for the pre-apoprotein. *J. Biol. Chem.* **262,** 9062–9067.

Michalowski, C. B., Schmitt, J. M., and Bohnert, H. J. (1989). Expression during salt stress and nucleotide sequence of cDNA for ferredoxin-NADP+ reductase from *Mesembryanthemum crystallinum*. *Plant Physiol.* **89,** 817–822.

Michel, H., Hung, D. F., Shabanovitz, J., and Bennett, J. (1988). Tandem mass spectrometry reveals that three photosystem II proteins of spinach chloroplasts contain N-acetyl-O-phosphothreonine at their NH_2 termini. *J. Biol. Chem.* **263,** 1123–1130.

Mishkind, M. L., and Scioli, S. E. (1988). Recent developments in chloroplast protein transport. *Photosyn. Res.* **19,** 153–184.

Mishkind, M. L., Branagan, A. J., Jensen, K. H., Plumley, F. G., and Schmidt, G. W. (1985a). Roles of proteases in chloroplast biogenesis. *Curr. Top. Plant Biochem. Physiol.* **4,** 34–50.

Mishkind, M. L., Wessler, S. R., and Schmidt, G. W. (1985b). Functional determinants in transit sequences: Import and partial maturation by vascular plant chloroplasts of the ribulose-1,5-bisphosphate carboxylase small subunit of *Chlamydomonas*. *J. Cell. Biol.* **100,** 226–234.

Musgrove, J. E., Elderfield, P. D., and Robinson, C. (1989). Endopeptidases in the stroma and thylakoids of pea chloroplasts. *Plant Physiol.* **90,** 1616–1621.

Newman, B. J., and Gray, J. C. (1988). Characterisation of a full-length cDNA clone for pea ferredoxin-NADP+ reductase. *Plant Mol. Biol.* **10,** 511–520.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., and Ozeki, H. (1986). Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature (London)* **322,** 572–574.

Olive, M. R., Ellis, R. J., and Schuch, W. W. (1989). Isolation and nucleotide sequences of cDNA clones encoding ADP-glucose pyrophosphorylase polypeptides from wheat leaf and endosperm. *Plant Mol. Biol.* **12,** 525–538.

Olsen, L. J., Theg, S. M., Selman, B. R., and Keegstra, K. (1989). ATP is required for the binding of precursor proteins to chloroplasts. *J. Biol. Chem.* **264,** 6724–6729.

Ostrem, J. A., Ramage, R. T., Bohnert, H. J., and Wasmann, C. C. (1989). Deletion of the carboxyl-terminal portion of the transit peptide affects processing but not import or assembly of the small subunit of ribulose-1,5-bisphosphate carboxylase. *J. Biol. Chem.* **264,** 3662–3665.

Pain, D., and Blobel, G. (1987). Protein import into chloroplasts requires a chloroplast ATPase. *Proc. Natl. Acad. Sci. USA* **84,** 3288–3292.

Pain, D., Kanwar, Y. S., and Blobel, G. (1988). Identification of a receptor for protein import into chloroplasts and its localization to envelope contact sites. *Nature (London)* **331,** 232–237.

Perl-Treves, R., Nacmias, B., Aviv, D., Zeelon, E. P., and Galun, E. (1988). Isolation of two cDNA clones from tomato containing two different superoxide dismutase sequences. *Plant Mol. Biol.* **11,** 609–623.

Pfisterer, J., Lachmann, P., Kloppstech, K. (1982). Transport of proteins into chloroplasts. Binding of nuclear-coded chloroplast proteins to the chloroplast envelope. *Eur. J. Biochem.* **126,** 143–148.

Philbrick, J. B., and Zilinskas, B. A. (1988). Cloning, nucleotide sequence and mutational analysis of the gene encoding the photosystem II manganese-stabilizing polypeptide of *Synechosystis* 6803. *Mol. Gen. Genet.* **212,** 418–425.

Phua, S. H., Srinivasa, B. R., and Subramanian, A. R. (1989). Chloroplast ribosomal protein L13 is encoded in the nucleus and is considerably larger than its bacterial homologue. *J. Biol. Chem.* **264,** 1968–1971.

Pichersky, E., Hoffman, N. E., Bernatzky, R., Piechulla, B., Tanksley, S. D., and Cashmore, A. R. (1987a). Molecular characterization and genetic mapping of DNA sequences encoding the type I chlorophyll a/b-binding polypeptide of photosystem I in *Lycopersicon esculentum* (tomato). *Plant Mol. Biol.* **9,** 205–216.

Pichersky, E., Hoffman, N. E., Malik, V. S., Bernatzky, R., Tanksley, S. D., Szabo, L., and Cashmore, A. R. (1987b). The tomato Cab-4 and Cab-5 genes encode a second type of CAB polypeptide localized in photosystem II. *Plant Mol. Biol.* **9,** 109–120.

Pichersky, E., Tanksley, S. D., Piechulla, B., Stayton, M. M., and Dunsmuir, P. (1988). Nucleotide sequence and chromosomal location of Cab-7, the tomato gene encoding the type II chlorophyll a/b-binding polypeptide of photosystem I. *Plant Mol. Biol.* **11,** 69–71.

Post-Beittenmiller, M. A., Hlousek-Radojcic, A., and Ohlrogge, J. B. (1989). DNA sequence of a genomic clone encoding an *Arabidopsis* acyl carrier protein (ACP). *Nucleic Acids Res.* **17,** 1777.

Quigley, F., Martin, W. F., and Cerff, R. (1988). Intron conservation across the prokaryote–eukaryote boundary: Structure of the nuclear gene for chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize. *Proc. Natl. Acad. Sci. USA* **85,** 2672–2676.

Raines, C. A., Lloyd, J. C., Longstaff, M., Bradley, D., and Dyer, T. (1988). Chloroplast fructose-1,6-bisphosphatase: The product of a mosaic gene. *Nucleic Acids Res.* **16,** 7931–7942.

Reiss, B., Wasmann, C. C., and Bohnert, H. J. (1987). Regions in the transit peptide of SSU essential for transport into chloroplasts. *Mol. Gen. Genet.* **209,** 116–121.

Reiss, B., Wasmann, C. C., Schell, J., and Bohnert, H. J. (1989). Effect of mutations on the binding and translocation functions of a chloroplast transit peptide. *Proc. Natl. Acad. Sci. USA* **86,** 886–890.

Reithmeier, R. A. F. (1988). Assembly of proteins into biological membranes—A brief history lesson based on 10 key papers. *Comp. Biochem. Physiol.* **91B,** 613–616.

Robinson, C., and Ellis, R. J. (1984a). Transport of proteins into chloroplasts: Partial purification of a chloroplast protease involved in the processing of imported precursor polypeptides. *Eur. J. Biochem.* **142,** 337–342.

Robinson, C., and Ellis, R. J. (1984b). Transport of proteins into chloroplasts: The precur-

sor of ribulose bisphosphate carboxylase small subunit is processed in two steps. *Eur. J. Biochem.* **142,** 343–342.

Roesler, K. R., and Ogren, W. L. (1988). Nucleotide sequence of spinach cDNA encoding phosphoribulokinase. *Nucleic Acids Res.* **16,** 7192.

Rohde, W., Becker, D., and Salamini, F. (1988). Structural analysis of the waxy locus from *Hordeum vulgare*. *Nucleic Acids Res.* **16,** 7185–7186.

Rose, R. E., DeJesus, C. E., Moylan, S. L., Ridge, N. P., Scherer, D. E., and Knauf, V. C. (1987). The nucleotide sequence of a cDNA clone encoding acyl carrier protein (ACP) from *Brassica campestris* seeds. *Nucleic Acids Res.* **15,** 7197.

Rother, C., Jansen, T., Tyagi, A., Tittgen, J., and Herrmann, R. G. (1986). Plastocyanin is encoded by an uninterrupted nuclear gene in spinach. *Curr. Genet.* **11,** 171–176.

Roy, H., Cannon, S., and Gilson, M. (1988). Assembly of Rubisco from native subunits. *Biochim. Biophys. Acta* **957,** 323–334.

Safford, R., Windust, J. H. C., Lucas, C., DeSilva, J., James, C. M., Hellyer, A., Smith, G., Slabas, A. R., and Hughes, S. G. (1988). Plastid-localised seed acyl-carrier protein of *Brassica napus* is encoded by a distinct, nuclear multigene family. *Eur. J. Biochem.* **174,** 287–295.

Scherer, D. E., and Knauf, V. C. (1987). Isolation of a cDNA clone for the acyl carrier protein-I of spinach. *Plant Mol. Biol.* **9,** 127–134.

Schindler, C., and Soll, J. (1986). Protein transport in intact, purified pea etioplasts. *Arch. Biochem. Biophys.* **247,** 211–220.

Schindler, C., Hracky, R., and Soll, J. (1986). Protein transport in chloroplasts: ATP is prerequisite. *Z. Naturforsch.* **42c,** 103–108.

Schleyer, M., and Neupert, W. (1985). Transport of proteins into mitochondria: Translocational intermediates spanning contact sites between outer and inner membranes. *Cell* **43,** 339–350.

Schmidt, B., Wachter, E., Sebald, W., and Neupert, W. (1984). Processing peptidase of *Neurospora* mitochondria. Two-step cleavage of imported ATPase subunit 9. *Eur. J. Biochem.* **144,** 581–588.

Schmidt, G. W., and Mishkind, M. L. (1986). The transport of proteins into chloroplasts. *Annu. Rev. Biochem.* **55,** 879–912.

Schmidt, G. W., Devillers-Thiery, A., Desruisseaux, H., Blobel, G., and Chua, N.-H. (1979). NH_2-terminal amino acid sequences of precursor and mature forms of the ribulose-1,5-bisphosphate carboxylase small subunit from *Chlamydomonas reinhardtii*. *J. Cell Biol.* **83,** 615–622.

Schmidt, G. W., Bartlett, S. G., Grossman, A. R., Cashmore, A. R., and Chua, N.-H. (1981). Synthesis, transport, processing and assembly of two polypeptide subunits of the light harvesting chlorophyll a/b protein complex. *J. Cell Biol.* **91,** 468–479.

Schmidt, R. J., Myers, A. M., Gillham, N. W., and Boynton, J. E. (1984). Chloroplast ribosomal proteins of *Chlamydomonas* synthesized in the cytoplasm are made as precursors. *J. Cell Biol.* **98,** 2014–2018.

Schmidt, R. J., Gillham, N. W., and Boynton, J. E. (1985). Processing of the precursor to a chloroplast ribosomal protein made in the cytosol occurs in two steps, one of which depends on a protein made in the chloroplast. *Mol. Cell Biol.* **5,** 1093–1099.

Schreier, P., Seftor, E. A., Schell, J., and Bohnert, H. J. (1985). The use of nuclear-encoded sequences to direct the light-regulated synthesis and transport of a foreign protein into plant chloroplasts. *EMBO J.* **4,** 25–32.

Scioli, J. R., and Zilinskas, B. A. (1988). Cloning and characterization of a cDNA encoding the chloroplastic copper/zinc-superoxide dismutase from pea. *Proc. Natl. Acad. Sci. USA* **85,** 7661–7665.

Shih, M.-C., Lazar, G., and Goodman, H. M. (1986). Evidence in favor of the symbiotic

origin of chloroplasts: Primary structure and evolution of tobacco glyceraldehyde-3-phosphate dehydrogenases. *Cell* **47,** 73–80.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashido, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kususda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., and Sugiura, M. (1986). The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. *EMBO J.* **5,** 2043–2049.

Smeekens, S., and Weisbeek, P. (1988). Protein transport towards the thylakoid lumen: Post-translational translocation in tandem. *Photosyn. Res.* **16,** 177–186.

Smeekens, S., de Groot, M., van Binsbergen, J., and Weisbeek, P. (1985a). Sequence of the precursor of the chloroplast thylakoid lumen protein plastocyanin. *Nature (London)* **317,** 456–458.

Smeekens, S., van Binsbergen, J., and Weisbeek, P. (1985b). The plant ferredoxin precursor: Nucleotide sequence of a full length cDNA clone. *Nucleic Acids Res.* **13,** 3179–3194.

Smeekens, S., Bauerle, C., Hageman, J., Keegstra, K., and Weisbeek, P. (1986). The role of the transit peptide in the routing of precursors toward different chloroplast compartments. *Cell* **46,** 365–375.

Smeekens, S., van Steeg, H., Bauerle, C., Bettenbroek, H., Keegstra, K., and Weisbeek, P. (1987). Import into chloroplasts of a yeast mitochondrial protein directed by ferredoxin and plastocyanin transit peptides. *Plant Mol. Biol.* **9,** 377–388.

Smeekens, S., Geerts, D., Bauerle, C., and Weisbeek, P. (1989). Essential function in chloroplast recognition of the ferredoxin transit peptide processing region. *Mol. Gen. Genet.* **216,** 178–182.

Smith, S. M., and Ellis, R. J. (1979). Processing of small subunit precursor of ribulose bisphosphate carboxylase and its assembly into whole enzyme are stromal events. *Nature (London)* **278,** 662–664.

Soll, J. (1988). Purification and characterization of a chloroplast outer-envelope-bound, ATP-dependent protein kinase. *Plant Physiol.* **88,** 898–903.

Steppuhn, J., Rother, C., Hermans, J., Jansen, T., Salnikow, J., Hauska, G., and Herrmann, R. G. (1987). The complete amino-acid sequence of the Rieske FeS-precursor protein from spinach chloroplasts deduced from cDNA analysis. *Mol. Gen. Genet.* **210,** 171–177.

Theg, S. M., Bauerle, C., Olsen, L. J., Selman, B. R., and Keegstra, K. (1989). Internal ATP is the only energy requirement for the translocation of precursor proteins across chloroplastic membranes. *J. Biol. Chem.* **264,** 6730–6736.

Tingey, S. V., Tsai, F.-Y., Edwards, J. W., Walker, E. L., and Coruzzi, G. M. (1988). Chloroplast and cytosolic glutamine synthetase are encoded by homologous nuclear genes which are differentially expressed in vivo. *J. Biol. Chem.* **263,** 9651–9657.

Tyagi, A., Hermans, J., Steppuhn, J., Jansson, Ch., Vater, F., and Herrmann, R. G. (1987). Nucleotide sequence of cDNA clones encoding the complete "33kDa" precursor protein associated with the photosynthetic oxygen-evolving complex from spinach. *Mol. Gen. Genet.* **207,** 288–293.

Van den Broeck, G., Timko, M. P., Kausch, A. P., Cashmore, A. R., Van Montagu, M., and Herrera-Estrella, L. (1985). Targeting of a foreign protein to chloroplasts by fusion to the transit peptide from the small subunit of ribulose 1,4-bisphosphate carboxylase. *Nature (London)* **313,** 358–363.

Vierling, E., Nagao, R. T., DeRocher, A. E., and Harris, L. M. (1988). A heat shock protein localized to chloroplasts is a member of a eukaryotic superfamily of heat shock proteins. *EMBO J.* **7,** 575–581.

Viitanen, P. V., Doran, E. R., and Dunsmuir, P. (1988). What is the role of the transit

peptide in thylakoid integration of the light-harvesting chlorophyll a/b protein? *J. Biol. Chem.* **263,** 15000–15007.

Vorst, O., Oosterhoff-Teertstra, R., Vankan, P., Smeekens, S., and Weisbeek, P. (1988). Plastocyanin of *Arabidopsis thaliana;* isolation and characterization of the gene and chloroplast import of the precursor protein. *Gene* **65,** 59–69.

Wales, R., Newman, B. J., Pappin, D., and Gray, J. C. (1989). The extrinsic 33 kDa polypeptide of the oxygen-evolving complex of photosystem II is a putative calcium-binding protein and is encoded by a multi-gene family in pea. *Plant Mol. Biol.* **12,** 439–451.

Wasmann, C. C., Reiss, B., Bartlett, S. G., and Bohnert, H. J. (1986). The importance of the transit peptide and the transported protein for protein import into chloroplasts. *Mol. Gen. Genet.* **205,** 446–453.

Wasmann, C. C., Reiss, B., and Bohnert, H. J. (1988). Complete processing of a small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from pea requires the amino acid sequence Ile-Thr-Ser. *J. Biol. Chem.* **263,** 617–619.

Werneke, J. M., Zielinski, R. E., and Ogren, W. L. (1988). Structure and expression of spinach leaf cDNA encoding ribulosebisphosphate carboxylase/oxygenase activase. *Proc. Natl. Acad. Sci. USA* **85,** 787–791.

Yu, L. M., and Selman, B. R. (1988). cDNA sequence and predicted primary structure of the τ subunit from the ATP synthase from *Chlamydomonas reinhardtii. J. Biol. Chem.* **263,** 19342–19345.

CHAPTER 11

Origin and Evolution of Plastid Genomes and Genes

Michael W. Gray

Department of Biochemistry
Faculty of Medicine
Dalhousie University
Halifax, Nova Scotia, Canada

I. INTRODUCTION

More than 100 years ago, primarily on the basis of morphological similarities, Schimper (1883) proposed that plastids were the evolutionary descendants of blue-green algal (cyanobacterial) endosymbionts. During the first two decades of this century, this idea was further elaborated by Mereschkowsky (1905, 1910, 1920). It was then largely ignored or for-

gotten in the ensuing 40 years, to be revived by the discovery of plastid DNA (plDNA) (Ris, 1961; Ris and Plaut, 1962; Sager and Ishida, 1963; Chun *et al.*, 1963) and repopularized in the early 1970s (e.g., Margulis, 1970) as molecular data about organelles began to accumulate (for more detailed historical summaries, see Taylor, 1980, 1987; Margulis, 1981). During the past two decades, a convergence of ultrastructural, biochemical, and molecular biological evidence has made it virtually impossible to deny the validity of Schimper's surmise about the origin of plastids. His theory, and Mereschkowsky's extension of it, are now part of what is commonly called the endosymbiont hypothesis, an integrated scheme that sees symbiosis as the evolutionary driving force in the establishment of eukaryotic organelles, notably plastids and mitochondria (see Margulis, 1981; Gray and Doolittle, 1982).

During the 1970s, a vigorous debate ensued between advocates (e.g., Margulis, 1975) and critics (e.g., Raff and Mahler, 1972, 1975; Uzzell and Spolsky, 1974) of the endosymbiont hypothesis. During this period, much data and many arguments were brought to bear on opposing autogenous (origin from within) and xenogenous (origin from outside) theories of organelle evolution, detailed accounts of which may be found in other reviews and commentaries (Gillham and Boynton, 1981; Gray and Doolittle, 1982; Wallace, 1982; Gray, 1983, 1988, 1989). There is now such a solid body of support for a eubacterial (specifically cyanobacterial) endosymbiotic origin of the plastid and its genome that it seems pointless to consider seriously alternative explanations. This chapter focuses on the most recent molecular data (particularly gene sequence, mapping, and expression) that highlight the questions that have now assumed prominence in this field. In particular, I will consider at some length what is perhaps the most contentious issue at the moment: Did plastids arise only once (monophyletic origin) or more than once (polyphyletic origin)? Because the topic of plastid chromosome structure and evolution is dealt with elsewhere in this volume (Chapter 2A), some other evolutionary questions (e.g., What evolutionary processes have led to the differences in genome organization and expression that exist among the plDNAs of the various photosynthetic eukaryotes [land plants and algae]?, What molecular mechanisms are operating in plDNA evolution?) will not be discussed in great depth here.

Although cyanobacteria and other photosynthetic bacteria can both respire aerobically and carry out photosynthesis, the latter do not evolve oxygen and have only one light reaction; the former, in concert with the plastids of all eukaryotic algae and land plants, have two light reactions and evolve oxygen. These considerations alone initially suggested that

plastids originated from within the cyanobacterial group of eubacteria and not from within a group of nonoxygenic photosynthesizers. Before considering some of the most recent evidence that supports this conclusion, I will briefly review current concepts and controversies concerning the origin of plastids. I will also emphasize the variety of plastid types that exist, because such variation has not only been used as a basis for classification of the eukaryotic algae, but also as an indication of a possible polyphyletic origin of plastids.

II. ENDOSYMBIOTIC ORIGIN OF PLASTIDS AND MITOCHONDRIA: SERIAL OR SIMULTANEOUS?

Molecular data have provided one particularly compelling argument in favor of an endosymbiotic origin of plastids and mitochondria, namely that the evolutionary ancestry of the two organelles can be traced to distinctly different phyla of eubacteria: cyanobacteria in the case of plastids and the so-called purple nonsulfur (PNS) bacteria (Woese, 1987) in the case of mitochondria (see Gray and Doolittle, 1982; Gray, 1988, 1989). As Doolittle (1980) has pointed out, this clear distinction makes it impossible to entertain an autogenous origin for both organelles from within the same host cell. In an autogenous scenario, the progenitor would have to have been either a cyanobacterialike organism (to account for the plastid) or an organism like a PNS bacterium (to account for the mitochondrion), but clearly it could not have been both at the same time. Thus, in an autogenous scheme, either the plastid or the mitochondrion would still have to have been acquired xenogenously. Considering the gross similarities between the two organelles, it seems unnecessarily *ad hoc* to postulate an autogenous origin for one but a xenogenous origin for the other. As argued elsewhere (Gray and Doolittle, 1982; Gray, 1983, 1988, 1989), the endosymbiotic acquisition of *both* the plastid (from one group of eubacteria) *and* the mitochondrion (from another group) is the most straightforward explanation, consistent with available evidence, of how these eukaryotic organelles originated.

Although it is safe to assert that acceptance of this thesis is overwhelming, there is much less agreement on when the original endosymbioses were established, how many there were in the case of each organelle, and how the transition from symbiont to organelle proceeded. The serial endosymbiosis theory (Taylor, 1974; Margulis, 1981) proposes a

stepwise acquisition of organelles, with mitochondrion probably entering first and plastid sometime later; this theory can readily accommodate multiple origins of each organelle in different eukaryotic lineages. A less widely appreciated view, but one no less forcefully presented, is that of Cavalier-Smith (1987b), who proposes a simultaneous symbiotic origin of mitochondria and chloroplasts. Cavalier-Smith (1987b) argues that the reason cyanobacteria became chloroplasts whereas purple bacteria became mitochondria, and not the reverse, is that they were taken up by the very same host (a primitive eukaryotic protozoan, or Archezoan [Cavalier-Smith, 1987a] and evolved into organelles simultaneously and synergistically. He supposes that because PNS bacteria cannot photosynthesize in the presence of oxygen, the protomitochondrion lost photosynthesis; conversely, because cyanobacteria cannot control respiration separately from photosynthesis, the protoplastid lost respiration and relied in the dark on the more complete and efficient respiratory system of the PNS symbiont. According to this scenario, all extant mitochondrial and plastid types are descendants of these single and essentially simultaneous symbioses.

III. PHYLOGENETIC DISTRIBUTION AND DIVERSITY OF PLASTIDS

Table I lists plastid types of eukaryotic algae classified according to pigment composition and gives the various plant and algal groups in which these types occur. In plastids as in cyanobacteria, the photosynthetic apparatus always contains chlorophyll *a* as the primary pigment associated with photosystem I and II reaction centers. However, differences in accessory pigments serve to divide the eukaryotic algae into three generally recognized groups (Christensen, 1964): (1) the red algae (Rhodophyta), which have phycobilins (chl *a*/PB algae); (2) the green algae (Chlorophyta and Gamophyta) and euglenoid algae (Euglenophyta), which contain chlorophyll *b* (chl *a*/*b* algae); and (3) the chromophyte algae, a diverse assemblage of eight different phyla (using the classification of Margulis and Schwartz [1982]), all of which have chlorophyll *c* (chl *a*/*c* algae). These pigment differences are paralleled by substantial differences in thylakoid structure (e.g., whether or not thylakoids are stacked and how they are interconnected) (Gibbs, 1970). The above three groupings may be subdivided further according to number of

TABLE I
Structural and Biochemical Diversity of Plastids in Eukaryotic Algae[a]

Plastid Type	No. of Bounding Membranes		
	Two	Three	Four
Green (chl *a*/*b*)	Chlorophyta Gamophyta	Euglenophyta	
Red (chl *a*/PB)	Rhodophyta		
Yellow-brown (chl *a*/*c*) ("chromophytes")		Dinoflagellata[b]	Chrysophyta Haptophyta Cryptophyta[c] Xanthophyta Eustigmatophyta[c] Bacillariophyta Phaeophyta

[a] Phyla as designated by Margulis and Schwartz (1982).
[b] Certain dinoflagellate groups have plastids with only two membranes.
[c] Also contain phycobilins.

membranes surrounding the plastid (Table I), the evolutionary significance of which is discussed below.

Like chlorophyte algae, multicellular land plants (metaphytes) have chlorophylls *a* and *b*. This and much other evidence supports a chlorophytan ancestry of the metaphytes (Ragan and Chapman, 1978; Mattox and Stewart, 1984). Indeed, chloroplasts from a green alga such as *Chara* [a member of the Charophyceae, from which land plants are thought to have arisen (Mattox and Stewart, 1984)] and an angiosperm such as *Phaesolus vulgaris* are virtually indistinguishable in structure (Whatley, 1983). Molecular genetic evidence in support of the charophycean origin of the land plants has recently been published (Manhart and Palmer, 1990; Devereux *et al.*, 1990).

IV. SINGLE OR MULTIPLE ORIGIN(S) OF PLASTIDS?

Mereshkowsky (1910) first suggested that plastids of different pigment composition developed from "cyanophytes" having the corresponding pigment arrays, through a series of separate endosymbioses. This theme of a multiple origin of plastids was taken up more recently by Raven (1970) and subsequently by others (Whatley and Whatley, 1981; Mirab-

dullaev, 1985). At the time of Raven's proposal, the only known oxygen-evolving eubacteria were the cyanobacteria, which (like rhodophyte plastids) have chlorophyll *a* and phycobilins, the latter localized in phycobilisomes that protrude from single thylakoids. Among prokaryotes, the phycobilins are found only in the cyanobacteria, within which group they are universal (Gray and Doolittle, 1982). On the basis of these similarities, Raven proposed that a cyanobacterialike (chl *a*/PB) progenitor was the likely endosymbiotic precursor of rhodophyte plastids. However, in the case of chlorophyte–metaphyte–euglenophyte (chl *a*/*b*) and chromophyte (chl *a*/*c*) plastids, no potential eubacterial relatives of the appropriate pigment composition were known. The discovery of "prochlorophytes" (Lewin, 1976, 1977, 1981; Lewin and Withers, 1975; Burger-Wiersma *et al.*, 1986; Chisholm *et al.*, 1988), eubacteria having chlorophyll *b* rather than phycobilins, raised the possibility that contemporary relatives of Raven's postulated "green" (chl *a*/*b*) prokaryotes had been found. Similarly, characterization of the brownish photoheterotroph, *Heliobacterium chlorum* (Gest and Favinger, 1983), prompted the suggestion that this eubacterium might be a modern counterpart of Raven's "yellow" (chl a/c) endosymbiont (Margulis and Obar, 1985). Recent molecular data addressing these possibilities are discussed below.

Proposals for a single, primary origin of plastids from a cyanobacteriumlike endosymbiont have been put foward by Cavalier-Smith (1982, 1987b) and Taylor (1979, 1987). These authors argue that diversity in pigment composition, membrane structure, and other features is due to evolutionary divergence occurring subsequent to the primary endosymbiotic event, rather than being the consequence of multiple origins from correspondingly pigmented and structured bacterial endosymbionts. Differences between chlorophylls *a* and *b* (replacement of a side-chain methyl group by formyl) and *a* and *c* (acquisition of one [c_1] or two [c_2] double bonds) are biochemically rather simple, and the reactions involved are secondary steps in the chlorophyll biosynthetic pathway. Thus, pigment composition differences might readily be accounted for in a monophyletic scheme by the evolutionary gain or loss of individual genes specifying reactions that modify a primary biochemical pathway. Whether differences in membrane ultrastructure can be so readily rationalized is another matter (Cox, 1986; Miller and Jacob, 1989). The issues of diversity in pigment composition and membrane ultrastructure, and whether they occurred prior to or subsequent to the endosymbiosis(es) that gave rise to the various plastid types, are addressed in detail in alternative monophyletic (Cavalier-Smith, 1982) and polyphyletic (Whatley, 1981; Whatley and Whatley, 1981) schemes of plastid origin.

V. PLASTIDS FROM EUKARYOTE RATHER THAN PROKARYOTE ENDOSYMBIONTS

In all rhodophytes, chlorophytes, and metaphytes, and in some dinoflagellates, the plastid is bounded by two membranes. In other groups of algae, the plastid is surrounded by additional membranes (Table I). The double membrane is usually considered to be an evolutionary chimera, with the inner membrane representing the endosymbiont's plasma membrane and the outer membrane the phagosome (food vacuole) membrane of the host (Schnepf, 1964). However, because both cyanobacteria and PNS bacteria have a second lipoprotein outer membrane in addition to the plasma membrane (Benz, 1985), Cavalier-Smith (1987b) has argued that the plastid double membrane actually represents the two membranes of the original endosymbiont, the host phagosome membrane having been lost. This view is consistent with similarities in permeability properties displayed by the outer membranes of mitochondria and plastids on the one hand and bacteria such as *Escherichia coli* on the other and the presence of porins in all three (Benz, 1985). This scenario does demand the selective loss of any peptidoglycan layer that existed between the two lipoprotein membranes of the original endosymbiont.

A particularly intriguing organism in this regard is the photoautotrophic protist *Cyanophora paradoxa*, which harbors photosynthetic "cyanelles" that lack chlorophyll *b* but contain phycobilins. The cyanelle functions as a chloroplast, but unlike *bona fide* plastids, it possesses a rudimentary peptidoglycan wall. This has prompted the suggestion that the cyanelle may represent a transitional form between cyanobacterium and plastid, the product of a relatively recent endosymbiotic event (Hall and Claus, 1963). In Cavalier-Smith's view, however, the cyanelle is simply one of the diverged products of a single original endosymbiosis, having retained the peptidoglycan layer that he assumes was present in the cyanobacterialike progenitor of plastids (Cavalier-Smith, 1982). Consistent with the idea that the *Cyanophora* cyanelle is more plastid than cyanobacterium is the fact that its genome is the same size as that of "regular" plastids (Herdman and Stanier, 1977; Lambert *et al.*, 1985).

Cavalier-Smith supposes that in those algae whose plastids have three membranes (euglenophytes and some dinoflagellates), the host phagosome membrane has been retained. On the other hand, John and the Whatleys (Whatley *et al.*, 1979; Whatley, 1981) consider that euglenophyte plastids evolved by the selective uptake of plastids from other (chl

a/b) eukaryotes, while Gibbs (1978, 1981) has argued that *Euglena* plastids were acquired by uptake of an entire chl *a/b* alga in a secondary symbiosis. According to Gibbs' hypothesis, the engulfed eukaryote endosymbiont was subsequently reduced to the point where only its plastid and plasmalemma remained, the surrounding phagosome membrane of the host being lost.

The secondary acquisition of eukaryote symbionts, followed by their reduction to different degrees, is widely believed to account for the origin of the four-membraned plastids of chromophytes (Whatley, 1981; Whatley and Whatley, 1981; Gibbs, 1981). In these cases, the inner two membranes are thought to represent the double membrane of the symbiont's plastid, the next membrane the plasmalemma of the symbiont, and the outermost membrane the phagosome membrane of the host eukaryotic cell. The latter two membranes constitute what has been termed the chloroplast endoplasmic reticulum (CER)(Gibbs, 1962).

Provocative support for the foregoing scheme comes from the discovery and characterization of the nucleomorph (Greenwood *et al.*, 1977), a body in cryptophyte (cryptomonad) algae that lies within the periplastidal compartment (between the CER and the inner two plastid membranes) and is postulated to be the vestigial nucleus of a eukaryote symbiont (Greenwood *et al.*, 1977; Whatley *et al.*, 1979; Gillott and Gibbs, 1980). The presumption that the nucleomorph is the evolutionary remnant of a nucleus has been strengthened by the demonstration that it contains DNA (Ludwig and Gibbs, 1985, 1987) although it remains to be determined what genetic information, if any, this DNA encodes. If nucleomorph DNA does contain genes (e.g., for the RNA components of the 80S-sized ribosomes that are present within the periplastidal compartment [Ludwig and Gibbs, 1985]), such genes should yield valuable information about the nature of the endosymbiont that provided the cryptomonad plastid. In fact, it has recently been demonstrated (McFadden, 1990) that eukaryotic nuclear-type rRNA genes are present in a nucleoluslike structure in the nucleomorph of the cryptomonad *Chroomonas caudata*.

Loss of the nucleomorph-containing compartment is supposed to have led directly to the four concentric membranes surrounding the plastids of other chromophyte algae, such as the chrysophytes (Gibbs, 1981; Cavalier-Smith, 1982). The demonstration of symbiotic associations in which a dinoflagellate is host to either a crysophyte (Tomas and Cox, 1973; Jeffrey and Vesk, 1976) or cryptomonad (Wilcox and Wedemayer, 1984) symbiont opens the possibility that plastid types have originated via a number of different routes involving eukaryote rather than prokaryote endosymbionts (e.g., Wilcox and Wedemayer, 1985).

VI. FROM ENDOSYMBIONT GENOME TO HOST NUCLEAR DNA: TRANSFER OF GENETIC INFORMATION

In size and genetic information content, plastid genomes are at least an order of magnitude smaller than the average eubacterial genome (Cattolico, 1986). Assuming that the genome of the protoplastid endosymbiont had many more genes than currently exist in plDNA, and recognizing that the majority of the genes specifying plastid structure and function are in fact nuclear genes, it seems clear that massive loss and/or transfer of genetic information from the endosymbiont genome must have occurred in the course of organelle evolution. Nuclear genes for organellar biogenesis may have had either of two origins: (1) they may represent genes that were already present in the host genome and that happened to duplicate endosymbiont genes, in which case the redundant endosymbiont genes could have been lost, or (2) they may have been transferred from endosymbiont genome to the host nuclear genome. Pathways and possible mechanisms for such transfer have been discussed (Bogorad, 1975, 1982; Weeden, 1981; Harington and Thornley, 1982; Obar and Green, 1985) and are postulated to have involved duplication of an endosymbiont gene in the host nuclear genome (perhaps initially by random incorporation of portions of the endosymbiont genome into the host genome), acquisition of function by the nuclear copy (whose activation would have involved elaboration of a transport mechanism to move the gene product back into the organelle), selection of the nuclear gene over the organellar gene for ultimate functioning, and eventual loss of the nonfunctional organellar gene. In the case of plastids, such transfer can account for the existence of nucleus-encoded, plastid-specific isozymes that are distinct from their nucleus-encoded, cytosol counterparts and that do not exist in animal cells (Weeden, 1981). This scenario is strongly supported in the case of glyceraldehyde-3-phosphate dehydrogenase by sequence analysis of nuclear genes encoding the cytosolic and plastid forms of this enzyme (Shih *et al.*, 1986; Martin and Cerff, 1986; Brinkmann *et al.*, 1987).

The lower genetic information content of mitochondrial DNA (mtDNA) compared with plDNA is often taken as a reflection of a longer residence of the mitochondrion in the eukaryotic host, with a consequent greater loss or transfer to the nucleus of genetic information originally present in the protomitochondrial endosymbiont. Cavalier-Smith (1987b), however, argues that most of the symbionts' genetic information was transferred to the host in a "burst of quantum evolution during the origin of the organelles," and he explains the smaller number of

genes remaining in mtDNA than in plDNA by supposing that fewer transfers of genes from protomitochondrion to nucleus were necessary to allow their eventual deletion from mtDNA (the host already having most of the proteins now present in the mitochondrion). Because so few plastid genomes have been analyzed in detail (particularly among non-chlorophyte algae), it is a little risky at this point to assume that little change in plDNA gene content occurred subsequent to a massive initial loss and/or transfer of endosymbiont genes, especially in view of the fact that both unusually large (Padmanabhan and Green, 1978) and unusually small (Hedberg *et al.*, 1981; Manhart *et al.*, 1989) plastid genomes have been reported. Indeed, evidence of a relatively recent evolutionary transfer of a chloroplast gene to the nucleus has been reported (Baldauf and Palmer, 1990; Baldauf *et al.*, 1990).

VII. MOLECULAR DATA BEARING ON THE ORIGIN OF PLASTIDS

The foregoing discussion provides an overview of some of the unresolved issues in plastid evolution and the sometimes disparate views expressed about them. This summary serves to emphasize that current models of serial versus simultaneous and single versus multiple origins of plastids rely heavily on comparison of morphological and ultrastructural characteristics. In these instances, it is difficult to distinguish between primitive and derived character states, and to determine whether shared characteristics reflect descent from a common ancestor (homology) or convergent evolution (analogy) (for a succinct commentary on these issues, see Uzzell and Spolsky, 1981). A case in point is the presumptive evolutionary origin of the various membranes surrounding the plastid in different eukaryotic groups, where assertions of "homology" between a particular plastid membrane and a symbiont or host membrane are often made on the basis of little or no structural and/or functional data to support such inferences. Just as molecular data have been instrumental in affirming a eubacterial (specifically cyanobacterial) origin of plastids, it seems that decisions about subsequent pathways and mechanisms of plastid evolution will ultimately have to be based on judicious comparative analyses of the organization, expression, and sequence of a diversity of plDNAs. This theme will be taken up again following a brief review of the cyanobacterial character of the plastid genetic system.

A. Gene Organization

Much information on plastid genome organization and expression has been compiled in a number of recent reviews (Bohnert *et al.*, 1982; Whitfeld and Bottomley, 1983; Palmer, 1985a,b, 1987; Weil, 1987; Hanley-Bowdoin and Chua, 1987; Gruissem, 1989) and elsewhere in this volume (Section I). Our knowledge of the plastid genome has increased dramatically as a result of the determination of the complete nucleotide sequences of liverwort (*Marchantia polymorpha*) (Ohyama *et al.*, 1986), tobacco (*Nicotiana tabacum*) (Shinozaki *et al.*, 1986) and rice (*Oryza sativa*) (Hiratsuka *et al.*, 1989) plDNAs. The nature and overall organization of genetic information in these plDNAs is essentially prokaryotic, with a number of examples of clustering and cotranscription of functionally related genes. More significantly, in several instances a striking similarity exists in gene order between corresponding chloroplast and eubacterial genes, with homologs of several *E. coli* operons (e.g., *rrn* [rRNA genes], S10-*spc*-α [mostly ribosomal protein genes], and *unc* [genes for ATP synthase subunits]) readily identifiable. Similarly conserved linkage groups have been found not only in metaphyte plDNAs (Palmer, 1985b), but also in those of chlorophytes (Palmer, 1985b) euglenophytes (Christopher *et al.*, 1988) and chromophytes (Janssen *et al.*, 1987; Markowicz *et al.*, 1988b; Delaney and Cattolico, 1989; Evrard *et al.*, 1990; Douglas, 1991; Douglas and Durnford, 1990). Coupled with the presence of many photosynthetic genes and a number of other open reading frames that appear to be homologous with identified bacterial genes, it is easy to view the plastid genome as a highly condensed (cyano)bacterial genome.

B. Gene Expression

Besides being organized into polycistronic transcription units, plastid genes are expressed by way of bacterialike promoter motifs (Hanley-Bowdoin and Chua, 1987; Gruissem, 1989). The 5′ regions of many of the identified transcription units contain sequences that resemble the "−10" and "−35" consensus promoter elements of eubacteria (Steinmetz *et al.*, 1983; Briat *et al.*, 1986; Hanley-Bowdoin and Chua, 1987), and the functional importance of such sequences has been demonstrated *in vitro* (Link, 1984; see Gruissem, 1989). The finding in plDNA of genes homologous to the *rpoA*, *rpoB*, and *rpoC* loci of *E. coli*, which encode the α, β, and β' subunits of the eubacterial RNA polymerase, provides suggestive evidence that a eubacterialike RNA polymerase

functions in transcription in plastids (Gruissem, 1989; also see Chapter 4A).

The plastid translation system is strikingly eubacterial in its properties, and these strong similarities provided some of the earliest molecular evidence pointing to a bacterial origin of plastids (Margulis, 1970; Gillham and Boynton, 1981; Gray and Doolittle, 1982). Plastid ribosomes (70S) are composed of 50S and 30S subunits containing, respectively, high-molecular weight RNA species 23S and 16S in size. Both structural and immunological similarity has been demonstrated between eubacterial and plastid ribosomal proteins (Schmidt *et al.*, 1984; Chapter 8A). Functional similarity between plastid and eubacterial ribosomes is evidenced by interchangeability of ribosomal subunits and translation factors, use of *N*-formylmethionyl-tRNA as initiator of protein synthesis, and a similar spectrum of antibiotic sensitivity and resistance (Gray and Doolittle, 1982; see also Chapter 8A).

C. Gene Structure

Various aspects of plastid genome organization and expression testify to its essentially prokaryotic nature. However, a number of these features, such as similarities in the organization of certain operons and ribosome/rRNA size, are shared by both eubacteria and archaebacteria (e.g., Auer *et al.*, 1989), and so are presumably primitive traits for these groups. Gene sequence, on the other hand, differentiates convincingly between eubacteria plus plastids on the one hand and archaebacteria on the other. Sequence information, both nucleotide and amino acid, also clearly places the plastids within the cyanobacteria phylum of eubacteria (see Section VIII).

Protein sequence data have been used to construct phylogenetic trees (e.g., of *c*-type cytochromes) that show a specific clustering of plastids with cyanobacteria (Schwartz and Dayhoff, 1978; Hunt *et al.*, 1985), to the exclusion of other types of eubacteria, including other photosynthetic species. Where clear homologs of plastid-encoded genes exist in other eubacteria as well as in cyanobacteria, the plastid genes display highest sequence similarity with their cyanobacterial counterparts (e.g., Cozens *et al.*, 1986). The cyanobacterial character of plastid 16S, 23S, and 5S rRNAs and tRNAs is documented elsewhere in this volume (Chapters 6A and 7A). The specific similarities between homologous plastid and cyanobacterial rRNAs are evident at the level of secondary structure as well as primary sequence (Delihas *et al.*, 1982, 1985; Tomioka and Sugiura, 1983; Kumano *et al.*, 1983; Douglas and Doolittle, 1984a,b; Delihas and Fox, 1987).

VIII. PLASTID AND NUCLEAR PHYLOGENIES BASED ON RIBOSOMAL RNA SEQUENCE COMPARISONS

Molecular data used to infer plastid evolutionary descent may be grouped into several categories, according to whether the character states being compared distinguish (1) between prokaryotes (archaebacteria and eubacteria) and eukaryotes, (2) between eubacteria and archaebacteria, (3) among different eubacterial phyla, and (4) among different divisions of cyanobacteria. Data in categories (1) and (2) have been used to argue in support of a prokaryotic (specifically eubacterial) origin of both plastids and mitochondria, whereas data in category (3) have served to trace the evolutionary origin of mitochondria and plastids to different eubacterial phyla. Category (4) data, of which very little exists at present, should ultimately allow us to decide whether the plastid genome is monophyletic or polyphyletic in origin.

Of the types of molecular evidence that have been used to assess organelle origins and evolution, only a few potentially can encompass the four categories outlined above, i.e., to define both long-range and short-range evolutionary relationships. Among the approaches being used to generate this kind of incisive information, rRNA sequence comparison is particularly attractive. Small subunit (16S-like) and large subunit (23S-like) rRNAs have a number of properties that make them especially useful as phylogenetic probes (the "ultimate molecular chronometers," according to Woese [1987], including the fact that they are evolutionarily ancient, are ubiquitous (present in all genomes), and contain a conserved core of secondary structure that provides a basis for accurate alignment of primary sequence (Gray *et al.*, 1984; Gray, 1988; Cedergren *et al.*, 1988). In the particular context of exploring plastid evolution, rRNA sequence comparisons allow one to determine, simultaneously, nuclear and plastid phylogenies; i.e., they offer the possibility of assessing whether or not the genomes of host and organelle have evolved in parallel.

Observations on and conclusions about algal genome evolution that have emerged from small and large subunit rRNA tree construction include the following:

1. In global trees that define three separate and distinct lines of descent (archaebacterial, eubacterial, and eukaryotic nuclear), plastids clearly fall within the eubacterial lineage and specifically cluster with cyanobacteria (Gray *et al.*, 1984; Giovannoni *et al.*, 1988; Cedergren *et al.*, 1988; Turner *et al.*, 1989).

2. Plastids of the chl *a/b* type form a monophyletic group that is contained within the cyanobacterial radiation of eubacteria (Giovannoni *et*

al., 1988). This includes plastids of representative metaphytes (maize, tobacco, and liverwort), a chlorophyte (*C. reinhardtii*), and a euglenophyte (*E. gracilis*). This strongly indicates a common evolutionary ancestry of chlorophyte and euglenophyte plastids, despite some notable differences (Palmer, 1985a) in genome organization in the two.

3. *Prochlorothrix hollandica*, a free-living "prochlorophyte" that contains chlorophylls *a* and *b* but no phycobilins (Burger-Wiersma *et al.*, 1986), is a deeply branching member of the cyanobacteria, not specifically related to the coherent phylogenetic grouping of the green (chl *a*/*b*) chloroplasts (Turner *et al.*, 1989). This conclusion is consistent with analyses of 5S (MacKay *et al.*, 1982) and 16S (Seewaldt and Stackebrandt, 1982) rRNA from an exoymbiotic prochlorophyte, *Prochloron* (although the interpretation of the 16S results, based on RNase T1-generated oligonucleotide catalogs, has been challenged on methodological grounds [Van Valen, 1982]). The conclusion of Turner *et al.* (1989) was initially at odds with that of Morden and Golden (1989), whose analysis of *psbA* genes seemed to indiicate a common ancestry of prochlorophytes and *a*/*b* plastids. However, a subsequent reassessment of the *psbA* analysis does not support this interpretation (see Gray, 1989). The separate branching of prochlorophytes and green plastids implies that the ability to synthesize chlorophyll *b* arose independently in the ancestors of these two groups or that it was acquired by lateral transfer (Turner *et al.*, 1989).

4. Other rRNA sequence comparisons (Witt and Stackebrandt, 1988) have demonstrated that there is not a specific relationship between the chl *a*/*c* plastid of the chromophyte *Ochromonas danica* (phylum Chrysophyta) and the brownish photoheterotrophic eubacterium *Heliobacterium chlorum*. These results do not support the proposal (Margulis and Obar, 1985) of an endosymbiotic origin of the chrysophyte plastid from an ancestor that also gave rise to *H. chlorum*.

5. The *C. paradoxa* cyanelle (which contains chlorophyll *a* and phycobilins) is the closest relative of the green chloroplast lineage (Giovannoni *et al.*, 1988; Turner *et al.*, 1989). Significantly, the cyanelle represents an early diverging line within the plastid branch, splitting off before the diversification of the chl *a*/*b* plastids. This supports Cavalier-Smith's contention that the cyanelle, although structurally more conservative than other plastids, is not of recent endosymbiotic origin. As well, these 16S rRNA trees (Turner *et al.*, 1989) provide no evidence as yet of a polyphyletic origin of the plastid genome. A more recent and comprehensive phylogenetic analysis of plastid 16S rRNA sequences (Douglas and Turner, 1991) also suggests a common origin of all plastids from a cyanobacteriumlike ancestor, as well as an early dichotomy between the metaphytes/chlorophytes on the one hand and rhodophytes/chromophytes

on the other. In this tree, *C. paradoxa* represents the earliest branching in the latter group.

6. The chl *a/c* plastid of a brown alga, *Pylaiella littoralis*, appears phylogenetically closest to the chl *a/b* plastid of *E. gracilis* in a 16S rRNA tree that also includes chlorophyte (*C. reinhardtii*) and metaphyte (maize) plastids and the cyanobacterium *Anacystis nidulans* (Markowicz *et al.*, 1988a). This tree shows chlorophyte–metaphyte plastids emerging from the cyanobacterial branch earlier than the euglenophyte plastid, in contradistinction to the 16S rRNA tree of Turner *et al.* (1989). The plastid 16S rRNA tree of Witt and Stackebrandt (1988) also places *E. gracilis* together with a chromophyte (*O. danica*), but in this case the two branch *before* the chlorophyte/metaphyte clade. A specific euglenophyte/chromophyte association is also seen in the plastid 16S rRNA tree of Douglas and Turner (1991). Taken at face value, these data would seem to imply that chlorophyll *b* evolved twice within the plastid lineage: once in the chlorophyte/metaphyte line and once in the euglenoids. Until there is a more phylogentically representative set of complete 16S rRNA sequences (from photosynthetic bacteria, particularly cyanobacteria, as well as from plastids), and until statistical evaluation of the robustness of branching topologies becomes the norm, we will have to be cautious about the specific conclusions we draw from rRNA trees about plastid interrelationships.

7. Trees based on algal *nuclear* rRNA sequences, i.e., tracing the evolutionary descent of the host rather than the organelle genome, indicate that the eukaryotic algae do not form a coherent group relative to nonphotosynthetic eukaryotes (Gunderson *et al.*, 1987; Bhattacharya and Druehl, 1988; Perasso *et al.*, 1989; Bhattacharya *et al.*, 1990). Two features in these trees are of particular interest. (1) *E. gracilis* is a very early diverging protist, which separated from the main eukaryotic line well before chlorophytes and metaphytes, and whose closest relatives are the trypanosomatid protozoa. The noncongruent phylogenetic position of *Euglena* nuclear and plastid genomes are most readily explained by postulating that the *Euglena* plastid originated from a eukaryotic cell directly on the line to chlorophytes and metaphytes, and not from a free-living cyanophyte. (2) Red algae (specifically *Porphyrideum purpureum* (Perasso *et al.*, 1989) and *Gracilaria lemaneiformis* (Bhattacharya *et al.*, 1990)) are relatively advanced, diverging from the eukaryotic line of descent at about the same time as plants, fungi, animals, and most protists. Rhodophytes appear to radiate at approximately the same time as other algal groups, an observation seemingly at odds with schemes (e.g., Cavalier-Smith, 1982) in which red algae are ancestral to all other algal types. On the other hand, it may still be possible to maintain a single primary ori-

gin of algal *plastids* followed by secondary loss in some lines and marked diversification in plastid-containing lineages, especially if eukaryotic endosymbioses are invoked to account for the origin of plastids having more than two membranes (Table I). As noted above, there is still no evidence from rRNA sequence such comparisons of a polyphyletic orgin of the plastid genome itself, although provocative indications of plastid polyphyly have come from recent analyses of chloroplast-encoded *rbcS* and *rbcL* (Rubisco subunit) genes in chromophyte algae (Boczar *et al.*, 1989; Assali *et al.*, 1990; Douglas *et al.*, 1990; Valentin and Zetsche, 1990) (see Section X). Based on the fact that the majority of eukaryotic algal phyla branch after a number of nonphotosynthetic eukaryotic groups, Perasso *et al.* (1989) concluded that, whether plastids are monophyletic or polyphyletic, "these models of symbiotic events appear to have occurred at an advanced stage of protistan diversification and hence much later than for the mitochondrial counterpart." This conclusion is clearly at odds with a simultaneous symbiotic origin of mitochondria and chloroplasts (Cavalier-Smith, 1987b).

8. Two distinctly different eukaryotic-type 18S rRNA genes (Nu and Nm), both of which are expressed, have recently been isolated and completely sequenced from the cryptomonad alga, *Cryptomonas* Φ (Douglas *et al.*, 1991). The Nm (presumptive nucleomorph) sequence was found to cluster specifically with nuclear 18S rRNA sequences from red algae. These new data strongly support the idea that cryptophyte algae are evolutionary chimeras of two phylogenetically distinct eukaryotic cells, with the symbiont partner being a rhodophyte-type alga. In the same study, *Cryptomonas* Φ (Nu) and chromophytes were shown to occupy separate branches of the 18S tree, an observation that does not support the hypothesis (Cavalier-Smith, 1982) that chromophyte algae evolved directly from a cryptomonadlike ancestor.

IX. EVOLUTIONARY DIVERGENCE OF PLASTID GENOMES

Accumulating data, especially in metaphytes, are starting to reveal the way in which both overall genome organization and gene sequence diverge in plDNA. Complete plDNA sequences, three of which are now available, provide a particularly rich data base for comparative analysis of genome divergence between monocotyledons (e.g., rice) and dicotyledons (e.g., tobacco) and between angiosperms and nonangiosperms (such as liverwort). Such comparisons, supplemented by mapping and

partial sequence data from other genomes, not only document specific changes that have occurred in plDNA since the divergence of two plant species from a common ancestor, but they also suggest molecular mechanisms by which such changes have occurred (e.g., Ogihara *et al.*, 1988; Hiratsuka *et al.*, 1989). Specific aspects of plDNA divergence in land plants have been discussed in detail in previous reviews (Bohnert *et al.*, 1982; Whitfeld and Bottomley, 1983; Palmer, 1985a,b, 1987) and are addressed elsewhere in this volume (Chapter 2A). Documented changes include loss of one copy of the inverted repeat that characterizes most plDNAs, with concomitant genome rearrangement, in several species of legume (Palmer and Thompson, 1982) and in conifers (Strauss *et al.*, 1988; Lidholm *et al.*, 1988); expansion of the inverted repeat into single-copy regions (Palmer *et al.*, 1987; Hiratsuka *et al.*, 1989); and a variety of additions, deletions, and inversions (Howe, 1985; Michalowski *et al.*, 1987; see also Palmer, 1985b). Such mapping and sequencing data have also provided evidence for selective loss (transfer?) of plDNA-encoded genes (e.g., tobacco plDNA encodes ribosomal protein S16, which is not present in liverwort plDNA, whereas the latter carries a gene for ribosomal protein L21, which is not present in the tobacco plastid genome [Gray, 1986]).

Information about plastid genome arrangement and divergence in the eukaryotic algae is still quite limited, and only a few rather tentative evolutionary conclusions can be drawn at present. In various species of the chlorophyte *Chlamydomonas*, the plDNA displays the basic arrangement of two inverted rRNA gene-containing repeats found in metaphyte plDNA (Palmer, 1985b). However, even within this genus, extensive sequence rearrangements of plDNA are evident (as, for example, between *C. reinhardtii* and *C. eugametos* [Lemieux and Lemieux, 1985]). In *E. gracilis* plDNA, rRNA genes are arrayed tandemly rather than as inverted repeats, an arrangement so far unique among characterized plastid genomes (Palmer, 1985b). This difference in rRNA gene organization is consistent with a deep evolutionary split between the chlorophyte–metaphyte and euglenophyte plastid lineages, as suggested by rRNA phylogenetic trees (Turner *et al.*, 1989). However, as noted previously, the plDNA of another chlorophyte, that of *Codium fragile*, is unusually small and contains only one copy of each rRNA gene, lacking any large repeats at all (Manhart *et al.*, 1989), whereas in *Chlorella ellipsoidea* plDNA, the 23S rRNA gene is inverted relative to the 16S gene (Yamada and Shimaji, 1987). Obviously, definitive evolutionary conclusions cannot be drawn from such a sparse data base.

The same limitations apply when considering the plastid genomes of

chl *a*/PB and chl *a*/*c* algae: Only fragmentary data on genome organization and divergence exist (Cattolico, 1985, 1986; Kuhsel and Kowallik, 1987; Li and Cattolico, 1987; Douglas, 1988), and so far this information provides few solid evolutionary insights; however, two features are noteworthy. (1) The gene for the small subunit of ribulose-1,5-bisphosphate carboxylase–oxygenase (*rbcS*) is encoded in plDNA in those chl *a*/PB and chl *a*/*c* algae that have been examined and appears to be cotranscribed with *rbcL*, the gene for the large subunit of the same enzyme (Starnes *et al.*, 1985; Reith and Cattolico, 1986; Douglas and Durnford, 1989; Valentin and Zetsche, 1989, 1990), as in cyanobacteria (Shinozaki and Sugiura, 1983; Nierzwicki-Bauer *et al.*, 1984). This contrasts with the situation in chlorophytes and metaphytes, where *rbcS* is a nuclear gene. (2) The plastid 16S–23S rDNA spacer region is highly condensed in chl *a*/PB and chl *a*/*c* algae, being only 287 bp in *C. paradoxa* (Janssen *et al.*, 1987), 322 bp in *P. littoralis* (Markowicz *et al.*, 1988b), and 265 bp in *Olisthodiscus luteus* (Delaney and Cattolico, 1989) and containing unsplit isoleucine and alanine tRNA genes. This arrangement is very similar to rDNA spacer organization in eubacteria, including *A. nidulans* (Williamson and Doolittle, 1983), whereas in chlorophyte algae and land plants the rDNA spacer is considerably enlarged by the presence of repetitive elements, introns, and open reading frames (see Markowicz *et al.*, 1988b; Delaney and Cattolico, 1989). Although it is premature to make too much of these data, they do give the impression that plDNA may be organized in a more "primitive" fashion in chromophyte algae than in chlorophytes, euglenophytes, and metaphytes.

X. FUTURE PROSPECTS

Are plastids monophyletic or polyphyletic? Despite a great deal of argumentation and speculation over the past 25 years, I submit that this question cannot be answered definitively at the present time. Much of the earlier molecular data have been interpreted as supporting a polyphyletic origin of plastids; this includes comparisons of RNase T1 oligonucleotide catalogs of 16S rRNA (Bonen and Doolittle, 1975; Doolittle and Bonen, 1981; but see Zablen *et al.*, 1975) and protein sequences (Schwartz and Dayhoff, 1978, 1981; Hunt *et al.*, 1985). Phylogenetic trees based on more complete rRNA sequence information seem to favor the idea of a monophyletic origin of plastids (Turner *et al.*, 1989; Douglas

and Turner, 1991; but see Markowicz *et al.*, 1988a). Not infrequently, contradictory indications of phylogenetic origin have been obtained from analyses of different plastid genes in the same organism. Particularly intriguing in this regard are recent analyses of plastid-encoded *rbcS* and *rbcL* genes in chl *a*/*c* algae. These genes appear specifically related to their homologs in β-purple bacteria, whereas the *rbcL* gene in chl *a*/*b* plastids, as well as the rRNA genes in all three plastid types, are specifically related to their homologs in cyanobacteria (Boczar *et al.*, 1989; Assali *et al.*, 1990; Douglas *et al.*, 1990; Valentin and Zetsche, 1990). These observations are perhaps the strongest indications yet of a polyphyletic origin of at least some plastid genes (Douglas *et al.*, 1990; Valentin and Zetsche, 1990), suggesting that the chromophyte plastid genome is a phylogenetic mosaic, resulting either from mixed endosymbiosis events or from horizontal gene transfer (Assali *et al.*, 1990).

Although certain aspects of plastid genome evolution have been greatly illuminated by the molecular data that have been generated in recent years, the continuing uncertainty about plastid origin(s) is to a great extent a reflection of the spotty nature of the current data base. With few exceptions, molecular data have been produced for other than evolutionary reasons, and much of the information has been derived from relatively few organisms (heavily skewed toward chl *a*/*b* types). So far, a systematic approach to the problem of reconstructing a phylogeny of plastid genomes, one that fully recognizes the evolutionary diversity of plastids and their hosts, is lacking.

There is, however, every reason to believe that a sufficient depth and breadth of molecular information will eventually provide the insights we currently lack. To provide a solid framework for future informed speculation about plastid genome origin and evolution, the most productive approach would seem to be phylogenetic tree construction based on comparison of small subunit and large subunit rRNA sequences. This approach, as noted previously, will yield parallel phylogenies of plastid and nuclear genomes, as well as provide definitive connections between plastids and various branches of the cyanobacterial evolutionary tree. Such information should be supplemented by trees of conserved protein sequences, also generated in a systematic manner. Plastid ribosomal protein genes, for which there are also nuclear and bacterial homologs, would seem to be a particularly good choice (e.g., Evrard *et al.*, 1990), but plastid-specific protein genes such as *rbcS* and *rbcL* should also prove extremely informative, as recent reports have indicated (Boczar *et al.*, 1989; Assali *et al.*, 1990; Douglas *et al.*, 1990; Valentin and Zetsche, 1990). The evolutionary framework provided by rRNA and protein trees

is, I believe, an essential prerequisite for rational integration of existing morphological, ultrastructural, and biochemical data into a compelling scheme of plastid evolution.

Finally, a major effort should be mounted to determine the complete sequences of a number of algal plDNAs, for only in this way will we obtain a detailed picture of what has happened to the plastid genome during its evolution, and by what mechanisms. The complete sequences of three plant plDNAs have provided a wealth of information that is relevant to our understanding of both function and evolution. The complete sequences of a dozen or more key algal plDNAs would have no less an impact. In view of the nature of genetic information encoded in plastid genomes, their relatively small size, and their relative conservatism, we have an unparalleled opportunity (and the tools at hand) to chart the evolutionary course of an entire cellular genome.

REFERENCES

Assali, N.-E., Mache, R., and Loiseaux-de Göer, S. (1990). Evidence for a composite phylogenetic origin of the plastid genome of the brown alga *Pylaiella littoralis* (L.) Kjellm. *Plant Mol. Biol.* **15,** 307–315.

Auer, J., Spicker, S., and Böck, A. (1989). Organization and structure of the *Methanococcus* transcriptional unit homologous to the *Escherichia coli* "spectinomycin" operon. Implications for the evolutionary relationship of 70 S and 80 S ribosomes. *J. Mol. Biol.* **209,** 21–36.

Baldauf, S. L., and Palmer, J. D. (1990). Evolutionary transfer of the chloroplast *tufA* gene to the nucleus. *Nature (London)* **344,** 262–265.

Baldauf, S. L., Manhart, J. R., and Palmer, J. D. (1990). Different fates of the chloroplast *tufA* gene following its transfer to the nucleus in green algae. *Proc. Natl. Acad. Sci. USA* **87,** 5317–5321.

Benz, R. (1985). Porin from bacterial and mitochondrial outer membranes. *CRC Crit. Rev. Biochem.* **19,** 145–190.

Bhattacharya, D., and Druehl, L. D. (1988). Phylogenetic comparison of the small-subunit ribosomal DNA sequence of *Costaria costata* (Phaeophyta) with those of other algae, vascular plants and oomycetes. *J. Phycol.* **24,** 539–543.

Bhattacharya, D., Elwood, H. J., Goff, L. J., and Sogin, M. L. (1990). Phylogeny of *Gracilaria lemaneiformis* (Rhodophyta) based on sequence analysis of its small subunit ribosomal RNA coding region. *J. Phycol.* **26,** 181–186.

Boczar, B. A., Delaney, T. P., and Cattolico, R. A. (1989). Gene for the ribulose-1,5-bisphosphate carboxylase small subunit protein of the marine chromophyte *Olisthodiscus luteus* is similar to that of a chemoautotrophic bacterium. *Proc. Natl. Acad. Sci. USA* **86,** 4996–4999.

Bogorad, L. (1975). Evolution of organelles and eukaryotic genomes. *Science* **188,** 891–898.

Bogorad, L. (1982). Regulation of intracellular gene flow in the evolution of eukaryotic genomes. *In* "On the Origins of Chloroplasts," (J. A. Schiff, ed.) pp. 277–295. Elsevier/North-Holland, Amsterdam.

Bohnert, H. J., Crouse, E. J., and Schmitt, J. M. (1982). Organization and expression of plastid genomes. *In* "Nucleic Acids and Proteins in Plants II. Structure, Biochemistry and Physiology of Nucleic Acids," Vol. 14B (B. Parthier and D. Boulter, eds.), pp. 475–530. *Encycl. Plant Physiol., New Ser.* Springer-Verlag, Berlin, Heidelberg, and New York.

Bonen, L., and Doolittle, W. F. (1975). On the prokaryotic nature of red algal chloroplasts. *Proc. Natl. Acad. Sci. USA* **72,** 2310–2314.

Briat, J. F., Lescure, A. M., and Mache, R. (1986). Transcription of the chloroplast DNA: A review. *Biochimie* **68.** 981–990.

Brinkmann, H., Martinez, P., Quigley, F., Martin, W., and Cerff, R. (1987). Endosymbiotic origin and codon bias of the nuclear gene for chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize. *J. Mol. Evol.* **26,** 320–328.

Burger-Wiersma, T., Veenhuis, M., Korthals, H. J., Van de Wiel, C. C. M., and Mur, L. R. (1986). A new prokaryote containing chlorophylls *a* and *b*. *Nature* **320,** 262–264.

Cattolico, R. A. (1985). Chloroplast biosystematics: Chloroplast DNA as a molecular probe. *BioSystems* **18,** 299–306.

Cattolico, R. A. (1986). Chloroplast evolution in algae and land plants. *Trends Ecol. Evol.* **1,** 64–67.

Cavalier-Smith, T. (1982). The origins of plastids. *Biol. J. Linn. Soc.* **17,** 289–306.

Cavalier-Smith, T. (1987a). The origin of eukaryote and archaebacterial cells. *Ann. N.Y. Acad. Sci.* **503,** 17–54.

Cavalier-Smith, T. (1987b). The simultaneous symbiotic origin of mitochondria, chloroplasts, and microbodies. *Ann. N.Y. Acad. Sci.* **503,** 55–71.

Cedergren, R., Gray, M. W., Abel, Y., and Sankoff, D. (1988). The evolutionary relationships among known life forms. *J. Mol. Evol.* **28,** 98–112.

Chisholm, S. W., Olson, R. J., Zettler, E. R., Goericke, R., Waterbury, J. B., and Welschmeyer, N. A. (1988). A novel free-living prochlorophyte abundant in the oceanic euphotic zone. *Nature* **334,** 340–343.

Christensen, T. (1964). The gross classification of algae. *In* "Algae and Man" (D. F. Jackson, ed.), pp. 59–64. Plenum, New York.

Christopher, D. A., Cushman, J. C., Price, C. A., and Hallick, R. B. (1988). Organization of ribosomal protein genes *rpl23*, *rpl2*, *rps19*, *rpl22* and *rps3* on the *Euglena gracilis* chloroplast genome. *Curr. Genet.* **14,** 275–286.

Chun, E. H. L., Vaughan, M. H., Jr., and Rich, A. (1963). The isolation and characterization of DNA associated with chloroplast preparations. *J. Mol. Biol.* **7,** 130–141.

Cox, G. (1986). The origins of chloroplasts in eukaryotes. *Nature* **322,** 412.

Cozens, A. L., Walker, J. E., Phillips, A. L., Huttly, A. K., and Gray, J. C. (1986). A sixth subunit of ATP synthase, an F_0 component, is encoded in the pea chloroplast genome. *EMBO J.* **5,** 217–222.

Delaney, T. P., and Cattolico, R. A. (1989). Chloroplast ribosomal DNA organization in the chromophytic alga *Olisthodiscus luteus*. *Curr. Genet.* **15,** 221–229.

Delihas, N., and Fox, G. E. (1987). Origins of the plant chloroplasts and mitochondria based on comparisons of 5S ribosomal RNAs. *Ann. N.Y. Acad. Sci.* **503,** 92–102.

Delihas, N., Andresini, W., Andersen, J., and Berns, D. (1982). Structural features unique to the 5 S ribosomal RNAs of the thermophilic cyanobacterium *Synechococcus lividus* III and the green plant chloroplasts. *J. Mol. Biol.* **162,** 721–727.

Delihas, N., Andersen, J., and Berns, D. (1985). Phylogeny of the 5S ribosomal RNA from *Synechococcus lividis* II: The cyanobacterial/chloroplast 5S RNAs form a common structural class. *J. Mol. Evol.* **21,** 334–337.

Devereux, R., Loeblich, A. R. III, and Fox, G. E. (1990). Higher plant origins and the phylogeny of green algae. *J. Mol. Evol.* **31,** 18–24.

Doolittle, W. F. (1980). Revolutionary concepts in evolutionary cell biology. *Trends Biochem. Sci.* **5,** 146–149.

Doolittle, W. F., and Bonen, L. (1981). Molecular sequence data indicating an endosymbiotic origin for plastids. *Ann. N.Y. Acad. Sci.* **361,** 248–259.

Douglas, S. E. (1988). Physical mapping of the plastid genome from the chlorophyll *c*-containing alga, *Cryptomonas* Φ. *Curr. Genet.* **14,** 591–598.

Douglas, S. E. (1991). Unusual organization of the *str* operon genes in the plastid genome of *Cryptomonas* Φ: evolutionary considerations. *Curr. Genet.*, in press.

Douglas, S. E., and Doolittle, W. F. (1984a). Complete nucleotide sequence of the 23S rRNA gene of the Cyanobacterium, *Anacystis nidulans*. *Nucleic Acids Res.* **12,** 3373–3386.

Douglas, S. E., and Doolittle, W. F. (1984b). Nucleotide sequence of the 5 S rRNA gene and flanking regions in the cyanobacterium. *Anacystis nidulans*. *FEBS Lett.* **166,** 307–310.

Douglas, S. E., and Durnford, D. G. (1989). The small subunit of ribulose-1,5-bisphosphate carboxylase is plastid-encoded in the chlorophyll *c*-containing alga *Cryptomonas* Φ. *Plant Mol. Biol.* **13,** 13–20.

Douglas, S. E., and Durnford, D. G. (1990). Sequence analysis of the plastid rDNA spacer region of the chlorophyll *c*-containing alga *Cryptomonas* Φ. *DNA Sequence-J. DNA Sequencing and Mapping* **1,** 55–62.

Douglas, S. E., and Turner, S. (1991). Molecular evidence for the origin of plastids from a cyanobacterium-like ancestor. Submitted for publication.

Douglas, S. E., Durnford, D. G., and Morden, C. W. (1990). Nucleotide sequence of the gene for the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Cryptomonas* Φ: Evidence supporting the polyphyletic origin of plastids. *J. Phycol.* **26,** 500–508.

Douglas, S. E., Murphy, C. A., Spencer, D. F., and Gray, M. W. (1991). Cryptomonad algae are evolutionary chimaeras of two phylogenetically distinct unicellular eukaryotes. *Nature (London)*, in press.

Evrad, J. L., Kuntz, M., and Weil, J. H. (1990). The nucleotide sequence of five ribosomal protein genes from the cyanelles of *Cyanophora paradoxa:* Implications concerning the phylogenetic relationship between cyanelles and chloroplasts. *J. Mol. Evol.* **30,** 16–25.

Gest, H., and Favinger, J. L. (1983). *Heliobacterium chlorum*, an anoxygenic brownish-green photosynthetic bacterium containing a "new" form of bacteriochlorophyll. *Arch. Microbiol.* **136,** 11–16.

Gibbs, S. P. (1962). Nuclear envelope-chloroplast relationships in algae. *J. Cell Biol.* **14,** 433–444.

Gibbs, S. P. (1970). The comparative ultrastructure of the algal chloroplast. *Ann. N.Y. Acad. Sci.* **175,** 454–473.

Gibbs, S. P. (1978). The chloroplasts of *Euglena* may have evolved from symbiotic green algae. *Can. J. Bot.* **56,** 2883–2889.

Gibbs, S. P. (1981). The chloroplasts of some algal groups may have evolved from endosymbiotic eukaryotic algae. *Ann. N.Y. Acad. Sci.* **361,** 193–208.

Gillham, N. W., and Boynton, J. E. (1981). Evolution of organelle genomes and protein-synthesizing systems. *Ann. N.Y. Acad. Sci. USA* **361,** 20–43.

Gillott, M. A., and Gibbs, S. P. (1980). The cryptomonad nucleomorph: Its ultrastructure and evolutionary significance. *J. Phycol.* **16,** 558–568.

Giovannoni, S. J., Turner, S., Olsen, G. J., Barns, S., Lane, D. J., and Pace, N. R. (1988). Evolutionary relationships among cyanobacteria and green chloroplasts. *J. Bacteriol.* **170,** 3584–3592.

Gray, J. (1986). Wonders of chloroplast DNA. *Nature* **322,** 501–502.

Gray, M. W. (1983). The bacterial ancestry of plastids and mitochondria. *BioScience* **33,** 693–699.

Gray, M. W. (1988). Organelle origins and ribosomal RNA. *Biochem. Cell Biol.* **66,** 325–348.

Gray, M. W. (1989). The evolutionary origins of organelles. *Trends Genet.* **5,** 294–299.

Gray, M. W., and Doolittle, W. F. (1982). Has the endosymbiont hypothesis been proven? *Microbiol. Rev.* **46,** 1–42.

Gray, M. W., Sankoff, D., and Cedergren, R. J. (1984). On the evolutionary descent of organisms and organelles: A global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. *Nucleic Acids Res.* **12,** 5837–5852.

Greenwood, A. D., Griffiths, H. B., and Santore, U. J. (1977). Chloroplasts and cell compartments in Cryptophyceae. *Br. Phycol. J.* **12,** 119.

Gruissem, W. (1989). Chloroplast gene expression: How plants turn their plastids on. *Cell* **56,** 161–170.

Gunderson, J. H., Elwood, H., Ingold, A., Kindle, K., and Sogin, M. L. (1987). Phylogenetic relationships between chlorophytes, chrysophytes, and oomycetes. *Proc. Natl. Acad. Sci. USA* **84,** 5823–5827.

Hall, W. T., and Claus, G. (1963). Ultrastructural studies on the blue-green algal symbiont in *Cyanophora paradoxa* Korschikoff. *J. Cell Biol.* **19,** 551–563.

Hanley-Bowdoin, L., and Chua, N.-H. (1987). Chloroplast promoters. *Trends Biochem. Sci.* **12,** 67–70.

Harington, A., and Thornley, A. L. (1982). Biochemical and genetic consequences of gene transfer from endosymbiont to host genome. *J. Mol. Evol.* **18,** 287–292.

Hedberg, M. F., Huang, Y.-S., and Hommersand, M. H. (1981). Size of the chloroplast genome in *Codium fragile. Science* **213,** 445–447.

Herdman, M., and Stanier, R. Y. (1977). The cyanelle: Chloroplast or endosymbiotic prokaryote? *FEMS Lett.* **1,** 7–12.

Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.-R., Meng, B.-Y., Li, Y.-Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., and Sugiura, M. (1989). The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.* **217,** 185–194.

Howe, C. J. (1985). The endpoints of an inversion in wheat chloroplast DNA are associated with short repeated sequences containing homology to *att*-lambda. *Curr. Genet.* **10,** 139–145.

Hunt, L. T., George, D. G., and Barker, W. C. (1985). The prokaryote–eukaryote interface. *BioSystems* **18,** 223–240.

Janssen, I., Mucke, H., Löffelhardt, W., and Bohnert, H. J. (1987). The central part of the cyanelle rDNA unit of *Cyanophora paradoxa:* Sequence comparison with chloroplasts and cyanobacteria. *Plant Mol. Biol.* **9,** 479–484.

Jeffrey, S. W., and Vesk, M. (1976). Further evidence for a membrane-bound endosymbiont within the dinoflagellate *Peridinium foliaceum. J. Phycol.* **12,** 450–455.

Kuhsel, M., and Kowallik, K. V. (1987). The plastome of a brown alga, *Dichtyota dichotoma.* II. Location of structural genes coding for ribosomal RNAs, the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase and for polypeptides of photosystems I and II. *Mol. Gen. Genet.* **207,** 361–368.

Kumano, M., Tomioka, N., and Sugiura, M. (1983). The complete nucleotide sequence of a 23S rRNA gene from a blue-green alga, *Anacystis nidulans*. *Gene* **24,** 219–225.

Lambert, D. H., Bryant, D. A., Stirewalt, V. L., Dubbs, J. M., Stevens, S. E. Jr., and Porter, R. D. (1985). Gene map for the *Cyanophora paradoxa* cyanelle genome. *J. Bacteriol.* **164,** 659–664.

Lemieux, B., and Lemieux, C. (1985). Extensive sequence rearrangements in the chloroplast genomes of the green algae *Chlamydomonas eugametos* and *Chlamydomonas reinhardtii*. *Curr. Genet.* **10,** 213–219.

Lewin, R. A. (1976). Prochlorophyta as a proposed new division of algae. *Nature* **261,** 697–698.

Lewin, R. A. (1977). *Prochloron*, type genus of the Prochlorophyta. *Phycologia* **16,** 217.

Lewin, R. A. (1981). *Prochloron* and the theory of symbiogenesis. *Ann. N.Y. Acad. Sci.* **361,** 325–329.

Lewin, R. A., and Withers, N. W. (1975). Extraordinary pigment composition of a prokaryotic alga. *Nature* **256,** 735–737.

Li, N., and Cattolico, R. A. (1987). Chloroplast genome characterization in the red alga *Griffithsia pacifica*. *Mol. Gen. Genet.* **209,** 343–351.

Lidholm, J., Szmidt, A. E. Hällgren, J.-E., and Gustafsson, P. (1988). The chloroplast genomes of conifers lack one of the rRNA-encoding inverted repeats. *Mol. Gen. Genet.* **212,** 6–10.

Link, G. (1984). DNA sequence requirements for the accurate transcription of a protein-coding plastid gene in a plastid *in vitro* system from mustard (*Sinapis alba* L.). *EMBO J.* **3,** 1697–1704.

Ludwig, M., and Gibbs, S. P. (1985). DNA is present in the nucleomorph of cryptomonads: Further evidence that the chloroplast evolved from a eukaryotic endosymbiont. *Protoplasma* **127,** 9–20.

Ludwig, M., and Gibbs, S. P. (1987). Are the nucleomorphs of cryptomonads and *Chlorarachnion* the vestigal nuclei of eukaryotic endosymbionts? *Ann. N.Y. Acad. Sci.* **503,** 198–211.

MacKay, R. M., Salgado, D., Bonen, L., Stackebrandt, E., and Doolittle, W. F. (1982). The 5S ribosomal RNAs of *Paracoccus denitrificans* and *Prochloron*. *Nucleic Acids Res.* **10,** 2963–2970.

Manhart, J. R., and Palmer, J. D. (1990). The gain of two chloroplast tRNA introns marks the green algal ancestors of land plants. *Nature (London)* **345,** 268–270.

Manhart, J. R., Kelly, K., Dudock, B. S., and Palmer, J. D. (1989). Unusual characteristics of *Codium fragile* chloroplast DNA revealed by physical and gene mapping. *Mol. Gen. Genet.* **216,** 417–421.

Margulis, L. (1970). "Origin of Eukaryotic Cells." Yale University Press, New Haven, Connecticut.

Margulis, L. (1975). Symbiotic theory of the origin of eukaryotic organelles: Criteria for proof. *Symp. Soc. Exp. Biol.* **29,** 21–38.

Margulis, L. (1981). "Symbiosis in Cell Evolution." W. H. Freeman and Co., San Francisco.

Margulis, L., and Obar, R. (1985). *Heliobacterium* and the origin of chrysoplasts. *BioSystems* **17,** 317–325.

Margulis, L., and Schwartz, K. V. (1982). "Five Kingdoms." W. H. Freeman and Co., New York.

Markowicz, Y., Loiseaux-de Göer, S., and Mache, R. (1988a). Presence of a 16S rRNA pseudogene in the bi-molecular plastid genome of the primitive brown alga *Pylaiella littoralis*. Evolutionary implications. *Curr. Genet.* **14,** 599–608.

Markowicz, Y., Mache, R., and Loiseaux-De Göer, S. (1988b). Sequence of the plastid

rDNA spacer region of the brown alga *Pylaiella littoralis* (L.) Kjellm. Evolutionary significance. *Plant Mol. Biol.* **10,** 465–469.

Martin, W., and Cerff, R. (1986). Prokaryotic features of a nucleus-encoded enzyme. cDNA sequences for chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenases from mustard (*Sinapis alba*). *Eur. J. Biochem.* **159,** 323–331.

Mattox, K. R., and Stewart, K. D. (1984). *Systematics of the Green Algae.* Academic Press, London.

McFadden, G. I. (1990). Evidence that cryptomonad chloroplasts evolved from photosynthetic eukaryotic endosymbionts. *J. Cell Science* **95,** 303–308.

Mereschkowsky, C. (1905). Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol. Centr.* **25,** 593–604.

Mereschkowsky, C. (1910). Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Entstehung der Organismen. *Biol. Centr.* **30,** 278–303, 321–347, 353–367.

Mereschkowsky, C. (1920). La plante considérée comme une complex symbiotique. *Bull. Soc. Nat. Sci. Ouest France* **6,** 17–21.

Michalowski, C., Breunig, K. D., and Bohnert, H. J. (1987). Points of rearrangements between plastid chromosomes: Location of protein coding regions on broad bean chloroplast DNA. *Curr. Genet.* **11,** 265–274.

Miller, K. R., and Jacob, J. S. (1989). On *Prochlorothrix. Nature* **338,** 303–304.

Mirabdullaev, I. M. (1985). The evolution of plastids and origin of cyanobacteria. *Zhurn. Obshch. Biol.* **46,** 483–490.

Morden, C. W., and Golden, S. S. (1989). *psbA* genes indicate common ancestry of prochlorophytes and chloroplasts. *Nature* **337,** 382–385.

Nierzwicki-Bauer, S. A., Curtis, S. E., and Haselkorn, R. (1984). Cotranscription of genes encoding the small and large subunits of ribulose-1,5-bisphosphate carboxylase in the cyanobacterium *Anabaena* 7120. *Proc. Natl. Acad. Sci. USA* **81,** 5961–5965.

Obar, R., and Green, J. (1985). Molecular archaeology of the mitochondrial genome. *J. Mol. Evol.* **22,** 243–251.

Ogihara, Y., Terachi, T., and Sasakuma, T. (1988). Intramolecular recombination of chloroplast genome mediated by short direct-repeat sequences in wheat species. *Proc. Natl. Acad. Sci. USA* **85,** 8573–8577.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S.-i., Inokuchi, H., and Ozeki, H. (1986). Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* **322,** 572–574.

Padmanabhan, U., and Green, B. R. (1978). The kinetic complexity of *Acetabularia* chloroplast DNAs *Biochim. Biophys. Acta* **521,** 67–73.

Palmer, J. D. (1985a). Comparative organization of chloroplast genomes. *Ann. Rev. Genet.* **19,** 325–354.

Palmer, J. D. (1985b). Evolution of chloroplast and mitochondrial DNA in plants and algae. *In* "Molecular Evolutionary Genetics" (R. J. MacIntyre, ed.), pp. 131–240. Plenum Press, New York and London.

Palmer, J. D. (1987). Chloroplast DNA evolution and biosystematic uses of chloroplast DNA variation. *Amer. Natur.* (Suppl.) **130,** S6–S29.

Palmer, J. D., and Thompson, W. F. (1982). Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. *Cell* **29,** 537–550.

Palmer, J. D., Nugent, J. M., and Herbon, L. A. (1987). Unusual structure of geranium chloroplast DNA: A triple-sized inverted repeat, extensive gene duplications, multiple inversions, and two repeat families. *Proc. Natl. Acad. Sci. USA* **84,** 769–773.

Perasso, R., Baroin, A., Qu, L. H., Bachellerie, J. P., and Adoutte, A. (1989). Origin of the algae. *Nature* **339,** 142–144.

Raff, R. A., and Mahler, H. R. (1972). The non symbiotic origin of mitochondria. *Science* **177,** 575–582.

Raff, R. A., and Mahler, H. R. (1975). The symbiont that never was: An inquiry into the evolutionary origin of the mitochondrion. *Symp. Soc. Exp. Biol.* **29,** 41–92.

Ragan, M. A., and Chapman, D. J. (1978). "A Biochemical Phylogeny of the Protists." Academic Press, New York.

Raven, P. H., (1970). A multiple origin for plastids and mitochondria. *Science* **169,** 641–646.

Reith, M., and Cattolico, R. A. (1986). Inverted repeat of *Olisthodiscus luteus* chloroplast DNA contains genes for both subunits of ribulose-1,5-bisphosphate carboxylase and the 32,000-dalton Q_B protein: Phylogenetic implications. *Proc. Natl. Acad. Sci. USA* **83,** 8599–8603.

Ris, H. (1961). Ultrastructure and molecular organization of genetic systems. *Can. J. Genet. Cytol.* **3,** 95–120.

Ris, H., and Plaut, W. (1962). The ultrastructure of DNA-containing areas in the chloroplast of *Chlamydomonas*. *J. Cell. Biol.* **13,** 383–391.

Sager, R., and Ishida, M. R. (1963). Chloroplast DNA in *Chlamydomonas*. *Proc. Natl. Acad. Sci. USA* **50,** 725–730.

Schimper, A. F. W. (1883). Über die Entwicklung der Chlorophyll Körner und Farb-Körner. *Bot. Zeit.* **41,** 105–114.

Schmidt, R. J., Myers, A. M., Gillham, N. W., and Boynton, J. E. (1984). Immunological similarities between specific chloroplast ribosomal proteins from *Chlamydomonas reinhardtii* and ribosomal proteins from *Escherichia coli*. *Mol. Biol. Evol.* **1,** 317–334.

Schnepf, E. (1964). Zur Feinstruktur von *Geosiphon pyriforme*. Ein versuch zur Deutung cytoplasmicher membranen und Kompartimente. *Archiv Mikrobiol.* **49,** 112–131.

Schwartz, R. M., and Dayhoff, M. O. (1978). Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. *Science* **199,** 395–403.

Schwartz, R. M., and Dayhoff, M. O. (1981). Chloroplast origins: Inferences from protein and nucleic acid sequences. *Ann. N.Y. Acad. Sci.* **361,** 260–272.

Seewaldt, E., and Stackebrandt, E. (1982). Partial sequence of 16S ribosomal RNA and the phylogeny of *Prochloron*. *Nature* **295** 618–620.

Shih, M.-C., Lazar, G., and Goodman, H. M. (1986). Evidence in favor of the symbiotic origin of chloroplasts: Primary structure and evolution of tobacco glyceraldehyde-3-phosphate dehydrogenases. *Cell* **47,** 73–80.

Shinozaki, K., and Sugiura, M. (1983). The gene for the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase is located close to the gene for the large subunit in the cyanobacterium *Anacystis nidulans* 6301. *Nucleic Acids Res.* **11,** 6957–6964.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matasubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., and Sugiura, M. (1986). The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. *EMBO J.* **5,** 2043–2049.

Starnes, S. M., Lambert, D. H., Maxwell, E. S., Stevens, S. E., Jr., Porter, R. D., and Shively, J. M. (1985). Cotranscription of the large and small subunit genes of ribulose-1,5-bisphosphate carboxylase/oxygenase in *Cyanophora paradoxa*. *FEMS Lett.* **28,** 165–169.

Steinmetz, A. A., Krebbers, E. T., Schwarz, Z., Gubbins, E. J., and Bogorad, L. (1983). Nucleotide sequences of five maize chloroplast transfer RNA genes and their flanking regions. *J. Biol. Chem.* **258,** 5503–5511.

Strauss, S. H., Palmer, J. D., Howe, G. T., and Doerksen, A. H. (1988). Chloroplast genomes of two conifers lack a large inverted repeat and are extensively rearranged. *Proc. Natl. Acad. Sci. USA* **85,** 3898–3902.

Taylor, F. J. R. (1974). Implications and extensions of the serial endosymbiosis theory of the origin of eukaryotes. *Taxon* **23,** 229–258.

Taylor, F. J. R. (1979). Symbionticism revisited: A discussion of the evolutionary impact of intracellular symbioses. *Proc. R. Soc. Lond. B* 267–286.

Taylor, F. J. R. (1980). The stimulation of cell research by endosymbiotic hypotheses for the origin of eukaryotes. *In* "Endocytobiology" (W. Schwemmler and H. E. A. Schenk, eds.), pp. 917–942, Walter de Gruyter and Co., Berlin, F.R.G.

Taylor, F. J. R. (1987). An overview of the status of evolutionary cell symbiosis theories. *Ann. N.Y. Acad. Sci.* **503,** 1–16.

Tomas, R. N., and Cox, E. R. (1973). Observations on the symbiosis of *Peridinium balticum* and its intracellular alga I. Ultrastructure. *J. Phycol.* **9,** 304–323.

Tomioka, N., and Sugiura, M. (1983). The complete nucleotide sequence of a 16S ribosomal RNA gene from a blue-green alga, *Anacystis nidulans*. *Mol. Gen. Genet.* **191,** 46–50.

Turner, S., Burger-Wiersma, T., Giovannoni, S. J., Muir, L. R., and Pace, N. R. (1989). The relationship of a prochlorophyte *Prochlorothrix hollandica* to green chloroplasts. *Nature* **337,** 380–382.

Uzzell, T., and Spolsky, C. (1974). Mitochondria and plastids as endosymbionts: A revival of special creation? *Am. Sci.* **62,** 334–343.

Uzzell, T., and Spolsky, C. (1981). Two data sets: Alternative explanations and interpretations. *Ann. N.Y. Acad. Sci.* **361,** 481–499.

Valentin, K., and Zetsche, K. (1989). The genes of both subunits of ribulose-1,5-bisphosphate carboxylase constitute an operon on the plastome of a red alga. *Curr. Genet.* **16,** 203–209.

Valentin, K., and Zetsche, K. (1990) Structure of the Rubisco operon from the unicellular red alga *Cyanidium caldarium:* Evidence for a polyphyletic origin of the plastids. *Mol. Gen. Genet.* **222,** 425–430.

Van Valen, L. M. (1982). Phylogenies in molecular evolution: *Prochloron*. *Nature* **298,** 493–494.

Wallace, D. C. (1982). Structure and evolution of organelle genomes. *Microbiol. Rev.* **46,** 208–240.

Weeden, N. F. (1981). Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. *J. Mol. Evol.* **17,** 133–139.

Weil, J. H. (1987). Organization and expression of the chloroplast genome. *Plant Science* **49,** 149–157.

Whatley, J. M. (1981). Chloroplast evolution—Ancient and modern. *Ann. N.Y. Acad. Sci.* **361,** 154–165.

Whatley, J. M. (1983). Plastids—Past, present, and future. *Int. Rev. Cytol.* (Suppl.) **14,** 329–373.

Whatley, J. M., and Whatley, F. R. (1981). Chloroplast evolution. *New Phytol.* **87,** 233–247.

Whatley, J. M., John, P., and Whatley, F. R. (1979). From extracellular to intracellular: The establishment of mitochondria and chloroplasts. *Proc. R. Soc. Lond. B* **204,** 165–187.

Whitfeld, P. R., and Bottomley, W. (1983). Organization and structure of chloroplast genes. *Annu. Rev. Plant Physiol.* **34,** 279–310.

Wilcox, L. W., and Wedemayer, G. J. (1984). *Gymnodinium acidotum* Nygaard (Pyrrophta), a dinoflagellate with an endosymbiotic cryptomonad. *J. Phycol.* **20,** 236–242.

Wilcox, L. W., and Wedemayer, G. J. (1985). Dinoflagellate with blue-green chloroplasts derived from an endosymbiotic eukaryote. *Science* **227,** 192–194.

Williamson, S. E., and Doolittle, W. F. (1983). Genes for $tRNA^{Ile}$ and $tRNA^{Ala}$ in the spacer

between the 16S and 23S rRNA genes of a blue-green alga: Strong homology to chloroplast tRNA genes and tRNA of the *E. coli rrnD* gene cluster. *Nucleic Acids Res.* **11,** 225–235.

Witt, D., and Stackebrandt, E. (1988). Disproving the hypothesis of a common ancestry for the *Ochromonas danica* chrysoplast and *Heliobacterium chlorum*. *Arch. Microbiol.* **150,** 244–248.

Woese, C. R. (1987). Bacterial evolution. *Microbiol. Rev.* **51,** 221–271.

Yamada, T., and Shimaji, M. (1987). Splitting of the ribosomal RNA operons on chloroplast DNA from *Chlorella ellipsoidea*. *Mol. Gen. Genet.* **208,** 377–383.

Zablen, L. B., Kissil, M. S., Woese, C. R., and Buetow, D. E. (1975). Phylogenetic origin of the chloroplast and prokaryotic nature of its ribosomal RNA. *Proc. Natl. Acad. Sci. USA* **72,** 2418–2422.

Index