Topics in Current Genetics

R. Bock
(Ed.)

19

Cell and Molecular Biology of Plastids

Springer

Topics in Current Genetics

19

Series Editor: *Stefan Hohmann*

Ralph Bock (Ed.)

Cell and Molecular Biology of Plastids

With 51 Figures, 11 in Color; and 23 Tables

Springer

Dr. Ralph Bock
Max-Planck-Institut für Molekulare Pflanzenphysiologie
Am Mühlenberg 1
D-14476 Potsdam-Golm
Germany
e-mail: rbock@mpimp-golm.mpg.de

The cover illustration depicts pseudohyphal filaments of the ascomycete *Saccharomyces cerevisiae* that enable this organism to forage for nutrients. Pseudohyphal filaments were induced here in a wild-type haploid MATa Σ1278b strain by an unknown readily diffusible factor provided by growth in confrontation with an isogenic petite yeast strain in a sealed petri dish for two weeks and photographed at 100X magnification (provided by Xuewen Pan and Joseph Heitman).

ISBN 978-3-540-75375-9 ISBN 978-3-540-75376-6 (eBook)

DOI 10.1007/978-3-540-75376-6

Topics in Current Genetics ISSN 1610-2096

Library of Congress Control Number: 2007936025

Typesetting: Camera ready by editor
Production: LE-TEX Jelonek, Schmidt & Vöckler GbR, Leipzig, Germany
Cover: WMXDesign GmbH, Heidelberg, Germany

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Topics in Current Genetics publishes review articles of wide interest in volumes that center around specific topics in genetics, genomics as well as cell, molecular and developmental biology. Particular emphasis is placed on the comparison of several model organisms. Volume editors are invited by the series editor for specific topics, but further suggestions for volume topics are highly welcomed.

Each volume is edited by one or several acknowledged leaders in the field, who invite authors and ensure the highest standard of content and presentation. Only solicited manuscript will be considered. All contributions are peer-reviewed. All papers are published online prior to the print version. Individual DOIs (digital object identifiers) make each article fully citable from the moment of online publication.

All volumes of Topics in Current Genetics are part of the Springer eBook Collection. The collection includes online access to more than 3,000 newly released books, book series volumes and reference works each year. In addition to the traditional print version, this new, state-of-the-art format of book publications gives every book a global readership and a better visibility.

Editorial office:
Topics in Current Genetics
Series Editor: Stefan Hohmann
Cell and Molecular Biology
Göteborg University
Box 462
40530 Göteborg, Sweden
FAX: +46 31 7862599
E-mail: editor@topics-current-genetics.se

Preface

Ralph Bock

Standard textbooks of genetics and molecular biology pay scant attention to plastids, although the chloroplast is arguably the best-studied genetic compartment in eukaryotic cells. The past two decades have witnessed an enormous progress in our understanding of plastid biogenesis, genome structure and function, gene expression and its regulation as well as plastid-nuclear interaction and communication pathways. In addition, research on plastids has benefited enormously from the development and continuous refinement of transgenic technologies. The possibility to directly alter the genetic information of the plastid has facilitated the study of virtually all aspects of plastid biology *in vivo* and, moreover, has paved the way to diverse applications of transgenic plastids in biotechnology.

It was with this in mind that we approached the writing of the present volume of *Topics in Current Genetics* entitled *Cell and Molecular Biology of Plastids*. The book begins with a chapter on plastid biogenesis, differentiation and division written by Kevin Pyke. The following chapter (contributed by Ralph Bock) covers plastid genome structure and function as well as the inheritance of plastids and their genetic material. Anil Day and Panagiotis Madesis portray the processes and mechanisms involved in both maintenance and structural dynamics of plastid genomes: recombination, DNA replication, and repair. The following four chapters cover the various steps of gene expression in plastids, their molecular components, and how they are regulated: transcription (by Karsten Liere and Thomas Börner), RNA stability and degradation (by David Stern and colleagues), the diverse RNA processing mechanisms operating in plastids, including intron splicing and RNA editing (by Christian Schmitz-Linneweber and Alice Barkan), and protein biosynthesis (by Hadas Peled-Zehavi and Avihai Danon). Three chapters are dedicated to key posttranslational processes in plastid biogenesis and function: protein processing and the assembly of multiprotein complexes (by Eva-Mari Aro and colleagues), protein stability and degradation (by Zach Adam), and protein import and sorting (by Birgit Agne and Felix Kessler). Many of these processes are described using chloroplasts and the photosynthetic apparatus as model system, not least because research on non-green plastid types is still far less advanced.

The chapter written by Bianca Naumann and Michael Hippler provides on overview of plastid proteomics research. It covers both methodological and functional aspects and demonstrates how a highly complex proteome can be dissected by splitting it up into analyzable subproteomes. The multifarious communication pathways between plastids and the nucleo-cytosolic compartment of the plant cell are dealt with in the contribution by Thomas Pfannschmidt and colleagues. Our current knowledge about anterograde (nucleus-to-plastid) and retrograde (plastid-to-nucleus) signalling processes is summarized illustrating the great complexity of the regulatory mechanisms that have evolved to coordinate the activities of the prokaryotic-type genome in the plastid and the eukaryotic-type genome in the nucleus of the plant cell. Last but not least, the chapter by Hans-Ulrich Koop and colleagues describes the state of the art in engineering plastid genomes of algae and higher plants and highlights selected applications of plastid transformation technology in basic research and plant biotechnology.

Cell and Molecular Biology of Plastids is written primarily for those working directly in the fields of plastid biology, organelle genetics and gene expression, photosynthesis research and biotechnology. The authors of the individual chapters have tried to discuss concepts and emphasize general principles that are accepted and proven. Inevitably, there is some overlap between the contributions, which, however, has been limited to the extent needed to ensure that the individual chapters can be read in isolation. Authors and editor hope that this volume will serve as a stepping-stone for graduate students becoming interested in organelle biology and new researchers entering the field.

In closing, I express my sincere thanks to the authors of each chapter – their thoroughness and commitment made this volume possible. I am also very grateful to the many colleagues who willingly acted as reviewers and to Springer Publishers and the editorial office of *Topics in Current Genetics* for their help in editing and formatting this book.

Bock, Ralph
Max Planck Institute for Molecular Plant Physiology, Am Muehlenberg 1, D-14476 Potsdam-Golm, Germany
rbock@mpimp-golm.mpg.de

Table of contents

List of contributors

Adam, Zach
The Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture, The Hebrew University, Rehovot 76100, Israel.
zach@agri.huji.ac.il

Agne, Birgit
Laboratoire de Physiologie Végétale, Institut de Biologie, Université de Neuchâtel, Rue Emile-Argand 11, 2009 Neuchâtel, Switzerland

Aro, Eva-Mari
Department of Biology, University of Turku, FIN-20014 Turku, Finland
evaaro@utu.fi

Barkan, Alice
Institute of Molecular Biology, University of Oregon, Eugene, OR 97403, USA

Bock, Ralph
Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany
rbock@mpimp-golm.mpg.de

Bollenbach, Thomas J.
Boyce Thompson Institute for Plant Research, Tower Rd. Ithaca NY 14853, USA

Börner, Thomas
Institut für Biologie / Genetik, Humboldt-Universität zu Berlin, Chausseestr. 117, 10115 Berlin, Germany
thomas.boerner@rz.hu-berlin.de

Bräutigam, Katharina
Institute for General Botany and Plant Physiology, Junior Research Group, Friedrich-Schiller-University Jena, Dornburger Str. 159, 07743 Jena, Germany

Danon, Avihai
Department of Plant Sciences, Weizmann Institute of Science, Rehovot 76100, Israel
avihai.danon@weizmann.ac.il

Day, Anil
Faculty of Life Sciences, The University of Manchester, Oxford Road, Manchester M13 9PT, UK
anil.day@manchester.ac.uk

Dietzel, Lars
Institute for General Botany and Plant Physiology, Junior Research Group, Friedrich-Schiller-University Jena, Dornburger Str. 159, 07743 Jena, Germany

Golds, Timothy J
Research Centre Freising, Icon Genetics AG, Lise-Meitner-Straße 30, D 85354 Freising, Germany

Herz, Stefan
Research Centre Freising, Icon Genetics AG, Lise-Meitner-Straße 30, D 85354 Freising, Germany

Hippler, Michael
Institute of Plant Biochemistry and Biotechnology, University of Muenster, Hindenburgplatz 55, 48143 Muenster, Germany
mhippler@uni-muenster.de

Kanervo, Eira
Department of Biology, University of Turku, FIN-20014 Turku, Finland

Kessler, Felix
Laboratoire de Physiologie Végétale, Institut de Biologie, Université de Neuchâtel, Rue Emile-Argand 11, 2009 Neuchâtel, Switzerland
felix.kessler@unine.ch

Koop, Hans-Ulrich
Faculty of Biology, Department I, Botany, Ludwig-Maximilians-Universität München, Menzinger Straße 67, D 80638 München, Germany
koop@lmu.de

Liere, Karsten
Institut für Biologie / Genetik, Humboldt-Universität zu Berlin, Chausseestr. 117, 10115 Berlin, Germany

Madesis, Panagiotis
Faculty of Life Sciences, The University of Manchester, Oxford Road, Manchester M13 9PT, UK

Naumann, Bianca
Institute of Plant Biochemistry and Biotechnology, University of Muenster, Hindenburgplatz 55, 48143 Muenster, Germany

Peled-Zehavi, Hadas
Department of Plant Sciences, Weizmann Institute of Science, Rehovot 76100, Israel

Pfannschmidt, Thomas
Institute for General Botany and Plant Physiology, Junior Research Group, Friedrich-Schiller-University Jena, Dornburger Str. 159, 07743 Jena, Germany
Thomas.Pfannschmidt@uni-jena.de

Portnoy, Victoria
Department of Biology, Technion-Israel Institute of Technology, Haifa 32000, Israel

Pyke, Kevin
Plant Sciences Division, School of Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough, Leicestershire LE12 7RD
Kevin.Pyke@nottingham.ac.uk

Schmitz-Linneweber, Christian
Institute of Biology, Humboldt-University Berlin, Chausseestr. 117, 10115 Berlin, Germany
christian.schmitz-linneweber@rz.hu-berlin.de

Schuster, Gadi
Department of Biology, Technion-Israel Institute of Technology, Haifa 32000, Israel

Stern, David B.
Boyce Thompson Institute for Plant Research, Tower Rd. Ithaca NY 14853
ds28@cornell.edu

Suorsa, Marjaana
Department of Biology, University of Turku, FIN-20014 Turku, Finland

Nickelsen, Jörg
Faculty of Biology, Department I, Botany, Ludwig-Maximilians-Universität München, Menzinger Straße 67, D 80638 München, Germany

Plastid biogenesis and differentiation

Kevin Pyke

Abstract

Plastids are crucial to plant functionality and develop from proplastids in meristem cells to generate different plastid forms in different types of plant cells. In addition to the photosynthesis of leaf mesophyll cell chloroplasts, plastids contribute to storage and pigmentation capacities in many different specialised cells as well as contributing essential metabolic pathways within the cell in general. Plastids also have the capacity to interconvert between types according to environmental and molecular signals. Progress in understanding the cell biology and morphological control of different plastid types is considered in the light of modern imaging techniques, which have revealed new aspects of plastid morphology. As well as considering molecular aspects of how plastids control their division, this article discusses also how cell-specific differentiation might be controlled and whether master control genes for plastid biogenesis might be in charge.

1 Introduction

Plastids form a distinct group of organelles in higher and lower plants and are one of the defining characteristics by which plants are different to animals. For many years, most plastid based research focused on the chloroplast and trying to understand the mechanism of photosynthesis and the biochemical interactions of the chloroplast with the cell. With the advent of molecular biology and more recently, a variety of novel imaging techniques, a better understanding of how the chloroplast and other plastid types function within the cell in a truly biological manner is starting to emerge. Even so, the chloroplast remains dominant in providing the bulk of our knowledge about plastid biology. In this article, I consider the structure and morphology of the chloroplast and a range of other plastid types as well as how plastids differentiate and undergo interconversions. Finally, I discuss two fields in plastid biology, which have progressed significantly in recent years, namely plastid division and the biology of stromules.

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0226 / Published online: 22 May 2007

2 Proplastids

All plastids within a plant are ultimately derived from those progenitor plastids, which are found in meristem cells called proplastids. These in turn have been derived from the few proplastids, which were present in the zygote and derived potentially from both the maternal egg cell and the paternal pollen grain. However, most Angiosperms have mechanisms to exclude or degrade proplastids in the pollen line and hence the plastids present in the majority of plants are inherited maternally (Mogensen 1996; Corriveau and Coleman 1998; Zhang et al. 2003). In those species in which biparental inheritance occurs, plastids within the zygote constitute a mixed population derived from both parent egg and pollen. However, many factors appear to bias the relative proportion of maternally and paternally-derived plastids and plastid populations in resulting plants can be highly variable with respect to the origins of plastids within them (Mogensen 1996).

Considering the fundamental importance of proplastids to plastid biology, the knowledge of proplastid cell biology and their fine ultrastructure is limited, mostly because of the difficulties with analysing small organelles with no pigment in small regions of dense tissue. Knowledge of the physical appearance of proplastids has been derived largely from electron micrographs (Chaley and Possingham 1981; Akita and Sagisaka 1995; Robertson et al. 1995; Gunning 2004), which show proplastids as small organelles containing limited internal structure that are dispersed throughout the cytoplasm. Most proplastids contain rudimentary pieces of thylakoid membrane, but are unpigmented although those in shoot apical meristems appear to contain more thylakoid in a more organized state than those in the root apical meristem (Gunning 2004). In addition, ingrowths from the inner plastid envelope membrane into the proplastid stroma can also be seen occasionally, as well as ribosomes. Starch grains may be present, especially in proplastids of seeds where starch was laid down in the proplastid during seed development (Gunning 2004). In wheat plumules and potato stolons, starch content of proplastids is variable with some containing significant starch grains and others with no starch. This difference in starch content appears to result from differences in the capacity for starch synthesis since immunogold labelling of the enzyme starch synthase reveals two types of proplastids: those with and those without the enzyme (Akita and Sagisaka 1995).

Estimating proplastid numbers is difficult and to date no studies have definitively counted proplastid populations in meristem cells. However, various studies of shoot meristem cells estimate that they contain 10-20 proplastids per cell (Cran and Possingham 1972; Lyndon and Robertson 1976; Pyke and Leech 1992). Using modern fluorescent protein technology, imaging of proplastids in meristems and during cytokinesis should be feasible, although proplastid dynamics during meristematic cell divisions have yet to be studied in detail. Proplastids with fluorescent marker proteins on board, such as GFP, can be imaged in root meristems (Kohler and Hanson 2000) and those in shoot apical meristems can be observed also (Trynka and Pyke, unpublished), although experiments to determine population sizes and segregation patterns in different parts of the meristem could prove technically demanding.

Differences in proplastid number according to cell position within the meristem or in organs derived from it may well exist (Lyndon and Robertson 1976), but whether such differences are significant to cellular function are unclear and they may simply reflect differences in proplastid division rate compared to local rates of cell division. Differences in proplastid DNA content and morphology have been shown to exist between cell layers within a meristem, suggesting that tissue-specific characteristics of proplastids within a meristem may be important (Fujie et al. 1994). During the cell divisions of embryogenesis and the cell divisions within meristems, proplastids must divide to ensure continuity within cell lines and to ensure that all cells within the plant contain plastids. Little is known of a distinct mechanism by which a correct proplastid segregation is achieved at cytokinesis (Sheahan et al. 2004) and it would appear that aplastidic daughter cells are prevented simply because proplastids are generally distributed in the cytoplasm, thus ensuring segregation into both daughter cells, but also because they locate more particularly in positions close to the nucleus prior to the onset of mitosis. Positioning in the peri-nuclear cytoplasm during protoplast division is driven by the actin cytoskeleton leading to entrapment of plastids close to the nucleus (Sheahan et al. 2004). Whether a similar process happens during cytokinesis in meristems is unknown. Nuclear mutations, which affect proplastid division and give rise to populations of few, enlarged proplastids in meristematic cells (Robertson et al. 1995) do not result in the appearance of aplastidic cells in meristems, which implies that giant proplastids can still maintain a mechanism by which they segregate correctly. Giant plastids in tomato fruit cells appear able to replicate by a budding/fragmentation mechanism (Forth and Pyke 2006) and therefore it is feasible that the generation of small budded proplastids could ensure correct segregation in meristematic cells containing giant proplastids.

Efforts to study the extent of proplastid metabolism and DNA transcription and translation have been limited but those which have examined proplastid transcription at the tissue level have shown such activity to be low and that the initiation of a differentiation pathway, such as chloroplast differentiation, is necessary to upregulate transcriptional activity (Harak et al. 1995; Mache et al. 1997; Sakai et al. 1998; Baumgartner et al. 1989). Indeed, expression of nuclear genes for proplastid ribosomes is required prior to the expression of those genes, which are plastid encoded. Overall proplastids remain an exasperating organelle, occupying a pivotal place in plastid cell biology but yet about which there is so much still to learn.

3 The morphology and structure of different plastid types

As cells within developing seedlings and developing plant organs differentiate, plastids embark on different patterns of differentiation according to the differentiation pathway that the cell itself takes. Proplastids have the ability to give rise to a variety of different types of plastid, which form in different types of tissue. Plastids can also interconvert between their different forms in many situations. Thus, for most plastid types, there are two different pathways by which they can arise: directly or by re-differentiation of an existing plastid type. Traditionally, characterisation and naming of different types of mature plastids has largely been based on the types of molecules they store or the types of pigments they accumulate, although this may not necessarily be the best system for plastid taxonomy since often plastids show a mixture of features from different types making precise naming difficult. Although distinct types of plastid differentiation do exist, a better system for their classification could be based on the biochemical and physiological properties or maybe the extent of their proteome or metabolome. Such a system could ease the difficulties by which plastids displaying intermediate phenotypes have to be named. In this chapter, the basic structure and morphology of the major types of differentiated plastids found in higher plants will be consider and subsequently, what is known of the differentiation pathways which give rise to each of the types will be discussed.

3.1 Chloroplast structure and morphology

Chloroplasts are the most prominent form of plastid occurring in all green plant tissues and enable photosynthetic carbon fixation to occur in addition to a variety of other biochemical processes central to cellular metabolism. Like all plastids, they are bounded by a double plastid envelope membrane, which acts as a major control point for chloroplast import and export as well as being a major site for biochemical synthesis (Joyard et al. 1998). Chloroplasts in leaf mesophyll cells are typically ellipsoidal in shape but with defined poles, a feature that is crucial to their division. However, chloroplasts can also be highly pleiomorphic and can take up irregular morphologies in different cell types. Indeed, the potential plasticity in plastid shape has become clear in recent years with the analysis of giant plastids, which occur when plastid division is perturbed. In these giant plastids, which are up to 50-fold larger than normal chloroplasts, the morphology is highly irregular (Pyke et al. 1994) yet apparently stable when osmotically challenged (Pyke 2006) suggesting that a mechanism exists which controls and exerts stability on plastid morphology. A suggestion that an FtsZ-based internal plastoskeleton might function in controlling plastid morphology (Reski 2002) needs further experimentation since most of the FtsZ molecules within the plastid appear to be involved in division rather than morphological control. The recent discovery of mechanosensory proteins within the plastid envelope (Haswell and Meyerowitz 2006) showed that perturbation of such proteins by mutation affects plastid morphology, implying that tension monitoring in the plastid envelope somehow plays a role in morphological control.

A major structural component, which typifies the chloroplast, is the extensive thylakoid membrane system, which extends throughout the body of the chloroplast and dominates its internal architecture. Thylakoid membranes are the site of photosynthetic electron transport and ATP synthesis and delimit a distinct compartment within the chloroplasts: the thylakoid lumen. Thylakoids are composed of lamellae, which are arranged into a highly complex system of stacked lamellae called grana interconnected by single lamellae called stromal lamellae.

Models for thylakoid membrane structure have been developed largely from analysis of electron micrographs of sectioned chloroplasts, a system that is fraught with difficulty in interpretation in generating three-dimensional models from two-dimensional images. Three different models have been proposed (Arvidsson and Sundby 1999; Mustardy and Garab 2003; Shimoni et al. 2005) but differ in their conclusions, although all show the highly complex nature of thylakoid membrane arrangement within the grana. The model of Mustardy and Garab (2003) shows the grana as fused stacks of membrane which look like fan blades, with stromal lamellae joining stacks together at alternating levels within the stack, and the whole structure forming a right handed helix. The reason for this complex thylakoid membrane morphology is to provide a large surface within the plastid on which light capture by chlorophyll and electron transport can occur. Consequently, the area of thylakoid membrane within a mature plastid is large and much greater than simple invaginations from the plastid envelope membrane.

Surprisingly, the mechanisms by which the construction of the thylakoid membrane system is initiated, synthesised in large amounts and then built into a complex three-dimensional architecture is poorly understood. Electron micrographs showing invaginations of the inner plastid envelope into the stroma gave credence to the hypothesis that thylakoid membrane is derived, at least initially, from such invaginations as proplastids differentiate into chloroplasts. Proplastids usually contain small amounts of thylakoid membrane and the extensive biogenesis of more thylakoid membrane may simply involve building off of extant membrane. However, recent studies have clearly shown that both chloroplasts and proplastids contain vesicles within the stroma (Westphal et al. 2003; Gunning 2004) and that a vesicle trafficking system occurs in plastids primarily between the plastid envelope and the stroma (Westphal et al. 2003). Vesicles are budded from the inner plastid envelope and accumulate close to the inner membrane, particularly when fusion processing at the thylakoid membrane is curtailed by low temperature (Morre et al. 1991). The main purpose of plastid vesicle transport is probably that of providing galactolipids, which are synthesised in the plastid envelope membranes (Joyard et al. 1998), for continued synthesis of thylakoid membrane, although they could also deliver hydrophobic proteins, which reside in the thylakoid membrane. Plastid vesicle trafficking appears to utilize several homologous components of the cytosolic ER Golgi trafficking system, encoded by nuclear genes, in that the chloroplast contains both ARF1 and Sar1 GTPases (Andersson and Sandelius 2004), which are involved in vesicle assembly. In addition, the chloroplast also contains dynamin (Park et al. 1998) and proteins required for vesicle fusion (Hugeney et al. 1995). Two other nuclear-encoded proteins involved in the vesicle directed thylakoid biogenesis are VIPP1 (Kroll et al. 2001) and Thf1 (Wang et al. 2004). Mutations in either gene result in abolition of vesicles and perturbed synthesis of the thylakoid membrane. VIPP1 forms a high molecular weight complex on the inner envelope membrane, which could conceivably be involved in vesicle production (Aseeva et al. 2004). An intriguing problem for the future will be to understand how vesicle directed thylakoid synthesis is controlled to facilitate the construction of thylakoid architecture and biogenesis of the correct three-dimensional arrangement of the thylakoid membrane network. FZL is a dynamin-related membrane remodelling protein and is located inside the chloroplast in punctate foci on the plastid envelope and on the thylakoid membrane (Gao et al. 2006). Perturbation of this protein results in altered thylakoid morphology and changes in patterns of granal stacking suggesting that it plays an important role in thylakoid organisation and possibly in the dynamic continuum of membrane synthesis between the plastid envelope and the thylakoid.

During chloroplast development there is a significant increase in size of the plastid organelle from proplastid to mature chloroplast. There is also significant variation in mature chloroplast size in different cell types and also within the population of chloroplasts within individual leaf mesophyll cells. An important question yet to be addressed is what mechanisms control chloroplast size? Within a population of chloroplasts in a leaf mesophyll cell there is a trade-off between plastid density and size such that permutations of more small ones or fewer larger ones can be observed in cells of differing sizes and in different species where average leaf mesophyll cell size varies (Ellis and Leech 1983; Pyke 1999). However, the expansion process by which chloroplasts increase the surface area of their envelope membrane and the extent of the stroma and the thylakoid membrane must have a control system which shuts down further expansion at maturity. Conceivably a mechanosensing mechanism (Haswell and Meyerowitz 2006) could achieve this so that as chloroplasts become more densely packed and start squashing each other, as happens in leaf mesophyll cells, mechanosensing feedback shuts down further plastid replication and plastid expansion.

3.2 Amyloplast structure and morphology

All chloroplasts seem to have the ability to accumulate starch grains within the stroma as a transient store of photosynthetic assimilate. Normally these starch grains are degraded through the dark part of the photoperiod and the products exported. Amyloplasts, however, are a plastid type in which starch accumulation is long term and are mostly found in storage tissues such as tubers and seed endosperm where they are highly abundant. All plant starch is synthesised in the plastid and produced either directly from photosynthate, as occurs in leaf plastids or indirectly from photosynthate transported to heterotrophic tissues within the plant. The latter process occurs in amyloplasts, which are the dominant organelle in storage tissues and are of great agricultural and economic significance since 75% of the energy contained in the average human diet is derived from starch (Duffus 1984). Starch is an insoluble, complex, semi-crystalline polymer of glucose synthesised in amyloplasts by the polymerisation of ADP glucose, producing highly branched amylopectin and relatively unbranched amylose in proportions of 70:30% (Smith et al. 1997). Starch is present within amyloplasts as grains, which have a distinct structure consisting of a series of concentric rings alternating between semi-crystalline and amorphous zones. These zones are a result of differences in organisation of the amylopectin chains (Smith et al. 1997). Two different size classes of starch grains are present in endosperm amyloplasts; the A-type, of up to 45 μm in diameter, and the smaller B-type of up to 10 μm in diameter. The ratio of these two types is variable and under environmental control but has a major effect on the processing qualities of the resultant starch in the food industry (Langeveld et al. 2000).

Amyloplasts most often form from proplastids during the early development of storage organs such as tubers or seed endosperm. In red winter wheat, proplastids are present within the coencytic endosperm but when cellularisation is complete, starch deposition commences and amyloplasts are formed (Bechtel and Wilson 2003). Some understanding of the way in which amyloplast differentiation is controlled has come mostly from exploiting cell cultures, in which amyloplast differentiation can be induced by adding phytohormones. Tobacco BY-2 cells grown in the dark are undifferentiated and contain proplastids. The presence of auxin causes these cells to proliferate. When auxin is replaced with benzyladenine, rates of cell division decline and proplastids accumulate starch and form amyloplasts (Sakai et al. 1992, 1999).

Amyloplasts are also present in a specialised cell type in the root tip; the collumella cells. Collumella cells form the gravisensing system in roots enabling gravitropic responses in growth in relation to the gravity vector. The sinking of amyloplasts, called statoliths in these cells, in the cytoplasm under the influence of gravity is thought to initiate a signal transduction pathway involving auxin redistribution (Swarup et al. 2005), which results in differential cell expansion and downward growth in a positively gravitropic manner. The reverse happens in stems of shoots where statoliths are present in a sheath of cells around the vasculature and cause upward growth of shoots in negatively gravitropic manner (Yamamoto et al. 2002). How these amyloplasts form specifically in these two cell types whilst cells around them contain different plastid types is unclear but undoubtedly the control is more complex than a simple change in the type of phytohormone as suggested by differentiation in cell cultures.

Since chloroplasts in many plant species accumulate significant amounts of starch during the light period, it may be pertinent to consider how these starch-laden chloroplasts differ from amyloplasts where starch storage is more long term. Transient starch in chloroplasts has a lower amylose content, forms as flattened plate-like structures rather than the more spherical grains of amyloplasts and does not possess the growth rings of amyloplast starch (A. Smith, personal communication). In some species such a tobacco and cotton, starch breakdown in older leaves is not complete by dawn and starch accumulates in these older leaf cells, taking on some properties of long-term storage starch. Conversely some species synthesise very little starch in the chloroplast at all during the light period and export their photosynthate and synthesise sucrose in the cytosol (Zeeman et al. 2004).

3.3 Chromoplast structure and morphology

During the evolution of higher plants, a necessity arose in that plants needed to attract insects and mammals to them in order to facilitate flower pollination and to aid in the dispersal of seeds within fleshy fruits. In order for higher plants to be prominent visually within the flora, the development of brightly coloured structures occurred, primarily in the petals of flowers and in the tissues of fleshy fruits. The accumulation of pigments within the plastids in these tissues led to the formation of a distinct type of plastid, namely the chromoplast. Most of the pigments that are laid down in chromoplasts are carotenoids, which are synthesised from the C_{40} molecule phytoene, and constitute several different types, namely carotenes, lycopene, lutein, violaxanthin, and neoxanthin (Camara et al. 1995; Cunningham and Gantt 1998; Bramley 2002). These classes of molecules are not the sole preserve of chromoplasts, since several are commonly found on thylakoid membranes in the chloroplast where they function as accessory pigments in light capture and energy dissipation. In addition, the carotenoid- related pigment astaxanthin is the basis of the red-pink colouration in several animals including flamingo, lobster, and shrimp (Armstrong and Hearst 1996). Other types of soluble pigments, which are found in the cell's vacuole, also contribute to colouration of plant parts and in many cases a mixture of pigment types is present (Kay et al. 1981; Weston and Pyke 1999).

Detailed structural analysis of chromoplasts in different tissues and species shows great heterogeneity in their structure, which probably reflects differences in the profile of coloured pigments present. There have been efforts to categorise chromoplasts into distinct types according to the types of storage structures present within the chromoplasts, i.e. globular, membranous, or crystalline (Thomson and Whatley 1980; Camara et al. 1995). Although laudable, such a classification system can be difficult to apply to the vast range of chromoplast types found throughout nature in different tissues in different plant species. Knowledge of chromoplast biogenesis has been gained largely from a detailed study of agriculturally important fleshy fruits, primarily in tomato (*Lycopersicon esculentum*) (Fraser et al. 1994; Camara et al. 1995; Cunningham and Gantt 1998), and bell pepper (Hugueny et al. 1995a). During the formation of pigmented chromoplasts from green chloroplasts in unripe fleshy fruit, a controlled breakdown of chlorophyll and the thylakoid membrane occurs concurrent with a significant increase in carotenoid pigment biosynthesis. Increased expression of the ELIP gene is associated with the chloroplast to chromoplast transition and may play a role in the regulated breakdown of the extensive thylakoid membranes of the chloroplast (Bruno and Wetzel 2004). Associated with increased carotenoid biosynthesis is the upregulated expression of several nuclear genes, which are required for chromoplast differentiation (Lawrence et al. 1993, 1997; Summer and Cline 1999). Plastid DNA appears to play a minor role in chromoplast differentiation and there is increased methylation of plastid DNA in chromoplasts (Kobayashi et al. 1990). Exactly how a chromoplast differentiation pathway is initiated in green chloroplasts in ripening fruit is unclear, even though a significant amount is known about the basic biochemistry and molecular biology of fruit ripening and the role of the hormone ethylene (Alexander and Grierson 2002). It could be argued that the chromoplast is no more than a bag into which carotenoid pigment is loaded and indeed increased transcription in a variety of carotenoid biosynthetic enzymes is a key phase in chromoplast biogenesis. Increases in enzyme activity of phytoene synthase and phytoene desaturase (Fraser et al. 1994), 1-deoxy-D-xylulose 5-phosphate synthase (Lois et al. 2000) and a plastid terminal oxidase associated with phytoene desaturation (Josse et al. 2000) are observed as chloroplast differentiate into chromoplasts. There are also increases in other proteins not associated with carotenoid metabolism and which could be best viewed as chromoplast specific differentiation genes. These include enzymes in response to oxidative stress (Livne and Gepstein 1988; Romer et al. 1992), and carotenoid sequestration proteins including fibrillin (Vishnevetsky et al. 1999).

Surprisingly the cell biology of chromoplast differentiation has been poorly described and until recently was dependent upon electron microscopy descriptions of their structure (Harris and Spurr 1969a, 1969b; Thomson and Whatley 1980; Bathgate et al. 1985). Internally, chromoplast structure appears dependant on which type of carotenoids are sequestered within them since the internal architecture is highly variable and can consist of either plastoglobules of pigment, crystalline structures of carotenoids, microfibrillar structures with sequestered carotenoids, extensive internal membranous structures or a mixture of these. Gunning (2004) shows particularly beautiful colour images of red and yellow chromoplasts in a variety of petals and fruits.

In recent years, the exploitation of green fluorescent proteins targeted to the plastid compartment has enabled chromoplasts to be observed within the whole cell and some aspects of their cell biology have been revealed, primarily in the ripening fruits of tomato (Fig. 1). In the light microscope, these red tomato chromoplasts appear as small heterogeneously shaped organelles with little clear structure. The production of thin membranous tubules from the chromoplasts called stromules, a feature of plastids in general, has been well studied and will be discussed in Section 6. Occasionally, the membrane of tomato chromoplasts is distorted by long thin crystals of lycopene (Pyke and Howells 2002). Mature pericarp cells in the fleshy part of the ripe tomato fruit are large and may contain up to 2000 red pigmented chromoplasts, which are generated from populations of dividing chloroplasts, that accumulate during the green phase of fruit development. During the differentiation of chromoplasts from chloroplasts, a heterogeneous array of small bodies within the cell can be observed, some of which appear to be broken pieces of stromule or even vesicles which appear to bud from the chloroplast body and are revealed by the GFP they carry (Waters et al. 2004; Forth and Pyke 2006). Thus, two different processes could give rise to large populations of differentiated chromoplasts within the cell and gives support to the idea that chromoplasts are little more than storage sacs with high levels of carotenoid biosynthetic enzyme activity.

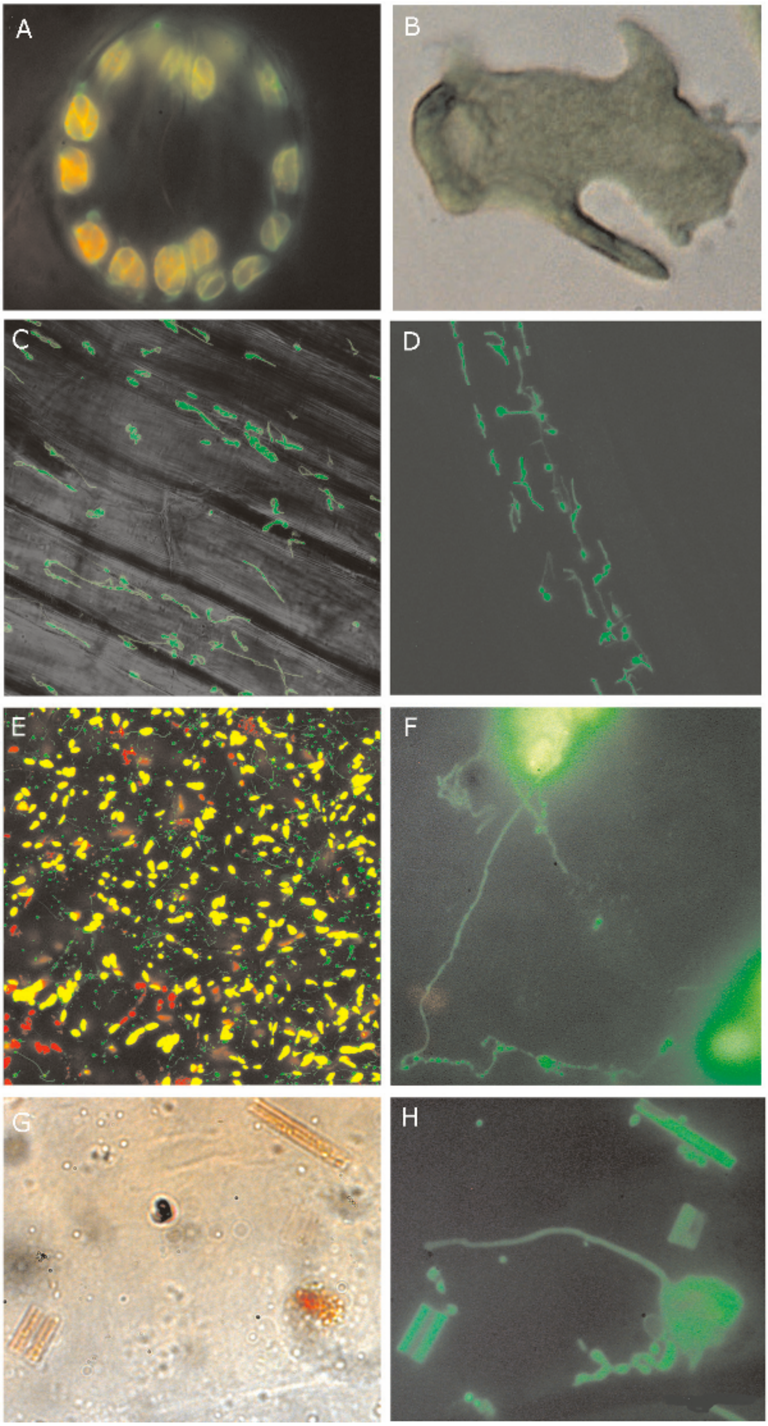

A
B
C
D
E
F
G
H

Fig. 1 (overleaf). Variation in plastid morphology. (A) Chloroplasts in a leaf stomatal guard cells containing GFP, which fluoresces green on a background of red chlorophyll fluorescence. These chloroplasts have a conventional chloroplast morphology and have only small stromule protrusions. (B) Isolated giant mesophyll cell chloroplasts from *Arabidopsis* leaves expressing antisense copies of the FtsZ plastid division protein. These chloroplasts are highly variable in shape but maintain their complex morphologies when isolated from the cell. (C) Plastids in the hypocotyl cells of a tobacco seedling illuminated by targeting green fluorescent protein to the plastids. These plastids show extensive stromules and complex looping. (D) Plastids in the epidermal cell of an *Arabidopsis* root illuminated by targeting green fluorescent protein to the plastids. These plastids are highly variable in morphology, at the most extreme showing thin stromules. (E) Image of a pericarp cell in a tomato fruit at the onset of chromoplast differentiation from chloroplasts. Green fluorescent protein has been targeted to the plastids. Yellow plastid bodies exhibit both green fluorescent protein fluorescence as well as red chlorophyll fluorescence. Some red plastid bodies contain little GFP. A large number of plastid-derived structures, which contain bright green GFP are visible, both as stromules and distinct vesicle-like structures which appear to bud off from the main plastid bodies and lack chlorophyll. (F) Two chromoplasts in a tomato pericarp cell, which contain extensive GFP in the main chromoplasts bodies and are connected by two long thin stromules which show significant beading. (G and H) The cytoplasm of a pericarp cell from a ripe tomato view with brightfield (G) and with GFP fluorescence targeted to the plastid. A stromule emanating from a plastid body is obvious as are crystals of lycopene, which contain GFP and presumably are surrounded by a chromoplast membrane.

3.4 Leucoplasts and root plastids

Leucoplast is the name given to a general group of plastids, which lack any pigment and are often referred to as non-green plastids. Leucoplasts are very widely distributed in different plant tissues and have a wide range of morphologies and content, the latter being primarily a variation in the type of storage molecules that they accumulate. In fact amyloplasts could be considered a form of leucoplast that has specialised in storing starch. Whilst they are widespread in plant tissues, the general cell biology of leucoplasts has not been extensively investigated although many aspects of their biochemistry have been examined under the umbrella of non-green plastids (Emes and Neuhaus 1997; Eastmond et al. 1997). Leucoplasts can be isolated in clean populations from seed endosperm tissues and their biochemical characteristics examined (Negm et al. 1995).

A major class of leucoplasts are those found in different types of root cells of and often referred to as root plastids. These undoubtedly play a central role in root metabolism and function and many aspects of their biochemistry and metabolism have been described in detail (Emes and Neuhaus 1997; Debnam and Emes 1999; Fox et al. 2001). Early work examining electron micrographs of root plastids showed that proplastids in cells leaving the root apical meristem lose any thylakoid-like structures and pass through a transient phase of starch accumulation before becoming highly amoeboid in shape and then finally discoid with significant amounts of pregranal structures (Whatley 1983).

As with chromoplasts, the targeting of GFP to the plastid compartment has allowed the direct observation of root plastids in living tissue and they appear highly variable in morphology and exhibit many structures, which are reminiscent of stromules (Fig. 1). Indeed it is difficult to separate the presence of stromules on these plastids with variation in their morphology to the extent that stromules might be regarded as the most extreme characteristic of their morphological form. The cellular distribution of root plastids and leucoplasts in other tissues appears to be directed and non-random, since targeting of GFP to the plastid compartment reveals that leucoplasts commonly associate with the nucleus in an intimate manner, in that they surround the nucleus and are even found to lie within grooves in the nuclear membrane (Kwok and Hanson 2004). Such an association would seem to facilitate efficient signalling between plastid and nucleus and may also be a strategy for ensuring correct plastid segregation at cytokinesis (see Section 2).

Recent studies have revealed a novel role for plastids in directing the interaction of the root cells with symbiotic fungi and bacteria. Firstly, extensive plastid stromule networks develop in cells in arbscules where they interact with the fungal surface (Fester et al. 2001; Hans et al. 2004). Secondly, there is a major upregulation in plastid metabolic activity in these cells, as shown by transcript and metabolite profiling which provides a variety of metabolites which are central to the symbiotic interaction with the invading symbiont and the synthesis of the symbiotic structures such as the peri-arbuscular membrane (Lohse et al. 2005). Moreover two plastid membrane proteins, CASTOR and POLLUX, are crucial to the microbial admission into root cells, which forms the very first stage of the symbiotic relationship (Imaizumi-Anraku et al. 2005). Thus, it appears that a pre-existing endosymbiont in root cells, the plastid, and aids the integration of free-living soil bacteria into a symbiotic relationship with plants.

3.5 Other types of storage plastids

In addition to coloured pigments and starch, plastids are capable of accumulating other types of storage material. These can include lipids, which accumulate in elaioplasts and proteins, which accumulate in proteinoplasts. In both cases such plastids are found often in specialised cells within complex tissues. For instance, elaioplasts are commonly formed in the tapetal cells of the anther where they accumulate large amounts of neutral esters (Ting et al. 1998), which are released by elaioplast breakdown and contribute to the lipid component of the pollen wall (Clement and Pacini 2001). Storage lipids in plastids occur in structures called plastoglobules, which are commonly found in all plastid types. It is the extent of plastoglobule production, which essentially defines an elaioplast from any other plastid type, since elaioplasts are generally packed full of plastoglobuli. A recent proteome analysis of plastoglobuli reveals they contain several proteins involved in metabolism of isoprenoid derived molecules as well as fibrillins, which form a protein coat around the exterior of the plastoglobulus preventing coalescence (Ytterberg et al. 2006). This suggests that plastoglobuli have a metabolic role in the plastid rather than simply being a storage sac. It is unclear whether such a proteome and metabolome profile varies significantly between plastoglobuli in elaioplasts and those plastoglobuli, which appear less abundantly in other plastid types such as chloroplasts.

4 The control of plastid differentiation

The type of plastid present in a given type of cell is dictated by the nature of that cell type. Exactly how this developmental system is controlled by the host cell is largely unknown. It is normally assumed that the differentiation of proplastids into mature chloroplasts is the default pathway of plastid development, occurring in much of the above ground tissues in most plants. The leaf is a good organ in which to consider variations in tissue-specific chloroplast development. The fine tuning of this developmental process is significant since different cell types in a leaf all contain chloroplasts but these chloroplasts vary significantly in size, the extent of chlorophyll accumulation and membrane synthesis as well as large differences in their abundance within the cell. The most authentic development occurs in palisade and spongy mesophyll cells, where chloroplasts pack the cytoplasm and individual chloroplasts are fully photosynthetic with extensive thylakoid membrane and high levels of chlorophyll. In all other types of leaf cells, chloroplast development is less extensive and could be considered repressed. Although all other cell types in the leaf such as bundle sheath cells, epidermal pavement cells, vascular tissue, stomata, and hair cells have chlorophyll-containing chloroplast, the chloroplasts are all smaller, less well developed, and less abundant per cell. The implication is that a cell-specific repressive signal perturbs normal chloroplast development in these cells, resulting in poorly developed chloroplasts. Although photosynthetically compromised, these chloroplasts perform a crucial role in cellular metabolism in these different cell types and without them cellular function would be highly compromised.

A fundamental point of control in chloroplast differentiation is the presence of light, which initiates a complex chain of events inducing gene expression and protein synthesis, which in turn generates the proteome and the resulting metabolome of the mature chloroplast. A tight interaction between the developing chloroplast and cellular differentiation is crucial during this stage and a key part of this is a retrograde signalling pathway from the developing chloroplast back to the nucleus, which induces patterns of expression for genes, which encode plastid-destined proteins. Details of these molecular processes have been discussed extensively in recent times (Moller 2004; Lopez-Juez and Pyke 2005; Lopez-Juez 2007) and are also considered in other chapters in this book and will be considered only briefly here.

The big question remains as to what are the major control genes, which enable chloroplast differentiation to occur in a light–induced manner in mesophyll cells but not to the same extent, for instance, in neighbouring epidermal cells. Mutant screens for chloroplast biogenesis genes have identified a vast array of lines, mutant in genes which are critical for normal chloroplast function and which result in pale compromised seedlings. Many studies have shown that perturbation of genes which have basic functions in the chloroplast, result in pale compromised chloroplasts; for instance, mutation of the RpoTp RNA polymerase (Hricova et al. 2006) or components of the Clp protease core (Rudella et al. 2006). Sifting out from such collections, mutants that represent the major control genes in this system is very difficult, although directed efforts in this direction are being made (Gutierrez-Nava et al. 2004). Indeed one might forecast that mutation in a global master switch for plastid development would be embryo lethal and therefore unlikely to figure in screens for pale mutants. Lopez-Juez (2007) considers the possibility of global master switches, which facilitate chloroplast development from proplastids. Several candidates are possible although none have compelling evidence to merit them being in complete charge. Maybe the most likely candidates at present are *GLK* genes, which encode transcription factors and appear to be conserved in all land plants but not in single-celled photosynthetic organisms. Maize and *Arabidopsis* contain two *GLK* genes and when both are mutated, chloroplast development and thylakoid biogenesis is dramatically perturbed (Fitter et al. 2002; Yasamura et al. 2005). Intriguingly, *GLK* genes are not sufficient to overcome the general repression of chloroplast development in non-green tissues, as *GLK* overexpressing plants fail to develop green roots, for example, and thus function only in the correct developmental context.

Progress in understanding how plastids develop and the precise differences between differentiated plastid states will likely come from proteomic analyses of distinct cell types and the plastids within them (Kleffmann et al. 2006). Such technology has the potential to pinpoint subtle differences between plastids that currently are unknown. For instance, differences between chloroplasts in bundle sheath and mesophyll cells in leaves as highlighted by proteomic analysis, reveals subtle differences in addition to the basic known differences in photosynthetic metabolism (Majeran et al. 2005). It seems likely that progress using such strategies may well reveal that even the chloroplast actually represents a collection of subtly different organelles reflecting their precise development in different types of cell.

4.1 Plastid interconversions

Although chloroplast differentiation from proplastids, as directed by light, appears the central tenet of plastid biogenesis, there are many examples in which plastids can redifferentiate from pre-existing plastid types and form a different type of plastid (Fig. 2). Such interconversions are controlled by cellular developmental processes as well as environmental or hormonal signals and demonstrate an extreme plasticity in the plastid's functionality within the cell. Although several of these interconversion processes have been described, little is known of the exact molecular control of such redifferentiation processes. The best studied interconversion is that of chloroplasts redifferentiating into chromoplasts during fruit ripening, as discussed previously in this article. In tomatoes and peppers, the chromoplast differentiation pathway has a clear endpoint in mature ripe fruit, but in other systems such as orange citrus fruit ripening and maturation of pumpkins, the

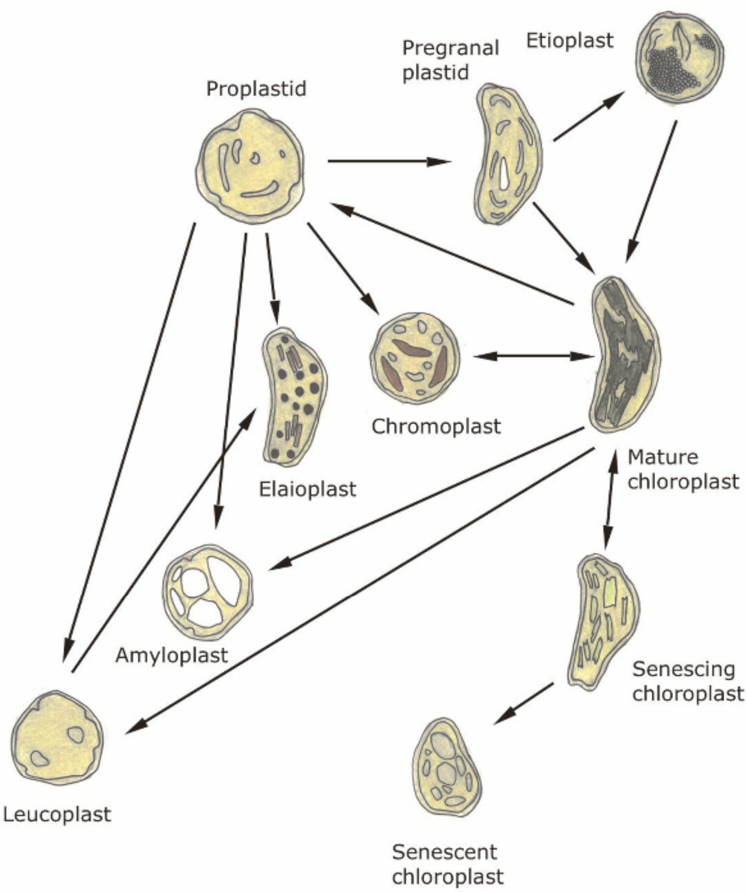


Fig. 2. A general scheme for interconversions of plastid types in different plant tissues. Although various routes for plastid interconversions are arrowed, it is likely that in various specific instances in different tissues, the majority of plastid types can interconvert to a different type.

orange chromoplasts are capable of reverting back to green chloroplasts. Application of the hormone gibberellin further promotes this process in oranges (Thomson et al. 1967). Another plastid redifferentiation pathway, which has major agronomic consequences, is the formation of chloroplasts from amyloplasts in the tissues of potato tubers as a result of illumination (Virgin and Sundquist 1992; Ljubicic et al. 1998). Although significant efforts are made to prevent such tuber greening during potato storage, the reason why this amyoplast-chloroplast interconversion is enabled in potato storage cells but prevented in other amyloplasts containing tissues, such as endosperm, is unclear. In reality, plastid interconversion is a common process during the development of complex tissues. For instance during the development of the anther, there is a complex pattern of interconversion between proplastids, amyloplasts, chromoplasts, chloroplasts, leucoplasts, and elaioplasts which varies in its nature according to the specific tissue type within the anther (Clement and Pacini 2001). In the face of such complex interactions, it could be more prudent to consider that no plastid differentiation pathways are terminal and that all plastids have the ability to change between different states according to the precise information derived from the cell.

5 Plastid division

The fact that plastids can divide as distinct organelles within the cytoplasm of the eukaryotic plant cell was confirmed by several studies in the late 1960s in which populations of plastids were counted and changes in their population size were established in correlation with cell expansion in developing leaves (see Pyke 1997). These studies clearly showed that there were two different points in plastid development where division takes place. Firstly in dividing cells in the meristem, proplastids are required to divide in order to maintain their lineage in newly divided cells. Without such a division, proplastids would likely be lost and aplastidic cells would be generated. Secondly, during the expansion phase of leaves, mesophyll cells increase in volume and the young chloroplasts divide during this period in order to maintain a population in an ever-enlarging cell. The final outcome of this process is mature mesophyll cells containing large populations of individual chloroplasts. The actual number of chloroplasts present is mainly related to the size of the cell, a relationship that extends across different species. In mature leaves in most species, the mesophyll cells contain between 50 and 200 chloroplasts. It is normally assumed that the basic mechanism by which plastids divide is the same for proplastids and for young chloroplasts although the control factors for these two processes are likely to differ. Cells in other green tissues in plants also accumulate chloroplasts in a similar manner although the end point of plastid population size and the size of individual plastids in different tissues and cell types varies greatly.

The plastid division process involves the constriction of the plastid centrally, which eventually leads to a pinching of the envelope membrane and fusion producing two separate daughter plastids, a process termed binary fission.

Progress in understanding the molecular basis of the plastid division machinery has been significant in the last 15 years due to approaches on two fronts. Firstly mutants of *Arabidopsis* were identified in which chloroplast numbers in leaf cells were altered significantly and secondly, genes involved in prokaryotic cell division were discovered in plant genomes and shown to function in plastid division (Pyke and Leech 1994; Pyke 1999). These two approaches have revealed many nuclear genes and their associated proteins and have enabled working models to be developed of how plastids divide (Aldridge et al. 2005). Central to the division process is the formation of a constriction ring composed primarily of FtsZ proteins that resides on the inside of the plastid envelope in the stroma (Osteryoung and Vierling 1995; Osteryoung et al. 1998; Vitha et al. 2001; McAndrew et al. 2001;

Kuroiwa et al. 2002). FtsZ proteins have characteristics of the cytoskeletal protein tubulin, and plant FtsZ proteins are homologues of those present in prokaryotic bacteria, which function in bacterial cell division. The FtsZ ring is stabilized by the protein ARC6, originally identified from an *Arabidopsis* mutant with few giant chloroplasts (Pyke et al. 1994; Vitha et al. 2003). FtsZ proteins assemble in the ring structure at the onset of plastid division and constriction of the ring and force generation appears to be controlled by the protein ARC5, which is a dynamin-like protein (Gao et al. 2003; Miyagishima et al. 2003) which functions on the outer surface of the plastid envelope. Evidence that ARC5 generates force and constricts the FtsZ ring complex comes from viewing isolated FtsZ rings and inducing constriction by adding ARC5 protein to them (Yoshida et al. 2006). Coordination of events on either side of the plastid envelope as the division process progresses appears to be controlled by PDV1 proteins, which form foci in the outer plastid envelope overlying the stromal FtsZ ring (Miyagishima et al. 2006). Undoubtedly, the complete plastid division machinery is a complex structure and probably contains other unknown proteins which function in a combinatorial fashion to facilitate the division process (Maple et al. 2005) especially since imaging of isolated plastid division rings containing FtsZ show distinct rings on the outer and inner surfaces of the plastid envelope (Kuroiwa et al. 2002; Miyagishima et al. 2001, 2003). Figure 3 shows a tentative model of how these proteins and the plastid division rings could be arranged. Plastid division normally occurs at the midpoint of the plastid such that the two daughter plastids, which result from the division process, are equally sized. The mechanism that ensures this equality is based on the system of Min genes, which function for a similar purpose in bacterial cell division. MinD (Colletti et al. 2000) and MinE (Itoh et al. 2001) both dictate that the FtsZ ring is allowed to form only in the middle of the plastid's long axis and is prevented from forming at either pole (Fujiwara et al. 2004). Interestingly these genes define the fact that plastids have distinct poles and are not unpolarised organelles as has been generally believed. Although the third member of the bacterial family of Min genes, MinC, appears to be absent from plant genomes, expression of the prokaryotic MinC gene in *Arabidopsis* interferes with the plastid division machinery and results in abnormally large chloroplasts (Tavva et al. 2006). Whether this is a direct interaction between the Min proteins or an effect of MinC directly on FtsZ functionality is unclear.

A clear theme which has arisen from the recent knowledge about the molecular basis of plastid division is that the division machinery currently used by chloroplasts involves proteins originally involved in prokaryotic cell division, reflecting the plastid's ancestry, and new genes which have been hijacked from the plant's genome. In addition to the genes already mentioned, *ARTEMIS* (Fulgosi et al. 2002) and *GIANT CHLOROPLAST 1* (Maple et al. 2004) are both related to prokaryotic proteins and both function in plastid division, since perturbation of them results in abnormal plastids. Conversely, ARC5 has no prokaryotic relations and ARC3 is a chimera of an FtsZ gene and a eukaryotic gene, phosphatidylinositol-4-phosphate 5-kinase (Shimada et al. 2004). It is clear that during the evolution of the plastid replication process, plant nuclear genes were recruited to interact with

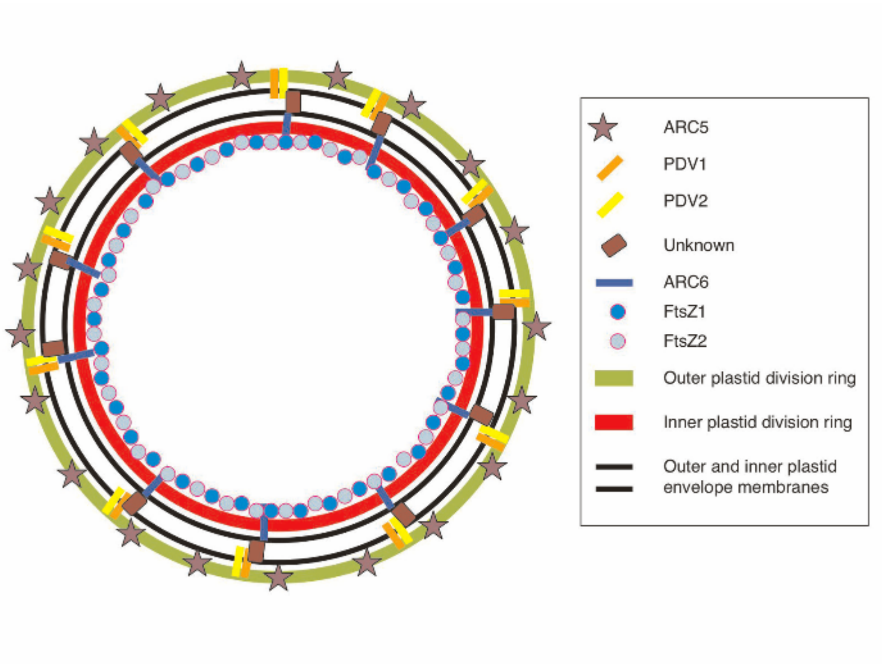


Fig. 3. A tentative model for the arrangement of proteins and the plastid dividing rings at the midpoint of a plastid about to commence division. ARC5 is associated with the outer plastid division ring and the proteins PDV1 and PDV2 link the outer plastid division ring to the plastid envelope membrane. An unknown protein spans the lumen of the envelope membrane and provides attachment points for ARC6, which links the inner envelope membrane to the inner plastid division ring and the FtsZ ring, composed of FtsZ1 and FtsZ2 proteins. Adapted from Glynn et al. (2007).

the solely prokaryotic process in order to enable control of the process in the endosymbiotic organelles by the plant nucleus. There are many questions still to be answered concerning the control of the plastid division machinery including how it is activated and stopped and how is the division of large populations of organelles during leaf cell expansion coordinated? Another often overlooked question is what suppresses the division machinery in cells where plastid replication rarely occurs and where plastid populations are relatively sparse, as in leaf epidermal cells. In addition, it is normally assumed that the binary fission type of plastid division as discussed here is the sole type of mechanism by which plastids divide. However, replication by a budding type of mechanism, which could be regarded as an extreme asymmetric type of binary division, does occur occasionally in plants (Kulandaielu and Gnanam 1985) and has been shown clearly in giant plastids of the *suffulta* mutant in tomato (Forth and Pyke 2006) where small budding vesicles bud off from the large plastid body as the chloroplasts differentiate into chromoplasts. Highly asymmetric chloroplast division has been observed in plants of *arc*11 (Marrison et al. 1999), which contain a mutation in the MinD gene (Fijiwara et al. 2004), so it is conceivable that a budding type mechanism could result from a breakdown in the Min centralizing system. To date, plastid division mechanisms and cell biology have only been studied in relatively few plants species and it will be interesting to ascertain the degree of variation in division mechanisms that might exist in all higher plants.

6 Stromules

Over the last decade, several important developments have occurred in our understanding of plastids. In addition to major developments in the understanding of molecular processes which occur during plastid development, a subject considered in several other chapters in this book, a renewed consideration of plastid morphology and the dynamic nature of changes in plastid morphology has also taken place. Central to this latter consideration has been the exploitation of green fluorescent protein targeted to the plastid compartment, which has revealed dramatic new aspects of plastid morphology called stromules (Fig. 1). These long thin membranous tubules containing stroma but not thylakoid membrane or chlorophyll were rediscovered in the late 1990s (Köhler et al. 1997) by imaging GFP fluorescence in plastids of tobacco and petunia containing GFP. These stromules were between 350 and 850 nm in diameter and were highly dynamic in nature extending from and retracting into the plastid body and occasionally interacting with a stromule from a neighbouring plastid. In this case, the movement of GFP from one plastid to another by stromule transfer was shown using photobleaching (Köhler et al. 2000). Ironically, the modern day observation of stromules emanating from plastids was a reconfirmation of many observations made through the last century in which microscopists have observed various protrusions and dynamic extensions of plastids in many different types of tissue (Gray et al. 2001; Kwok and Hanson 2004). Wildmann's laboratory at the University of California was famous in the 1960s for images and movies of highly dynamic plastids producing long thin extensions in the cytoplasm which can fragment, leading to the improbable suggestion that these smaller structures become mitochondria (Wildmann et al. 1962; Wildmann 1967). What we now call stromules are clearly seen in his pictures. Perhaps not surprisingly, stromule-like structures were not considered seriously within the plastid community until their rediscovery 30 years later (Tobin 1997). So how do stromules form and what do they do?

Stromules form by dynamic out growth of the plastid envelope membranes and their movement within the cytosol is controlled in part by the actin microfilament cytoskeletal system in which myosin motors link stromules and plastid bodies to the actin microfilaments (Kwok and Hanson 2003, 2004a). Careful observation of stromules with DIC optics (Gunning 2004, 2005) has revealed a great deal about the precise dynamics of stromule interaction with the microfilament tracks and clearly shows how stromules are pulled out from plastid bodies by attachment to microfilament tracks at points of attachment, not only at stromule tips but also at points along the stromule length. Sudden loss of attachment causes rapid recoil of the stromule. In addition, stromules can also branch and rejoin forming closed loops as well as forming distinct bead-like structures along their length. Beads are particularly clear in stromules on chromoplasts in tomato fruit (Pyke and Howells 2002) although there is little evidence that such structures actually move along the stromule length. Whether the extension of stromules is entire due to pulling by the microfilament strands rather than a pushing out by a stromal pressure is unclear as is the exact source of the new membrane needed to produce a new stromule.

So what do stromules do to aid plastid function? At present the precise role of stromules is unclear but several considerations have been made. It is obvious that production of a stromule by a plastid will increase its surface area significantly and thereby increase the surface of interaction with the cytosol. Since plastids are highly active in cellular biochemistry and are sites of synthesis of many molecules important in cellular function, an increased surface area should potentially improve this interaction. This suggestion makes the assumption that the envelope membranes in the stromule have similar import capacity to that of the plastid body, a fact that has yet to be clearly addressed. The potential for movement of molecules between plastids has been demonstrated but how relevant this process might be to what actually occurs within the cell is difficult to determine. Certainly observation of plastids and stromules in the majority of cell types suggests that such joining is relatively rare and probably transitory in nature. A key point in trying to understand what stromules do is a clear distinction between their propensities in different types of cells and in particular their relative rarity in cells containing mature green chloroplasts. Thus, in mesophyll cells, which are packed with chloroplasts, stromules are rarely seen whereas in other cells containing non-green plastids such as in root cells, petal cells, epidermal cells and cultured suspension cells, stromules are much more abundant. Waters et al. (2004) showed that a decline in plastid density in the epidermal cells of expanding tobacco hypocotyls is correlated with a significant increase in stromule length raising the possibility that stromules act as a density sensing mechanism for plastids which are far apart. This could also tie in with mechanosensing proteins in the plastid envelope which sense when plastids are squashed together (Haswell and Meyerowitz 2006). In many cells containing non-green plastids, stromule networks are extensive and appear to link plastids, which are closely associated with the nucleus and surround it, to the peripheral cell membrane (Kwok and Hanson 2004b). Maybe stromules are involved with intracellular communication in some way. Fragmentation of stromules into distinct vesicles has also been suggested as a method of plastid replication since pieces of broken stromule in ripening tomato fruit cells appear to differentiate as chromoplasts. More work on stromules will be required to understand more fully these enigmatic interesting structures associated with plastids.

7 Conclusion

Our understanding of some of the cell biology aspects of plastids have improved significantly in the last two decades and the plastid has risen above the status of an organelle that carries out only photosynthesis. The advent of omic technology has the potential for describing subtle differences between different types of plastids and may give clues as to how master controlling genes work, if they exist. Even so, we are still a long way from a clear understanding of what determines a particular plastid type in a particular type of cell and what facilitates the interconversions of different plastid types. Maybe the next decade will see big advances in addressing these questions.

Acknowledgement

Thanks to Mark Waters (University of Oxford) for critical reading of the manuscript and helpful discussions and to Deborah Ballinger for the artwork in Figure 2.

References

Akita T, Sagisaka S (1995) Functional and structural differences among proplastids as seen by immunogold staining and electron microscopy. Biosci Biotech Biochem 59:1477-1484

Aldridge C, Maple J, Moller SG (2005) The molecular biology of plastid division in higher plants. J Exp Bot 56:1061-1077

Alexander L, Grierson D (2002) Ethylene biosynthesis and action in tomato: a model for climacteric fruit ripening. J Exp Bot 53:2039-2055

Andersson MX, Sandelius AS (2004) A chloroplast-localized vesicular transport system: a bioinformatics approach. BMC Genomics 5:40

Armstrong GA, Hearst JE (1996) Genetics and molecular biology of carotenoid pigment biosynthesis. FASEB Review 10:228-237

Arvidsson PO, Sundby C (1999) A model for the topology of the chloroplast thylakoid membrane. Aust J Plant Physiol 26:687-694

Aseeva E, Ossenbuhl F, Eichacker LA, Wanner G, Soll J, Vothknecht UC (2004) Complex formation of Vipp1 depends on its ã helical PspA-like domain. J Biol Chem 279:35535-35541

Bathgate B, Purton ME, Grierson D, Goodenough PW (1985) Plastid changes during the conversion of chloroplasts to chromoplasts in ripening tomatoes. Planta 165:197-204

Baumgartner BJ, Rapp JC, Mullet JE (1989) Plastid transcription activity and DNA copy number increase early in barley chloroplast development. Plant Physiol 89:1011-1018

Bechtel DB, Wilson JD (2003) Amyloplast formation and starch grain development in hard winter wheat. Cereal Chem 80:175-183

Bramley PM (2002) Regulation of carotenoid formation during tomato fruit formation and development. J Exp Bot 53:2107-2113

Bruno AK, Wetzel CM (2004) The early light-inducible protein (ELIP) gene is expressed during the chloroplast-to-chromoplast transition in ripening fruit. J Exp Bot 55:2541-2548

Camara B, Hugueney P, Bouvier F, Kuntz M, Moneger R (1995) Biochemistry and molecular biology of chromoplast development. Int Rev Cytol 163:175-247

Chaley N, Possingham JV (1981) Structure of constricted proplastids in meristematic plant tissues. Biologie Cellulaire 41:203-210

Clement C, Pacini E (2001) Anther plastids in Angiosperms. Bot Rev 67:54-73

Colletti KS, Tattersall EA, Pyke KA, Froelich JE, Stokes KD, Osteryoung KW (2000) A homologue of the bacterial cell division site-determining factor MinD mediates placement of the chloroplast division apparatus. Curr Biol 10:507-516

Corriveau JL, Coleman AW (1988) Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 Angiosperm species. Am J Bot 75:1443-1458

Cran DG, Possingham JV (1972) Variation in plastid types in spinach. Protoplasma 74:345-356

Cunningham FX, Gantt E (1998) Genes and enzymes of carotenoid biosynthesis in plants. Ann Rev Plant Physiol Plant Mol Biol 49:557-583

Debnam PM, Emes MJ (1999) Subcellular distribution of enzymes of the oxidative pentose phosphate pathway in root and leaf tissues. J Exp Bot 40:1653-1661

Duffus CM (1984) Metabolism of reserve starch. In: Lewis DH (ed) Storage carbohydrates in vascular plants. Cambridge UK: Cambridge University Press

Eastmond PJ, Dennis DT, Rawsthorne S (1997) evidence that a malate/inorganic phosphate exchange translocator imports carbon across the leucoplast envelope for fatty acid synthesis in developing castor seed endosperm. Plant Physiol 114:851-856

Ellis JR, Leech RM (1983) Cell size and chloroplast size in relation to chloroplast replication in light-grown wheat leaves. Planta 165:120-125

Emes MJ, Neuhaus HE (1997) Metabolism and transport in non-photosynthetic plastids. J Exp Bot 48:1995-2005

Fester T, Strack D, Hause B (2001) Reorganization of tobacco root plastids during arbuscle development. Planta 213:864-868

Fitter DW, Martin DJ, Copley MJ, Scotland RW, Langdale JA (2002) GLK gene pairs regulate chloroplast development in diverse plant species. Plant J 31:713-727

Forth D, Pyke KA (2006) The *suffulta* mutation in tomato reveals a novel method of plastid replication during fruit ripening. J Exp Bot 57:1971-1979

Fox SR, Rawsthorne S, Hills MJ (2001) Fatty acid synthesis in pea root plastids is inhibited by the action of long-chain acyl coenzyme as on metabolite transporters. Plant Physiol 126:1259-1265

Fraser PD, Truesdale MR, Bird CR, Schuch W, Bramley PM (1994) Carotenoid biosynthesis during tomato fruit development. Plant Physiol 105:405-413

Fujie M, Kuroiwa H, Kawano S, Kuroiwa T (1994) Behaviour of organelles and their nucleoids in the shoot apical meristem during leaf development in *Arabidopsis thaliana* L. Planta 194:395-405

Fijiwara M, Nakamura A, Itoh R, Shimada Y, Yoshida S, Moller SG (2004) Chloroplast division site placement requires dimerisation of the ARC11/AtMind1 protein in *Arabidopsis*. J Cell Sci 117:2399-2410

Fulgosi H, Gerdes L, Westphal S, Glockmann C, Soll J (2002) Cell and chloroplast division requires ARTEMIS. Proc Nat Acad Sci 99:11501-11506

Gao H, Kadirjan-Kalbach D, Froehlich JE, Osteryoung KW (2003) ARC5, a cytosolic dynamin-like protein from plants, is part of the chloroplast division machinery. Proc Natl Acad Sci USA 100:4328-4333

Gao H, Sage TL, Osteryoung KW (2006) FZL, an FZO-like protein in plants, is a determinant of thylakoid and chloroplast morphology. Proc Natl Acad Sci USA 103:6759-6764

Glynn JM, Miyagishima S, Yoder DW, Osteryoung KW (2007) Chloroplast division. Traffic 8:451-461

Gutierrez-Nava Mde L, Gillmor CS, Jimenez LF, Guevara-Garcia A, Leon P (2004) *CHLOROPLAST BIOGENESIS* genes act cell and noncell autonomously in early chloroplast development. Plant Physiol 135:471-482

Gray JC, Sullivan JA, Hibberd JM, Hanson MR (2001) Stromules: mobile protrusions and interconnections between plastids. Plant Biol 3:223-233

Gunning BES (2004) Plant cell biology on CD: information for students and a resource for teachers. www.plantcellbiologyoncd.com

Gunning BES (2005) Plastid stromules: video microscopy of their outgrowth, retraction, tensioning, anchoring, branching, bridging, and tip-shedding. Protoplasma 225:33-42

Hans J, Hause B, Strack D, Walter MH (2004) Cloning, characterisation, and immunolocalization of a mycorrhiza-inducible 1-deoxy-D-xylulose-5-phosphate reuctoisomerase in arbscule-containing cells of maize. Plant Physiol 134:614-624

Harak H, Lagrange T, Bisanz-Seyer C, Lerbs-Mache S, Mache R (1995) The expression of nuclear genes encoding plastid ribosomal proteins precedes the expression of chloroplast genes during early phases of chloroplast development. Plant Physiol 108:685-692

Harris WM, Spurr AR (1969a) Chromoplasts of tomato fruits I. Ultrastructure of low-pigment and high beta mutants. Carotene analyses. Am J Bot 56:369-379

Harris WM, Spurr AR (1969b) Chromoplasts of tomato fruits II. The red tomato. Am J Bot 56:380-389

Haswell ES, Meyerowitz EM (2006) MscS-like proteins control plastid size and shape in *Arabidopsis thaliana*. Curr Biol 16:1-11

Hricova A, Quesada V, Micol JL (2006) The *SCABRA3* nuclear gene encodes the plastid RpoTp RNA polymerase, which is required for chloroplast biogenesis and mesophyll cell proliferation in *Arabidopsis*. Plant Physiol 141:942-956

Hugeney P, Bouvier F, Badillo A, Dharlingue A, Kuntz M, Camara B (1995) Identification of a plastid protein involved in vesicle formation and/or membrane-protein translocation. Proc Nat Acad Sci USA 92:5630-5634

Hugueny P, Badillo A, Chen HC, Klein A, Hirschberg J, Camara B, Kuntz M (1995a) Metabolism of cyclic carotenoids: model for the alteration of this biosynthetic pathway in *Capsicum annuum*. Plant J 8:417-424

Imaizumi-Anraku H, Takeda N, Charpentier M, Perry J, Miwa H, Umehara Y, Kouchi H, Murakami Y, Mulder L, Vickers K, Pike J, Downie JA, Wang T, Sato S, Asamizu E, Tabata S, Yoshikawa M, Murooka Y, Wu GJ, Kawaguchi M, Kawasaki S, Parniske M, Hayashi M (2005) Plastid proteins crucial for symbiotic fungal and bacterial entry into plant roots. Nature 433:527-531

Itoh R, Fujiwara M, Nagata N, Yoshida S (2001) A chloroplast protein homologous to the eubacterial topological specificity factor *minE* plays a role in chloroplast division. Plant Physiol 127:1644-1655

Josse E-M, Simkin AJ, Gaffe J, Laboure A-M, Kuntz M, Carol P (2000) A plastid terminal oxidase associated with carotenoid desaturation during chromoplast differentiation. Plant Physiol 123:1427-1436

Joyard J, Teyssier E, Miege C, Berny-Seigneurin D, Marechal E, Block MA, Dorne AJ, Rolland N, Ajlani G, Douce R (1998) The biochemical machinery of plastid envelope membranes. Plant Physiol 118:715-723

Kay QON, Daoud HS, Stirton CH (1981) Pigment distribution, light reflection and cell structure in petals. Bot J Linn Soc 83:57-84

Kleffmann T, Hirsch-Hoffmann M, Gruissem W, Baginsky S (2006) plprot: a comprehensive proteome database for different plastid types. Plant Cell Physiol 47:432-436

Kobayashi H, Ngernprasirtsiri J, Akazawa T (1990) Transcriptional regulation and DNA methylation in plastids during transitional conversion of chloroplasts to chromoplasts. EMBO J 9:307-313

Köhler R, Cao J, Zipfel W, Webb W, Hanson M (1997) Exchange of protein molecules through connections of higher plant plastids. Science 276:2039-2042

Köhler R, Hanson M (2000) Plastid tubules of higher plants are tissue-specific and developmentally regulated. J Cell Sci 113:81-89

Kulandaivelu G, Gnanam A (1985) Scanning electron microscopic evidence for a budding mode of chloroplast multiplication in higher plants. Physiol Plant 63:299–302

Kroll D, Meierhoff K, Bechtold N, Kinoshita M, Westphal S, Vothknecht UC, Soll J, Westhoff P (2001) *VIPP1*, a nuclear gene in *Arabidopsis thaliana* essential for thylakoid membrane formation. Proc Nat Acad Sci USA 98:4238-4242

Kuroiwa H, Mori T, Takahara M, Miyagishima S, Kuroiwa T (2002) Chloroplast division machinery as revealed by immunofluorescence and electron microscopy. Planta 215:185-190

Kwok EY, Hanson MR (2003) Microfilaments and microtubules control the morphology and movement of non-green plastids and stromules in *Nicotiana tabacum*. Plant J 35:16-26

Kwok EY, Hanson MR (2004) Stromules and the dynamic nature of plastid morphology. J Microsc 214:124-137

Kwok EY, Hanson MR (2004a) *In vivo* analysis of interactions between GFP-labelled microfilaments and plastid stromules BMC. Plant Biol 4:2

Kwok EY, Hanson MR (2004b) Plastids and stromules interact with the nucleus and cell membrane in vascular plants. Plant Cell Rep 23:188-195

Langeveld SMJ, Van Wijk R, Stuurman N, Kijne JW, de Pater S (2000) B-type granule containing protrusions and interconnections between amyloplasts in developing wheat endosperm revealed by transmission electron microscopy and GFP expression. J Exp Bot 51:1357-1361

Lawrence SD, Cline K, Moore GA (1993) Chromoplast targeted proteins in tomato (*Lycopersicon esculentum* Mill.) fruit. Plant Physiol 102:789-794

Lawrence SD, Cline K, Moore GA (1997) Chromoplast development in ripening tomato fruit: identification of cDNAs for chromoplast-targeted proteins and characterisation of a cDNA encoding a plastid-localised low molecular weight heat shock protein. Plant Mol Biol 33:483-492

Livne A, Gepstein S (1988) Abundance of the major chloroplast polypeptides during development and ripeninjg of tomato fruits. Plant Physiol 87:239-243

Lohse S, Schliemann W, Ammer C, Kopka J, Strack D, Fester T (2005) Organization and metabolism of plastids and mitochondria in arbuscular mycorrhizal roots of *Medicago truncatula*. Plant Physiol 139:329-340

Lois LM, Rodriguez-Concepcion M, Gallego F, Campos N, Boronat A (2000) Carotenoid biosynthesis during tomato fruit development: regulatory role of 1-deoxy-D-xylulose 5-phosphate synthase. Plant J 22:503-513

Lopez EJ, Pyke KA (2005) Plastids unleashed: their development and their integration in plant development. Int J Dev Biol 49:557-577

Lopez EJ (2007) Plastid biogenesis, between shadows and light. J Exp Bot 58:11-26

Ljubicic JM, Wrischer M, Ljubicic N (1998) Formation of the photosynthetic apparatus in plastids during greening of potato microtubules. Plant Physiol Biochem 36:747-752

Lyndon RF, Robertson ES (1976) The quantitative ultrastructure of the pea shoot apex in relation to leaf initiation. Protoplasma 87:387-402

Mache R, Zhou D-X, Lerbs-Mache S, Harrak H, Villain PM, Gauvin S (1997) Nuclear control of early plastid differentiation. Plant Physiol Biochem 35:199-203

Majeran W, Cai Y, Sun Q, van Wijk KJ (2005) Functional differentiation of bundle sheath and mesophyll maize chloroplasts determined by comparative proteomics. Plant Cell 17:3111-3140

Maple J, Fujiwara MT, Kitahata N, Lawson T, Baker NR, Yoshida S, Moller SG (2004) GIANT CHLOROPLAST 1 essential for correct chloroplast division site placement in *Arabidopsis*. Curr Biol 14:776-781

Maple J, Aldrige C, Møller SG (2005) Plastid division is mediated by combinatorial assembly of plastid division proteins. Plant J 43:811–823

Marrison JL, Rutherford SM, Robertson EJ, Lister C, Dean C, Leech RM (1999) The distinctive roles of five different ARC genes in the chloroplast division process in *Arabidopsis*. Plant J 18:776-781

McAndrew RS, Froehlich JE, Vitha S, Stokes KD, Osteryoung KW (2001) Co-localization of plastid division proteins in the chloroplast stromal compartment establishes a new functional relationship between FtsZ1 and FtsZ2 in higher plants. Plant Physiol 127:1656-1666

Miyagishima S, Takahara M, Kuroiwa T (2001) Novel filaments 5 nm in diameter constitute the cytosolic ring of the plastid division apparatus. Plant Cell 13:707–721

Miyagishima S, Nishida K, Mori T, Matsuzaki M, Higashiyama T, Kuroiwa H, Kuroiwa T (2003) A plant-specific dynamin-related protein forms a ring at the chloroplast division site. Plant Cell 15:655-665

Miyagishimaa S, Froehlichb JE, Osteryoung KW (2006) The outer envelope protein PDV, together with its paralogue PDV, mediates recruitment of the dynamin-related protein ARC5 to the plastid division site in *Arabidopsis*. Plant Cell 18:2517-2530

Mogensen HL (1996) The hows and whys of cytoplasmic inheritance in seed plants. Am J Bot 83:383-404

Moller SG (2004) Plastids. Annual Plant Reviews Vol. 13. Oxford: Blackwell Publishing

Morre DJ, Sellden G, Sundqvist C, Sandelius AS (1991) Stromal low temperature compartment derived from inner membrane of the chloroplast envelope. Plant Physiol 97:1558-1564

Mustardy L, Garab G (2003) Granum revisited. A three-dimensional model – where things fall into place. Trends Plant Sci 8:117-122

Negm FB, Cornel FA, Plaxton WC (1995) Suborganellar organization and molecular characterization of nonproteolytic degraded leucoplast pyruvate kinase from developing endosperm of castor oil seeds (*Ricinus communis*). J Exp Bot 34:712-718

Osteryoung KW, Vierling E (1995) Conserved cell and organelle division. Nature 376:473–474

Osteryoung KW, Stokes KD, Rutherford SM, Percival AL, Lee WY (1998) Chloroplast division in higher plants requires members of two functionally divergent gene families with homology to bacterial *ftsZ*. Plant Cell 10:1991-2004

Park JM, Cho JH, Kang SG, Jang HJ, Pih KT, Piao HL, Cho MJ, Hwang I (1998) A dynamin-like protein in *Arabidopsis thaliana* is involved in biogenesis of thylakoid membrane. EMBO J 17:859-867

Pyke KA (1997) The genetic control of plastid division in higher plants. Am J Bot 84:1017-1027

Pyke KA (1999) Plastid division and development. Plant Cell 11:549-556

Pyke KA (2006) Plastid division: the squeezing gets tense. Curr Biol 16:R60-R62

Pyke K, Leech RM (1992) Chloroplast division and expansion is radically altered by nuclear mutations in *Arabidopsis thaliana*. Plant Physiol 99:1005-1008

Pyke KA, Leech RM (1994) A genetic analysis of chloroplast division and expansion in *Arabidopsis thaliana*. Plant Physiol 104:201-207

Pyke KA, Howells CA (2002) Plastid and stromule morphogenesis in tomato. Ann Bot 90:559-566

Pyke KA, Rutherford SM, Robertson EJ, Leech RM (1994) *arc6*, a fertile *Arabidopsis* mutant with only two mesophyll cell chloroplasts. Plant Physiol 106:1169-1177

Reski R (2002) Rings and networks: the amazing complexity of FtsZ in chloroplasts. Trends Plant Sci 7:103-105

Robertson EJ, Pyke KA, Leech RM (1995) *arc6*, a radical chloroplast division mutant of *Arabidopsis* also alters proplastid proliferation and morphology in shoot and root apices. J Cell Sci 108:2937-2944

Romer S, d'Harlingue A, Camara B, Schantz R, Kuntz M (1992) Cysteine synthase from *Capsicum annum* chromoplasts. Characterization and cDNA cloning of an up-regulated enzyme during fruit development. J Biol Chem 267:17466-17470

Rudellaa A, Frisoa G, Alonsob JM, Ecker JR, van Wijk KJ (2006) Downregulation of ClpR2 leads to reduced accumulation of the ClpPRS protease complex and defects in chloroplast biogenesis in *Arabidopsis*. Plant Cell 18:1704-1721

Sakai A, Kawano S, Kuroiwa T (1992) Conversion of proplastids to amyloplasts in tobacco cultured cells is accompanied by changes in the transcriptional activities of plastid genes. Plant Physiol 100:1062-1066

Sakai A, Susuki T, Sasaki N, Kuroiwa T (1999) Plastid gene expression during amyloplast formation in cultured tobacco cells. J Plant Physiol 154:71-78

Sakai A, Susuki T, Miyazawa Y, Kawano S, Nagata T, Kuroiwa T (1998) Comparative analysis of plastid gene expression in tobacco chloroplasts and proplastids: relationship between transcription and transcript accumulation. Plant Cell Physiol 39:581-589

Sheahan MB, Rose RJ, McCurdy DW (2004) Organelle inheritance in plant cell division: the actin cytoskeleton is required for unbiased inheritance of chloroplasts, mitochondria and endoplasmic reticulum in dividing protoplasts. Plant J 37:379-390

Shimada H, Koizumi M, Kuroki K, Mochizuki M, Fujimoto H, Ohta H, Masuda T, Takamiya K (2004) ARC3, a chloroplast division factor, is a chimera of prokaryotic FtsZ and part of eukaryotic phosphatidylinositol-4-phosphate 5-kinase. Plant Cell Physiol 45:960-967

Shimoni E, Rav-Hon O, Ohad I, Brumfeld V, Reich Z (2005) Three-dimensional organization of higher plant chloroplast thylakoid membranes revealed by electron tomography. Plant Cell 17:2580-2586

Smith AM, Denyer K, Martin C (1997) The synthesis of the starch granule. Ann Rev Plant Physiol Mol Biol 48:67-87

Swarup R, Kramer EM, Knox K, Leyser HMO, Haseloff J, Beemster GTS, Bhalerao R, Bennett MJ (2005) Root gravitropism requires lateral root cap and epidermal cells for transport and response to a mobile auxin signal. Nat Cell Biol 7:1057-1065

Summer EJ, Cline K (1999) Red bell pepper chromoplasts exhibit *in vitro* import competency and membrane targeting of passenger proteins from the thylakoid Sec and delta pH pathways but not the chloroplast signal recognition particle pathway. Plant Physiol 119:575-584

Tavva VS, Collins GB, Dinkins RD (2006) Targeted overexpression of the *Escherichia coli* MinC protein in higher plants results in abnormal chloroplasts. Plant Cell Rep 25:341-348

Thomson WW, Whatley JM (1980) Development of non-green plastids. Ann Rev Plant Physiol 31:375-394

Ting JTL, Wu SSH, Ratnayake C, Huang AHC (1998) Constituents of the tapetosomes and elaioplasts in *Brassica campestris* tapetum and their degradation and retention during microsporogenesis. Plant J 16:541-551

Thomson WW, Lewis N, Coggins CW (1967) The reversion of chromoplasts to chloroplasts in Valencia oranges. Cytologia 32:117-124

Tobin EM (1997) Renewing an old view of chloroplasts. Trends Plant Sci 2:405-406

Vishnevetsky M, Ovadis M, Vainstein A (1999) Carotenoid sequestration in plants: the role of carotenoid associated proteins. Trends Plant Sci 4:232-235

Vitha S, McAndrew RS, Osteryoung KW (2001) FtsZ ring formation at the chloroplast division site in plants. J Cell Biol 153:111–119

Vitha S, Froehlich JE, Koksharova O, Pyke KA, van Erp H, Osteryoung KW (2003) ARC6 is a J-domain plastid division protein and an evolutionary descendant of the cyanobacterial cell division protein Ftn2. Plant Cell 15:1918-1933

Virgin HI, Sundquist C (1992) Pigment formation in potato tubers (*Solanum tuberosum*) exposed to light following darkness. Physiol Plant 86:587-592

Waters M, Fray R, Pyke K (2004) Stromule formation is dependent upon plastid size, plastid differentiation status and the density of plastids within the cell. Plant J 39:655-667

Wang Q, Sullivan RW, Kight A, Henry RL, Huang J, Jones AM, Korth KL (2004) Deletion of the chloroplast-localized Thylakoid formation1 gene product in *Arabidopsis* leads to deficient thylakoid formation and variegated leaves. Plant Physiol 136:3594-3604

Whatley JM (1983) The ultrastructure of plastids in roots. Int Rev Cytol 85:175-220

Weston EA, Pyke KA (1999) Developmental ultrastructure of cells and plastids in the petals of Wallflower (*Erysimum cheiri*). Ann Bot 84:763-769

Westphal S, Soll J, Vothknecht UC (2003) Evolution of chloroplast vesicle transport. Plant Cell Physiol 44:217-222

Wildman S, Hongladarom T, Honda S (1962) Chloroplasts and mitochondria in living plant cells: cinephotomicrographic studies. Science 138:434-435

Wildman SG (1967) The organization of grana-containing chloroplasts in relation to location of some enzymatic systems concerned with photosynthesis, protein synthesis, and ribonucleic acid synthesis. In: Goodwin TW (ed) Biochemistry of Chloroplasts Vol. 2. London: Academic Press pp 295-319

Yamamoto K, Pyke KA, Kiss JZ (2002) Reduced gravitropism in inflorescence stems and hypocotyls, but not roots, of *Arabidopsis* mutants with large plastids. Physiol Planta 114:627-636

Yasumura Y, Moylan EC, Langdale JA (2005) A conserved transcription factor mediates nuclear control of organelle biogenesis in anciently diverged land plants. Plant Cell 17:1894-1907

Yoshida Y, Kuroiwa H, Misumi O, Nishida K, Yagisawa F, Fujiwara T, Nanamiya H, Kawamura F, Kuroiwa T (2006) Isolated chloroplast division machinery can actively constrict after stretching. Science 313:1435-1438

Ytterberg JA, Peltier J-B, van Wijk KJ (2006) Protein profiling of plastoglobules in chloroplasts and chromoplasts. A surprising site for differential accumulation of metabolic enzymes. Plant Physiol 140:984-997

Zeeman SC, Smith SM, Smith AM (2004) The breakdown of starch in leaves. New Phytol 163:247-261

Zhang Q, Liu Y, Sodmergen (2003) Examination of the cytoplasmic DNA in male reproductive cells to determine the potential for cytoplasmic inheritance in 295 Angiosperm species. Plant Cell Physiol 44:941-951

Pyke, Kevin

Plant Sciences Division, School of Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough, Leicestershire LE12 7RD

Kevin.Pyke@nottingham.ac.uk

Structure, function, and inheritance of plastid genomes

Ralph Bock

Abstract

Plastids (chloroplasts) possess their own genetic information and consequently, express heritable traits. The plastid genome (plastome) occurs at high copy numbers, with up to thousands of genome copies being present in a single cell. Although mapping as a single circular molecule, the plastid DNA shows great structural dynamics. Multiple copies of the plastome are packed together in large nucleoprotein bodies, referred to as plastid nucleoids. The plastomes of land plants harbor a rather conserved set of approximately 100-120 genes in a genome of 120-160 kilobase pairs (kb). In contrast, size and coding capacity of plastomes in algae are much more variable. In most plant species, plastids and their genetic information are inherited maternally and thus excluded from sexual recombination. The cytological mechanisms leading to uniparentally maternal inheritance are surprisingly diverse and can involve organelle exclusion by unequal cell division, plastid destruction or selective degradation of the plastid DNA from the paternal parent. Exceptions from maternal inheritance, i.e., biparental or paternal plastid transmission, have arisen multiple times during evolution.

1 Introduction

Already at the beginning of the last century, the German geneticist and plant breeder Erwin Baur proposed that the non-Mendelian inheritance of leaf variegations can be explained with the assumption that chloroplasts (plastids) contain their own genetic material (Baur 1909, 1910; reviewed in Hagemann 2000, 2002). More than half a century later, the discovery of plastid DNA (Chun et al. 1963; Sager and Ishida 1963; Tewari and Wildman 1966) ultimately confirmed Baur's ingenious hypothesis. During the following decades, the plastid genome (plastome), its coding capacity and gene expression have been the subject of extensive molecular studies and today, the chloroplast represents the by far best-studied genetic compartment of the plant cell.

Due to its cyanobacterial ancestry, the plastome has retained numerous prokaryotic features, including a bacterial-type circular genome structure, genome packaging in nucleoids, organization of genes in operons, and a prokaryotic gene

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0223 / Published online: 14 April 2007

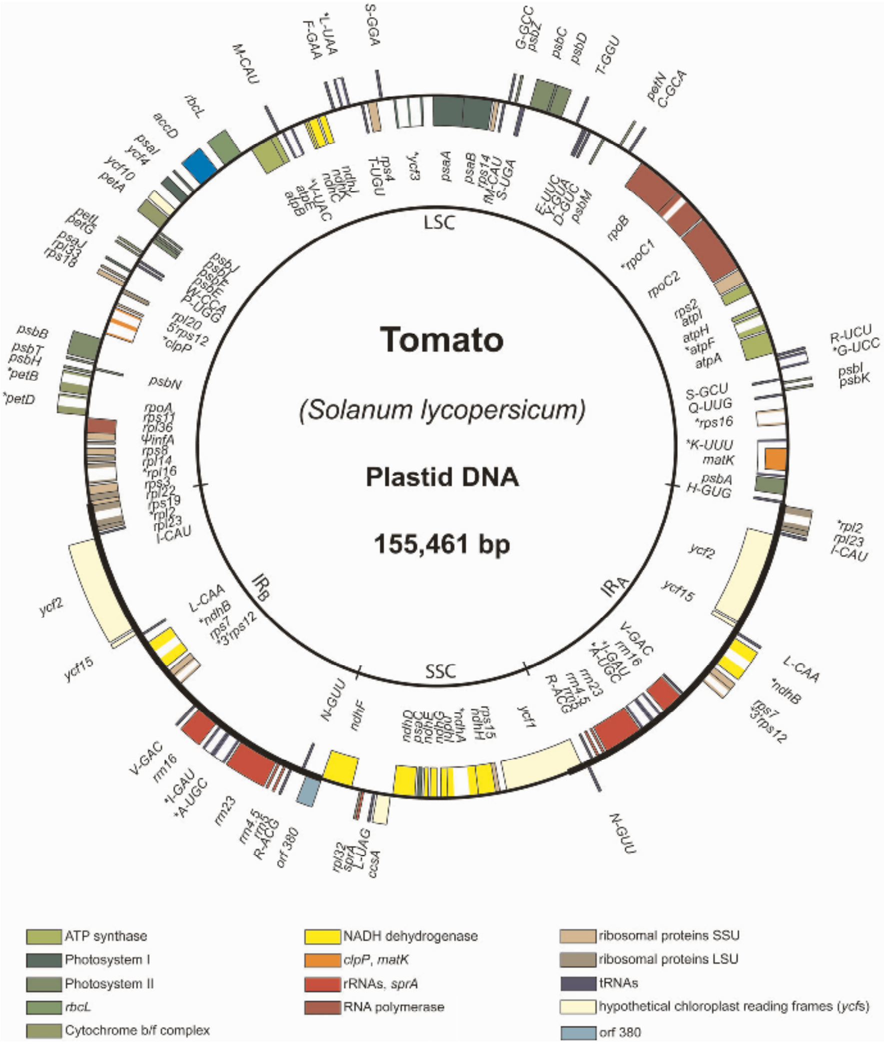


Fig. 1. Physical map of the tomato (*Solanum lycopersicum*, formerly *Lycopersicon esculentum*) plastid genome as a typical example of a plastid genome in higher plants (modified from Kahlau et al. 2006). Genes inside the circle are transcribed clockwise; genes outside the circle are transcribed counterclockwise. The two large inverted repeat regions IR_A and IR_B are shown as fat lines. Asterisks indicate intron-containing genes; introns are depicted as open boxes. For gene products and their functions, compare Table 1.

expression machinery. This chapter provides an overview of our current understanding of (i) the structural properties of the plastid DNA, (ii) structure and function of the plastome and (iii) the inheritance of plastids and their genomes.

Table 1. Plastid-encoded genes and conserved open reading frames (*ycf* = hypothetical chloroplast reading frame) in higher plants

Gene	Gene product	Functions and remarks
psaA	A subunit of PSI	reaction center subunit, essential for PSI function
psaB	B subunit of PSI	reaction center subunit, essential for PSI function
psaC	C subunit of PSI	essential cofactor-binding subunit
psaI	I subunit of PSI	small subunit, not essential for PSI function
psaJ	J subunit of PSI	small subunit, not essential for PSI function
ycf3	Ycf3 protein	essential PSI assembly factor, contains three tetratricopeptide (TPR) repeats
ycf4	Ycf4 protein	essential PSI assembly factor
psbA	D1 protein of PSII	reaction center, also termed 'herbicide-binding protein', essential for PSII function
psbB	CP47 subunit of PSII	inner antenna protein, essential for PSII function
psbC	CP43 subunit of PSII	inner antenna protein, essential for PSII function
psbD	D2 protein of PSII	reaction center, essential for PSII function
psbE	α-subunit of cytochrome b_{559}	essential for PSII assembly/stability/function, protection of PSII against photoinhibition, dark oxidation of plastoquinol
psbF	β-subunit of cytochrome b_{559}	essential for PSII assembly/stability/function, protection of PSII against photoinhibition, dark oxidation of plastoquinol
psbH	H subunit of PSII	small subunit associated with CP47, involved in PSII assembly, stabilization and photoprotection
psbI	I subunit of PSII	small subunit, involved in stabilization of PSII dimers and PSII-LHCII supercomplexes
psbJ	J subunit of PSII	small subunit, involved in assembly of the water-splitting complex and intra-complex electron transfer
psbK	K subunit of PSII	small subunit associated with CP43, presumably involved in PSII assembly/stability
psbL	L subunit of PSII	small subunit, involved in PSII dimerization and PSII-LHCII supercomplex formation, required for assembly of the water-splitting complex
psbM	M subunit of PSII	small subunit, function unknown
psbN	N subunit of PSII	function unknown, assignment as PSII subunit uncertain
psbT	T subunit of PSII	small subunit, involved in repair of photodamaged PSII reaction centers
psbZ	Z subunit of PSII	small subunit, couples the light-harvesting complex protein CP26 to PSII
petA	cytochrome f	core subunit of cyt b_6f complex, essential for cyt b_6f function
petB	cytochrome b_6	core subunit of cyt b_6f complex, essential for cyt b_6f function
petD	subunit IV of cyt b_6f	essential for cyt b_6f function

Table 1 continued

Gene	Gene product	Functions and remarks
petG	G subunit of cyt b_6f	small subunit, essential for cyt b_6f assembly/stability in *Chlamydomonas*
petL	L subunit of cyt b_6f	small subunit, not essential for cyt b_6f function, involved in complex stabilization
petN	N subunit of cyt b_6f	small subunit, essential for cyt b_6f assembly/stability
atpA	ATP synthase α-subunit	CF_1, nucleotide-binding site
atpB	ATP synthase β-subunit	CF_1, catalytic site
atpE	ATP synthase ε-subunit	CF_1, regulation of CF_1CF_0 activation, required for proton gating
atpF	ATP synthase b-subunit	CF_0, binding of CF_1
atpH	ATP synthase c-subunit	CF_0, proton translocation
atpI	ATP synthase a-subunit	CF_0, proton translocation
ndhA	A subunit of NAD(P)H dehydrogenase	chlororespiration, cyclic electron transfer, absent from gymnosperm plastomes
ndhB	B subunit of NAD(P)H dehydrogenase	chlororespiration, cyclic electron transfer, absent from gymnosperm plastomes
ndhC	C subunit of NAD(P)H dehydrogenase	chlororespiration, cyclic electron transfer, absent from gymnosperm plastomes
ndhD	D subunit of NAD(P)H dehydrogenase	chlororespiration, cyclic electron transfer, absent from gymnosperm plastomes
ndhE	E subunit of NAD(P)H dehydrogenase	chlororespiration, cyclic electron transfer, absent from gymnosperm plastomes
ndhF	F subunit of NAD(P)H dehydrogenase	chlororespiration, cyclic electron transfer, absent from gymnosperm plastomes
ndhG	G subunit of NAD(P)H dehydrogenase	chlororespiration, cyclic electron transfer, absent from gymnosperm plastomes
ndhH	H subunit of NAD(P)H dehydrogenase	chlororespiration, cyclic electron transfer, absent from gymnosperm plastomes
ndhI	I subunit of NAD(P)H dehydrogenase	chlororespiration, cyclic electron transfer, absent from gymnosperm plastomes
ndhJ	J subunit of NAD(P)H dehydrogenase	chlororespiration, cyclic electron transfer, absent from gymnosperm plastomes
ndhK	K subunit of NAD(P)H dehydrogenase	chlororespiration, cyclic electron transfer, absent from gymnosperm plastomes
rbcL	Rubisco large subunit	CO_2 fixation
rpoA	RNA polymerase α-subunit	transcription, *E. coli*-like plastid RNA polymerase (PEP)
rpoB	RNA polymerase β-subunit	transcription, *E. coli*-like plastid RNA polymerase (PEP)
rpoC1	RNA polymerase β'-subunit	transcription, *E. coli*-like plastid RNA polymerase (PEP)
rpoC2	RNA polymerase β''-subunit	transcription, *E. coli*-like plastid RNA polymerase (PEP)
matK	intron maturase	splicing factor for group II introns
rrn16	16S ribosomal RNA	translation, small ribosomal subunit

Table 1 continued

Gene	Gene product	Functions and remarks
rrn23	23S ribosomal RNA	translation, large ribosomal subunit
rrn5	23S ribosomal RNA	translation, large ribosomal subunit
rrn4.5	23S ribosomal RNA	translation, large ribosomal subunit
trnA-UGC	tRNA-Alanine(UGC)	translation
trnC-GCA	tRNA-Cysteine(GCA)	translation
trnD-GUC	tRNA-Aspartate(GUC)	translation
trnE-UUC	tRNA-Glutamate(UUC)	translation, tetrapyrrole biosynthesis
trnF-GAA	tRNA-Phenylalanine(GAA)	translation
trnG-GCC	tRNA-Glycine(GCC)	translation
trnG-UCC	tRNA-Glycine(UCC)	translation
trnH-GUG	tRNA-Histidine(GUG)	translation
trnI-CAU	tRNA-Isoleucine(CAU)	translation
trnI-GAU	tRNA-Isoleucine(GAU)	translation
trnK-UUU	tRNA-Lysine(UUU)	translation
trnL-CAA	tRNA-Leucine(CAA)	translation
trnL-UAA	tRNA-Leucine(UAA)	translation
trnL-UAG	tRNA-Leucine(UAG)	translation
trnM-CAU	tRNA-Methionine(CAU)	translation
trnfM-CAU	tRNA-N-Formyl-methionine(CAU)	translation initiation
trnN-GUU	tRNA-Asparagine(GUU)	translation
trnP-UGG	tRNA-Proline(UGG)	translation
trnQ-UUG	tRNA-Glutamine(UUG)	translation
trnR-ACG	tRNA-Arginine(ACG)	translation
trnR-UCU	tRNA-Arginine(UCU)	translation
trnS-GCU	tRNA-Serine(GCU)	translation
trnS-GGA	tRNA-Serine(GGA)	translation
trnS-UGA	tRNA-Serine(UGA)	translation
trnT-GGU	tRNA-Threonine(GGU)	translation
trnT-UGU	tRNA-Threonine(UGU)	translation
trnV-GAC	tRNA-Valine(GAC)	translation
trnV-UAC	tRNA-Valine(UAC)	translation
trnW-CCA	tRNA-Tryptophan(CCA)	translation
trnY-GUA	tRNA-Tyrosine(GUA)	translation
rps2	ribosomal protein S2	translation, small ribosomal subunit
rps3	ribosomal protein S3	translation, small ribosomal subunit

Table 1 continued

Gene	Gene product	Functions and remarks
rps4	ribosomal protein S4	translation, small ribosomal subunit
rps7	ribosomal protein S7	translation, small ribosomal subunit
rps8	ribosomal protein S8	translation, small ribosomal subunit
rps11	ribosomal protein S11	translation, small ribosomal subunit
rps12	ribosomal protein S12	translation, small ribosomal subunit
rps14	ribosomal protein S14	translation, small ribosomal subunit
rps15	ribosomal protein S15	translation, small ribosomal subunit
rps16	ribosomal protein S16	translation, small ribosomal subunit
rps18	ribosomal protein S18	translation, small ribosomal subunit
rps19	ribosomal protein S19	translation, small ribosomal subunit
rpl2	ribosomal protein L2	translation, large ribosomal subunit
rpl14	ribosomal protein L14	translation, large ribosomal subunit
rpl16	ribosomal protein L16	translation, large ribosomal subunit
rpl20	ribosomal protein L20	translation, large ribosomal subunit
rpl22	ribosomal protein L22	translation, large ribosomal subunit
rpl23	ribosomal protein L23	translation, large ribosomal subunit, inactive pseudogene in *Caryophyllidae*
rpl32	ribosomal protein L32	translation, large ribosomal subunit
rpl33	ribosomal protein L33	translation, large ribosomal subunit
rpl36	ribosomal protein L36	translation, large ribosomal subunit
infA	translation initiation factor 1	translation, inactive pseudogene or gene lost (and transferred to the nucleus) in several lineages
clpP	catalytic subunit of the protease Clp	ATP-dependent protein degradation, essential for cell survival
accD	acetyl-CoA carboxylase subunit	fatty acid biosynthesis, essential for cell survival
ycf5 / ccsA	subunit A of the system II complex for c-type cytochrome biogenesis	required for heme attachment to chloroplast c-type cytochromes
ycf10	inner envelope protein	presumably involved in the uptake of inorganic carbon
ycf1	putative Ycf1 protein	essential gene, function unknown
ycf2	putative Ycf2 protein	essential gene, function unknown, contains a putative nucleotide-binding domain
ycf15	unknown	ORF with unclear functional significance
sprA	small RNA	function unknown

2 Physical properties of plastid genomes

The plastid genome maps as a circular molecule of double-stranded DNA (ptDNA). In land plants, the genome size is typically in the range of 120-160 kb (Fig. 1), although some exceptions have been noted (see 3.1 and 3.2.4). Identical copies of this genome are present in all plastid types: the undifferentiated proplastids of meristematic tissues, the green chloroplasts in photosynthesis-performing cells, the colored chromoplasts of flowers and fruits and other plastid types specialized in storage of starch, proteins, or lipids.

Chloroplast DNA can be extracted from isolated organelles (which are purified by gradient centrifugation; Jansen et al. 2005) and was found to have physical properties distinct from nuclear DNA. The distinguishing features include different buoyant density in CsCl gradients, different melting and renaturation behavior, different GC content and the absence of 5-methylcytosine from plastid DNA (Tewari and Wildman 1966). In spite of its small genome size, plastid DNA can make up a significant fraction of the total cellular DNA which is due to its presence in high copy numbers. For tobacco leaves, it was estimated that about 9% of the total DNA is chloroplast DNA representing about 4.7×10^{-15} g DNA per chloroplast (Tewari and Wildman 1966).

2.1 Copy number of plastid genomes

A single plant cell contains many plastids and each plastid contains numerous (identical) plastome copies. Thus, in contrast to the two copies of each gene in the nucleus of a diploid plant, the cell is highly polyploid for its plastid genome. Depending on species, tissue, developmental stage and environmental conditions, the ploidy level can easily reach more than 10,000 identical copies of the plastid genome per cell (Bendich 1987). In land plants, plastome copy numbers are usually highest in photosynthetically active cells, where plastids are present as green chloroplasts. In contrast, non-green plastid types often possess fewer plastomes. The copy number in root plastids, for example, is only about one fifth of that in chloroplasts (Aguettaz et al. 1987; Isono et al. 1997). Likewise, chloroplast development from proplastids and etioplasts is associated with an increase in plastome copy number (from about 2000 to more than 8000 copies per cell in barley; Baumgartner et al. 1988). Changes in plastid genome copy numbers per cell during plastid differentiation and plant development most likely come from the combined action of two processes: changes in organelle number per cell and changes in the plastome copy number per plastid. For example, the copy number per plastids almost doubles during etioplast to chloroplast differentiation in barley leaf development (Baumgartner et al. 1988). Once plastid differentiation is completed, plastome copy numbers remain remarkably constant and do not vary significantly with leaf age or the plant's developmental stage (Li et al. 2006; Zoschke et al. 2007).

As plastids are asexual genetic systems and, in most species, excluded from sexual recombination (see 4), an intriguing question has been how plastid genomes can avoid evolutionary deterioration. Asexual reproduction is believed to be detrimental because of the accumulation of deleterious mutations over time, a hypothesis known as Muller's ratchet (Muller 1964). Since the vast majority of mutations are deleterious, an asexual genetic system is expected to suffer a continuous decline in fitness. Surprisingly, in spite of their asexual mode of reproduction, plastid genomes even have considerably lower mutation rates than nuclear genomes (Wolfe et al. 1987). A recent study has provided experimental evidence that it is the plastid's high degree of polyploidy which, together with a very active mutation-correcting activity by gene conversion, counteracts the detrimental effects of Muller's ratchet and keeps mutation rates in plastid genomes very low (Khakhlova and Bock 2006). These findings suggest a molecular link between asexual reproduction, high genome copy numbers, and low mutation rates.

2.2 Organization of plastid genomes in nucleoids

The plastid genomes do not swim around as naked DNA in the plasmatic compartment (stroma) of the organelle. Instead, several copies of the plastome are densely packed together in large nucleoprotein bodies called plastid nucleoids (Kuroiwa 1989, 1991). Nucleoids can be visualized by fluorescence microscopy after staining of cells or tissues with the DNA-intercalating fluorochrome DAPI (4',6-diamidino-2-phenylindole; Kuroiwa 1991). Number, shape, and size of the nucleoids as well as their distribution in the chloroplast vary depending on the species. In algae and higher plants, five different subtypes of nucleoid morphology have been described, ranging from spherical to ring-like structures (Kuroiwa 1989). Likewise, plastome copy numbers per nucleoid are variable between species and in dependence on plastid differentiation. Proplastids, for example, often contain only a single nucleoid, whereas mature chloroplasts can easily contain several or even dozens of nucleoids.

The nucleoid, and probably each individual plastid genome, is membrane bound. In higher plants, evidence has been provided for both an association with thylakoid membranes and an anchoring to the inner envelope of the chloroplast (Liu and Rose 1992; Sato et al. 1993). Isolated nucleoids retain transcriptional activity *in vitro* (Sakai et al. 1991) suggesting that the transcriptional apparatus (RNA polymerases and sigma factors) is tightly associated with the plastid genome (Krause and Krupinska 2000). Little is known about the molecular processes and mechanisms that organize plastid nucleoids. Notably, it is entirely unclear, how a defined number of genome copies are packed into one nucleoid and how higher-order structures of the plastid genome are built and regulated (Salvador et al. 1998). It is known, however, that nucleoid size and number per cell are controlled by nuclear genes. In the unicellular green alga *Chlamydomonas reinhardtii*, mutants were obtained that had increased or drastically decreased numbers of nucleoids (Ikehara et al. 1996; Misumi et al. 1999). While wild type cells contain on average seven nucleoids, mutants with increased nucleoid number had 14-23 nucleoids. At the other extreme, a mutant called *moc* (for 'monokaryotic chloroplast') had only a single huge nucleoid per chloroplast (Misumi et al. 1999). However, the genes and mutations responsible for these interesting nucleoid phenotypes have not been identified to date.

Recently, several DNA-binding nucleoid proteins have been identified and/or biochemically characterized to some extent (Nakano et al. 1997; Sekine et al. 2002; Jeong et al. 2003; Cannon et al. 1999). The arguably best-studied nucleoid constituent is Hlp (also called HU), a histone-like DNA-binding protein that, similar to Hlp homologs in eubacteria, is believed to serve as a general architectural nucleoid protein (Kobayashi et al. 2002). Recently, a first step towards determining the proteome of chloroplast nucleoids has been taken in *Arabidopsis* (Pfalz et al. 2006). Nucleoid preparations obtained by a two-step chromatographic purification were subjected to mass spectrometric protein identification. Although the preparations were not absolutely pure, several good candidates for genuine nucleoid proteins could be identified, including RNA polymerase subunits, topoisomerases and DNA polymerase subunits (along with a number of novel proteins of unknown function; Pfalz et al. 2006). Interestingly, Hlp, the most abundant and major architectural nucleoid protein in bacteria and algae (Kobayashi et al. 2002) was not identified, possibly suggesting that nucleoid organization in higher plants is fundamentally different from that in eubacteria and algae. The systematic identification of nucleoid proteins and their functional characterization using the power of *Arabidopsis* nuclear genetics should pave the way to a better understanding of the higher-order structure of plastid DNA, its dynamics and impact on the regulation of plastid gene expression.

2.3 Structural conformations of plastid genomes

The finding that plastid genomes map as circular molecules and the identification of circular ptDNA molecules by electron microscopy (the contour length of which corresponded to the determined size of the plastome) in many independent studies led to the long-held belief that the structural conformation of the plastome *in vivo* is a simple circle of double-stranded DNA. However, more recent investigations have revealed that the ptDNA displays a surprisingly great structural plasticity with only a minority of the genome molecules being circular (25 to 45% in developing leaf tissue; Lilly et al. 2001). In addition to circles, both electron microscopic investigations and pulsed-field gel electrophoretic analyses have identified various linear genome conformations, including plastome multimers (resembling concatemers as arising during rolling-circle replication of bacteriophage genomes) and branched multimers (Bendich and Smith 1990; Lilly et al. 2001; Oldenburg and Bendich 2004). Linear ptDNA molecules are unlikely to originate just from randomly broken circles: they were demonstrated to possess defined ends, some of which correspond to known origins of DNA replication (Oldenburg and Bendich 2004; Scharff and Koop 2006).

Interestingly, the fraction of genome molecules that is circular also shows a variety of different conformations. A substantial amount of the circles (>30% in tobacco; Lilly et al. 2001) is multimeric in that two or more copies of the genome form a single large circle (Deng et al. 1989). Multimers can come, for example, from rolling-circle replication and/or fusion of monomeric circles by homologous recombination. In addition to circles, lasso-like structures and suspected partially single-stranded molecules (showing D-loop-like bubbles) have been seen (Lilly et al. 2001). Similar to other circular genomes and episomes found in nature, the ptDNA can also adopt various supercoiled conformations and form catenanes (interlocking circles; Mukherjee et al. 1994; Kumar et al. 1995; Ahlert et al. 2003; Cho et al. 2004; Bendich 2004). Another structural peculiarity of the plastid genome is its presence in two isoforms due to flip-flop recombination of the two inverted repeat regions (3.1).

The functional relevance of most of the many different conformations of the ptDNA is still unclear. Some of them may simply represent replication intermediates; others may lack any functional significance. An interesting exception may be the degree of ptDNA supercoiling: studies in the unicellular green alga *Chlamydomonas reinhardtii* have revealed that DNA topology fluctuates in dependence on the diurnal rhythm and that these fluctuations correlate with changes in the transcriptional activity (Salvador 1998). This finding may suggest that conformational changes of the ptDNA are involved in the regulation of plastid gene expression.

3 Fine structure of plastid genomes

The complete sequencing of two plastid genomes more than twenty years ago (Ohyama et al. 1986; Shinozaki et al. 1986) marks a milestone in structural genomics and has had a profound influence on our understanding of the genetics and molecular biology of plastids. In the following years, dozens of additional plastid genomes have been sequenced. The 88 plastomes fully sequenced by the end of 2006 (http://www.ncbi.nlm.nih.gov/genomes/static/euk_o.html) represent all major lineages of plant evolution. The picture that has emerged from these studies is that the plastome of land plants is a conservative genome while considerable variation in genome organization and coding capacity exists in algae (see 3.2.5).

In general, plastid genomes have a low GC content which is typically in the range of 30-40% (Ohyama et al. 1988; Shimada and Sugiura 1991). The low GC content is particularly pronounced in non-coding intergenic spacer regions where AT richness is often extreme and can reach values above 80% AT (Ohyama et al. 1988). Within coding regions, AT richness manifests as strong bias in codon usage, in that synonymous codons with an A or T in third codon position are strongly preferred over those with G or C in third position (Shimada and Sugiura 1991).

3.1 Inverted repeats and single-copy regions

In land plants, most plastomes display a tetrapartite genome organization with a large single copy region (LSC) and a small single copy region (SSC) separating tow inverted repeat regions (IR_A and IR_B; Fig. 1). The two IRs are identical in their nucleotide sequence, so that every gene contained within them is present in two copies per genome which only differ in their relative orientation (Fig. 1). The boarders between the IRs and the single copy regions are somewhat variable even between closely related species (Goulding et al. 1996). Expansion of the inverted repeat region is extreme in *Pelargonium*, the flowering plant species with the largest plastome (217 kb; Palmer et al. 1987; Chumley et al. 2006). Here, the IRs are 75 kb in size each and thus about three times as big as in most other higher plants. The functional significance of the presence of the IR region in two copies is not quite clear. Increasing the gene dosage of highly expressed genes (such as the ribosomal RNA genes; Fig. 1) and genome stabilization (Palmer and Thompson 1982) have been proposed as possible reasons why having this large inverted duplication could be beneficial. Its absence from some algal (Reith 1995) and even some higher plant plastomes (Palmer and Thompson 1982), however, indicates that the IR is not essential for plastome maintenance and/or function.

The presence of two large identical regions in the plastome facilitates two types of genetic interactions between homologous sequences: intramolecular recombination and gene conversion (Birky and Walsh 1992; Khakhlova and Bock 2006). Homologous recombination between the two IRs produces two isoforms of the plastid genome (dubbed flip-flop recombination; Palmer 1983; Stein et al. 1986), which differ in the relative orientations of LSC and SSC. Circumstantial evidence for the action of gene conversion in the IRs has come from the observation that the mutation frequency of genes in the IR regions is significantly lower than for genes located in the two single copy regions of the plastome (Wolfe et al. 1987; Maier et al. 1995). Gene conversion biased on average towards the wild type sequence has been proposed to account for the lower mutation rate in the inverted repeats (Birky and Walsh 1992). The recent experimental demonstration of high gene conversion activity in plastids (Khakhlova and Bock 2006) lends support to this hypothesis.

3.2 Information content of plastid genomes

Among the three genomes of the plant cell, the plastome is the most gene-dense one with more than 100 genes in a genome of typically only 120 to 160 kb (Sugiura 1989, 1992; Wakasugi et al. 2001; Fig.1; Table 1). The plastid genome is the evolutionary remnant of a cyanobacterial genome. After endosymbiosis, the genome has undergone a dramatic size reduction and, thus, contemporary plastomes contain only a small proportion of the genes of their free-living cyanobacterial ancestors. Whereas the genome of the cyanobacterium *Synechocystis* contains more than 3000 genes (Kaneko et al. 1996; Kaneko and Tabata 1997), the plastid genomes of land plants harbor only approximately 115 genes.

Very obviously, the limited coding capacity of the plastome is by far insufficient to provide the thousands of components required to support its own gene expression system, photosynthesis and all the many other plastid-localized metabolic functions. Therefore, all cellular functions fulfilled by present-day plastids are strictly dependent upon the products of nuclear genes that are synthesized on cytoplasmic ribosomes and post-translationally imported into the organelle. Nuclear-encoded proteins make up the by far largest fraction of the plastid proteome (Abdallah et al. 2000; Rujan and Martin 2001; Martin et al. 2002; Hippler and Bock 2004) and it is estimated that chloroplasts import more than 95 % of their proteins from the cytosol. Consequently, the spatial and temporal expression of nuclear and organellar genes must be tightly coordinated.

Plastid-encoded genes can be roughly classified into three major groups (Shimada and Sugiura 1991; Kahlau et al. 2006): genetic system genes, photosynthesis-related genes and other genes. The approximately 60 genetic system genes contained in land plant plastomes encode RNA and protein components of the plastid's gene expression machinery (Fig. 1; Table 1). Approximately 50 plastid genes encode protein products involved in photosynthesis (Fig. 1; Table 1). The heterogeneous third gene group comprises all other genes and conserved open reading frames of unknown function.

3.2.1 Photosynthesis genes

Chloroplasts are the site of photosynthesis, the conversion of solar energy to chemical energy. Photosynthesis consists of two stages, the light reactions and the dark reactions, both of which involve complex molecular machineries. A substantial number of plastome-encoded genes (47 genes in angiosperms; Table 1) is dedicated to the photosynthetic apparatus. These include fifteen genes for subunits of photosystem II (PSII), the membrane protein complex catalyzing the light-driven oxidation of water. The products of another seven genes are required for photosystem I (PSI) function, the membrane protein complex that catalyzes the light-driven transmembrane electron transfer from plastocyanin (or cytochrome c_6) to the ferredoxin-NADP complex. In addition to five genes for subunits of the PSI complex, the seven PSI-related genes also include *ycf3* and *ycf4*, two genes for proteins involved in PSI assembly (Ruf et al. 1997; Boudreau et al. 1997). Six plastid genes encode subunits of the cytochrome b_6f complex, the redox-coupling protein complex interconnecting the two photosystems. Another six genes encode subunits of the chloroplast ATP synthase, the enzyme that catalyzes the conversion of phosphate and adenosine diphosphate into adenosine triphosphate utilizing a proton gradient across the thylakoid membrane as energy source. Eleven genes on the plastome encode subunits of a chloroplast NAD(P)H dehydrogenase, a thylakoid protein complex suggested to be involved in chlororespiration and cyclic electron flow around PSI (Burrows et al. 1998; Shikanai et al. 1998; Joet et al. 2001; Munekage et al. 2004). This complex is non-essential for photosynthesis and all genes for its subunits were found to be absent from the fully sequenced plastid genomes of the gymnosperm *Pinus thunbergii* and the green alga *Chlamydomonas reinhardtii* (Wakasugi et al. 1994; Maul et al. 2002). Finally, two plastid-encoded gene products are directly or indirectly involved in the dark reactions: *rbcL* encoding the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) and *ycf10*, a conserved open reading frame encoding a chloroplast inner envelope membrane protein reportedly involved in inorganic carbon uptake (Sasaki et al. 1993a; Rolland et al. 1997).

None of the protein complexes involved in photosynthesis is composed entirely of plastome-encoded subunits. Instead, all of them require the products of nuclear genes which are of cyanobacterial origin and, during evolution, have been transferred from the plastid to the nuclear genome. The two-subunit enzyme Rubisco provides the classical example for this intimate plastid-nuclear cooperation. In all flowering plants, the large subunit of Rubisco is encoded in the plastome whereas the small subunit is encoded in the nuclear genome, typically by a small gene family.

3.2.2 Genetic system genes

The genetic system genes comprise the largest group of genes located on higher plant plastomes (62 genes; Table 1). To this group belong all genes whose products are involved in plastid gene expression (i.e. transcription, RNA processing, translation, protein degradation): 30 tRNA genes, four rRNA genes, 21 genes for ribosomal proteins (nine proteins of the large subunit and twelve proteins of the small subunit of the plastid 70S ribosome), four genes for subunits of the *E. coli*-like plastid RNA polymerase (PEP), *matK* suggested to encode an RNA maturase (i.e. a splicing factor involved in the removal of a subset of chloroplast group II introns; Hess et al. 1994; Liere and Link 1995; Mohr et al. 1993; Jenkins et al. 1997), *clpP* encoding a subunit of a chloroplast protease (Shanklin et al. 1995; Majeran et al. 2000) and *infA* encoding translation initiation factor IF-1 (Sijben-Müller et al. 1986).

A complete set of tRNAs for decoding all triplets in protein-coding genes is thought to comprise 32 tRNA species. Although only 30 tRNA genes are encoded in the plastome, they are nonetheless believed to be sufficient to read all codons (Jukes and Osawa 1990; Osawa et al. 1992). This is presumably achieved by an extended wobbling (referred to as 'four-way wobble') between the third codon position and the 5' nucleotide of the anticodon in the tRNA. In the case of the four alanine codons (GCU, GCC, GCA and GCG), for example, this means that the U in the first anticodon position of the single tRNA-Ala species (*trnA-UGC*; Table 1) can probably basepair with all four possible nucleotides in third codon position of the alanine triplets (Jukes and Osawa 1990; Osawa et al. 1992).

Remarkable differences between species exist concerning the essentiality of the plastid gene expression apparatus. Plastid translation has been demonstrated to be essential for cell survival in tobacco (Ahlert et al. 2003; Rogalski et al. 2006), but appears to be non-essential under heterotrophic culture conditions in at least some Brassicaceae species (Zubko and Day 1998, 2002) and probably also in some cereals (Hess et al. 1993, 1994).

While the RNA components of the gene expression machinery (rRNAs and tRNAs) are exclusively encoded in the plastid genome (Lung et al. 2006), many of the protein components are encoded by nuclear genes. For example, only about one third of the plastid ribosomal proteins is plastome encoded in higher plants, the other two thirds are nuclear-encoded, made in the cytosol and imported into the plastid. A similar division of labor between the nucleus and the plastid occurs in the coding of the transcriptional apparatus. The four core subunits of the *E. coli*-

like plastid RNA polymerase (plastid-encoded RNA polymerase, PEP) are encoded in the plastome, but the sigma factors, which are required for promoter recognition, are encoded in the nuclear genome. In addition, a second RNA-synthesizing activity in the plastid (nuclear-encoded RNA polymerase, NEP) provided by bacteriophage-type enzymes is encoded by nuclear genes (Hedtke et al. 1997; Hess and Börner 1999).

While in angiosperm plants, the set of genes encoded in the plastome is usually highly conserved between species, a small number of genetic system genes, including *rpl23* and *infA* (Table 1), provide notable exceptions in that they have been transferred to the nucleus or replaced by nuclear genes of non-plastid origin in some lineages of evolution (Bubunenko et al. 1994; Millen et al. 2001). The presence in the plastome of pseudogenic remnants of the genes suggests that these events occurred only relatively recently in evolution. The *infA* gene encoding the plastid translation initiation factor 1 provides a particularly striking example. It had long been known that *infA*, while being a functional gene in the plastome of the liverwort *Marchantia polymorpha* and the higher plant rice (Ohyama et al. 1986; Hiratsuka et al. 1989), exists only as a pseudogene in the tobacco ptDNA (Shinozaki et al. 1986; Shimada and Sugiura 1991). A systematic phylogenetic analysis of *infA* structure in the plastomes of angiosperms revealed that the gene has repeatedly become non-functional in approximately 24 separate lineages of angiosperm evolution. Search for nuclear *infA* copies in four of these lineages identified expressed nuclear *infA* genes whose gene products are targeted to plastids (Millen et al. 2001).

3.2.3 Other genes and conserved open reading frames

A small number of genes on the plastome of land plants are not directly involved in photosynthesis or gene expression. These include the *accD* gene, which encodes a subunit of acetyl-CoA carboxylase, the key enzyme in fatty acid biosynthesis (Sasaki et al. 1993b, 1995). A second example is *ccsA* (*ycf5*), the protein product of which is required for heme attachment to chloroplast c-type cytochromes (Orsat et al. 1992; Xie et al. 1998; Xie and Merchant 1996).

The plastome of pland plants also harbors a few conserved open reading frames (ORFs) of unknown function (Table 1). Interspecific conservation of an ORF is usually taken as good indication that it constitutes a genuine gene. In plastids, such conserved ORFs are referred to as *ycf* (hypothetical chloroplast reading frame). Although during recent years, the functions of most *ycf*s could be determined by reverse genetics in *Chlamydomonas reinhardtii* and tobacco (e.g. Monod et al. 1994; Ruf et al. 1997; Hager et al. 1999; reviewed in Rochaix 1997; Bock and Hippler 2002), there are a few left whose functions have not yet been elucidated. Among them are two giant open reading frame, *ycf1* and *ycf2*, which in tobacco, potentially encode protein products of 1901 and 2280 amino acids, respectively. Attempts to inactivate *ycf1* and *ycf2* in tobacco have revealed that both reading frames are essential genes (Drescher et al. 2000). This excludes a function of the gene products in photosynthesis (because photosynthesis is non-essential under *in vitro* culture conditions), but leaves the possibility of a function in plastid gene expression (which has been demonstrated to be essential in tobacco; Ahlert et al. 2003; Rogalski et al. 2006) or in some essential metabolic pathway.

All plastid genomes also harbor a number of non-conserved open reading frames, most of which are shorter than 150 codons. Lack of evolutionary conservation even among closely related species is usually interpreted as evidence for these ORFs fortuitously forming contiguous reading frames that have no functional significance (Kahlau et al. 2006).

3.2.4 Plastid genomes of parasitic plants

With the exception of only very few genes, the above-described set of plastome-encoded genes (Table 1) is highly conserved among angiosperm plant species. There is, however, one group of angiosperms whose plastid genomes differ radically in gene content: parasitic plants. A limited number of plant species grows heterotrophically by exploiting green plants as carbon source. Many of these holoparasites have lost the capacity to carry out photosynthesis and also lack photosynthetic pigments. The ability to obtain sugars from a host plant releases the selective pressure on the maintenance of photosynthesis-related genes. Consequently, the plastomes of such parasites suffer dramatic size reductions, mainly caused by the loss of photosynthesis genes or their degeneration to pseudogenes (dePamphilis and Palmer 1990; Wimpee et al. 1991; Bungard 2004). For example, the plastome of the root holoparasite *Epifagus virginiana* (an Orobanchaceae species) is less than half the size of that in photoautotrophic angiosperms (dePamphilis and Palmer 1990; Wolfe et al. 1992). It contains only 21 intact protein-coding genes, 18 of which belong to the genetic system genes and the remaining three falling into the category ‘other genes’ (*accD*, *ycf1*, and *ycf2*; see 3.2.3 and Table 1). Remarkably, also some genetic system genes have been lost or degraded to pseudogenes (Morden et al. 1991; Wolfe et al. 1992). It is currently unclear, whether or not these missing genes have been substituted by functional nuclear gene copies the protein products of which are imported into plastids. Nonetheless, plastid genes in *Epifagus* are actively transcribed and their mRNAs are faithfully processed by intron splicing and RNA editing suggesting that the vestigial plastid genome is indeed expressed (Ems et al. 1995). However, whether or not also the translational apparatus in these plastids has remained functional, is not yet clear.

Parasitism in seed plants has evolved several times independently (presumably at least ten times; Bungard 2004). Interestingly, not all parasitic plants grow exclusively heterotrophically. A number of parasitic species have retained at least some photosynthetic activity and thus, strictly speaking, grow mixotrophically. They fix a limited amount of carbon by themselves through photosynthesis, while obtaining the bulk of carbon as sugars from their host plant. Such species are believed to represent evolutionary intermediates that are in the process of losing their ability to photosynthesize. The genus Cuscuta (dodders) provides a particularly striking example for this evolutionary transition. Its more than 150 species vary greatly in their residual photosynthetic activities and also show great interspecific variation with respect to the extent of plastid genome degeneration by gene deletion or gene decay to inactive pseudogenes (Berg et al. 2004; Revill et al. 2005).

Thus, the analysis of ptDNA evolution in Cuscuta provides a unique opportunity to follow the molecular changes associated with the gradual transition to heterotrophy and to study the mechanisms of plastid genome streamlining as triggered by the loss of photosynthesis.

3.2.5 Plastid genomes of algae

While structure and coding capacity of the plastid genome are highly conserved among land plants, the plastome has experienced many architectural changes during the evolution of algae (Simpson and Stern 2002). The most dramatic change appears to have occurred in some dinoflagellates, where single genes are contained on small (2-3 kb) minicircles and the sum of the minicircles may make up the plastid genome (Zhang et al. 1999; Barbrook and Howe 2000). This unusual multipartite plastid genome structure seems to be confined to dinoflagellates: the genomes of most other algae map as circular molecules of 100-300 kb (Reith 1995; Simpson and Stern 2002).

The inverted repeat region present in most higher plant plastomes and containing the ribosomal RNA (rRNA) operon (3.1; Fig.1) has undergone many structural changes in algae. While, for example, many green and red algae have rRNA operon-containing IRs, some green and red algae have lost one of the IRs and, thus, possess only a single copy of the rRNA operon. Again other green and red algal species have direct repeats rather than inverted repeats (Reith 1995). The perhaps most unusual structure of the rRNA operon is found in *Euglena gracilis*, a unicellular flagellate-like protist with a green algal plastid acquired via secondary endosymbiosis. In *Euglena*, multiple copies of the rRNA operon are arranged as a tandem array of three complete and one partial operons (Hallick et al. 1993).

The plastomes of algae also vary greatly with respect to gene density and information content. While in some algal lineages, plastid genomes are extremely compact and gene-dense (such as the plastome of the cryptophyte alga *Guillardia theta* with 180 genes in a genome of only 122 kb; Douglas and Penny 1999), genome expansion by accumulation of non-coding DNA has occurred in other algae. The model green alga *Chlamydomonas reinhardtii* provides a particularly striking example for such a genome expansion: its plastome is 203 kb large, but contains only 99 genes. The presence of repetitive DNA (i.e. short dispersed repeats) in intergenic regions accounts for more than 20% of the genome size (Maul et al. 2002). The selective forces that have led to extreme genome streamlining in some algae, but genome expansion in others are currently totally unknown.

Green algae share a common ancestry with land plants and it is therefore unsurprising that the gene content of their plastomes is similar to that of higher plants. Exceptions include, for example, *petN* which is nuclear encoded in green algae (Hager et al. 1999; Table 1), and two genes for proteins involved in plastid division which are present in the plastome of the green alga *Chlorella vulgaris*, but absent from the ptDNA of higher plants and another green alga, *Chalmydomonas reinhardtii* (Wakasugi et al. 1997). In contrast to green algae, the plastomes of non-green algae have retained many genes that are absent from the plastid genomes of land plants (Reith and Munholland 1993, 1995; Kowallik et al. 1995;

Reith 1995; Ohta et al. 2003). It is generally assumed that these additional genes were transferred to the nucleus in the common ancestor of land plants and green algae which, in this respect, represent a more derived state of plastome evolution. Among the algal plastomes fully sequenced to date, the red alga *Porphyra purpurea* has the highest number of genes (250 genes in a genome of 191 kb, Reith and Munholland 1995). Expansion of the two main gene classes in plastid genomes (photosynthesis genes and genetic system genes; see 3.2.1 and 3.2.2) as well as presence on the plastome of additional groups of genes (e.g. for amino acid, fatty acid, pigment and vitamin biosyntheses) account for this enlarged gene repertoire.

4 Inheritance of plastid genomes

DNA-containing organelles (plastids and mitochondria) are inherited in a non-Mendelian fashion in all eukaryotes. In most organisms, organellar genomes are inherited from only one parent, with maternal inheritance being much more widespread than paternal inheritance. Plastids and their DNA can be inherited maternally, paternally or biparentally (reviewed, e.g., in Mogensen 1996; Birky Jr 1995; Hagemann 2002). At least in higher plants, plastid genomes do not normally undergo sexual recombination, even when they are inherited biparentally. This means that, except in very rare cases (which may be considered accidents; Medgyesy et al. 1985; Thanh and Medgyesy 1989; Baldev et al. 1998), chloroplast fusion and genetic recombination do not occur (an exception is the green alga *Chlamydomonas reinhardtii*; see 4.1.4).

Uniparentally maternal plastid inheritance has long been considered the rule, although the first exception (*Pelargonium* as a species with biparental inheritance; Baur 1909) was discovered simultaneously with the rule (maternal inheritance of plastids in several angiosperm species; Correns 1909; Baur 1910). Although it is still true that the majority of flowering plants transmit their plastids uniparentally from the female parent to the progeny, exceptions are found in nearly all major lineages of plant evolution (Mogensen 1996; Birky Jr 1995). This suggests that maternal inheritance, as the presumably ancestral mode of plastid transmission has been broken many times independently in plant evolution (Birky Jr 1995).

While the different modes of plastid inheritance (maternal, paternal, biparental) are cytologically reasonably well described (see below), the characterization of the molecular mechanisms underlying plastid inheritance is still in its infancy. While in *Chlamydomonas*, a model alga in which plastid inheritance is genetically tractable (but mechanistically very different from higher plants; see 4.1.4), some of the molecular components involved in maternal inheritance have been identified, next to nothing is known about the factors involved in the various modes of plastid inheritance existing in flowering plants. Plastid transmission is very likely controlled by nuclear genes (Tilney-Bassett 1984, 1994), but to date, not a single gene involved in plastid inheritance has been identified in any higher plant.

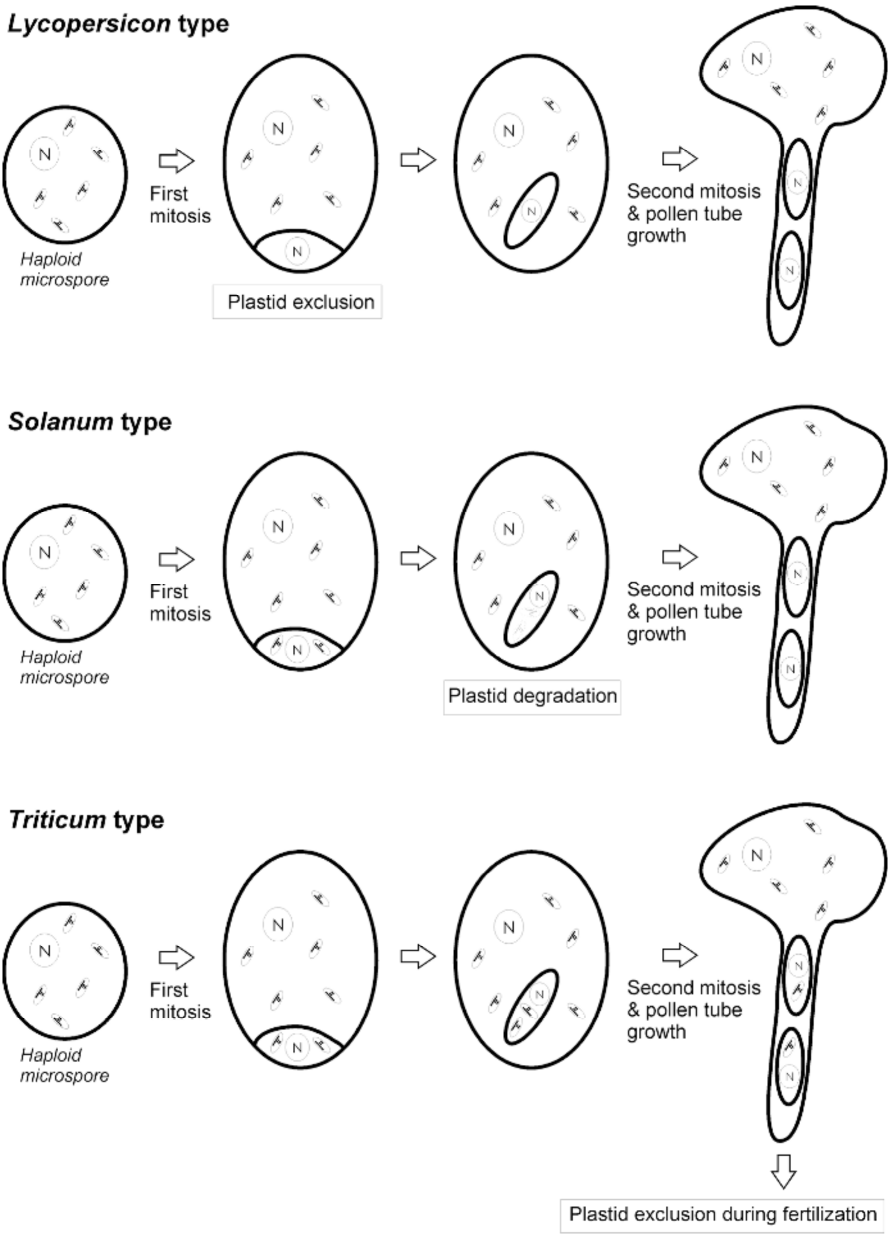


Fig. 2. Cytological mechanisms leading to maternal plastid inheritance in higher plants. See text for details and Table 2 for example species.

Table 2. Modes and mechanisms of plastid inheritance.

Mode of inheritance	Cytological mechanism	Examples
Maternal (Lycopersicon type)	Plastid exclusion from the generative cell during the first haploid pollen mitosis	*Beta vulgaris, Brassica campestris, Gossypium hirsutum, Solanum lycopersicum, Nicotiana alata, Petunia hybrida, Prunus avium, Spinacia oleracea*
Maternal (Solanum type)	Plastid degradation in the generative cell	*Convallaria majalis, Epilobium* spec.*, Fritillaria imperialis, Hosta japonica, Solanum tuberosum*
Maternal (Triticum type)	Plastid exclusion during fertilization	*Hordeum vulgare, Pisum sativum, Triticum aestivum, Zea mays*
Maternal (Chlamydomonas type)	Selective degradation of the paternal ptDNA after syngamy, fusion of maternal and paternal plastids	*Chlamydomonas reinhardtii*
Biparental (Pelargonium type)	Presence of plastids in sperm cells, transmission into the zygote	*Medicago sativa, Oenothera* spec.*, Pelargonium* spec.*, Rhododendron* spec.*, Hypericum* spec.
Paternal	Presence of plastids in sperm cells and transmission into the zygote, disintegration of maternal plastids in the egg cell	*Abies alba, Actinidia deliciosa, Larix decidua, Pinus taeda, Pseudotsuga menziesii, Sequoia sempervirens*

4.1 Maternal inheritance

The vast majority of angiosperms and at least some gymnosperms display a maternal mode of plastid inheritance and thus do not regularly transmit plastids and plastid genes through pollen. Cytological investigations have revealed that there is not a unique mechanism how maternal inheritance of plastids is brought about. Instead, different species can utilize very different mechanisms of eliminating paternal plastids and/or paternal plastid genomes. The correlation of the cytological mechanisms leading to maternal inheritance with plant phylogeny is rather poor, and therefore, the mechanism operating in a given species is hardly predictable. Similarly to the mode of plastid inheritance, the cytological mechanism of maternal inheritance must be determined on a species-by-species basis.

According to the mechanism of paternal plastid elimination, at least four different subtypes of maternal inheritance can be distinguished (Hagemann and Schröder 1989; Hagemann 2002). This classification is largely based on electron microscopic investigations of plastid fate during male gametophyte development.

The subtypes are named after the first species discovered to realize the respective cytological mechanism.

4.1.1 Maternal inheritance: Lycopersicon type

In angiosperms, development of the male gametophyte starts out with meiotic division of the microspore mother cell generating four haploid microspores. Subsequently, the haploid microspores undergo two mitotic divisions, referred to as pollen mitoses. The first pollen mitosis involves an asymmetric division of the haploid microspore resulting in a large vegetative cell and a small generative cell. The vegetative cell receives most of the cytoplasm from the microspore and completely envelopes the generative cell (Fig. 2). The generative cell then gives rise to two sperm cells by another mitotic division (second pollen mitosis). During fertilization, the two sperm cells move towards the ovule through the growing pollen tube. One of them fuses with the egg cell giving rise to the zygote, whereas the other fuses with the central cell to produce the precursor cell of the endosperm (double fertilization).

The Lycopersicon type of maternal plastid inheritance involves plastid exclusion during the first pollen mitosis. The extremely asymmetric division of the microspore results in a vegetative cell that contains all plastids and a generative cell that is free of plastids (Fig. 2). Consequently, also both sperm cells lack plastids. It is generally assumed that plastid inheritance in the majority of angiosperm species follows the exclusion mechanism of the Lycopersicon type (Hagemann and Schröder 1989; Hagemann 2002; Table 2).

4.1.2 Maternal inheritance: Solanum type

In a number of angiosperm species, maternal inheritance is brought about by degradation of paternal plastids. During male gametophyte development in species belonging to the Solanum type of maternal inheritance, plastids in the generative cells are selectively destructed whereas plastids in the vegetative cell remain intact (Fig. 2). Consequently, as in the Lycopersicon type, the two sperm cells carrying out the double fertilization are free of plastids.

The examples of species utilizing the plastid exclusion mechanism of the Lycopersicon type and those utilizing the plastid degradation mechanism of the Solanum type (Table 2) illustrate that even closely related species can differ in the cytological mechanisms conferring maternal plastid inheritance: tomato (*Solanum lycopersicum*, formerly *Lycopersicon esculentum*) and potato (*Solanum tuberosum*) belong to the same genus of Solanaceae, but yet exhibit different modes of maternal plastid inheritance.

4.1.3 Maternal inheritance: Triticum type

In pea and at least some monocotyledonous species, both the generative cell and the sperm cells regularly contain plastids. Nonetheless, these species do not normally transmit paternal plastids into the zygote. It is generally believed that, during fertilization of the egg cell by one of the two sperm cells, the plastids are stripped off together with most of the cytoplasm and do not enter the zygote along with the sperm cell's nucleus (Hagemann and Schröder 1989; Fig. 2; Table 2). Such a mechanism would be somewhat reminiscent of the exclusion of sperm cell mitochondria during fertilization of the egg in animals. However, 'smoking gun' evidence for a stripping-off mechanism underlying maternal inheritance in the Triticum type is largely lacking. This is mainly due to the difficulty to catch in the act sperm and egg by electron microscopy. Therefore, alternative mechanisms, such as degeneration of the cytoplasm surrounding the sperm cell nucleus (including the demise of plastids and mitochondria) shortly before the fertilization process, presently cannot be excluded.

4.1.4 Maternal inheritance: Chlamydomonas type

In no other plant, chloroplast inheritance has been as thoroughly studied as in the unicellular green alga *Chlamydomonas reinhardtii*. *Chlamydomonas* has a single large (cup-shaped) chloroplast per cell. There exist 'male' and 'female' algae which are morphologically indistinguishable and commonly referred to as mating type + (mt+, 'female') and mating type – (mt–, 'male'). Organelle inheritance in *Chlamydomonas* exhibits several interesting features (Umen and Goodenough 2001). First, chloroplast and mitochondrial genomes are oppositely inherited: the chloroplast DNA is transmitted maternally whereas the mitochondrial DNA is transmitted paternally. Second, during syngamy the maternal chloroplast fuses with the paternal chloroplast. Third, chloroplast DNA in mt+ *Chlamydomonas* gametes is methylated by a DNA methyltransferase converting cytosine to 5-methylcytosine (Nishiyama et al. 2002, 2004). In contrast, plastid DNA in higher plants is nowadays believed to be unmethylated (at least in somatic tissues; Marano and Carrillo 1991; Fojtová et al. 2001), although some early reports had suggested that cytosine methylation can occur also in higher plant plastomes (Ngernprasirtsiri et al. 1988a, 1988b).

If the two parental chloroplasts fuse upon mating, how then is maternal inheritance of chloroplast DNA in *Chlamydomonas* secured? Following syngamy, a zygotic maturation program sets in which leads to selective destruction of chloroplast DNA from the mt– parent, while the mt+ chloroplast genomes survive (Nishimura et al. 1999). Degradation of the paternal chloroplast genomes (by a specific endonuclease; Nishimura et al. 2002) is largely completed before fusion of the two parental chloroplasts occurs, thus resulting in uniparental inheritance of the maternal plastid DNA. It was reasonable to speculate that the difference in DNA methylation could be causally responsible for the selective degradation of the chloroplast genomes in mt– chloroplasts: by analogy to the restriction-methylation systems operating in eubacteria, this model posited that cytosine methylation protects mt+ plastid genomes from endonucleolytic degradation. However, recent studies have cast considerable doubt on this idea. Apparently, DNA methylation is not necessary for protection of mt+ plastid genomes in early zygotes and instead, may affect the relative rates of plastid genome replication in mt– and mt+ cells (Umen and Goodenough 2001). Thus the mechanistic details of how mt+ plastid genomes are protected from decay remain to be elucidated.

Maternal plastid DNA inheritance in *Chlamydomonas* is not absolute in that occasionally, paternal plastid DNA molecules (or fragments thereof) survive until chloroplast fusion occurs and thus can recombine with the maternal plastid genomes. These so-called 'exceptional zygotes' occur spontaneously at a frequency of a few percent (1-9%, depending on the genotype of the algal strains and on the experimental conditions). Interestingly, UV irradiation of mt+ gametes can significantly increase the intake rate of paternal ptDNA into the zygote. This discovery made in the mid-sixties of the last century has facilitated the recombination mapping of the chloroplast genome in *Chlamydomonas* (by R. Sager, J. Boynton, N. Gillham, and E. Harris; e.g. Sager and Ramanis 1976) and, in this way, contributed greatly to the development of plastid genetics in the pre-genomics era.

Together with the availability of antibiotic resistance markers encoded in the plastid genome, the low-level transmission of paternal plastid genes provides a powerful tool to quantify plastid inheritance (Bolen et al. 1982) and moreover, makes the unique system of chloroplast DNA inheritance in *Chlamydomonas* amenable to rigorous genetic analysis by selecting mutants with altered chloroplast genome transmission.

4.2 Biparental inheritance

A small number of angiosperms transmit their plastids biparentally. Working with *Pelargonium, Mirabilis, Melandrium, Antirrhinum* and *Aquilegia*, already Erwin Baur and Carl Correns noted in their first experiments on the inheritance of leaf variegations almost hundred years ago (Baur 1909, 1910; Correns 1909) that the mode of plastid inheritance may differ between species. While *Melandrium, Antirrhinum* and *Aquilegia* mutants transmitted their altered leaf color (which, as we now know, represented plastome mutations) purely maternally, similar traits could also be transmitted via pollen in *Pelargonium* (for review see Hagemann 2000). Baur concluded that the plastids (or the 'chromatophors', as they were called at that time) must be biparentally inherited in *Pelargonium zonale*. Later, other examples of species with biparental chloroplast inheritance were found (Table 2), including *Oenothera* (evening primrose), *Hypericum* (St. John's wort), and *Medicago* (alfalfa).

Extensive genetic work has determined the relative contributions of maternal and paternal plastids to the organelle population in the progeny in these species and revealed striking differences. In *Oenothera* and *Hypericum*, the rate of paternal transmission is relatively low, as evidenced by reciprocal crosses between white plastome mutants and green wild type plants. When the plastome mutant served as maternal parent, many white and variegated seedlings were obtained, but almost no green progeny. In contrast, when the plastome mutant was the paternal parent (i.e. the pollen donor), most F1 seedlings were uniformly green or variegated and only very few were white. In *Pelargonium* and *Medicago*, the paternal contributions are much greater. Whereas in *Pelargonium*, sperm and egg seem to make about equal plastid contributions to the zygote, paternal plastids are even predominantly inherited in alfalfa (Shi et al. 1991; Hagemann 2002).

Cytological investigations confirmed that, as expected, biparental plastid inheritance correlates with (i) the distribution of microspore plastids between vegetative cell and generative cell during the first pollen mitosis, (ii) the regular presence of viable plastids in sperm cells and (iii) their entry into the zygote.

4.3 Paternal inheritance

Thus far, only a single angiosperm species has been found to inherit its plastids uniparentally paternally: the kiwi plant *Actinidia deliciosa* (Testolin and Cipriani 1997). By contrast, in gymnosperms, paternal inheritance (or biparental inheritance with a strong predominance of paternal transmission) seems to be widespread (Szmidt et al. 1987; Neale et al. 1989; Mogensen 1996). Distinction between purely paternal inheritance and biparental inheritance with a strongly prevailing paternal component has been difficult, because most studies on plastid inheritance in gymnosperms suffer from statistically limited datasets. This is due to the lack of suitable phenotypic markers (i.e. plastome mutations resulting in pigment deficiencies and, thus, providing visible markers) in most species analyzed to date, which restricts the assay of progeny plants to RFLP analysis employing phenotypically neutral polymorphisms in the paternal and maternal ptDNAs. Naturally, this limits the number of progeny seedlings that can be analyzed and makes it difficult to exclude maternal plastid transmission below a certain level (Hagemann 2004).

Electron microscopic investigations confirmed the absence of plastids from egg cells (and the presence of them in sperm cells) in gymnosperm species displaying paternal plastid inheritance. In analogy to the diverse cytological mechanisms leading to maternal plastid inheritance (Fig. 2), at least two distinct mechanisms can contribute to paternal inheritance: plastid exclusion by unequal organelle distribution during female gametophyte development and/or plastid degradation in the egg cell (Mogensen 1996; Hagemann 2004).

4.4 Paternal leakage

As evident from the above-mentioned exceptional transmission of paternal plastid genes in *Chlamydomonas* and the discussion of paternal vs. biparental inheritance in gymnosperms, there is a grey zone between uniparental inheritance and biparental inheritance. In most instances, the conclusion that a given species transmits its plastids uniparentally is based on the phenotypic analysis of at most a few thousand progeny plants from reciprocal crosses (see 4.2). Failure to detect variegated seedlings is usually interpreted as uniparental mode of inheritance. However, in this approach, occasional plastid transmission from the other parent goes undetected if it occurs only at a very low level. How strict maternal plastid inheritance can be has been a highly controversial issue, particularly in the context of the level of transgene containment provided by plastid transformation technology (see below). Paternal leakage, the low-level paternal transmission of plastids in species believed to inherit their plastids maternally, is known to occur at least in some plant species (Avni and Edelman 1991; Medgyesy et al. 1986; Horlow et al. 1990; Wang et al. 2004). A large-scale genetic study in foxtail millet, *Setaria italica*, employed crosses between male-sterile yellow- or green-leafed herbicide susceptible lines (as maternal parent) and a line with chloroplast-inherited atrazine resistance as pollen donor (Wang et al. 2004). Assaying more than 780,000 hybrid offspring for atrazine resistance as it would be caused by paternally transmitted plastid genomes, paternal leakage was detected at a frequency of $3x10^{-4}$. Unfortunately, similarly reliable quantitative data in other plant species are largely lacking. It seems reasonable to suspect that the rate of paternal leakage can be very different in species representing the different subtypes of maternal inheritance (Fig. 2; Table 2), but this remains to be established experimentally.

The laborious and time-consuming genetic analyses required to establish low-level paternal leakage make it desirable to develop faster assays suitable to assess a species' potential to occasionally transmit paternal plastids via pollen. A rapid screening method that has been widely used employs staining of pollen with the DNA fluorochrome DAPI to identify plastid DNA in generative cells (Corriveau and Coleman 1988; Zhang et al. 2003). DAPI stains intensely plastid nucleoids which then can be readily detected by fluorescence microscopy. Absence of stainable plastid DNA from generative and sperm cells was taken as evidence for strictly maternal inheritance, whereas species with detectable ptDNA in generative and/or sperm cells were classified as potentially capable of occasional or regular biparental plastid transmission (which, however, does not mean that these species indeed display biparental plastid transmission: species of the Triticum type regularly have plastids in their sperm cells, but yet transmit their plastids maternally; Fig. 2 and 4.1.3). The latter was the case for roughly one fifth of the species investigated (Corriveau and Coleman 1988). Generally, how reliable DAPI staining of pollen grains can predict paternal leakage will require confirmation by rigorous genetic analysis.

Finally, it seems possible that environmental factors influence the rate of occasional paternal plastid transmission. Experimental evidence for this has been obtained already in *Chlamydomonas* where exposure of mt+ gametes to UV light increases the rate of occasional paternal chloroplast DNA transmission (Hagemann 2004). Whether or not abiotic stress conditions also affect plastid inheritance in higher plants (which is mechanistically very different from *Chlamydomonas*; Table 2) remains to be investigated.

4.5 Biotechnological implications of plastid inheritance

With very few exceptions (e.g. alfalfa), all major food and fodder crop species fall into the large group of angiosperm plants exhibiting maternal plastid transmission. Maternal inheritance excludes plastid genes from pollen transmission. Consequently, putting transgenes into the plastid genome instead of the nuclear genome (as done in conventional transgenic plants) can greatly reduce the risk of unwanted transgene spreading via pollen. Uncontrolled transgene transmission through pollen dispersal represents a major concern in the public debate on transgenic technologies in agriculture and plant biotechnology. In this respect, two scenarios are frequently discussed: (i) pollen flow from fields with genetically modified (GM) cultivars to neighboring fields with non-GM cultivars and (ii) unwanted transgene spreading via pollen from GM plants to related plant species (through hybridization with sexually compatible wild or weed species). As maternal transgene inheritance can potentially prevent outcrossing via pollen flow, plastid genetic engineering has recently stirred tremendous interest among plant biotechnologists (reviewed, e.g., in Bock 2001, 2007; Bock and Khan 2004; Maliga 2004).

To critically assess the level of transgene confinement attainable by chloroplast transformation technology, knowledge about the reliability of maternal inheritance and the possible frequency of paternal leakage in a given crop species is of paramount importance. In view of the many different cytological and molecular mechanisms involved in maternal plastid inheritance (Fig. 2; Table 2) and the significant variation in them even between closely related species, general conclusions and statements are inappropriate here. How strict maternal inheritance is and whether or not paternal leakage occurs must be assessed on a species-by-species basis and requires genetic analyses (crosses and phenotypic analysis of the progeny) at a very large scale (Wang et al. 2004). The possibility of occasional paternal leakage notwithstanding, it is self-evident that chloroplast transformation offers greatly increased transgene containment compared with conventional (nuclear-transgenic) plants which would transmit the transgene with every single pollen grain. However, if paternal leakage occurs in a given species and pollen transmission of the transgene must be prevented altogether, stacking of plastid transformation with other containment methods will be necessary to eliminate the residual outcrossing risk (Daniell 2002; Lee and Natesan 2006).

Acknowledgement

I thank Dr. Mark A. Schöttler and Dr. Stephanie Ruf for discussion and help with the artwork and Professors Rudolf Hagemann and Pal Maliga for helpful comments on the manuscript. Research on plastid genetics in my laboratory is supported by the Max Planck Society, the European Commission (Integrated Projects "Co-Extra" and "Plastomics" under the 6th Framework Programme) and the Deutsche Forschungsgemeinschaft.

References

Abdallah F, Salamini F, Leister D (2000) A prediction of the size and evolutionary origin of the proteome of chloroplasts of *Arabidopsis*. Trends Plant Sci 5:141-142

Aguettaz P, Seyer P, Pesey H, Lescure A-M (1987) Relation between the plastid gene dosage and the levels of 16S rRNA and rbcL gene transcripts during amyloplast to chloroplast change in mixotrophic spinach cell suspensions. Plant Mol Biol 8:169-177

Ahlert D, Ruf S, Bock R (2003) Plastid protein synthesis is required for plant development in tobacco. Proc Natl Acad Sci USA 100:15730-15735

Avni A, Edelman M (1991) Direct selection for paternal inheritance of chloroplasts in sexual progeny of *Nicotiana*. Mol Gen Genet 225:273-277

Baldev A, Gaikwad K, Kirti PB, Mohapatra T, Prakash S, Chopra VL (1998) Recombination between chloroplast genomes of *Trachystoma ballii* and *Brassica juncea* following protoplast fusion. Mol Gen Genet 260:357-361

Barbrook AC, Howe CJ (2000) Minicircular plastid DNA in the dinoflagellate Amphidinium operculatum. Mol Gen Genet 263:152-158

Baumgartner BJ, Rapp JC, Mullet JE (1988) Plastid transcription activity and DNA copy number increase early in barley chloroplast development. Plant Physiol 89:1011-1018

Baur E (1909) Das Wesen und die Erblichkeitsverhältnisse der "Varietates albomarginatae hort." von Pelargonium zonale. Z indukt Abstammungs- u Vererbungslehre 1:330-351

Baur E (1910) Untersuchungen über die Vererbung von Chromatophorenmerkmalen bei Melandrium, Antirrhinum und Aquilegia. Z indukt Abstammungs- u Vererbungslehre 4:81-102

Bendich AJ (1987) Why do chloroplasts and mitochondria contain so many copies of their genome? BioEssays 6:279-282

Bendich AJ (2004) Circular chloroplast chromosomes: the grand illusion. Plant Cell 16:1661-1666

Bendich AJ, Smith SB (1990) Moving pictures and pulsed-field gel electrophoresis show linear DNA molecules from chloroplasts and mitochondria. Curr Genet 17:421-425

Berg S, Krause K, Krupinska K (2004) The rbcL genes of two Cuscuta species, *C. gronovii* and *C. subinclusa*, are transcribed by the nuclear-encoded plastid RNA polymerase (NEP). Planta 219:541-546

Birky CW Jr, Walsh JB (1992) Biased gene conversion, copy number, and apparent mutation rate differences within chloroplast and bacterial genomes. Genetics 130:677-683

Birky CW Jr (1995) Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. Proc Natl Acad Sci USA 92:11331-11338

Bock R (2001) Transgenic chloroplasts in basic research and plant biotechnology. J Mol Biol 312:425-438

Bock R (2007) Plastid biotechnology: prospects for herbicide and insect resistance, metabolic engineering and molecular farming. Curr Opin Biotechnol: in press

Bock R, Hippler M (2002) Extranuclear inheritance: Functional genomics in chloroplasts. Prog Bot 63:106-131

Bock R, Khan MS (2004) Taming plastids for a green future. Trends Biotechnol 22:311-318

Bolen PL, Grant DM, Swinton D, Boynton JE, Gillham NW (1982) Extensive methylation of chloroplast DNA by a nuclear gene mutation does not affect chloroplast gene transmission in Chlamydomonas. Cell 28:335-343

Boudreau E, Takahashi Y, Lemieux C, Turmel M, Rochaix J-D (1997) The chloroplast ycf3 and ycf4 open reading frames of *Chlamydomonas reinhardtii* are required for the accumulation of the photosystem I complex. EMBO J 16:6095-6104

Bubunenko MG, Schmidt J, Subramanian AR (1994) Protein substitution in chloroplast ribosome evolution. A eukaryotic cytosolic protein has replaced its organelle homologue (L23) in spinach. J Mol Biol 240:28-41

Bungard RA (2004) Photosynthetic evolution in parasitic plants: insight from the chloroplast genome. BioEssays 26:235-247

Burrows PA, Sazanov LA, Svab Z, Maliga P, Nixon PJ (1998) Identification of a functional respiratory complex in chloroplasts through analysis of tobacco mutants containing disrupted plastid ndh genes. EMBO J 17:868-876

Cannon GC, Ward LN, Case CI, Heinhorst S (1999) The 68 kDa compacting nucleoid protein from soybean chloroplasts inhibits DNA synthesis *in vitro*. Plant Mol Biol 39:835-845

Cho HS, Lee SS, Kim KD, Hwang I, Lim J-S, Park Y-I (2004) DNA gyrase is involved in chloroplast nucleoid partitioning. Plant Cell 16:2665-2682

Chumley TW, Palmer JD, Mower JP, Fourcade HM, Calie PJ, Boore JL, Jansen RK (2006) The complete chloroplast genome sequence of Pelargonium x hortorum: Organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. Mol Biol Evol 23:2175-2190

Chun EHL, Vaugham MH, Rich A (1963) The isolation and characterization of DNA associated with chloroplast preparations. J Mol Biol 7:130-141

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

Corriveau JL, Coleman AW (1988) Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. Am J Bot 75:1443-1458

Daniell H (2002) Molecular strategies for gene containment in transgenic crops. Nature Biotechnol 20:581-586

Deng X-W, Wing RA, Gruissem W (1989) The chloroplast genome exists in multimeric forms. Proc Natl Acad Sci USA 86:4156-4160

dePamphilis CW, Palmer JD (1990) Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. Nature 348:337-339

Douglas SE, Penny SL (1999) The plastid genome of the cryptophyte alga, Guillardia theta: complete sequence and conserved synteny groups confirm its common ancestry with red algae. J Mol Evol 48:236-244

Drescher A, Ruf S, Calsa T Jr, Carrer H, Bock R (2000) The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. Plant J 22:97-104

Ems SC, Morden CW, Dixon CK, Wolfe KH, dePamphilis CW, Palmer JD (1995) Transcription, splicing and editing of plastid RNAs in the nonphotosynthetic plant *Epifagus virginiana*. Plant Mol Biol 29:721-733

Fojtová M, Kovarik A, Matyásek R (2001) Cytosine methylation of plastid genome in higher plants. Fact or artefact? Plant Sci 160:585-593

Goulding SE, Olmstead RG, Morden CW, Wolfe KH (1996) Ebb and flow of the chloroplast inverted repeat. Mol Gen Genet 252:195-206

Hagemann R (2000) Erwin Baur or Carl Correns: Who really created the theory of plastid inheritance? J Hered 91:435-440

Hagemann R (2002) Milestones in plastid genetics of higher plants. Prog Bot 63:1-51

Hagemann R (2004) The sexual inheritance of plant organelles. In: Daniell H, Chase C (eds) Molecular Biology and Biotechnology of Plant Organelles. Springer, Heidelberg, pp. 93-113

Hagemann R, Schröder M-B (1989) The cytological basis of the plastid inheritance in angiosperms. Protoplasma 152:57-64

Hager M, Biehler K, Illerhaus J, Ruf S, Bock R (1999) Targeted inactivation of the smallest plastid genome-encoded open reading frame reveals a novel and essential subunit of the cytochrome b6f complex. EMBO J 18:5834-5842

Hallick RB, Hong L, Drager RG, Favreau MR, Monfort A, Orsat B, Spielmann A, Stutz E (1993) Complete sequence of Euglena gracilis chloroplast DNA. Nucleic Acids Res 21:3537-3544

Hedtke B, Börner T, Weihe A (1997) Mitochondrial and chloroplast phage-type RNA polymerases in *Arabidopsis*. Science 277:809-811

Hess WR, Börner T (1999) Organellar RNA polymerases of higher plants. Inter Rev Cytol 190:1-59

Hess WR, Prombona A, Fieder B, Subramanian AR, Börner T (1993) Chloroplast rps15 and the rpoB/C1/C2 gene cluster are strongly transcribed in ribosome-deficient plastids: evidence for a functioning non-chloroplast-encoded RNA polymerase. EMBO J 12:563-571

Hess WR, Hoch B, Zeltz P, Hübschmann T, Kössel H, Börner T (1994) Inefficient rpl2 splicing in barley mutants with ribosome-deficient plastids. Plant Cell 6:1455-1465

Hippler M, Bock R (2004) Chloroplast proteomics. Prog Bot 65:90-105

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, 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

Horlow C, Goujaud J, Lépingle A, Missonier C, Bourgin J-P (1990) Transmission of paternal chloroplasts in tobacco (*Nicotiana tabacum*). Plant Cell Rep 9:249-252

Ikehara T, Uchida H, Suzuki L, Nakamura S (1996) Chloroplast nucleoids in large number and large DNA amount with regard to maternal inheritance in *Chlamydomonas reinhardtii*. Protoplasma 194:11-17

Isono K, Niwa Y, Satoh K, Kobayashi H (1997) Evidence for transcriptional regulation of plastid photosynthesis genes in *Arabidopsis thaliana* roots. Plant Physiol 114:623-630

Jansen RK, Raubeson LA, Boore JL, de Pamphilis CW, Chumley TW, Haberle RC, Wyman SK, Alverson AJ, Peery R, Herman SJ, Fourcade HM, Kuehl JV, McNeal JR, Leebens-Mack J, Cui L (2005) Methods for obtaining and analyzing whole chloroplast genome sequences. Methods Enzymol 395:348-384

Jenkins BD, Kulhanek DJ, Barkan A (1997) Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. Plant Cell 9:283-296

Jeong SY, Rose A, Meier I (2003) MFP1 is a thylakoid-associated, nucleoid-binding protein with a coiled-coil structure. Nucleic Acids Res 31:5175-5185

Joet T, Cournac L, Horvath EM, Medgyesy P, Peltier G (2001) Increased sensitivity of photosynthesis to antimycin A induced by inactivation of the chloroplast ndhB gene. Evidence for a participation of the NADH-dehydrogenase complex to cyclic electron flow around photosystem I. Plant Physiol 125:1919-1929

Jukes TH, Osawa S (1990) The genetic code in mitochondria and chloroplasts. Experientia 46:1117-1126

Kahlau S, Aspinall S, Gray JC, Bock R (2006) Sequence of the tomato chloroplast DNA and evolutionary comparison of solanaceous plastid genomes. J Mol Evol 63:194-207

Kaneko T, Tabata S (1997) Complete genome structure of the unicellular cyanobacterium Synechocystis sp. PCC6803. Plant Cell Physiol 38:1171-1176

Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, Nakamura Y, Miyajima N, Hirosawa M, Sugiura M, Sasamoto S, Kimura T, Hosouchi T, Matsuno A, Muraki A, Nakazaki N, Naruo K, Okumura S, Shimpo S, Takeuchi C, Wada T, Watanabe A, Yamada M, Yasuda M, Tabata S (1996) Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res 3:109-136

Khakhlova O, Bock R (2006) Elimination of deleterious mutations in plastid genomes by gene conversion. Plant J 46:85-94

Kobayashi T, Takahara M, Miyagishima S-y, Kuroiwa H, Sasaki N, Ohta N, Matsuzaki M, Kuroiwa T (2002) Detection and localization of a chloroplast-encoded HU-like protein that organizes chloroplast nucleoids. Plant Cell 14:1579-1589

Kowallik KV, Stoebe B, Schaffran I, Kroth-Pancic P, Freier U (1995) The chloroplast genome of a chlorophyll a+c-containing alga, *Odontella sinensis*. Plant Mol Biol Rep 13:336-342

Krause K, Krupinska K (2000) Molecular and functional properties of highly purified transcriptionally active chromosomes from spinach chloroplasts. Physiol Plant 109:188-195

Kumar D, Mukherjee S, Reddy MK, Tewari KK (1995) A novel single-stranded DNA-specific endonuclease from pea chloroplasts. J Exp Bot 46:767-776

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 (1991) The replication, differentiation, and inheritance of plastids with emphasis on the concept of organelle nuclei. Int Rev Cytol 128:1-62

Lee D, Natesan E (2006) Evaluating genetic containment strategies for transgenic plants. Trends Biotechnol 24:109-114

Li W, Ruf S, Bock R (2006) Constancy of organellar genome copy numbers during leaf development and senescence in higher plants. Mol Gen Genomics 275:185-192

Liere K, Link G (1995) RNA-binding activity of the matK protein encoded by the chloroplast trnK intron from mustard (Sinapis alba L.). Nucleic Acids Res 23:917-921

Lilly JW, Havey MJ, Jackson SA, Jiang J (2001) Cytogenomic analyses reveal the structural plasticity of the chloroplast genome in higher plants. Plant Cell 13:245-254

Liu J-W, Rose RJ (1992) The spinach chloroplast chromosome is bound to the thylakoid membrane in the region of the inverted repeat. Biochem Biophys Res Commun 184:993-1000

Lung B, Zemann A, Madej MJ, Schuelke M, Techritz S, Ruf S, Bock R, Hüttenhofer A (2006) Identification of small non-coding RNAs from mitochondria and chloroplasts. Nucleic Acids Res 34:3842-3852

Maier RM, Neckermann K, Igloi GL, Kössel H (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. J Mol Biol 251:614-628

Majeran W, Wollman F-A, Vallon O (2000) Evidence for a role of ClpP in the degradation of the chloroplast cytochrome b6f complex. Plant Cell 12:137-149

Maliga P (2004) Plastid transformation in higher plants. Annu Rev Plant Biol 55:289-313

Marano MR, Carrillo N (1991) Chromoplast formation during tomato fruit ripening. No evidence for plastid methylation. Plant Mol Biol 16:11-19

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci USA 99:12246-12251

Maul JE, Lilly JW, Cui L, dePamphilis CW, Miller W, Harris EH, Stern DB (2002) The *Chlamydomonas reinhardtii* plastid chromosome: islands of genes in a sea of repeats. Plant Cell 14:2659-2679

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

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

Millen RS, Olmstead RG, Adams KL, Palmer JD, Lao NT, Heggie L, Kavanagh TA, Hibberd JM, Gray JC, Morden CW, Calie PJ, Jermiin LS, Wolfe KH (2001) Many parallel losses of infA from chloroplast DNA during angiosperm evolution with multiple independent transfers to the nucleus. Plant Cell 13:645-658

Misumi O, Suzuki L, Nishimura Y, Sakai A, Kawano S, Kuroiwa H, Kuroiwa T (1999) Isolation and phenotypic characterization of *Chlamydomonas reinhardtii* mutants defective in chloroplast DNA segregation. Protoplasma 209:273-282

Mogensen HL (1996) The hows and whys of cytoplasmic inheritance in seed plants. Am J Bot 83:383-404

Mohr G, Perlman PS, Lambowitz AM (1993) Evolutionary relationships among group II intron-encoded proteins and identification of a conserved domain that may be related to maturase function. Nucleic Acids Res 21:4991-4997

Monod C, Takahashi Y, Goldschmidt-Clermont M, Rochaix J-D (1994) The chloroplast ycf8 open reading frame encodes a photosystem II polypeptide which maintains photosynthetic activity under adverse growth conditions. EMBO J 13:2747-2754

Morden CW, Wolfe KH, dePamphilis CW, Palmer JD (1991) Plastid translation and transcription genes in a non-photosynthetic plant: intact, missing and pseudo genes. EMBO J 10:3281-3288

Mukherjee SK, Reddy MK, Kumar D, Tewari KK (1994) Purification and characterization of a eukaryotic type 1 topoisomerase from pea chloroplasts. J Biol Chem 269:3793-3801

Muller HJ (1964) The relation of recombination to mutational advance. Mutat Res 1:2-9

Munekage Y, Hashimoto M, Miyake C, Tomizawa K-i, Endo T, Tasaka M, Shikanai T (2004) Cyclic electron flow around photosystem I is essential for photosynthesis. Nature 429:579-582

Nakano T, Murakami S, Shoji T, Yoshida S, Yamada Y, Sato F (1997) A novel protein with DNA binding activity from tobacco chloroplast nucleoids. Plant Cell 9:1673-1682

Neale DB, Marshall KA, Sederoff RR (1989) Chloroplast and mitochondrial DNA are paternally inherited in *Sequoia sempervirens* D. Don Endl. Proc Natl Acad Sci USA 86:9347-9349

Ngernprasirtsiri J, Kobayashi H, Akazawa T (1988a) DNA methylation occurred around lowly expressed genes of plastid DNA during tomato fruit development. Plant Physiol 88:16-20

Ngernprasirtsiri J, Kobayashi H, Akazawa T (1988b) DNA methylation as a mechanism of transcriptional regulation in nonphotosynthetic plastids in plant cells. Proc Natl Acad Sci USA 85:4750-4754

Nishimura Y, Misumi O, Matsunaga S, Higashiyama T, Yokota A, Kuroiwa T (1999) The active digestion of uniparental chloroplast DNA in a single zygote of *Chlamydomonas reinhardtii* is revealed by using the optical tweezer. Proc Natl Acad Sci USA 96:12577-12582

Nishimura Y, Misumi O, Kato K, Inada N, Higashiyma T, Momoyama Y, Kuroiwa T (2002) An mt^{+} gamete-specific nuclease that targets mt^{-} chloroplasts during sexual reproduction in *C. reinhardtii*. Genes & Dev 16:1116-1128

Nishiyama R, Ito M, Yamaguchi Y, Koizumi N, Sano H (2002) A chloroplast-resident DNA methyltransferase is responsible for hypermethylation of chloroplast genes in *Chlamydomonas* maternal gametes. Proc Natl Acad Sci USA 99:5925-5930

Nishiyama R, Wada Y, Mibu M, Yamaguchi Y, Shimogawara K, Sano H (2004) Role of a nonselective de novo DNA methyltransferase in maternal inheritance of chloroplast genes in the green alga, *Chlamydomonas reinhardtii*. Genetics 168:809-816

Ohta N, Matsuzaki M, Misumi O, Miyagishima S-y, Nozaki H, Tanaka K, Shin-I T, Kohara Y, Kuroiwa T (2003) Complete sequence and analysis of the plastid genome of the unicellular red alga *Cyanidioschyzon merolae*. DNA Res 10:67-77

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, 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 T, Takeuchi M, Chang Z, Aota S-i, Inokuchi H, Ozeki H (1988) Structure and organization of *Marchantia polymorpha* chloroplast genome. I. Cloning and gene identification. J Mol Biol 203:281-298

Oldenburg DJ, Bendich AJ (2004) Most chloroplast DNA of maize seedlings in linear molecules with defined ends and branched forms. J Mol Biol 335:953-970

Orsat B, Monfort A, Chatellard P, Stutz E (1992) Mapping and sequencing of an actively transcribed Euglena gracilis chloroplast gene (ccsA) homologous to the *Arabidopsis thaliana* nuclear gene cs (ch-42). FEBS Lett 303:181-184

Osawa S, Jukes TH, Watanabe K, Muto A (1992) Recent evidence for evolution of the genetic code. Microbiol Rev 56:229-264

Palmer JD (1983) Chloroplast DNA exists in two orientations. Nature 301:92-93

Palmer JD, Thompson WF (1982) Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. Cell 29:537-550

Palmer JD, Nugent JM, Hebron LA (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

Pfalz J, Liere K, Kandlbinder A, Dietz K-J, Oelmüller R (2006) pTAC2, -6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. Plant Cell 18:176-197

Reith M (1995) Molecular biology of rhodophyte and chromophyte plastids. Annu Rev Plant Physiol Plant Mol Biol 46:549-575

Reith M, Munholland J (1993) A high-resolution gene map of the chloroplast genome of the red alga *Porphyra purpurea*. Plant Cell 5:465-475

Reith M, Munholland J (1995) Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. Plant Mol Biol Rep 13:333-335

Revill MJW, Stanley S, Hibberd JM (2005) Plastid genome structure and loss of photosynthetic ability in the parasitic genus Cuscuta. J Exp Bot 56:2477-2486

Rochaix J-D (1997) Chloroplast reverse genetics: new insights into the function of plastid genes. Trends Plant Sci 2:419-425

Rogalski M, Ruf S, Bock R (2006) Tobacco plastid ribosomal protein S18 is essential for cell survival. Nucleic Acids Res 34:4537-4545

Rolland N, Dorne A-J, Amoroso G, Sültemeyer DF, Joyard J, Rochaix J-D (1997) Disruption of the plastid ycf10 open reading frame affects uptake of inorganic carbon in the chloroplast of *Chlamydomonas*. EMBO J 16:6713-6726

Ruf S, Kössel H, Bock R (1997) Targeted inactivation of a tobacco intron-containing open reading frame reveals a novel chloroplast-encoded photosystem I-related gene. J Cell Biol 139:95-102

Rujan T, Martin W (2001) How many genes in *Arabidopsis* come from cyanobacteria ? An estimate from 386 protein phylogenies. Trends Genet 17:113-121

Sager R, Ishida MR (1963) Chloroplast DNA in *Chlamydomonas*. Proc Natl Acad Sci USA 50:725-730

Sager R, Ramanis Z (1976) Chloroplast genetics of *Chlamydomonas*. II. Mapping by cosegregation frequency analysis. Genetics 83:323-340

Sakai A, Yamashita H, Nemoto Y, Kawano S, Kuroiwa T (1991) Transcriptional activity of morphologically intact proplastid-nuclei (nucleoids) isolated from tobacco cultured cells. Plant Cell Physiol 32:835-843

Salvador ML, Klein U, Bogorad L (1998) Endogenous fluctuations of DNA topology in the chloroplast of *Chlamydomonas reinhardtii*. Mol Cell Biol 18:7235-7242

Sasaki Y, Sekiguchi K, Nagano Y, Matsuno R (1993a) Chloroplast envelope protein encoded by chloroplast genome. FEBS Lett 316:93-98

Sasaki Y, Hakamada K, Suama Y, Nagano Y, Furusawa I, Matsuno R (1993b) Chloroplast-encoded protein as a subunit of acetyl-CoA carboxylase in pea plant. J Biol Chem 268:25118-25123

Sasaki Y, Konishi T, Nagano Y (1995) The compartmentation of acetyl-coenzyme A carboxylase in plants. Plant Physiol 108:445-449

Sato N, Albrieux C, Joyard J, Douce R, Kuroiwa T (1993) Detection and characterization of a plastid envelope DNA-binding protein which may anchor plastid nucleoids. EMBO J 12:555-561

Scharff LB, Koop H-U (2006) Linear molecules of tobacco ptDNA end at known replication origins and additional loci. Plant Mol Biol 62:611-621

Sekine K, Hase T, Sato N (2002) Reversible DNA compaction by sulfite reductase regulates transcriptional activity of chloroplast nucleoids. J Biol Chem 277:24399-24404

Shanklin J, DeWitt ND, Flanagan JM (1995) The stroma of higher plant plastids contain CplP and CplC, functional homologs of *Escherichia coli* ClpP and ClpA: an archetypal two-component ATP-dependent protease. Plant Cell 7:1713-1722

Shi L, Zhu T, Mogensen HL, Smith SE (1991) Paternal plastid inheritance in alfalfa: plastid nucleoid number within generative cells correlates poorly with plastid number and male plastid transmission strength. Curr Genet 19:399-401

Shikanai T, Endo T, Hashimoto T, Yamada Y, Asada K, Yokota A (1998) Directed disruption of the tobacco ndhB gene impairs cyclic electron flow around photosystem I. Proc Natl Acad Sci USA 95:9705-9709

Shimada H, Sugiura M (1991) Fine structural features of the chloroplast genome: comparison of the sequenced chloroplast genomes. Nucleic Acids Res 19:983-995

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 BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5:2043-2049

Sijben-Müller G, Hallick RB, Alt J, Westhoff P, Herrmann RG (1986) Spinach plastid gene coding for initiation factor IF-1, ribosomal protein S11 and RNA polymerase α-subunit. Nucleic Acids Res 14:1029-1042

Simpson CL, Stern DB (2002) The treasure trove of algal chloroplast genomes. Surprises in architecture and gene content, and their functional implications. Plant Physiol 129:957-966

Stein DB, Palmer JD, Thompson WF (1986) Structural evolution and flip-flop recombination of chloroplast DNA in the fern genus Osmunda. Curr Genet 10:835-841

Sugiura M (1989) The chloroplast chromosomes in land plants. Annu Rev Cell Biol 5:51-70

Sugiura M (1992) The chloroplast genome. Plant Mol Biol 19:149-168

Szmidt AE, Alden T, Hällgren J-E (1987) Paternal inheritance of chloroplast DNA in Larix. Plant Mol Biol 9:59-64

Testolin R, Cipriani G (1997) Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in the genus Actinidia. Theor Appl Genet 94:897-903

Tewari KK, Wildman SG (1966) Chloroplast DNA from tobacco leaves. Science 153:1269-1271

Thanh ND, 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

Tilney-Bassett RAE (1984) The genetic evidence for nuclear control of chloroplast biogenesis in higher plants. In: Ellis, R J (ed) Chloroplast Biogenesis. Cambridge Univ Press, Cambridge, pp13-50

Tilney-Bassett RAE (1994) Nuclear control of chloroplast inheritance in higher plants. J Heredity 85:347-354

Umen JG, Goodenough UW (2001) Chloroplast DNA methylation and inheritance in Chlamydomonas. Genes & Dev 15:2585-2597

Wakasugi T, Tsudzuki J, Ito S, Nakashima K, Tsudzuki T, Sugiura M (1994) Loss of all ndh genes as determined by sequencing the entire chloroplast genome of the black pine Pinus thunbergii. Proc Natl Acad Sci USA 91:9794-9798

Wakasugi T, Nagai T, Kapoor M, Sugita M, Ito M, Ito S, Tsudzuki J, Nakashima K, Tsudzuki T, Suzuki Y, Hamada A, Ohta T, Inamura A, Yoshinaga K, Sugiura M (1997) Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: The existence of genes possibly involved in chloroplast division. Proc Natl Acad Sci USA 94:5967-5972

Wakasugi T, Tsudzuki T, Sugiura M (2001) The genomics of land plant chloroplasts: gene content and alteration of genomic information by RNA editing. Photosynthesis Res 70:107-118

Wang T, Li Y, Shi Y, Reboud X, Darmency H, Gressel J (2004) Low frequency transmission of a plastid-encoded trait in *Setaria italica*. Theor Appl Genet 108:315-320

Wimpee CF, Wrobel RL, Garvin DK (1991) A divergent plastid genome in *Conopholis americana*, an achlorophyllous parasitic plant. Plant Mol Biol 17:161-166

Wolfe KH, Li W-H, Sharp PM (1987) Rates of nucleotide substitutions vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054-9058

Wolfe KH, Morden CW, Ems SC, Palmer JD (1992) Rapid evolution of the plastid translational apparatus in a nonphotosynthetic plant: Loss or accelerated sequence evolution of tRNA and ribosomal protein genes. J Mol Evol 35:304-317

Wolfe KH, Morden CW, Palmer JD (1992) Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. Proc Natl Acad Sci USA 89:10648-10652

Xie Z, Merchant S (1996) The plastid-encoded ccsA gene is required for heme attachment to chloroplast c-type cytochromes. J Biol Chem 271:4632-4639

Xie Z, Culler D, Dreyfuss BW, Kuras R, Wollman F-A, Girard-Bascou J, Merchant S (1998) Genetic analysis of chloroplast c-type cytochrome assembly in *Chlamydomonas reinhardtii*: one chloroplast locus and at least four nuclear loci are required for heme attachment. Genetics 148:681-692

Zhang Q, Liu Y, Sodmergen (2003) Examination of the cytoplasmic DNA in male reproductive cells to determine the potential for cytoplasmic inheritance in 295 angiosperm species. Plant Cell Physiol 44:941-951

Zhang Z, Green BR, Cavalier-Smith T (1999) Single gene circles in dinoflagellate chloroplast genomes. Nature 400:155-159

Zoschke R, Liere K, Börner T (2007) From seedling to mature plant: *Arabidopsis* plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. Plant J: in press

Zubko MK, Day A (1998) Stable albinism induced without mutagenesis: a model for ribosome-free plastid inheritance. Plant J 15:265-271

Zubko MK, Day A (2002) Differential regulation of genes transcribed by nucleus-encoded plastid RNA polymerase, and DNA amplification, within ribosome-deficient plastids in stable phenocopies of cereal albino mutants. Mol Genet Genomics 267:27-37

Bock, Ralph

Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany

rbock@mpimp-golm.mpg.de

List of abbreviations

CF: coupling factor
cyt b_6f: cytochrome b_6f complex
DAPI: 4',6-diamidino-2-phenylindole
GM: genetically modified
IR: inverted repeat
kb: kilobase pairs
LSC: large single copy region
LSU: large subunit
mt: mating type
NEP: nuclear-encoded RNA polymerase
ORF: open reading frame
PEP: plastid-encoded RNA polymerase
PSI: photosystem I
PSII: photosystem II
ptDNA: plastid DNA
RFLP: restriction fragment length polymorphism
Rubisco: ribulose 1,5-bisphosphate carboxylase/oxygenase
SSC: small single copy region
SSU: small subunit
ycf: hypothetical chloroplast reading frame

DNA replication, recombination, and repair in plastids

Anil Day and Panagiotis Madesis

Abstract

Plastid DNA is conserved, highly polyploid and uniform within a plant reflecting efficient plastid DNA replication/recombination/repair (DNA-RRR) pathways. We will review the current understanding of the DNA sequences, proteins, and mechanisms involved in plastid genome maintenance. This includes analysis of the topological forms of plastid DNA, models of plastid DNA replication, homologous recombination, replication slippage, DNA repair, and plastid DNA-RRR-proteins. We will focus on flowering plants but include information from algae when relevant. Plastid DNA is comprised of a multimeric series of circular, linear, and branched forms. Variant plastid DNA molecules include small linear palindromes with hairpin ends. Plastid transformation has demonstrated an efficient homologous recombination pathway, acting on short ~200 bp sequences, that is active throughout shoot development. These functional studies involving plastid transformation to manipulate DNA sequences, combined with genomics and reverse genetics to isolate mutants in plastid DNA-RRR proteins, will be particularly important for making progress in this field.

1 The importance of DNA replication, recombination, and repair pathways in plastids

All life on earth relies on DNA-replication, recombination, and repair (DNA-RRR) pathways for stable maintenance and propagation of DNA. Plants are dependent on light for growth and are exposed to the damaging effects of radiation. Chloroplasts, the light harvesting plastids of plants and algae, are the sites at which radiation might be expected to cause the greatest damage. Radiation itself and toxic reactive oxygen species, produced as by products of photosynthesis, are examples of destructive agents that can damage plastid DNA (Fig. 1). Despite these damaging agents plastid DNA is relatively well conserved with respect to sequence and gene content (Palmer 1990) and is widely used for phylogenetic studies (Soltis et al. 1999). Plastid genomes are present in multiple copies per cell, which are identical resulting in a uniform population of DNA molecules within an

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0231 / Published online: 13 June 2007

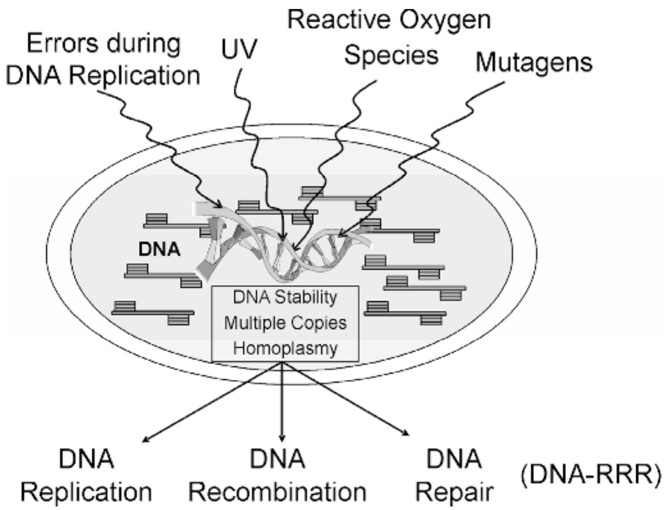


Fig. 1. DNA-RRR pathways are responsible for the high copy number, uniformity, and stable maintenance of plastid genomes.

individual cell or multicellular plant. Highly effective DNA maintenance pathways in plastids must underpin the evolutionary stability and uniformity of plastid DNA.

DNA-RRR pathways must overcome two potential problems associated with the mode of inheritance and ploidy of plastid DNA. First, in sexual crosses plastid DNA often exhibits uniparental inheritance (Corriveau and Coleman 1988; Reboud and Zeyl 1994) reducing the possibility of DNA recombination between parental plastid genomes (Chapter 3). Moreover, in flowering plants, when two plastid types are present in the same cell they rarely recombine (Medgyesy et al. 1985) and segregate away from each other to form cells with pure populations of each plastid type. Segregation of plastids during vegetative growth is known as cytoplasmic sorting or vegetative segregation (Birky 1994). Lack of DNA recombination between different plastid types means that plastids propagate asexually and do not have the benefits of sex and DNA recombination between parental alleles (applicable to nuclear genes) to eliminate deleterious mutations. Muller's ratchet (Muller 1964) would operate leading to an accumulation of mutations in plastid DNA. Second, because plastid DNA is present in multiple copies any new mutations in plastid genes would be masked by the wild type (WT) alleles present in the cell. Asexual propagation and a high degree of polyploidy are two features of plastid DNA that would be expected to promote the accumulation of mutations. Without effective plastid DNA-RRR pathways, plastid mutations would accumulate with time resulting in loss of fitness and death.

1.1 Proteins and DNA targets of plastid DNA-RRR pathways

Plastid DNA maintenance is governed by cis-acting plastid DNA sequences which are the targets for trans-acting proteins that replicate, recombine, and repair plastid genomes. The plastid genomes of green algae and plants that have been characterised do not encode any known DNA-RRR proteins. A number of non-green algae, including diatoms, red and cryptomonad algae contain a plastid *dnaB*-like gene (Kowallik et al. 1995; Reith and Munholland 1995; Douglas and Penny 1999). The *dnaB* gene encodes a DNA helicase involved in replication (Nakayama et al. 1984). In angiosperms the absence of plastid-encoded DNA-RRR proteins is demonstrated by the observation that plastid DNA is replicated in albino cereal (Hess et al. 1994; Zubko and Day 2002) and *Brassica* plants (Zubko and Day 1998) lacking plastid-encoded proteins.

DNA replication, recombination, and repair were once considered to be distinct pathways but more recent work in bacteria has shown they are interrelated processes (Kreuzer 2005). Pathways for recombination-dependent DNA replication and DNA replication-dependent recombination have been described (Kowalczykowski 2000; Kreuzer 2000, 2005) and are applicable to plastid DNA (see Section 7 below). This review will summarise our current knowledge on the mechanisms, DNA sequences and proteins involved in the maintenance of plastid DNA. We will focus on plastid DNA in flowering plants but will include relevant work from algal plastids where appropriate.

2 Plastid DNA polyploidy, packaging, and segregation

2.1 Plastid DNA copy number

One thousand to 1,700 copies of plastid DNA are present per cell in *Arabidopsis thaliana* leaves (Zoschke et al. 2007) whilst five thousand to over ten thousand copies of plastid DNA per cell are present in leaves of *Pisum sativum* (Lamppa and Bendich 1979), *Triticum aestivum* (Day and Ellis 1984), *Spinacia oleracea* (Lawrence and Possingham 1986), and *Hordeum vulgare* (Baumgartner et al. 1989). Fewer plastid genomes per cell are found in other organs containing non-green plastids, such as the roots of *P. sativum* (~500 copies per cell, Lamppa and Bendich 1979) and *T. aestivum* (~300 copies per cell, Day and Ellis 1984). An increase in plastid genome copies is associated with the development of chloroplasts from precursor plastids. Copy number estimates based on quantifying the DNA present in purified plastids from leaf cells of different ages indicate the number of genomes per chloroplast reaches a maximum value in young leaves and then decreases in older cells well before senescence. For example, in the developing primary leaf of four-day-old *H. vulgare* seedlings, plastids in the basal meristem were estimated to contain ~130 genomes, this increased to ~210 genomes in chloroplasts in older cells located one to three cm above the meristem, and decreased to ~50 genomes per chloroplast in the oldest cells in the leaf tip

(Baumgartner et al. 1989). More recent publications also report decreases in genomes per chloroplasts in mature leaves compared to young leaves. Decreases observed include 225 to 106 genomes per chloroplast in *Zea mays* (Oldenburg and Bendich 2004a; Shaver et al. 2006), 135 to 53 genomes per chloroplast in *P. sativum*, 122 to 47 genomes per chloroplast in *Medicago truncatula*, and 190 to 70 genomes per chloroplast in *Nicotiana tabacum* (Shaver et al. 2006). These results are consistent with the idea that replication of plastid DNA takes place predominantly in meristematic cells and leaf primordia (Kuroiwa 1991; Fujie et al. 1994; see Section 13.1) and as plastids divide during leaf development the number of genomes per plastid falls. We know very little about the replication mechanisms regulating the copy number of plastid DNA. Some progress has been made in this field with the recent finding that copy number is influenced by specific plastid DNA sequences. Deletion of the OriA plastid DNA sequence implicated in DNA replication (Section 4.1 below) reduces the copy number of plastid DNA in developing leaves of *N. tabacum* (Scharff and Koop 2007).

Recent publications detailing two to threefold reductions in DNA levels per plastid during leaf maturation also suggest the apparent absence of DNA in some mature chloroplasts: DNA was not observed in approximately 11% of *M. truncatula*, 9% of *P. sativum*, 80-90% of *Z. mays* (Shaver et al. 2006) and 29% of *A. thaliana* chloroplasts (Rowan et al. 2004). Loss of DNA from chloroplasts during leaf ageing was not observed in *N. tabacum* (Shaver et al. 2006). Based on the results obtained in *Z. mays* and *A. thaliana* a mechanism that actively degrades DNA in maturing chloroplasts was proposed by the authors (Oldenburg and Bendich 2004a; Rowan et al. 2004). These results appear to suggest that chloroplasts lacking DNA retain photosynthetic activity for long periods (Oldenburg and Bendich 2004a; Rowan et al. 2004), which conflicts with our current understanding of the importance of plastid gene expression for maintaining chloroplast functions. An alternative explanation for the apparent absence of DNA in isolated plastids is that it is an artefact resulting from the experimental approaches used (Li et al. 2006). In particular, degradation of plastid DNA during the purification of plastids or during treatment of plastids with DNase I (to remove contaminating extra-plastidic DNA outside plastids) will give rise to low copy number estimates. This might be more problematic for old leaf cells of some species where the release of DNA nucleases during homogenisation and changes in plastid porosity might allow nucleases to enter plastids. DNA fluoresces when stained with the DNA-binding dye 4',6-diamidino-2-phenylindole (DAPI). DAPI stained chloroplasts in leaf sections of *A. thaliana* and *N. tabacum* fixed immediately after sectioning appeared to show the same pattern of DNA reduction or loss obtained with isolated chloroplasts supporting the data with isolated chloroplasts (Shaver et al. 2006). Quantitation of plastid DNA levels in total DNA preparations from liquid nitrogen frozen leaves using Southern blot analysis is less sensitive to plastid DNA degradation during sample preparation and provides a robust method for estimating plastid DNA levels. Such an approach using rapidly extracted total DNA can be used to confirm or dismiss the findings based on isolated chloroplasts. Copy number estimates can also be obtained using quantitative real-time PCR on total DNA with controls to rule out contamination or amplification of ’promiscuous’ plastid

DNA sequences present in mitochondria and nuclei (Zoschke et al. 2007). Using these methods it appears that once chloroplast development is completed the levels of plastid DNA in total DNA appear to remain relatively constant during further leaf development in *A. thaliana* (Zoschke et al. 2007; Li et al. 2006) and *N. tabacum* (Li et al. 2006). These results do not support the idea of a dramatic reduction in plastid DNA levels during leaf development in *A. thaliana*.

Replication of plastid and nuclear DNA do not appear to be tightly co-ordinated in *N. tabacum* (Heinhorst et al. 1985) or the green alga *Chlamydomonas reinhardtii* (Chiang and Sueoka 1967). In contrast to the stringent controls restricting nuclear DNA synthesis to one round of replication during the S phase of each cell cycle, plastid DNA replication appears to be less stringent and is not limited to the S phase (Heinhorst and Cannon 1993). Moreover, plastid genomes appear to be chosen randomly for replication (Birky 1994). The prereplication factor CDT1 appears to affect both nuclear DNA replication and plastid division in *A. thaliana* and provides a possible link between the cell cycle and plastid division (Raynaud et al. 2005). In synchronous cultures of *C. reinhardtii*, duplication of plastid DNA could be localised to a particular time period (Chiang and Sueoka 1967) whereas plastid DNA synthesis, monitored by ^{32}P incorporation, was observed throughout the cell cycle (Grant et al. 1978). The ^{32}P-incorporation was suggested to be due to DNA repair activities which were required to maintain plastid genomes throughout the cell cycle (Grant et al. 1978).

2.2 Packaging of plastid DNA

Within plastids, the DNA is not dispersed but localised into aggregates of DNA and protein called nucleoids (Kuroiwa 1991; Sakai et al. 2004). The uniformity of plastid DNA is governed by DNA-RRR pathways that are likely to be carried out in nucleoids. The organisation of multiple plastid genome copies into a smaller number of units will govern the segregation of plastid DNA during plastid and cell divisions (VanWinkle-Swift 1980) and will facilitate cytoplasmic sorting. The number, sizes, morphologies, and distribution of nucleoids, visualised by DAPI staining, vary during development of chloroplasts from proplastids (Miyamura et al. 1986). *T. aestivum* proplastids contain one to ten nucleoids and 30 to 40 plastid genomes whereas chloroplasts contain ten to thirty nucleoids and 70 to 100 plastid genomes (Miyamura et al. 1990). In *Nicotiana*, mature chloroplasts contain eight to forty nucleoids, each with about ten plastid genomes (Kuroiwa 1991). Nucleoids appear to be located in the stroma or attached to the envelope or thylakoids depending on the plastid type (Sato et al. 2003; Sakai et al. 2004). The functional significance of changes in the intra-plastidic location of nucleoids is not known. However, it is interesting to note that constitutive expression of the *B. napus* homologue of the *P. sativum* plastid envelope DNA (PEND) binding protein in *N. tabacum* nuclear transformants leads to an albino phenotype possibly due to a lack of release of DNA from the envelope (Wycliffe et al. 2005). The PEND protein is targeted to plastids and might be involved in anchoring plastid DNA to the inner envelope during early chloroplast development (Sakai et al. 2004),

2.3 Segregation of plastid genomes

The replication and segregation mechanisms in plastids prevent the persistence of two different plastid genomes in cells. Heteroplasmy can only be maintained for long periods of time by selection for both plastid genomes (Drescher et al. 2000; Shikanai et al. 2001; Kode et al. 2005). The fine details of heteroplasmy are not known and the two plastid DNA types might be mixed within single nucleoids, or localised to separate nucleoids within a plastid or be separated into two populations of plastids within a cell. Heteroplasmy within a plastid is required when a lethal mutation is plastid autonomous (Kode et al. 2005) and cannot be rescued by import of cytoplasmic metabolites.

3 Topological forms of plastid DNA

The mechanisms of plastid DNA replication and maintenance will be reflected in the topologies of DNA molecules found in plastids. Plastids are most likely to be descendents of ancient cyanobacteria (Martin et al. 2002), which contain circular double-stranded DNA genomes. Circular DNA overcomes the problems of replicating gaps at the ends of linear DNA molecules following RNA primer removal at the 5’ ends of newly synthesized DNA (Cavalier-Smith 1974). The sequence maps of all the plastid genomes that have been characterized are circular (Chapter 3). In the majority of species the genomes can be represented as a single circular double-stranded DNA molecule containing all genes. Dinoflagellates are an exception and contain genes dispersed over a number of DNA mini-circles each with one to three genes (Koumandou et al. 2004). A circular sequence or restriction map does not necessarily imply the physical structure of a DNA species is a circle (Streisinger et al. 1964). Tandemly repeated DNA sequences (Fig. 2a), such as nuclear ribosomal RNA genes, on a linear chromosome or circularly permuted sequences arranged on separate linear DNA molecules of defined (Fig. 2b) or varying lengths (Fig. 2c) will also give rise to circular maps (Fig. 2d). To study the structure of plastid DNA requires the analysis of intact DNA isolated from chloroplasts. Because double-stranded DNA is prone to breakage by shearing, the analysis of plastid DNA topology requires distinguishing breakage products of the extraction process from intact plastid DNA molecules. Most studies have involved chloroplasts, which are easily identified, abundant in leaves and relatively easy to purify. The structure of chloroplast DNA has been studied by microscopic and gel electrophoretic methods.

An early electron microscopic study showed that monomer circles corresponding in size to a single set of plastid genes represented 37% of the DNA extracted from *P. sativum* chloroplasts (Kolodner and Tewari 1972). The remaining DNA

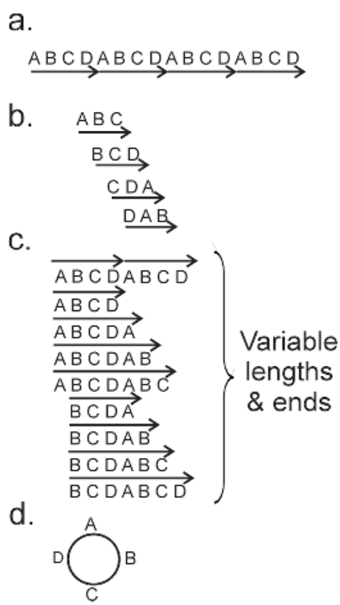


Fig. 2. a) Tandemly repeated linear DNA sequences, b) Circularly permuted linear DNA sequences of fixed length, and c) Multimeric linear DNA molecules of variable sizes and a number of dispersed ends will give rise to d) Circular sequence maps. Arrows indicate orientation of sequences. Molecules starting with the letters c or d are not shown in c).

was mainly comprised of sub-genomic linear DNA forms, which could represent breakage products of circular plastid DNA molecules or even contaminating nuclear DNA (Fig. 3a). Breakage during extraction would give rise to variable ratios of circles to linear products with each preparation. In a later study, three to four percent of circular species were found to be dimers (Kolodner and Tewari 1979). These were arranged head-to-tail in *P. sativum* chloroplasts, and both head-to-head and head-to-tail in *S. oleracea* and *Lactuca sativa* chloroplasts. The majority of *L. sativa* and *S. oleracea* chloroplast DNA dimers (about 80%) were arranged head-to-head (Kolodner and Tewari 1979; see also Section 8 and Fig. 8a). In an independent study on *S. olereacea* and other dicots 80% of chloroplast DNA molecules were found as monomer circles and 10 to 15% as dimers. About 15% of circles were supercoiled (Herrmann et al. 1975).

An elegant more recent study utilised fluorescence *in situ* hybridization (FISH) involving extended DNA fibres and plastid DNA probes (Lilly et al. 2001). At the time of writing this single report remains the only published source for FISH-based analysis of plastid DNA in flowering plants. Purified chloroplasts were lysed and DNA fixed directly on a slide before hybridization providing less opportunity for DNA breakage. Using this method, chloroplast DNA from *A. thaliana* and *N. tabacum* was found to be comprised of a multimeric series of circular and linear DNA molecules (Fig. 3b). Circles comprised about 40-50% and linear DNA species about 20-25% of chloroplast DNA molecules. The remaining molecules

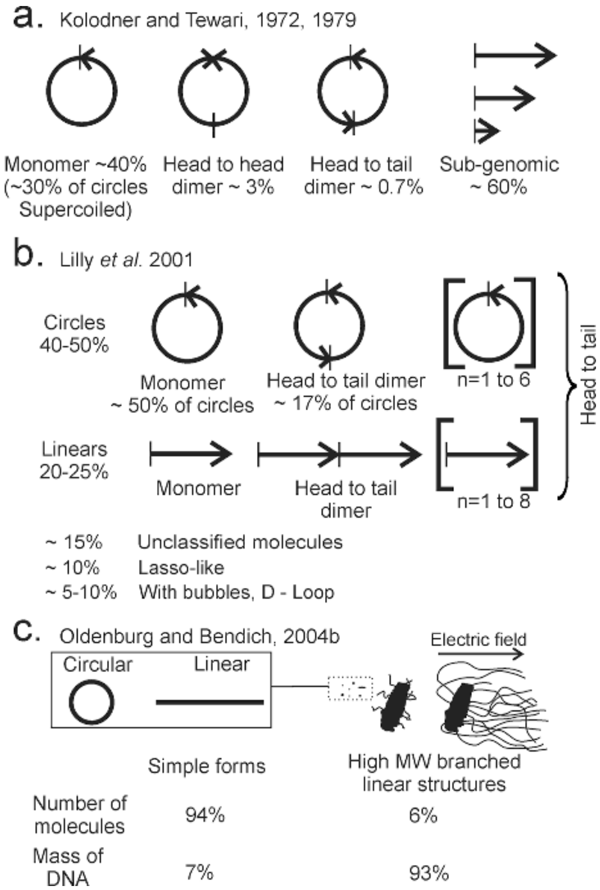


Fig. 3. a) Topological forms of plastid DNA revealed by electron microscopy (Kolodner and Tewari 1972). Only head-to-tail dimers were found in *P. sativum* while head-to-head dimers were predominant in *L. sativa* and *S. oleracea* (Kolodner and Tewari 1979). Breakage of circles during extraction will give rise to a variable percentage of circles and sub-genomic linear forms. b) Topological forms of plastid DNA in *A. thaliana* and *N. tabacum* revealed by DNA fibre-based FISH with plastid DNA probes (Lilly et al. 2001). Only the monomer and dimer are shown in the circular and linear multimeric series. Arrows indicate sequence orientation. c) Structures of plastid DNA molecules in fluorescent images of ethidium-stained DNA from purified *Z. mays* chloroplasts lysed in agarose plugs. DNA fibres in compact high MW structures were extended by an electric field or flowing liquid (Oldenburg and Bendich 2004b).

were unclassified (15%), lasso-like (~10%) or contained bubbles or D-loops (5-10%). In *N. tabacum* chloroplasts (Lilly et al. 2001), monomeric circles are the most abundant (55%) followed by dimers (17%), trimers (10%), tetramers (7%), pentamers (5%), and hexameric circles (1%). Rare higher-order multimers of unit genome sized (genome size = 156 kbp) linear DNA molecules were found up to the octomer. In *A. thaliana* and *N. tabacum* multimers were comprised of monomers arranged head-to-tail (Lilly et al. 2001) in contrast to the earlier results of predominantly head-to-head dimers in *L. sativa* and *S. oleracea* (Kolodner and Tewari 1979). The ends of the linear DNA molecules were not mapped and they could represent real linear plastid DNA species or the breakage products of large circles (Lilly et al. 2001). A number of lines of evidence indicate that linear DNA molecules found in plastids are not simply breakage products of large circles (see Section 3.1, 3.2, and 7 below). Using a similar DNA fibre-based FISH method, *C. reinhardtii* plastid DNA was found to be mainly comprised of monomeric and dimeric linear and circular forms (Maul et al. 2002). The multimeric series of linear and circular DNA molecules found in plastids (Fig. 3b) must result from the action of plastid DNA-RRR pathways.

Embedding cells in agarose plugs prior to cell and chloroplast lysis reduces DNA breakage and allows the isolation of large DNA molecules. DNA in agarose plugs can then be analysed by pulsed-field gel electrophoresis or microscopy after staining with ethidium bromide. Circular DNA does not enter pulsed field gels and remains within the agarose plugs at the origin at the relatively short pulse times used to fractionate DNA in the 100-1000 kbp range (Bendich and Smith 1990; Backert et al. 1995). Linear chloroplast DNA molecules enter the agarose gel and can be identified by blot hybridization with chloroplast DNA probes. Some of these linear DNA molecules might result from breakage of circular DNA molecules (Backert et al. 1995; Bendich 2004). A multimeric series comprised of monomer (most abundant) and higher molecular weight (MW) linear plastid DNA forms can be visualised on pulsed-field gels. The largest multimers found were tetramers for *S. oleracea* (Deng et al. 1989) and *N. tabacum* (Lilly et al. 2001), dimers (Lilly et al. 2001) or trimers for *P. sativum*, and up to the octomer for *Citrullus vulgaris* plastid DNA (watermelon, Bendich and Smith 1990). The banding pattern can be disrupted by altering the activities of plastid DNA-RRR proteins. Inhibition of plastid-targeted gyrase (see Section 13.4 below), which is required to decatenate newly replicated DNA, reduces the levels of discrete bands corresponding to the monomer and dimer, and gives rise to a heterogeneous mixture of plastid DNA molecules, some of which are greater than 1000 kbp in size on pulsed field gels (Cho et al. 2004).

Whilst the bands seen on pulsed-field gels were useful for visualising multimers of plastid DNA they represent a minor proportion of plastid DNA and give a distorted view of the topological forms of plastid DNA molecules (Bendich 2004). The bulk of plastid DNA molecules including circles, high MW linear branched forms (Bendich 2004; Oldenburg and Bendich 2004b), tangled DNA fibres and any DNA in unlysed plastids remains immobile in the agarose plugs at the origin and does not enter pulsed field gels. *N. tabacum* leaf chloroplast DNA remained at the origin (migration into the gel was not detected) whereas about 35% of *Chenopidium album* plastid DNA from a non-green suspension-culture entered the gel revealing monomer and dimer bands (pulse times of 30-60 seconds, Backert et al. 1995). The presence of electrophoretically-mobile linear plastid DNA molecules in *C. album* non-pigmented plastids but not in *N. tabacum* chloroplasts might reflect changes in plastid DNA topologies in different plastid types, also indicated from other studies on *Z. mays* (Oldenburg and Bendich 2004a), or result from breakage during extraction. Mild DNase I treatment of high MW *N. tabacum* chloroplast DNA and blot hybridization with a plastid probe revealed a smear of DNA (representing molecules of different lengths) within which discrete monomer to tetramer bands were clearly visualised (Backert et al. 1995). These discrete linear bands are likely to be derived from circular DNA because a single double strand break mediated by DNase I will convert a circle to its linear form.

The structures of plastid DNA molecules in agarose plugs prepared from 10-14 day old *Z. mays* seedlings have been studied by fluorescence microscopy following ethidium bromide staining (Oldenburg and Bendich 2004b). The DNA was present as simple DNA molecules and high MW DNA complexes with a central core and attached DNA fibres (Fig. 3c). In the presence of an electric field or liquid flow the simple molecules migrate whereas linear fibres extend from the immobile cores of the high MW complexes. Simple DNA molecules are comprised of circles and linear molecules and represent 94% of the DNA molecules but only 7% of the mass of DNA in plastids due to their small sizes relative to the high MW DNA complexes. The high MW complexes contained on average a minimum of eight plastid genomes (not including bright fluorescent cores) and were suggested to be largely comprised of linear and complex-branched molecules (Oldenburg and Bendich 2004b). A reduction in high MW complexes and an increase in simple forms were correlated with chloroplast maturation during leaf development in *Z. mays* (Oldenburg and Bendich 2004a). The multigenome complexes were reported to represent 93% of plastid DNA by mass. Following removal of linear DNA molecules from multigenome complexes by pulsed-field gel electrophoresis the immobile high MW core was suggested to be comprised of complex-branched DNA structures representing 50% of the mass of DNA in plastids (Oldenburg and Bendich 2004b). The complex high MW branched forms have been suggested to represent replication intermediates and their analysis is particularly important (Oldenburg and Bendich 2004b; Scharff and Koop 2006). The ~15% of tangled DNA fibres that were unclassified (Fig. 3b) by Lilly et al. (2001) might correspond to these high MW DNA complexes identified by Oldenburg and Bendich (2004b). DNA fibre-based FISH using plastid DNA probes would confirm the presence of plastid DNA in these high MW complexes and might be a useful tool to study their sequence organisation.

3.1 Linear hairpin DNA molecules in plastids

Genuine linear plastid DNA molecules can be distinguished from linear products of broken DNA circles by studying their ends. Breakage products would be expected to possess ends that map to randomly selected regions of the plastid genome, and these ends would be expected to be indistinguishable from double-strand breaks with flush or short single-stranded 5’ or 3’ DNA extensions. Analysis of plastid DNA deletion mutants in albino cereal plants regenerated from pollen provided the first evidence for the presence of linear plastid DNA molecules with special ends (Day and Ellis 1985; Ellis and Day 1986). Small linear subgenomic molecules have also been found in albino somatic cells from cereals (Kawata et al. 1997; Zubko and Day 2002) and can represent the predominant plastid DNA species in albino cereal plants. Their abundance facilitates the analysis of their ends, which have been examined in detail.

Small linear plastid DNA molecules are inverted repeat palindromes with hairpin ends, which map to a number of sites in the large single copy region of plastid DNA near the *trn*E(UUC) gene (Ellis and Day 1986; Harada et al. 1992; Kawata et al. 1997). The centres of a subset of these linear palindromes are located between *trn*G(GCC) and *trnf*M(CAU) of the 135 to 140 kbp cereal plastid genome (Ogihara et al. 2000) and retain only 5.2 kbp of plastid DNA (Zubko and Day 2002). Small linear DNA molecules all contain the plastid *trn*E(UUC) gene, which is probably essential for heme synthesis (Howe and Smith 1991; Zubko and Day 2002). Linear hairpin DNA molecules are also found in eubacteria including the spirochete genus *Borrelia* (Casjens 1999) and prophage N15 of *Escherichia coli* (Rybchin and Svarchevsky 1999). Models for the origin of these linear DNA molecules include: strand switching during DNA replication (Ellis and Day 1986), possibly promoted by short inverted repeats (Fig. 4a); repair of double strand DNA breaks by intra-strand annealing at inverted repeats (Fig. 4b; Qin and Cohen 2000); and an *E. coli* linear prophage N15-like mechanism involving two cleavages, sealing DNA ends to form hairpins and resolution of replicated DNA into a linear palindrome (Fig. 4c; Rybchin and Svarchevsky 1999). Hairpin ends provide a mechanism to overcome the end-replication-problem and stabilise the ends of linear DNA molecules by protecting them from nucleases (Cavalier-Smith 1974).

Studies on albino cereal plants demonstrate that plastids contain the enzymes required to maintain and replicate linear DNA molecules. If hairpin molecules play a role in maintenance of intact plastid DNA they should also be found in the green chloroplasts of WT plants. Revealingly, hairpin molecules are found in WT *H. vulgare* chloroplasts. The hairpin ends do not appear to be localised but map to various sites within the plastid genome (Collin and Ellis 1991), which is consistent with their derivation from a population of linear DNA molecules with heterogeneous ends. Dispersed ends that are not defined in location are also found in the heterogeneous populations of linear DNA molecules in plant (Backert and Börner 2000; Oldenburg and Bendich 2001) and *Saccharomyces cerevisiae* (bakers' yeast) mitochondria. Mitochondrial DNA in *S. cerevisiae* is comprised of a polydisperse population of linear DNA molecules ranging in size between the 75 kb monomer and 150 kbp dimer (Williamson 2002).

3.2 Linear plastid DNA molecules with discrete ends in WT plastids

Restriction enzymes that cleave plastid DNA rarely (once or twice) have been used to map the ends of linear molecules in high MW DNA prepared in agarose plugs. When *Z. mays* plastid DNA was cleaved with an enzyme that cuts once, the predicted linear 140 kbp genome band was observed. In addition, discrete smaller

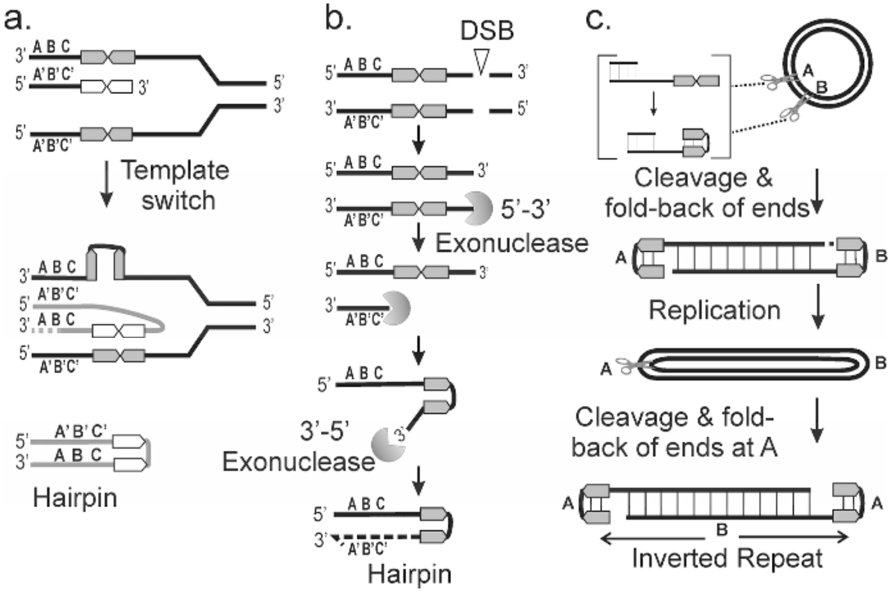


Fig. 4. Models for generation of linear hairpin plastid DNA molecules by a) Template strand switching at the replication fork (Ellis and Day 1986), b) Intra-strand annealing at inverted repeat sequences (Qin and Cohen 2000). A double-strand break (DSB) initiates the pathway, which involves exonuclease and fold-back of single-stranded DNA at inverted repeats. c) A bacteriophage N15-like mechanism involving DNA cleavage at two sites, fold-back and repair of ends to form hairpins and replication followed by cleavage at one site to form a linear palindrome (Rybchin and Svarchevsky 1999). Converging box arrows indicate inverted repeats.

sub-genomic bands were found (Oldenburg and Bendich 2004b). These sub-genomic bands can be explained if they arise from long linear DNA molecules containing one natural end found *in vivo* and one site created by restriction enzyme cleavage. These natural ends in sub-genomic fragments map to the large inverted repeats of *Z. mays* plastid DNA. A similar but more detailed analysis on high MW *N. tabacum* plastid DNA identified eleven natural ends (Scharff and Koop 2006). The majority of breaks mapped to the large inverted repeats but ends were also found in the large and small single copy regions (Scharff and Koop 2006). Some of the *Z. mays* and *N. tabacum* ends map close to plastid DNA sequences promoting DNA synthesis or exhibiting features resembling D-loops or replication bubbles (see Section 4.1 below). Only one end corresponded to the proposed site for initiation of rolling circle replication (Kolodner and Tewari 1975) located at 180˚ from the two D-loops (Section 4; Fig. 5a) in *N. tabacum* (Scharff and Koop 2006). These mapped ends define the termini of linear DNA present in high MW plastid DNA complexes (Oldenburg and Bendich 2004b; Scharff and Koop 2006). The structures of the ends of these linear sections of *Z. mays* and *N. tabacum* plastid DNA are not known but their elucidation (e.g. protected or exposed, hairpin or secondary DNA structure, or simply double-strand

DNA breaks with flush or 5' or 3' protruding ends) is likely to provide information on the mechanisms underlying their formation.

4 A replicon model for plastid genome maintenance

Research on plastid DNA replication has been heavily influenced by the 'replicon model' put forward by Jacob, Cuzin, and Brenner (Jacob et al. 1963). The model proposes a specific DNA element that is recognised by an initiator protein. If plastid DNA replication conforms to the model it would predict initiation of replication at specific sites in plastid DNA. Replication of bacterial genomes and plasmids in bacteria and *S. cerevisiae* conform to the 'replicon model' and involve origins of replication recognised by specific origin recognition proteins (Gilbert 2004). A variety of methods have been used to try and localise origins of replication in plastid genomes. Early electron microscopy (EM) studies on DNA isolated from *P. sativum* and *Z. mays* chloroplasts identified structures resembling D-loops and rolling circles (Kolodner and Tewari 1975). These are well known DNA replication intermediates and provided early models for plastid DNA replication. Unidirectional replication from an origin of replication creates a displacement loop (D-loop), comprised of double stranded DNA and a displaced single stranded DNA loop. Two D-loops spaced 7 kbp apart were found in monomer circles of *P. sativum* plastid DNA and gave rise to the dual D-loop model (Fig. 5a) for initiation of chloroplast DNA replication (Kolodner and Tewari 1975). Convergent replication forks from the two D-loops pass each other and a bidirectional replication bubble is formed once the forks pass the starting points of replication. The complete genome is replicated by the replication forks continuing round the circle in opposite directions, with discontinuous replication on the lagging strands, until they meet at 180^{o} from the origin of D-loop synthesis (Fig. 5a). For rolling circle replication, a replication fork displaces the lagging strand at a nick and continues round the circle (Fig. 5b). Rolling circle replication enables multiple tandem head-to-tail copies of plastid DNA to be made from a single round of replication initiation. The rolling circles appeared to be initiated at the terminus of bidirectional replication (Kolodner and Tewari 1975). This early EM study stimulated research to locate D-loops on sequence maps of plastid DNA.

4.1 Replication origins mapped to the large inverted repeat

Electron-microscopy combined with restriction enzyme digestion enables D-loops and replication bubbles to be mapped onto restriction fragments of plastid DNA. The *P. sativum* D-loops (OriA and OriB) flank the 23S ribosomal RNA gene (Fig. 6a; Meeker et al. 1988). Unlike most angiosperm plastid genomes *P. sativum* lacks a large inverted repeat (Chapter 3). Restriction fragments of proplastid DNA with a high frequency of D-loops from *N. tabacum* BY2 suspension culture cells

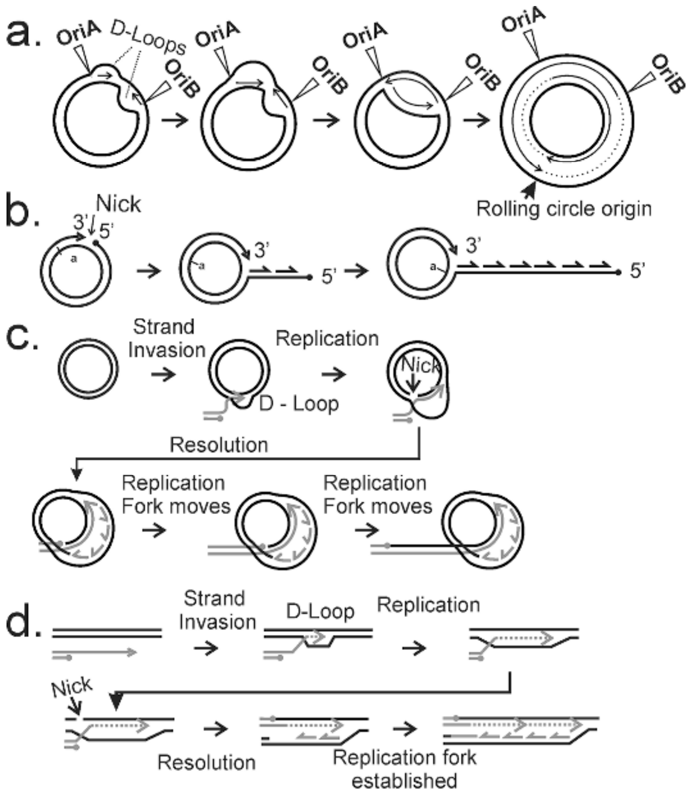


Fig. 5. Plastid DNA replication models. a) Displacement-loop (D-loop) model of plastid DNA replication. Two D-loops converge to give rise to bidirectional replication (Kolodner and Tewari 1975). b) Rolling circle replication arising from strand displacement at a nick. Movement of the replication fork is shown by anti-clockwise rotation of the circle marked by "a" c) Recombination-dependent DNA replication (Kowalczykowski 2000) on a circular template with D-loop gives rise to a bubble-containing circle with tail. d) Recombination-dependent DNA replication on a linear DNA template gives rise to a branched molecule.

mapped close to the end of the 23S rRNA gene in the large inverted repeat (Fig. 6b Nt (pro)) and a less active D-loop mapped to a 2.3 kbp Stu I fragment containing part of the *psaA* and *psaB* genes in the large single copy region (Takeda et al. 1992). Later work, using two dimensional agarose gel electrophoresis to map bubbles in cloned plastid DNA templates replicated in chloroplast fractions, *in vitro* DNA replication assays and primer extension on nascent DNA strands, suggested different positions for two D-loops (named OriA and OriB) in plastid DNA from *N. tabacum* leaves. *N. tabacum* OriA mapped to the intron of the *trnI* (GAU) gene located between the 16S and 23S rRNA genes (Lu et al. 1996; Kunnimalaiyaan and Nielsen 1997a). OriB mapped to the large inverted repeat close to the border of the small single copy region in orf350 or *ycf1* (Kunnimalaiyaan and Nielsen 1997a; Kunnimalaiyaan et al. 1997b).

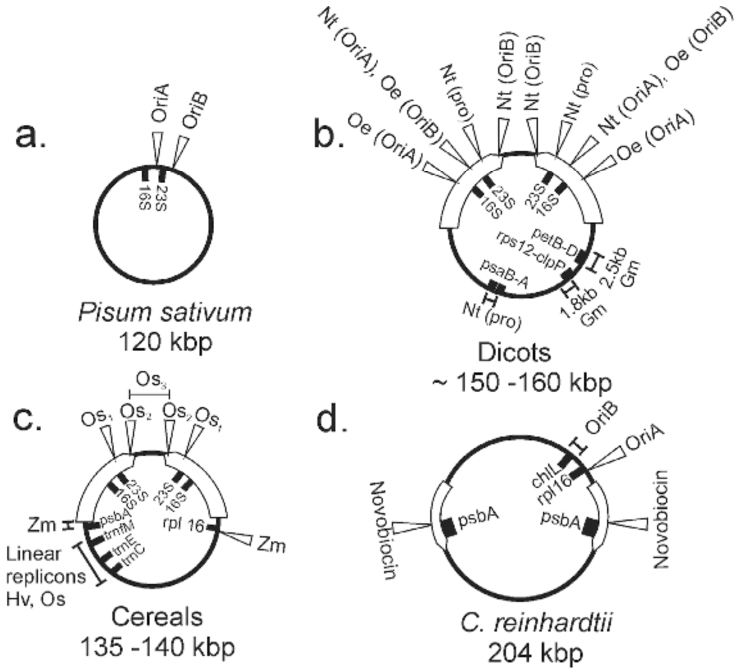


Fig. 6. Schematic diagram showing the locations of potential replication origins. Origin locations are shown outside the circular maps as triangles or bar-ended lines. a) *P. sativum* (Meeker et al. 1988) lacks a large inverted repeat, b) Dicots containing large inverted repeats. *N. tabacum* Nt (Ori A) and Nt (Ori B) (Kunnimalaiyaan and Nielsen 1997; Kunnimalaiyaan et al. 1997), *N. tabacum* D-loops Nt (pro) in proplastids (Takeda et al. 1992), *Oenothera hookeri* Oe (Ori A) and Oe (Ori B) (Chiu and Sears 1992; Sears et al. 1996), *Glycine max* bubbles Gm (Hedrick et al. 1993). c) *Oryza sativa* (Os) replication origins in suspension culture cells (Os_1), leaf blades (Os_2), and coleoptiles (Os_3) mapped by Wang et al. 2003, *Z. mays* (Zm, Gold et al. 1987), linear DNA replicons (Ellis and Day 1986; Harada et al. 1992; Zubko and Day 2002), *Hordeum vulgare* (Hv), *Oryza sativa* (Os). d) *C. reinhardtii* Ori A and Ori B (Waddell et al. 1984; Chang and Wu 2000) and initiation of novobiocin-resistant replication (Woelfle et al. 1993). The large inverted repeats are shown as converging box arrows on maps b-d; arrow orientation according to rrn operon transcription direction.

Two D-loops separated by 4 kbp (Chiu and Sears 1992; Sears et al. 1996) were found in the large inverted repeat of *Oenothera hookeri* plastid DNA, where they flank the 16S ribosomal genes (Fig. 6b). The locations of two origins in the large inverted repeat suggest four potential replication origins in *N. tabacum* and *O. hookeri* plastid DNA; two in each inverted repeat. The complexity of mapping plastid origins is illustrated by the observation that the locations of origins appear to vary in different cells, tissues and organs from the same species. Differences in the mechanism of DNA replication and locations of plastid origins have been observed in suspension culture cells, coleoptiles, and leaves of *Oryza sativa* (rice, Wang et al. 2003). *O. sativa* plastid replication origins were mapped to the small single copy region and to two positions in the large inverted repeat (Fig. 6c; Wang et al. 2003). Comparisons of mapped origins show that the location of only one is conserved in *N. tabacum* (OriA), *O. hookeri* (OriB), and *P. sativum* (OriA). This is located in the intergenic region between the 16S and 23S rRNA genes (Fig. 6a, 6b; Lu et al. 1996). Conservation in location might suggest this region is important for plastid genome maintenance. However, deletion of OriA in *N. tabacum* using plastid transformation has revealed that it is not essential for plastid DNA replication and maintenance (Mühlbauer et al. 2002). Of the two copies of OriB present in the large inverted repeat the one located in orf350 (OriB2) could be deleted. The other copy of OriB (OriB1) cannot be removed without mutating the essential *ycf1* gene and hence OriB1 dispensability cannot be addressed by a deletion that removes *ycf1* function (Mühlbauer et al. 2002). In a recent study, a stem-loop in OriB1 was mutated that left *ycf*1 intact (Scharff and Koop 2007). This allowed the isolation of plants in which OriB1 was mutated and OriB2 was deleted indicating that neither OriB sequence was essential. The copy number of plastid DNA appeared to be the same in shoot tips but lower in young and older leaves of deleted OriA lines (down ~1.5-fold) and lines lacking both OriA and OriB2 (down ~2-fold) compared to WT. The plastid DNA copy number in young and older leaves of OriB mutated lines (OriB2 deleted, OriB1 mutated) was higher (up ~1.7-fold) than WT plants (Scharff and Koop 2007).

4.2 Replication origins located in the single copy regions

Potential replication origins have also been located well away from the ribosomal genes. In *C. reinhardtii* two D-loops (*Ori*A and *Ori*B) spaced 7 kb apart (Waddell et al. 1984) map to the single copy region of plastid DNA (Fig. 6d). *C. reinhardtii* OriA was localised to a 224 bp region containing the *rpl16* gene (Chang and Wu 2000) whereas ”OriB is located in or adjacent to *chl* L” (Wu personal communication). Replication at OriA is influenced by transcription across *rpl16* (Chang and Wu 2000). Novobiocin inhibition of D-loop replication led to novobiocin-resistant replication starting close to a ‘hot spot’ of recombination near the 3’ end of the *psbA* gene in the large inverted repeat of *C. reinhardtii* plastid DNA (Woelfle et al. 1993). In *Z. mays*, preferential DNA synthesis from a 1368 bp plastid DNA sequence amongst cloned templates representing 94% of the *Z. mays* plastid genome by a partially purified *P. sativum* plastid DNA polymerase suggested this region promoted replication (Gold et al. 1987). The *Z. mays* region promoting DNA synthesis was found to promote bidirectional replication using a partially purified ~90 kDa *Z. mays* DNA polymerase. This was localised to a 455 bp sequence containing the 3’ end of the *rpl16* gene in the large single copy region of plastid DNA (Fig. 6c; Carrillo and Bogorad 1988). A one kbp region including the 3’ end of the *psbA* gene also promoted DNA synthesis using the *Z. mays* DNA polymerase fraction (Fig. 6c; Carrillo and Bogorad 1988). Whilst the sequence organizations of *Z. mays* and *C. reinhardtii* plastid DNA are not conserved it is interesting that a potential replication origin overlaps with the *rpl16* genes in both species (Gold et al. 1987; Chang and Wu 2000).

Two replication bubbles were mapped to 1.8 kbp *Sac* I-*Bam* HI and 2.5 kbp *Bam* HI fragments in the large single copy region of *Glycine max* (soybean) chloroplast DNA by two-dimensional gel electrophoresis (Fig. 6b; Hedrick et al. 1993). The recently published *G. max* plastid genome (Saski et al. 2005) locates these bubble-containing fragments to regions containing the *rps12-clp* P1 and *pet* B-*petD* genes in the large single copy region. *Petunia hybrida* and *N. tabacum* plastid DNA sequences located in the small copy region near the location of *N. tabacum Ori*B promote autonomous replication in yeast nuclei (Ohtani et al. 1984; Dehaas et al. 1986). Although interesting, the finding that some A + T rich plastid sequences resembling yeast autonomously replicating elements (ARS) promote replication in *S. cerevisiae* nuclei would appear to provide weak evidence for locating plastid origins of replication. The locations of these ARS regions are not shown in Figure 6. The lack of agreement between different experimental approaches in locating origins of replication (Fig. 6) has hindered progress in this difficult area of research. It might indicate a lack of conservation in the location of plastid replication origins between different species, the presence of multiple origins, differences in the accuracy of locating origins or limitations of the plastid dual D-loop replication model.

5 Maintenance of small DNA molecules in plastids

Promotion of autonomous replication of plasmids within plastids themselves would provide a function-based assay in a homologous system for locating plastid origins of replication. Free monomeric plasmids are known to persist for a short period after transformation in *C. reinhardtii* (Boynton et al. 1988). This persistence might reflect replication and/or excision of complete plasmids integrated by single recombination events into the plastid genome. Homologous recombination between 16 bp direct repeats gave rise to an 868-bp DNA minicircle in *N. tabacum* plastids that was unstable but persisted as a multimeric series for several months during growth and development of transplastomic plants (Staub and Maliga 1994). The 868 bp excised sequence extends from the *trnI*(GAU) intron to the *trnA*(UGC) intron in the large inverted repeat and is located about 300 bases from the 82 bp region containing Ori A (Kunnimalaiyaan and Nielsen 1997a).

In *C. reinhardtii*, rescue of an *atpB* photosynthetic mutant with a partial function *atpB* allele led to amplification of the transforming plasmid in the form of large tandem arrays that appeared to be episomal (Suzuki et al. 1997). The maintenance of these episomal tandem arrays in plastids required a region of homology with the resident plastid genome. Revealingly, plastid transformation was not observed if there was no homology between resident plastid genome and transforming plasmid. Moreover, the inclusion of *C. reinhardtii* Ori A promoted integration rather than persistent autonomous replication (Suzuki et al. 1997). These results demonstrate that small plasmids containing mapped plastid replication origins do not contain sufficient sequence information for replication and partition to exist as stable autonomous replicons in plastids. Recombination appears to play a role in plasmid maintenance in plastids and this may also be true of WT plastid genomes (see Section 7 below).

6 Deletion mapping delimits DNA sequences capable of self-replication in plastids

In cereals, most plastid genes are dispensable allowing the isolation of deletion mutants lacking most of the plastid genome (see Section 3.1 above). This natural deletion mapping identifies a region of plastid DNA located around the *trnE*(UUC) gene (Fig. 6c) that is self-replicating as linear DNA molecules (Day and Ellis 1985; Ellis and Day 1986; Harada et al. 1992; Kawata et al. 1997; Zubko and Day 2002; Cahoon et al. 2003). The retained region does not contain any known plastid replication origins mapped by electron microscopy, gel electrophoresis or biochemical methods (Fig. 6b, 6c). Maintenance of these linear DNA molecules might involve a replication origin that is not highly active in WT shoots and leaves. The relationship between the replication mechanisms maintaining small linear molecules and WT plastid DNA is not known but hairpin ends have been found at low frequency in WT *H. vulgare* plastids (Collin and Ellis 1990; Section 3.1).

Recombination events between sequences in the large single copy region and those located in either the large inverted repeat or small single copy region give rise to deleted plastid DNA molecules with circular maps (Day and Ellis 1984) and circular structures (Day and Ellis 1985) that have been found in albino *T. aestivum* plants from anther culture. The region present in the smallest circles (39 kbp, Day and Ellis 1985), containing only 30% of the plastid genome, stretches from the *trnE* gene region to the end of the adjacent large inverted repeat (Fig. 6c). These molecules contain only one large inverted repeat sequence and lack the replication origin mapped near *rpl16* in *Z. mays* (Fig. 6c; Gold et al. 1987). Deleted circular plastid DNA molecules containing only one large inverted repeat have also been found in WT *N. tabacum* chloroplasts by DNA fibre-based FISH (Lilly et al. 2001). These results show that sub-genomic plastid DNA molecules lacking the small single copy region, most of the large single copy region and one large inverted repeat can be maintained as circular DNA molecules in plastids.

7 A recombination-dependent DNA replication model of plastid DNA

The replicon model (Jacob et al. 1963) has been very successful and been substantiated in bacteria, animal viruses and budding yeasts such as *S. cerevisiae*. Problems in localising *bona fide* replication origins in plastids (Section 4.1, 4.2) and the nuclear genomes of multicellular organisms (Gilbert 2004) have hindered universal application of the replicon model. The apparent failure of plastid genomes to conform to the replicon model with one or two well-defined replication origins might suggest the standard model (Fig. 5a, 5b) for replication of plastid genomes (Kolodner and Tewari 1975) requires revision (Bendich 2004). The lack of progress in understanding plastid genome replication has been matched by illuminating advances in bacterial genetics, particularly by Kogoma and colleagues (Asai et al. 1994), that have identified new replication mechanisms initiated by recombination for genome maintenance. Recombination-dependent DNA replication allows stalled replication forks at double-strand DNA breaks to re-establish and enables initiation of replication in the absence of a defined origin of replication (Asai et al. 1994; Kowalczykowski 2000). The mechanism requires a linear DNA end and strand invasion to prime DNA synthesis on a circular (Fig. 5c) or linear DNA (Fig. 5d) template followed by resolution of the recombination-junction to re-establish the replication fork. Replication forks resulting from D-loops primed by strand invasion were first described in bacteriophage T4 DNA replication (Kreuzer 2000, 2005). The role of recombination in maintaining replication forks has been suggested to be the main function of recombination systems (Cox et al. 2000; Goodman 2000). Plastids are known to contain a highly active homologous recombination pathway (see Section 8 below), which is a requirement for recombination-dependent DNA replication.

The only requirement for recombination-dependent DNA replication is a free end that could be located at any position on plastid DNA. As mentioned above (Section 3.1), the linear genomes in *S. cerevisiae* and plant mitochondria appear to have heterogeneous ends rather than a limited number of defined ends. Recombination-dependent DNA replication has been put forward as a mechanism for replication of fungal and plant mitochondrial genomes (Oldenburg and Bendich 1998; Williamson 2002), and the minicircles present in the plastids of dinoflagellates (Nelson and Green 2005). Linear ends with 3’ overhangs would allow strand invasion to prime DNA replication on other plastid DNA molecules acting as templates. Induction of double-strand breaks at specific sites in *E. coli* gives rise to recombination-dependent DNA replication origins that can be mapped (Asai and Kogoma 1994). Linear DNA molecules with defined ends have been found in *Z. mays* (Oldenburg and Bendich 2004b) and *N. tabacum* (Scharff and Koop 2006, 2007) plastids (Section 3.2 above). One common end in both species maps close to OriA in the large inverted repeat of *N. tabacum* between the 16S and 23S rRNA genes (Fig. 6b). An origin of replication has not yet been located in this position in cereal plastid genomes (Fig. 6c). These natural ends have been suggested to invade template DNA and prime DNA synthesis by recombination-dependent DNA replication (Bendich 2004; Oldenburg and Bendich 2004b). As a result they could define sites at which replication of the plastid genome is initiated. Interestingly, the locations of these ends appeared to change when OriA and OriB2 were deleted (Scharff and Koop 2007). Recombination-based-replication will lead to circular DNA molecules with tails (Fig. 5c) and linear branched structures (Fig. 5d). Linear DNA molecules are extended when their ends invade and replicate template genomes. The process leads to multimerization of linear DNA molecules and has been documented in detail during replication of bacteriophage T4 in *E. coli* (Kreuzer 2000). Highly branched complexes will be formed if several independent DNA molecules are connected by recombination-dependent DNA replication events. The complexity of branching increases if resolution (see resolution step shown in Fig. 5c, 5d) is not completed in some of the branches (Kreuzer 2000). Complex branched networks arising from recombination-dependent DNA replication might explain the 90-95% (by mass) of *Z. mays* plastid DNA found in high MW complexes (Oldenburg and Bendich 2004b; Section 3; Fig. 3c).

8 DNA recombination in plastids

Plastid fusion and DNA recombination between different plastid types are rare in flowering plants. Rapid segregation is observed when two plastid types with different genomes are forced into the same cell by protoplast fusion (Morgan and Maliga 1987). In *C. reinhardtii*, recombination between parental plastid genomes in exceptional zygotes is well established (Gillham 1974). The development of plastid transformation has demonstrated an active homologous DNA recombination pathway in *C. reinhardtii* (Boynton et al. 1988) and flowering plant plastids (Svab et al. 1990). The rarity of plastid fusion in angiosperms probably explains the lack of DNA recombination between ”parent” plastid genomes in protoplast fusion experiments. In one successful protoplast fusion experiment a single plant with a recombinant plastid genome resulting from at least six crossover events between parental genomes was isolated (Medgyesy et al. 1985).

Most characterised plastid genomes contain a large inverted repeat sequence. Recombination between the large inverted repeat sequences (flip-flop recombination) is responsible for the two isomers of plastid DNA, which differ with respect to the orientation of the single copy regions (Palmer 1983). Flip-flop recombination giving rise to the two isomers can take place between circular (Fig. 7a) or linear DNA substrates (Fig. 7b). The head-to-head circular dimers (Fig. 3a) in *L. sativa* and *S. oleracea* plastids observed by Kolodner and Tewari (1979) were explained by intermolecular recombination between opposite large inverted repeats in circular DNA substrates (Fig. 8a). These head-to-head dimers are comprised of an inverted sequence representing ~90% of the unit genome size separated by small spacer loops comprised of the small single copy sequences. *P. sativum* plastid DNA lacks a large inverted repeat providing an explanation for the lack of head-to-head dimers in plastids from this species (Kolodner and Tewari 1979). Head-to-head inverted sequences representing ~90% of the unit genome length will also be produced by recombination events between large inverted repeat sequences involving linear DNA substrate (Fig. 8b). Homologous recombination that is not limited to specific sequences appears to be responsible for generating these isomers. Intermolecular recombination between inverted repeats in long chain multimers of plastid DNA would be expected to place at any point and would give rise to a large number of isomers. Intramolecular recombination events between tandemly repeated copies of the unit genome in linear multimers will give

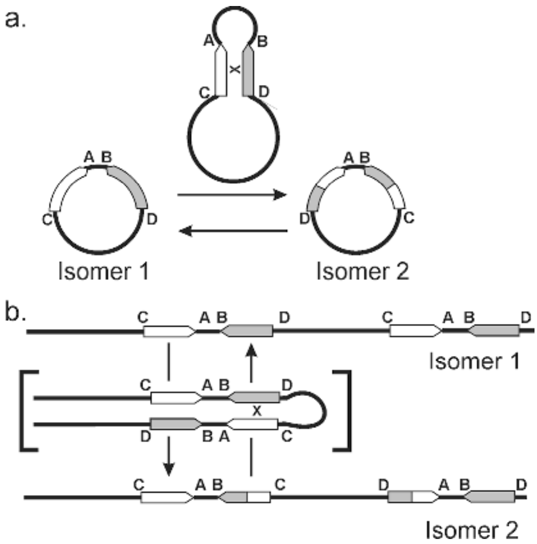


Fig. 7. Intramolecular flip-flop recombination between large inverted repeat sequences in a) Circular, and b) Linear DNA molecules. The large inverted repeat sequences are shown as converging grey and white box arrows. Note that the linear product in b contains one of the head-to-head inverted sequences high-lighted in Fig. 8b.

rise to circular DNA molecules. Oldenburg and Bendich (2004b) have pointed out that recombination-dependent DNA replication primed by an end within the large inverted repeat will also result in head-to-head inverted sequences and flipping of single copy regions. This can be visualised by looking at Figure 8b where the products of reciprocal recombination can also be obtained by strand-invasion by the top molecule on the bottom template at the crossover site followed by D-loop replication, resolution, and replication fork movement (see Fig. 5d) to the end of the template molecule.

Figure 9 shows a recombination event between large inverted repeats following replication of one copy of the repeat (Futcher 1986). This switches the direction of the replication fork allowing many identical head-to-tail copies of the unit genome to be made in a multimeric circle from a single template without re-initiation of DNA replication. The absence of a large inverted repeat in some plastid genomes (Palmer and Thompson 1982) would suggest that this double rolling circle mechanism is not essential for amplification of plastid DNA. A linear multimeric chain replicated from a circular template will be formed if the lagging strand is not replicated following recombination between the duplicated and unreplicated copies of the large inverted repeats (Ellis and Day 1985).

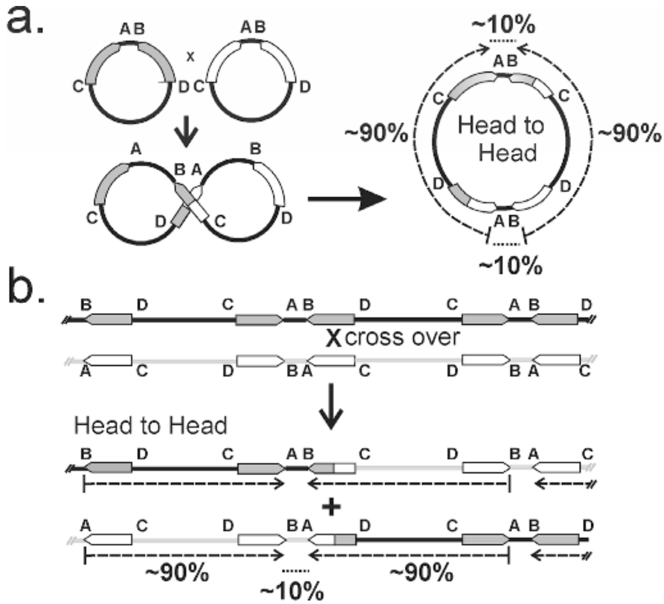


Fig. 8. Intermolecular recombination between opposite large inverted repeat sequences gives rise to plastid genomes orientated head-to-head. a) Circular head-to-head and b) Linear head-to-head DNA molecules. Head-to-head inverted sequences are shown as dotted arrowed lines. Length is expressed as a percentage of the unit genome size.

8.1 Integration of foreign genes by homologous recombination

The era of plastid transformation enables the detailed study of recombination events in plastids. In algae and plants foreign DNA integrates by homologous recombination between common DNA sequences in the transforming vector and resident plastid genome. Large regions of donor plastid DNA integrate into the resident plastid genome well beyond the markers used to select transformants resulting in incorporation of all or almost all of a 6.2 kb recombinant plastid sequence in *N. tabacum* plastid transformants (Staub and Maliga 1992). Reciprocal recombination or gene conversion events between transforming plasmid and resident plastid genome will result in transgene integration (Fig. 10a). A variety of plastid DNA sequences have been used to target integration of foreign genes to different sites in the plastid genome (Chapter 14). This indicates the plastid recombination machinery is not limited to specific substrates but can act on a wide selection of DNA sequences. When a vector containing a gene-of-interest flanked by targeting DNA is introduced into angiosperm plastids, double recombination events in both arms will insert the transgene into the plastid genome. Alternatively, the entire plasmid can integrate as a result of homologous recombination in

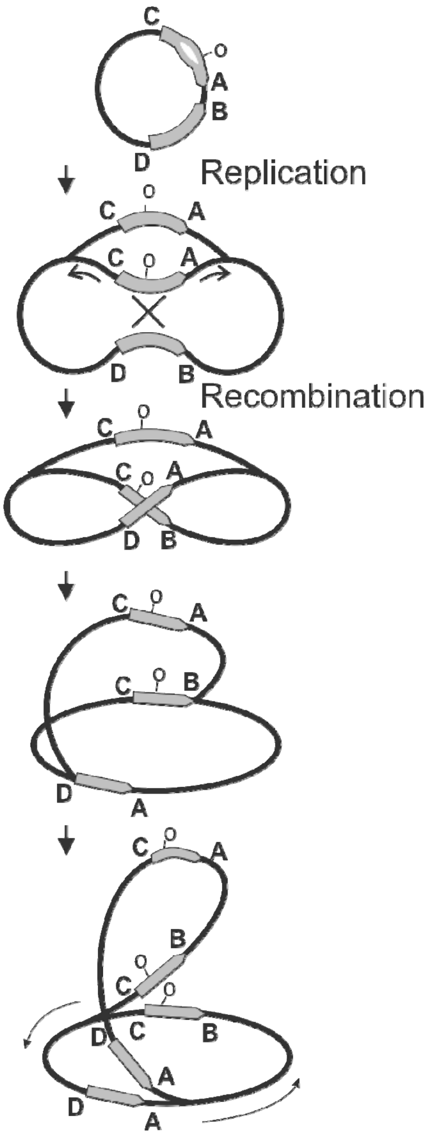


Fig. 9. Double rolling circle model of DNA replication gives rise to large circular multimers from a single round of replication initiation (Futcher 1986). Not to scale: the small single copy region is enlarged to illustrate the model. O= replication origin.

one flanking region of plastid DNA. The resulting co-integrate contains duplications of left and right targeting regions. The co-integrate can be selected by placing the marker gene in the vector sequences (Klaus et al. 2004). When selection is removed further homologous recombination events between these duplicated sequences excise vector and marker sequences to leave either the gene-of-interest or

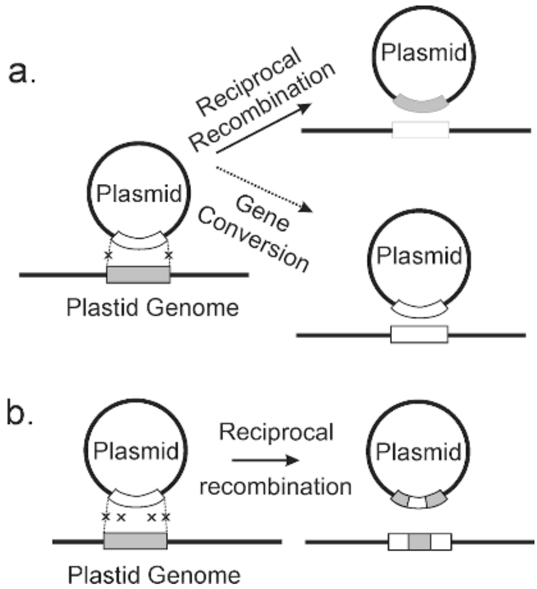


Fig. 10. Integration of foreign DNA into the plastid genome by homologous recombination. a) Integration requires homology and involves reciprocal recombination and possibly gene conversion events. b) Integration of homeologous DNA reveals multiple crossover events between target and donor DNA (Kavanagh et al. 1999).

WT plastid sequences (see Chapter 14; Klaus et al. 2004). These co-integrate experiments show that the crossover events in both arms required for integration are not tightly linked and this integration pathway must reflect properties of the DNA-RRR pathways in angiosperm plastids.

In *E. coli,* homologous recombination is stimulated by 8 base 5'GCTGGTGG chi sequences which are recognised by the *recBCD* complex (Kowalczykowski 2000). The eight base chi motif is absent in *N. tabacum* plastids (Shinozaki et al. 1986) but is present in the 16S ribosomal RNA genes in grass plastid genomes (Hiratsuka et al. 1989). Any role for chi sequences in plastids would appear to be ruled out by the finding that cyanobacterial and angiosperm genomes do not appear to contain homologues of genes encoding the *recBCD* complex. In *C. reinhardtii,* sequences that appear to stimulate recombination have been identified within the large inverted repeat and have been localised to a 400 bp region of plastid DNA containing the 3' end of the *psbA* gene (Newman et al. 1992).

8.2 Homologous recombination between short DNA repeats

The introduction of short repeated sequences into plastid genomes by transformation has demonstrated that they are effective substrates for homologous recombination (Table 1). Recombination between direct repeats excises the intervening

Table 1. Recombination events between engineered DNA repeats in transgenic plastids

Species	Repeat length (bp)	Repeated sequence	Reference
Nicotiana tabacum	174[1]	16S rrn promoter & *rblC* RBS	Iamtham and Day 2000
	418[1]	*psbA* 3' UTR	Iamtham and Day 2000
	210	*rbcL* 3' UTR	Dufourmantel et al. 2007
	232 (IR)	*psbA* 3' UTR	Rogalski et al. 2006
	649	atpB 5'UTR & promoter	Kode et al. 2006
Glycine max	403	HPPD gene	Dufourmantel et al. 2007
Lactuca sativa	~200	16S rrn promoter & RBS	Lelivelt et al. 2005
Chlamydomonas reinhardtii	216	*chlL* coding region	Cerutti et al. 1995
	483	pACYC184 NruI-BspHI[2]	Fischer et al. 1996

Repeats were in direct orientation apart from Rogalski et al. 2006. Most recombinant genomes contained two engineered repeats apart from [1]three 418 bp and two 174 bp repeats (Iamtham and Day 2000). [2]Restriction fragment from pACYC184 plasmid. IR = inverted repeat. RBS = ribosome binding site. HPPD = 4-hydroxyphenylpyruvate dioxygenase.

DNA (Fig. 11a) while recombination between inverted repeats reverses the orientation of the intervening DNA. Both length and number of repeated sequences influence recombination frequency. Whereas two 418 bp direct repeats were ineffective in deleting intervening DNA, three 418 bp repeats promoted high frequencies of excision to leave a single 418 bp direct repeat (Iamtham and Day 2000). Intermediate forms containing two copies of the 418 bp repeat were not detected indicating that once activated the homologous recombination pathway goes to completion (Fig. 11b). A variety of direct repeats promote excision (Table 1) but a systematic study on the relationship between repeat length, DNA sequence and recombination frequency has not been carried out. Whereas recombination was barely detected between two 3'UTR *psb* A direct repeats of 418 bp (Iamtham and Day 2000), excision was promoted by repeated ~200 bp 16S *rrn* promoter elements in *L. sativa* (Lelivelt et al. 2005), and *N. tabacum* (Zou et al. 2003), and 210 bp *rbc* L 3' UTR repeats in *N. tabacum* (Dufourmantel et al. 2007). Recombination between 232 bp inverted *psbA* 3' UTR repeats has been shown in *N. tabacum* plastids (Rogalski et al. 2006). Excision can take place at any time during the transformation process and non-deleterious transgenes flanked by direct repeats, might be stabilised by the high copy number of plastid DNA once homoplasmy has been reached. Promoter regions of ~120 bp are duplicated in a number of plastid transformation vectors with no apparent reports of instability (Zoubenko et al. 1994) hinting at a lower size limit for efficient recombination.

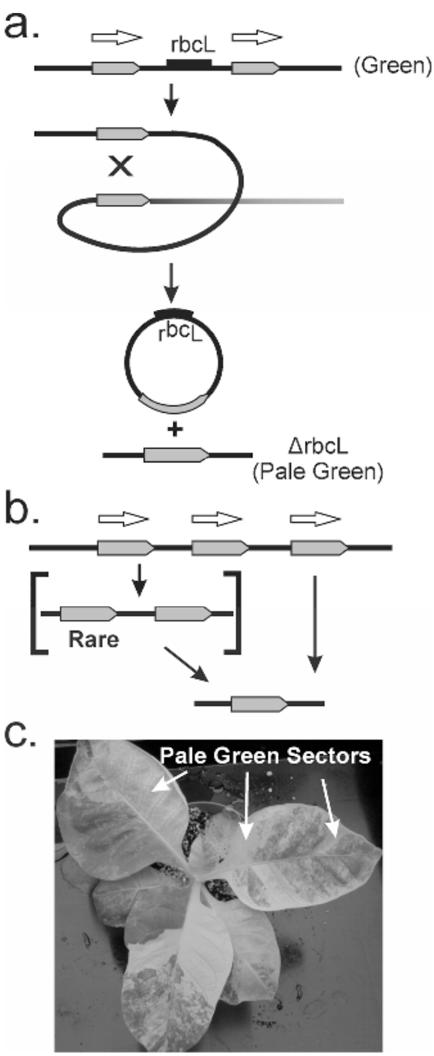


Fig. 11. Homologous recombination between engineered direct repeats (arrows) in plastid DNA. a) Recombination between two 649 base *atp*B promoter repeats (Kode et al. 2006) excises the *rbcL* gene and a marker gene (not shown). b) Recombination between three 418 bp 3' *psb*A repeats leaves a single 418 bp 3' *psbA* sequence. Once recombination is activated the pathway appears to go to completion because intermediates containing two repeats are rare and were not found (Iamtham and Day 2000). c) The recombination event shown in a) gives rise to pale green sectors (light areas) which appear randomly throughout leaf development.

The types of intervening genes excised by flanking direct repeats might also influence the accumulation of recombination products. Selection on plastid fitness might be expected to promote the division of plastids from which foreign genes with a negative impact on plastid functions and division have been deleted. The finding that excision mediated by recombination between duplicated 649 bp *atpB* 5’ regulatory regions allows the isolation of mutant defective plastid genomes lacking the *rbcL* gene (Fig. 11a; Kode et al. 2006) indicates that products of deleterious recombination events can be isolated under suitable conditions; in this case sucrose was provided in the media to allow non-photosynthetic growth. Recombination-mediated excision of *rbcL* and segregation of plastid genomes gives rise to pale-green sectors allowing the process to be monitored. Pale-green sectors of varying sizes are visualised in leaves (Fig. 11c, lighter areas represent pale-green sectors) indicating the recombination and segregation pathways are active throughout leaf development. The variable sizes and random appearance of pale-green sectors reflect random spontaneous excision events combined with stochastic replication and segregation of plastid genomes.

In *E. coli*, a minimum sequence length of 23-27 bp is considered to be required for efficient homologous recombination via the recBC-dependent pathway (Shen and Huang 1986). A minimum identical stretch of 150 to 200 bases appears to be required for homologous recombination in *C. reinhardtii* (Newman et al. 1992). Recombination between engineered direct repeats of 216 bp (Cerutti et al. 1995) has been observed in *C. reinhardtii* (Cerutti et al. 1995). In another study recombination was not detected between 100 bp or 230 bp direct repeats but frequent recombination was detected between 483 bp repeats (Fischer et al. 1996). This is probably because the recombination assay for the 216 bp repeats relied on restoration of gene function giving rise to green sectors and was more sensitive than the loss of antibiotic resistance assay involving the 230 bp direct repeats flanking *aad*A (Fischer et al. 1996). However, sequence-dependent differences in recombination rates between direct repeats cannot be ruled out (Fischer et al. 1996).

Site-specific recombination is an alternative to homologous recombination for manipulating plastid genomes. The Cre site-specific recombinase from the P1 bacteriophage of *E. coli* mediates strand-exchange between 34 bp LoxP sites (Sternberg and Hamilton 1981). When Cre is introduced into plastids it recombines LoxP sites as intended (Corneille et al. 2001; Hajdukiewicz et al. 2001). Unexpectedly, Cre recombinase also appears to stimulate recombination between ~120 bp direct repeats comprised of the 16S rrn promoter region (Corneille et al. 2001; Hajdukiewicz et al. 2001). The WT *N. tabacum* plastid genome lacking LoxP sites is stable in the presence of Cre recombinase (Corneille et al. 2003). This suggests that creation of double strand breaks at LoxP sites by Cre recombinase stimulates native recombination events in plastids. Cre-stimulated illegitimate recombination events between a LoxP site and a recombination hotspot in the promoter region of the rps7/3’rps12 operon were also reported (Hajdukiewicz et al. 2001; Corneille et al. 2003). The recombination hotspot contained multiple copies of a TATTA sequence (Hajdukiewicz et al. 2001). Short repeats are often associated with recombination hotspots in plastid DNA. The role of short multiple 18 to 37 bp repeats near a recombination hot spot in *C. reinhardtii* was addressed by deleting them. Their deletion did not reduce recombination frequency (Newman et al. 1992) indicating the repeats were not responsible for increased recombination. The observation that Cre-cleaved DNA ends are recombinogenic might suggest that the natural ends of the linear DNA molecules found in plastids (Oldenburg and Bendich 2004b; Scharff and Koop 2006) are protected by proteins or secondary structures, for example, DNA loops or hairpins.

Spontaneous excision of an 868 bp sequence following apparent recombination between 16 bp imperfect direct repeats (5’ GTACTGc/tGCTCTCCAA) was reported to accompany plastid transformation in *N. tabacum* (see Section 5 above; Staub and Maliga 1994). This might indicate that some plastid sequences of less that 20 bp are effective substrates for recombination. Evolutionary comparisons of plastid genomes have found DNA sequence inversions in the large single copy region that distinguish related species of flowering plants (Doyle et al. 1992). Analyses of the end points of an inversion in rice plastid DNA relative to *N. tabacum* suggest recombination events between sequences as short as 16 bp in length (Hiratsuka et al. 1989). In another study, short inverted repeats of 7-11 bp were found to be associated with inversions of the intervening 4 bp region (Kelchner and Wendel 1996). These studies indicate recombination events between short repeated stretches of nucleotides. Whether the same recombination pathway acts on the very short (10-20 bp) and longer (~200 bp and above) substrates remains to be determined. These questions can only be addressed once mutants in specific plastid recombination pathways have been isolated.

9 Recombination and plastid genome stability

An active homologous recombination pathway in plastids must underlie maintenance of the plastid genome and shape its evolution. Plastid transformation experiments have demonstrated recombination events between short repeats of ~200 bp in length in flowering plant plastids resulting in the deletion of genes (Table 1). Such excision events would be deleterious to WT plastid genomes. Grossly deleted *N. tabacum* transgenic plastid genomes resulting from recombination between distant *psbA* 3’UTR sequences do not persist (Svab and Maliga 1993) presumably because of strong selection against dysfunctional plastids with defective plastid genomes. Deleterious recombination events in WT plastid genomes would be avoided if they lacked DNA repeats or contained small repeats that were below the size needed for homologous recombination. One set of inverted repeats would be allowed because they would only flip the relative orientation of the single copy regions (Fig. 7). Any repeats that lie on either side of the large inverted repeat would be converted into direct repeats by large inverted repeat mediated flipping of single copy regions and destabilise plastid genomes (Day and Ellis 1984). Loss of the large inverted repeat in *P. sativum* is associated with rearrangements in plastid gene order (Palmer and Thompson 1982); presumably because inversions are not restricted by the presence of a large inverted repeat (Day and Ellis 1984).

Whilst most angiosperm plastid genomes contain a large inverted repeat of 20-76 kbp they are deficient in repeated DNA sequences over 100 bp in size. The plastid genome of *Pelargonium* x *hortorum* contains a large number of repeated sequences, including nine pairs of dispersed repeats of 31-101 bp in size

(Chumley et al. 2006). Over 20% of the *C. reinhardtii* plastid genome is comprised of repetitive DNA of less than 50 bp in length (Maul et al. 2002). Presumably these repeated sequences of 100 bp or less are too small to act as efficient substrates for homologous recombination. Plastid DNA is uniform within a plant and recombination events that would destabilize and fragment plastid genomes are not normally observed. However, low frequency recombination events between short repeats might give rise to the plastid DNA rearrangements observed during evolution (Hiratsuka et al. 1989; Chumley et al. 2006). Deleted plastid genomes with circular maps (Day and Ellis 1984; Cuzzoni et al. 1995) and circular topologies (Day and Ellis 1985) have been described in albino cereal plants from anther culture demonstrating the instability of the plastid genome when cells are rescued by heterotrophic growth *in vitro*. A recent report suggests albino *Bambusa edulis* (bamboo) plants contain deleted plastid genomes (Liu et al. 2007). Aberrant subgenomic circles have also been described in WT *N. tabacum* chloroplasts by DNA fibre-based FISH analysis (Lilly et al. 2000). Two factors appear to be required to maintain plastid genomes, first, DNA-RRR surveillance mechanisms that either reduce the frequency of deleterious recombination events or repair deleted plastid genomes, and second, selection for functional plastids with an intact plastid genome.

10 Homeologous recombination in plastids

In *E. coli*, sequence divergence dramatically reduces the rate of homologous recombination. A 10% reduction in identity between DNA sequences reduces homologous recombination frequency by 40-fold (Shen and Huang 1986). The effect of mismatches on integration of transgenes can be studied by using heterologous plastid DNA to target integration of foreign genes into plastids. *N. tabacum* plastid DNA has been used to target integration of foreign genes in *Lycopersicon esculentum* (tomato; Ruf et al. 2001), *Petunia hybrida* (Zubko et al. 2004) and *Solanum tuberosum* (potato; Sidorov et al. 1999). Because the transformation frequency of homeologous *N. tabacum* plastid DNA was not compared with homologous plastid DNA the influence of mismatch on plastid transformation frequency in these species is not known. In a more detailed study, a 7.8 kbp region of *Solanum nigrum* plastid DNA was introduced into *N. tabacum* plastids (Kavanagh et al. 1999). Recombinant plastid genomes exhibited a mosaic structure comprised of several patches of *S. nigrum* DNA interspersed with *N. tabacum* plastid DNA sequences (Fig. 10b). This is consistent with multiple recombination events during integration of 7.8 kbp *S. nigrum* plastid DNA and random resolution of Holliday junctions. Although *S. nigrum* and *N. tabacum* plastid DNA showed 2.4% sequence divergence, plastid transformation frequencies were not reduced relative to using homologous sequences for *N. tabacum* plastid transformation. Stringent mismatch repair processes which reduce recombination between diverged DNA sequences (Evans and Alani 2000) might be suppressed in flowering plant plastids (Kavanagh et al. 1999). In contrast, homeologous plastid DNA exhibiting around

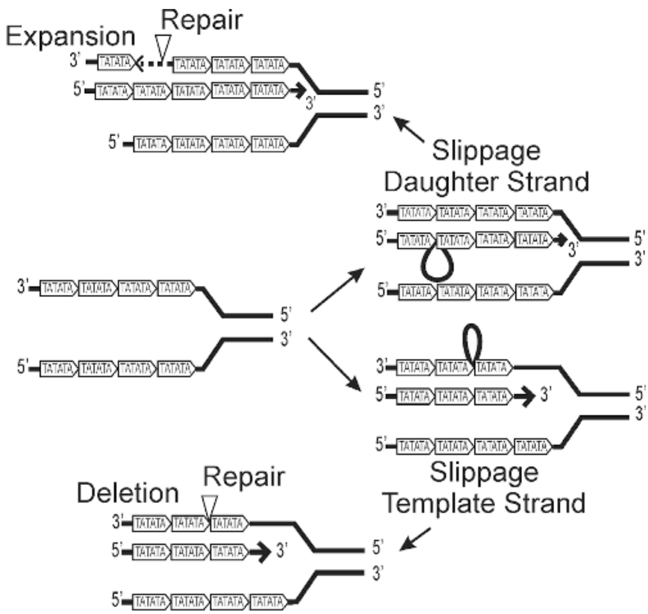


Fig. 12. Expansion and deletion of short direct repeats by replication slippage (Lovett 2004). Slippage in the daughter strand increases the number of repeat units while slippage in the template strand reduces the number of repeat units.

2-4% mismatch decreased transformation frequency by two to fivefold (Newman et al. 1990) in *C. reinhardtii*. This might indicate a more efficient plastid mismatch repair system in this green alga.

11 Replication slippage in plastids

Very short tandem repeats, based on mononucleotide, dinucleotide, trinucleotide, and consecutive nucleotide repeats up to the ~30-mer, are found in plastid genomes. Very short tandem repeats are considered to result from slippage of the replication fork. Replication slippage in the newly replicated daughter strand inserts a repeat whilst slippage of the template strand deletes a repeat (Fig. 12). A number of hot spots of variation in plastid genomes are associated with short tandem repeats (Newman et al. 1992; Sears et al. 1996; Stoike and Sears 1998; Ogihara et al. 2002). Plastome-mutator is a nuclear mutation in *O. hookeri* associated with a 200 to 1000-fold increase in pigment-deficient sectors and changes in plastid DNA (Epp 1973; Stoike and Sears 1998). The product of the plastome-mutator gene is not known but it has been suggested to be involved in plastid DNA-RRR pathways (Stoike and Sears 1998). Examination of alterations induced by plastome-mutator in the intergenic region between the 16S and 23S ribosomal RNA

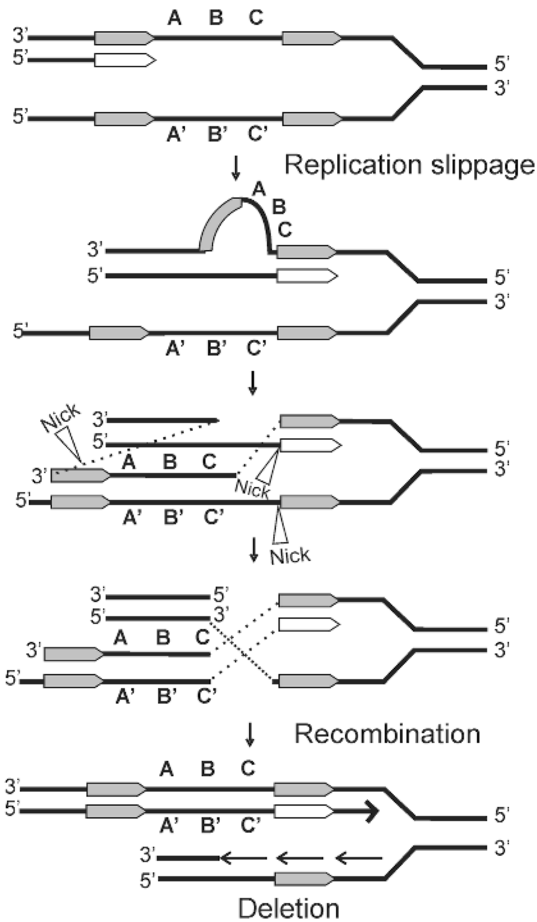


Fig. 13. Direct repeat mediated deletion by replication slippage and recombination (Bi and Liu 1996).

genes suggested they were the result of replication slippage rather than recombination events (Stoike and Sears 1998).

Replication slippage between direct repeats and recombination (Fig. 13; Bi and Liu 1996) provides an alternative mechanism to homologous recombination (Fig. 11a) for excision of genes from plastid DNA. Replication slippage induced recombination is reduced with increasing distance between repeats and does not increase in frequency when repeat length is increased above ~100 bp in *E. coli* (Bi and Liu 1996). The observation that recombination between engineered repeats in plastids increases with longer repeats favours a mechanism involving homologous recombination rather than slippage (see Section 8.2 above). Replication slippage can give rise to insertions and deletions (Fig. 12) and appears to be a major force in plastid genome evolution. The 70 kb plastid genome of the non-photosynthetic parasite *Epifagus virginiana* is less than half the size of the 156 kb *N. tabacum*

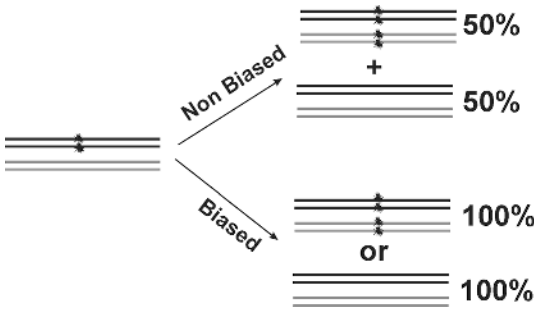


Fig. 14. Recombination events between WT and mutant (*) DNA duplexes can be non-biased or biased. Biased gene conversion events will favour one of the alleles and either fix or remove mutations.

plastid genome (Wolfe et al. 1992). Gene order in *E. virginiana* and *N. tabacum* is conserved despite the loss of all photosynthesis-related genes. A large number of deletion events on a local scale mediated by replication slippage rather than recombination events between distant parts of the plastid genome would appear to be the major mechanism underlying selective gene loss and genome reduction in *E. virginiana* (Wolfe et al. 1992).

12 DNA repair in plastids

DNA replication, recombination, and repair are interrelated processes and the homologous recombination pathway in plastids is likely to play a role in DNA repair (Cerutti et al. 1995). The estimated mutation rate of plastid genes is approximately two-fold lower than that of nuclear genes (Wolfe et al. 1987). Within plastids the synonymous substitution rate of genes located in the large inverted repeat is about two-fold lower than that for genes located in the single copy regions (Perry and Wolfe 2002). This has been interpreted to be the result of the two-fold higher dosage of inverted repeat sequences and biased gene conversion in favour of WT plastid DNA sequences (Birky and Walsh 1992; Perry and Wolfe 2002). Non-biased repair will either correct the mutation to WT or fix the mutation (convert WT to mutant) and give rise to both outcomes in equal proportions. Biased repair favours one of these outcomes to give 100% of only one product (Fig. 14). Direct experimental confirmation for biased gene conversion has been obtained by monitoring correction of mutations tightly linked (31 bp distance) to an *aad*A insertion in transgenic *N. tabacum* plastids (Khakhlova and Bock 2006). Whilst the *aad*A gene was retained by spectinomycin selection the mutations were repaired to WT with a bias for reversing AT to GC changes more efficiently than GC to AT mutations. It has been suggested that this bias towards AT might underlie the high overall AT content (>70% AT) of plastid genomes (Khakhlova and Bock 2006). Multiple copies of plastid DNA and biased gene conversion in favour of WT would reduce the rate at which mutations are fixed.

Alternatives to RecA-based recombination repair include photoreactivation, base excision repair, nucleotide excision repair, and mismatch repair (Kimura and Sakaguchi 2006). Little is known on these alternative repair pathways in plastids. A putative plastid-localised uracil-DNA glycosylase activity probably involved in base excision repair was partially purified from *Z. mays* chloroplasts (Bensen and Warner 1987). UV-induced lesions in the plastid *psbA* gene of *G. max* suspension culture cells were repaired in the light (but not in the dark) with kinetics that were considerably slower than expected for photoreactivation by photolyases (Cannon et al. 1995). Experiments on purified *S. oleracea* chloroplasts (Hada et al. 2000) and the lack of identification of plastid transit peptides in the products of plant genes encoding photolyases (Draper and Hays 2000) led to the possibility that plastids might be deficient in photolyase-mediated photoreactivation. However, tolerance of plastid DNA replication to UV-B lesions in *A. thaliana* plants grown in blue (photorepair-compatible) light might suggest the presence of as yet unidentified photolyases in plastids (Draper and Hays 2000). Under gold light where light-dependent photorepair does not take place, a UV-B fluence rate of 5 kJ m-2 inhibits replication of plastid DNA but not nuclear and mitochondrial DNA indicating a deficiency in light-independent (dark) repair pathways in chloroplasts (Draper and Hays 2000).

Endonuclease activities that could act on apurinic sites following base removal were purified from *H. vulgare* chloroplasts (Velemínskíý et al. 1980). A single-strand specific nuclease activity from *T. aestivum* chloroplasts cleaves single stranded DNA or RNA regions including 5’ flaps, 5’ overhangs, and 3’ pseudoflaps and has been suggested to be involved in DNA repair (Przykorska et al. 2004). The multi-subunit replication protein A (RPA) binds to single-stranded DNA and is involved in pathways including nucleotide excision repair. RPA subunits appear to be targeted to plastids (Kimura and Sakaguchi 2006). Plastid-localised homologues of RecQ have been implicated in DNA repair (see Section13.5 below). Nitroso-methy-urea and nitroso-guanidine are particularly effective for inducing mutations in flowering plant plastid genomes (Hagemann 1976). Methyl transferases reverse the damage to bases caused by these alkylating agents. The presence of methyl transferases in nuclei but their absence in plastids might explain the utility of these mutagens for inducing plastid mutations (Sears 1998).

13 Identification of proteins involved in plastid DNA RRR-pathways

Studies on bacteria have identified a suite of DNA metabolism enzymes including DNA polymerases, DNA primase, RecA, topoisomerases, and helicases (CameriniOtero and Hsieh 1995). The cyanobacterium *Anabaena* contains 93 genes encoding proteins with significant similarity to known DNA-RRR proteins (Kaneko et al. 2001). Early studies to identify plastid DNA-RRR proteins involved purifying enzymes with DNA-RRR activities, such as DNA synthesis, from chloroplasts (McKown and Tewari 1984). More recently, whole genome databases have been used to identify candidate plastid-targeted proteins with significant matches to well-known bacterial DNA-RRR proteins. Further experimental support is then required to confirm *in silico* predictions of plastid location. The list of genes encoding homologues of DNA-RRR proteins for which there is experimental support for a plastid location is short and is reviewed below. Proteomics of purified chloroplasts (Chapter 12) provides an alternative approach to identify DNA-RRR proteins. However, the limited abundance of plastid DNA-RRR proteins hinders their identification in whole chloroplast preparations. Further purification of sub-chloroplast fractions containing DNA-protein complexes is required to identify plastid DNA-RRR proteins (Sakai et al. 1999; Phinney and Thelen 2005).

13.1 Plastid DNA polymerases

Eukaryotic cellular template-dependent DNA polymerases can be classified into α, β, γ, δ, and ε DNA polymerases based largely on work in vertebrates and yeast (Wang 1991). The α, β, δ, and ε DNA polymerases are located in nuclei whereas the γ–DNA polymerase is located in mitochondria. This original list of five eukaryotic DNA polymerases has expanded to fourteen (Mori et al. 2005) as new DNA polymerases are identified including those involved in translesion DNA synthesis (Hubscher et al. 2000). Nine classes of DNA polymerase have been identified by analyses of plant genomes (Mori et al. 2005). The activity of γ–DNA polymerases is resistant to aphidicolin, sensitive to N-ethylmaleimide, and relatively resistant to low concentrations of dideoxynucleoside 5' triphosphates (Wang 1991). This γ–DNA polymerase activity can be distinguished from α, δ, and ε DNA polymerases which are sensitive to aphidicolin but resistant to dideoxynucleoside 5' triphosphates, and β–DNA polymerase, which is resistant to both aphidocolin and N-ethylmaleimide but sensitive to low concentrations of dideoxynucleoside 5' triphosphates (Wang 1991). Plastid DNA polymerases that have been characterised from *S. oleracea* (Spencer and Whitfeld 1969; Sala et al. 1980), *P. sativum* (McKown and Tewari 1984), *G. max* (Heinhorst et al. 1990; Bailey et al. 1995), *N. tabacum* (Sakai et al. 1999), and *C. reinhardtii* (Wang et al. 1991) resemble γ-DNA polymerases based on their resistance to aphidocolin and sensitivity to N-ethylmaleimide. Furthermore, the plant but not the *C. reinhardtii* plastid DNA polymerases are resistant to low concentrations of dideoxynucleoside 5' triphosphates (Heinhorst and Cannon 1993). The activities of purified plastid DNA polymerases appear to be stimulated by KCl (Spencer and Whitfeld 1969; Sala et al. 1980; McKown and Tewari 1984; Heinhorst et al. 1990; Sakai et al. 1999) and inhibited by ethidium bromide (McKown and Tewari 1984; Wang et al. 1991; Sakai et al. 1999).

Estimated sizes of plastid DNA polymerases were 87 kDa (McKown and Tewari 1984) and 70 kDa for *P. sativum* (Gaikwad et al. 2002), 85-90 kDa for *G. max* (Heinhorst et al. 1990), 116 kDa for *N. tabacum* (Sakai et al. 1999), and

Table 2. Plastid DNA-RRR proteins encoded by *Arabidopsis thaliana* nuclear genes.

Protein	Length (aa)	Acc. no.	Gene no.	Reference
DNA polymerase	1049[1]	AAL58915	At1g50840	Christensen et al. 2005
	1034[1]	BAE98907	At3g20540	
RecA	439	Q39199	At1g79050	Cao et al. 1997
Gyrase A	950[1]	AAG51377	At3g10690	Wall et al. 2004
Gyrase B	732[1]	Q94BZ7	At5g04130	Wall et al. 2004;
	657[1]	Q9SS38	At3g10270	Christensen et al. 2005
RecQ	858[2]	Q9FT69	At5g27680	Saotome et al. 2006;
	606[2,3]	Q9FT74	At3g05740;	Hartung and Puchta 2006

Dual targeting or alternative translation start sites target product to both plastids and mitochondria. Length will vary depending on translation initiation at alternative start sites or at non-AUG start codons in the 5' UTR (Christensen et al. 2005). 2. Homologues of *Oryza sativa* (rice) plastid-localised proteins (Saotome et al. 2006). 3. Targeted to plastids and nuclei in *O. sativa.*

possibly two catalytic subunits of 80 and 116 kDa for *C. reinhardtii* (Wang et al. 1991). The size discrepancies might be explained by proteolytic cleavage of proteins during purification. In the absence of primary sequence information the relationships between these proteins are not known and it is unclear whether they are different DNA polymerases or homologues of the same protein. Analyses of DNA polymerase genes present in sequenced genomes (discussed below; Mori et al. 2005) are likely to help resolve some of the discrepancies encountered in the earlier biochemical work. A number of proteins have been found to be associated with purified plastid DNA polymerases. These include a 43 kDa protein (related to ribonuclease T2, GenBank Acc. P93845) that stimulates the activity and processivity of an 87 kDa (Chen et al. 1996) and a 70 kDa *P. sativum* DNA polymerase (Gaikwad et al. 2002), and a 3' to 5' 20 kDa exonuclease subunit of a 105 kDa *S. oleracea* DNA polymerase complex (Keim and Mosbaugh 1991).

There are similarities in the sizes and properties of DNA polymerases purified from plant mitochondria and chloroplasts (Heinhorst et al. 1990; Sakai et al. 1999). The idea of related DNA polymerases in mitochondria and plastids is supported by analyses of genes in sequenced genomes. The *A. thaliana* genome encodes at least two genes encoding organelle DNA polymerases (Table 2) sharing 70% amino acid identity that are expressed in the shoot apical meristem (Mori et al. 2005). The 116 kDa 1034-long *A. thaliana* DNA polymerase-like protein (gene no. At3g20540) contains a presequence that is predicted (Emanuelsson et al. 2000) to target mitochondria (TargetP score 0.74) and chloroplasts (TargetP score 0.59). A related 117 kDa, 1049 amino acid *A. thaliana* protein (gene no. At1g50840) contains a putative plastid targeting presequence (TargetP score 0.933). Both N-termini deliver GFP to chloroplasts but the 116 kDa presequence appears to also target GFP to the mitochondria (Mori et al. 2005). Interestingly, the 117 kDa protein is targeted to mitochondria when translation initiates upstream of the predicted AUG start codon most probably at an in frame CUG lying seven codons upstream of AUG (Christensen et al. 2005). Use of this non-AUG start codon suggests both *A. thaliana* organelle DNA polymerases are dual-targeted to plastids and mitochondria adding complexity to the regulation of organelle DNA polymerases.

Homologues of the *A. thaliana* organelle targeted DNA polymerases are present in *O. sativa*. Two *O. sativa* organellar DNA polymerases are predicted to be dual-targeted to plastids and mitochondria (Christensen et al. 2005). Polyclonal antibodies raised against one of these polymerases cross-reacted with a protein in isolated chloroplasts (Kimura et al. 2002). Expression of this DNA polymerase was studied by *in situ* hybridization. RNA was detected in tissues with dividing cells including leaf primordia, and the apical meristem of shoots and roots but not in mature leaves (Kimura et al. 2002). These genome based studies suggest plastids contain at least two ~110 kDa DNA polymerases that appear to be dual-targeted to mitochondria and plastids (Christensen et al. 2005; Mori et al. 2005). Future experiments should delineate the roles of these polymerases in organelle DNA replication versus repair. Recombinant forms of these enzymes will facilitate the characterization of their properties (Mori et al. 2005). These *A. thaliana* and *O. sativa* organelle DNA polymerases contain 3'-5' exonuclease and DNA polA-like domains. The distantly related apicoplast in the malaria parasite *Plasmodium falciparum* appears to contain a different DNA polymerase. A 235 kDa multidomain protein with DNA primase, DNA helicase, DNA polymerase and 3' to 5' exonuclease regions is implicated in replication of apicoplast DNA (Seow et al. 2005).

13.2 DNA primase activities in plastids

A primase activity purified from *P. sativum* chloroplasts was linked to a 115-120 kDa protein that was distinct from the 90 kDa DNA polymerase. The primase was resistant to tagetitoxin, an inhibitor of RNA polymerase, and able to initiate replication on poly dT, poly dC and cloned ssDNA templates containing *P. sativum* OriA (Nielsen et al. 1991). No primary sequence information is available on this primase.

13.3 Plastid localised RecA

The RecA protein promotes strand transfer and heteroduplex formation between DNA duplexes in prokaryotes (CameriniOtero and Hsieh 1995). RecA is required for DNA replication, recombination and repair pathways (Kowalczykowski 2000). The central role of RecA protein in homologous recombination is illustrated by its position in the double-strand break recombination model (Szostak et al. 1983) in Figure 15. Absence of *recA* function reduces the frequency of homologous recombination by 10,000–fold in *E. coli* (CameriniOtero and Hsieh 1995). Plastids contain a homologue of RecA, which was first identified as a 39 kDa protein in *P. sativum* chloroplasts using polyclonal antibodies against *E. coli* RecA protein

(Cerutti et al. 1992). This provided the first evidence for a RecA-mediated homologous recombination pathway in plastids. Accumulation of the *P. sativum* RecA-like protein appears to increase following exposure of protoplasts to DNA damaging agents (Cerutti et al. 1993). The presence of a RecA homologue in *P. sativum* chloroplasts has been recently confirmed by proteomics of protein-nucleic acid particles from purified chloroplasts (Phinney and Thelen 2005). The *A. thaliana* nuclear genome contains four genes encoding RecA-like proteins related to the bacterial RecA proteins (Khazi et al. 2003). Two of these genes encode organelle-targeted proteins (Khazi et al. 2003). Gene At1g79050 encodes a 439 amino acid protein (Table 2) that is predicted to be targeted to chloroplasts (TargetP score 0.841; Emanuelsson et al. 2000) and is imported into isolated *P. sativum* chloroplasts (Cao et al. 1997), whereas gene At3g10140 encodes a 389 amino acid protein that is targeted to mitochondria (Khazi et al. 2003). Both RecA-like proteins share 36% identity excluding N- and C-terminal extensions (Khazi et al. 2003). The At3g10140 mitochondrial RecA-like protein partially complements an *E. coli recA* mutant and provides tolerance to the methyl methane sulfonate (MMS) and mitomycin C (Khazi et al. 2003), which are DNA damaging agents.

C. reinhardtii plastids contain a RecA homologue (Nakazato et al. 2003). The influence of altered RecA activity on plastid recombination was addressed by expressing *E. coli* WT and dominant negative RecA proteins in *C. reinhardtii* chloroplasts (Cerutti et al. 1995). Expression of WT *E. coli* RecA in *C. reinhardtii* chloroplasts increased the frequency of homologous recombination between 216 bp direct repeats by over 15-fold. This indicates that recombination between direct repeats appears to be limited by RecA-mediated strand exchange. In contrast, overexpressed WT *E. coli* RecA did not enhance survival of cells exposed to DNA damaging agents indicating that chloroplast DNA repair pathways are not limited by RecA activity but by processing of DNA substrates generated by DNA damaging agents. An *E. coli* dominant-negative RecA protein reduced recombination between direct repeats and DNA repair in chloroplasts consistent with a negative influence on a RecA-mediated DNA-RRR pathway in plastids (Cerutti et al. 1995).

13.4 DNA topoisomerases

Topoisomerases regulate DNA supercoiling, DNA catenation and knotting, and are important enzymes required for DNA replication, recombination and repair. Type I topoisomerases relax supercoiled DNA whereas type II topoisomerases, such as DNA gyrase, not only relax supercoiled DNA but also introduce supercoils using ATP (Singh et al. 2004). Several reports have identified type I topoisomerases in chloroplasts including a 115 kDa protein in *S. oleracea* (Siedlecki et al. 1983), a 54 kDa protein in *Brassica oleracea* (cauliflower; Fukata et al. 1991), a 69 kDa protein in *P. sativum* (Mukherjee et al. 1994), and a 70 kDa type IB topoisomerase in *Sinapsis alba* (white mustard; Belkina et al. 2004). Evidence for type II topoisomerases in plastids include the observation that a *P. sativum* chloroplast transcription extract contained a gyrase-like activity which was sensitive to the gyrase inhibitor novobiocin (Lam and Chua 1987). Furthermore, an antibody raised against yeast topoisomerase II cross-reacted with 96 and 101 kDa proteins in *T. aestivum* chloroplasts (Pyke et al. 1989); *E. coli* gyrase is comprised of 95 kDa gyrase B and 105 kDa gyrase A subunits (Reece and Maxwell 1991). A more recent proteomic study identified gyrase A- and B-like subunits in protein-nucleic acid particles from purified *P. sativum* chloroplasts (Phinney and Thelen 2005).

The *A. thaliana* nucleus contains one gene encoding a gyrase A-like subunit and three genes for gyrase B-like subunits (Wall et al. 2004). The gyrase A subunit (At3g10690) coding sequence has alternative start sites giving rise to plastid and mitochondrial targeted forms. T-DNA knockouts of the gyrase A subunit were embryo-lethal. One gyrase B subunit appeared to be targeted to plastids, the second to mitochondria, whereas the location of the third was unclear but was possibly located in the nucleus/cytosol (Wall et al. 2004). Knockouts of either organelle-targeted gyrase B subunit were seedling-lethal rather than the more severe embryo-lethal phenotype of gyrase A mutants. This suggested the gyrase B subunits complement each other to a limited extent indicating their products might be targeted to both organelles (Wall et al. 2004). This appears to be the case. An alternative upstream non-AUG start codon (most probably CUG) in one gyrase B subunit (gene At3g10270) gives rise to an N-terminus that confers dual-targeting to mitochondria and plastids. Alternative translation start sites in the coding sequence for the second organelle-targeted gyrase B subunit (gene At5g04130) give rise to either mitochondrial or plastid targeted proteins (Christensen et al. 2005). In summary, one gyrase A subunit and two gyrase B subunits appear to be targeted to both plastids and mitochondria in *A. thaliana* (Table 2). Presequences that confer dual-targeting of gyrase A and B subunits to mitochondria and chloroplasts have been found in *Nicotiana benthamiana* (Cho et al. 2004).

The effects of transient downregulation of organelle-targeted gyrase A and B subunit expression were studied in *N. benthamiana* by virus-induced gene silencing (Cho et al. 2004). Downregulation of gyrase A or B subunits prevents chloroplast development giving rise to white or yellow leaf sectors. Larger nucleoids and a mixture of heterogeneous high MW DNA molecules in plastids, possibly representing tangled DNA and their breakage products, are consistent with a crucial role for gyrase in untangling plastid DNA following replication and recombination (Cho et al. 2004). A role for gyrase in plastid DNA maintenance is supported by an earlier study where the gyrase inhibitors novobiocin and naladixic acid were shown to reduce the copy number of plastid DNA in *Solanum nigrum* suspension cultures (Ye and Sayre 1990).

In addition to roles in DNA-RRR pathways gyrase activity can also influence transcription (Chapter 5) through changes in supercoiling (Reece and Maxwell 1991). Mutations in gyrase activity might therefore impact on plastid gene expression as well as genome maintenance. The gyrase inhibitors novobiocin and naladixic acid were found to alter the accumulation of plastid transcripts in *C. reinhardtii* (Thompson and Mosig 1985). Addition of novobiocin to a *P. sativum* chloroplast transcription system containing cloned plastid genes inhibited the expression of the *atpB* gene to a larger extent than the *rbcL* gene. This raised the possibility that template topology may enable differential regulation of plastid genes (Lam and Chua 1987).

13.5 DNA helicases

The DNA unwinding steps of DNA-RRR pathways include the generation of single-stranded recombination substrates (Fig. 15) and are carried out by ATP-dependent helicases. In an early study, a helicase fraction containing 6-8 protein bands was purified from *G. max* chloroplasts and shown to remove a 28 base oligomer from a single-stranded circular M13 template (Cannon and Heinhorst 1990). A similar biochemical approach identified 68 and 78 kDa helicases in purified *P. sativum* chloroplasts (Tuteja 2003). The 78 kDa helicase was stimulated by DNA fork structures indicating a role in replication (Tuteja and Phan 1998). Unwinding was inhibited by nogalamycin and ATPase activity by daunorubicin (Tuteja and Phan 1998). Both nogalamycin and daunorubicin are major groove intercalating agents. Chloroplast helicases appear to be sensitive to actinomycin C1 and resistant to ellipticine whereas the converse is true of nuclear helicases (Tuteja 2003).

A genomics based study has identified two *O. sativa* RecQ helicase homologues that are likely to be present in plastids (Saotome et al. 2006). Transient expression of GFP-fusions in onion epidermal cells indicated one 588 amino acid RecQ-like protein (OsRecQ1, Nucleotide Acc No. AK101124) was targeted to nuclei and plastids, whereas the second 844 amino acid Rec Q-like protein (OsRecQsim, Nucleotide Acc No. AK072977) was targeted predominantly to plastids. RNA levels for these RecQ-like proteins appeared highest in meristematic tissues containing immature plastids and did not increase in response to light and chloroplast development. A role for these proteins in repair was suggested by the observations that RNA encoding the 588 amino acid RecQ-like protein increased in levels in response to the four DNA damaging agents, mitomycin C, H_2O_2, MMS and bleomycin. Expression of the RNA encoding the 844 amino acid protein increased following treatment with mitomycin C and bleomycin and increased slightly with MMS (Saotome et al. 2006).

14 Identifying DNA-RRR proteins by complementation of *E. coli* mutants

Most of our knowledge of eubacterial DNA-RRR pathways is based on *E. coli* (CameriniOtero and Hsieh 1995; Kowalczykowski 2000). Plastids are probably descendants of an ancient cyanobacterium (Martin et al. 2002; Chapter 1) and have retained components of eubacterial DNA-RRR pathways including homologues of DNA polymerase, gyrase, and RecA (Table 2). In some cases, these plastid proteins have been shown to complement mutations in the homologous

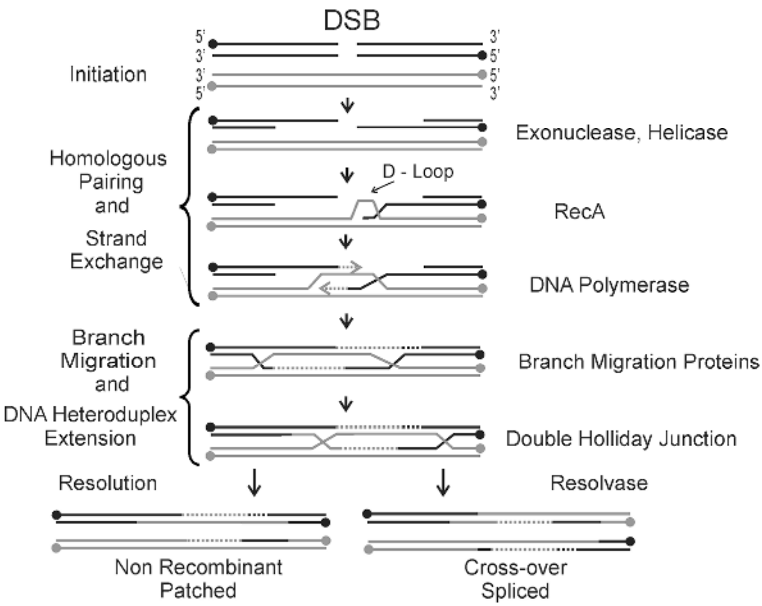


Fig. 15. Double-strand-break model of homologous recombination showing main proteins involved (Szostak et al. 1983; Kowalczykowski 2000).

E. coli proteins (Cho et al. 2004). Complementation of *E. coli* mutations provides a functional assay for identifying cDNAs encoding plastid-DNA RRR proteins andwould appear to be an attractive method for isolating plant homologues of *E. coli* DNA-RRR proteins. Using cDNA libraries, plant cDNAs that complement a number of *E. coli* mutations in DNA-RRR genes were isolated (Pang et al. 1993b) including *RecA* (Pang et al. 1992), and *ruvC* and *recG* mutants (Pang et al. 1993a). Unfortunately, the isolated cDNAs did not encode proteins related to characterized DNA-RRR proteins hindering progress. One of the isolated cDNAs complementing *ruvC* and *recG* double mutants was subsequently identified as plastocyanin (acc number P42699) raising questions on the validity of these library screening experiments. Screening cDNA libraries for functional complementation of *E. coli* mutants is technically difficult especially if complementation is not strong. Pitfalls include transformation or transfection of the tiny number of *E. coli* cells in which the mutation has reverted or been suppressed, which will give rise to viable cells on antibiotic medium. Alternatively, the cDNA may encode a protein that rescues the mutation indirectly, for example, by stabilising a temperature sensitive *E. coli* mutant protein.

15 Conclusions and outlook

Plastid transformation experiments have demonstrated an efficient homologous recombination pathway in plastids mediated by a RecA-homologue that appears to be active throughout shoot development. The presence of this pathway is consistent with a new emerging view of plastid DNA maintenance in which recombination plays a predominant role. WT plastid DNA is comprised of a mixture of circular and linear DNA molecules, which form a multimeric series from monomer to at least the octomer, and high MW DNA complexes (Section 3). Deleted plastid genomes in grasses contain sub-genomic circles and linear hairpin DNA molecules (Section 3.1, 6). The relationship between the mechanisms responsible for the maintenance of WT plastid DNA and the formation and replication of small linear DNA molecules in plastids is not understood. Replication models to account for the different topological forms of WT plastid DNA (circular DNA, linear DNA, branched complex DNA) have been proposed (Fig. 5). To identify which of these models are correct requires experimental confirmation beyond further descriptions of topological forms in WT chloroplasts. Progress in this research area requires the identification of proteins involved and mutants to determine the impact of loss or downregulation of these DNA-RRR proteins on plastid DNA levels and topological forms.

Whilst several approaches have localised putative replication origins in plastid DNA from flowering plants they have been mapped to different positions (Fig. 6) hindering the application of a universal model. Multiple locations for replication origins might reflect multiple origins in plastid DNA and differential usage of replication origins in different cells or differences in the accuracies of the methods used. The possibility of alternative modes of replication in different plastid types (Wang et al. 2003) increases the complexity of studying plastid genome maintenance. Distinguishing between these replication pathways might require the isolation and analysis of mutants affecting specific pathways. Recombination-dependent DNA replication plays an important role in genome maintenance in bacteria and has been suggested to be active in plastids to account for the complex branched DNA structures found in *Z. mays* plastids (Bendich 2004; Oldenburg and Bendich 2004b). Linear DNA molecules with heterodisperse or defined ends could invade template DNA molecules to prime DNA replication by recombination (Fig. 5d). The identification of linear DNA molecules with fixed ends that map to potential origins in the large inverted repeat (Oldenburg and Bendich 2004b; Scharff and Koop 2006) is interesting and characterisation of the structures of these ends might reveal the mechanisms involved in their formation. Whether recombination-dependent DNA replication is limited to a relatively small number of specific sites in plastids, possibly corresponding to the natural ends mapped in *Z. mays* (Oldenburg and Bendich 2004b) and *N. tabacum* (Scharff and Koop 2006), can be addressed by mutating these sites in recombinant plastid genomes (Scharff and Koop 2007). The finding that a high proportion of plastid DNA (50%) is comprised of complex branched DNA molecules in *Z. mays* seedlings (Oldenburg and Bendich 2004b) warrants further investigation in other species using additional techniques such as DNA fibre-based FISH with plastid DNA probes (Lilly et al. 2001) to study the organisation of these complexes.

A highly active homologous recombination pathway in plastids is consistent with recombination-dependent DNA replication. Widespread inter-molecular and intra-molecular recombination between large inverted repeat sequences or between repeated copies of the unit genome would be expected to produce a large number of isomers. If the molecules are linked by strand-invasion and recombination-dependent DNA replication this will give rise to a complex mixture of interconnected high-molecular-weight complexes (Oldenburg and Bendich 2004). The organisation of plastid DNA as high MW multi-genome complexes (containing linear and branched forms) has been suggested to underlie the packaging of plastid DNA into nucleoids (Bendich 2004). Random replication and recombination events are thought to contribute to the random segregation patterns observed for plastid genomes (Birky 1994, 2001). This raises the question of whether all genomes and topological forms have an equal chance of being replicated? Our current knowledge is too limited to address such a question. Other interesting areas worth exploring in future work include the relationship between topological forms and transcription, and the maintenance of heteroplasmic states. Distinct plastid genomes in heteroplasmic plants, where both genomes are required for survival, might be expected to segregate to different high MW DNA complexes within a plastid. However, the maintenance of different plastid genomes in the same high MW complexes might be possible and shed light on the dynamics of plastid genome maintenance. In normal WT plants copy-correction involving DNA repair pathways would be expected to ensure the maintenance of a uniform population of plastid DNA molecules.

Plastids have been evolving in the cytoplasm of their eukaryotic hosts for several billion years and have acquired proteins of nuclear or mitochondrial origin that were not present in the original cyanobacterial symbiont (see for example Wagner and Pfannschmidt 2006). Elucidation of DNA-RRR pathways in plastids should confirm roles for eukaryotic proteins (Mukherjee et al. 1994) in addition to roles for homologues of well known prokaryotic DNA-RRR proteins. It seems likely that the proteins, mechanisms, and regulation of plastid DNA-RRR pathways will have diverged substantially from the eubacterial DNA-RRR model. The availability of whole plant genome sequences allow genomic approaches to identify genes encoding proteins of prokaryotic (Table 2) and eukaryotic origin involved in plastid genome maintenance. A major problem is the prediction of plastid location due to the difficulty in identifying the N-termini of proteins from gene and cDNA sequences, and because computer programs (Emanuelsson et al. 2000) are only partially successful in predicting plastid-targeted proteins. Approximately 30% of chloroplast proteins do not contain recognisable plastid targeting signals (Kleffmann et al. 2004). Proteomics provides an alternative method to identify plastid DNA-RRR proteins. However, proteomic studies on chloroplasts (Kleffmann et al. 2004) have not uncovered the suite of DNA-RRR present in plastids possibly because of their limited abundance. More success has been achieved by proteomics of purified subfractions enriched in nucleoids (Sakai et al. 1999; Phinney and Thelen 2005). Alternatively, shoot tissues with actively dividing cells express a number of DNA-RRR proteins (Sakai et al. 1999; Saotome et al. 2006) and might provide better material for proteomic studies on plastid DNA-RRR proteins.

Reverse genetics provides a powerful tool to elucidate the roles of candidate DNA-RRR proteins in plastid genome maintenance. Knockout, using T-DNA insertions or transposons, and knockdown approaches, using RNAi, can be used to identify genes with important roles in plastid DNA maintenance. Knockdowns are particularly suitable for studying essential genes by allowing the isolation of viable plants. This provides the plant material in which to study the impact of DNA-RRR protein deficiencies on plastid genome maintenance. Plastid transformation allows the DNA substrates of plastid DNA-RRR pathways to be manipulated. Combining plastid transformation technologies with knockouts and knockdowns in nuclear genes is a particularly attractive method for studying plastid DNA-RRR pathways. Functional assays (see for example Fig. 11c) based on recombinant plastid genomes (Mühlbauer et al. 2002; Khakhlova and Bock 2006; Kode et al. 2006) will enable the impact of plastid DNA-RRR deficiencies on plastid DNA maintenance to be monitored. These new experimental approaches involving genomics, reverse genetics and transplastomic technologies, where both the trans-acting proteins and cis-acting DNA sequences can be manipulated, are likely to provide the functional studies needed to improve our understanding of the DNA-RRR pathways responsible for the maintenance of plastid genomes.

Acknowledgement

We are grateful to Professors Hans-Ulrich Koop (Munich) and Madeline Wu (Hong Kong) for providing information prior to publication and improvements to the text suggested by two anonymous reviewers. Work in the authors' laboratory was supported by the EU-FP6 Plastomics STREP consortium and the Biotechnology and Biological Sciences Research Council (UK).

References

Asai T, Bates DB, Kogoma T (1994) DNA replication triggered by double stranded breaks in *Escherichia coli*: dependence on homologous recombination functions. Cell 78:1051-1061

Backert S, Dorfel P, Börner T (1995) Investigation of plant organellar DNAs by pulsed-field gel-electrophoresis. Curr Genet 28:390-399

Bailey JC, Heinhorst S, Cannon GC (1995) Accuracy of deoxynucleotide incorporation by soybean chloroplast DNA polymerases is independent of the presence of a 3' to 5' exonuclease. Plant Physiol 107:1277-1284

Baumgartner BJ, Rapp JC, Mullet JE (1989) Plastid transcription activity and DNA copy number increase early in barley chloroplast development. Plant Physiol 89:1011-1018

Belkina GG, Pogul'skaya EV, Yurina NP (2004) Isolation and partial characterization of DNA topoisomerase I from the nucleoids of white mustard chloroplasts. Appl Biochem Microbiol 40:231-235
Bendich AJ (1987) Why do chloroplasts and mitochondria contain so many copies of their genome? Bioessays 6:279-282
Bendich AJ (2004) Circular chloroplast chromosomes: the grand illusion. Plant Cell 16:1661-1666
Bendich AJ, Smith SB (1990) Moving-pictures and pulsed-field gel-electrophoresis show linear DNA-molecules from chloroplasts and mitochondria. Curr Genet 17:421-425
Bensen RJ, Warner HR (1987) Partial purification and characterization of uracil-DNA glycosylase activity from chloroplasts of *Zea mays* Seedlings. Plant Physiol 84:1102-1106
Bi X, Liu LF (1996) Replicational model for DNA recombination between direct repeats. J Mol Biol 256:849-858
Birky CW (1994) Relaxed and stringent genomes: why cytoplasmic genes don't obey Mendel's Laws. J Heredity 85:355-365
Birky CW (2001) The inheritance of genes in mitochondria and chloroplasts: Laws, mechanisms, and models. Ann Rev Genet 35:125-148
Birky CW, Walsh JB (1992) Biased gene conversion, copy number, and apparent mutation rate differences within chloroplast and bacterial genomes. Genetics 130:677-683
Boynton JE, Gillham NW, Harris EH, Hosler JP, Johnson AM, Jones AR, Randolph-Anderson BL, Robertson D, Klein TM, Shark KB, Sanford JC (1988) Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. Science 240:1534-1538
Cahoon AB, Cunningham KA, Bollenbach TJ, Stern DB (2003) Maize BMS cultured cell lines survive with massive plastid gene loss. Curr Genet 44:104-113
CameriniOtero RD, Hsieh P (1995) Homologous recombination proteins in prokaryotes and eukaryotes. Ann Rev Genet 29:509-552
Cannon GC, Heinhorst S (1990) Partial purification and characterization of a DNA helicase from chloroplasts of *Glycine max*. Plant Mol Biol 15:457-464
Cannon GC, Hedrick LA, Heinhorst S (1995) Repair mechanisms of UV-induced DNA damage in soybean chloroplasts. Plant Mol Biol 29:1267-1277
Cao J, Combs C, Jagendorf AT (1997) The chloroplast located homolog of bacterial DNA recombinase. Plant Cell Physiol 38:1319-1325
Carrillo N, Bogorad L (1988) Chloroplast DNA replication *in vitro*: site-specific initiation from preferred templates. Nucl Acids Res 16:5603-5620
Casjens S (1999) Evolution of the linear DNA replicons of the *Borrelia* spirochetes. Curr Opin Microbiol 2:529-534
Cavalier-Smith T (1974) Palindromic base sequences and replication of eukaryote chromosome ends. Nature 250:467-470
Cerutti H, Ibrahim HZ, Jagendorf AT (1993) Treatment of pea (*Pisum sativum* L) Protoplasts with DNA-damaging agents induces a 39-kilodalton chloroplast protein immunologically related to *Escherichia coli* RecA. Plant Physiol 102:155-163
Cerutti H, Johnson AM, Boynton JE, Gillham NW (1995) Inhibition of chloroplast DNA recombination and repair by dominant negative mutants of *Escherichia coli* RecA. Mol Cell Biol 15:3003-3011
Cerutti H, Osman M, Grandoni P, Jagendorf AT (1992) A homolog of *Escherichia coli* RecA protein in plastids of higher plants. Proc Natl Acad Sci USA 89:8068-8072

Chang CH, Wu M (2000) The effects of transcription and RNA processing on the initiation of chloroplast DNA replication in *Chlamydomonas reinhardtii*. Mol Gen Genet 263:320-327

Chen WL, Gaikwad A, Mukherjee SK, Choudhary NR, Kumar D, Tewari KK (1996) A 43 kDa DNA binding protein from the pea chloroplast interacts with and stimulates the cognate DNA polymerase. Nucl Acids Res 24:3953-3961

Chiang KS, Sueoka N (1967) Replication of chloroplast DNA in *Chlamydomonas reinhardi* during vegetative cell cycle: its mode and regulation. Proc Natl Acad Sci USA 57:1506-1513

Chiu WL, Sears BB (1992) Electron microscopic localization of replication origins in *Oenothera* chloroplast DNA. Mol Gen Genet 232:33-39

Cho HS, Lee SS, Kim KD, Hwang I, Lim JS, Park YI, Pai HS (2004) DNA gyrase is involved in chloroplast nucleoid partitioning. Plant Cell 16:2665-2682

Christensen AC, Lyznik A, Mohammed S, Elowsky CG, Elo A, Yule R, Mackenzie SA (2005) Dual-domain, dual-targeting organellar protein presequences in *Arabidopsis* can use non-AUG start codons. Plant Cell 17:2805-2816

Chumley TW, Palmer JD, Mower JP, Fourcade HM, Calie PJ, Boore JL, Jansen RK (2006) The complete chloroplast genome sequence of *Pelargonium* x hortorum: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. Mol Biol Evol 23:2175-2190

Collin S, Ellis THN (1991) Evidence for the presence of hairpin chloroplast DNA molecules in barley cultivars. Curr Genet 20:253-258

Corneille S, Lutz K, Svab Z, Maliga P (2001) Efficient elimination of selectable marker genes from the plastid genome by the CRE-lox site-specific recombination system. Plant J 27:171-178

Corneille S, Lutz KA, Azhagiri AK, Maliga P (2003) Identification of functional lox sites in the plastid genome. Plant J 35:753-762

Corriveau JL, Coleman AW (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

Cox MM, Goodman MF, Kreuzer KN, Sherratt DJ, Sandler SJ, Marians KJ (2000) The importance of repairing stalled replication forks. Nature 404:37-41

Cuzzoni E, Giordani C, Stampacchia O, Bolchi A, Malcevschi A, Ottonello S, Ferretti L, Sala F (1995) Presence of a chloroplast DNA sequence in an autonomous circular DNA molecule in cultures rice cell (*Oryza sativa*). Plant Cell Physiol 36:717-72

Day A, Ellis THN (1984) Chloroplast DNA deletions associated with wheat plants regenerated from pollen: possible basis for maternal inheritance of chloroplasts. Cell 39:359-368

Day A, Ellis THN (1985) Deleted forms of plastid DNA in albino plants from cereal anther culture. Curr Genet 9:671-678

Dehaas JM, Boot KJM, Haring MA, Kool AJ, Nijkamp HJJ (1986) A *Petunia hybrida* chloroplast DNA region, close to one of the inverted repeats, shows sequence homology with the *Euglena gracilis* chloroplast DNA region that carries the putative replication origin. Mol Gen Genet 202:48-54

Deng XW, Wing RA, Gruissem W (1989) The chloroplast genome exists in multimeric forms. Proc Natl Acad Sci USA 86:4156-4160

Douglas SE, Penny SL (1999) The plastid genome of the cryptophyte alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae. J Mol Evol 48:236-244

Doyle JJ, Davis JI, Soreng RJ, Garvin D, Anderson MJ (1992) Chloroplast DNA inversions and the origin of the grass family (*Poaceae*). Proc Natl Acad Sci USA 89:7722-7726

Draper CK, Hays JB (2000) Replication of chloroplast, mitochondrial and nuclear DNA during growth of unirradiated and UVB-irradiated *Arabidopsis* leaves. Plant J 23:255-265

Drescher A, Ruf S, Calsa T, Carrer H, Bock R (2000) The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. Plant J 22:97-104

Dufourmantel N, Dubald M, Matringe M, Canard H, Garcon F, Job C, Kay E, Wisniewski JP, Ferullo JM, Pelissier B, Sailland A, Tissot G (2007) Generation and characterization of soybean and marker-free tobacco plastid transformants over-expressing a bacterial 4-hydroxyphenylpyruvate dioxygenase which provides strong herbicide tolerance. Plant Biotechnol J 5:118-133

Dufourmantel N, Tissot G, Goutorbe F, Garcon F, Muhr C, Jansens S, Pelissier B, Peltier G, Dubald M (2005) Generation and analysis of soybean plastid transformants expressing *Bacillus thuringiensis* Cry1Ab protoxin. Plant Mol Biol 58:659-668

Ellis THN, Day A (1986) A hairpin plastid genome in barley. EMBO J 5:2769-2774

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting sub-cellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol 300:1005-1016

Epp MD (1973) Nuclear gene-induced plastome mutations in *Oenothera hookeri*. I. Genetic analysis. Genetics 75:465-483

Evans E, Alani E (2000) Roles for mismatch repair factors in regulating genetic recombination. Mol Cell Biol 20:7839-7844

Fischer N, Stampacchia O, Redding K, Rochaix JD (1996) Selectable marker recycling in the chloroplast. Mol Gen Genet 251:373-380

Fujie M, Kuroiwa H, Kawano S, Mutoh S, Kuroiwa T (1994) Behavior of organelles and their nucleoids in the shoot apical meristem during leaf development in *Arabidopsis thaliana* L. Planta 194:395-405

Fukata H, Mochida A, Maruyama N, Fukasawa H (1991) Chloroplast DNA topoisomerase I from cauliflower. J Biochem 109:127-131

Futcher AB (1986) Copy number amplification of the 2-μm circle plasmid of *Saccharomyces cerevisiae*. J Theor Biol 119:197-204

Gaikwad A, Hop DV, Mukherjee SK (2002) A 70-kDa chloroplast DNA polymerase from pea (*Pisum sativum*) that shows high processivity and displays moderate fidelity. Mol Genet Genomics 267:45-56

Gilbert DM (2004) In search of the holy replicator. Nat Rev Mol Cell Biol 5:848-854

Gillham NW (1974) Genetic analysis of chloroplast and mitochondrial genomes. Ann Rev Genet 8:347-391

Gold B, Carrillo N, Tewari KK, 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

Goodman MF (2000) Coping with replication 'train wrecks' in *Escherichia coli* using Pol V, Pol II and RecA proteins. Trends Biochem Sci 25:189-195

Grant D, Swinton DC, Chiang KS (1978) Differential patterns of mitochondrial, chloroplastic and nuclear-DNA synthesis in synchronous cell-cycle of *Chlamydomonas reinhardtii*. Planta 141:259-267

Hada M, Hino K, Buchholz G, Goss J, Wellmann E, Shin M (2000) Assay of DNA photolyase activity in spinach leaves in relation to cell compartmentation: evidence for lack of DNA photolyase in chloroplasts. Biosci Biotechnol Biochem 64:1288-1291

Hagemann R (1976) Plastid distribution and plastid competition in higher plants and the induction of plastom mutations in higher plants by nitroso-urea compounts. In: Bücher T, Neupert W, Sebald W, Werner S (eds) Genetics and biogenesis of chloroplasts and mitochondria. Amsterdam: Elsevier, pp 331-338

Hajdukiewicz PTJ, Gilbertson L, Staub JM (2001) Multiple pathways for Cre/lox-mediated recombination in plastids. Plant J 27:161-170

Harada T, Ishikawa R, Niizeki M, Saito KI (1992) Pollen derived rice calli that have large deletions in plastid DNA do not require protein-synthesis in plastids for growth. Mol Gen Genet 233:145-150

Hartung F, Puchta H (2006) The Rec Q gene family in plants. J Plant Physiol 163:287-296

Hedrick LA, Heinhorst S, White MA, Cannon GC (1993) Analysis of soybean chloroplast DNA replication by 2-dimensional gel electrophoresis. Plant Mol Biol 23:779-792

Heinhorst S, Cannon G, Weissbach A (1985) Plastid and nuclear-DNA synthesis are not coupled in suspension cells of *Nicotiana tabacum*. Plant Mol Biol 4:3-12

Heinhorst S, Cannon GC, Weissbach A (1990) Chloroplast and mitochondrial DNA polymerases from cultured soybean cells. Plant Physiol 92:939-945

Herrmann RG, Bohnert HJ, Kowallik KV, Schmitt JM (1975) Size, conformation and purity of chloroplast DNA of some higher plants. Biochim Biophys Acta 378:305-317

Hess WR, Hubschmann T, Börner T (1994) Ribosome deficient plastids of albostrians barley: extreme representatives of nonphotosynthetic plastids. Endocytobiosis Cell Res 10:65-80

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

Howe CJ, Smith AG (1991) Plants without chlorophyll. Nature 349:109-109

Hubscher U, Nasheuer HP, Syvaoja JE (2000) Eukaryotic DNA polymerases, a growing family. Trends Biochem Sci 25:143-147

Iamtham S, Day A (2000) Removal of antibiotic resistance genes from transgenic tobacco plastids. Nature Biotechnol 18:1172-1176

Jacob F, Cuzin F, Brenner S (1963) On regulation of DNA replication in bacteria. Cold Spring Harbor Symp Quant Biol 28:329-348

Kaneko T, Nakamura Y, Wolk CP, Kuritz T, Sasamoto S, Watanabe A, Iriguchi M, Ishikawa A, Kawashima K, Kimura T, Kishida Y, Kohara M, Matsumoto M, Matsuno A, Muraki A, Nakazaki N, Shimpo S, Sugimoto M, Takazawa M, Yamada M, Yasuda M, Tabata S (2001) Complete genomic sequence of the filamentous nitrogen-fixing *Cyanobacterium anabaena* sp strain PCC 7120. DNA Res 8:205-213

Kavanagh TA, Thanh ND, Lao NT, McGrath N, Peter SO, Horvath EM, Dix PJ, Medgyesy P (1999) Homeologous plastid DNA transformation in tobacco is mediated by multiple recombination events. Genetics 152:1111-1122

Kawata M, Harada T, Shimamoto Y, Oono K, Takaiwa F (1997) Short inverted repeats function as hotspots of intermolecular recombination giving rise to oligomers of deleted plastid DNAs (ptDNAs). Curr Genet 31:179-184

Keim CA, Mosbaugh DW (1991) Identification and characterization of a 3' to 5' exonuclease Associated with spinach chloroplast DNA-polymerase. Biochemistry 30:11109-11118

Kelchner SA, Wendel JF (1996) Hairpins create minute inversions in non-coding regions of chloroplast DNA. Curr Genet 30:259-262

Khakhlova O, Bock R (2006) Elimination of deleterious mutations in plastid genomes by gene conversion. Plant J 46:85-94

Khazi FR, Edmondson AC, Nielsen BL (2003) An *Arabidopsis* homologue of bacterial RecA that complements an *E. coli* RecA deletion is targeted to plant mitochondria. Mol Genet Genomics 269:454-463

Kimura S, Sakaguchi K (2006) DNA repair in plants. Chem Rev 106:753-766

Kimura S, Uchiyama Y, Kasai N, Namekawa S, Saotome A, Ueda T, Ando T, Ishibashi T, Oshige M, Furukawa T, Yamamoto T, Hashimoto J, Sakaguchi K (2002) A novel DNA polymerase homologous to *Escherichia coli* DNA polymerase I from a higher plant, rice (*Oryza sativa* L.). Nucl Acids Res 30:1585-1592

Klaus SMJ, Huang FC, Golds TJ, Koop HU (2004) Generation of marker-free plastid transformants using a transiently cointegrated selection gene. Nature Biotechnol 22:225-229

Kleffmann T, Russenberger D, von Zychlinski A, Christopher W, Sjolander K, Gruissem W, Baginsky S (2004) The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. Curr Biol 14:354-362

Kode V, Mudd EA, Iamtham S, Day A (2005) The tobacco *acc*D gene is essential and is required for leaf development. Plant J 44:237-244

Kode V, Mudd EA, Iamtham S, Day A (2006) Isolation of precise plastid deletion mutants by homology-based excision: a resource for site-directed mutagenesis, multi-gene changes and high-throughput plastid transformation. Plant J 46:901-909

Kolodner R, Tewari KK (1972) Molecular-size and conformation of chloroplast deoxyribonucleic acid from pea leaves. J Biol Chem 247:6355-6364

Kolodner RD, Tewari KK (1975) Chloroplast DNA from higher plants replicates by both Cairns and rolling circle mechanism. Nature 256:708-711

Kolodner R, Tewari KK (1979) Inverted repeats in chloroplast DNA from higher plants. Proc Natl Acad Sci USA 76:41-45

Koumandou VL, Nisbet RER, Barbrook AC, Howe CJ (2004) Dinoflagellate chloroplasts: where have all the genes gone? Trends Genet 20:261-267

Kowalczykowski SC (2000) Initiation of genetic recombination and recombination-dependent replication. Trends Biochem Sci 25:156-165

Kowallik KV, Stoebe B, Schaffran I, KrothPancic P, Freier U (1995) The chloroplast genome of a chlorophyll a+c-containing alga, *Odontella sinensis*. Plant Mol Biol Rep 13:336-342

Kreuzer KN (2000) Recombination-dependent DNA replication in phage T4. Trends Biochem Sci 25:165-173

Kreuzer KN (2005) Interplay between DNA replication and recombination in prokaryotes. Ann Rev Microbiol 59:43-67

Kunnimalaiyaan M, Nielsen BL (1997a) Fine mapping of replication origins (oriA and oriB) in *Nicotiana tabacum* chloroplast DNA. Nucl Acids Res 25:3681-3686

Kunnimalaiyaan M, Shi F, Nielsen BL (1997b) Analysis of the tobacco chloroplast DNA replication origin (oriB) downstream of the 23 S rRNA gene. J Mol Biol 268:273-283

Kuroiwa T (1991) The replication, differentiation, and inheritance of plastids with emphasis on the concept of organelle nuclei. Int Rev Cytol 128:1-62

Lam E, Chua NH (1987) Chloroplast DNA gyrase and *in vitro* regulation of transcription by template topology and novobiocin. Plant Mol Biol 8:415-424

Lamppa GK, Bendich AJ (1979) Changes in chloroplast DNA levels during development of pea (*Pisum sativum*). Plant Physiol 64:126-130

Lawrence ME, Possingham JV (1986) Microspectrofluorometric measurement of chloroplast DNA in dividing and expanding leaf cells of *Spinacia oleracea*. Plant Physiol 81:708-710

Lelivelt CLC, McCabe MS, Newell CA, deSnoo CB, van Dun KMP, Birch-Machin I, Gray JC, Mills KHG, Nugent JM (2005) Stable plastid transformation in lettuce (*Lactuca sativa* L.). Plant Mol Biol 58:763-774

Li WM, Ruf S, Bock R (2006) Constancy of organellar genome copy numbers during leaf development and senescence in higher plants. Mol Genet Genomics 275:185-192

Lilly JW, Havey MJ, Jackson SA, Jiang JM (2001) Cytogenomic analyses reveal the structural plasticity of the chloroplast genome in higher plants. Plant Cell 13:245-254

Liu NT, Jane WN, Tsay HS, Wu H, Chang WC, Lin CS (2007) Chloroplast genome aberration in micropropagation-derived albino *Bambusa edulis* mutants, ab1 and ab2. Plant Cell Tiss Org Culture 88:147-156

Lovett ST (2004) Encoded errors: mutations and rearrangements mediated by misalignment at repetitive DNA sequences. Mol Microbiol 52:1243-1253

Lu Z, Kunnimalaiyaan M, Nielsen BL (1996) Characterization of replication origins flanking the 23S rRNA gene in tobacco chloroplast DNA. Plant Mol Biol 32:693-706

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci USA 99:12246-12251

Maul JE, Lilly JW, Cui LY, dePamphilis CW, Miller W, Harris EH, Stern DB (2002) The *Chlamydomonas reinhardtti* plastid chromosome: Islands of genes in a sea of repeats. Plant Cell 14:2659-2679

McKown RL, Tewari KK (1984) Purification and properties of a pea chloroplast DNA polymerase. Proc Natl Acad Sci USA 81:2354-2358

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

Meeker R, Nielsen B, Tewari KK (1988) Localization of replication origins in pea chloroplast DNA. Mol Cell Biol 8:1216-1223

Miyamura S, Kuroiwa T, Nagata T (1990) Multiplication and differentiation of plastid nucleoids during development of chloroplasts and etioplasts from proplastids in *Triticum aestivum*. Plant Cell Physiol 31:597-602

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

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

Mori Y, Kimura S, Saotome A, Kasai N, Sakaguchi N, Uchiyama Y, Ishibashi T, Yamamoto T, Chiku H, Sakaguchi K (2005) Plastid DNA polymerases from higher plants, *Arabidopsis thaliana*. Biochem Biophys Res Comm 334:43-50

Mühlbauer SK, Lossl A, Tzekova L, Zou ZR, Koop HU (2002) Functional analysis of plastid DNA replication origins in tobacco by targeted inactivation. Plant J 32:175-184
Mukherjee SK, Reddy MK, Kumar D, Tewari KK (1994) Purification and characterization of a eukaryotic type I topoisomerase from pea chloroplast. J Biol Chem 269:3793-3801
Muller HJ (1964) The relation of recombination to mutational advance. Mutation Res 1:2-9
Nakayama N, Arai N, Bond MW, Kaziro Y, Arai K (1984) Nucleotide sequence of *dna* B and the primary structure of the dna B protein from *Escherichia coli*. J Biol Chem 259:97-101
Nakazato E, Fukuzawa H, Tabata S, Takahashi H, Tanaka K (2003) Identification and expression analysis of cDNA encoding a chloroplast recombination protein REC1, the chloroplast RecA homologue in *Chlamydomonas reinhardtii*. Biosci Biotechnol Biochem 67:2608-2613
Nelson MJ, Green BR (2005) Double hairpin elements and tandem repeats in the noncoding region of *Adenoides eludens* chloroplast gene minicircles. Gene 358:102-110
Newman SM, Boynton JE, Gillham NW, Randolphanderson BL, Johnson AM, Harris EH (1990) Transformation of chloroplast ribosomal RNA genes in *Chlamydomonas*: molecular and genetic-characterization of integration events. Genetics 126:875-888
Newman SM, Harris EH, Johnson AM, Boynton JE, Gillham NW (1992) Nonrandom distribution of chloroplast recombination events in *Chlamydomonas reinhardtii*: evidence for a hotspot and an adjacent cold region. Genetics 132:413-429
Nielsen BL, Rajasekhar VK, Tewari KK (1991) Pea chloroplast DNA primase: characterization and role in initiation of replication. Plant Mol Biol 16:1019-1034
Ogihara Y, Isono K, Kojima T, Endo A, Hanaoka M, Shiina T, Terachi T, Utsugi S, Murata M, Mori N, Takumi S, Ikeo K, Gojobori T, Murai R, Murai K, Matsuoka Y, Ohnishi Y, Tajiri H, Tsunewaki K (2000) Chinese spring wheat (*Triticum aestivum* L.) chloroplast genome: Complete sequence and contig clones. Plant Mol Biol Rep 18:243-253
Ogihara Y, Isono K, Kojima T, Endo A, Hanaoka M, Shiina T, Terachi T, Utsugi S, Murata M, Mori N, Takumi S, Ikeo K, Gojobori T, Murai R, Murai K, Matsuoka Y, Ohnishi Y, Tajiri H, Tsunewaki K (2002) Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA. Mol Genet Genomics 266:740-746
Ohtani T, Uchimiya H, Kato A, Harada H, Sugita M, Sugiura M (1984) Location and nucleotide-sequence of a tobacco chloroplast DNA segment capable of replication in yeast. Mol Gen Genet 195:1-4
Oldenburg DJ, Bendich AJ (1998) The structure of mitochondrial DNA from the liverwort, *Marchantia polymorpha*. J Mol Biol 276:745-758
Oldenburg DJ, Bendich AJ (2001) Mitochondrial DNA from the liverwort *Marchantia polymorpha*: Circularly permuted linear molecules, head-to-tail concatemers, and a 5 ' protein. J Mol Biol 310:549-56
Oldenburg DJ, Bendich AJ (2004a) Changes in the structure of DNA molecules and the amount of DNA per plastid during chloroplast development in maize. J Mol Biol 344:1311-1330
Oldenburg DJ, Bendich AJ (2004b) Most chloroplast DNA of maize seedlings in linear molecules with defined ends and branched forms. J Mol Biol 335:953-970
Oldenburg DJ, Rowan BA, Zhao L, Watcher CL, Schleh M, Bendich AJ (2006) Loss or retention of chloroplast DNA in maize seedlings is affected by both light and genotype. Planta 225:41-55
Palmer JD (1983) Chloroplast DNA exists in 2 orientations. Nature 301: 92-93

Palmer JD (1990) Contrasting modes and tempos of genome evolution in land plant organelles. Trends Genet 6:115-120

Palmer JD, Thompson WF (1982) Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. Cell 29:537-550

Pang QS, Hays JB, Rajagopal I (1992) A plant cDNA that partially complements *Escherichia coli* RecA mutations predicts a polypeptide not strongly homologous to RecA proteins. Proc Natl Acad Sci USA 89:8073-8077

Pang QS, Hays JB, Rajagopal I (1993a) Two cDNAs from the plant *Arabidopsis thaliana* that partially restore recombination proficiency and DNA damage resistance to *Escherichia coli* mutants lacking recombination intermediate resolution activities. Nucl Acids Res 21:1647-1653

Pang QS, Hays JB, Rajagopal I, Schaefer TS (1993b) Selection of *Arabidopsis* cDNAs that partially correct phenotypes of *Escherichia coli* DNA damage sensitive mutants and analysis of 2 plant cDNAs that appear to express UV specific dark repair activities. Plant Mol Biol 22:411-426

Perry AS, Wolfe KH (2002) Nucleotide substitution rates in legume chloroplast DNA depend on the presence of the inverted repeat. J Mol Evol 55:501-508

Phinney BS, Thelen JJ (2005) Proteomic characterization of a triton-insoluble fraction from chloroplasts defines a novel group of proteins associated with macromolecular structures. J Proteome Res 4:497-506

Przykorska A, Solecka K, Olszak K, Keith G, Nawrot B, Kuligowska E (2004) Wheat (*Triticum vulgare*) chloroplast nuclease ChSI exhibits 5 ' flap structure-specific endonuclease activity. Biochemistry 43:11283-11294

Pyke KA, Marrison J, Leech RM (1989) Evidence for a Type-II topoisomerase in wheat chloroplasts. FEBS Letts 242:305-308

Qin ZJ, Cohen SN (2000) Long palindromes formed in *Streptomyces* by nonrecombinational intra-strand annealing. Genes Develop 14:1789-1796

Raynaud C, Perennes C, Reuzeau C, Catrice O, Brown S, Bergounioux C (2005) Cell and plastid division are coordinated through the prereplication factor AtCDT1. Proc Natl Acad Sci USA 102:8216-8221

Reboud X, Zeyl C (1994) Organelle inheritance in plants. Heredity 72:132-140

Reece RJ, Maxwell A (1991) DNA gyrase - structure and function. Crit Rev Biochem Mol Biol 26:335-375

Reith M, Munholland J (1995) Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. Plant Mol Biol Rep 13:333-335

Rogalski M, Ruf S, Bock R (2006) Tobacco plastid ribosomal protein S18 is essential for cell survival. Nucl Acids Res 34:4537-4545

Rowan BA, Oldenburg DJ, Bendich AJ (2004) The demise of chloroplast DNA in *Arabidopsis*. Curr Genet 46:176-181

Ruf S, Hermann M, Berger IJ, Carrer H, Bock R (2001) Stable genetic transformation of tomato plastids and expression of a foreign protein in fruit. Nat Biotechnol 19:870-875

Rybchin VN, Svarchevsky AN (1999) The plasmid prophage N15: a linear DNA with covalently closed ends. Mol Microbiol 33:895-903

Sakai A, Suzuki T, Nagata N, Sasaki N, Miyazawa Y, Saito C, Inada N, Nishimura Y, Kuroiwa T (1999) Comparative analysis of DNA synthesis activity in plastid-nuclei and mitochondrial-nuclei simultaneously isolated from cultured tobacco cells. Plant Sci 140:9-19

Sakai A, Takano H, Kuroiwa T (2004) Organelle nuclei in higher plants: Structure, composition, function, and evolution. Int Rev Cytol 238:59-118
Sala F, Amileni AR, Parisi B, Spadari S (1980) A gamma-like DNA polymerase in spinach chloroplasts. Eur J Biochem 112:211-217
Saotome A, Kimura S, Mori Y, Uchiyama Y, Morohashi K, Sakaguchi K (2006) Characterization of four Rec Q homologues from rice (*Oryza sativa* L. cv. Nipponbare). Biochem Biophys Res Commun 345:1283-1291
Saski C, Lee SB, Daniell H, Wood TC, Tomkins J, Kim HG, Jansen RK (2005) Complete chloroplast genome sequence of *Glycine max* and comparative analyses with other legume genomes. Plant Mol Biol 59:309-322
Sato N, Terasawa K, Miyajima K, Kabeya Y (2003) Organization, developmental dynamics, and evolution of plastid nucleoids. Int Rev Cytol 232:217-262
Scharff LB, Koop HU (2006) Linear molecules of tobacco ptDNA end at known replication origins and additional loci. Plant Mol Biol 62:611-621
Scharff LB, Koop HU (2007) Targeted inactivation of the tobacco plastome origins of replication A and B. Plant J: in press
Sears BB (1998) Replication, recombination and repair in the chloroplast genetic system of *Chlamydomonas*. In: Rochaix JD, Goldschmidt-Clermont M, Merchant S (eds) The molecular biology of chloroplasts and mitochondria in *Chlamydomonas*. Dordrecht-Boston-London: Kluwer, 115-138
Sears BB, Stoike LL, Chiu WL (1996) Proliferation of direct repeats near the *Oenothera* chloroplast DNA origin of replication. Mol Biol Evol 13:850-863
Seow F, Sato S, Janssen CS, Riehle MO, Mukhopadhyay A, Phillips RS, Wilson RJM, Barrett MP (2005) The plastidic DNA replication enzyme complex of *Plasmodium falciparum*. Mol Biochem Parasitol 141:145-153
Shaver JM, Oldenburg DJ, Bendich AJ (2006) Changes in chloroplast DNA during development in tobacco, *Medicago truncatula*, pea, and maize. Planta 224:72-82
Shen P, Huang HV (1986) Homologous recombination in *Escherichia coli* dependence on substrate length and homology. Genetics 112:441-457
Shikanai T, Shimizu K, Ueda K, Nishimura Y, Kuroiwa T, Hashimoto T (2001) The chloroplast *clp* P gene, encoding a proteolytic subunit of ATP-dependent protease, is indispensable for chloroplast development in tobacco. Plant Cell Physiol 42:264-273
Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchishinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5:2043-2049
Sidorov VA, Kasten D, Pang SZ, Hajdukiewicz PTJ, Staub JM, Nehra NS (1999) Stable chloroplast transformation in potato: use of green fluorescent protein as a plastid marker. Plant J 19:209-216
Siedlecki J, Zimmermann W, Weissbach A (1983) Characterization of a prokaryotic topoisomerase I activity in chloroplast cxtracts from spinach. Nucl Acids Res 11:1523-1536
Singh BN, Sopory SK, Reddy MK (2004) Plant DNA topoisomerases: Structure, function, and cellular roles in plant development. Crit Rev Plant Sci 23:251-269
Soltis PS, Soltis DE, Chase MW (1999) Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. Nature 402:402-404
Spencer D, Whitfeld PR (1969) Characteristics of spinach chloroplast DNA polymerase. Arch Biochem Biophy 132:477-488

Staub JM, Maliga P (1992) Long regions of homologous DNA are incorporated into the tobacco plastid genome by transformation. Plant Cell 4:39-45
Staub JM, Maliga P (1994) Extrachromosomal elements in tobacco plastids. Proc Natl Acad Sci USA 91:7468-7472
Sternberg N, Hamilton D (1981) Bacteriophage-P1 site-specific recombination. I. Recombination between LoxP sites. J Mol Biol 150:467-486
Stoike LL, Sears BB (1998) Plastome mutator-induced alterations arise in Oenothera chloroplast DNA through template slippage. Genetics 149:347-353
Streisinger G, Edgar RS, Denhardt GH (1964) Chromosome structure in phage T4, I. circularity of linkage map. Proc Natl Acad Sci USA 51:775-779
Suzuki H, Ingersoll J, Stern DB, Kindle KL (1997) Generation and maintenance of tandemly repeated extrachromosomal plasmid DNA in *Chlamydomonas* chloroplasts. Plant J 11:635-648
Svab Z, Hajdukiewicz P, Maliga P (1990) Stable transformation of plastids in higher plants. Proc Natl Acad Sci USA 87:8526-8530
Svab Z, Maliga P (1993) High frequency plastid transformation in tobacco by selection for a chimeric *aad*A Gene. Proc Natl Acad Sci USA 90:913-917
Szostak JW, Orrweaver TL, Rothstein RJ, Stahl FW (1983) The double strand break repair model for recombination. Cell 33:25-35
Takeda Y, Hirokawa H, Nagata T (1992) The replication origin of proplastid DNA in cultured cells of tobacco. Mol Gen Genet 232:191-198
Thompson RJ, Mosig G (1985) An ATP dependent supercoiling topoisomerase of *Chlamydomonas reinhardtii* affects accumulation of specific chloroplast transcripts. Nucleic Acids Res 13:873-891
Tuteja N (2003) Plant DNA helicases: the long unwinding road. J Exp Bot 54:2201-2214
Tuteja N, Phan TN (1998) A chloroplast DNA helicase II from pea that prefers fork-like replication structures. Plant Physiol 118:1029-1039
VanWinkle-Swift KP (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
Velemínský J, Švachulová J, Šatava J (1980) Endonucleases for UV-irradiated and depurinated DNA in barley chloroplasts. Nucl Acids Res 8:1373-1381
Waddell J, Wang XM, Wu M (1984) Electron microscopic localization of the chloroplast DNA replicative origins in *Chlamydomonas reinhardii*. Nucl Acids Res 12:3843-3856
Wagner R, Pfannschmidt T (2006) Eukaryotic transcription factors in plastids: bioinformatic assessment and implications for the evolution of gene expression machineries in plants. Gene 381:62-70
Wall MK, Mitchenall LA, Maxwell A (2004) *Arabidopsis thaliana* DNA gyrase is targeted to chloroplasts and mitochondria. Proc Natl Acad Sci USA 101:7821-7826
Wang TSF (1991) Eukaryotic DNA polymerases. Ann Rev Biochem 60:513-552
Wang Y, Saitoh Y, Sato T, Hidaka S, Tsutsumi K (2003) Comparison of plastid DNA replication in different cells and tissues of the rice plant. Plant Mol Biol 52:905-913
Wang ZF, Yang JM, Nie ZQ, Wu M (1991) Purification and characterization of a gamma-like DNA-polymerase from *Chlamydomonas reinhardtii*. Biochemistry 30:1127-1131
Williamson D (2002) Timeline-The curious history of yeast mitochondrial DNA. Nat Rev Genet 3:475-481

Woelfle MA, Thompson RJ, Mosig G (1993) Roles of novobiocin sensitive topoisomerases in chloroplast DNA replication in *Chlamydomonas reinhardtii*. Nucl Acids Res 21:4231-4238

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054-9058

Wolfe KH, Morden CW, Palmer JD (1992) Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. Proc Natl Acad Sci USA 89:10648-10652

Wycliffe P, Sitbon F, Wernersson J, Ezcurra I, Ellerstrom M, Rask L (2005) Continuous expression in tobacco leaves of a *Brassica napus* PEND homologue blocks differentiation of plastids and development of palisade cells. Plant J 44:1-15

Ye JS, Sayre RT (1990) Reduction of chloroplast DNA content in *Solanum nigrum* suspension cells by treatment with chloroplast DNA synthesis inhibitors. Plant Physiol 94:1477-1483

Zoschke R, Liere K, Börner T (2007) From seedling to mature plant: *Arabidopsis* plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. Plant J 50:710-722

Zoubenko OV, Allison LA, Svab Z, Maliga P (1994) Efficient targeting of foreign genes into the tobacco plastid genome. Nucl Acids Res 22:3819-3824

Zubko MK, Day A (1998) Stable albinism induced without mutagenesis: a model for ribosome-free plastid inheritance. Plant J 15:265-271

Zubko MK, Day A (2002) Differential regulation of genes transcribed by nucleus-encoded plastid RNA polymerase, and DNA amplification, within ribosome- deficient plastids in stable phenocopies of cereal albino mutants. Mol Gen Genomics 267:27-37

Zubko MK, Zubko EI, van Zuilen K, Meyer P, Day A (2004) Stable transformation of petunia plastids. Transgenic Res 13:523-530

Day, Anil

Faculty of Life Sciences, The University of Manchester, Oxford Road, Manchester M13 9PT, UK

anil.day@manchester.ac.uk

Madesis, Panagiotis

Faculty of Life Sciences, The University of Manchester, Oxford Road, Manchester M13 9PT, UK

List of abbreviations

DAPI: 4',6-diamidino-2-phenylindole
DSB: Double strand break
D-loop: Displacement loop
DNA-RRR: DNA replication, recombination, and repair
FISH: Fluorescent *in situ* hybridization
G. max: *Glycine max* (soybean)
H. vulgare: Hordeum vulgare (barley)
IR: Inverted repeat
L. sativa: *Lactuca sativa* (lettuce)
MMS: methyl methane sulfonate
MW: molecular weight
N. tabacum: Nicotiana tabacum (tobacco)
O. hookeri: Oenothera hookeri (evening primrose)
O. sativa: *Oryza sativa* (rice)
Ori: origin of replication
P. sativum: *Pisum sativum* (pea)
RBS: Ribosome binding site
S. oleracea: Spinacia oleracea (spinach)
T. aestivum: Triticum aestivum (wheat)
UTR: Untranslated region
WT: wild type
Z. mays: Zea mays (maize)

Transcription and transcriptional regulation in plastids

Karsten Liere and Thomas Börner

Abstract

This chapter describes the components of the transcriptional apparatus in plastids (RNA polymerases, promoters, transcription factors) and their roles in transcription. The chromosomes of plastids from nearly all plants contain genes for core subunits of PEP, a bacterial-type RNA polymerase which might be responsible for transcription of all plastid genes in algae but shares responsibility for transcription with one or more nuclear encoded transcriptases (NEP) in higher plants. There is increasing evidence that the catalytic subunit of NEP is related to RNA polymerases of bacteriophages like T7. NEP and PEP are active throughout leaf development. Transcription of plastid genes and operons by multiple promoters is common. Promoter recognition by PEP is mediated by σ-factors. Factors supporting NEP in promoter binding are not known yet. Examples of regulation of transcription are described demonstrating promoter selection by σ-factors and activation/repression of gene activity by transcription factors.

1 Introduction

Plastids divide in a similar manner as bacteria. Each plastid in a plant contains identical circular copies of the plastid chromosome, the plastome. In addition to monomeric circles, dimers, trimers, and tetramers exist, but also numerous linear and even more complex molecules of different sizes. The number of plastids per cell and of plastomes per plastid changes species-specifically from cell-type to cell-type and may vary during the development of plants (Butterfass 1980; Herrmann and Possingham 1980; López-Juez and Pyke 2005; see Chapters 2, 3, 4). Adjusting the copy number of plastomes per cell could be a way to respond to different needs for plastid gene products, in particular of rRNAs as Bendich (1987) suggested. The striking increase of plastome copies at the beginning of the development of chloroplasts from proplastids is certainly a precondition for the biogenesis of the photosynthetic apparatus in young leaf cells. Other ways to control gene expression at the DNA level could be *via* alteration of the DNA conformation (Stirdivant et al. 1985; Gauly and Kössel 1989; Sekine et al. 2002) or differential methylation (Ngernprasirtsiri et al. 1989; Kobayashi et al. 1990; Ngernprasirtsiri and Akazawa 1990). Both ways are investigated in only a few

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0232 / Published online: 13 June 2007

studies and at least the latter one may be an exception rather than the rule (Hess et al. 1993; Isono et al. 1997a).

Early studies on gene expression in chloroplasts revealed specific effects of light on the expression of the *psbA* gene and of the cell type (mesophyll vs. bundle sheath cells) on the expression of *rbcL* (Bedbrook et al. 1978; Link et al. 1978) suggesting an important role of differential transcription like in bacteria. Further studies during the 1980's, however, revealed important contributions of posttranscriptional processes in controlling the levels of gene products (plastid RNAs and proteins) and only a minor role for transcription in the regulation of gene expression in plastids during plant development (Deng and Gruissem 1987; Gruissem et al. 1988). While the importance of posttranscriptional processes in the control of RNA and protein levels remained undisputed until today (see Chapters 6-10), the view on the role of transcription changed again during the 1990's with the discovery of differential transcription of house-keeping vs. photosynthesis genes, of light-induced differential transcription of several genes, and of additional promoters within operons (Mullet 1993). Moreover, the machinery for transcription in plastids of angiosperms unexpectedly turned out to be more complex as known from bacteria and to need different RNA polymerases, although plastids possess a much smaller genome than their cyanobacterial ancestors (Stern et al. 1997; Gray and Lang 1998; Hess and Börner 1999; Liere and Maliga 2001). This chapter describes the components of the transcriptional apparatus in plastids and their roles in transcription and its regulation.

2 RNA polymerases

2.1 NEP: nuclear-encoded plastid RNA polymerase

2.1.1 Evidence for the existence of PEP and NEP

Transcription of all bacterial genes is performed by one core RNA polymerase consisting of 4 subunits (two α, one β, and one β'; e.g. in *E. coli*), or 5 subunits in case of the cyanobacteria which have β' split into two subunits, β' and β'' (Kaneko et al. 1996). The plastid chromosomes of algae and higher plants posses genes for core subunits of a cyanobacterial-type RNA polymerase, first reported for *Marchantia*, tobacco and spinach (Ohyama et al. 1986; Shinozaki et al. 1986; Sijben-Muller et al. 1986), which is commonly abbreviated as PEP (for *p*lastid-*e*ncoded *p*lastid RNA *p*olymerase; Hajdukiewicz et al. 1997; see Section 1.2; Fig. 1). The existence of one or more *n*uclear-*e*ncoded *p*lastid RNA *p*olymerase(s) (NEP) was suggested by comparing the effects of inhibitors of translation on cytoplasmic and plastidial ribosomes, respectively (Ellis and Hartley 1971). Detection of RNA polymerase in plastids with impaired protein synthesis implies a nuclear location of the gene(s) encoding this activity. Ribosome-deficient plastids isolated from heat-bleached rye leaves were found to exhibit RNA polymerase activity (Bünger and Feierabend 1980), and the detection of mature rRNAs in plastids that lack

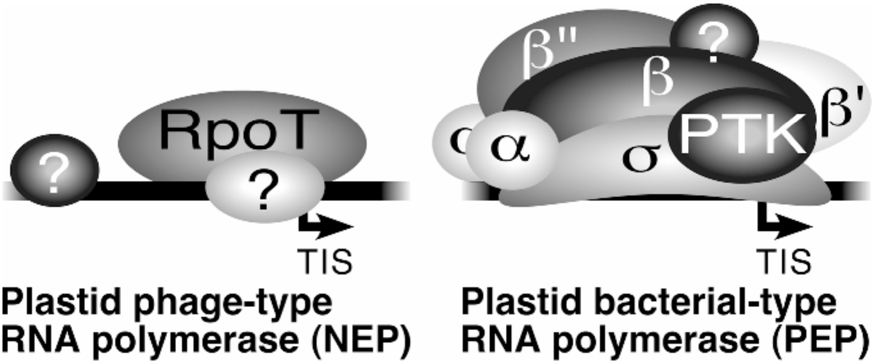


Fig. 1. RNA polymerases in plastids. The nuclear-encoded plastid RNA polymerase (NEP) is related to phage-type single-subunit enzymes and may need additional, yet unknown protein factors for promoter recognition. The plastid-encoded plastid RNA polymerase (PEP) is a multisubunit enzyme homologous to bacterial RNA polymerases and consists of the core α_2, β, β', and β'' subunits and the nuclear-encoded σ-like factor required for promoter recognition. In chloroplasts, PEP associates with additional factors, which are thought to be involved in regulation of PEP transcription activity, including the plastid transcription kinase (PTK). The transcription initiation sites (TIS) are indicated by arrows.

ribosomes was reported as proof for accurately functioning nuclear-encoded RNA polymerase and rRNA processing activities in plastids of the barley mutant *albostrians* (Siemenroth et al. 1981). These early data have later been confirmed by demonstrating the expression of several genes in ribosome-free plastids of barley mutants (Hess et al. 1993), of heat-bleached rye leaves (Falk et al. 1993; Hess et al. 1993), and of the *iojap* mutant of maize (Han et al. 1993). Further evidence for NEP activity came from the detection of RNA synthesis in nonphotosynthetic plastids of the parasitic plant *Epifagus virginiana* (Ems et al. 1995). *E. virginiana* has a relatively small plastid genome that lacks genes for proteins involved in photosynthesis and, important in this context, for the core subunits of PEP (Morden et al. 1992). Similar observations have been made more recently with other parasitic plants, where transcription has to rely solely on NEP activity as their plastomes lack PEP genes (Lusson et al. 1998; Krause et al. 2003; Berg et al. 2004). The invention of genetic manipulation of plastid genes allowed for the directed inactivation of PEP genes. Plants with deleted PEP genes still were able to transcribe their plastid genes, i.e., provided additional evidence for the existence of NEP. Moreover, the albino phenotype of these plants indicated that NEP activity alone is not sufficient for the development of photosynthetically active chloroplasts (Allison et al. 1996; Hajdukiewicz et al. 1997; Krause et al. 2000; Legen et al. 2002).

2.1.2 Phage-type RNA polymerases in plants

There is increasing evidence that the catalytic subunit of NEP is related to RNA polymerases of bacteriophages like T3, T7, or SP6 (Fig. 1). Lerbs-Mache (1993)

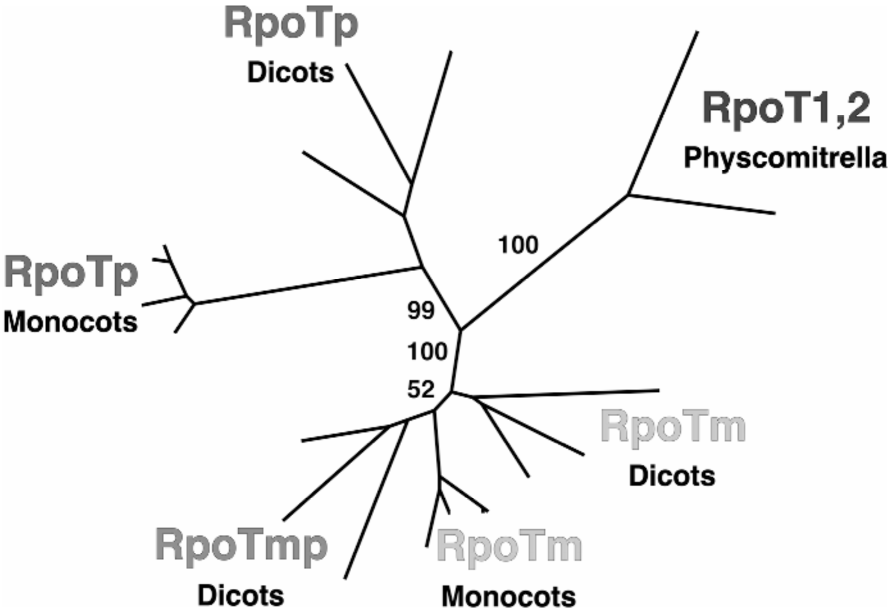


Fig. 2. Schematic representation of RpoT phylogeny. Distances are proportional to relative sequence divergence; numbers at nodes show branch support values. *RpoT* gene duplications occurred several times during the evolution of land plants. Reconstruction of phylogeny was done by quartet puzzling (based on data published in Emanuel et al. 2006).

observed an RNA polymerase activity of a 110 kDa protein (the size expected for RpoT products) that was prepared from spinach chloroplasts and initiated transcription from a T7 promoter but not from an *rbcL* PEP promoter. It has been known for many years that the mitochondrial genes of baker's yeast, *Saccharomyces cerevisiae*, are transcribed by a nuclear encoded phage-type RNA polymerase (Masters et al. 1987). It is now evident that related phage-type polymerases are responsible for mitochondrial transcription in nearly all eukaryotes. The only exceptions from this rule are freshwater protozoa like *Reclinomonas* belonging to the jakobids. These lower eukaryotes still possess genes for a bacterial-type RNA polymerase in their mitochondrial genomes which became lost during the evolution of this organelle in the other lineages of eukaryotes (Lang et al. 1997). Genes potentially coding for RNA polymerases of the phage-type are also found on so-called 'linear plasmids', double-stranded DNAs of around 10 kb that have been detected in the mitochondria of several protozoa, fungi, and plants. Neither the origin of these 'plasmids', that exhibit features of viral genomes, nor the functional roles of their genes are known (Meinhardt et al. 1997). Phylogenetic trees of phage-type RNA polymerases suggest that the nuclear gene for the mitochondrial RNA polymerase evolved independently of the plasmid-localized RNA polymerase genes probably from an ancestral bacteriophage gene (Lysenko and Kuznetsov 2005; Azevedo et al. 2006; Emanuel et al. 2006). Genes coding for phage-type RNA polymerases duplicated several times during the evolution of plants (Fig. 2). They were first discovered in *Arabidopsis* and *Chenopodium* and designated as *RpoT* genes (for *R*NA *po*lymerase of the phage *T*3/*T*7 type; Hedtke et al. 1997; Weihe et al. 1997). Meanwhile it is evident that the nuclear genomes of dicotyledonous and monocotyledonous plants contain more than one *RpoT* gene (Fig. 3). The diploid genomes of the eudicots *Arabidopsis* (Hedtke et al. 1997, 2000), *Nicotiana silvestris* (Kobayashi et al. 2001a, 2001b, 2002), and *Populus* (deduced from the sequences data in http://genome.jgi-psf.org; Tuskan et al. 2006) contain 3 *RpoT* genes. The amphidiploid genome of tobacco, *N. tabacum*, contains 6 *RpoT* genes (two sets of three genes, one set each from the two diploid parental species; Hedtke et al. 2002).

The N-termini of the different RpoT polymerases (RpoTm, RpoTp, and RpoTmp) fused to GFP (green fluorescence protein) target the protein to mitochondria, plastids, and both organelles, respectively (Hedtke et al. 1997, 1999, 2000, 2002; Kobayashi et al. 2001a, 2001b, 2002). It has therefore been suggested that the *RpoT* genes encode mitochondrial (RpoTm; the *Arabidopsis* gene was originally designated as *RpoY* and *RpoT;1*; Hedtke et al. 1997, 2000), plastid (RpoTp; originally RpoZ and RpoT;3), and dual-targeted RNA polymerases (RpoTmp; originally RpoT;2). Targeting to one or the other organelle might be regulated at the level of translation as the *RpoTmp* mRNAs contain two potential start codons for translation, a feature which is conserved for all RpoTmp messengers of dicots and even *Physcomitrella* (see below). If the first start codon with a position more close to the 5′ end is used, a transit peptide is synthesized that allowed for transportation of the protein into both organelles. If the exclusive usage of the second start codon was forced by deletion of the first one, the smaller transit peptide imported GFP only into mitochondria (Hedtke et al. 2000, 2002; Kobayashi et al. 2001a; Richter et al. 2002).

Monocots (only cereals have so far been investigated) have only two *RpoT* genes, one coding for a mitochondrial (*RpoTm*), the other for a plastidial RNA polymerase (*RpoTp*; Chang et al. 1999; Ikeda and Gray 1999; Emanuel et al. 2004; Kusumi et al. 2004; Fig. 3). Also the moss *Physcomitrella patens* contains two *RpoT* genes (Kabeya et al. 2002; Richter et al. 2002). Other plants have not been studied so far. The *Physcomitrella* genes were named *RpoTmp1* and *RpoTmp2* by Richter et al. (2002), since it was observed that the putative transit peptides encoded by both genes mediated dual targeting of GFP to plastids and mitochondria like in the case of the RpoTmp polymerases of dicots. However, RpoTmp localization is still a matter of debate. Targeting of GFP to mitochondria but not to plastids was observed in *Arabidopsis* or *Physcomitrella* when the protein was fused not only with the putative RpoTmp targeting sequence but also with the 5′-flanking UTR. For yet unknown reasons, the presence of the 5′-UTR prevents usage of the first start codon during translation of the *Arabidopsis* and *Physcomitrella* RpoTmp mRNAs. As mentioned above, translation from the second start codon produces a transit peptide for import into mitochondria (Kabeya and Sato 2005). The authors proposed, therefore, that dual targeting of GFP fused to the putative RpoTmp transit peptide alone, i.e. without the 5′-UTR, was an experimental artifact and these genes encode mitochondrial RNA polymerases. On the other hand, exclusive targeting to mitochondria is not in agreement with the observation

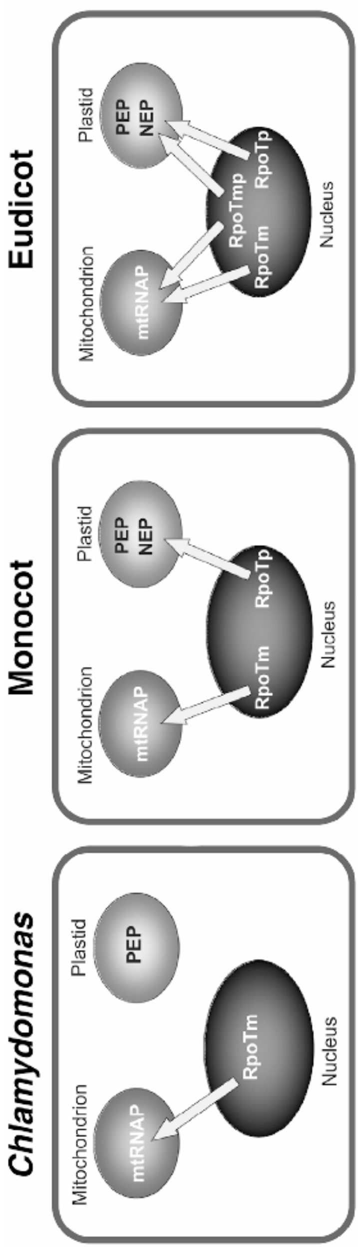


Fig. 3. *Chlamydomonas* possesses only one nuclear *RpoT* gene that is proposed to encode the mitochondrial RNA polymerase (mtRNAP). Cereals have two RpoT genes, one encoding the mtRNAP, the other a plastid RNA polymerase supposed to represent NEP, whereas *Arabidopsis* and other eudicots additionally acquired RpoTmp, which may contribute to transcription in mitochondria and plastids.

that the RpoTmp homolog of spinach was detected in chloroplasts but not in mitochondria with antibodies reacting specifically with this enzyme (Azevedo et al. 2006). Moreover, studies on mutants lacking RpoTmp suggested a function in plastid transcription (Baba et al. 2004; Hricova et al. 2006; see below). Obviously, more detailed investigations into the localization of RpoTmp are required and the possibility of a regulation of targeting to mitochondria and/or plastids at the level of translation should be considered (Christensen et al. 2005).

Whether algae need NEP in addition to PEP to transcribe their plastid genes is not known yet (see review by Smith and Purton 2002). Like higher plants, algae bear genes for the core subunits of PEP in their plastid genomes. In contrast to *Epifagus* (see above), even the nonphotosynthetic alga *Astasia longa*, the malaria parasite *Plasmodium falciparum*, and related organisms have plastid (apicoplast) PEP genes (Wilson et al. 1996; Gockel and Hachtel 2000; Sheveleva et al. 2002) suggesting that a NEP activity is lacking. The nuclear genome of *Chlamydomonas* contains only one *RpoT* gene (A. Weihe et al., unpublished data). There are no experimental data on the subcellular localization of the *Chlamydomonas RpoT* gene product, but it likely codes for the mitochondrial RNA polymerase as in the other eukaryotes that possess only a single gene for a phage-type RNA polymerase (Fig. 3). Inhibition of transcription in *Chlamydomonas* chloroplasts by inhibitors which are specific for the bacterial-type RNA polymerase and would not affect the activity of phage-type RNA polymerases led to a complete block of plastid gene expression arguing against the presence of NEP activity in this alga (Surzycki 1969; Guertin and Bellemare 1979). Also another alga, *Osteococcus tauri*, possesses only one *RpoT* gene, which likely encodes the mitochondrial RNA polymerase (W. Hess, T. Börner, H. Moreau, unpublished; Derelle et al. 2006).

2.1.3 Function of RpoT polymerases in higher plants

Only little information is available on the function of RpoT polymerases in plants. Heterologously expressed RpoTp, RpoTmp, and RpoTm enzymes of *Arabidopsis* are active RNA polymerases that prefer circular over linear template DNA. RpoTm and RpoTp (not RpoTmp) exhibit an inherent ability to recognize several mitochondrial and at least one NEP promoter *in vitro* (Kühn et al. 2007). In monocots with two *RpoT* genes, RpoTm is assumed to represent the catalytic subunit of the mitochondrial RNA polymerase and RpoTp the catalytic subunit of NEP. RpoTp has been detected by specific antibodies in the chloroplasts of rice and maize (Chang et al. 1999; Kusumi et al. 2004) and RpoTm in maize mitochondria (Chang et al. 1999). *RpoTp* mRNAs are particularly abundant in very young cells of cereal leaves (Chang et al. 1999; Emanuel et al. 2004; Kusumi et al. 2004) in agreement with the proposed importance of NEP activity early in chloroplast development for transcription of the PEP genes (Mullet 1993; see Section 4.1). Expression of *RpoTm* and *RpoTp* in monocots is under control of light (Chang et al. 1999) and plastid signal(s) (Emanuel et al. 2004). In *Arabidopsis*, *RpoTm* and *RpoTmp* promoters showed identical expression patterns with highest levels in tissues known for their requirement of high respiration activity (e.g. meristems, tapetum) suggesting a function of both polymerases in mitochondria, whereas

RpoTp expression was highest in green tissues of leaves, stems, and sepals (Emanuel et al. 2006). Like in monocots, transcription of the *RpoT* genes is stimulated by light in *Arabidopsis* leaves (T. Preuten, K. Liere, T. Börner, unpublished results), i.e., light-activated expression of phage-type RNA polymerases may be a general phenomenon in angiosperms. Evidence for NEP being represented by RpoTp (probably together with RpoTmp; see below) was provided by studies on transgenic *Nicotiana* and *Arabidopsis* plants that overexpressed RpoTp and exhibited an increased usage of certain NEP promoters (Liere et al. 2004). Mutation of the *Arabidopsis RpoTp* gene led to impaired chloroplast biogenesis and altered accumulation of plastid transcripts (Hricova et al. 2006). Similar observations were made on *Arabidopsis* plants with reduced RpoTp mRNA levels due to expression of antisense RNA (Emanuel et al., unpublished data). Although the localization of RpoTmp in mitochondria is not in doubt (Kabeya and Sato 2005), its function for this organelle remains obscure so far. RpoTmp was supposed, however, to play a role in plastid gene expression (Baba et al. 2004; Hricova et al. 2006). *Arabidopsis* lines with impaired *RpoTmp* function were delayed in chloroplast biogenesis and showed altered plastid transcript levels (Baba et al. 2004). *RpoTp/RpoTmp* double mutants exhibited a more severe phenotype than both of the single mutants and were extremely retarded in growth (Hricova et al. 2006).

Clearly, more studies are needed to exactly define the function of the different organellar RNA polymerases. First insights into the division of labor between PEP (the bacterial type RNA polymerase) and NEP (probably represented by RpoTp in monocots and RpoTp and RpoTmp in dicots) were obtained from investigations on the use of PEP *vs.* NEP promoters in different tissues and under the influence of different endogenous and exogenous factors as discussed below.

2.2 PEP: plastid-encoded plastid RNA polymerase

The chromosomes of plastids from nearly all plants contain genes for core subunits of PEP, a bacterial-type RNA polymerase, which might be responsible for transcription of all plastidial genes in algae but shares responsibility for plastid transcription with one or more NEP enzymes in higher plants (see above). The *rpoA* gene codes for the 38-kDa α-subunit of PEP (Little and Hallick 1988; Ruf and Kössel 1988; Hu and Bogorad 1990). Like in bacteria, it forms an operon together with several ribosomal protein-encoding genes (Purton and Gray 1989). The *Physcomitrella* plastome lacks this gene. Instead, *rpoA* is found in the nuclear genome (Sugiura et al. 2003). A similar situation was reported for the bacterial-type RNA polymerase in chloroplasts of several algae (Smith and Purton 2002) and in the plastid-like organelles (apicoplasts) of *Plasmodium* (Wilson et al. 1996). The ß- (120 kDa), ß'- (85 kDa), and ß''-subunits (185 kDa) are encoded by the *rpoB*, *rpoC1*, and *rpoC2* genes, respectively, which together form an operon, exactly as known from cyanobacteria (Ohyama et al. 1986; Hudson et al. 1988; Little and Hallick 1988; Hu et al. 1991; Kaneko et al. 1996; Shinozaki et al. 1986; reviewed in Lysenko and Kuznetsov 2005). The structural relationship of the *E. coli* RNA polymerase and PEP was confirmed by reconstituting a functional *E.*

coli enzyme with polypeptides truncated as in PEP (Severinov et al. 1996). The high degree of conservation kept by PEP during evolution from the bacterial RNA polymerase is also demonstrated by its sensitivity to tagetitoxin (e.g. Mathews and Durbin 1990; Sakai et al. 1998). Other potent inhibitors of transcription in bacteria, rifampicin, and its related drugs, were also shown to inhibit transcription by the *E. coli*-like form of PEP found in etioplasts but not by the more complex form in chloroplasts (Fig. 1; e.g. Surzycki 1969; Loiseaux et al. 1975; Pfannschmidt and Link 1997). Furthermore, replacing the PEP α-subunit with the *E. coli* homologue in transplastomic tobacco resulted in a non-functional PEP enzyme, indicating that the evolutionary conservation of both α-subunits is insufficient to allow such an exchange (Suzuki and Maliga 2000).

The *rpoBC* operon is under control of a NEP promoter in monocotyledonous and dicotyledonous plants (see Section 3.1). Transcript levels of *rpo* genes are low compared with genes for proteins involved in photosynthesis (e.g. Hess et al. 1993; Legen et al. 2002). For promoter recognition, the core subunits have to be complemented by a sigma (σ) factor. Sigma factors are encoded by nuclear genes in all embryophytes (see Section 4.2.2) ensuring together with NEP a control of plastid transcription by the nucleus.

While NEP activity (demonstrated by recognition of NEP promoters) could be found hitherto only in soluble fractions of plastid lysates, PEP can be isolated from plastids as a 'soluble' (DNA-dependent) enzyme and in a 'insoluble' (DNA-associated) form together with DNA and other proteins of unknown function as the so-called 'transcriptionally active chromosome' (TAC; e.g. Briat et al. 1979; Greenberg et al. 1984; Little and Hallick 1988; Suck et al. 1996; Krause and Krupinska 2000; Pfalz et al. 2006). In the case of *Euglena*, the soluble RNA polymerase fraction and TAC were reported to transcribe different sets of genes. If this is due to the presence of different RNA polymerases, as discussed by Little and Hallick (1988), different transcriptions factors, or has other reasons is unclear yet (Smith and Purton 2002). The soluble PEP fraction contains different proteins and exhibits different sensitivity against rifampicin if prepared from etioplasts *vs.* chloroplasts. PEP isolated from etioplasts of mustard seedlings consists mainly of the core subunits (Pfannschmidt and Link 1997), whereas PEP preparations from chloroplasts were found to be more complex and contain additional proteins that might be needed for transcription and regulation of transcription under the conditions of light and active photosynthesis (Pfannschmidt and Link 1994, 1997; Link 1996; Baginsky et al. 1999; Pfannschmidt et al. 2000; Ogrzewalla et al. 2002) as discussed below.

3 Plastidial Promoters

3.1 NEP promoters

Unambiguous identification of transcription initiation sites for a nuclear-encoded transcription activity (i.e. NEP) became feasible in plants with reduced or eliminated transcriptional activity by PEP. Such systems comprise the ribosome-deficient plastids of the monocot *albostrians* barley and *iojap* maize mutants (Hübschmann and Börner 1998; Silhavy and Maliga 1998a), tobacco Δ*rpo* plants (Allison et al. 1996; Hajdukiewicz et al. 1997; Serino and Maliga 1998), *Arabidopsis* lacking PEP due to the action of spectinomycin which blocks plastidial protein synthesis (Swiatecka-Hagenbruch et al. 2007), and photosynthetically inactive tobacco and rice suspension cultures, with elevated levels of NEP activity (Vera et al. 1996; Kapoor et al. 1997; Miyagi et al. 1998; Silhavy and Maliga 1998b).

Most non-photosynthetic genes involved in housekeeping functions such as transcription and translation have promoters for both RNA polymerases NEP and PEP. NEP transcripts of these genes are, with a few exceptions, rarely detectable in chloroplasts and were therefore mostly analyzed in PEP-deficient plants (see above). Only a few genes are known to be transcribed exclusively from a NEP promoter, i.e. *accD,* encoding a subunit of the acetyl-CoA carboxylase in dicots; *ycf2*, encoding a protein with a yet unknown function; *rpl23*, encoding a ribosomal protein; *clpP*, encoding the proteolytic subunit of the Clp ATP-dependent protease, in monocots; and, most interestingly, the *rpoB* operon encoding three of the four PEP core subunits in all higher plants (Hajdukiewicz et al. 1997; Hübschmann and Börner 1998; Silhavy and Maliga 1998a; Swiatecka-Hagenbruch et al. 2007). Consequently, PEP abundance and activity depends on the nuclear-encoded RNA polymerase.

NEP promoters analyzed thus far resemble mitochondrial and phage promoters in their structural organization. Based on their sequence properties they can be grouped into three types (Fig. 4; Weihe and Börner 1999; Liere and Maliga 2001). Type-I promoters are characterized by a conserved YRTa-motif critical for *rpoB* promoter recognition embedded in a small DNA fragment (-15 to +5) upstream of the transcription initiation site (+1) (P*atpB*-289; Kapoor and Sugiura 1999; Xie and Allison 2002; P*accD*-129; Liere and Maliga 1999b; P*rpoB*-345; Liere and Maliga 1999a). Transient expression of chimeric *Arabidopsis rpoB* 5'-flanking region::GUS deletion-constructs in cultured tobacco cells suggested upstream regulatory regions for *rpoB* expression (Inada et al. 1997). However, no additional sequence elements outside the promoter core altered *rpoB* transcription *in vitro* (Liere and Maliga 1999a). Similar transient transcription assays to examine the 5'-flanking region of the tobacco *accD* gene revealed putative sequence elements up- and downstream of the promoter to determine its strength (Hirata et al. 2004). A subset of Type-I NEP promoters possesses a second conserved sequence motif (ATAN$_{0\text{-}1}$GAA) ~18 to 20 bp upstream of the YRTa-motif, designated box II or GAA-box (Fig. 4; Silhavy and Maliga 1998a; Kapoor and Sugiura 1999). Mutational analyses of the tobacco P*atpB*-289 promoter in *in vitro* and *in vivo* transcription experiments suggested a functional role of this element in promoter recognition (Kapoor and Sugiura 1999; Xie and Allison 2002). Hence, Type-I promoters are grouped into two subgroups: Ia, with only the YRTa-motif, and Ib, carrying both YRTa- and GAA-box (Weihe and Börner 1999; Liere and Börner 2007).

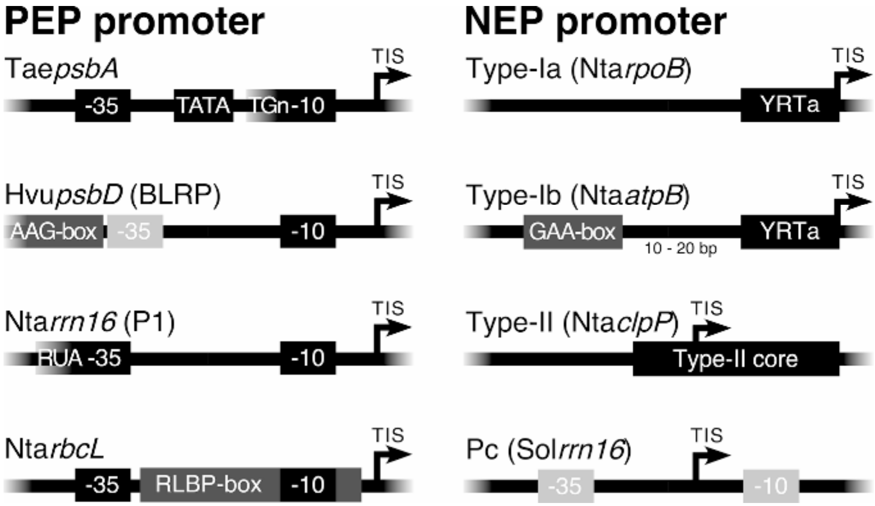


Fig. 4. Schematic overview of different types of PEP and NEP promoters. PEP promoter: the wheat *psbA* (Tae*psbA*), barley *psbD* BLRP (Hvu*psbD*), tobacco *rrn16* (Nta*rrn16*), and the tobacco *rbcL* PEP promoters (Nta*rbcL*) are shown. Conserved -10/-35 consensus elements, as well as individual promoter elements as the TATA-box (Eisermann et al. 1990), extended -10 sequence (TGn; Satoh et al. 1999), AAG-box (Kim et al. 1999), RUA-element (Suzuki et al. 2003), and RLBP-binding region (Kim et al. 2002) are indicated. The less conserved -35 element in the barley *psbD* BRLP is shown in grey. NEP promoter: typical architectures of Type-I, Type-II, and Pc NEP promoters from tobacco and spinach are shown with their names in brackets. The YRTa promoter core and GAA-box are marked (Hübschmann and Börner 1998; Kapoor and Sugiura 1999; Liere and Maliga 1999). TIS: transcription initiation site, indicated by arrows. The -35 and -10 elements not used in spinach *rrn16* promoter recognition are shown in grey.

Type-II NEP promoters lack the YRTa-motif and differ completely in sequence and organization from Type-I promoters. So far this class is represented by a single example, a promoter of the ClpP protease subunit gene (Fig. 4). The tobacco P*clpP*-53 was characterized using a transplastomic *in vivo* approach demonstrating that critical promoter sequences are located mainly downstream of the transcription initiation site (–5 to +25; Sriraman et al. 1998a). The *clpP*-53 promoter motif and transcription initiation site are conserved among monocots, dicots, conifers, and liverworts. But, although present, the tobacco P*clpP*-53 sequence motif is not used as a promoter in rice and *Chlamydomonas*. If the rice sequence is introduced into tobacco plastids, the tobacco NEP recognizes this conserved Type-II promoter. Therefore, the lack of transcription in rice from the P*clpP*-53 homologue may be resulting from either the lack of a Type-II specificity factor or to the lack of a distinct NEP enzyme not present in monocots (e.g. RpoTmp, see Chapter 1.1; Sriraman et al. 1998a; Liere et al. 2004). However, experimental data supporting one or the other of these scenarios are still missing.

Another non-YRTa-type promoter, which is attributed to be recognized by a NEP transcription activity is the *rrn* operon Pc promoter in spinach and *Arabidop-*

sis (Fig. 4, 6; Pc promoter; Baeza et al. 1991; Iratni et al. 1994, 1997; Sriraman et al. 1998a; Swiatecka-Hagenbruch et al. 2007). In spinach, the Pc promoter solely drives *rrn* operon transcription. Although it contains typical σ^{70}-elements which are active as the *rrn* operon promoter in other species, transcription initiates from a site between the conserved –10/–35 hexamers. However, sequences relevant for transcription initiation from Pc have yet to be identified.

3.2 PEP promoters

Having coevolved with the bacterial-type RNA polymerase (PEP), many plastidial promoters contain a variant of the -35 (TTGaca) and -10 (TAtaaT) consensus sequences of typical σ^{70}-type *E. coli* promoters (Reznikov et al. 1985; for reviews see Gruissem and Tonkyn 1993; Link 1994; Hess and Börner 1999; Liere and Maliga 2001; Weihe 2004). In fact, the *E. coli* RNA polymerase is able to accurately recognize plastidial σ^{70}-type promoters (e.g. Gatenby et al. 1981; Bradley and Gatenby 1985; Boyer and Mullet 1988; Eisermann et al. 1990). Since plastidial σ^{70}-type promoters are recognized by PEP, they are also often termed PEP promoters. In addition to the core motifs, some PEP promoters contain regulatory *cis*-elements. One of the best-characterized PEP promoters ensures transcription of the *psbA* gene, which encodes the D1 photosystem II reaction center polypeptide (Link 1984; Gruissem and Zurawski 1985; Boyer and Mullet 1986, 1988). *In vivo psbA* transcription is developmentally timed and activated by light (Klein and Mullet 1990; Schrubar et al. 1990; Baumgartner et al. 1993). *In vitro* characterization of the mustard *psbA* promoter identified a TATATA promoter element between the –10 and –35 hexamers resembling the TATA-box of nuclear genes transcribed by RNA polymerase II (Fig. 4; Eisermann et al. 1990; Link 1994). Basic transcription levels in plastidial extracts prepared from both dark and light grown plants were obtained *in vitro* with both the TATATA element together with the -10 region. Nonetheless, presence of the -35 element was essential for enhanced transcription rates characteristic of chloroplasts of light-grown plants (Link 1984; Eisermann et al. 1990). In barley, the *psbA* promoter also contains the TATA-motif between the -35/-10-elements. But, unlike in mustard, the –35 sequence is absolutely required for transcription *in vitro* (Kim et al. 1999b). Similarly, such TATA-box is also present in the wheat *psbA* promoter, but does not seem to be important. Light-independent (constitutive) transcription by PEP isolated from the leaf base (base-type PEP; young plastids) required both the –10 and –35 elements for promoter activity. However, PEP isolated from the leaf tip (tip-type PEP; mature plastids) employed only the –10 region with an additional TGn motif upstream of the –10 element (Fig. 4; extended –10 sequence; Bown et al. 1997; Satoh et al. 1999). The extended –10 sequence may be involved in promoter recognition by the tip-type PEP in mature plastids indicating that basal- and tip-type PEPs may differ by their associated transcription factors (Satoh et al. 1999). Since the mustard, barley and wheat *psbA* promoter sequences are highly conserved, differences in the utilization of *cis*-elements possibly are the result of a divergent evolution of *trans*-factors in these species.

Interestingly, it seems that most plastidial promoters in *Chlamydomonas* do not possess a valid -35 element, but rather a downstream extended -10 box (Klein et al. 1992). Furthermore, even remote sequences such as the coding regions are needed for full promoter strength of the *rbcL* and *psbA* but not *psbD*, *atpA*, and *atpB* genes (Blowers et al. 1990; Klein et al. 1994; Ishikura et al. 1999; Kasai et al. 2003). However, the mechanism of transcriptional enhancement by these *cis*-acting elements within the coding regions is not yet examined and might be unique for *Chlamydomonas* (Shiina et al. 1998; Kasai et al. 2003). Further regulatory sequences in addition to the core promoter regions were identified in the proximity of the *psbD-psbC* and *rbcL* promoters in higher plants (Fig. 4; see Section 4.1 for details).

3.3 Internal promoters of tRNAs

Most plastidial tRNAs are transcribed by the PEP from upstream σ^{70}-type promoters. However, transcription from internal promoters is assumed for some tRNA genes such as the spinach *trnS*, *trnR*, and *trnT* (Gruissem et al. 1986; Cheng et al. 1997b) as well as *trnS*, *trnH*, and *trnR* from mustard (Neuhaus and Link 1990; Nickelsen and Link 1990; Liere and Link 1994), and the *Chlamydomonas trnE* gene (Jahn 1992). Transcription of the spinach *trnS* gene is initiated twelve nucleotides upstream of the mature tRNA coding region (Wu et al. 1997). *In vitro* assays demonstrated that the coding region (+1/+93) promoted basal levels (8%) of transcription. Inclusion of an AT-rich sequence stretch between -31 and -11 upstream of the coding region restored wild type promoter strength. However, no sequences resembling either NEP or PEP promoters were found in this region. As most tRNAs, the *trnS* coding region contains sequences resembling the A and B blocks of nuclear tRNA promoters transcribed by the eukaryotic RNA polymerase III (Galli et al. 1981; Geiduschek et al. 1995). The $tRNA^{Arg}$(ACG) gene from *Pelargonium zonale* was efficiently transcribed in *Xenopus* oocyte nuclei (Hellmund et al. 1984), suggesting that the plastidial tRNAs may be transcribed by an RNA polymerase III-type enzyme. The biochemical properties and enzyme composition of such a transcription activity, however, remain to be determined. Thus far, *in silico* analyses of the *Arabidopsis* genome did not reveal a plastid-targeted polymerase of this type (Liere and Börner, unpublished). Alternatively, such tRNAs may be transcribed by specialized NEP or PEP enzymes associated with distinct transcription factors recognizing internal promoter structures.

4 Regulation of transcription in plastids

Expression of nuclear-encoded plastid-localized gene products is thought to be managed by transcriptional control (Kuhlemeier 1992). While posttranscriptional events contribute significantly to regulation of plastidial gene expression (see Chapters 6, 7; Deng and Gruissem 1987; Stern et al. 1997; Barkan and Goldschmidt-Clermont 2000; Monde et al. 2000), transcription of plastid genes was also shown to react to exogenous and endogenous factors such as light and plastid type (Rapp et al. 1992; Mullet 1993; Mayfield et al. 1995; Link 1996).

The circadian rhythm of plastidial gene expression in *Chlamydomonas* is regulated by transcriptional activity (Salvador et al. 1993; Hwang et al. 1996). Kawazoe et al. (2000) could show that the circadian clock-induced transcription is sensitive to cycloheximide, an inhibitor of cytoplasmic translation. However, basal plastidial transcription activity was still maintained. The identity of the cycloheximide-sensitive factor(s) needed for circadian peaks of plastidial transcription is still unknown. Expression of the sole σ-factor gene *CreRpoD* (Section 3.2.2; Carter et al. 2004; Bohne et al. 2006) seems also to be under circadian control (Carter et al. 2004). Therefore, a possible dual role of CreRpoD, which might be assisted by topological fluctuations of the plastome (Thompson and Mosig 1990; Salvador et al. 1998), in regulating plastidial gene transcription in *Chlamydomonas* has been discussed (Misquitta and Herrin 2005).

Transcription activities of most plastid-encoded genes in higher plants increase at an early stage of light-induced plastid development to support rapid construction of the photosynthesis apparatus. Moreover, light-dependent plastid transcription occurs in mature leaves as well as leaves under greening (Greenberg et al. 1989; Schrubar et al. 1990; Baumgartner et al. 1993; DuBell and Mullet 1995; Hoffer and Christopher 1997; Shiina et al. 1998; Satoh et al. 1999; Baena-Gonzalez et al. 2001; Chun et al. 2001; Nakamura et al. 2003). Most prominent examples are photosynthesis-related genes as *psbA*, *psbD-psbC*, *petG*, *rbcL*, but also housekeeping genes as *atpB* (Klein et al. 1988; Haley and Bogorad 1990; Klein and Mullet 1990; Sexton et al. 1990; Isono et al. 1997a). Distinctive photoreceptors involved in transcriptional activation of photosynthesis-related genes have been analyzed (Chun et al. 2001; Thum et al. 2001). The developmental stage may influence perception of the light quality. While red light only partially increased plastid transcription, blue light further enhanced overall plastid transcription activity in dark-adapted mature leaves. Therefore, global activation of plastidial transcription after dark adaptation is likely to be mediated by cryptochromes. When exposed to blue light/UV-A an *Arabidopsis phyA*-mutant displayed lower *psbA* and *rrn16* transcript activities than the wild type suggesting a further role for PhyA in light reception (Chun et al. 2001). Recently, Dhingra et al. (2006) furthermore showed that green light plays a balancing/antagonistic role in controlling gene expression during early photomorphogenic development by downregulating plastidial transcription of genes normally induced by light. As illustrated before, transcriptional response to developmental and environmental changes is likely to involve interaction of the core RNA polymerase with specific regulatory molecules (e.g. σ-factors), which may be available only under certain conditions. *In silico* analyses of nuclear *Arabidopsis* and rice genes with putative chloroplast transit peptides revealed many putative transcription factors likely to be imported into plastids (Wagner and Pfannschmidt 2006; Schwacke et al. 2007).

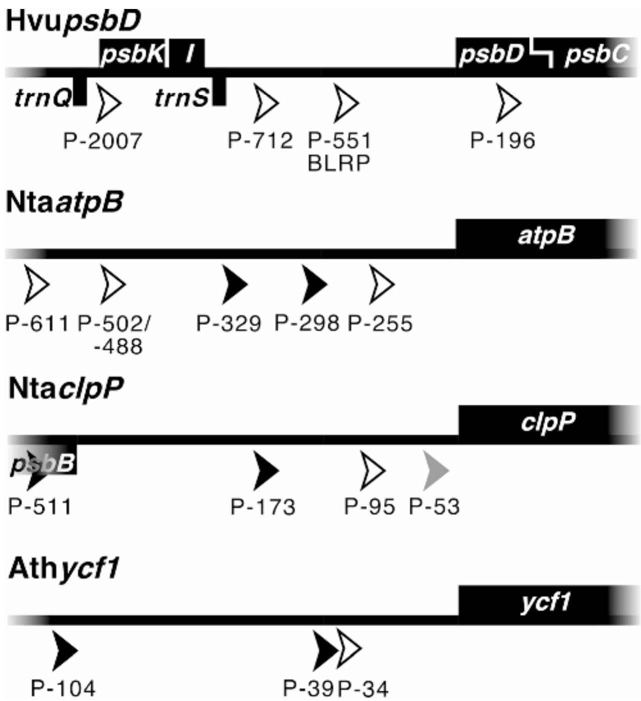


Fig. 5. Genes with multiple promoters. Schematic synopsis which shows the multiple PEP promoters of the barley *psbD/C* operon (Hvu*psbD*), as well as PEP and NEP promoters of the tobacco *atpB* (Nta*atpB*), *clpP* (Nta*clpP*), and *Arabidopsis ycf1* genes (Ath*ycf1*). Boxes below the line represent genes on the opposite strand, while open arrowheads denote PEP promoters, filled black arrowheads Type-I NEP promoters, and filled gray arrowheads Type-II NEP promoters. The promoters are named based on their transcription initiation sites in respect to the translation initiation site (+1).

These factors may represent such candidates to expand the actually known capacity of the chloroplast to regulate its transcription machinery. Additionally, various pathways routing developmental and environmental cues may regulate these factors.

4.1 Role of multiple and diverse promoters

Although genes exist that are transcribed from a single promoter, transcription of plastidial genes and operons by multiple promoters seems to be a common feature. For example, the *psbD-psbC* operon is transcribed from up to three different PEP promoters (Fig. 5; Yao et al. 1989; Berends Sexton et al. 1990; Christopher et al. 1992; Wada et al. 1994; To et al. 1996; Hoffer and Christopher 1997) and the tobacco *rpl32* gene from two promoters far upstream of the coding region (NEP-P*rpl32*-1101, PEP-P*rpl32*-1030; Vera et al. 1996). Similarly, the tobacco *atpB*

Table 1. Diversity of promoter usage in different species.

Gene	Arabidopsis	tobacco	maize	barley
accD	NEP: P*accD*-251; P*accD*-172	NEP: P*accD*-129	n.d.	n.d.
atpB	PEP: P*atpB*-520 NEP: P*atpB*-318	PEP: P*atpB*-611; P*atpB*-502; P*atpB*-255 NEP: P*atpB*-329; P*atpB*-289	PEP: P*atpB*-298 NEP: P*atpB*-601	PEP: n.d. NEP: P*atpB*-593
atpI	PEP: P*atpI*-229	PEP: P*atpI*-130 NEP: P*atpI*-207	n.d.	n.d.
clpP	PEP: P*clpP*-115 NEP II: P*clpP*-57	PEP: P*clpP*-95 NEP: P*clpP*-511; P*clpP*-173 NEP II: P*clpP*-53	NEP: P*clpP*-111	NEP: P*clpP*-133
psaA	PEP: P*psaA*-188	PEP: P*psaA*-194	PEP: P*psaA*-175	n.d.
psbA	PEP: P*psbA*-77	PEP: P*psbA*-85	PEP: P*psbA*-86	PEP: P*psbA*-80
rpoB	NEP: P*rpoB*-300	NEP: P*rpoB*-345	NEP: P*rpoB*-147	NEP: P*rpoB*-147
rrn16	PEP: P*rrn16*-112 NEP Pc: P*rrn16*-139	PEP: P*rrn16*-114 NEP: P*rrn16*-62	PEP: P*rrn16*-117	PEP: P*rrn16*-118
ycf1	PEP: P*ycf1*-34 NEP: P*ycf1*-39; P*ycf1*-104	NEP: P*ycf1*-41	n.d.	n.d.

PEP denotes PEP promoters, NEP represents NEP Type-I promoters, NEP II indicates NEP Type-II promoters, and NEP Pc denotes Pc promoters; n.d. indicates not yet identified promoters.

gene is transcribed from at least three NEP (P*atpB*-255, -502/-488, -611) and two PEP promoters (Fig. 5; P*atpB*-289, -329; Hajdukiewicz et al. 1997), but only one PEP and one NEP promoter are driving this gene in *Arabidopsis* (PEP-P*atpB*-520, NEP-P*atpB*-318; Swiatecka-Hagenbruch et al. 2007) and maize (NEP-P*atpB*-601, PEP-P*atpB*-298; Silhavy and Maliga 1998a). In case of *clpP*, the tobacco gene has two Type-I NEP (P*clpP*-173, -511), one PEP (P*clpP*-95) and the main Type-II NEP initiation sites (Fig. 5; P*clpP*-53; Hajdukiewicz et al. 1997; Sriraman et al. 1998a). The *Arabidopsis clpP* gene has a PEP (P*clpP*-115) and a Type-II NEP initiation site (P*clpP*-58; Sriraman et al. 1998a; Swiatecka-Hagenbruch et al. 2007). The maize gene, however, is transcribed from a sole Type-I NEP promoter (P*clpP*-111; Silhavy and Maliga 1998a) indicating a high diversity in promoter usage in different species (see Table 1 for a comparison of promoters of more plastidial genes in different plants). Furthermore, an increasing number of genes are reported to be co-transcribed with other genes within an operon and to additionally possess an individual promoter upstream of their coding region (e.g. *trnG* and *psbA*; Meng et al. 1991; Nickelsen and Link 1991; Kapoor et al. 1994; Liere and Link 1994; Liere et al. 1995).

The *rrn16* promoters are an interesting and well investigated example of the diversity of promoter usage within a highly conserved DNA sequence even in closely related species. The main *rrn* operon promoter in tobacco is a σ^{70}-type PEP promoter (P1 or Nt-P*rrn*-114; Vera and Sugiura 1995; Allison et al. 1996). In barley, maize, and pea the *rrn* operon is also transcribed from the P1 σ^{70}-type PEP promoter (Fig. 6; Strittmatter et al. 1985; Sun et al. 1989; Hübschmann and Börner 1998). Additionally, the *rrn* operon in tobacco has a NEP promoter (Fig. 6; P2 or Nta-P*rrn*-62), inactive in chloroplasts, but functional in BY2 tissue culture cells and in plants lacking PEP (Vera and Sugiura 1995; Allison et al. 1996). Conversely, there is no active NEP promoter directly upstream of the *rrn* operon in maize plastids (Silhavy and Maliga 1998a). In spinach chloroplasts transcription of the *rrn* operon initiates within a region between the promoter elements of P1 (Fig. 4, 6; Pc promoter; Baeza et al. 1991; Iratni et al. 1994, 1997). However, the σ^{70}-type promoter sequences are not utilized *in vivo*. Interestingly, the Pc site appears to be faithfully recognized by partially purified mustard PEP *in vitro* (Pfannschmidt and Link 1997). A good candidate for the Pc activating factor in spinach is CDF2 (see Section 4.2.1; Iratni et al. 1994, 1997; Bligny et al. 2000).

In *Arabidopsis*, *rrn* operon transcripts were mapped to both the major PEP P1 and the spinach Pc initiation sites (Fig. 6; Sriraman et al. 1998b; Swiatecka-Hagenbruch et al. 2007). A study of *rrn* promoters in heterologous plastids indicates that tobacco plastids lack the factor required for transcription from Pc, while spinach has an intact P1 promoter but lacks the cognate P1 activator (Sriraman et al. 1998b). However, in tobacco an rRNA operon upstream activator region (RUA) that is conserved in monocot and dicot species has been identified (Fig. 4; Suzuki et al. 2003). It has been suggested that the -10 element plays only a limited role in *rrn16* P1 recognition and that σ-factor interaction is replaced in part by direct PEP-RUA (protein-DNA) interaction or by protein-protein interaction between the PEP and a putative RUA-binding transcription factor.

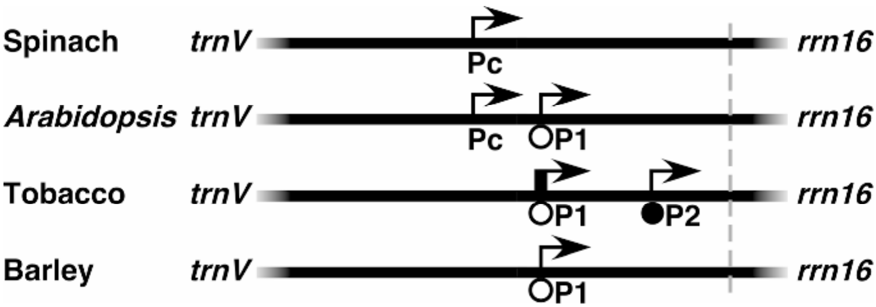


Fig. 6. Diverse promoters of the *rrn* operon in spinach, *Arabidopsis*, tobacco, and barley. The distinct promoters that are used in different species are shown by the schematic representation of the transcription initiation sites between *trnV* and *rrn16* (marked by arrows). P1 and P2 mark transcription initiation sites by PEP (open circle) and NEP (filled circle). Transcript initiation in spinach and *Arabidopsis* from a yet uncharacterized NEP promoter is indicated by Pc. A dashed vertical line indicates an RNA processing site.

Interestingly, *ycf1* in *Arabidopsis* is transcribed from a strongly conserved NEP promoter as in tobacco (NEP-AthP*ycf1*-39; Swiatecka-Hagenbruch et al. 2007). However, a PEP promoter located at the NEP promoter position takes over transcription in green leaves (PEP-AthP*ycf1*-34). With the *rrn16* Pc and PEP promoters in *Arabidopsis*, this is a rare incident where a defined DNA sequence serves as a promoter for both NEP and PEP.

The role of most multiple promoters upstream of plastidial genes and operons is not fully understood, however, some are well characterized. The blue-light-responsive promoter (BRLP) amongst the three PEP promoters of the *psbD-psbC* operon, for example, is thought to differentially maintain the ability to re-synthesize and replace damaged D2 and CP43 photosystem components in mature chloroplasts.

In spite of the observed diversity of plastidial promoter usage between different species of higher plants, the data support also the existence of common themes in promoter usage that have been deduced mainly from studies on transcription in tobacco plastids. Mixed NEP and PEP promoters typically are found upstream of housekeeping genes which need to be transcribed during full plastidial development (Maliga 1998). Consequently, both promoter types are believed to differentially express their cognate gene during plant development (reviewed in Liere and Maliga 2001). NEP promoters are generally recognized in youngest and non-green tissues early in plant development, while PEP takes over in maturating, photosynthetically active chloroplasts (Bisanz-Seyer et al. 1989; Baumgartner et al. 1993; Hajdukiewicz et al. 1997; Kapoor et al. 1997; Emanuel et al. 2004).

This simple model has been challenged by results from transcriptional re-analyses of tobacco Δ*rpo* mutants lacking PEP (Krause et al. 2000; Legen et al. 2002). Large spurious transcripts initiated by NEP cover the entire plastome in these mutants, suggesting that besides selective promoter utilization, posttranscriptional processes also determine the transcript pattern of plastids. Furthermore, it has been shown in maize, that although the NEP enzyme becomes less abundant as chloroplasts mature its transcriptional activity increases (Cahoon et al. 2004). The stability of the RNA generated by NEP, however, declines during chloroplast development. For transcripts generated by PEP, transcription rates increase as chloroplasts develop, whereas RNA stability remains constant or increases. Hence, in a proposed model for maize plastidial biogenesis, NEP-controlled transcript accumulation changes little during plastidial development while PEP-controlled transcript accumulation increases (Cahoon et al. 2004). In other species, a strong correlation between the transcribing enzyme (NEP or PEP) and the pattern of transcript accumulation was not observed (Zoschke et al. 2007).

Since genes exclusively transcribed by NEP encode housekeeping functions like the *rpoB* gene/operon and *rps15*, NEP should be still necessary for proper gene expression and regulation also in mature chloroplasts. Furthermore, an additional role of NEP as an SOS-enzyme in plastidial transcription has been proposed by Schweer et al. (2006). Analyses of transcript accumulation of *atpB* in an Ath-Sig6 knockout mutant suggested that a further upstream located NEP promoter compensates for failing transcription from the main PEP promoter. Indeed, NEP and PEP are active throughout leaf development in *Arabidopsis*, although PEP plays a major role in mature leaves (Cahoon et al. 2004; Zoschke et al. 2007). Interestingly, exclusively PEP-transcribed genes code for proteins with a role in photosynthesis. As the major active polymerase in mature chloroplasts, present data point to PEP as a prominent target for regulation signals including redox control, not yet determined for NEP (for review see Forsberg et al. 2001; Liere and Maliga 2001; Pfannschmidt and Liere 2005). Since plants that turn to a parasitic lifestyle lost photosynthetic genes as well as PEP promoters (Wolfe et al. 1992a; Wolfe et al. 1992b; Krause et al. 2003; Berg et al. 2004), transcription and regulation of plastidial gene expression by PEP might be connected to photosynthesis.

4.2 Transcription factors involved in promoter recognition in plastids

4.2.1 NEP transcription factors

Recent *in vitro* studies of the yeast mitochondrial transcription machinery unexpectedly revealed promoter specificity to be conferred by the core RNAP rather than mtTFB (Matsunaga and Jaehning 2004). Similarly, *in vitro* transcription assays with recombinant AthRpoTm and AthRpoTp enzymes showed accurate initiation of transcription from overlapping subsets of mitochondrial and plastidial promoters without auxiliary factors, therefore retaining a characteristic feature of the T7 RNAP. However, AthRpoTm and AthRpoTp failed to recognize some of the investigated promoters and AthRpoTmp displayed no significant promoter specificity while showing high non-specific transcription activity. Therefore, it is evident that the *Arabidopsis* enzymes need auxiliary factors for transcription *in organello* like the mitochondrial RNA polymerases of other organisms (Kühn et al. 2007).

Thus far, identification of factors involved in specific promoter recognition and transcription initiation by NEP has failed. Based on information on such factors interacting with the related mitochondrial phage-type RNA polymerases from humans, mice, *Xenopus laevis*, and yeast one can only speculate. These mitochondrial transcription complexes consist of a minimum of two components: the catalytic core enzyme (mtRPO, ~ 120–150 kDa), and a specificity factor, which confers promoter recognition (mtTFB, ~ 40–45 kDa). Despite poor overall sequence similarity, it recently has been shown that mtTFB factors belong to a family of RNA-methyltransferases (Falkenberg et al. 2002; McCulloch et al. 2002; Rantanen et al. 2003; Seidel-Rogol et al. 2003). An additional component, which binds the DNA further upstream, enhances mitochondrial transcription *in vitro* (mtTFA, 20–25 kDa). This DNA-binding protein belongs to the HMG (high mobility group) family and may also facilitate the interaction with other trans-acting factors (reviewed in Jaehning 1993; Shadel and Clayton 1993; Tracy and Stern 1995; Hess and Börner 1999). To date, no functional mtTFA or mtTFB homologues have been isolated from plant mitochondria or plastids, and the presence of such proteins in plant organelles is unclear. BLAST searches of the *Arabidopsis* genome revealed a TFB-like dimethyladenosine transferase gene, which possesses an N-terminal transit peptide mediating protein import into plastids of isolated tobacco protoplasts (B. Kuhla, K. Liere, T. Börner; unpublished data). This gene corresponds to the previously characterized *PFC1* gene encoding a plastid 16S rRNA dimethylase homologous to the yeast nucleolar 18S rRNA dimethylase Dim1 (Tokuhisa et al. 1998). The phenotype of *PFC1*-knockout mutants, however, does not support the idea that this TFB-like dimethyladenosine transferase may act as a primary transcription factor for the phage-type RNA polymerases (M. Swiatecka-Hagenbruch, K. Liere, T. Börner, unpublished data).

A good candidate for an activating factor for NEP transcription in spinach is CDF2, which has been reported to stimulate transcription of the *rrn* operon Pc promoter by NEP-2, a yet to be characterized nuclear-encoded transcription activity (Table 3; Bligny et al. 2000). CDF2 is supposed to exist in two distinct forms, CDF2-A and CDF2-B. CDF2-A might repress transcription initiation of PEP at the *rrn16* P1 promoter (termed P2 in spinach), while CDF2-B possibly binds NEP-2 and initiates specific transcription from the *rrn16* Pc promoter.

Another factor that is discussed to be involved in NEP transcription is the plastidial ribosomal protein L4 (RPL4; encoded by the nuclear *Rpl4* gene). A role for RPL4 in NEP transcription was proposed, as it co-purifies with the T7-like transcription complex in spinach (Trifa et al. 1998). In prokaryotes the ribosomal protein L4 was shown to have extra-ribosomal functions in transcriptional regulation (Zengel et al. 1980). The spinach and *Arabidopsis Rpl4* genes have acquired remarkable 3' extensions during evolutionary transfer to the nuclear genome, which resemble highly acidic C-terminal ends of certain transcription factors. A function for this protein in NEP or PEP transcription, however, has yet to be demonstrated.

Besides, some nucleus-encoded σ-factors for the bacterial-type PEP in plastids were found to additionally localize to mitochondria (see Section 4.2.2). So far phage-type RNA polymerases are the sole transcription activity in mitochondria of higher plants. One may speculate that these σ-factors may have an additional role in regulating mitochondrial transcription by these RNA polymerases (H. Tandara and K. Liere, unpublished data; Beardslee et al. 2002; Yao et al. 2003). Yet, experimental data to link the activity of the bacterial-type plastidial σ-factors to the phage-type enzymes in mitochondria or plastids are still lacking.

4.2.2 Nuclear-encoded plastidial σ-factors

Specific transcription initiation in bacteria requires a transcription factor (σ), which is responsible for promoter recognition and contributes to DNA melting around the initiation site. Most bacterial genomes contain genes for several σ-factors recognizing distinct promoters. Bacterial σ-factors possess conserved functional regions and are grouped into two families, σ^{70} and σ^{54} (Wösten 1998; Ishihama 2000). The σ^{70}-factors are furthermore categorized into primary (group 1, essential for cell growth), non-essential primary (group 2), and alternative σ-factors (group 3), responsible for recognition of certain promoters in response to environmental signals (Lonetto et al. 1992; Gruber and Bryant 1997). Cyanobacteria, the ancestors of plastids, have also multiple σ-factors with distinct promoter specificity (Kaneko et al. 1996).

Early on, biochemically purified σ-like activities in plant plastids were reported in *Chlamydomonas* (Surzycki and Shellenbarger 1976), spinach (Lerbs et al. 1983), and mustard (Bülow and Link 1988; Tiller and Link 1993b). Furthermore, immunological evidence for σ-like factors was obtained in chloroplast RNA polymerase preparations of maize, rice, *Chlamydomonas reinhardtii*, and *Cyanidium caldarium* (Troxler et al. 1994). Moreover, multiple nuclear-encoded genes encoding bacterial σ^{70}-type factors were identified in the red algae *Cyanidium caldarium* (*CcaA-C*; Liu and Troxler 1996; Tanaka et al. 1996) and *Cyandioschyzon merolae* (*CmeSig1-4*; Matsuzaki et al. 2004) suggesting specialized promoter recognition as in bacteria. Correspondingly, σ-factor families were identified in genomes of land plants such as *Arabidopsis* (*AthSig1-6*; Isono et al. 1997b; Tanaka et al. 1997; Fujiwara et al. 2000; Hakimi et al. 2000), mustard (*SalSig1-3*; Kestermann et al. 1998; Homann and Link 2003), tobacco (*NtaSigA1*, *-A2*; Oikawa et al. 2000), rice (*OsaSig1-4*; Tozawa et al. 1998; Kasai et al. 2004), maize (*ZmaSig1-5*; Lahiri et al. 1999; Tan and Troxler 1999; Lahiri and Allison 2000), *Physcomitrella patens* (*PpaSig1, -2, -5*; Hara et al. 2001a, 2001b; Ichikawa et al. 2004), as well as wheat (*TaeSigA*; Morikawa et al. 1999), and *Sorghum* (*SbiSig1*; Kroll et al. 1999). Interestingly, the genome of the unicellular green algae *Chlamydomonas reinhardtii* harbors only a single gene encoding a σ-factor (*CreRpoD*; Carter et al. 2004; Bohne et al. 2006). The N-termini of these σ-factors show sequences typical for plastid-targeting transit peptides and indeed have been demonstrated to confer plastidial targeting either of GFP-fusion proteins *in vivo* (Isono et al. 1997b; Tanaka et al. 1997; Kanamaru et al. 1999; Fujiwara et al. 2000; Lahiri and Allison 2000; Oikawa et al. 2000; Hara et al. 2001a) or with radio-labeled proteins *in vitro* (Kestermann et al. 1998). Surprisingly, targeting of some plant σ-factors occurred not only into plastids but also into mitochondria. Alternative splicing of *AthSig5* transcripts within intron 1 establishes two initiation methionines (M1 and M2). Shorter peptides starting with M2 showed exclusive GFP targeting into plastids. However, GFP fusion proteins starting with M1 were localized to mitochondria.

RNA analyses revealed that the longer (plastidial) *AthSig5* transcripts are exclusively located in flowers, whereas the shorter (mitochondrial) transcripts were detectable in both flower and leaf tissue (Yao et al. 2003). Furthermore, AthSig1::GFP fusion proteins as well are co-localized to both plastids and mitochondria in tobacco protoplast import assays (H. Tandara and K. Liere, unpublished data). Similarly, dual targeting was shown for the maize ZmaSig2B protein by immunological and GFP-fusion protein import studies. Interestingly, ZmaSig2B was biochemically co-purified with RpoTm, the mitochondrial phage-type RNA polymerase (Beardslee et al. 2002), suggesting a possible role of these mitochondrial localized σ-factors in regulation of plant mitochondrial transcription.

Historically, plastidial σ-factors were designated either alphabetically or by numbers. Thus, in *Arabidopsis* SigA, SigB, and SigC (Tanaka et al. 1997) were also named SIG2, SIG1, and SIG3 (Isono et al. 1997b), respectively. In an effort to unify the nomenclature, σ-factors sequences were subjected to phylogenetic analyses and distinguished by numbers (http://sfns.u-shizuoka-ken.ac.jp/pctech; Shiina et al. 2005). Higher plant σ-factors belong into a monophyletic group (Lysenko 2006). They are related to bacterial primary (group 1) and non-essential primary (group 2) σ^{70}-factors. However, none fit into alternative group 3 nor are related to σ^{54}-factors. Phylogenetic analyses revealed that plastidial σ-factors are split into at least 5 subgroups: Sig1, Sig2, Sig3, Sig5, and Sig6. Interestingly, the monocot and dicot σ-factors within the Sig1 and Sig2 groups are located on separate branches. Most sequenced higher plant and moss genomes contain at least one gene for a Sig1-type σ-factor. Since the *Arabidopsis Sig1* homologues are highly expressed during chloroplast biogenesis, it is assumed that Sig1 represents the principal σ-factor in chloroplasts (Tanaka et al. 1997; Kestermann et al. 1998; Tozawa et al. 1998; Kanamaru et al. 1999; Morikawa et al. 1999). Similarly, *Sig2*, *Sig3*, and *Sig5* genes have been identified in various plant organisms, suggesting a correspondingly important role in plastidial transcription. Conversely, to date *AthSig4* is the only *Arabidopsis Sig* gene without known ortholog in other plants, and in comparison to the other σ-factors its transcription is rather low in light-grown plants (Tsunoyama et al. 2002). Supported by the observation that intron sites of *AthSig1*, *AthSig2*, *AthSig3*, *AthSig4*, and *AthSig6* are almost identical (Fujiwara et al. 2000), phylogenetic analysis suggests that the Sig3, AthSig4, and Sig6 groups are related to Sig2 (Shiina et al. 2005; Lysenko 2006). Although closely related, the Sig1 and Sig2 groups possess different number of introns. These σ-factors, therefore, may originate from gene duplication events of one or more ancestral genes. Albeit only partially, the Sig5 group seems to be phylogenetically related to the bacterial alternative σ-factors (Tsunoyama et al. 2002; Shiina et al. 2005; Lysenko 2006). *AtSig4* is suggested to have originated from partly processed transcript of *AthSig2*, *AthSig3*, or *AthSig6* inserted as cDNA into the genome, since it is the only *Sig* gene in higher plants that has lost an intron (Lysenko 2006).

Bacterial σ^{70}-factors contain three conserved domains involved in binding the core RNA polymerase (domains 2.1 and 3), hydrophobic core formation (2.2), DNA melting (2.3), recognition of the -10 promoter motif (2.4), and recognition of the -35 promoter motif (Section 4.1, 4.2; Wösten 1998; Paget and Helmann 2003).

Since these domains are as well present in all known plastidial σ-factors it is to be expected that they are responsible for transcription from σ^{70}-type promoters in plastids. However, structural analysis seems not to provide answers if the role of the different plastidial σ-factors is to selectively activate promoters and if they possess distinct or overlapping promoter specificities. Based on the phylogenetic analyses one might presume that plastidial σ-factors group into general σ-factors involved in transcription of standard σ^{70}-type promoters and specialized σ-factors responsible for recognition of exceptional promoters in response to developmental and/or environmental cues (Shiina et al. 2005; Lysenko 2006).

4.2.3 Role of σ-factor diversity in transcriptional regulation

To address the question of a specific role of σ-factor diversity in transcriptional regulation (see Table 2 summarizing putative roles of σ-factors), *in vitro* reconstitution and transcription experiments using recombinant σ-factors and the *E. coli* core RNA polymerase were carried out by several groups. These again demonstrated that plant σ-factor genes encode functional plastidial σ-factors (Kestermann et al. 1998; Hakimi et al. 2000; Beardslee et al. 2002; Homann and Link 2003; Privat et al. 2003). While the three mustard σ-factors SalSig1, SalSig2, and SalSig3 recognized the *psbA* promoter, only SalSig1 and SalSig2 recognized the *rbcL* promoter. However, *trnK*, *trnQ*, *rps16*, and *rrn16* (PEP-P1) promoters were rather recognized by SalSig1 and SalSig3, but less efficiently by SalSig2 (Homann and Link 2003). Similar experiments with *Arabidopsis* σ-factors suggested that rather AthSig2 and AthSig3 confer specific recognition of the *rbcL* and *psbA* promoters than AthSig1 (Hakimi et al. 2000; Privat et al. 2003). The observed discrepancies in promoter recognition may be due to the heterologous transcription systems with hindered abilities to identify species- and/or PEP-specific regulatory elements at *cis*- and *trans*-factor level.

Further efforts to specify distinct functionality of plant σ-factors in regulation of plastidial gene expression employed characterization of their expression profiles. Profiling of light-dependent transcription in the red algae *Cyanidioschyzon merolae* and *Cyanidium caldarium* revealed light induced accumulation of the mRNAs of σ-factor genes (*CmeSig1-4*; *CcaSigB,C*; Oikawa et al. 1998; Minoda et al. 2005). Furthermore, *CmeSig2* transcript accumulation was additionally increased by high light, indicating that CmeSig2 might be a high-light responsive σ-factor (Minoda et al. 2005). Consistent with a prominent role of PEP in leaves, most plastidial σ-factor genes of higher plants are expressed in light-dependent manner in green tissue but are silent in non-photosynthetic roots (Isono et al. 1997b; Tanaka et al. 1997; Fujiwara et al. 2000; Oikawa et al. 2000). Moreover, expression of plastidial σ-factors seems to be differentially regulated during early *Arabidopsis* development. *AthSig2, AthSig3, AthSig4,* and *AthSig6* but not *AthSig1* and *AthSig5* transcripts accumulate in four day old seedlings (Ishizaki et al. 2005), while in eight day old seedlings transcript levels increase for all σ-factors (Nagashima et al. 2004a). Additionally, expression of *Sig2* transcripts prior to *Sig1*

Table 2. Roles of σ-factors in higher plants.

σ-factor	Target	Function	Plant	Gene	Reference
Sig1	general, *rbcL*	alternative σ-factor, may possibly need activating factor(s)	*Arabidopsis*	At1g64860	(Hakimi et al. 2000) (Privat et al. 2003)
			mustard		(Kestermann et al. 1998) (Homann and Link 2003)
Sig2	general, *psbA*, *trnV/E*, sole σ-factor to control *psaJ*	primary / alternative σ-factor (?)	*Arabidopsis*	At1g08540	(Hakimi et al. 2000) (Hanaoka et al. 2003) (Kanamaru et al. 2001) (Privat et al. 2003) (Shirano et al. 2000) (Tsunoyama et al. 2002)
			mustard		(Homann and Link 2003)
Sig3	general, sole σ-factor to control *psbN*	light-independent early primary / alternative σ-factor (?)	*Arabidopsis*	At3g53920	(Privat et al. 2003) (Zghidi et al. 2006)
			mustard		(Homann and Link 2003)
Sig4	sole σ-factor to control *ndhF*	σ-factor in plant stress response (?)	*Arabidopsis*	At5g13730	(Favory et al. 2005)
Sig5	*psbA*, sole σ-factor to control *psbD*	σ-factor in plant stress response (?) and regulating *psbD* BLRP *via* blue- /UVA-light	*Arabidopsis*	At5g24120	(Nagashima et al. 2004a) (Tsunoyama et al. 2002, 2004)
Sig6	general, *psbA*, *rbcL*, *atpB*, *trnV/E*, *ndhC*	light-independent early primary / alternative σ-factor (?)	*Arabidopsis*	At2g36990	(Ishizaki et al. 2005) (Loschelder et al. 2006)
			maize		(Lahiri and Allison 2000)

Question marks signify a proposed yet unproved function.

in developing leaves was reported for both *Arabidopsis* and rice, suggesting an early function of Sig2 in seedling development (Kanamaru et al. 1999; Kasai et al. 2004). This was supported by recent findings by Demarsy et al. (2006) showing that the mRNAs of AthSig2 and AthSig5 are already present in dry *Arabidopsis* seeds. Interestingly, unlike AthSig1 and AthSig2, AthSig3 protein accumulates in seeds and during early germination (Homann and Link 2003; Privat et al. 2003) as was shown for SolSig2 in spinach (Demarsy et al. 2006). A similar expression pattern was observed for the mustard SalSig3 factor, which accumulates rather in the dark than in light grown seedlings (Homann and Link 2003). Hence, Sig3 may play a distinctive role in regulation of gene expression in etio- and/or proplastids, and might be regulated by posttranslational processes (Homann and Link 2003; Privat et al. 2003). Similarly, ZmaSig6 was detected in root, leaf base, and etiolated leaf tissue in maize (Lahiri and Allison 2000). Therefore, it might be possible that Sig3 and Sig6 represent light-independent, early σ-factors regulating plastid gene expression during seedling growth and development. In opposite, AthSig5 transcripts are expressed later in plant development, controlled *via* the plastidial redox state (Fey et al. 2005). Furthermore, *AthSig5* is rapidly induced by blue, but not red light, which coincides with the blue-light-activated expression of *psbD* (Tsunoyama et al. 2002, 2004). *AthSig5* expression is also activated by various stress cues (Nagashima et al. 2004b). Expression of some plastid genes in higher plants seems to be regulated by circadian rhythms (Nakahira et al. 1998). Circadian timing of plastid gene expression is expected to be mediated by nuclear factors. σ-factors are good candidates to represent such factors. Indeed, *TaeSig1*, *NtaSig1*, *AthSig1*, *AthSig2*, and *PpaSig5* transcripts were shown to exhibit circadian or diurnal expression patterns (Kanamaru et al. 1999; Morikawa et al. 1999; Oikawa et al. 2000; Ichikawa et al. 2004).

Increasingly, functions of σ-factor genes in plants are investigated by analyses of knockout mutants, overexpression, or anti-sense lines. If plastidial gene expression would be controlled by a principal σ-factor similar to the situation in most bacteria, one would assume that inactivation of this gene would result in a drastic, most likely albino phenotype by causing defects in PEP-dependent transcription of photosynthesis related genes. However, examination of various *Arabidopsis* mutants of *AthSig2*, *AthSig3*, *AthSig4*, *AthSig5*, and *AthSig6* did not reveal such a severe phenotype (Shirano et al. 2000; Kanamaru et al. 2001; Hanaoka et al. 2003; Privat et al. 2003; Nagashima et al. 2004b; Tsunoyama et al. 2004; Favory et al. 2005; Ishizaki et al. 2005; Loschelder et al. 2006; Zghidi et al. 2006). Yet, a major break-through in revealing the specificity of σ-factors in transcription came by characterization of these plants.

***AthSig2* knockout mutants.** *AthSig2* mutants displayed a pale green phenotype accompanied by reduced accumulation of some plastid-encoded photosynthesis genes (Shirano et al. 2000; Kanamaru et al. 2001; Privat et al. 2003; Nagashima et al. 2004a). Furthermore, several PEP-transcribed tRNAs including *trnD*-GUC, *trnE*-UUC, *trnM*-CAU, and *trnV*-UAC were prominently reduced in *AthSig2* knockout mutants (Kanamaru et al. 2001; Hanaoka et al. 2003) and anti-sense plants (Privat et al. 2003). *Vice versa*, overexpression of AthSig2 enhanced transcription of *trnE-trnD* (Tsunoyama et al. 2004). It has been suggested that reduction of the photosynthesis-related components is caused by defects in chlorophyll biosynthesis and plastid translation due to the decrease of *trnE*, an initiator of ALA and consequently chlorophyll synthesis. Hence, AthSig2 may have a primary role in driving transcription of certain plastid tRNAs. It cannot be excluded, however, that Sig2 is able to recognize other PEP promoters as suggested for *psbA, psbD,* and *rbcL* (Kanamaru et al. 2001; Hanaoka et al. 2003; Tsunoyama et al. 2004).

***AthSig3* knockout mutants.** In opposite, characterization of *AthSig3* knockout mutants revealed a distinct reduction of transcript levels of the plastid *psbN* gene (Zghidi et al. 2006). Further analyses of transcript initiation sites in these mutants not only showed a loss of transcription initiation from AthP*psbN*-32 but also from AthP*atpH*-413, one of the two PEP promoters upstream of *atpH* in *Arabidopsis*. Therefore, it seems likely that AthSig3 directly controls *psbN* and partially *atpH* gene expression. The function of PsbN is still unknown and its suggested presence in photosystem II has been challenged (Kashino et al. 2002).

***AthSig4* knockout mutants.** Similarly, characterization of an *AthSig4* knockout mutant revealed a specific reduction in transcription of the plastid *ndhF* gene resulting in a strong downregulation of the plastid NDH activity (Favory et al. 2005). Therefore, *ndhF* expression and thus NDH activity seems to be regulated at transcriptional level, controlled by specific σ-factor AthSig4. Interestingly, NDH is involved in plant stress response (Casano et al. 2001) and leaf senescence (Zapata et al. 2005). Whether AthSig4 expression is modulated by such environmental or developmental parameters remains to be investigated.

***AthSig5* knockout mutants.** Apart from AthSig3 and AthSig4, AthSig5 might be an additional σ-factor tied to a specific function in regulation of plastid gene expression. As shown by analyses of transcription in light-treated plants (Tsunoyama et al. 2002; Nagashima et al. 2004b), *AthSig5* knockout plants, and overexpression studies (Nagashima et al. 2004b; Tsunoyama et al. 2004), AthSig5 is regulated by blue light and specifically activates transcription from the *psbD* blue-light responsive promoter (BLRP). Interestingly, analysis of a further *AthSig5* knockout mutant showed embryo lethality (Yao et al. 2003). *AthSig5* has recently been identified as one of 250 genes required for normal embryo development in *Arabidopsis* (Tzafrir et al. 2004) and its mRNA is present in seeds (Demarsy et al. 2006) indicating a substantial role of AthSig5 in seed development. However, it is not yet understood why the different *AthSig5* mutants exhibit these diverse phenotypes.

***AthSig6* knockout mutants.** Cotyledons of *AthSig6* knockout mutants displayed a transient pale green phenotype during early plant development combined with a delay in light-dependent chloroplast development (Ishizaki et al. 2005; Loschelder et al. 2006). During this developmental stage the transcript pattern was found to be similar to that of Δ*rpo* mutants, since transcript levels of most PEP-dependent genes for photosynthesis components, rRNAs, and some tRNAs were decreased. Since the maize homologue ZmSig6 is expressed exclusively in tissue containing immature plastids (Lahiri and Allison 2000), it was proposed that (Ath)Sig6 might be a general σ-factor serving PEP in an early, initial developmental stage. Nonetheless, given that after eight days the mutant phenotype is restored to wild type it is plausible that other σ-factor(s) are able to take over AthSig6 function later in seedling development and plant growth (Shiina et al. 2005). However, characterization of a second *Arabidopsis* knockout line with a *Sig6* mutant allele throughout leaf development (*sig6-2*) suggested a second (persistent or long-term) role of AthSig6 (Loschelder et al. 2006). While transcript accumulation of genes such as *psbA* and *rbcL* was only affected early in development, RNA levels of *atpB* and *ndhC* originating from their corresponding PEP promoters declined during plant development. Interestingly, emerging transcripts which originated further upstream of *atpB* suggested a SOS promoter switch (Schweer et al. 2006).

***AthSig1* overexpressing mutants.** Knockout or anti-sense mutants of *AthSig1* have yet to be characterized. Thus far, data on the role of AthSig1 in plastidial gene expression have been derived from mutant plants overexpressing the *AthSig1* gene (Tsunoyama et al. 2001). Investigation of transcription activity by run-on analyses revealed enhanced initiation from *psaA*, *psbB*, *psbE*, and *rbcL* promoters indicating a more general role of this σ-factor in transcription of genes encoding components of the photosynthesis complexes.

Taken together, only five genes in *Arabidopsis* seem to be controlled by a distinct σ-factor with specific function: *psaJ* by AthSig2, *psbN* by AthSig3, *ndhF* by AthSig4, and *psbD* (BLRP) by AthSig5 (Table 2). However, some other genes appear to be controlled by several σ-factors thereby possessing overlapping functions. Most prominent are genes such as *psbA*, controlled by AthSig2, AthSig5, and AthSig6; *rbcL* controlled by AthSig1 and AthSig6; *trnV*-UAC and *trnE*-UUC, controlled by AthSig2 and AthSig6. Consequently, overlapping functions of σ-factors are generally believed to be the reason for the weak phenotype of σ-factor knockout mutants.

Regulation of σ-factors. PEP activity depends on the developmental stage of the plastids: it is down regulated in etioplasts and is more active in chloroplasts (Rapp et al. 1992; DuBell and Mullet 1995). Furthermore, rates of PEP transcription are higher in the light than in the dark (Shiina et al. 1998). Changes in PEP transcription activity have been suggested to be partly resulting from changes in the phosphorylation state of σ-factors. Phosphorylation of σ-factors and the PEP enzyme itself have been shown to be an important regulatory event in chloroplast transcription (Tiller and Link 1993a; Baginsky et al. 1997; Christopher et al. 1997). In mustard, a CK2-type kinase has been identified to be part of the chloroplast PEP-A complex (Ogrzewalla et al. 2002). This plastid transcription kinase activity (PTK), termed cpCK2, is able to phosphorylate purified sigma-like factors (SLFs) as well as subunits of the PEP-A complex *in vitro*. Based on the observation that cpCK2 itself is differentially regulated by phosphorylation and redox state, cpCK2 was proposed to be part of a signaling pathway controlling PEP activity (Baginsky et al. 1999). Phosphorylation and SH-group redox state were shown to work antagonistically. A non-phosphorylated cpCK2 appears to be more active, but is inhibited by treatment with reduced glutathione (GSH). *Vice versa*, a phosphorylated non-active enzyme could be re-activated by adding GSH. In opposite to cpCK2 isolated from plants grown under high light conditions, cpCK2 isolated from plants grown under moderate light conditions effectively phosphorylated the associated PEP-A, therefore corroborating these findings (Baena-Gonzalez et al. 2001). Thus, light dependent reduction of GSH would inactivate cpCK2, while dephosphorylation of PEP under high light conditions would enhance PEP-dependent transcription. It remains unknown whether cpCK2 is also regulated *via* extraplastidic signal chains mediated by phyto- and/or chryptochromes. Since cpCK2 orthologs have been identified in various plant species (Loschelder et al. 2004) it might well be that this kinase has an evolutionary conserved role in plastid redox-sensitive signal transduction.

In bacteria, σ-factor activity is controlled by anti-σ factors (Ishihama 2000). Plastid σ-factor AthSig1 associated proteins with plastid localization were identified in *Arabidopsis* (SibI and T3K9.5; Morikawa et al. 2002). They are not related to any proteins of known function and are light-dependent, developmental, and tissue-specifically expressed, and thus may be involved in regulation of AthSig1 activity.

4.3 Exogenous and endogenous factors controlling plastidial transcription

Plant development is highly influenced by environmental factors. Plastid gene expression was shown to differentially respond to environmental cues (Chory et al. 1995; Link 1996; Goldschmidt-Clermont 1998; Barkan and Goldschmidt-Clermont 2000). Therefore, *cis-* and *trans*-elements regulating differential gene expression in plastids were in the center of attention in the last decades (see Table 3 for summary). Regulatory sequence motifs upstream the -35 core promoter region were found in the promoters of *rbcL* and *psbD-psbC*. The *rbcL* gene is transcribed from a single PEP promoter with well conserved -35 and -10-elements and canonical spacing by 18 nucleotides (Shinozaki and Sugiura 1982; Mullet et al. 1985; Reinbothe et al. 1993; Isono et al. 1997a). *In vitro* studies demonstrated the importance of both the -35/-10 box spacing and sequence for *rbcL* promoter strength (Gruissem and Zurawski 1985; Hanley-Bowdoin et al. 1985). An upstream element, conserved between maize, pea, spinach, and tobacco was proposed to function as a binding site for the chloroplast DNA-binding factor 1 (CDF1) in maize (Lam et al. 1988). Interestingly, a segment of CDF1, region II, is reminiscent of the AT-rich UP element stimulating transcription in *E. coli* (Ross et al. 1993). However, analyses of transplastomic plants expressing chimeric P*rbcL::uidA* constructs demonstrated, that the *rbcL* core promoter is sufficient to obtain wild type rates of transcription (Shiina et al. 1998). Interestingly, another DNA-binding protein (RLBP, *rbcL* promoter-binding protein) binds specifically to the *rbcL* promoter core in tobacco (Fig. 4; -3 to -32; Kim et al. 2002). Only detectable in light-grown seedlings, RLBP is suggested to play a role in light-dependent *rbcL* transcription. However, stabilization of the *rbcL* mRNA *via* its 5' UTR is compensating for reduced rates of transcription in the dark and leads to light-dependent transcript accumulation (Shiina et al. 1998).

Table 3. Transcription regulating factors in higher plants.

Protein	Target	Function	Plant	Gene	Reference
PEP-regulating factors					
AGF	*psbD* BLRP	binds AAG-box; transcription enhancer	barley wheat		(Kim and Mullet 1995) (Nakahira et al. 1998)
PTF1	*psbD* BLRP	binds AAG-box; transcription enhancer part of AGF	*Arabidopsis*	At3g02150	(Baba et al. 2001)
PGTF	*psbD* BLRP	binds to PGT-box	barley		(Kim and Mullet 1995)
CDF1	*rbcL*	DNA-bdg.; transcription regulation	pea maize		(Lam et al. 1988)
CDF2-A	*rrn16*(P1)	DNA-bdg.; transcription repression	spinach		(Baeza et al. 1991) (Bligny et al. 2000)
Region U-bdg. protein	*psaA*	DNA-bdg.; transcription regulation	spinach		(Cheng et al. 1997a)
Region D-bdg. protein	*psaA*	DNA-bdg.; transcription regulation	spinach		(Cheng et al. 1997a)
RLBP	*rbcL*	DNA-bdg.; transcription regulation	tobacco		(Kim et al. 2002)
SibI		AthSig1-bdg. protein	*Arabidopsis*	At3g56710	(Morikawa et al. 2002)
T3K9.5		AthSig1-bdg. protein	*Arabidopsis*	At2g41180	(Morikawa et al. 2002)
cpCK2 (PTK)		plastid transcription kinase, phosphorylates sigma-like factors (SLFs) and subunits of PEP-A	*Arabidopsis* mustard	At5g67380	(Baginsky and Gruissem 2002) (Baginsky et al. 1997) (Ogrzewalla et al. 2002)
NEP-regulating factors					
RPL4		interaction with CDF2 (?); transcription regulation	spinach	X93160 (gi2792019)	(Trifa et al. 1998)
CDF-2B	*rrn16*	interaction with a NEP-2 transcription activity (?) DNA-bdg.; transcription regulation	spinach		(Bligny et al. 2000)
tRNA^{Glu}		inactivates NEP activity by binding to RpoTp	*Arabidopsis*	*trnE* (ArthCt097)	(Hanaoka et al. 2005)

Question marks signifies a proposed yet unproved function.

Blue-light control of the *psbD-psbC* operon. Contrary to most photosynthetic genes, the rate of transcription of *psbD-psbC* remains high in mature chloroplasts (Klein and Mullet 1990; Baumgartner et al. 1993; DuBell and Mullet 1995). Responsible is a specific activation of one of the *psbD* promoters, the blue light-responsive promoter (BLRP; Sexton et al. 1990), which is found upstream of the *psbD* gene of various species (Fig. 4; Christopher et al. 1992; Wada et al. 1994; Allison and Maliga 1995; Kim and Mullet 1995; To et al. 1996; Hoffer and Christopher 1997; Kim et al. 1999b; Thum et al. 2001). The architecture of the *psbD* BLRP promoter consists of two conserved upstream elements (PGT-box, AAG-box) and poorly conserved and closely spaced -35/-10-elements. *In vivo* studies in transplastomic tobacco revealed that deletion of parts of the PGT-box reduced mRNA levels, while subsequent deletion the AAG-box sequences even further reduced transcript levels (Allison and Maliga 1995). In tobacco, therefore, the conserved sequence elements upstream of the *psbD* promoter core are accountable for light-activated transcript accumulation. *In vitro* transcription from the *psbD* promoter in rice, wheat, and barley depends on the -10, but not on the -35 promoter element (To et al. 1996; Satoh et al. 1997; Nakahira et al. 1998; Kim et al. 1999b). The AAG-box of the barley promoter was shown to be the binding site for a nuclear-encoded AAG-binding complex *in vitro* (AGF; Kim and Mullet 1995). However, binding activity of AGF to the AAG-box is not correlated with transcriptional activation of the *psbD* BLRP (Nakahira et al. 1998). One of the components of AGF of *Arabidopsis* was cloned and designated plastid transcription factor 1 (PTF1; Baba et al. 2001). Studies on *PTF1*-deficient mutants revealed that PTF1 is rather involved in general transcriptional enhancement than in light-dependent activation of *psbD* transcription. Correspondingly, the PGT-box is the binding site for PGTF, the PGT-binding factor. Its DNA-binding activity is regulated by an ADP-dependent kinase (Kim et al. 1999a). A model based on these *in vitro* experiments in barley explains that constitutively binding of AGF to the upstream AAG-element may assist promoter recognition by PEP, whereas light-dependent transcriptional activation of *psbD* transcription is mediated by binding of PGTF to the PGT-box. In the dark, PGTF is phosphorylated and loses its affinity for the PGT element, thereby decreasing transcription. Although the *psbD* promoter architecture is highly conserved, it is unlikely that PGT is required for light-dependent transcription in various other plants. It was shown for rice (To et al. 1996), wheat (Satoh et al. 1997), and barley *in vitro* (Kim et al. 1999b) and in transplastomic tobacco *in vivo* (Thum et al. 2001) that the PGF-box is not required for light-dependent activation in these plants. Therefore, the roles of PGT and PGTF remain largely unknown.

It has been proposed that AthSig5 might act as a mediator of blue-light signaling in activating *psbD* BLRP transcription in blue light (see Section 4.2.3; Tsunoyama et al. 2002; Nagashima et al. 2004b; Tsunoyama et al. 2004), whereas AGF enhances *psbD* BLRP transcription by constitutively binding to the AAG-box (Shiina et al. 2005). It is assumed that the signal transduction pathway involves reception of blue light by cryptochromes and PhyA (Thum et al. 2001; Mochizuki et al. 2004), further mediation by a protein phosphatase PP7 (Moller et al. 2003), and subsequent induction of *Sig5* expression (Mochizuki et al. 2004).

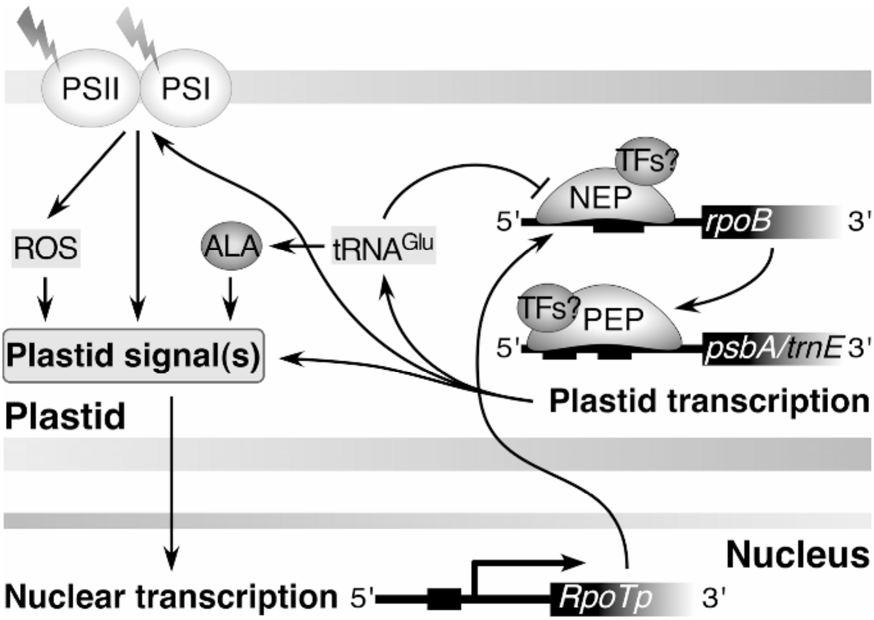


Fig. 7. The role of nuclear-encoded phage-type RNA polymerases in regulation of plastidial gene expression. NEP transcription activity is in part represented by a phage-type RNA polymerase encoded by the nuclear located *RpoTp* gene. NEP transcribes and therefore may regulate expression of the plastidial *rpoB* operon encoding subunits of the plastid-encoded RNA polymerase (PEP). PEP in turn transcribes genes encoding components of the photosynthetic complexes (PSI, PSII) that regulate nuclear transcription by generating diverse 'plastid signals' (ROS, reactive oxygen species). The *trnE* gene encoding trnA^{Glu} which is required for the synthesis of δ-aminolevulinic acid (ALA) is also transcribed by PEP (Hess et al. 1992; Walter et al. 1995). ALA is a precursor of the chlorophyll and heme biosynthesis thought to provide 'plastid signals' which influence nuclear transcription. Furthermore, tRNA^{Glu} is assumed to developmentally inhibit NEP transcription by binding to RpoTp (Hanaoka et al. 2005). In turn, the expression and activity of nuclear-encoded, plastid phage-type RNA polymerase regulates the transcription of plastidial genes and consequently the developmental stage of the plastid (RpoTp; Emanuel et al. 2004). Thus, the regulated network of the nuclear and plastidial transcription machineries may be a key element for a concerted expression of genes located within compartments of the plant cell.

After import into plastids, Sig5 associates with AGF (PTF1) and initiates *psbD* transcription. Furthermore, *psbD* BLRP activity is also regulated in a developmental and tissue-specific manner, since the *Arabidopsis* DET1 gene product down-regulates the activity of *psbD* BLRP in young seedlings (Christopher and Hoffer 1998).

Plastid-to-nucleus signaling. Environmental control of plastidial gene expression is most intense in differentiation from proplastids to either etioplasts (dark) or chloroplasts (light). Analyses of photomorphogenic mutants established the existence of different pathways to communicate light perception to plastids in order to control their development (Leon et al. 1998; Rodermel 2001; Gray et al. 2003; López-Juez and Pyke 2005). However, these analyses also showed that retrograde or 'plastid signals' are controlling nuclear gene expression depending on the developmental status of the plastid (Fig. 7; see Chapter 13; Rodermel 2001; Gray 2003; Beck 2005; Leister 2005; Nott et al. 2006). Both plastid transcription and translation are necessary for the production of a 'plastid signal'. However, it is not an immediate translational product of a plastid gene (Oelmüller et al. 1986; Lukens et al. 1987), but rather part(s) of signal transduction pathways in plastids.

The barley mutant *albostrians*, with alternating stripes of white and green tissue, contains no detectable ribosomes in plastids of white tissue cells (Siemenroth et al. 1981; Hess et al. 1993). Transcript levels of some photosynthesis-related plastidial and nuclear genes are reduced or missing suggesting the existence of 'plastid signals' controlling nuclear gene expression (Bradbeer et al. 1979; Hess et al. 1994). Recently, transcript levels of the nuclear-encoded *RpoTp*, which is likely to represent NEP activity, and its plastidial target genes were analyzed throughout the developmental gradient of *albostrians* leaves (Emanuel et al. 2004). The results revealed a significant influence of the developmental stage of plastids on expression and activity of RpoTp, indicating a plastid-to-nucleus signaling to coordinate expression of plastidial and nuclear-encoded RNA polymerases as a prerequisite of a concerted gene expression in both plastids and nucleus (Fig. 7).

Redox control of plastid gene expression. Light is not only the energy source for photosynthesis, but also an environmental signal to regulate plant biogenesis and environmental adaptation. Apart from blue/UVA-light, illumination has been early hypothesized to control plastid gene expression *via* the physiological status of the plastid, e.g., redox conditions (Link 2003; Pfannschmidt and Liere 2005). Redox control of plastidial gene expression has been interpreted as a selection force throughout evolution to retaining their genomes (Allen 1993). First confirmation for such a redox control was obtained by demonstrating that light supported incorporation of radioactive-labeled NADH into the RNA fraction of lettuce plastids (Pearson et al. 1993). Plastidial gene expression is controlled at different levels by photosynthetic activity such as RNA maturation (Deshpande et al. 1997; Liere and Link 1997; Salvador and Klein 1999) and translation (Danon and Mayfield 1994; Bruick and Mayfield 1999; Trebitsh et al. 2000; Zhang et al. 2000). Effects of the redox state on plastidial gene transcription were furthermore demonstrated by growing plants under light conditions generating an imbalance in excitation energy distribution between photosystems (PSII- and PSI-light, 680 and 700 nm, respectively; Pfannschmidt et al. 1999a, 1999b; Fey et al. 2005). Preferential excitation of PSII results in a reduction of the electron transport chain while a preferential excitation of PSI results in its oxidation. The change in photosystem stoichiometry correlated with respective changes in the transcriptional rates and transcript amounts of the plastidial genes for the reaction centre proteins of PSII and PSI, *psbA* and *psaAB*. Indeed, the redox state of the plastoquinone pool (PQ) is the major determinant for the changes in gene expression. A reduced PQ pool promotes transcription of the *psaAB* operon. In reverse, an oxidized PQ pool increases *psbA* transcription. Opposite regulation of these genes has been recently found also in pea (Tullberg et al. 2000), *Chlamydobotrys stellata* (Kovacs et al. 2000) and *Synechocystis* PCC 6803 (Li and Sherman 2000; El Bissati and Kirilovsky 2001) suggesting that this mechanism represents an evolutionary old means of regulating gene expression. These data provide a first model on how plants adapt to light quality gradients occurring in natural environments under low light intensities. Still, the signal transduction pathway connecting the PQ pool with transcription is yet unknown. However, a long-term response may represent an extended branch of the short-term response (the state transition), which is also regulated by the redox state of the PQ pool (Allen and Forsberg 2001; Pursiheimo et al. 2001). The PQ oxidation site at the cyt b_6f complex functions as a sensor for the PQ redox state during state transition (Vener et al. 1997; Zito et al. 1999). A putative DNA-binding protein of PS II, TSP9, is partially released from PSII upon PQ reduction in spinach and may represent such a signal transducer towards transcription (Carlberg et al. 2003; Zer and Ohad 2003). Identification of an additional protein of 31 kDa capable of sequence-specific binding between positions + 64 to +83 (region D) of the light dependent *psaAB* PEP promoter region (Chen et al. 1993; Cheng et al. 1997a) suggests the existence of yet unidentified transcription factors that transmit redox signals. Furthermore, the *Arabidopsis* high chlorophyll fluorescence mutant *hcf145* shows decreased mRNA stability and transcription of *psaA* (Lezhneva and Meurer 2004). Thus, HCF145 might be involved in transcriptional regulation of the *psaA* operon. Further analysis of this promoter has yet to be reported.

PEP is not only responsible for the redox regulation at the *psbA* and *psaAB* promoters, but apparently is also regulated *via* redox control. A regulatory impact on steady-state levels of transcripts of genes for PEP components was observed by microarray analyses (Fey et al. 2005): *rpoB* (plastid-encoded β-subunit), *AthSig5* (nuclear-encoded σ-factor), and *SibI* (nuclear-encoded Sig1-binding protein; Morikawa et al. 2002). Interestingly, *rpoB* is transcribed by a nuclear-encoded phage-type RNA polymerase (Fig. 7, RpoTp; Liere et al. 2004), suggesting a redox regulation of this enzyme (see Chapter 13).

Developmental switch from NEP to PEP. A regulatory role, which links chlorophyll synthesis and the developmental switch from nucleus-encoded RNA polymerases to the plastid-encoded bacterial-type enzyme, has been proposed for the plastid-encoded $tRNA^{Glu}$ in *Arabidopsis* (Hanaoka et al. 2005). $tRNA^{Glu}$ is not only required for translation, but also for synthesis of δ-aminolevulinic acid, a precursor of chlorophyll (Schön et al. 1986). In gel mobility shift experiments recombinant RpoTp specifically bound this tRNA. Additionally, transcription from a putative plastidial *accD* NEP promoter sequence was inhibited by addition of $tRNA^{Glu}$ to *in vitro* transcription reactions with proplastid extracts from *Arabidopsis*. Hence, the authors suggested $tRNA^{Glu}$ to developmentally inhibit transcription by RpoTp (Fig. 7).

Bacterial-like stringent control. In bacteria, one of the most important processes to regulate gene expression is the so-called ‘stringent control’ enabling adaptation to nutrient-limiting conditions (Cashel et al. 1996). The effector molecule is guanosine 5'-diphosphate 3'-diphosphate (ppGpp), which binds to the core RNA polymerase modifying its promoter specificity (Toulokhonov et al. 2001). Stress-induced synthesis is mediated by ppGpp synthetases, RelA and SpoT, homologues of which were found in *Chlamydomonas reinhardtii* (Kasai et al. 2002), *Arabidopsis* (van der Biezen et al. 2000), and tobacco (Givens et al. 2004). Plastidial targeting has been demonstrated for some of these RSH termed proteins, suggesting an implication in ppGpp signaling in plastids. RSH expression and plastidial ppGpp levels are clearly elevated by light and various abiotic and biotic stress conditions. Furthermore, PEP activity is inhibited by ppGpp *in vitro* (Givens et al. 2004; Takahashi et al. 2004). Thus, it is conceivable that PEP might indeed be under control of a bacterial-like stringent response mediated by ppGpp. Interestingly, stress signals specifically induce transcription initiation from the *psbD* BRLP conferred by a special σ-factor, AthSig5 (see Section 3.2.3; Nagashima et al. 2004b; Tsunoyama et al. 2004). However, target genes that are regulated by a plastidial stringent control have yet to be identified, which might help to elucidate the molecular mechanisms of transcriptional responses to plant hormones and environmental stress situations.

Acknowledgement

Andreas Weihe kindly provided Figure 2. The work of the authors is supported by Deutsche Forschungsgemeinschaft (SFB 429).

References

Allen JF (1993) Control of gene expression by redox potential and the requirement for chloroplast and mitochondrial genomes. J Theor Biol 165:609-631

Allen JF, Forsberg J (2001) Molecular recognition in thylakoid structure and function. Trends Plant Sci 6:317-326

Allison LA, Maliga P (1995) Light-responsive and transcription-enhancing elements regulate the plastid *psbD* core promoter. EMBO J 14:3721-3730

Allison LA, Simon LD, Maliga P (1996) Deletion of *rpoB* reveals a second distinct transcription system in plastids of higher plants. EMBO J 15:2802-2809

Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS, Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD, Jurka J, Kapitonov VV, Kroger N, Lau WWY, Lane TW, Larimer FW, Lippmeier JC, Lucas S, Medina M, Montsant A, Obornik M, Parker MS, Palenik B, Pazour GJ, Richardson PM, Rynearson TA, Saito MA, Schwartz DC, Thamatrakoln K, Valentin K, Vardi A, Wilkerson FP, Rokhsar DS (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. Science 306:79-86

Azevedo J, Courtois F, Lerbs-Mache S (2006) Sub-plastidial localization of two different phage-type RNA polymerases in spinach chloroplasts. Nucleic Acids Res 34:436-444

Baba K, Nakano T, Yamagishi K, Yoshida S (2001) Involvement of a nuclear-encoded basic helix-loop-helix protein in transcription of the light-responsive promoter of *psbD*. Plant Physiol 125:595-603

Baba K, Schmidt J, Espinosa-Ruiz A, Villarejo A, Shiina T, Gardestrom P, Sane AP, Bhalerao RP (2004) Organellar gene transcription and early seedling development are affected in the *RpoT;2* mutant of *Arabidopsis*. Plant J 38:38-48

Baena-Gonzalez E, Baginsky S, Mulo P, Summer H, Aro E-M, Link G (2001) Chloroplast transcription at different light intensities. Glutathione-mediated phosphorylation of the major RNA polymerase involved in redox-regulated organellar gene expression. Plant Physiol 127:1044-1052

Baeza L, Bertrand A, Mache R, Lerbs-Mache S (1991) Characterization of a protein binding sequence in the promoter region of the 16S rRNA gene of the spinach chloroplast genome. Nucleic Acids Res 19:3577-3581

Baginsky S, Gruissem W (2002) Endonucleolytic activation directs dark-induced chloroplast mRNA degradation. Nucleic Acids Res 30:4527-4533

Baginsky S, Tiller K, Link G (1997) Transcription factor phosphorylation by a protein kinase associated with chloroplast RNA polymerase from mustard (*Sinapis alba*). Plant Mol Biol 34:181-189

Baginsky S, Tiller K, Pfannschmidt T, Link G (1999) PTK, the chloroplast RNA polymerase-associated protein kinase from mustard (*Sinapis alba*), mediates redox control of plastid *in vitro* transcription. Plant Mol Biol 39:1013-1023

Barkan A, Goldschmidt-Clermont M (2000) Participation of nuclear genes in chloroplast gene expression. Biochimie 82:559-572

Baumgartner BJ, Rapp JC, Mullet JE (1993) Plastid genes encoding the transcription/translation apparatus are differentially transcribed early in barley (*Hordeum vulgare*) chloroplast development: evidence for selective stabilization of *psbA* mRNA. Plant Physiol 101:781-791

Beardslee TA, Roy-Chowdhury S, Jaiswal P, Buhot L, Lerbs-Mache S, Stern DB, Allison LA (2002) A nucleus-encoded maize protein with sigma factor activity accumulates in mitochondria and chloroplasts. Plant J 31:199-209

Beck CF (2005) Signaling pathways from the chloroplast to the nucleus. Planta 222:743-756

Bedbrook JR, Link G, Coen DM, Bogorad L, Rich A (1978) Maize plastid gene expressed during photoregulated development. Proc Natl Acad Sci USA 75:3060-3064

Bendich AJ (1987) Why do chloroplasts and mitochondria contain so many copies of their genome? Bioessays 6:279-282

Berends Sexton T, Jones JT, Mullet JE (1990) Sequence and transcriptional analysis of the barley ctDNA region upstream of *psbD-psbC* encoding *trnK* (UUU), *rps16*, *trnQ* (UUG), *psbK*, *psbI*, and *trnS* (GCU). Curr Genet 17:445-454

Berg S, Krause K, Krupinska K (2004) The *rbcL* genes of two *Cuscuta* species, *C. gronovii* and *C. subinclusa*, are transcribed by the nuclear-encoded plastid RNA polymerase (NEP). Planta 219:541-546

Bisanz-Seyer C, Li Y-F, Seyer P, Mache R (1989) The components of the plastid ribosome are not accumulated synchronously during the early development of spinach plants. Plant Mol Biol 12:201-211

Bligny M, Courtois F, Thaminy S, Chang CC, Lagrange T, Baruah-Wolff J, Stern D, Lerbs-Mache S (2000) Regulation of plastid rDNA transcription by interaction of CDF2 with two different RNA polymerases. EMBO J 19:1851-1860

Blowers AD, Ellmore GS, Klein U, Bogorad L (1990) Transcriptional analysis of endogenous and foreign genes in chloroplast transformants of *Chlamydomonas*. Plant Cell 2:1059-1070

Bohne A-V, Irihimovitch V, Weihe A, Stern D (2006) *Chlamydomonas reinhardii* encodes a single sigma^{70}-like factor which likely functions in chloroplast transcription. Curr Genet 49:333-340

Bown J, Barne K, Minchin S, Busby S (1997) Extended -10 promoters. Nucleic Acids Mol Biol 11:41-52

Boyer SK, Mullet JE (1986) Characterization of *P. sativum* chloroplast *psbA* transcripts produced *in vivo* and *in vitro* and in *E. coli*. Plant Mol Biol 6:229-243

Boyer SK, Mullet JE (1988) Sequence and transcript map of barley chloroplast *psbA* gene. Nucleic Acids Res 16:8184

Bradbeer JW, Atkinson YE, Börner T, Hagemann R (1979) Cytoplasmic synthesis of plastid polypeptides may be controlled by plastid-synthesized RNA. Nature 279: 816-817

Bradley D, Gatenby AA (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

Briat JF, Laulhere JP, Mache R (1979) Transcription activity of a DNA-protein complex isolated from spinach plastids. Eur J Biochem 98:285-292

Bruick RK, Mayfield SP (1999) Light-activated translation of chloroplast mRNAs. Trends Plant Sci 4:190-195

Bülow S, Link G (1988) Sigma-like activity from mustard (*Sinapis alba* L.) chloroplasts confering DNA-binding and transcription specificity to *E. coli* core RNA polymerase. Plant Mol Biol 10:349-357

Bünger W, Feierabend J (1980) Capacity for RNA synthesis in 70S ribosome-deficient plastids of heat-bleached rye leaves. Planta 149:163-169

Butterfass T (1980) The continuity of plastids and the differentiation of plastid populations. In: Reinert J (ed) Results and problems in cell differentiation. Berlin-Heidelberg-New York: Springer-Verlag, pp 29-44

Cahoon AB, Harris FM, Stern DB (2004) Analysis of developing maize plastids reveals two mRNA stability classes correlating with RNA polymerase type. EMBO Rep 5:801–806

Carlberg I, Hansson M, Kieselbach T, Schroder WP, Andersson B, Vener AV (2003) A novel plant protein undergoing light-induced phosphorylation and release from the photosynthetic thylakoid membranes. Proc Natl Acad Sci USA 100:757-762

Carter ML, Smith AC, Kobayashi H, Purton S, Herrin DL (2004) Structure, circadian regulation and bioinformatic analysis of the unique sigma factor gene in *Chlamydomonas reinhardtii*. Photosynth Res 82:339-349

Casano LM, Martin M, Sabater B (2001) Hydrogen peroxide mediates the induction of chloroplastic NDH complex under photooxidative stress in barley. Plant Physiol 125:1450-1458

Cashel M, Gentry DM, Hernandez VJ, Vinella D (1996) The stringent response. In: Neidhardt FC (ed) *Escherichia coli* and *Salmonella typhimurium:* cellular and molecular biology. Washington D.C.: ASM Press, pp 1458-1496

Chang C-C, Sheen J, Bligny M, Niwa Y, Lerbs-Mache S, Stern DB (1999) Functional analysis of two maize cDNAs encoding T7-like RNA polymerases. Plant Cell 11:911-926

Chen MC, Cheng MC, Chen SC (1993) Characterization of the promoter of rice plastid *psaA-psaB-rps14* operon and the DNA-specific binding proteins. Plant Cell Physiol 34:577-584

Cheng MC, Wu SP, Chen LFO, Chen SCG (1997a) Identification and purification of a spinach chloroplast DNA-binding protein that interacts specifically with the plastid *psaA-psaB-rps14* promoter region. Planta 203:373-380

Cheng YS, Lin CH, Chen LJ (1997b) Transcription and processing of the gene for spinach chloroplast threonine tRNA in a homologous *in vitro* system. Biochem Biophys Res Commun 233:380-385

Chory J, Cook RK, Dixon R, Elich T, Li HM, Lopez E, Mochizuki N, Nagpal P, Pepper A, Poole D, Reed J (1995) Signal-transduction pathways controlling light-regulated development in *Arabidopsis*. Philos Trans R Soc Lond B Biol Sci 350:59-65

Christensen AC, Lyznik A, Mohammed S, Elowsky CG, Elo A, Yule R, Mackenzie SA (2005) Dual-domain, dual-targeting organellar protein presequences in *Arabidopsis* can use non-AUG start codons. Plant Cell 17:2805-2816

Christopher DA, Hoffer PH (1998) DET1 represses a chloroplast blue light-responsive promoter in a developmental and tissue-specific manner in *Arabidopsis thaliana*. Plant J 14:1-11

Christopher DA, Kim M, Mullet JE (1992) A novel light-regulated promoter is conserved in cereal and dicot chloroplasts. Plant Cell 4:785-798

Christopher DA, Li XL, Kim M, Mullet JE (1997) Involvement of protein kinase and extraplastidic Serine/Threonine protein phosphatases in signaling pathways regulating plastid transcription and the *psbD* blue light-responsive promoter in barley. Plant Physiol 113:1273-1282

Chun L, Kawakami A, Christopher DA (2001) Phytochrome A mediates blue light and UV-A-dependent chloroplast gene transcription in green leaves. Plant Physiol 125:1957-1966

Danon A, Mayfield SP (1994) Light-regulated translation of chloroplast messenger RNAs through redox potential. Science 266:1717-1719

Demarsy E, Courtois F, Azevedo J, Buhot L, Lerbs-Mache S (2006) Building up of the plastid transcriptional machinery during germination and early plant development. Plant Physiology 142:993-1003

Deng XW, Gruissem W (1987) Control of plastid gene expression during development: the limited role of transcriptional regulation. Cell 49:379-387

Derelle E, Ferraz C, Rombauts S, Rouze P, Worden AZ, Robbens S, Partensky F, Degroeve S, Echeynie S, Cooke R, Saeys Y, Wuyts J, Jabbari K, Bowler C, Panaud O, Piegu B, Ball SG, Ral J-P, Bouget F-Y, Piganeau G, De Baets B, Picard A, Delseny M, Demaille J, Van de Peer Y, Moreau H (2006) Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. Proc Natl Acad Sci USA 103:11647-11652

Deshpande NN, Bao Y, Herrin DL (1997) Evidence for light/redox-regulated splicing of *psbA* pre-RNAs in *Chlamydomonas* chloroplasts. RNA 3:37-48

Dhingra A, Bies DH, Lehner KR, Folta KM (2006) Green light adjusts the plastid transcriptome during early photomorphogenic development. Plant Physiol 142:1256-1266

DuBell AN, Mullet JE (1995) Differential transcription of pea chloroplast genes during light-induced leaf development. Plant Physiol 109:105-112

Eisermann A, Tiller K, Link G (1990) *In vitro* transcription and DNA binding characteristics of chloroplast and etioplast extracts from mustard (*Sinapis alba*) indicate differential usage of the *psbA* promoter. EMBO J 9:3981-3987

El Bissati K, Kirilovsky D (2001) Regulation of *psbA* and *psaE* expression by light quality in *Synechocystis* species PCC 6803. A redox control mechanism. Plant Physiol 125:1988-2000

Ellis RJ, Hartley MR (1971) Sites of synthesis of chloroplast proteins. Nature 233:193-196

Emanuel C, von Groll U, Müller M, Börner T, Weihe A (2006) Development- and tissue-specific expression of the *RpoT* gene family of *Arabidopsis* encoding mitochondrial and plastid RNA polymerases. Planta 223:998-1009

Emanuel C, Weihe A, Graner A, Hess WR, Börner T (2004) Chloroplast development affects expression of phage-type RNA polymerases in barley leaves. Plant J 38:460-472

Ems SC, Morden CW, Dixon CK, Wolfe KH, dePamphilis CW, Palmer JD (1995) Transcription, splicing and editing of plastid RNAs in the nonphotosynthetic plant *Epifagus virginiana*. Plant Mol Biol 29:721-733

Falk J, Schmidt A, Krupinska K (1993) Characterization of plastid DNA transcription in ribosome deficient plastids of heat-bleached barley leaves. J Plant Physiol 141:176-181

Falkenberg M, Gaspari M, Rantanen A, Trifunovic A, Larsson N-G, Gustafsson CM (2002) Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. Nature Genet 31:289-294

Favory J-J, Kobayshi M, Tanaka K, Peltier G, Kreis M, Valay J-G, Lerbs-Mache S (2005) Specific function of a plastid sigma factor for *ndhF* gene transcription. Nucleic Acids Res 33:5991-5999

Fey V, Wagner R, Brautigam K, Wirtz M, Hell R, Dietzmann A, Leister D, Oelmuller R, Pfannschmidt T (2005) Retrograde plastid redox signals in the expression of nuclear genes for chloroplast proteins of *Arabidopsis thaliana*. J Biol Chem 280:5318-5328

Forsberg J, Rosenquist M, Fraysse L, Allen JF (2001) Redox signalling in chloroplasts and mitochondria: genomic and biochemical evidence for two-component regulatory systems in bioenergetic organelles. Biochem Soc Trans 29:403-407

Fujiwara M, Nagashima A, Kanamaru K, Tanaka K, Takahashi H (2000) Three new nuclear genes, *sigD*, *sigE* and *sigF*, encoding putative plastid RNA polymerase σ factors in *Arabidopsis thaliana*. FEBS Lett 481:47-52

Galli G, Hofstetter H, Birnstil ML (1981) Two conserved blocks within eukaryotic tRNA genes are major promoter elements. Nature 294:626-631

Gatenby AA, Castleton JA, Saul MW (1981) Expression in *E. coli* of maize and wheat chloroplast genes for large subunit of ribulose bisphosphate carboxylase. Nature 291:117-121

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

Geiduschek EP, Bardeleben C, Joazeiro CA, Kassavetis GA, Whitehall S (1995) Yeast RNA polymerase III: transcription complexes and RNA synthesis. Braz J Med Biol Res 28:147-159

Givens RM, Lin MH, Taylor DJ, Mechold U, Berry JO, Hernandez VJ (2004) Inducible expression, enzymatic activity, and origin of higher plant homologues of bacterial RelA/SpoT stress proteins in *Nicotiana tabacum*. J Biol Chem 279:7495-7504

Gockel G, Hachtel W (2000) Complete gene map of the plastid genome of the nonphotosynthetic euglenoid flagellate *Astasia longa*. Protist 151:347-351

Goldschmidt-Clermont M (1998) Coordination of nuclear and chloroplast gene expression in plant cells. Int Rev Cytol 177:115-180

Gray JC (2003) Chloroplast-to-nucleus signalling: a role for Mg-protoporphyrin. Trends Genet 19:526-529

Gray MW, Lang BF (1998) Transcription in chloroplasts and mitochondria: a tale of two polymerases. Trends Microbiol 6:1-3

Greenberg BM, Gaba V, Canaani O, Malkin S, Mattoo AK, Edelman M (1989) Separate photosensitizers mediate degradation of the 32-kDa photosystem II reaction center protein in the visible and UV spectral regions. Proc Natl Acad Sci USA 86:6617-6620

Greenberg BM, Narita JO, DeLuca-Flaherty C, Gruissem W, Rushlow KA, Hallick RB (1984) Evidence for two RNA polymerase activities in *Euglena gracilis* chloroplasts. J Biol Chem 259:14880-14887

Gruber TM, Bryant DA (1997) Molecular systematic studies of eubacteria, using sigma^{70}-type sigma factors of group 1 and group 2. J Bacteriol 179:1734-1747

Gruissem W, Barkan A, Deng XW, Stern D (1988) Transcriptional and post-transcriptional control of plastid mRNA levels in higher plants. Trends Genet 4:258-263

Gruissem W, Elsner-Menzel C, Latshaw S, Narita JO, Schaffer MA, 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, Tonkyn JC (1993) Control mechanisms of plastid gene expression. Crit Rev Plant Sci 12:19-55

Gruissem W, Zurawski G (1985) Analysis of promoter regions for the spinach chloroplast *rbcL*, *atpB* and *psbA* genes. EMBO J 4:3375-3383

Guertin M, Bellemare G (1979) Synthesis of chloroplast ribonucleic acid in *Chlamydomonas reinhardtii* toluene-treated cells. Eur J Biochem 96:125-129

Hajdukiewicz PTJ, Allison LA, Maliga P (1997) The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. EMBO J 16:4041-4048

Hakimi MA, Privat I, Valay JG, Lerbs-Mache S (2000) Evolutionary conservation of C-terminal domains of primary sigma^{70}-type transcription factors between plants and bacteria. J Biol Chem 275:9215-9221

Haley J, Bogorad L (1990) Alternative promoters are used for genes within maize chloroplast polycistronic transcription units. Plant Cell 2:323-333

Han CD, Patrie W, Polacco M, Coe EH (1993) Abberations in plastid transcripts and deficiency of plastid DNA in striped and albino mutants in maize. Planta 191:552-563

Hanaoka M, Kanamaru K, Fujiwara M, Takahashi H, Tanaka K (2005) Glutamyl-tRNA mediates a switch in RNA polymerase use during chloroplast biogenesis. EMBO Rep 6:545-550

Hanaoka M, Kanamaru K, Takahashi H, Tanaka K (2003) Molecular genetic analysis of chloroplast gene promoters dependent on SIG2, a nucleus-encoded sigma factor for the plastid-encoded RNA polymerase, in *Arabidopsis thaliana*. Nucleic Acids Res 31:7090-7098

Hanley-Bowdoin L, Orozco EMJ, Chua NH (1985) *In vitro* synthesis and processing of a maize chloroplast transcript encoded by the ribulose 1,5-bisphosphate carboxylase large subunit gene. Mol Cell Biol 5:2733-2745

Hara K, Morita M, Takahashi R, Sugita M, Kato S, Aoki S (2001a) Characterization of two genes, *Sig1* and *Sig2*, encoding distinct plastid sigma factors(1) in the moss *Physcomi-*

trella patens: phylogenetic relationships to plastid sigma factors in higher plants. FEBS Lett 499:87-91

Hara K, Sugita M, Aoki S (2001b) Cloning and characterization of the cDNA for a plastid sigma factor from the moss *Physcomitrella patens*. Biochim Biophys Acta 1517:302-306

Hedtke B, Börner T, Weihe A (1997) Mitochondrial and chloroplast phage-type RNA polymerases in *Arabidopsis*. Science 277:809-811

Hedtke B, Börner T, Weihe A (2000) One RNA polymerase serving two genomes. EMBO Rep 1:435-440

Hedtke B, Legen J, Weihe A, Herrmann RG, Börner T (2002) Six active phage-type RNA polymerase genes in *Nicotiana tabacum*. Plant J 30:625-637

Hedtke B, Meixner M, Gillandt S, Richter E, Börner T, Weihe A (1999) Green fluorescent protein as a marker to investigate targeting of organellar RNA polymerases of higher plants *in vivo*. Plant J 17:557-561

Hellmund D, Metzlaff M, 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

Herrmann RG, Possingham JV (1980) Plastid DNA - the plastome. In: Reinert J (ed) Results and problems in cell differentiation. Berlin-Heidelberg-New York: Springer Verlag, pp 45-96

Hess WR, Börner T (1999) Organellar RNA polymerases of higher plants. Int Rev Cytol 190:1-59

Hess WR, Muller A, Nagy F, Börner T (1994) Ribosome-deficient plastids affect transcription of light-induced nuclear genes: genetic evidence for a plastid-derived signal. Mol Gen Genet 242:305-312

Hess WR, Prombona A, Fieder B, Subramanian AR, Börner T (1993) Chloroplast *rps15* and the *rpoB/C1/C2* gene cluster are strongly transcribed in ribosome-deficient plastids: evidence for a functioning non-chloroplast-encoded RNA polymerase. EMBO J 12:563-571

Hess WR, Schendel R, Rüdiger W, Börner T (1992) Protochlorophyllide oxidoreductase and chlorophyll synthetase are present in a barley albina mutant unable to synthesize δ-aminolevulinic acid by utilizing the transfer RNA for glutamic acid. Planta 188:19-27

Hirata N, Yonekura D, Yanagisawa S, Iba K (2004) Possible involvement of the 5'-flanking region and the 5'UTR of plastid *accD* gene in NEP-dependent transcription. Plant Cell Physiol 45:176-186

Hoffer PH, Christopher DA (1997) Structure and blue-light-responsive transcription of a chloroplast *psbD* promoter from *Arabidopsis thaliana*. Plant Physiol 115:213-222

Homann A, Link G (2003) DNA-binding and transcription characteristics of three cloned sigma factors from mustard (*Sinapis alba* L.) suggest overlapping and distinct roles in plastid gene expression. Eur J Biochem 270:1288-1300

Hricova A, Quesada V, Micol JL (2006) The *SCABRA3* nuclear gene encodes the plastid RpoTp RNA polymerase, which is required for chloroplast biogenesis and mesophyll cell proliferation in *Arabidopsis*. Plant Physiol 141:942-956

Hu J, 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

Hu J, Troxler RF, Bogorad L (1991) Maize chloroplast RNA polymerase: the 78-kilodalton polypeptide is encoded by the plastid *rpoC*1 gene. Nucleic Acids Res 19:3431-3434

Hübschmann T, Börner T (1998) Characterisation of transcript initiation sites in ribosome-deficient barley plastids. Plant Mol Biol 36:493-496

Hudson GS, Holton DA, Whitfeld PR, Bottomley W (1988) Spinach chloroplast *rpoBC* genes encode three subunits of the chloroplast RNA polymerase. Proc Natl Acad Sci USA 10:525-558

Hwang S, Kawazoe R, Herrin DL (1996) Transcription of *tufA* and other chloroplast-encoded genes is controlled by a circadian clock in *Chlamydomonas*. Proc Natl Acad Sci USA 93:996-1000

Ichikawa K, Sugita M, Imaizumi T, Wada M, Aoki S (2004) Differential expression on a daily basis of plastid sigma factor genes from the moss *Physcomitrella patens*. Regulatory interactions among *PpSig5*, the circadian clock, and blue light signaling mediated by cryptochromes. Plant Physiol 136:4285-4298

Ikeda T, Gray M (1999) Characterization of a DNA-binding protein implicated in transcription in wheat mitochondria. Mol Cell Biol 19:8113-8122

Inada H, Seki M, Morikawa H, Nishimura M, Iba K (1997) Existence of three regulatory regions each containing a highly conserved motif in the promoter of plastid-encoded RNA polymerase gene (*rpoB*). Plant J 11:883-890

Iratni R, Baeza L, Andreeva A, Mache R, Lerbs-Mache S (1994) Regulation of rDNA transcripion in chloroplasts: promoter exclusion by constitutive repression. Genes Dev 8:2928-2938

Iratni R, Diederich L, Harrak H, Bligny M, Lerbs-Mache S (1997) Organ-specific transcription of the *rrn* operon in spinach plastids. J Biol Chem 272:13676-13682

Ishihama A (2000) Functional modulation of *Escherichia coli* RNA polymerase. Annu Rev Microbiol 54:499-518

Ishikura K, Takaoka Y, Kato K, Sekine M, Yoshida K, Shinmyo A (1999) Expression of a foreign gene in *Chlamydomonas reinhardtii* chloroplast. J Biosci Bioeng 87:307-314

Ishizaki Y, Tsunoyama Y, Hatano K, Ando K, Kato K, Shinmyo A, Kobori M, Takeba G, Nakahira Y, Shiina T (2005) A nuclear-encoded sigma factor, *Arabidopsis* SIG6, recognizes sigma^{70} type chloroplast promoters and regulates early chloroplast development in cotyledons. Plant J 42:133-144

Isono K, Niwa Y, Satoh K, Kobayashi H (1997a) Evidence for transcriptional regulation of plastid photosynthesis genes in *Arabidopsis thaliana* roots. Plant Physiol 114:623-630

Isono K, Shimizu M, Yoshimoto K, Niwa Y, Satoh K, Yokota A, Kobayashi H (1997b) Leaf-specifically expressed genes for polypeptides destined for chloroplasts with domains of sigma^{70} factors of bacterial RNA polymerases in *Arabidopsis thaliana*. Proc Natl Acad Sci USA 94:14948-14953

Jaehning JA (1993) Mitochondrial transcription: is a pattern emerging? Mol Microbiol 8:1-4

Jahn D (1992) Expression of the *Chlamydomonas reinhardtii* chloroplast tRNA(Glu) gene in a homologous *in vitro* transcription system is independent of upstream promoter elements. Arch Biochem Biophys 298:505-513

Kabeya Y, Hashimoto K, Sato N (2002) Identification and characterization of two phage-type RNA polymerase cDNAs in the moss *Physcomitrella patens*: implication of recent evolution of nuclear-encoded RNA polymerase of plastids in plants. Plant Cell Physiol 43:245-255

Kabeya Y, Sato N (2005) Unique translation initiation at the second AUG codon determines mitochondrial localization of the phage-type RNA polymerases in the moss *Physcomitrella patens*. Plant Physiol 138:369-382

Kanamaru K, Fujiwara M, Seki M, Katagiri T, Nakamura M, Mochizuki N, Nagatani A, Shinozaki K, Tanaka K, Takahashi H (1999) Plastidic RNA polymerase sigma factors in *Arabidopsis*. Plant Cell Physiol 40:832-842

Kanamaru K, Nagashima A, Fujiwara M, Shimada H, Shirano Y, Nakabayashi K, Shibata D, Tanaka K, Takahashi H (2001) An *Arabidopsis* sigma factor (SIG2)-dependent expression of plastid-encoded tRNAs in chloroplasts. Plant Cell Physiol 42:1034-1043

Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, Nakamura Y, Miyajima N, Hirosawa M, Sugiura M, Sasamoto S, Kimura T, Hosouchi T, Matsuno A, Muraki A, Nakazaki N, Naruo K, Okumura S, Shimpo S, Takeuchi C, Wada T, Watanabe A, Yamada M, Yasuda M, Tabata S (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res 3:109-136

Kapoor S, Sugiura M (1999) Identification of two essential sequence elements in the non-consensus Type II P*atpB*-290 plastid promoter by using plastid transcription extracts from cultured tobacco BY-2 cells. Plant Cell 11:1799-1810

Kapoor S, Suzuki JY, Sugiura M (1997) Identification and functional significance of a new class of non-consensus-type plastid promoters. Plant J 11:327-337

Kapoor S, Wakasugi T, Deno H, Sugiura M (1994) An *atpE*-specific promoter within the coding region of the *atpB* gene in tobacco chloroplast DNA. Curr Genet 26:263-268

Kasai K, Kawagishi-Kobayashi M, Teraishi M, Ito Y, Ochi K, Wakasa K, Tozawa Y (2004) Differential expression of three plastidial sigma factors, *OsSIG1, OsSIG2A*, and *OsSIG2B*, during leaf development in rice. Biosci Biotechnol Biochem 68:973-977

Kasai K, Usami S, Yamada T, Endo Y, Ochi K, Tozawa Y (2002) A RelA-SpoT homolog (Cr-RSH) identified in *Chlamydomonas reinhardtii* generates stringent factor *in vivo* and localizes to chloroplasts *in vitro*. Nucleic Acids Res 30:4985-4992

Kasai S, Yoshimura S, Ishikura K, Takaoka Y, Kobayashi K, Kato K, Shinmyo A (2003) Effect of coding regions on chloroplast gene expression in *Chlamydomonas reinhardtii*. J Biosci Bioeng 95:276-282

Kashino Y, Koike H, Yoshio M, Egashira H, Ikeuchi M, Pakrasi HB, Satoh K (2002) Low-molecular-mass polypeptide components of a photosystem II preparation from the thermophilic cyanobacterium *Thermosynechococcus vulcanus*. Plant Cell Physiol 43:1366-1373

Kawazoe R, Hwang S, Herrin DL (2000) Requirement for cytoplasmic protein synthesis during circadian peaks of transcription of chloroplast-encoded genes in *Chlamydomonas*. Plant Mol Biol 44:699-709

Kestermann M, Neukirchen S, Kloppstech K, Link G (1998) Sequence and expression characteristics of a nuclear-encoded chloroplast sigma factor from mustard (*Sinapis alba*). Nucleic Acids Res 26:2747-2753

Kim JW, Park JK, Kim BH, Lee J-S, Sim WS (2002) Molecular analysis of the accumulation of the transcripts of the large subunit gene of ribulose-1,5-bisphosphate carboxylase/oxygenase by light. Mol Cells 14:281-287

Kim M, Christopher DA, Mullet JE (1999a) ADP-Dependent phosphorylation regulates association of a DNA-binding complex with the barley chloroplast *psbD* blue-light-responsive promoter. Plant Physiol 119:663-670

Kim M, Mullet JE (1995) Identification of a sequence-specific DNA binding factor required for transcription of the barley chloroplast blue light-responsive *psbD-psbC* promoter. Plant Cell 7:1445-1457

Kim M, Thum KE, Morishige DT, Mullet JE (1999b) Detailed architecture of the barley chloroplast *psbD-psbC* blue light-responsive promoter. J Biol Chem 274:4684-4692

Klein RR, Mason HS, Mullet JE (1988) Light-regulated translation of chloroplast proteins. I. Transcripts of *psaA-psaB*, *psbA*, and *rbcL* are associated with polysomes in dark-grown and illuminated barley seedlings. J Cell Biol 106:289-301

Klein RR, Mullet JE (1990) Light-induced transcription of chloroplast genes. *psbA* transcription is differentially enhanced in illuminated barley. J Biol Chem 265:1895-1902

Klein U, De Camp JD, Bogorad L (1992) Two types of chloroplast gene promoters in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA 89:3453-3457

Klein U, Salvador ML, Bogorad L (1994) Activity of the *Chlamydomonas* chloroplast *rbcL* gene promoter is enhanced by a remote sequence element. Proc Natl Acad Sci USA 91:10819-10823

Kobayashi H, Ngernprasirtsiri J, Akazawa T (1990) Transcriptional regulation and DNA methylation in plastids during transitional conversion of chloroplasts to chromoplasts. EMBO J 9:307-313

Kobayashi Y, Dokiya Y, Kumazawa Y, Sugita M (2002) Non-AUG translation initiation of mRNA encoding plastid-targeted phage-type RNA polymerase in *Nicotiana sylvestris*. Biochem Biophys Res Commun 299:57-61

Kobayashi Y, Dokiya Y, Sugita M (2001a) Dual targeting of phage-type RNA polymerase to both mitochondria and plastids is due to alternative translation initiation in single transcripts. Biochem Biophys Res Commun 289:1106-1113

Kobayashi Y, Dokiya Y, Sugiura M, Niwa Y, Sugita M (2001b) Genomic organization and organ-specific expression of a nuclear gene encoding phage-type RNA polymerase in *Nicotiana sylvestris*. Gene 279:33-40

Kovacs L, Wiessner W, Kis M, Nagy F, Mende D, Demeter S (2000) Short- and long-term redox regulation of photosynthetic light energy distribution and photosystem stoichiometry by acetate metabolism in the green alga, *Chlamydobotrys stellata*. Photosynth Res 65:231-247

Krause K, Berg S, Krupinska K (2003) Plastid transcription in the holoparasitic plant genus *Cuscuta*: parallel loss of the *rrn16* PEP-promoter and of the *rpoA* and *rpoB* genes coding for the plastid-encoded RNA polymerase. Planta 216:815-823

Krause K, Krupinska K (2000) Molecular and functional properties of highly purified transcriptionally active chromosomes from spinach chloroplasts. Physiol Plant 109:188-195

Krause K, Maier RM, Kofer W, Krupinska K, Herrmann RG (2000) Disruption of plastid-encoded RNA polymerase genes in tobacco: expression of only a distinct set of genes is not based on selective transcription of the plastid chromosome. Mol Gen Genet 263:1022-1030

Kroll D, Streubel M, Westhoff P (1999) A plastid sigma factor sequence from the C4 monocot *Sorghum bicolor*. Plant Biol 1:180-186

Kuhlemeier C (1992) Transcriptional and post-transcriptional regulation of gene expression in plants. Plant Mol Biol 19:1-14

Kühn K, Bohne A-V, Liere K, Weihe A, Börner T (2007) *Arabidopsis* single-polypeptide RNA polymerases: accurate *in vitro* transcription of organellar genes. Plant Cell 19:959-971

Kusumi K, Yara A, Mitsui N, Tozawa Y, Iba K (2004) Characterization of a rice nuclear-encoded plastid RNA polymerase gene *OsRpoTp*. Plant Cell Physiol 45:1194-1201

Lahiri SD, Allison LA (2000) Complementary expression of two plastid-localized sigma-like factors in maize. Plant Physiol 123:883-894

Lahiri SD, Yao J, McCumbers C, Allison LA (1999) Tissue-specific and light-dependent expression within a family of nuclear-encoded sigma-like factors from *Zea mays*. Mol Cell Biol Res Commun 1:14-20

Lam E, Hanley-Bowdoin L, Chua NH (1988) Characterization of a chloroplast sequence-specific DNA binding factor. J Biol Chem 263:8288-8293

Lang BF, Burger G, O'Kelly CJ, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Gray MW (1997) An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. Nature 387:493-497

Legen J, Kemp S, Krause K, Profanter B, Herrmann RG, Maier RM (2002) Comparative analysis of plastid transcription profiles of entire plastid chromosomes from tobacco attributed to wild-type and PEP-deficient transcription machineries. Plant J 31:171-188

Leister D (2005) Genomics-based dissection of the cross-talk of chloroplasts with the nucleus and mitochondria in *Arabidopsis*. Gene 354:110-116

Leon P, Arroyo A, Mackenzie S (1998) Nuclear control of plastid and mitochondrial development in higher plants. Annu Rev Plant Physiol Plant Mol Biol 49:453-480

Lerbs S, Briat JF, Mache R (1983) Chloroplast RNA polymerase from spinach: purification and DNA-binding proteins. Plant Mol Biol 2:67-74

Lerbs-Mache S (1993) The 110-kDa polypeptide of spinach plastid DNA-dependent RNA polymerase: single-subunit enzyme or catalytic core of multimeric enzyme complexes? Proc Natl Acad Sci USA 90:5509-5513

Lezhneva L, Meurer J (2004) The nuclear factor HCF145 affects chloroplast *psaA-psaB-rps14* transcript abundance in *Arabidopsis thaliana*. Plant J 38:740-753

Li H, Sherman LA (2000) A redox-responsive regulator of photosynthesis gene expression in the cyanobacterium *Synechocystis* sp. Strain PCC 6803. J Bacteriol 182:4268-4277

Liere K, Börner T (2007) Transcription of plastid genes. In: Grasser KD (ed) Regulation of transcription in plants. Oxford: Blackwell Publishing, pp 184-224

Liere K, Kaden D, Maliga P, Börner T (2004) Overexpression of phage-type RNA polymerase RpoTp in tobacco demonstrates its role in chloroplast transcription by recognizing a distinct promoter type. Nucleic Acids Res 32:1159-1165

Liere K, Kestermann M, Müller U, Link G (1995) Identification and characterization of the *Arabidopsis thaliana* chloroplast DNA region containing the genes *psbA*, *trnH* and *rps19'*. Curr Genet 28:128-130

Liere K, Link G (1994) Structure and expression characteristics of the chloroplast DNA region containing the split gene for tRNA(Gly) (UCC) from mustard (*Sinapis alba* L.). Curr Genet 26:557-563

Liere K, Link G (1997) Chloroplast endoribonuclease p54 involved in RNA 3'-end processing is regulated by phosphorylation and redox state. Nucleic Acids Res 25:2403-2408

Liere K, Maliga P (1999a) *In vitro* characterization of the tobacco *rpoB* promoter reveals a core sequence motif conserved between phage-type plastid and plant mitochondrial promoters. EMBO J 18:249-257

Liere K, Maliga P (1999b) Novel *in vitro* transcription assay indicates that the *accD* NEP promoter is contained in a 19 bp fragment. In: Argyroudi-Akoyunoglou JH, Senger H (eds) The chloroplast: from molecular biology to biotechnology. Amsterdam: Kluwer Academic Publishers, pp 79-84

Liere K, Maliga P (2001) Plastid RNA Polymerases. In: Andersson B, Aro E-M (eds) Regulation of photosynthesis. Kluwer Academic Publishers, Netherlands, Dordrecht, pp 29-49

Link G (1984) DNA sequence requirements for the accurate transcription of a protein-coding plastid gene in a plastid *in vitro* transcription system from mustard (*Sinapis alba* L.). EMBO J. 3:1697-1704

Link G (1994) Plastid differentiation: organelle promoters and transcription factors. In: Nover L (ed) Plant promoters and transcription factors - results & problems in cell differentiation. Berlin: Springer Verlag, pp 65-85

Link G (1996) Green life: control of chloroplast gene transcription. BioEssays 18:465-471

Link G (2003) Redox regulation of chloroplast transcription. Antioxid Redox Signal 5:79-87

Link G, Coen DM, Bogorad L (1978) Differential expression of the gene for the large subunit of ribulose bisphosphate carboxylase in maize leaf cell types. Cell 15:725-731

Little MC, Hallick RB (1988) Chloroplast *rpoA*, *rpoB*, and *rpoC* 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

Liu B, Troxler RF (1996) Molecular characterization of a positively photoregulated nuclear gene for a chloroplast RNA polymerase sigma factor in *Cyanidium caldarium*. Proc Natl Acad Sci USA 93:3313-3318

Loiseaux S, Mache R, Rozier C (1975) Rifampicin inhibition of the plastid rRNA synthesis of *Marchantia polymorpha*. J Cell Sci 17:327-335

Lonetto M, Gribskov M, Gross CA (1992) The sigma^{70} family: sequence conservation and evolutionary relationships. J Bacteriol 174:3843-3849

López-Juez E, Pyke KA (2005) Plastids unleashed: their development and their integration in plant development. Int J Dev Biol 49:557-577

Loschelder H, Homann A, Ogrzewalla K, Link G (2004) Proteomics-based sequence analysis of plant gene expression - the chloroplast transcription apparatus. Phytochem 65:1785-1793

Loschelder H, Schweer J, Link B, Link G (2006) Dual temporal role of plastid sigma factor 6 in *Arabidopsis* development. Plant Physiol 142:642-650

Lukens JH, Mathews DE, Durbin RD (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

Lusson NA, Delavault PM, Thalouarn PA (1998) The *rbcL* gene from the non-photosynthetic parasite *Lathraea clandestina* is not transcribed by a plastid-encoded RNA polymerase. Curr Genet 34:212-215

Lysenko EA (2006) Analysis of the evolution of the family of the *Sig* genes encoding plant sigma factors. Russ J Plant Physiol 53:605-614

Lysenko EA, Kuznetsov VV (2005) Plastid RNA Polymerases. Mol Biol 39:661-674

Maliga P (1998) Two plastid polymerases of higher plants: an evolving story. Trends Plant Sci 3:4-6

Masters BS, Stohl LL, Clayton DA (1987) Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. Cell 51:89-99

Mathews DE, Durbin RD (1990) Tagetitoxin inhibits RNA synthesis directed by RNA polymerases from chloroplasts and *Escherichia coli*. J Biol Chem 265:493-498

Matsunaga M, Jaehning JA (2004) Intrinsic promoter recognition by a "core" RNA polymerase. J Biol Chem 279:44239-44242

Matsuzaki M, Misumi O, Shin IT, Maruyama S, Takahara M, Miyagishima SY, Mori T, Nishida K, Yagisawa F, Nishida K, Yoshida Y, Nishimura Y, Nakao S, Kobayashi T, Momoyama Y, Higashiyama T, Minoda A, Sano M, Nomoto H, Oishi K, Hayashi H, Ohta F, Nishizaka S, Haga S, Miura S, Morishita T, Kabeya Y, Terasawa K, Suzuki Y, Ishii Y, Asakawa S, Takano H, Ohta N, Kuroiwa H, Tanaka K, Shimizu N, Sugano S, Sato N, Nozaki H, Ogasawara N, Kohara Y, Kuroiwa T (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. Nature 428:653-657

Mayfield SP, Yohn CB, Cohen A, Danon A (1995) Regulation of chloroplast gene expression. Annu Rev Plant Physiol Plant Mol Biol 46:147-166

McCulloch V, Seidel-Rogol BL, Shadel GS (2002) A human mitochondrial transcription factor is related to RNA adenine methyltransferases and binds S-adenosylmethionine. Mol Cell Biol 22:1116-1125

Meinhardt F, Schaffrath R, Larsen M (1997) Microbial linear plasmids. Appl Microbiol Biotechnol 47:329-336

Meng BY, Wakasugi T, Sugiura M (1991) Two promoters within the *psbK-psbI-trnG* gene cluster in tobacco chloroplast DNA. Curr Genet 20:259-264

Minoda A, Nagasawa K, Hanaoka M, Horiuchi M, Takahashi H, Tanaka K (2005) Microarray profiling of plastid gene expression in a unicellular red alga, *Cyanidioschyzon merolae*. Plant Mol Biol 59:375-385

Misquitta RW, Herrin DL (2005) Circadian regulation of chloroplast gene transcription: a review. Plant Tissue Cult 15:83-101

Miyagi T, Kapoor S, Sugita M, Sugiura M (1998) Transcript analysis of the tobacco plastid operon *rps2/atpI/H/F/A* reveals the existence of a non-consensus type II (NCII) promoter upstream of the *atpI* coding sequence. Mol Gen Genet 257:299-307

Mochizuki T, Onda Y, Fujiwara E, Wada M, Toyoshima Y (2004) Two independent light signals cooperate in the activation of the plastid *psbD* blue light-responsive promoter in *Arabidopsis*. FEBS Lett 571:26-30

Moller SG, Kim YS, Kunkel T, Chua NH (2003) PP7 is a positive regulator of blue light signaling in *Arabidopsis*. Plant Cell 15:1111-1119

Monde RA, Schuster G, Stern DB (2000) Processing and degradation of chloroplast mRNA. Biochimie 82:573-582

Morden CW, Delwiche CF, Kuhsel M, Palmer JD (1992) Gene phylogenies and the endosymbiotic origin of plastids. Biosystems 28:75-90

Morikawa K, Ito S, Tsunoyama Y, Nakahira Y, Shiina T, Toyoshima Y (1999) Circadian-regulated expression of a nuclear encoded plastid sigma factor gene (*sigA*) in wheat seedlings. FEBS Lett 451:275-278

Morikawa K, Shiina T, Murakami S, Toyoshima Y (2002) Novel nuclear-encoded proteins interacting with a plastid sigma factor, Sig1, in *Arabidopsis thaliana*. FEBS Lett 514:300-304

Mullet JE (1993) Dynamic regulation of chloroplast transcription. Plant Physiol 103:309-313

Mullet JE, Orozco EM, Chua N-H (1985) Multiple transcripts for higher plant *rbcL* and *atpB* genes and localization of the transcription initiation sites of the *rbcL* gene. Plant Mol Biol 4:39-54

Nagashima A, Hanaoka M, Motohashi R, Seki M, Shinozaki K, Kanamaru K, Takahashi H, Tanaka K (2004a) DNA microarray analysis of plastid gene expression in an *Arabidopsis* mutant deficient in a plastid transcription factor sigma, SIG2. Biosci Biotechnol Biochem 68:694-704

Nagashima A, Hanaoka M, Shikanai T, Fujiwara M, Kanamaru K, Takahashi H, Tanaka K (2004b) The multiple-stress responsive plastid sigma factor, SIG5, directs activation of the *psbD* blue light-responsive promoter (BLRP) in *Arabidopsis thaliana*. Plant Cell Physiol 45:357-368

Nakahira Y, Baba K, Yoneda A, Shiina T, Toyoshima Y (1998) Circadian-regulated transcription of the *psbD* light-responsive promoter in wheat chloroplasts. Plant Physiol 118:1079-1088

Nakamura T, Furuhashi Y, Hasegawa K, Hashimoto H, Watanabe K, Obokata J, Sugita M, Sugiura M (2003) Array-based analysis on tobacco plastid transcripts: preparation of a genomic microarray containing all genes and all intergenic regions. Plant Cell Physiol 44:861-867

Neuhaus H, Link G (1990) The chloroplast *psbK* operon from mustard (*Sinapis alba* L.): multiple transcripts during seedling development and evidence for divergent overlapping transcription. Curr Genet 18:377-383

Ngernprasirtsiri J, Akazawa T (1990) Modulation of DNA methylation and gene expression in cultured sycamore cells treated by hypomethylating base analog. Eur J Biochem 194:513-520

Ngernprasirtsiri J, Chollet R, Kobayashi H, Sugiyama T, Akazawa T (1989) DNA methylation and the differential expression of C4 photosynthesis genes in mesophyll and bundle sheath cells of greening maize leaves. J Biol Chem 264:8241-8248

Nickelsen J, Link G (1990) Nucleotide sequence of the mustard chloroplast genes *trnH* and *rps19'*. Nucleic Acids Res 18:1051

Nickelsen J, Link G (1991) RNA-protein interactions at transcript 3' ends and evidence for *trnK-psbA* cotranscription in mustard chloroplasts. Mol Gen Genet 228:89-96

Nott A, Jung HS, Koussevitzky S, Chory J (2006) Plastid-to-nucleus retrograde signaling. Annu Rev Plant Biol. 57:739-759

Oelmüller R, Levitan I, Bergfeld R, Rajasekhar VK, Mohr H (1986) Expression of nuclear genes as affected by treatments acting on the plastids. Planta 168:482-492

Ogrzewalla K, Piotrowski M, Reinbothe S, Link G (2002) The plastid transcription kinase from mustard (*Sinapis alba* L.). A nuclear-encoded CK2-type chloroplast enzyme with redox-sensitive function. Eur J Biochem 269:3329-3337

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, Ozeki H (1986) Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322:572-574

Oikawa K, Fujiwara M, Nakazato E, Tanaka K, Takahashi H (2000) Characterization of two plastid sigma factors, SigA1 and SigA2, that mainly function in matured chloroplasts in *Nicotiana tabacum*. Gene 261:221-228

Oikawa K, Tanaka K, Takahashi H (1998) Two types of differentially photo-regulated nuclear genes that encode sigma factors for chloroplast RNA polymerase in the red alga *Cyanidium caldarium* strain RK-1. Gene 210:277-285

Paget MS, Helmann JD (2003) The σ^{70} family of sigma factors. Genome Biol 4:203.1-203.6

Pearson CK, Wilson SB, Schaffer R, Ross AW (1993) NAD turnover and utilisation of metabolites for RNA synthesis in a reaction sensing the redox state of the cytochrome b_6f complex in isolated chloroplasts. Eur J Biochem 218:397-404

Pfalz J, Liere K, Kandlbinder A, Dietz K-J, Oelmüller R (2006) pTAC2, -6 and -12 are components of the Transcriptionally Active Plastid Chromosome that are required for plastid gene expression. Plant Cell 18:176-197

Pfannschmidt T, Liere K (2005) Redox regulation and modification of proteins controlling chloroplast gene expression. Antioxid Redox Signal 7:607-618

Pfannschmidt T, Link G (1994) Separation of two classes of plastid DNA-dependent RNA polymerases that are differentially expressed in mustard (*Sinapis alba* L.) seedlings. Plant Mol Biol 25:69-81

Pfannschmidt T, Link G (1997) The A and B forms of plastid DNA-dependent RNA polymerase from mustard (*Sinapis alba* L.) transcribe the same genes in a different developmental context. Mol Gen Genet 257:35-44

Pfannschmidt T, Nilsson A, Allen JF (1999a) Photosynthetic control of chloroplast gene expression. Nature 397:625-628

Pfannschmidt T, Nilsson A, Tullberg A, Link G, Allen JF (1999b) Direct transcriptional control of the chloroplast genes *psbA* and *psaAB* adjusts photosynthesis to light energy distribution in plants. IUBMB Life 48:271-276

Pfannschmidt T, Ogrzewalla K, Baginsky S, Sickmann A, Meyer HE, Link G (2000) The multisubunit chloroplast RNA polymerase A from mustard (*Sinapis alba* L.): Integration of a prokaryotic core into a larger complex with organelle-specific functions. Eur J Biochem 267:253-261

Privat I, Hakimi MA, Buhot L, Favory JJ, Mache-Lerbs S (2003) Characterization of *Arabidopsis* plastid sigma-like transcription factors SIG1, SIG2 and SIG3. Plant Mol Biol 51:385-399

Pursiheimo S, Mulo P, Rintamaki E, Aro EM (2001) Coregulation of light-harvesting complex II phosphorylation and *lhcb* mRNA accumulation in winter rye. Plant J 26:317-327

Purton S, Gray JC (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

Rantanen A, Gaspari M, Falkenberg M, Gustafsson C, M., Larsson N-G (2003) Characterization of the mouse genes for mitochondrial transcription factors B1 and B2. Mamm Genome 14:1-6

Rapp JC, Baumgartner BJ, Mullet J (1992) Quantitative analysis of transcription and RNA levels of 15 barley chloroplast genes. Transcription rates and mRNA levels vary over 300-fold; predicted mRNA stabilities vary 30-fold. J Biol Chem 267:21404-21411

Reinbothe S, Reinbothe C, Heintzen C, Seidenbecher C, Parthier B (1993) A methyl jasmonate-induced shift in the length of the 5' untranslated region impairs translation of the plastid *rbcL* transcript in barley. EMBO J 12:1505-1512

Reznikov W, Siegle DA, Cowing DW, Gross CA (1985) The regulation of transcription initiation in bacteria. Annu Rev Genet 19:355-387

Richter U, Kiessling J, Hedtke B, Decker E, Reski R, Börner T, Weihe A (2002) Two *RpoT* genes of *Physcomitrella patens* encode phage-type RNA polymerases with dual targeting to mitochondria and plastids. Gene 290:95-105

Rodermel S (2001) Pathways of plastid-to-nucleus signaling. Trends Plant Sci 6:471-478

Ross W, Gosink KK, Salomon J, Igarashi K, Zou C, Ishihama A, Severinov K, Gourse RL (1993) A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. Science 262:1407-1413

Ruf M, Kössel 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

Sakai A, Saito C, Inada N, Kuroiwa T (1998) Transcriptional activities of the chloroplast-nuclei and proplastid-nuclei isolated from tobacco exhibit different sensitivities to tagetitoxin: implication of the presence of distinct RNA polymerases. Plant Cell Physiol 39:928-934

Salvador ML, Klein U (1999) The redox state regulates RNA degradation in the chloroplast of *Chlamydomonas reinhardtii*. Plant Physiol 121:1367-1374

Salvador ML, Klein U, Bogorad L (1993) Light-regulated and endogenous fluctuations of chloroplast transcript levels in *Chlamydomonas*. Regulation by transcription and RNA degradation. Plant J 3:213-219

Salvador ML, Klein U, Bogorad L (1998) Endogenous fluctuations of DNA topology in the chloroplast of *Chlamydomonas reinhardtii*. Mol Cell Biol 18:7235-7242

Satoh J, Baba K, Nakahira Y, Shiina T, Toyoshima Y (1997) Characterization of dynamics of the *psbD* light-induced transcription in mature wheat chloroplasts. Plant Mol Biol 33:267-278

Satoh J, Baba K, Nakahira Y, Tsunoyama Y, Shiina T, Toyoshima Y (1999) Devolpmental stage-specific multi-subunit plastid RNA polymerases (PEP) in wheat. Plant J 18:407-416

Schön A, Krupp G, Gough S, Berry-Lowe S, Kannangara CG, Söll D (1986) The RNA required in the first step of chlorophyll biosynthesis is a chloroplast glutamate tRNA. Nature 322:281-284

Schrubar H, Wanner G, Westhoff P (1990) Transcriptional control of plastid gene expression in greening *sorghum* seedlings. Planta 183:101-111

Schwacke R, Fischer K, Ketelsen B, Krupinska K, Krause K. (2007) Comparative survey of plastid and mitochondrial targeting properties of transcription factors in *Arabidopsis* and rice. Mol Genet Genomics 277:631-646

Schweer J, Loschelder H, Link G (2006) A promoter switch that can rescue a plant sigma factor mutant. FEBS Lett 580:6617-6622

Seidel-Rogol BL, McCulloch V, Shadel GS (2003) Human mitochondrial transcription factor B1 methylates ribosomal RNA at a conserved stem-loop. Nature Genet 33:23-24

Sekine K, Hase T, Sato N (2002) Reversible DNA compaction by sulfite reductase regulates transcriptional activity of chloroplast nucleoids. J Biol Chem 277:24399-24404

Serino G, Maliga P (1998) RNA polymerase subunits encoded by the plastid *rpo* genes are not shared with the nucleus-encoded plastid enzyme. Plant Physiol 117:1165-1170

Severinov K, Mustaev A, Kukarin A, Muzzin O, Bass I, Darst SA, Goldfarb A (1996) Structural modules of the large subunits of RNA polymerase. Introducing archaebacterial and chloroplast split sites in the beta and beta' subunits of *Escherichia coli* RNA polymerase. J Biol Chem 271:27969-27974

Sexton TB, Christopher DA, Mullet JE (1990) Light-induced switch in barley *psbD-psbC* promoter utilization: a novel mechanism regulating chloroplast gene expression. EMBO J 9:4485-4494

Shadel GS, Clayton DA (1993) Mitochondrial Transcription. J Biol Chem 268:16083-16086

Sheveleva EV, Giordani NV, Hallick RB (2002) Identification and comparative analysis of the chloroplast α-subunit gene of DNA-dependent RNA polymerase from seven *Euglena* species. Nucleic Acids Res 30:1247-1254

Shiina T, Allison L, Maliga P (1998) *rbcL* transcript levels in tobacco plastids are independent of light: reduced dark transcription rate is compensated by increased mRNA stability. Plant Cell 10:1713-1722

Shiina T, Tsunoyama Y, Nakahira Y, Khan MS (2005) Plastid RNA Polymerases, Promoters, and Transcription Regulators in Higher Plants. Int Rev Cytol 244:1-68

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsabayashi T, Zaita N, Chungwongse J, Obokata J, Yamaguchi-Shinozaki K, Deno H, Kamogashira T, Yamada K, Kasuda J, Takaiwa F, Kato A, Todoh N, Shimada H, Sugiura M (1986) The complete sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5:2043-2049

Shinozaki K, Sugiura M (1982) The nucleotide sequence of the tobacco chloroplast gene for the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. Gene 20:91-102

Shirano Y, Shimada H, Kanamaru K, Fujiwara M, Tanaka K, Takahashi H, Unno K, Sato S, Tabata S, Hayashi H, Miyake C, Yokota A, Shibata D (2000) Chloroplast development in *Arabidopsis thaliana* requires the nuclear-encoded transcription factor Sigma B. FEBS Lett 485:178-182

Siemenroth A, Wollgiehn R, Neumann D, Börner T (1981) Synthesis of ribosomal RNA in ribosome-feficient plastids of the mutant "albostrians" of *Hordeum vulgare* L. Planta 153:547-555

Sijben-Muller G, Hallick RB, Alt J, Westhoff P, Herrmann RG (1986) Spinach plastid genes coding for initiation factor IF-1, ribosomal protein S11 and RNA polymerase alpha-subunit. Nucleic Acids Res 14:1029-1044

Silhavy D, Maliga P (1998a) Mapping of the promoters for the nucleus-encoded plastid RNA polymerase (NEP) in the *iojap* maize mutant. Curr Genet 33:340-344

Silhavy D, Maliga P (1998b) Plastid promoter utilization in a rice embryonic cell culture. Curr Genet 34:67-70

Smith AC, Purton S (2002) The transcriptional apparatus of algal plastids. Eur J Phycol 37:301-311

Sriraman P, Silhavy D, Maliga P (1998a) The phage-type P*clpP*-53 plastid promoter comprises sequences downstream of the transcription initiation site. Nucleic Acids Res 26:4874-4879

Sriraman P, Silhavy D, Maliga P (1998b) Transcription from heterologous rRNA operon promoters in chloroplasts reveals requirement for specific activating factors. Plant Physiol 117:1495-1499

Stern DB, Higgs DC, Yang JJ (1997) Transcription and translation in chloroplasts. Trends Plant Sci 2:308-315

Stirdivant SM, Crossland LD, Bogorad L (1985) DNA supercoiling affects *in vitro* transcription of two maize chloroplast genes differently. Proc Natl Acad Sci USA 82:4886-4890

Strittmatter G, Godzicka-Josefiak A, Kössel H (1985) Identification of an rRNA operon promoter from *Zea mays* chloroplast which excludes the proximal $tRNA^{Val}$ from the primary transcript. EMBO J 4:599-604

Suck R, Zeltz P, Falk J, Acker A, Kössel H, Krupinska K (1996) Transcriptionally active chromosomes (TACs) of barley chloroplasts contain the α-subunit of plastome encoded RNA polymerase. Curr Genet 30:515-521

Sugiura C, Kobayashi Y, Aoki S, Sugita C, Sugita M (2003) Complete chloroplast DNA sequence of the moss *Physcomitrella patens*: evidence for the loss and relocation of *rpoA* from the chloroplast to the nucleus. Nucleic Acids Res 31:5324-5331

Sun E, Wu BW, Tewari KK (1989) *In vitro* analysis of the pea chloroplast 16S rRNA gene promoter. Mol Cell Biol 9:5650-5659

Surzycki SJ (1969) Genetic functions of the chloroplast of *Chlamydomonas reinhardii*: effect of rifampin on chloroplast DNA-dependent RNA polymerase. Proc Natl Acad Sci USA 63:1327-1334

Surzycki SJ, Shellenbarger DL (1976) Purification and characterization of a putative sigma factor from *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA 73:3961-3965

Suzuki JY, Maliga P (2000) Engineering of the *rpl23* gene cluster to replace the plastid RNA polymerase alpha subunit with the *Escherichia coli* homologue. Current Genetics 38:218-225

Suzuki JY, Sriraman P, Svab Z, Maliga P (2003) Unique architecture of the plastid ribosomal RNA operon promoter recognized by the multisubunit RNA polymerase in tobacco and other higher plants. Plant Cell 15:195-205

Swiatecka-Hagenbruch M, Liere K, Börner T (2007) High diversity of plastidial promoters in *Arabidopsis thaliana*. Mol Genet Genomics 277:725-734

Takahashi K, Kasai K, Ochi K (2004) Identification of the bacterial alarmone guanosine 5'-diphosphate 3'-diphosphate (ppGpp) in plants. Proc Natl Acad Sci USA 101:4320-4324

Tan S, Troxler RF (1999) Characterization of two chloroplast RNA polymerase sigma factors from *Zea mays*: photoregulation and differential expression. Proc Natl Acad Sci USA 96:5316-5321

Tanaka K, Oikawa K, Ohta N, Kuroiwa H, Kuroiwa T, Takahashi H (1996) Nuclear encoding of a chloroplast RNA polymerase sigma subunit in a red alga. Science 272:1932-1935

Tanaka K, Tozawa Y, Mochizuki N, Shinozaki K, Nagatani A, Wakasa K, Takahashi H (1997) Characterization of three cDNA species encoding plastid RNA polymerase sigma factors in *Arabidopsis thaliana*: evidence for the sigma factor heterogeneity in higher plant plastids. FEBS Lett 413:309-313

Thompson RJ, Mosig G (1990) Light affects the structure of *Chlamydomonas* chloroplast chromosomes. Nucleic Acids Res 18:2625-2631

Thum KE, Kim M, Morishige DT, Eibl C, Koop H-U, Mullet JE (2001) Analysis of barley chloroplast *psbD* light-responsive promoter elements in transplastomic tobacco. Plant Mol Biol 47:353-366

Tiller K, Link G (1993a) Phosphorylation and dephosphorylation affect functional characteristics of chloroplast and etioplast transcription systems from mustard (*Sinapis alba* L.). EMBO J 12:1745-1753

Tiller K, Link G (1993b) Sigma-like transcription factors from mustard (*Sinapis alba* L.) etioplast are similar in size to, but functionally distinct from, their chloroplast counterparts. Plant Mol Biol 21:503-513

To KY, Cheng MC, Suen DF, Mon DP, Chen LFO, Chen SCG (1996) Characterization of the light-responsive promoter of rice chloroplast *psbD-C* operon and the sequence-specific DNA binding factor. Plant Cell Physiol 37:660-666

Tokuhisa JG, Vijayan P, Feldmann KA, Browse JA (1998) Chloroplast development at low temperatures requires a homolog of DIM1, a yeast gene encoding the 18S rRNA dimethylase. Plant Cell 10:699-712

Toulokhonov, II, Shulgina I, Hernandez VJ (2001) Binding of the transcription effector ppGpp to *Escherichia coli* RNA polymerase is allosteric, modular, and occurs near the N terminus of the beta'-subunit. J Biol Chem 276:1220-1225

Tozawa Y, Tanaka K, Takahashi H, Wakasa K (1998) Nuclear encoding of a plastid sigma factor in rice and its tissue- and light-dependent expression. Nucleic Acids Res 26:415-419

Tracy RL, Stern DB (1995) Mitochondrial transcription initiation: promoter structures and RNA polymerases. Curr Genet 28:205-216

Trebitsh T, Levitan A, Sofer A, Danon A (2000) Translation of chloroplast *psbA* mRNA is modulated in the light by counteracting oxidizing and reducing activities. Mol Cell Biol 20:1116-1123

Trifa Y, Privat I, Gagnon J, Baeza L, Lerbs-Mache S (1998) The nuclear *RPL4* gene encodes a chloroplast protein that co-purifies with the T7-like transcription complex as well as plastid ribosomes. J Biol Chem 273:3980-3985

Troxler RF, Zhang F, Hu J, Bogorad L (1994) Evidence that sigma factors are components of chloroplast RNA polymerase. Plant Phys 104:753-759

Tsunoyama Y, Ishizaki Y, Morikawa K, Kobori M, Nakahira Y, Takeba G, Toyoshima Y, Shiina T (2004) Blue light-induced transcription of plastid-encoded *psbD* gene is mediated by a nuclear-encoded transcription initiation factor, AtSig5. Proc Natl Acad Sci USA 101:3304-3309

Tsunoyama Y, Morikawa K, Shiina T, Toyoshima Y (2002) Blue light specific and differential expression of a plastid sigma factor, Sig5 in *Arabidopsis thaliana*. FEBS Lett 516:225-228

Tsunoyama Y, Toyoshima Y, Shiina T, Toyoshima Y (2001) Analysis of promoter selectivity of plastid sigma factors in *Arabidopsis thaliana*. Science Access 3:1-4

Tullberg A, Alexciev K, Pfannschmidt T, Allen JF (2000) Photosynthetic electron flow regulates transcription of the *psaB* gene in pea (*Pisum sativum* L.) chloroplasts through the redox state of the plastoquinone pool. Plant Cell Physiol 41:1045-1054

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen GL, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Dejardin A, dePamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjarvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leple JC, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napoli C, Nelson DR, Nelson C, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouze P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shin H, Siddiqui A, Sterky F, Terry A, Tsai CJ, Uberbacher E, Unneberg P, Vahala J, Wall K, Wessler S, Yang G, Yin T, Douglas C, Marra M, Sandberg G, Van de Peer Y, Rokhsar D (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313:1596-1604

Tzafrir I, Pena-Muralla R, Dickerman A, Berg M, Rogers R, Hutchens S, Sweeney TC, McElver J, Aux G, Patton D, Meinke D (2004) Identification of genes required for embryo development in *Arabidopsis*. Plant Physiology 135:1206-1220 van der Biezen EA, Sun J, Coleman MJ, Bibb MJ, Jones JD (2000) *Arabidopsis* RelA/SpoT homologs implicate (p)ppGpp in plant signaling. Proc Natl Acad Sci USA 97:3747-3752

Vener AV, van Kan PJ, Rich PR, Ohad II, Andersson B (1997) Plastoquinol at the quinol oxidation site of reduced cytochrome *bf* mediates signal transduction between light and protein phosphorylation: Thylakoid protein kinase deactivation by a single-turnover flash. Proc Natl Acad Sci USA 94:1585-1590

Vera A, Hirose T, Sugiura M (1996) A ribosomal protein gene (*rpl32*) from tobacco chloroplast DNA is transcribed from alternative promoters - similarities in promoter region organization in plastid housekeeping genes. Mol Gen Genet 251:518-525

Vera A, Sugiura M (1995) Chloroplast rRNA transcription from structurally different tandem promoters: an additional novel-type promoter. Curr Genet 27:280-284

Wada T, Tunoyama Y, Shiina T, Toyoshima Y (1994) *In vitro* analysis of light-induced transcription in the wheat *psbD/C* gene cluster using plastid extracts from dark-grown and short-term-illuminated seedlings. Plant Phys 104:1259-1267

Wagner R, Pfannschmidt T (2006) Eukaryotic transcription factors in plastids--Bioinformatic assessment and implications for the evolution of gene expression machineries in plants. Gene 381:62-70

Walter G, Müller A, Hoffmann P, Börner T (1995) Tetrapyrrole synthesis in leaves and roots of a barley albina mutant with extremely low amount of plastid glutamyl-tRNA. In: Mathis P (ed) Photosynthesis: from light to biosphere. Dordrecht/Boston/London: Kluwer Academic Publishers, pp 937–940

Weihe A (2004) The transcription of plant organelle genomes. In: Daniell H, Chase CD (eds) Molecular biology and biotechnology of plant organelles. Dordrecht: Kluwer Academic Publishers, pp 213-237

Weihe A, Börner T (1999) Transcription and the architecture of promoters in chloroplasts. Trends Plant Sci 4:169-170

Weihe A, Hedtke B, Börner T (1997) Cloning and characterization of a cDNA encoding a bacteriophage-type RNA polymerase from the higher plant *Chenopodium album*. Nucleic Acids Res 25:2319-2325

Wilson RJ, Denny PW, Preiser PR, Rangachari K, Roberts K, Roy A, Whyte A, Strath M, Moore DJ, Moore PW, Williamson DH (1996) Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. J Mol Biol 261:155-172

Wolfe KH, Morden CW, Ems SC, Palmer JD (1992a) Rapid evolution of the plastid translational apparatus in a nonphotosynthetic plant: loss or accelerated sequence evolution of tRNA and ribosomal protein genes. J Mol Evol 35:304-317

Wolfe KH, Morden CW, Palmer JD (1992b) Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. Proc Natl Acad Sci USA 89:10648-10652

Wösten MM (1998) Eubacterial sigma-factors. FEMS Microbiol Rev 22:127-150

Wu CY, Lin CH, Chen LJ (1997) Identification of the transcription start site for the spinach chloroplast serine tRNA gene. FEBS Lett 418:157-161

Xie G, Allison LA (2002) Sequences upstream of the YRTA core region are essential for transcription of the tobacco *atpB* NEP promoter in chloroplasts *in vivo*. Curr Genet 41:176-182

Yao J, Roy-Chowdhury S, Allison LA (2003) AtSig5 is an essential nucleus-encoded *Arabidopsis* σ-Like factor. Plant Physiol 132:739-747

Yao WB, Meng BY, Tanaka M, 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

Zapata JM, Guera A, Esteban-Carrasco A, Martin M, Sabater B (2005) Chloroplasts regulate leaf senescence: delayed senescence in transgenic *ndhF*-defective tobacco. Cell Death Differ 12:1277-1284

Zengel JM, Mueckl D, Lindahl L (1980) Protein L4 of the *E. coli* ribosome regulates an eleven gene r protein operon. Cell 21:523-535

Zer H, Ohad I (2003) Light, redox state, thylakoid-protein phosphorylation and signaling gene expression. Trends Biochem Sci 28:467-470

Zghidi W, Merendino L, Cottet A, Mache R, Lerbs-Mache S (2006) Nucleus-encoded plastid sigma factor SIG3 transcribes specifically the *psbN* gene in plastids. Nucleic Acids Res:Online 1-10

Zhang L, Paakkarinen V, van Wijk KJ, Aro EM (2000) Biogenesis of the chloroplast-encoded D1 protein: regulation of translation elongation, insertion, and assembly into photosystem II. Plant Cell 12:1769-1782

Zito F, Finazzi G, Delosme R, Nitschke W, Picot D, Wollman FA (1999) The Qo site of cytochrome b_6f complexes controls the activation of the LHCII kinase. EMBO J 18:2961-2969

Zoschke R, Liere K, Börner T (2007) From seedling to mature plant: *Arabidopsis* plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. Plant J 50:710-722

Börner, Thomas
Institut für Biologie / Genetik, Humboldt-Universität zu Berlin, Chausseestr. 117, 10115 Berlin, Germany
thomas.boerner@rz.hu-berlin.de

Liere, Karsten
Institut für Biologie / Genetik, Humboldt-Universität zu Berlin, Chausseestr. 117, 10115 Berlin, Germany

Processing, degradation, and polyadenylation of chloroplast transcripts

Thomas J. Bollenbach, Gadi Schuster, Victoria Portnoy, and David B. Stern

Abstract

In this chapter, we describe the major enzymes and characteristics of transcript 5' and 3' end maturation, and polyadenylation-stimulated degradation. The picture which emerges is that maturation and degradation share many prokaryotic features, vestiges of the chloroplast endosymbiont ancestor. The major exoribonucleases are well-defined, being polynucleotide phosphorylase and RNase II/R. The endonucleases include CSP41, with largely informatic evidence for homologs of prokaryotic RNases E, J, and III. The polyadenylation-stimulated degradation pathway, which occurs in most living systems, is a major player in chloroplast RNA degradation. We discuss known or potential roles for polynucleotide phosphorylase and a prokaryotic-type poly(A) polymerase. Finally, we discuss nuclear mutations that affect RNA maturation and degradation, defining genes that are likely or known to encode regulatory factors. Major questions for future research include how the ribonucleases, which are inherently nonspecific, interact with these specificity factors, and whether newly-discovered noncoding RNAs in the chloroplast play any role in RNA metabolism.

1 Introduction

Chloroplasts originated from a cyanobacterial ancestor that entered a heterotrophically growing eukaryote some 1.5 billion years ago (Hoffmeister and Martin 2003). Ensuing gene transfer from the organelle to the nucleus has been extensive, resulting in a situation where the vast majority of the chloroplast proteome is encoded either by nuclear genes acquired from the endosymbiont, or by those that already existed in the nucleus of the mitochondriate host (Martin et al. 2002). Consequently, the chloroplast multisubunit complexes required for photosynthesis and gene expression contain both chloroplast- and nucleus-encoded components, necessitating coordinated gene expression in the two compartments. Plants and green algae have therefore evolved sophisticated intracellular communication systems that regulate chloroplast gene expression at multiple levels, many of which are reviewed in other chapters of this book.

This chapter concerns primarily posttranscription regulation of chloroplast gene expression, particularly RNA processing and degradation. RNA processing in

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0235 / Published online: 4 July 2007

chloroplasts is catalyzed by nucleus-encoded ribonucleases and includes 5' end maturation, which is catalyzed primarily by endoribonucleases and 3' end maturation, which is catalyzed by endonucleases and/or 3' to 5' exoribonucleases (Stern and Kindle 1993; Hayes et al. 1996). Like bacteria, chloroplasts often express genes from clusters or operons, leading to synthesis of polycistronic transcripts that are often cleaved intercistronically, requiring endoribonuclease activity and RNA-binding proteins (Barkan et al. 1994; Meierhoff et al. 2003). Although splicing and RNA editing are also important posttranscriptional processing events in the chloroplast, the reader is directed to the chapter by Christian Schmitz-Linneweber in this volume for a comprehensive discussion of these topics.

Although endo- and exoribonucleases feature prominently in RNA processing, these same activities are also important in catalyzing chloroplast RNA turnover. Chloroplast RNA accumulation increases significantly during leaf development and plastid differentiation. The accumulation of a specific transcript is controlled by the difference in its transcription and degradation rates, and can in principle be controlled at either one or both of these steps. Although global changes in plastid transcription are associated with leaf development and illumination (Deng and Gruissem 1987; Mullet and Klein 1987; Dhingra et al. 2006; Zoschke et al. 2007), chloroplast genes are rarely regulated individually at the transcriptional level, with the notable exception of *psbD*, which is regulated by a specialized promoter (Gamble and Mullet 1989; Kim et al. 1999; Thum et al. 2001). Instead, the significant differences in the accumulation of individual transcripts in various tissues and during leaf development and plastid differentiation are modulated in large part by transcript degradation rates, or at the level of RNA stability (Gruissem 1989; Monde et al. 2000b; Bollenbach et al. 2004). Chloroplast RNA stability is regulated primarily by its rate of degradation through a polyadenylation-stimulated turnover pathway, which is discussed in detail below. mRNA abundance for a handful of plant chloroplast genes has been shown to correlate with abundance of their respective proteins, consistent with the idea that regulation of mRNA accumulation is an important control point of chloroplast gene expression (Rapp et al. 1992; Mullet 1993). However, translation is also a key regulatory step, and *Chlamydomonas reinhardtii* chloroplasts maintain protein homeostasis even in the face of significant decreases in mRNA accumulation (Eberhard et al. 2002).

In this review, we describe the mechanisms of chloroplast RNA processing and degradation, including known and candidate endoribonucleases, exoribonucleases and regulatory proteins. The role of these nucleus-encoded factors, and the potential role of newly discovered chloroplast-encoded antisense RNAs in posttranscriptional regulation are discussed.

2 The enzymes of RNA degradation and maturation

2.1 Endoribonucleases

2.1.1 CSP41

CSP41a (Chloroplast Stem-loop binding Protein, 41 kDa) and CSP41b are widespread, highly conserved endoribonucleases, which are unique to photosynthetic organisms. The photosynthetic bacteria *Synechocystis* sp. PCC6803 and *Nostoc* sp. PCC7120 encode only a CSP41b homolog, whereas plant and algal nuclear genomes encode both CSP41a and CSP41b homologs (Yamaguchi et al. 2003). Phylogenetic and motif analyses have shown that CSP41a and CSP41b are paralogs of a cyanobacterial ancestor that diverged from a bacterial epimerase/dehydratase (Baker et al. 1998; Yamaguchi et al. 2003).

CSP41a and CSP41b are abundant proteins, and have been found in a number of chloroplast complexes by proteomics, including RNPs, chloroplast ribosomes, and the plastid-encoded RNA polymerase (Yang et al. 1996; Pfannschmidt et al. 2000; Yamaguchi et al. 2003; Suzuki et al. 2004; Peltier et al. 2006), although no primary function for these proteins in either transcription or translation has been demonstrated.

CSP41a was first purified from spinach chloroplasts as a *petD*-specific RNA-binding protein and a nonspecific endoribonuclease (Yang et al. 1996; Yang and Stern 1997). Spinach CSP41a was shown to cleave synthetic stem-loop-containing *petD*, *psbA*, and *rbcL* RNAs, and could cleave arbitrary single-stranded RNAs (Yang and Stern 1997), which suggested that it could initiate turnover of chloroplast transcripts by endonucleolytic cleavage, the first step in the poly(A)-stimulated turnover pathway (see Section 2). *In vitro* measurements of tobacco chloroplast mRNA degradation rates showed significant decreases in the rates of *rbcL*, *psbA*, and *petD* transcript turnover in CSP41a-deficient plants (Bollenbach et al. 2003), suggesting that CSP41a could participate broadly in chloroplast mRNA turnover.

Structure. A Hidden Markov model-based search of Genpept suggested that CSP41 proteins are homologous to sugar-nucleotide epimerases and hydroxysteroid reductases, and as such belong to the short-chain dehydrogenase/reductase (SDR) superfamily (Baker et al. 1998). This family comprises 1600 proteins, including more than 130 in *Arabidopsis* (Kallberg et al. 2002). Like other members of this family, CSP41 contains an N-terminal bidomain Rossman fold, including the βαβ-turn, which is responsible for binding the nucleotide portion of NAD(P)H in dehydrogenases. CSP41 homologs have, however, lost the conserved Gly-X-Gly-X_3-Gly NAD(P)H binding motif, and have therefore lost the ability to bind $NAD(P)^+$ or NAD(P)H (Baker et al. 1998; Bollenbach and Stern 2003a). Instead, deletion mutant analysis suggested that the N-terminal CSP41 Rossman fold is responsible for substrate (RNA) binding (Bollenbach and Stern 2003b).

Divalent metal requirement. Several SDR family proteins bind and cleave RNA, including glyceraldehyde phosphate dehydrogenase (GAPDH), and two endoribonucleases from the archaeon *Sulfolobus solfataricus*, but do not require divalent metal ions for activity (Evguenieva-Hackenberg et al. 2002). CSP41, a divalent metal-dependent ribonuclease, is therefore unique among RNA-cleaving SDR enzymes. CSP41a contains a single, broad specificity divalent metal binding site, but is optimally active in the presence of Mg^{2+}; the abundance of Mg^{2+} in the chloroplast suggests that this is the physiological activator of CSP41a (Bollenbach and Stern 2003a). Interestingly, the $K_{A,Mg}{}^{2+}$ for CSP41a is approximately 2 mM, a value that is within the physiological Mg^{2+} concentration range, which varies from 0.5 mM in etiolated leaves to 2-3 mM in young light-grown leaves and 10 mM in mature green leaves (Horlitz and Klaff 2000; Ishijima et al. 2003). Although CSP41b is known to catalyze a divalent metal-dependent reaction (Bollenbach and Stern, unpublished data), the biophysical parameters describing its interaction with Mg^{2+} remain to be tested.

The physiological variation in stromal Mg^{2+} concentration suggested that light-dependent and developmental fluctuations in Mg^{2+} could regulate CSP41a activity *in vivo* (Yang et al. 1996). This hypothesis was verified by experiments in which the turnover of *rbcL* in lysed WT and CSP41a-deficient chloroplasts was measured as a function of free Mg^{2+}, which was varied from <1 mM to 12.5 mM (Bollenbach et al. 2003). Whereas the rate of *rbcL* turnover was invariant in chloroplasts from WT plants, its rate of turnover increased as a function of decreasing Mg^{2+} in chloroplasts from CSP41a-depleted plants. Together, these experiments suggested that CSP41a provides the primary route for transcript cleavage at high stromal Mg^{2+} concentrations but that it is bypassed, possibly by another endoribonuclease such as RNase E, RNase J, p54 or CSP41b (see Sections 1.1.2-1.1.4), at lower Mg^{2+} concentrations where CSP41a is only minimally active.

Substrate specificity. Most chloroplast open reading frames encode inverted repeat (IR) sequences in their 3' untranslated regions that can fold into stable stem-loop structures. Prior research has shown that these IRs act as processing determinants and protect upstream sequences against 3' to 5' exonucleolytic degradation (Stern and Gruissem 1987). CSP41 has no sequence specificity, but displays a substrate preference for stem-loop containing RNAs from *petD*, *psbA* and *rbcL in vitro* (Yang and Stern 1997). This property would potentially target CSP41 to mature RNAs for turnover (Bollenbach et al. 2003).

CSP41 activity was shown to be optimal with substrates containing fully base-paired stem-loops, whereas deletion of part or all of a stem-loop structure resulted in a 100-fold decrease in activity (Bollenbach and Stern 2003b). Mutations at the scissile bond, and mutations or deletions of the terminal loop structure had only minor effects on activity, whereas changes in stem torsion, either by intercalation of ethidium or though the introduction of single base bulges into either arm of the stem-loop, had more drastic effects. Together with *in vitro* measurements of several mRNA degradation rates in WT and CSP41a-deficient chloroplasts, this suggests that CSP41 has a broad substrate specificity, and that stem-loop structure is a major determinant of CSP41 cleavage rates, and therefore of transcript half-life.

2.1.2 RNase E/G

Ribonuclease E is generally believed to initiate RNA degradation in *E. coli* and also mediates the processing of certain rRNAs and tRNAs (Kushner 2002). *E. coli* and some other bacteria also encode a homolog, RNase G, which lacks the C-terminal domain (Fig. 1). RNase E, but not RNase G, is essential in *E. coli* and *Synechocystis* (Cohen and McDowall 1997; Rott et al. 2003).

Full-length or partial ESTs have been found for rice, *Arabidopsis*, tomato, barley, cocoa, grape, ice plant, sorghum, wheat, maize, soybean, and *Medicago truncatula*. Each of these RNase E/G homologs resembles the *E. coli* enzyme in the catalytic region, but lacks the C-terminal domain and contains an N-terminal extension.

In *E. coli* and several other related bacteria, RNase E is a component of the degradosome (Vanzo et al. 1998), a multiprotein complex that also contains PNPase, the DEAD-box RNA helicase RhlB, and the glycolytic enzyme enolase (Blum et al. 1999), which is believed to be important for mRNA degradation and processing (Symmons et al. 2002; Marcaida et al. 2006). Degradosome assembly is dependent on the RNase E C-terminal domain (Coburn et al. 1999). The absence of the C-terminal domain in plant RNase E/G homologs correlates with the absence of a degradosome in chloroplasts (Baginsky et al. 2001). The N-terminal extension is reminiscent of a chloroplast transit peptide (Fig. 1), and when the "plant-specific" extension of the *Arabidopsis* protein is analyzed for possible chloroplast targeting using bioinformatic tools, chloroplast localization is predicted (PCLR, 68%; TargetP, 69%; Predotar, 58%). A partial sequence of this protein was also reported in a Triton-insoluble pea chloroplast fraction (Phinney and Thelen 2005). Given this information, and the fact that RNase E has never been found in mitochondria, support the hypothesis that the nucleus-encoded RNase E homolog functions in the chloroplast and is responsible for an initial step in RNA degradation and/or for intercistronic processing (see Section 3.1.2). However, the function(s) of RNase E alone and/or within the context of other chloroplast endoribonucleases such as CSP41 remains speculative and awaits further analysis.

2.1.3 RNase J

Many organisms lack an RNase E homolog, suggesting that another endoribonuclease is responsible for endonucleolytic processing and turnover. Recently, the purification and identification of two novel *B. subtilis* endoribonucleases, RNases J1 and J2, was described (Even et al. 2005). These RNases, like the tRNA 3' processing endonuclease RNase Z, belong to β–CASP family of zinc-dependent metallo β–lactamases (de la Sierra-Gallay et al. 2005; Even et al. 2005) and *in vitro* assays suggest they are functionally homologous to RNase E, since they have the same substrate specificity, both in terms of cleavage site selection and in their preference for 5' monophosphorylated RNA substrates (Even et al. 2005).

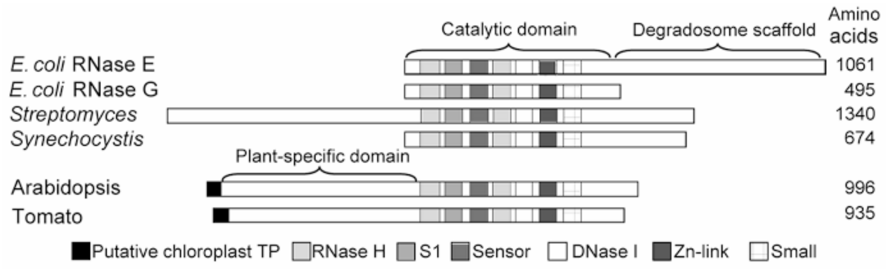


Fig. 1. Schematic amino acid alignment of RNase E homologs performed using MEME. Regions of significant homology are shown as textured boxes, with the catalytic subdomains named according to the recently solved structure (Callaghan et al. 2005). The catalytic and C-terminal degradosome scaffolding domains are highlighted by brackets at the top; the C-terminal domain is not conserved in any other protein shown. The "plant-specific" domains in *Arabidopsis* and tomato have no similarity to the *Streptomyces* N-terminal extension, and are preceded by putative plastid transit peptides (TP).

RNase J homologs are widespread in the eubacteria and archaea and although they appear to replace RNase E in many organisms, some encode both types of enzymes. The occurrence of both RNase E/G and RNase J in *Synechocystis* (Rott et al. 2003; Even et al. 2005) prompted us to search for RNase J homologs in the *Chlamydomonas* and *Arabidopsis* nuclear genomes. Each of these genomes contains a single *RNJ* gene (Positions 1136733-1144060, Scaffold 14 of the *Chlamydomonas* genome v3.0, and At5g63420, respectively), and the N-terminus of the *Arabidopsis* gene product targets GFP to chloroplasts in transient assays (Bollenbach and Stern, unpublished data). Any function of this enzyme in chloroplast RNA metabolism remains to be demonstrated, but it is essential for embryo development because plants heterozygous for a T-DNA insertion in the *RNJ* coding sequence produce siliques containing aborted embryos (www.seedgenes.org). This may be related to a function in 16S rRNA and/or ribosome assembly maturation, as was recently reported for the *B. subtilis* enzyme (Britton et al. 2007).

2.1.4 p54

RNase activities have been purified from chloroplasts for which no specific gene product has been associated (Nickelsen and Link 1989; Chen and Stern 1991). A well-characterized example is p54, a chloroplast RNA-binding protein and endoribonuclease originally identified by *in vitro* studies with mustard chloroplast protein extracts (Nickelsen and Link 1989, 1991). The interaction between p54 and RNA and its subsequent endonucleolytic cleavage were shown to be dependent on a heptamer motif located within the 3' non-coding regions of $tRNA^{Lys}$ and *rps16* mRNAs (Nickelsen and Link 1989). Therefore, p54 was hypothesized to be essential for $tRNA^{Lys}$ and *rps16* 3' processing, and *in vitro* cleavage sites correlated well with $tRNA^{Lys}$ and *rps16* mRNA 3' ends detected *in vivo* (Neuhaus et al. 1989; Nickelsen and Link 1991). Failure to bind $tRNA^{Gln}$, however, suggests that p54 is not a broadly specific in chloroplast tRNA 3' maturation (Nickelsen and

Link 1989); a role in tRNA 3' processing has recently ascribed to a chloroplast RNase Z homolog (Schiffer et al. 2002).

p54 is a divalent metal-independent ribonuclease and because its activity is not dependent on RNA secondary structure (Nickelsen and Link 1989, 1991) it has been suggested that it catalyzes RNA processing and/or turnover under conditions or on substrates where CSP41 is inactive (Bollenbach et al. 2003). Testing this hypothesis awaits identification of the p54 gene, and subsequent *in vivo* analysis. It cannot be ruled out, in fact, that p54 is none other than the Rubisco LS, which has recently been shown to have RNA-binding properties (Yosef et al. 2004), but which was not tested for endonuclease activity. Indeed, in our hands the two proteins co-purify (S. Preiss and D. Stern, unpublished results), and both p54 (Liere and Link 1997) and LS are redox-sensitive as RNA interactors.

2.2 Exoribonucleases

2.2.1 PNPase (polynucleotide phosphorylase)

PNPase (EC 2.7.7.8) was discovered during studies of biological phosphorylation in *Azotobacter vinelandii* (Grunberg-Manago and Ochoa 1955), and was later characterized in the context of its role in *E. coli* RNA synthesis (Littauer and Soreq 1982). In fact, PNPase was the first enzyme shown to catalyze the formation of polynucleotides from ribonucleotides; unlike RNA polymerases, PNPase catalyzes this reaction in a template-independent manner.

As a phosphorylase, PNPase catalyzes both processive 3' to 5' degradation and RNA polymerization, and in bacteria and organelles, participates in the degradation, processing and polyadenylation of RNA (Hayes et al. 1996; Grunberg-Manago 1999; Littauer and Grunberg-Manago 1999; Jarrige et al. 2002; Bollenbach et al. 2004; Slomovic et al. 2006a). PNPase was also reported to be a global regulator of virulence and persistency in *Salmonella enterica* (Clements et al. 2002), and its activity in some way regulates both chloroplast isoprenoid metabolism (Sauret-Gueto et al. 2006) and in *Chlamydomonas*, its ability to survive phosphate starvation (Yehudai-Resheff et al. 2007). Human PNPase was recently shown to be localized to the mitochondrial inter-membrane space (Chen et al. 2006; French et al. 2006; Rainey et al. 2006), and was identified in an overlapping-pathway screen to discover genes displaying coordinated expression as a consequence of terminal differentiation and senescence of melanoma cells (Leszczyniecka et al. 2002; Sarkar et al. 2003). Genes encoding PNPase homologs have been identified in almost all prokaryotes and eukaryotes with the exception of the *Mycoplasma*, trypanosomes and yeast (Slomovic et al. 2006a). In addition, there is no PNPase in archaea, though the hyperthermophiles and some methanogenic archaea contain an exosome that is very similar to the PNPase (Fig. 2) (Lorentzen et al. 2005; Portnoy et al. 2005; Slomovic et al. 2006a).

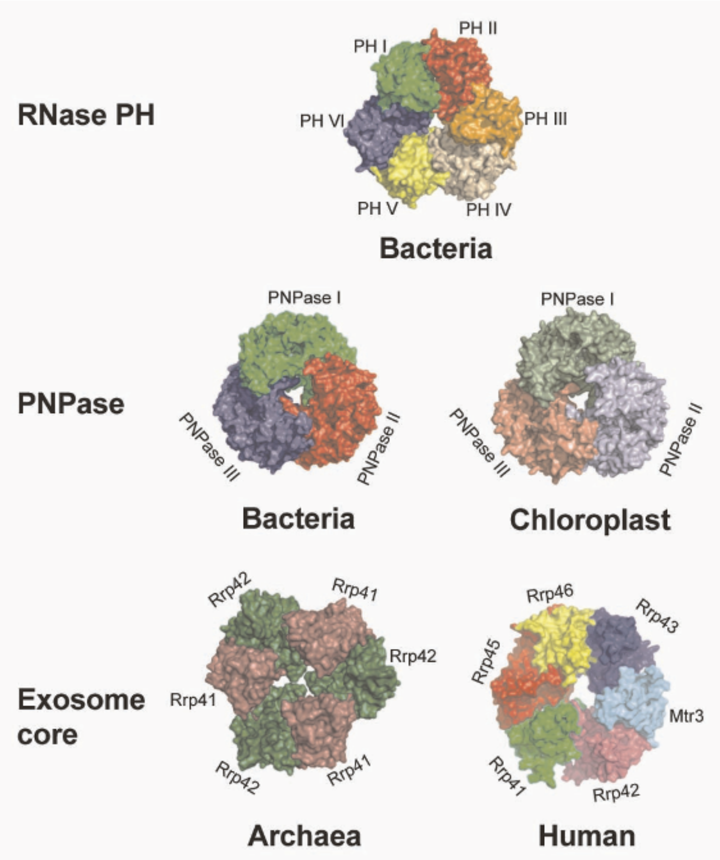


Fig. 2. Similarities in the structures of RNase PH, bacterial and chloroplast PNPase and the archaeal and human exosome cores. The bacterial RNase PH structure (Ishii et al. 2003; Harlow et al. 2004) and bacterial PNPase (Symmons et al. 2000a), archaeal (Buttner et al. 2005; Lorentzen et al. 2005) and human (Liu et al. 2006) exosomes, as well as the predicted structure of the chloroplast PNPase (Yehudai-Resheff et al. 2003), are shown in order to compare the ring shapes. The molecular surfaces are represented such that each protein subunit is differently colored. The structures were generated using PyMOL.

The primary structures of PNPases from bacteria and the nuclear genomes of plants and mammals comprise five domains, which are two N-terminal core domains homologous to the *E. coli* phosphorylase RNase PH, which are separated by an α-helical domain, and two C-terminal RNA-binding domains (KH and S1) (Symmons et al. 2000b, 2002; Zuo and Deutscher 2001; Raijmakers et al. 2002; Yehudai-Resheff et al. 2003). X-ray crystallographic analysis was used to reveal the three-dimensional structure of the PNPase from the bacterium *Streptomyces antibioticus*. The enzyme is arranged in a homotrimeric complex forming a circle (doughnut), which surrounds a central channel that can accommodate a single-stranded RNA molecule (Fig. 2) (Symmons et al. 2000b, 2002).

The domains of spinach chloroplast PNPase were analyzed in detail using a series of recombinant proteins (Yehudai-Resheff et al. 2003). It was found that the first core domain, which was predicted to be inactive in bacterial enzymes, was active in RNA degradation but not in polymerization. Surprisingly, the second core domain was found to be active only in degrading polyadenylated RNA, suggesting that non-polyadenylated molecules can be degraded by this domain only if tails are added, apparently by the same protein (see Section 2.4.2). The high-affinity poly(A) binding site was localized to the S1 domain.

Recent observations suggest the unexpected conclusion that bacterial and chloroplast PNPases are evolutionary related to the archaeal and eukaryotic exosomes. The exosome functions in 3’ to 5’ RNA degradation, processing, and quality control of gene expression in the cytoplasm and nucleus of eukaryotic cells (Houseley et al. 2006), and is comprised of 10-11 proteins including six related to the phosphorylase RNase PH and two to the S1 and KH RNA-binding domains. Overall, the exosome is structurally similar to trimeric PNPase (Fig. 2)(Aloy et al. 2002; Raijmakers et al. 2002; Yehudai-Resheff et al. 2003; Hernandez et al. 2006; Liu et al. 2006). Therefore, the PNPase/archaeal exosome/eukaryotic exosome represent a functionally and evolutionary conserved machine for 3’ to 5’ exonucleolytic degradation.

2.2.2 RNase II/R

The RNR exoribonuclease family, which is typified by *E. coli* RNase II and RNase R, are hydrolytic processive 3’ to 5’ exoribonucleases that release 5’ mononophosphates. These enzymes are widely distributed among eukaryotes, eubacteria, mycoplasma and archaea. While most eukaryotic nuclear genomes encode at least three RNR homologs, some encode only a single RNR-like enzyme, and exceptional ones such as *Mycoplasma* encode a single RNR homolog as the only exoribonuclease (Zuo and Deutscher 2001). The halophilic archaea also contain an RNR homolog, while hyperthermophiles and several methanogens contain the archaeal exosome, which is similar to PNPase (Portnoy et al. 2005; Portnoy and Schuster 2006). Interestingly, no homolog of could be detected in methanogens that do not contain the archael exosome (Ng et al. 2000; Portnoy and Schuster 2006). The *Arabidopsis* nuclear genome encodes three homologs including RNR1, which is both plastid and mitochondria-localized, and RNR2 and RNR3, which based on GFP fusions are localized to the nucleus and cytosol, respectively, and are therefore putative exosome subunits (Perrin et al. 2004; Bollenbach et al. 2005).

In *E. coli*, the RNR family enzymes differ in their ability to remain processive through secondary structures. For example, RNase II becomes distributive near stem-loops and is eventually inhibited by them, while RNase R can melt secondary structures (Cheng and Deutscher 2002). Therefore, although in *E. coli* both enzymes are nonspecific exonucleases, RNase II is more active on single-stranded homopolymeric transcripts such as poly(A), and RNase R has a preference for rRNAs (Cheng and Deutscher 2002).

An RNase II crystal structure has recently shed light on the catalytic activity and substrate specificity of RNR enzymes (Frazao et al. 2006; Zuo et al. 2006). RNase II folds into four domains comprising two N-terminal RNA-binding moieties, a central catalytic domain, and a C-terminal S1-like RNA binding region (Frazao et al. 2006; Zuo et al. 2006). The N- and C-terminal domains form a clamp atop the catalytic domain, which funnels the ssRNA substrate into a narrow channel that houses the active site. Although domain structure and sequence motifs are highly conserved among RNR family members, it is thought that differences in the clamp arrangement and thus RNA binding properties play an important role in regulating the activity on transcripts containing secondary structures.

Chloroplast RNR1 is inhibited by secondary structures when assayed *in vitro* (Perrin et al. 2004; Bollenbach et al. 2005). Therefore, it could participate in the processing of precursor RNAs, in particular 3' ends. Since mature transcripts often contain terminal stem-loops any degradative action of RNR1 would require prior endonucleolytic cleavage and polyadenylation, or recruitment of an RNA helicase. The latter tactic is employed by yeast mitochondrial Dss1, an RNase R homolog that digests secondary structures by complexing with a helicase. It should be noted that there is no PNPase in yeast mitochondria, thus Dss1 is the only exonuclease so far identified in that organelle (Dziembowski et al. 1998).

RNase II, RNase R, and PNPase, which represent the major exoribonuclease activities in *E. coli*, have significantly different substrate specificities and catalytic properties *in vitro* but share overlapping functions *in vivo*. In *Synechocystis*, there is a single RNase II/R homolog. In addition, PNPase functions as the only polyadenylation enzyme (in addition to its function in degradation). Accordingly, deletion of *Synechocystis* PNPase- or RNase II/R-encoding genes, unlike the situation in *E. coli* (Donovan and Kushner 1986), leads to inviability (Rott et al. 2003). Similarly, since there is no PNPase in yeast mitochondria, deletion of the RNase II/R homolog *DSS1* leads to mitochondrial dysfunction and eventually to loss of its genome (Dziembowski et al. 1998, 2003).

Plant chloroplast PNPase and RNR1 catalyze distinguishable reactions *in vivo*, but may functionally overlap. Repression of the *pnp1* gene, for example, leads to defects in mRNA and 23S rRNA 3' processing, but plants retaining only minimal amounts of chloroplast PNPase are viable and grow on soil (Walter et al. 2002). In contrast, *rnr1* null mutants are defective in rRNA 3' processing but not in mRNA 3' processing (Kishine et al. 2004; Bollenbach et al. 2005). RNR1 mutants are inviable on soil, owing to a dependence on RNR1 for chloroplast development in cotyledons, and perhaps an effect on mitochondrial mRNA metabolism (Perrin et al. 2004). On the other hand, *pnp1*/*rnr1* double null mutants have an embryo lethal phenotype (Bollenbach, Gutierrez, and Stern, unpublished data), suggesting either that these enzymes are redundant or additive for an essential processing or regulatory step(s).

2.2.3 Evidence for a 5' to 3' pathway

A major player in eukaryotic RNA decay is a 5' to 3' pathway catalyzed by the exonuclease Xrn1/Rat1. First described in *S. cerevisiae* and subsequently in animals (Newbury et al. 2006), Xrn1 is encoded by a small gene family in plants (Kastenmayer and Green 2000), with at least one member involved in miRNA metabolism (Souret et al. 2004). None of the family members, however, are known or suspected to be organelle-targeted.

It is therefore surprising that chloroplasts possess a 5' to 3' RNA degradation activity, which was revealed through the phenotypes of nuclear mutants affecting the stabilities of individual chloroplast transcripts (see Section 3.1.2). This suggests several possibilities: (1) one of the Xrn1-like proteins may be organelle-localized or dual targeted; (2) an organellar protein with Xrn1-like activity may exist but have little sequence homology; and/or (3) the apparent 5' to 3' RNA degradation maybe be a net activity, in fact catalyzed by a processive endonuclease.

Current literature best supports the concept of a net 5' to 3' pathway. Evidence for this comes from studies of endonuclease cleavage sites in the 3' UTRs of the *Chlamydomonas rbcL* and *atpB* mRNAs. When cleavage occurs, presumably as part of 3' end maturation (see Section 3.3), the downstream moiety is rapidly degraded (Stern and Kindle 1993). Subsequent studies showed that the degradation cannot be blocked using polyguanosine [poly(G)] or a stem-loop structure, which prevent exonuclease attack (Hicks et al. 2002). On the other hand, the 5' to 3' degradation found in RNA stability mutants can be blocked by poly(G), leaving open the possibility that chloroplasts have multiple 5' to 3' activities (Drager et al. 1998, 1999; Nickelsen et al. 1999).

If a vectorial endonuclease exists in chloroplasts, the best candidate would be an RNase E-like enzyme (Mackie 1998). As discussed in Section 1.1.2, however, its function in chloroplasts is still speculative. Furthermore, there is no evidence as yet that 5' to 3' pathway(s) occur in higher plant chloroplasts. Indeed, none of the plant nuclear mutants affecting cpRNA metabolism appear to mimic the RNA stability mutants of *Chlamydomonas* (see Section 4). Whether this is an artifact of the small number of mutants characterized to date or an evolutionary difference, remains to be established.

3 Polyadenylation

3.1 Historical perspective on polyadenylation

Polyadenylation is an important posttranscriptional modification of prokaryotic, eukaryotic and organellar RNA. In the cytoplasm and nucleus, the molecular mechanism of the addition of stable poly(A) tails to the 3' ends of most mRNAs and the importance of this process for translation initiation have been well established (Wickens et al. 1997; Dreyfus and Regnier 2002a; Edmonds 2002). In addition, transient polyadenylation was recently described for the yeast nucleus as part of an exosome-dependent RNA quality control mechanism (Lacava et al. 2005;

Vanacova et al. 2005; Wyers et al. 2005; Houseley et al. 2006). In bacteria, the major proteins involved in the polyadenylation-stimulated pathway have been identified and the relationship between polyadenylation and RNA decay has been characterized (Coburn and Mackie 1999). Polyadenylated RNA was first detected in the chloroplast more than 30 years ago (Haff and Bogorad 1976). Using hybridization experiments with cpDNA and ^{125}I-labeled RNA from maize seedlings, it was determined that about 6% of the poly(A)-containing RNA hybridized to cpDNA, and that the chloroplast poly(A) tracts averaged about 45 nucleotides in length.

Since polyadenylation is a phenomenon observed in almost all organisms, a major point is the assumption that a basal mechanism of polyadenylation-stimulated degradation of RNA was present in the last universal common ancestor of the three domains of life. During evolution, this basal mechanism was subjected to many modifications and variations that can be observed today in different organisms and organelles (Table 1) (Slomovic et al. 2006a). Moreover, different and perhaps conflicting biological functions for polyadenylation were acquired in several cases, such as transcript stabilization and translation initiation in the case of eukaryotic mRNA, and stimulation of turnover in the case of bacteria and organelles (Dreyfus and Regnier 2002a; Slomovic et al. 2006b).

The addition of a stable poly(A) tail to most nucleus-encoded mRNAs was first observed many years ago, and shown to occur following endonucleolytic cleavage in the 3' UTR, by a complex of several proteins providing enzymatic, RNA-binding and regulatory functions (Weiner 2005). Therefore, even though the first PAP was identified in *E. coli*, polyadenylation has long been considered a unique feature of eukaryotic cells and one of the major differences between eukaryotes and prokaryotes.

3.2 The polyadenylation-stimulated degradation pathway in bacteria

As mentioned above, even though the purification of *E. coli* PAP was reported many years ago, polyadenylation in bacteria was not studied extensively, perhaps because no biological role had been conceived (Sarkar 1997; Deutscher and Li 2001; Kushner 2004). However, attention was refocused on polyadenylation when it was discovered that mutations in *pcnB* (encoding PAP) resulted in a tenfold increase in accumulation of RNA I, which represses plasmid replication. These results suggested that polyadenylation targets RNA I for rapid degradation, in contrast to the stability and translational competence imparted by the stable poly(A) tails at the 3' ends of nuclear mRNA.

Considerable progress has subsequently been made in understanding bacterial RNA polyadenylation and degradation, mostly by analyzing *E. coli* (Deutscher 2006). The first step in RNA degradation is endonucleolytic cleavage, which is believed to be carried out mainly by RNase E or RNases J1 and J2 (see Section 1.1.3). In chloroplasts, CSP41a was also shown to be a key enzyme in endonucleolytic cleavage (see Section 1.1.1), thus chloroplasts may have two or even three endonucleases in the polyadenylation pathway.

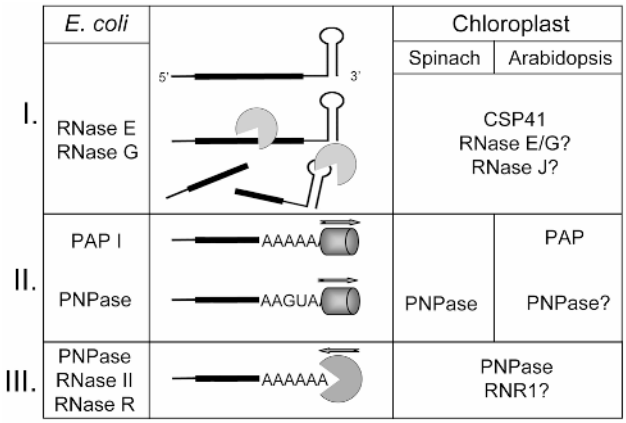


Fig. 3. A comparison of polyadenylation-stimulated RNA turnover pathways in *E. coli* and chloroplasts. The three stages of polyadenylation-stimulated RNA turnover are highlighted at left: endonucleolytic cleavage (I), polyadenylation (II), and exonucleolytic turnover (III).

In the second step, the cleavage product is polyadenylated and thus targeted for rapid exonucleolytic degradation (Fig. 3). In *E. coli*, polyadenylation is carried out mainly by a nucleotidyltransferase-type PAP (Ntr-PAP) producing homopolymeric poly(A) tails and to a certain extent by PNPase, which produces heteropolymeric poly(A)-rich tails containing all four nucleotides (Mohanty and Kushner 2000b). The protein Hfq, which resembles the eukaryotic Sm-like protein, was recently found to be involved in the modulation of polyadenylation activity between Ntr-PAP and PNPase (Mohanty et al. 2004; Folichon et al. 2005). The final step in the polyadenylation pathway is exonucleolytic degradation, which is performed by PNPase, RNase II, and RNase R in *E. coli* (Cheng and Deutscher 2005).

These findings along the way stimulated related research in other prokaryotes and in organelles. Indeed, evidence for the evolution and adaptation of the basic ancient polyadenylation-stimulated degradation process and the proteins involved have been revealed (Table 1) (Slomovic et al. 2006b), and the reader is referred to several recent reviews (Coburn and Mackie 1999; Grunberg-Manago 1999; Marujo et al. 2000; Deutscher and Li 2001; Dreyfus and Regnier 2002b; Kushner 2002, 2004; Condon 2003; Deutscher 2006).

3.3 PNPase as the major polyadenylating enzyme: variations from *E. coli*

Only limited studies have been carried out on Gram-positive bacteria. When *Streptomyces coelicolor* and *B. subtilis* transcripts were analyzed, heteropolymeric tails containing all four nucleotides were found, suggesting that PNPase and not Ntr-PAP is the major polyadenylating enzyme (Bralley and Jones 2002; Campos-Guillen et al. 2005). Accordingly, the sole Ntr proteins encoded by both these

Table 1. Similarities and differences between RNA polyadenylation systems among prokaryotes, chloroplasts, and eukaryotes.

	Prokaryotes			Chloroplast		Eukaryotes
	G-	G+	Cyano- bacteria	Plants	algae	Nucleus +Cytoplasm
	E. coli	*S. coe.* *B. sub.*	*Syn.*	Spinach, *Arabidopsis*	*Chlamydomonas*	Yeast Human
Endo.	E G	E J	E J	E CSP41 J	? J	?
Polyad- enylation	PAP I PNP	PNP PAP?	PNP	PNP PAP	PNP? PAP?	PAP TRAMP Exo.?
Exo.	PNP II R	PNP R	PNP R	PNP R	PNP R	3' → 5' Exo 5' → 3'
Poly(A)	Hom.	Het.	Het.	Het. Hom.	Hom.	Hom. Het.
Poly(A) RNA	Unstable					Stable +Unstable

Note: Within the bacteria, *E. coli* represents the Gram-negative (G-) and *Streptomyces coelicolor* (*S. coe.*) and *Bacillus subtilis* (*B. sub.*) the Gram positive (G+). Cyanobacteria are represented by *Synechocystis* (*Syn.*). Land plants are represented by spinach and *Arabidopsis* while algal data are from *Chlamydomonas*.
Symbols and abbreviations are: E, proteins homologous to RNase E or RNase G of *E. coli*; G, RNase G; PAP, poly(A) polymerase; PNP, polynucleotide phosphorylase; II and R, proteins homologous to RNase II and RNase R of *E. coli*; (?), Unknown or only based on prediction from genomic sequences. Hom., homopolymeric poly(A); Het., heteropolymeric poly(A)-rich. A gray background marks systems where both stable and unstable poly(A) tails are present.

organisms were active as Ntrs and not PAPs *in vitro* (Raynal et al. 1998; Sohlberg et al. 2003). Nevertheless, the analysis of PNPase-deficient *B. subtilis* revealed pronounced polyadenylation with homopolymeric poly(A) tails. This result suggested that *B. subtilis* has both PNPase and PAP-like activities, although the enzyme encoding the PAP-like activity has not been identified.

Cyanobacteria are related to the evolutionary ancestor of the chloroplast (Dyall et al. 2004), suggesting that an analysis of cyanobacterial RNA turnover could shed light on the ancient evolutionary form of the polyadenylation-stimulated pathway. Studies of *Synechocystis* revealed that mRNA, rRNA, tRNA and the single intron located at the $tRNA^{fMet}$ undergo polyadenylation (Rott et al. 2003), mirroring results for the same RNA classes in *E. coli* (Li et al. 1998), *Chlamydomonas* (Komine et al. 2000) and human mitochondria (Slomovic et al. 2005). The nature of the tails, which were poly(A)-rich and not homopolymeric, indicated that the polyadenylating enzyme is PNPase and not an Ntr. Therefore, PNPase is the major polyadenylating enzyme in cyanobacteria, spinach chloroplasts, and *Strep-*

tomyces. These results support the hypothesis that *E. coli,* other proteobacteria and *Arabidopsis* chloroplasts (see Section 2.4.2) acquired PAP relatively late in evolution through the conversion of a CCA-adding Ntr (Yue et al. 1996). Therefore, the RNA polyadenylation mechanism in cyanobacteria may represent a more ancient evolutionary state of the version found in *E. coli*.

3.4 Polyadenylation in the chloroplast

3.4.1 Discovery of heteropolymeric tails and relationship to degradation

Assuming that RNA metabolic pathways in the chloroplast were retained from its prokaryotic ancestor and following elucidation of the polyadenylation-degradation pathway in *E. coli*, the way was paved for dissecting this process in the chloroplast. RT-PCR analysis of oligo(dT)-primed cDNAs revealed polyadenylation in spinach chloroplasts (Kudla et al. 1996; Lisitsky et al. 1996). These studies revealed heteropolymeric, poly(A)-rich tails, the first observation of such tails in any organism. In addition, at the time of this discovery, there was still no explanation of how the heteropolymeric tails were formed. Nevertheless, heteropolymeric tails were later discovered in bacteria, archaea and human cells, as discussed above.

Several polyadenylation sites within the spinach *psbA* RNA matched endonucleolytic cleavage sites mapped by primer extension (Lisitsky et al. 1996). In addition, a polyadenylation site identified by RT-PCR in the spinach *petD* RNA was found to coincide with the cleavage site of a partially purified endoribonuclease when incubated with RNA resembling the *petD* transcript (Kudla et al. 1996). These results implied that the polyadenylation sites are produced by endonucleolytic cleavage of mature RNA and do not arise from polyadenylation of truncated molecules resulting from premature transcription termination (reviewed in Hayes et al. 1999; Schuster et al. 1999).

That polyadenylation stimulates degradation was observed by several biochemical and molecular approaches, as well as by experiments using the green alga *Chlamydomonas reinhardtii*. A DNA construct was engineered to express GFP mRNA and protein in *Chlamydomonas* chloroplasts such that the 3’ end poly(A) tail would be exposed after RNase P cleavage upstream of an ectopic *trnE* (Komine et al. 2002). Indeed, no GFP protein or polyadenylated *gfp* transcript could be detected in this strain. In contrast, the expression of GFP was relatively high in strains where the *gfp* mRNA either lacked a poly(A) tail or contained an arbitrary (A+U) tail (Komine et al. 2002). This result, together with those obtained using *in vitro* and lysed chloroplast assays demonstrated that polyadenylation-stimulated degradation in chloroplasts and bacteria were similar. Therefore, the next step was to identify the proteins responsible for initial endonucleolytic cleavage, polyadenylation and exoribonucleolytic degradation.

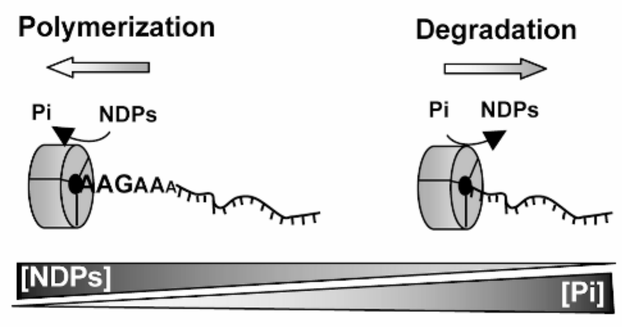


Fig. 4. PNPase acts as both a polymerase and a 3' to 5' exoribonuclease. PNPase is presented schematically as a homotrimer. When polymerizing RNA (left side), PNPase consumes nucleotide diphosphates (NDPs) and produces inorganic phosphate (P_i). When PNPase is an exoribonuclease and catalyzes RNA degradation (right side), it consumes P_i and produces NDPs. Because the equilibrium of this reaction lies close to unity, PNPase is exquisitely sensitive to P_i and NDP concentrations (grey wedges). Therefore, the reaction catalyzed by PNPase can theoretically be dictated by local concentrations of each substrate.

3.4.2 Different enzymes perform polyadenylation in spinach and Arabidopsis chloroplasts

Interestingly, species differences for polyadenylation enzymes were found in chloroplasts as they were in bacteria. In 2000, it was discovered that poly(A) tails in *pcnB* deletion strains of *E. coli* were heteropolymeric, very similar to those characterized before in spinach chloroplasts, and that these heteropolymeric tails were produced by PNPase (Mohanty and Kushner 2000a). This meant that PNPase was likely responsible for polyadenylation in spinach chloroplasts and indeed, purification of PAP activity from spinach chloroplasts yielded only PNPase, whose activity was the same as the stromal extracts from which it was isolated (Yehudai-Resheff et al. 2001).

How can one enzyme perform the opposing activities of polyadenylation and degradation? Biochemical and molecular analyses revealed that the directionality of the nearly freely reversible reaction that chloroplast PNPase catalyzes is directly influenced by the P_i/NDP ratio (Yehudai-Resheff et al. 2001, 2003; Bollenbach et al. 2004). This suggests that PNPase activity may be shifted towards net exonucleolytic or polymerization activities by shifting concentrations of its substrates (Fig. 4).

A different situation exists in *Arabidopsis* chloroplasts where as in *E. coli*, an Ntr-like PAP may be responsible for polyadenylation (Fig. 3). This is because the tails identified so far in *Arabidopsis* chloroplasts are virtually homopolymeric (our unpublished results). Moreover, several putative chloroplast- and mitochondrially-targeted PAPs were identified bioinformatically in the *Arabidopsis* genome (Martin and Keller 2004). If one or more of these PAPs can be confirmed experimentally to be chloroplast-localized and to act as a PAP rather than an Ntr, this would suggest that the conversion of Ntr to PAP occurred independently in the evolution of *E. coli* and *Arabidopsis* chloroplasts. The third observation suggestive of PAP activity in *Arabidopsis* chloroplasts came from the analysis of a transgenic line in which the amount of PNPase was significantly reduced, but chloroplast polyadenylation appeared to be undiminished or even increase (Walter et al. 2002).

Together, these observations show that while PNPase performs polyadenylation in spinach chloroplasts and in *Synechocystis*, PAP seems to be responsible for this process in *Arabidopsis* chloroplasts. This suggests that chloroplast lineages containing PAP vs. Ntr may have split relatively recently in evolutionary terms.

4 RNA maturation

4.1 5' end maturation

4.1.1 5' ends can be processed or primary transcripts

Chloroplast mRNAs are not capped but instead accumulate as unprocessed primary transcripts or processed transcripts, which are characterized by a 5' di- or triphosphate, or by a 5' hydroxyl group, respectively. 5' phosphorylated RNAs are cappable by GDP and guanylyltransferase, whereas hydroxylated 5' ends are not. In angiosperm chloroplasts, many RNAs accumulate both in primary and processed forms, whereas no cappable chloroplast RNAs have been detected in *Chlamydomonas*, suggesting that all transcripts are 5' processed. Although 5' processing sites and the mode of processing have been identified for a number of chloroplast RNAs, the enzymes that catalyze these reactions have not.

Chloroplast RNA 5' processing can result in differential translation efficiencies, as exemplified by tobacco *atpB*, *atpH*, *psbB*, and *rbcL*, which are processed within their 5' UTRs and accumulate in multiple forms (Tanaka et al. 1987; Orozco et al. 1990; Kapoor et al. 1997; Miyagi et al. 1998; Serino and Maliga 1998). *In vitro* assays suggested that translation efficiencies of unprocessed and processed tobacco *rbcL* and *atpH* 5' UTRs were comparable, while processing of *atpB* and *psbB* 5' UTRs resulted in enhanced translation efficiencies (Yukawa et al. 2006). In an extreme case, one of five spinach *atpB* transcripts, whose 5' end mapped to the start of the coding region, was associated with crude polysomes (Bennett et al. 1990). This variation is likely to reflect species differences, in particular *cis* elements in the 5' UTRs.

RNA processing in *Chlamydomonas* chloroplasts is also linked to translation when two different 5' ends are present. For example, mutagenesis experiments with *psbA* and *psbD* have suggested that only the shorter of the two transcripts that accumulate for each gene is competent for translation (Bruick and Mayfield 1998; Nickelsen et al. 1999). In at least one case, the differences in translation efficiency have been correlated with the presence of sequence elements in the 5' UTR that are present in the longer transcript, but not in the shorter one (Bruick and Mayfield 1998), while the causal relationship between processing and translation in other cases is not as clear-cut (Yukawa et al. 2006). This type of processing-dependent regulation is also true for 5' ends generated by intercistronic processing, as described below.

4.2 Intercistronic processing

Plastid-encoded genes are often clustered into transcription units, reflecting their post-endosymbiotic assembly from different cyanobacterial genes and operons (Douglas 1998, 1999). Typical transcript patterns from these regions are complex, the result of extensive posttranscriptional processing including 5' and 3' maturation, intercistronic cleavages, and splicing, which are catalyzed by nucleus-encoded enzymes and are regulated by nucleus-encoded proteins (Barkan and Goldschmidt-Clermont 2000; Nickelsen 2003).

4.2.1 Clusters encoding mRNAs

The *psbB* gene cluster has long been a paradigm for studying the processing of plastid transcription units (Barkan 1988; Westhoff and Herrmann 1988). This cluster encodes five thylakoid membrane proteins, three of which are PSII components (*psbB*, *psbT*, *psbH*) and two of which are components of the cytochrome b_6/f complex (*petB*, *petD*).

Significant evidence suggests that intercistronic processing of the *psbB* gene cluster is required for efficient translation. *hcf107* is an *Arabidopsis* mutant impaired in *psbH* 5' processing, which results in a decrease in accumulation of monocistronic *psbH* and therefore in a decrease in the PsbH protein (Felder et al. 2001). This is thought to arise because the cleavage at position -45 of the *psbH* 5' UTR is required to alleviate inhibition by an intramolecular base pairing interaction that obscures the ribosome binding site. Similarly, the maize *crp1* mutant is impaired in cytochrome b_6/f complex accumulation, which is thought to result from the masking of the *petD* ribosome binding site, which requires endonucleolytic cleavage and formation of a monocistronic *petD* RNA to alleviate an intramolecular base pairing interaction (Barkan et al. 1994). On the other hand, tobacco and *Arabidopsis* chloroplasts do not accumulate monocistronic *petD* RNA and therefore do not require this same type of processing for translation, even though the *petB-petD* intergenic spacer contains elements important for translation (Monde et al. 2000a).

Although not affected in translation initiation, a third mutant of note that affects *psbB* operon processing is *Arabidopsis hcf152*, which is defective in *petB* intron splicing and therefore in cytochrome b_6/f accumulation (Meierhoff et al. 2003). Although the endoribonucleases responsible for intercistronic cleavage and splicing have not been identified, *HCF107*, *CRP1*, and *HCF152* each encode TPR/PPR family proteins (see Section 4.3), suggesting that this abundant class of proteins plays an important role in regulating the processing of polycistronic RNAs in the chloroplast. Further supporting this conclusion is a recent report showing that a *Physcomitrella* PPR protein is required both for intercistronic cleavage between *clpP* and 5'-*rps12*, and for *clpP* splicing (Hattori et al. 2007).

A highly regulated chloroplast gene cluster is the *ndhH-D* operon. This operon encodes, in order, *ndhH*, *ndhA*, *ndhI*, *ndhG*, *ndhE*, *psaC*, and *ndhD*. The *ndh* genes encode components of the low abundance NADH dehydrogenase complex, and *psaC* encodes subunit VII of photosystem I. Despite being co-transcribed, the *psaC* message accumulates two orders of magnitude higher than the *ndh* messages (Meurer et al. 1996). In leek and barley, the *psaC-ndhD* dicistronic intermediate is cleaved within the *ndhD* coding sequence, which provides monocistronic *psaC* with a stabilizing 3' UTR and yields a non-translatable monocistronic *ndhD* (del Campo et al. 2002, 2006). Alternative *psaC-ndhD* intergenic cleavages produce translationally competent *ndhD* at low levels, but only from dicistronic messages in which C to U editing has restored the *ndhD* start codon (Hirose and Sugiura 1997; del Campo et al. 2002). *In vitro* evidence from tobacco translation extracts suggested that the *psaC-ndhD* dicistronic RNA is not translationally competent, and that production of monocistronic RNAs is required to alleviate a base pairing interaction between the *ndhD* 3' UTR and an 8 nt element contained within the *psaC* coding region, thus allowing translation to occur (Hirose and Sugiura 1997). Mutations of the negative control element destabilized this base pairing and resulted in the translation of *ndhD* from the dicistronic RNA. This highly regulated system apparently ensures that processing and accumulation of *ndhD* does not exceed that of other Ndh complex subunits, while still allowing the *psaC* message, and PSI subunit VII, to accumulate to high levels.

4.2.2 The chloroplast rrn operon

Chloroplast rRNA genes resemble those of bacteria, in that their coding sequences are conserved and that they are co-transcribed as part of an operon with the gene order 16S-23S-4.5S-5S. The operon also encodes two tRNAs within the 16S-23S spacer, and is flanked by tRNA genes. Chloroplast ribosome biogenesis requires considerable processing and maturation of rRNAs, which requires both endo- and exoribonuclease steps. The primary *rrn* transcript is cleaved endonucleolytically by an unidentified enzyme(s), which releases pre-tRNAs and pre-rRNAs. Pre-tRNAs are matured by chloroplast homologs of RNase P and RNase Z at their 5' and 3' ends, respectively (Wang et al. 1988; Schiffer et al. 2002). The pre-16S and 5S RNAs differ considerably from their bacterial counterparts in that they are not processed close to their mature termini and therefore accumulate long 3' tails, which require 3' to 5' exonucleolytic processing by RNR1 and/or PNPase (Yamamoto et al. 2000; Walter et al. 2002; Bollenbach et al. 2005).

The 23S rRNA in plants appears to co-migrate with the *E. coli* 23S rRNA under non-denaturing conditions, but migrates as smaller RNAs under denaturing conditions due to cleavage at the so-called "hidden breaks" (Leaver 1973). The 4.5S RNA, which is unique to angiosperms, is homologous to the bacterial 23S rRNA 3' end, and is separated from the remainder of the 23S sequence by a 100 nt internal transcribed spacer (ITS). The 23S-4.5S processing intermediate undergoes 3' maturation prior to cleavage at the 4.5S 5' end, in a series of steps that requires prior assembly into pre-50S ribosomal subunits, as evidenced by the accumulation of this transcript in mutants defective in both rRNA 3' processing and ribosome assembly (Bellaoui et al. 2003; Bisanz et al. 2003; Bellaoui and Gruissem 2004; Bollenbach et al. 2005). 23S rRNA then undergoes a two-step 3' maturation that in *Arabidopsis* requires both PNPase and RNR1 (Walter et al. 2002; Bollenbach et al. 2005), but appears to be PNPase-independent in *Chlamydomonas* (Yehudai-Resheff et al. 2007). The translational consequences of a failure to remove the 23S ITS in plants is unknown and may be phenotypically silent as it is in bacteria (Kordes et al. 1994; Gregory et al. 1996; Mattatall and Sanderson 1998). On the other hand, the *Chlamydomonas ac20* mutant, which accumulates unspliced 23S rRNA and fewer mature ribosomes, fails to grow photoautotropically (Holloway and Herrin 1998).

4.3 3' end maturation

The 3' IRs of bacterial mRNAs promote transcript stability and can act as rho-independent transcription terminators. In chloroplasts, transcription termination is not influenced by 3' IRs and is probably stochastic (Stern and Gruissem 1987, 1989). Therefore, chloroplast mRNAs require 3' processing for maturation by processive 3' to 5' exoribonucleases (Stern and Gruissem 1987; Rott et al. 1996). 3' end maturation and 3' IR function has been studied in detail in *Chlamydomonas* using the *atpB* mRNA as a model. Termination of *atpB* transcription by its 3' IR is less than 50% efficient (Rott et al. 1996) and the resultant heterogeneous pre-mRNAs undergo two-step processing that begins with cleavage at a specific endonucleolytic cleavage site (ECS), and is completed by 3' to 5' exonucleolytic trimming (Stern and Kindle 1993), and may involve polyadenylation (Komine et al. 2000). Recent analysis of the *atpB* and *rbcL* 3' IR and ECS, together referred to as the 3' processing determinant (PD), suggested that these elements contain a significant amount of redundancy, since deletion of one or the other *cis*-element did not cause changes in *atpB* maturation (Rymarquis et al. 2006b). Redundancy in 3' PDs may be fairly common. For example, the *Chlamydomonas* chloroplast *petA* gene has at least ten possible mature 3' termini (Jiao et al. 2004).

Genetic screens have identified at least two nuclear genes important to chloroplast 3' RNA processing, *CRP3* and *MCD4*. The *crp3* mutant was isolated as a suppressor of a chloroplast *atpB* 3' IR deletion mutant, and was later found to affect the 3' maturation of several chloroplast-encoded RNAs (Levy et al. 1997, 1999). The *mcd4* mutant, which has numerous chloroplast 3' processing defects, is described in Section 4.2. The genes encoding CRP3 and MCD4 have not been cloned, but evidence suggests that they either represent endoribonucleases or RNA-binding proteins that guide ribonucleases to the ECS.

PNPase has been shown to be important for mRNA processing in *Arabidopsis*, since plants in which PNPase expression was inhibited by co-suppression were defective in *rbcL* and *psbA* 3' maturation, and accumulated RNAs with multiple 3' ends. Unlike the case with *Chlamydomonas atpB*, however, these transcripts were not differentially polysome associated versus their processed counterparts (Walter et al. 2002). *Chlamydomonas* cells nearly lacking PNPase due to RNAi suppression, however, accumulated apparently normal chloroplast mRNAs, suggesting a redundancy in this organism (Yehudai-Resheff et al. 2007).

4.4 Non-coding RNAs

Antisense RNA (asRNA)-mediated gene regulation occurs widely in prokaryotes and eukaryotes and bacteria express both *cis*- and *trans*-encoded antisense transcripts (reviewed in Gottesman 2004; Storz et al. 2005). For example, accumulation of *Synechocystis isiA* was shown to correlate inversely with the *cis*-encoded asRNA *isiR*, and it was suggested that the *isiA*-*isiR* duplex could be targeted for turnover by a dsRNA-specific RNase, such as RNase III (Duhring et al. 2006). The *Arabidopsis* nuclear genome encodes two RNase III homologs with putative chloroplast transit peptides at the loci At4g37510 and At3g20420.

Because posttranscriptional regulation is important in chloroplasts, it stands to reason that antisense-mediated mechanisms may operate in these organelles, although a role for noncoding RNAs (ncRNAs) remains to be clearly established. The tobacco chloroplast-encoded *sprA* gene (Vera and Sugiura 1994) encodes a *trans*-encoded RNA that was hypothesized to control 16S rRNA 5' maturation, but this function could not be confirmed by further experimentation with transgenic plants (Sugita et al. 1997). More recently, a search for chloroplast-encoded ncRNAs in tobacco identified several short sequences including two *cis*-asRNAs, Ntr-5 and Ntr-7, which are complementary to *atpE* and the *rps16* intron, respectively (Lung et al. 2006). Thus chloroplasts, like their prokaryotic ancestors, may encode functional asRNAs.

Evidence that asRNAs can regulate their targets in chloroplasts is currently restricted to transgenic contexts. For example, fortuitous expression of a synthetic asRNA following a chloroplast genome rearrangement in *Chlamydomonas* resulted in the stabilization of an otherwise unstable polyadenylated *atpB* transcript (Nishimura et al. 2004). In another case, expression of asRNAs decreased the efficiency of sense RNA editing in tobacco chloroplasts (Hegeman et al. 2005). Thus, chloroplasts have the potential to utilize natural asRNAs for gene regulation.

5 Regulatory factors

Generalized screens have led to identification of cpRNA mutants. In *Chlamydomonas*, mutants were obtained by isolating colonies unable to grow on minimal medium (acetate-requiring). These nonphotosynthetic mutants affect all stages of gene expression, as well as metabolic functions (Harris 1989; Rochaix 1995). The analogous screens in higher plants are seedling lethality in maize and a sucrose requirement in *Arabidopsis* (Barkan 1998; Stern et al. 2004). These plants display chlorotic or ivory phenotypes and if blocked in photosynthetic electron transport, high chlorophyll fluorescence (hcf). The hcf screen has also been used in *Chlamydomonas*, simplified by a video imaging approach (Bennoun and Béal 1997). While some of these mutants have turned out to affect ribonucleases, as discussed above, most remain uncloned or encode regulatory proteins. In this section, we discuss mutant characteristics and the PPR/TPR protein families, which are emerging as key regulators of organellar RNA metabolism.

5.1 Mutations affecting single chloroplast loci

A mutant class essentially unique to *Chlamydomonas* is gene-specific RNA stability mutants. These recessive mutants lack factors that stabilize certain transcripts, generally against 5' to 3' degradation. Known targets include *petA*, *psbB-psbT*, *petD*, *psbC*, *atpB*, and *psbD* (Barkan and Goldschmidt-Clermont 2000). The specificity of such mutants is somewhat presumptive, since in only one case was each chloroplast transcript checked in the mutant background; a microarray analysis of the *petD* mutant *mcd1* confirmed its specificity (Erickson et al. 2005).

Several *Chlamydomonas* RNA stability factors have been cloned. *MCA1*, which stabilizes *petA* mRNA, encodes a pentatricopeptide repeat (PPR) protein (Lown et al. 2001), a motif which is discussed below. Some nomenclature confusion exists because *MCA1* was previously attributed to mitochondrial carbonic anhydrase (Eriksson et al. 1998). The *psbB/T* and *psbD* stability factors are encoded by *MBB1* and *MBD1/NAC2*, respectively, which both feature tetratricopeptide (TPR) repeats (Boudreau et al. 2000; Vaistij et al. 2000), another motif that is discussed below. The *petD* stability factor MCD1, however, possesses neither of these motifs nor any recognizable domains (Murakami et al. 2005). From just this small sample, it appears that even within *Chlamydomonas* various solutions have arisen to protect transcripts, and possibly to promote their translation.

While no higher plant mutants are fully analogous to the *Chlamydomonas* RNA stability mutants, in at least one case an orthologous gene has been found. The *Arabidopsis* mutant *hcf107* (Felder et al. 2001) has defects in the processing of *psbH* mRNA (see Section 3.2.1). The Hcf107 protein is homologous to Mbb1 (Sane et al. 2005), and the slightly different phenotypic consequence of its absence can be ascribed to the different gene arrangements in the respective chloroplast genomes. Other homologous pairs of genes have been identified for chloroplast biogenesis, such as *TAB2* (Dauvillee et al. 2003; Barneche et al. 2006), which functions gene-specifically in translation initiation in *Chlamydomonas* but appears to have multiple targets in *Arabidopsis*. This may suggest that evolution of these proteins has been more closely constrained by the RNA target, rather than interaction with cellular machinery such as ribosomes or nucleases. Otherwise, one might anticipate common motifs accompanied by a "gene specificity domain."

Several higher plant mutants, like *hcf107,* appear to have a single primary target. For example, *Arabidopsis HCF152* encodes a PPR protein that also affects *psbH* maturation (Meierhoff et al. 2003; Nakamura et al. 2003). Hcf152 has been reported to have structural similarity to Crp1, a maize protein whose primary target is cleavage between *petB* and *petD*, with a concomitant or secondary effect on PetA translation (Barkan et al. 1994). Because *psbH* and *petB-D* are in the same gene cluster, the functions of Crp1 and Hcf152 are in a sense related. In turn, cloning of Crp1 revealed sequence similarity to at least two fungal regulators of mitochondrial translation (Fisk et al. 1999), which is most related to its maize function for PetA. Crp1 was also reported to share homology with p67, an RNA-binding PPR protein from radish chloroplasts (Lahmy et al. 2000). The *Arabidopsis* homolog of p67 (At4g16390) and Hcf152 (At3g09650), however, are minimally related, making the situation somewhat ambiguous, and pointing to the difficulty of assigning correct homologies in large, degenerate gene families.

One nearly universal feature of the regulatory factors described above is that they are found in high molecular weight complexes. These have most often been revealed by gel filtration, and tend to show broad peaks in the 350 kDa – 600 kDa range, such as for Nac2, Crp1, and Mbb1 (cited above). A major unanswered question is the composition of these complexes, apart from the presence of the cognizant RNA, which has been detected in some cases (e.g. *psbD* mRNA in the Nac2 complex). One difficulty is their low abundance, which is a consequence of their single or dual-gene specificity. However, affinity methods are likely to lead to purification in the near future. The reader is also directed to the chapter by Schmitz-Linneweber and Barkan for a somewhat better-developed knowledge of chloroplast splicing complexes.

A final point regarding gene-specific regulators is the implication of co-evolution of the regulatory factor and the gene sequence. Evidence for this includes the lack of conservation between 5' UTRs of different chloroplast mRNAs, the targets of the vast majority of the regulators. Furthermore, small sequence motifs, when mutated, phenocopy the cognizant nuclear mutations. For example, 4-nt changes in the 5' UTR of the *Chlamydomonas petD* mRNA destabilize the transcript, phenocopying the *mcd1* mutation (Higgs et al. 1999); similar results were obtained for *psbD* (Nickelsen et al. 1999). Interestingly, the *petD* regulatory motifs tend to be highly conserved among *Chlamydomonas* species whose cpDNAs are otherwise highly divergent (Kramzar et al. 2006). This argues in favor of constraints on the *cis* elements in a given gene, most likely because of their interactions with specific motifs in the regulatory proteins.

Because transcript destabilization for these genes leads to a loss of photosynthetic capability, genetic screens can be carried out for restoration of photosynthetic growth. In the case of *psbD*, three unlinked nuclear suppressors were obtained which restored *psbD* expression, but did not affect *psbA* expression (Nickelsen 2000). For *petD* three suppressors were also obtained, again in unlinked nuclear loci. Most surprisingly, the restoration of *petD* expression was accompanied by pleiotropic effects on other chloroplast mRNAs (Rymarquis et al. 2006a), which are described in more detail in the next section. Direct screens for suppressors of the mutated nuclear factors have been less successful. A suppressor of an *mcd1* mutant was isolated and found to be allele-specific and semi-dominant, however it was revealed to encode a suppressor tRNA, rather than a new effector of *petD* expression (Murakami et al. 2005). In summary, studies of genetic interactions with gene-specific regulators is scattered, and understanding the basis of the specific interactions awaits knowledge of complex components and suitable *in vitro* systems.

5.2 Pleiotropic mutations

In principle, mutation of general RNA regulators should cause pleiotropic phenotypes, much as the maize nuclear mutant *crs1*, which is affected in the splicing of many chloroplast introns (Jenkins et al. 1997). Indeed, the *Arabidopsis rnr1* and PNP- lines have pleiotropic defects (Walter et al. 2002; Bollenbach et al. 2005). Another class of pleiotropic mutations affects mRNAs, and is exemplified by *mcd3*, *mcd4* and *mcd5*, which were isolated as suppressors of *petD* 5' UTR mutations as described above (Rymarquis et al. 2006a). Most pleiotropic were *mcd3* and *mcd4*, which accumulated numerous transcripts with extended 3' ends, particularly in gene clusters. This implicates the genes in 3' end formation, which is counterintuitive since they were isolated as suppressors of 5' UTR mutations causing RNA instability. Some resolution of this dilemma may be offered by the fact that 5' ends of chloroplast transcripts are often generated by endonucleolytic processing, which also occurs at the 3' end, as exemplified by *Chlamydomonas atpB* and *rbcL*, among others (Blowers et al. 1993; Stern and Kindle 1993).

5.3 The PPR/TPR protein superfamilies

As noted above, at least some of the RNA regulators are members of the PPR and TPR protein classes. While the *Chlamydomonas* nuclear only encodes about two dozen TPRs and less than ten PPR family members, the protein class has been highly amplified in flowering plants. Indeed, *Arabidopsis* was found to encode 441 PPR proteins, many of which appear to encode essential functions in mitochondria and chloroplasts (Lurin et al. 2004). Why the families would be expanded in plants vs. *Chlamydomonas* is not yet known, however it may related to the lack of RNA editing in *Chlamydomonas*, and also to the extreme simplicity of its mitochondrial genome relative to that of the flowering plants. As components of multiprotein RNA processing complexes, TPR/PPR proteins are very likely to interact with catalytically active complex members. It could be that chloroplast RNA processing complexes, or processosomes, will be analogous to the bacterial degradosome, which contains both ribonucleases and scaffolding factors, in particular the C-terminal part of ribonuclease E (Vanzo et al. 1998). The degradosome, however, is not gene-specific, so the analogy is likely to be imperfect.

6 Conclusions

The last few years has seen a number of advances in the understanding of cpRNA processing and turnover, including a broader knowledge of how chloroplast polyadenylation has evolved and its extant diversity in the broader organismal context, the identification of new enzymatic and regulatory components of RNA metabolizing pathways, and the identification of chloroplast-encoded ncRNAs.

Much of the understanding of the polyadenylation pathway has been underpinned by comparative genomics, which permitted correlations between polyadenylation mechanisms and its enzymatic machinery (Slomovic et al. 2006a). Candidate gene approaches have been key to establishing the basic enzymatic framework of the cpRNA processing and turnover, setting the stage for a phase in which regulation of their activities and specificities will be investigated. Whether these enzymes turn out to be regulated by metabolites, as in the case of CSP41a and PNPase (Yehudai-Resheff et al. 2001; Bollenbach and Stern 2003a; Bollenbach et al. 2003), or by plant-specific proteins such as members of the PPR family(Lurin et al. 2004), remains to be seen. Answering these questions will likely take a multidisciplinary strategy, combining forward and reverse genetics, biochemistry and enzymology.

Finally, we note the timely identification of antisense RNA-mediated gene regulation in *Synechocystis* (Duhring et al. 2006), and the recent identification of small, chloroplast-encoded ncRNAs (Lung et al. 2006). Whether these small RNAs turn out to be regulatory transcripts remains to be determined, as we move from cataloging them to determining the mechanisms by which they might regulate chloroplast gene expression.

Acknowledgements

This work was supported by Department of Energy Biosciences grant DE-FG02-90ER20015 to DBS, Israel Science Foundation grant 266/05 to GS, United States - Israel Binational Science Foundation (BSF) awards 2001090 and 2005184 and United States - Israel Binational Agriculture Research and Development Fund BARD award IS-3605-04CR to DBS and GS.

References

Aloy P, Ciccarelli FD, Leutwein C, Gavin AC, Superti-Furga G, Bork P, Bottcher B, Russell RB (2002) A complex prediction: three-dimensional model of the yeast exosome. EMBO Rep 3:628-635

Baginsky S, Shteiman-Kotler A, Liveanu V, Yehudai-Resheff S, Bellaoui M, Settlage RE, Shabanowitz J, Hunt DF, Schuster G, Gruissem W (2001) Chloroplast PNPase exists as a homo-multimer enzyme complex that is distinct from the *Escherichia coli* degradosome. RNA 7:1464-1475

Baker ME, Grundy WN, Elkan CP (1998) Spinach CSP41, an mRNA-binding protein and ribonuclease, is homologous to nucleotide-sugar epimerases and hydroxysteriod dehydrogenases. Biochem Biophys Res Comm 248:250-254

Barkan A (1988) Proteins encoded by a complex chloroplast transcription unit are each translated from both monocistronic and polycistronic RNAs. EMBO J 7:2637-2644

Barkan A (1998) Approaches to investigating nuclear genes that function in chloroplast biogenesis in land plants. Meths Enzymol 297:38-57

Barkan A, Goldschmidt-Clermont M (2000) Participation of nuclear genes in chloroplast gene expression. Biochimie 82:559-572

Barkan A, Walker M, Nolasco M, Johnson D (1994) A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. EMBO J 13:3170-3181

Barneche F, Winter V, Crevecoeur M, Rochaix JD (2006) ATAB2 is a novel factor in the signalling pathway of light-controlled synthesis of photosystem proteins. EMBO J 25:5907-5918

Bellaoui M, Gruissem W (2004) Altered expression of the *Arabidopsis* ortholog of DCL affects normal plant development. Planta 219:819-826

Bellaoui M, Keddie JS, Gruissem W (2003) DCL is a plant-specific protein required for plastid ribosomal RNA processing and embryo development. Plant Mol Biol 53:531-543

Bennett DC, Rogers SA, Chen LJ, Orozco M (1990) A primary transcript in spinach chloroplasts that completely lacks a 5' untranslated leader region. Plant Mol Biol 15:111-120

Bennoun P, Béal D (1997) Screening algal mutant colonies with altered thylakoid electrochemical gradient through fluorescence and delayed luminescence digital imaging. Photosynthesis Res 51:161-165

Bisanz C, Begot L, Carol P, Perez P, Bligny M, Pesey H, Gallois JL, Lerbs-Mache S, Mache R (2003) The *Arabidopsis* nuclear *DAL* gene encodes a chloroplast protein which is required for the maturation of the plastid ribosomal RNAs and is essential for chloroplast differentiation. Plant Mol Biol 51:651-663

Blowers AD, Klein U, Ellmore GS, Bogorad L (1993) Functional *in vivo* analyses of the 3' flanking sequences of the *Chlamydomonas* chloroplast *rbcL* and *psaB* genes. Mol Gen Genet 238:339-349

Blum E, Carpousis AJ, Higgins CF (1999) Polyadenylation promotes degradation of 3'-structured RNA by the *Escherichia coli* mRNA degradosome *in vitro*. J Biol Chem 274:4009-4016

Bollenbach TJ, Lange H, Gutierrez R, Erhardt M, Stern DB, Gagliardi D (2005) RNR1, a 3'-5' exoribonuclease belonging to the RNR superfamily, catalyzes 3' maturation of chloroplast ribosomal RNAs in *Arabidopsis thaliana*. Nucleic Acids Res 33:2751-2763

Bollenbach TJ, Schuster G, Stern DB (2004) Cooperation of endo- and exoribonucleases in chloroplast mRNA turnover. Prog Nucleic Acid Res Mol Biol 78:305-337

Bollenbach TJ, Stern DB (2003a) Divalent metal-dependent catalysis and cleavage specificity of CSP41, a chloroplast endoribonuclease belonging to the short chain dehydrogenase/reductase superfamily. Nucleic Acids Res 31:4317-4325

Bollenbach TJ, Stern DB (2003b) Secondary structures common to chloroplast mRNA 3'-untranslated regions direct cleavage by CSP41, an endoribonuclease belonging to the short chain dehydrogenase/reductase superfamily. J Biol Chem 278:25832-25838

Bollenbach TJ, Tatman DA, Stern DB (2003) CSP41a, a multifunctional RNA-binding protein, initiates mRNA turnover in tobacco chloroplasts. Plant J 36:842-852

Boudreau E, Nickelsen J, Lemaire SD, Ossenbuhl F, Rochaix J-D (2000) The *Nac2* gene of *Chlamydomonas reinhardtii* encodes a chloroplast TPR protein involved in *psbD* mRNA stability, processing and/or translation. EMBO J 19:3366-3376

Bralley P, Jones GH (2002) cDNA cloning confirms the polyadenylation of RNA decay intermediates in *Streptomyces coelicolor*. Microbiology 148:1421-1425

Britton RA, Wen T, Schaefer L, Pellegrini O, Uicker WC, Mathy N, Tobin C, Daou R, Szyk J, Condon C (2007) Maturation of the 5' end of *Bacillus subtilis* 16S rRNA by the essential ribonuclease YkqC/RNase J1. Mol Microbiol 63:127-138

Bruick RK, Mayfield SP (1998) Processing of the *psbA* 5' untranslated region in *Chlamydomonas reinhardtii* depends upon factors mediating ribosome association. J Cell Biol 143:1145-1153

Buttner K, Wenig K, Hopfner KP (2005) Structural framework for the mechanism of archaeal exosomes in RNA processing. Mol Cell 20:461-471

Callaghan AJ, Marcaida MJ, Stead JA, McDowall KJ, Scott WG, Luisi BF (2005) Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover. Nature 437:1187-1191

Campos-Guillen J, Bralley P, Jones GH, Bechhofer DH, Olmedo-Alvarez G (2005) Addition of poly(A) and heteropolymeric 3' ends in *Bacillus subtilis* wild-type and polynucleotide phosphorylase-deficient strains. J Bacteriol 187:4698-4706

Chen H, Stern DB (1991) Specific ribonuclease activities in spinach chloroplasts promote mRNA maturation and degradation. J Biol Chem 266:24205-24211

Chen HW, Rainey RN, Balatoni CE, Dawson DW, Troke JJ, Wasiak S, Hong J, McBride H, Koehler CM, Teitell MA, French SW (2006) Mammalian polynucleotide phosphorylase is an intermembrane space RNase that maintains mitochondrial homeostasis. Mol Cell Biol 26:8475-8487

Cheng ZF, Deutscher MP (2002) Purification and characterization of the *Escherichia coli* exoribonuclease RNase R. Comparison with RNase II. J Biol Chem 277:21624-21629

Cheng ZF, Deutscher MP (2005) An important role for RNase R in mRNA decay. Mol Cell 17:313-318

Clements MO, Eriksson S, Thompson A, Lucchini S, Hinton JC, Normark S, Rhen M (2002) Polynucleotide phosphorylase is a global regulator of virulence and persistency in *Salmonella enterica*. Proc Natl Acad Sci USA 99:8784-8789

Coburn GA, Mackie GA (1999) Degradation of mRNA in *Escherichia coli*: an old problem with some new twists. Prog Nucleic Acid Res Mol Biol 62:55-108

Coburn GA, Miao X, Briant DJ, Mackie GA (1999) Reconstitution of a minimal RNA degradosome demonstrates functional coordination between a 3' exonuclease and a DEAD-box RNA helicase. Genes Dev 13:2594-2603

Cohen SN, McDowall KJ (1997) RNase E: still a wonderfully mysterious enzyme. Mol Microbiol 23:1099-1106

Condon C (2003) RNA processing and degradation in *Bacillus subtilis*. Microbiol Mol Biol Rev 67:157-174

Dauvillee D, Stampacchia O, Girard-Bascou J, Rochaix JD (2003) Tab2 is a novel conserved RNA binding protein required for translation of the chloroplast *psaB* mRNA. EMBO J 22:6378-6388

de la Sierra-Gallay IL, Pellegrini O, Condon C (2005) Structural basis for substrate binding, cleavage and allostery in the tRNA maturase RNase Z. Nature 433:657-661

del Campo EM, Sabater B, Martin M (2006) Characterization of the 5'- and 3'-ends of mRNAs of *ndhH*, *ndhA* and *ndhI* genes of the plastid *ndhH-D* operon. Biochimie 88:347-357

del Campo EM, Sabater B, Martin M (2002) Post-transcriptional control of chloroplast gene expression. Accumulation of stable *psaC* mRNA is due to downstream RNA cleavages in the *ndhD* gene. J Biol Chem 277:36457-36464

Deng XW, Gruissem W (1987) Control of plastid gene expression during development: the limited role of transcriptional regulation. Cell 49:379-387

Deutscher MP (2006) Degradation of RNA in bacteria: comparison of mRNA and stable RNA. Nucleic Acids Res 34:659-666

Deutscher MP, Li Z (2001) Exoribonucleases and their multiple roles in RNA metabolism. Prog Nucleic Acid Res Mol Biol 66:67-105

Dhingra A, Bies DH, Lehner KR, Folta KM (2006) Green light adjusts the plastid transcriptome during early photomorphogenic development. Plant Physiol 142:1256-1266

Donovan WP, Kushner SR (1986) Polynucleotide phosphorylase and ribonuclease II are required for cell viability and mRNA turnover in *Escherichia coli* K-12. Proc Natl Acad Sci USA 83:120-124

Douglas SE (1998) Plastid evolution: origins, diversity, trends. Curr Opin Genet Dev 8:655-661

Douglas SE (1999) Evolutionary history of plastids. Biol Bull 196:397-399

Drager RG, Girard-Bascou J, Choquet Y, Kindle KL, Stern DB (1998) *In vivo* evidence for 5'-3' exoribonuclease degradation of an unstable chloroplast mRNA. Plant J 13:85-96

Drager RG, Higgs DC, Kindle KL, Stern DB (1999) 5' to 3' exoribonucleolytic activity is a normal component of chloroplast mRNA decay pathways. Plant J 19:521-532

Dreyfus M, Regnier P (2002a) The poly(A) tail of mRNAs: bodyguard in eukaryotes, scavenger in bacteria. Cell 111:611-613

Dreyfus M, Regnier P (2002b) The poly(A) tail of mRNAs. Bodyguard in eukaryotes, scavenger in bacteria. Cell 111:611-613

Duhring U, Axmann IM, Hess WR, Wilde A (2006) An internal antisense RNA regulates expression of the photosynthesis gene *isiA*. Proc Natl Acad Sci USA 103:7054-7058

Dyall SD, Brown MT, Johnson PJ (2004) Ancient invasions: from endosymbionts to organelles. Science 304:253-257

Dziembowski A, Malewicz M, Minczuk M, Golik P, Dmochowska A, Stepien PP (1998) The yeast nuclear gene *DSS1*, which codes for a putative RNase II, is necessary for the function of the mitochondrial degradosome in processing and turnover of RNA. Mol Gen Genet 260:108-114

Dziembowski A, Piwowarski J, Hoser R, Minczuk M, Dmochowska A, Siep M, van der Spek H, Grivell L, Stepien PP (2003) The yeast mitochondrial degradosome. Its composition, interplay between RNA helicase and RNase activities and the role in mitochondrial RNA metabolism. J Biol Chem 278:1603-1611

Eberhard S, Drapier D, Wollman FA (2002) Searching limiting steps in the expression of chloroplast-encoded proteins: relations between gene copy number, transcription, transcript abundance and translation rate in the chloroplast of *Chlamydomonas reinhardtii*. Plant J 31:149-160

Edmonds M (2002) A history of poly A sequences: from formation to factors to function. Prog Nucleic Acid Res Mol Biol 71:285-389

Erickson B, Stern DB, Higgs DC (2005) Microarray analysis confirms the specificity of a *Chlamydomonas reinhardtii* chloroplast RNA stability mutant. Plant Physiol 137:534-544

Eriksson M, Villand P, Gardestrom P, Samuelsson G (1998) Induction and regulation of expression of a low-CO_2-induced mitochondrial carbonic anhydrase in *Chlamydomonas reinhardtii*. Plant Physiol 116:637-641

Even S, Pellegrini O, Zig L, Labas V, Vinh J, Brechemmier-Baey D, Putzer H (2005) Ribonucleases J1 and J2: two novel endoribonucleases in *B. subtilis* with functional homology to *E. coli* RNase E. Nucleic Acids Res 33:2141-2152

Evguenieva-Hackenberg E, Schiltz E, Klug G (2002) Dehydrogenases from all three domains of life cleave RNA. J Biol Chem 277:46145-46150

Felder S, Meierhoff K, Sane AP, Meurer J, Driemel C, Plucken H, Klaff P, Stein B, Bechtold N, Westhoff P (2001) The nucleus-encoded *HCF107* gene of *Arabidopsis* provides a link between intercistronic RNA processing and the accumulation of translation-competent *psbH* transcripts in chloroplasts. Plant Cell 13:2127-2141

Fisk DG, Walker MB, Barkan A (1999) Molecular cloning of the maize gene *crp1* reveals similarity between regulators of mitochondrial and chloroplast gene expression. EMBO J 18:2621-2630

Folichon M, Allemand F, Regnier P, Hajnsdorf E (2005) Stimulation of poly(A) synthesis by *Escherichia coli* poly(A)polymerase I is correlated with Hfq binding to poly(A) tails. FEBS J 272:454-463

Frazao C, McVey CE, Amblar M, Barbas A, Vonrhein C, Arraiano CM, Carrondo MA (2006) Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. Nature 443:110-114

French SW, Dawson DW, Chen HW, Rainey RN, Sievers SA, Balatoni CE, Wong L, Troke JJ, Nguyen MT, Koehler CM, Teitell MA (2006) The TCL1 oncoprotein binds the RNase PH domains of the PNPase exoribonuclease without affecting its RNA degrading activity. Cancer Lett 248:198-210

Gamble PE, Mullet JE (1989) Blue light regulates the accumulation of two *psbD-psbC* transcripts in barley chloroplasts. EMBO J 8:2785-2794

Gottesman S (2004) The small RNA regulators of *Escherichia coli*: roles and mechanisms*. Annu Rev Microbiol 58:303-328

Gregory ST, O'Connor M, Dahlberg AE (1996) Functional *Escherichia coli* 23S rRNAs containing processed and unprocessed intervening sequences from *Salmonella typhimurium*. Nucleic Acids Res 24:4918-4923

Gruissem W (1989) Chloroplast gene expression: How plants turn their plastids on. Cell 56:161-170

Grunberg-Manago M (1999) Messenger RNA stability and its role in control of gene expression in bacteria and phages. Annu Rev Genet 33:193-227

Grunberg-Manago M, Ochoa S (1955) Enzymatic synthesis and breakdown of polynucleotides; Polynucleotide phosphorylase. J Am Chem Soc 77:3165 - 3166

Haff LA, Bogorad L (1976) Poly(adenylic acid)-containing RNA from plastids of maize. Biochemistry 15:4110-4115

Harlow LS, Kadziola A, Jensen KF, Larsen S (2004) Crystal structure of the phosphorolytic exoribonuclease RNase PH from *Bacillus subtilis* and implications for its quaternary structure and tRNA binding. Protein Sci 13:668-677

Harris EH (1989) The *Chlamydomonas* Sourcebook: A comprehensive guide to biology and laboratory use. San Diego: Academic Press

Hattori M, Miyake H, Sugita M (2007) A pentatricopeptide repeat protein is required for RNA processing of *clpP* pre-mRNA in moss chloroplasts. J Biol Chem 282:10773-10782

Hayes R, Kudla J, Gruissem W (1999) Degrading chloroplast mRNA: the role of polyadenylation. Trends Biochem Sci 24:199-202

Hayes R, Kudla J, Schuster G, Gabay L, Maliga P, Gruissem W (1996) Chloroplast mRNA 3'-end processing by a high molecular weight protein complex is regulated by nuclear encoded RNA binding proteins. EMBO J 15:1132-1141

Hegeman CE, Halter CP, Owens TG, Hanson MR (2005) Expression of complementary RNA from chloroplast transgenes affects editing efficiency of transgene and endogenous chloroplast transcripts. Nucleic Acids Res 33:1454-1464

Hernandez H, Dziembowski A, Taverner T, Seraphin B, Robinson CV (2006) Subunit architecture of multimeric complexes isolated directly from cells. EMBO Rep 7:605-610

Hicks A, Drager RG, Higgs DC, Stern DB (2002) An mRNA 3' processing site targets downstream sequences for rapid degradation in *Chlamydomonas* chloroplasts. J Biol Chem 277:3325-3333

Higgs DC, Shapiro RS, Kindle KL, Stern DB (1999) Small *cis*-acting sequences that specify secondary structures in a chloroplast mRNA are essential for RNA stability and translation. Mol Cell Biol 19:8479-8491

Hirose T, Sugiura M (1997) Both RNA editing and RNA cleavage are required for translation of tobacco chloroplast *ndhD* mRNA: a possible regulatory mechanism for the expression of a chloroplast operon consisting of functionally unrelated genes. EMBO J 16:6804-6811

Hoffmeister M, Martin W (2003) Interspecific evolution: microbial symbiosis, endosymbiosis and gene transfer. Environ Microbiol 5:641-649

Holloway SP, Herrin DL (1998) Processing of a composite large subunit rRNA. Studies with *Chlamydomonas* mutants deficient in maturation of the 23S-like rRNA. Plant Cell 10:1193-1206

Horlitz M, Klaff P (2000) Gene-specific *trans*-regulatory functions of magnesium for chloroplast mRNA stability in higher plants. J Biol Chem 275:35638-35645

Houseley J, LaCava J, Tollervey D (2006) RNA-quality control by the exosome. Nat Rev Mol Cell Biol 7:529-539

Ishii R, Nureki O, Yokoyama S (2003) Crystal structure of the tRNA processing enzyme RNase PH from *Aquifex aeolicus*. J Biol Chem 278:32397-32404

Ishijima S, Uchibori A, Takagi H, Maki R, Ohnishi M (2003) Light-induced increase in free Mg(2+) concentration in spinach chloroplasts: Measurement of free Mg(2+) by using a fluorescent probe and necessity of stromal alkalinization. Arch Biochem Biophys 412:126-132

Jarrige A, Brechemier-Baey D, Mathy N, Duche O, Portier C (2002) Mutational analysis of polynucleotide phosphorylase from *Escherichia coli*. J Mol Biol 321:397-409

Jenkins BD, Kulhanek DJ, Barkan A (1997) Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. Plant Cell 9:283-296

Jiao HS, Hicks A, Simpson C, Stern DB (2004) Short dispersed repeats in the *Chlamydomonas* chloroplast genome are collocated with sites for mRNA 3' end formation. Curr Genet 45:311-322

Kallberg Y, Oppermann U, Jornvall H, Persson B (2002) Short-chain dehydrogenase/reductase (SDR) relationships: A large family with eight clusters common to human, animal, and plant genomes. Protein Sci 11:636-641

Kapoor S, Suzuki JY, Sugiura M (1997) Identification and functional significance of a new class of non-consensus-type plastid promoters. Plant J 11:327-337

Kastenmayer JP, Green PJ (2000) Novel features of the XRN-family in *Arabidopsis*: evidence that AtXRN4, one of several orthologs of nuclear Xrn2p/Rat1p, functions in the cytoplasm. Proc Natl Acad Sci USA 97:13985-13990

Kim M, Thum KE, Morishige DT, Mullet JE (1999) Detailed architecture of the barley chloroplast *psbD-psbC* blue light-responsive promoter. J Biol Chem 274:4684-4692

Kishine M, Takabayashi A, Munekage Y, Shikanai T, Endo T, Sato F (2004) Ribosomal RNA processing and an RNase R family member in chloroplasts of *Arabidopsis*. Plant Mol Biol 55:595-606

Komine Y, Kikis E, Schuster G, Stern D (2002) Evidence for *in vivo* modulation of chloroplast RNA stability by 3'-UTR homopolymeric tails in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA 99:4085-4090

Komine Y, Kwong L, Anguera MC, Schuster G, Stern DB (2000) Polyadenylation of three classes of chloroplast RNA in *Chlamydomonas reinhardtii*. RNA 6:598-607

Kordes E, Jock S, Fritsch J, Bosch F, Klug G (1994) Cloning of a gene involved in rRNA precursor processing and 23S rRNA cleavage in *Rhodobacter capsulatus*. J Bacteriol 176:1121-1127

Kramzar L, Mueller T, Erickson B, Higgs D (2006) Regulatory sequences of orthologous *petD* chloroplast mRNAs are highly specific among *Chlamydomonas* species. Plant Mol Biol 60:405-422

Kudla J, Hayes R, Gruissem W (1996) Polyadenylation accelerates degradation of chloroplast mRNA. EMBO J 15:7137-7146

Kushner SR (2002) mRNA decay in *Escherichia coli* comes of age. J Bacteriol 184:4658-4665

Kushner SR (2004) mRNA decay in prokaryotes and eukaryotes: different approaches to a similar problem. IUBMB Life 56:585-594

Lacava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D (2005) RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell 121:713-724

Lahmy S, Barneche F, Derancourt J, Filipowicz W, Delseny M, Echeverria M (2000) A chloroplastic RNA-binding protein is a new member of the PPR family. FEBS Lett 480:255-260

Leaver CJ (1973) Molecular integrity of chloroplast ribosomal ribonucleic acid. Biochem J 135:237-240

Leszczyniecka M, Kang DC, Sarkar D, Su ZZ, Holmes M, Valerie K, Fisher PB (2002) Identification and cloning of human polynucleotide phosphorylase, hPNPase old-35, in the context of terminal differentiation and cellular senescence. Proc Natl Acad Sci USA 99:16636-16641

Levy H, Kindle KL, Stern DB (1997) A nuclear mutation that affects the 3' processing of several mRNAs in *Chlamydomonas* chloroplasts. Plant Cell 9:825-836

Levy H, Kindle KL, Stern DB (1999) Target and specificity of a nuclear gene product that participates in mRNA 3'-end formation in *Chlamydomonas* chloroplasts. J Biol Chem 274:35955-35962

Li Z, Pandit S, Deutscher MP (1998) Polyadenylation of stable RNA precursors *in vivo*. Proc Natl Acad Sci USA 95:12158-12162

Liere K, Link G (1997) Chloroplast endoribonuclease p54 involved in RNA *3*'-end processing is regulated by phosphorylation and redox state. Nucleic Acids Res 25:2403-2438

Lisitsky I, Klaff P, Schuster G (1996) Addition of poly(A)-rich sequences to endonucleolytic cleavage sites in the degradation of spinach chloroplast mRNA. Proc Natl Acad Sci USA 93:13398-13403

Littauer UZ, Grunberg-Manago M (1999) Polynucleotide Phosphorylase. John Wiley and Sons Inc., New York

Littauer UZ, Soreq H (1982) Polynucleotide Phosphorylase. In: Boyer PD (ed) The Enzymes, 3rd edn. New York: Academic Press Inc, pp 517-553

Liu Q, Greimann JC, Lima CD (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. Cell 127:1223-1237

Lorentzen E, Walter P, Fribourg S, Evguenieva-Hackenberg E, Klug G, Conti E (2005) The archaeal exosome core is a hexameric ring structure with three catalytic subunits. Nat Struct Mol Biol 12:575-581

Lown FJ, Watson AT, Purton S (2001) *Chlamydomonas* nuclear mutants that fail to assemble respiratory or photosynthetic electron transfer complexes. Biochem Soc Trans 29:452-455

Lung B, Zemann A, Madej MJ, Schuelke M, Techritz S, Ruf S, Bock R, Huttenhofer A (2006) Identification of small non-coding RNAs from mitochondria and chloroplasts. Nucleic Acids Res 34:3842-3852

Lurin C, Andres C, Aubourg S, Bellaoui M, Bitton F, Bruyere C, Caboche M, Debast C, Gualberto J, Hoffmann B, Lecharny A, Le Ret M, Martin-Magniette ML, Mireau H, Peeters N, Renou JP, Szurek B, Taconnat L, Small I (2004) Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell 16:2089-2103

Mackie GA (1998) Ribonuclease E is a 5'-end-dependent endonuclease. Nature 395:720-723

Marcaida MJ, DePristo MA, Chandran V, Carpousis AJ, Luisi BF (2006) The RNA degradosome: life in the fast lane of adaptive molecular evolution. Trends Biochem Sci 31:359-365

Martin G, Keller W (2004) Sequence motifs that distinguish ATP(CTP):tRNA nucleotidyl transferases from eubacterial poly(A) polymerases. RNA 10:899-906

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci USA 99:12246-12251

Marujo PE, Hajnsdorf E, Le Derout J, Andrade R, Arraiano CM, Regnier P (2000) RNase II removes the oligo(A) tails that destabilize the *rpsO* mRNA of *Escherichia coli*. RNA 6:1185-1193

Mattatall NR, Sanderson KE (1998) RNase III deficient *Salmonella typhimurium* LT2 contains intervening sequences (IVSs) in its 23S rRNA. FEMS Microbiol Lett 159:179-185

Meierhoff K, Felder S, Nakamura T, Bechtold N, Schuster G (2003) HCF152, an *Arabidopsis* RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast *psbB-psbT-psbH-petB-petD* RNAs. Plant Cell 15:1480-1495

Meurer J, Berger A, Westhoff P (1996) A nuclear mutant of *Arabidopsis* with impaired stability on distinct transcripts of the plastid *psbB*, *psbD/C*, *ndhH*, and *ndhC* operons. Plant Cell 8:1193-1207

Miyagi T, Kapoor S, Sugita M, Sugiura M (1998) Transcript analysis of the tobacco plastid operon *rps2/atp*I/H/F/A reveals the existence of a non-consensus type II (NCII) promoter upstream of the *atpI* coding sequence. Mol Gen Genet 257:299-307

Mohanty BK, Kushner SR (2000a) Polynucleotide phosphorylase functions both as a 3' to 5' exonuclease and a poly(A) polymerase in *Escherichia coli*. Proc Natl Acad Sci USA 97:11966-11971

Mohanty BK, Kushner SR (2000b) Polynucleotide phosphorylase functions both as a 3' to 5' exonuclease and a poly(A) polymerase in *Escherichia coli*. Proc Natl Acad Sci USA 97:11966-11971

Mohanty BK, Maples VF, Kushner SR (2004) The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in *Escherichia coli*. Mol Microbiol 54:905-920

Monde RA, Greene JC, Stern DB (2000a) Disruption of the *petB-petD* intergenic region in tobacco chloroplasts affects *petD* RNA accumulation and translation. Mol Gen Genet 263:610-618

Monde RA, Schuster G, Stern DB (2000b) Processing and degradation of chloroplast mRNA. Biochimie 82:573-582

Mullet JE (1993) Dynamic regulation of chloroplast transcription. Plant Physiol 103:309-313

Mullet JE, Klein RR (1987) Transcription and RNA stability are important determinants of higher plant chloroplast RNA levels. EMBO J 6:1571-1579

Murakami S, Kuehnle K, Stern DB (2005) A spontaneous tRNA suppressor of a mutation in the *Chlamydomonas reinhardtii* nuclear *MCD1* gene required for stability of the chloroplast *petD* mRNA. Nucleic Acids Res 33:3372-3380

Nakamura T, Meierhoff K, Westhoff P, Schuster G (2003) RNA-binding properties of HCF152, an *Arabidopsis* PPR protein involved in the processing of chloroplast RNA. Eur J Biochem 270:4070-4081

Neuhaus H, Scholz A, 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

Newbury SF, Muhlemann O, Stoecklin G (2006) Turnover in the Alps: an mRNA perspective. Workshops on mechanisms and regulation of mRNA turnover. EMBO Rep 7:143-148

Ng WV, Kennedy SP, Mahairas GG, Berquist B, Pan M, Shukla HD, Lasky SR, Baliga NS, Thorsson V, Sbrogna J, Swartzell S, Weir D, Hall J, Dahl TA, Welti R, Goo YA, Leithauser B, Keller K, Cruz R, Danson MJ, Hough DW, Maddocks DG, Jablonski PE, Krebs MP, Angevine CM, Dale H, Isenbarger TA, Peck RF, Pohlschroder M, Spudich JL, Jung K-H, Alam M, Freitas T, Hou S, Daniels CJ, Dennis PP, Omer AD, Ebhardt H, Lowe TM, Liang P, Riley M, Hood L, DasSarma S (2000) Genome sequence of *Halobacterium* species NRC-1. Proc Natl Acad Sci USA 97:12176-12181

Nickelsen J (2000) Mutations at three different nuclear loci of *Chlamydomonas* suppress a defect in chloroplast *psbD* mRNA accumulation. Curr Genet 37:136-142

Nickelsen J (2003) Chloroplast RNA-binding proteins. Curr Genet 43:392-399

Nickelsen J, Fleischmann M, Boudreau E, Rahire M, Rochaix J-D (1999) Identification of *cis*-acting RNA leader elements required for chloroplast *psbD* gene expression in *Chlamydomonas*. Plant Cell 11:957-970

Nickelsen J, Link G (1989) Interaction of a 3' RNA region of the mustard *trnK* gene with chloroplast proteins. Nucleic Acids Res 17:9637-9648

Nickelsen J, Link G (1991) RNA-protein interactions at transcript 3' ends and evidence for *trnK-psbA* cotranscription in mustard chloroplasts. Mol Gen Genet 228:89-96

Nishimura Y, Kikis EA, Zimmer SL, Komine Y, Stern DB (2004) Antisense transcript and RNA processing alterations suppress instability of polyadenylated mRNA in *Chlamydomonas* chloroplasts. Plant Cell 16:2849-2869

Orozco EM Jr, Chen LJ, Eilers RJ (1990) The divergently transcribed rbcL and atpB genes of tobacco plastid DNA are separated by nineteen base pairs. Curr Genet 17:65-71

Peltier JB, Cai Y, Sun Q, Zabrouskov V, Giacomelli L, Rudella A, Ytterberg AJ, Rutschow H, van Wijk KJ (2006) The oligomeric stromal proteome of *Arabidopsis thaliana* chloroplasts. Mol Cell Proteomics 5:114-133

Perrin R, Meyer EH, Zaepfel M, Kim YJ, Mache R, Grienenberger JM, Gualberto JM, Gagliardi D (2004) Two exoribonucleases act sequentially to process mature 3'-ends of *atp9* mRNAs in *Arabidopsis* mitochondria. J Biol Chem 279:25440-25446

Pfannschmidt T, Ogrzewalla K, Baginsky S, Sickmann A, Meyer HE, Link G (2000) The multisubunit chloroplast RNA polymerase A from mustard (*Sinapis alba* L.). Integration of a prokaryotic core into a larger complex with organelle-specific functions. Eur J Biochem 267:253-261

Phinney BS, Thelen JJ (2005) Proteomic characterization of a triton-insoluble fraction from chloroplasts defines a novel group of proteins associated with macromolecular structures. J Proteome Res 4:497-506

Portnoy V, Evguenieva-Hackenberg E, Klein F, Walter P, Lorentzen E, Klug G, Schuster G (2005) RNA polyadenylation in Archaea: not observed in *Haloferax* while the exosome polynucleotidylates RNA in *Sulfolobus*. EMBO Rep 6:1188-1193

Portnoy V, Schuster G (2006) RNA polyadenylation and degradation in different Archaea; roles of the exosome and RNase R. Nucleic Acids Res 34:5923-5931

Raijmakers R, Noordman YE, van Venrooij WJ, Pruijn GJ (2002) Protein-protein interactions of hCsl4p with other human exosome subunits. J Mol Biol 315:809-818

Rainey RN, Glavin JD, Chen HW, French SW, Teitell MA, Koehler CM (2006) A new function in translocation for the mitochondrial i-AAA protease Yme1: import of polynucleotide phosphorylase into the intermembrane space. Mol Cell Biol 26:8488-8497

Rapp JC, Baumgartner BJ, Mullet J (1992) Quantitative analysis of transcription and RNA levels of 15 barley chloroplast genes. Transcription rates and mRNA levels vary over 300-fold; predicted mRNA stabilities vary 30-fold. J Biol Chem 267:21404-21411

Raynal LC, Krisch HM, Carpousis AJ (1998) The *Bacillus subtilis* nucleotidyltransferase is a tRNA CCA-adding enzyme. J Bacteriol 180:6276-6282

Rochaix JD (1995) *Chlamydomonas reinhardtii* as the photosynthetic yeast. Annu Rev Genet 29:209-230

Rott R, Drager RG, Stern DB, Schuster G (1996) The 3' untranslated regions of chloroplast genes in *Chlamydomonas reinhardtii* do not serve as efficient transcriptional terminators. Mol Gen Genet 252:676-683

Rott R, Zipor G, Portnoy V, Liveanu V, Schuster G (2003) RNA polyadenylation and degradation in cyanobacteria are similar to the chloroplast but different from *Escherichia coli*. J Biol Chem 278:15771-15777

Rymarquis LA, Higgs DC, Stern DB (2006a) Nuclear suppressors define three factors that participate in both 5’ and 3’ end processing of mRNAs in *Chlamydomonas* chloroplasts. Plant J 46:448-461

Rymarquis LA, Webster BR, Stern DB (2006b) The nucleus-encoded factor MCD4 participates in degradation of nonfunctional 3' UTR sequences generated by cleavage of pre-mRNA in *Chlamydomonas* chloroplasts. Mol Genet Genomics 277:329-340

Sane AP, Stein B, Westhoff P (2005) The nuclear gene *HCF107* encodes a membrane-associated R-TPR (RNA tetratricopeptide repeat)-containing protein involved in expression of the plastidial *psbH* gene in *Arabidopsis*. Plant J 42:720-730

Sarkar D, Leszczyniecka M, Kang DC, Lebedeva IV, Valerie K, Dhar S, Pandita TK, Fisher PB (2003) Down-regulation of Myc as a potential target for growth arrest induced by human polynucleotide phosphorylase (hPNPaseold-35) in human melanoma cells. J Biol Chem 278:24542-24551

Sarkar N (1997) Polyadenylation of mRNA in prokaryotes. Annu Rev Biochem 66:173-197

Sauret-Gueto S, Botella-Pavia P, Flores-Perez U, Martinez-Garcia JF, San Roman C, Leon P, Boronat A, Rodriguez-Concepcion M (2006) Plastid cues posttranscriptionally regulate the accumulation of key enzymes of the methylerythritol phosphate pathway in *Arabidopsis*. Plant Physiol 141:75-84

Schiffer S, Rosch S, Marchfelder A (2002) Assigning a function to a conserved group of proteins: the tRNA 3'-processing enzymes. EMBO J 21:2769-2777

Schuster G, Lisitsky I, Klaff P (1999) Update on chloroplast molecular biology: Polyadenylation and degradation of mRNA in the chloroplast. Plant Physiol 120:937-944

Serino G, Maliga P (1998) RNA polymerase subunits encoded by the plastid *rpo* genes are not shared with the nucleus-encoded plastid enzyme. Plant Physiol 117:1165-1170

Slomovic S, Laufer D, Geiger D, Schuster G (2005) Polyadenylation and degradation of human mitochondrial RNA: the prokaryotic past leaves its mark. Mol Cell Biol 25:6427-6435

Slomovic S, Portnoy V, Liveanu V, Schuster G (2006a) RNA polyadenylation in prokaryotes and organelles; Different tails tell different tales. Crit Rev Plant Sci 25:65-77

Slomovic S, Portnoy V, Liveanu V, Schuster G (2006b) RNA Polyadenylation in prokaryotes and organelles; Different tails tell different tales. Crit Rev Plant Sci 25:65-77

Sohlberg B, Huang J, Cohen SN (2003) The *Streptomyces coelicolor* polynucleotide phosphorylase homologue, and not the putative poly(A) polymerase, can polyadenylate RNA. J Bacteriol 185:7273-7278

Souret FF, Kastenmayer JP, Green PJ (2004) AtXRN4 degrades mRNA in *Arabidopsis* and its substrates include selected miRNA targets. Mol Cell 15:173-183

Stern DB, 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 DB, Gruissem W (1989) Chloroplast mRNA 3' end maturation is biochemically distinct from prokaryotic mRNA processing. Plant Mol Biol 13:615-625

Stern DB, Hanson MR, Barkan A (2004) Genetics and genomics of chloroplast biogenesis: maize as a model system. Trends Plant Sci 9:293-301

Stern DB, Kindle KL (1993) 3' end maturation of the *Chlamydomonas reinhardtii* chloroplast *atpB* mRNA is a two-step process. Mol Cell Biol 13:2277-2285

Storz G, Altuvia S, Wassarman KM (2005) An abundance of RNA regulators. Annu Rev Biochem 74:199-217

Sugita M, Svab Z, Maliga P, Sugiura M (1997) Targeted deletion of *sprA* from the tobacco plastid genome indicates that the encoded small RNA is not essential for pre-16S rRNA maturation in plastids. Mol Gen Genet 257:23-27

Suzuki JY, Ytterberg AJ, Beardslee TA, Allison LA, Wijk KJ, Maliga P (2004) Affinity purification of the tobacco plastid RNA polymerase and *in vitro* reconstitution of the holoenzyme. Plant J 40:164-172

Symmons MF, Jones GH, Luisi BF (2000a) A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. Structure 8:1215-1226

Symmons MF, Jones GH, Luisi BF (2000b) A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. Structure Fold Des 8:1215-1226

Symmons MF, Williams MG, Luisi BF, Jones GH, Carpousis AJ (2002) Running rings around RNA: a superfamily of phosphate-dependent RNases. Trends Biochem Sci 27:11-18

Tanaka M, Obokata J, Chunwongse J, Shinozaki K, Sugiura M (1987) Rapid splicing and stepwise processing of a transcript from the psbB operon in tobacco chloroplasts: Determination of the intron sites in petB and petD. MGG 209:427-431

Thum KE, Kim M, Christopher DA, Mullet JE (2001) Cryptochrome 1, cryptochrome 2, and phytochrome a co-activate the chloroplast *psbD* blue light-responsive promoter. Plant Cell 13:2747-2760

Vaistij FE, Boudreau E, Lemaire SD, Goldschmidt-Clermont M, Rochaix JD (2000) Characterization of Mbb1, a nucleus-encoded tetratricopeptide-like repeat protein required for expression of the chloroplast *psbB/psbT/psbH* gene cluster in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA 97:14813-14818

Vanacova S, Wolf J, Martin G, Blank D, Dettwiler S, Friedlein A, Langen H, Keith G, Keller W (2005) A new yeast poly(A) polymerase complex involved in RNA quality control. PLoS Biol 3:e189

Vanzo NF, Li YS, Py B, Blum E, Higgins CF, Raynal LC, Krisch HM, Carpousis AJ (1998) Ribonuclease E organizes the protein interactions in the *Escherichia coli* RNA degradosome. Genes Dev 12:2770-2781

Vera A, Sugiura M (1994) A novel RNA gene in the tobacco plastid genome: its possible role in the maturation of 16S rRNA. EMBO J 13:2211-2217

Walter M, Kilian J, Kudla J (2002) PNPase activity determines the efficiency of mRNA 3'-end processing, the degradation of tRNA and the extent of polyadenylation in chloroplasts. EMBO J 21:6905-6914

Wang MJ, Davis NW, Gegenheimer P (1988) Novel mechanisms for maturation of chloroplast transfer RNA precursors. EMBO J 7:1567-1574

Weiner AM (2005) E Pluribus Unum: 3' end formation of polyadenylated mRNAs, Histone mRNAs, and U snRNAs. Mol Cell 20:168-170

Westhoff P, Herrmann RG (1988) Complex RNA maturation in chloroplasts: the *psbB* operon from spinach. Eur J Biochem 171:551-564

Wickens M, Anderson P, Jackson RJ (1997) Life and death in the cytoplasm: messages from the 3' end. Curr Opin Genet Dev 7:220-232

Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, Boulay J, Regnault B, Devaux F, Namane A, Seraphin B, Libri D, Jacquier A (2005) Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell 121:725-737

Yamaguchi K, Beligni MV, Prieto S, Haynes PA, McDonald WH, Yates JR, 3rd, Mayfield SP (2003) Proteomic characterization of the *Chlamydomonas reinhardtii* chloroplast ribosome. Identification of proteins unique to the 70S ribosome. J Biol Chem 278:33774-33785

Yamamoto YY, Puente P, Deng XW (2000) An *Arabidopsis* cotyledon-specific albino locus: a possible role in 16S rRNA maturation. Plant Cell Physiol 41:68-76

Yang J, Schuster G, Stern DB (1996) CSP41, a sequence-specific chloroplast mRNA binding protein, is an endoribonuclease. Plant Cell 8:1409-1420

Yang J, Stern DB (1997) The spinach chloroplast endoribonuclease CSP41 cleaves the 3' untranslated region of *petD* mRNA primarily within its terminal stem-loop structure. J Biol Chem 272:12784-12880

Yehudai-Resheff S, Hirsh M, Schuster G (2001) Polynucleotide phosphorylase functions as both an exonuclease and a poly(A) polymerase in spinach chloroplasts. Mol Cell Biol 21:5408-5416

Yehudai-Resheff S, Portnoy V, Yogev S, Adir N, Schuster G (2003) Domain analysis of the chloroplast polynucleotide phosphorylase reveals discrete functions in RNA degradation, polyadenylation, and sequence homology with exosome proteins. Plant Cell 15:2003-2019

Yehudai-Resheff S, Zimmer SL, Komine Y, Stern DB (2007) Integration of chloroplast nucleic acid metabolism into the phosphate deprivation response in *Chlamydomonas reinhardtii*. Plant Cell 19:1023-1038

Yosef I, Irihimovitch V, Knopf JA, Cohen I, Orr-Dahan I, Nahum E, Keasar C, Shapira M (2004) RNA binding activity of the ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit from *Chlamydomonas reinhardtii*. J Biol Chem 279:10148-10156

Yue D, Maizels N, Weiner AM (1996) CCA-adding enzymes and poly(A) polymerases are all members of the same nucleotidyltransferase superfamily: characterization of the CCA-adding enzyme from the archaeal hyperthermophile *Sulfolobus shibatae*. RNA 2:895-908

Yukawa M, Kuroda H, Sugiura M (2006) A new *in vitro* translation system for non-radioactive assay from tobacco chloroplasts: effect of pre-mRNA processing on translation *in vitro*. Plant J 49:367-376

Zoschke R, Liere K, Borner T (2007) From seedling to mature plant: *Arabidopsis* plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. Plant J 50:710-722

Zuo Y, Deutscher MP (2001) Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Res 29:1017-1026

Zuo Y, Vincent HA, Zhang J, Wang Y, Deutscher MP, Malhotra A (2006) Structural basis for processivity and single-strand specificity of RNase II. Mol Cell 24:149-156

Bollenbach, Thomas J.
Boyce Thompson Institute for Plant Research, Tower Rd. Ithaca NY 14853, USA

Portnoy, Victoria
Department of Biology, Technion-Israel Institute of Technology, Haifa 32000, Israel

Schuster, Gadi
Department of Biology, Technion-Israel Institute of Technology, Haifa 32000, Israel

Stern, David B.
Boyce Thompson Institute for Plant Research, Tower Rd. Ithaca NY 14853
ds28@cornell.edu

RNA splicing and RNA editing in chloroplasts

Christian Schmitz-Linneweber and Alice Barkan

Abstract

During the evolution of chloroplasts from their bacterial ancestor traits emerged that are absent or rare in bacteria. Prominent among these acquired traits are RNA splicing and RNA editing. The numbers and distribution of introns and editing sites in different taxa suggest that editing and splicing have taken different evolutionary pathways in different chloroplast lineages. Both processes are dependent on nuclear-encoded factors and, intriguingly, PPR (pentatricopeptide repeat) proteins have recently been recognized as a common player. This review summarizes recent progress in understanding the mechanisms, regulation, and *trans*-acting factors for these two types of RNA processing.

1 Introduction

As endosymbiotic descendants of cyanobacteria, chloroplasts share many features of their metabolism and biogenesis with prokaryotes. However, their coevolution with the eukaryotic host genome has led to the pronounced modification of prokaryotic features and the acquisition of novel features not present in their prokaryotic ancestors. Understanding these plastid-specific attributes is critical for understanding how the organelle was integrated into the regulatory circuits of the plant cell. Two features of chloroplast gene expression exemplify these acquired features. First, the chloroplast genome is rich in intervening sequences, whereas introns are rare in bacteria. Even more striking, chloroplasts display an RNA processing event called RNA editing that does not exist at all in prokaryotes. The acquisition of these two RNA processing mechanisms in chloroplasts necessitated the recruitment of pre-existing proteins and/or the evolution of novel proteins to participate in these processes. This review summarizes recent advances in understanding the molecular mechanisms, evolution and regulation of these two RNA maturation steps in chloroplasts.

2 Plastid RNA splicing

Accurate splicing of plastid introns is essential for the biogenesis of the chloroplast, as introns disrupt plastid genes encoding components of the gene expression

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0233 / Published online: 13 June 2007

machinery and the photosynthetic apparatus. Moreover, RNA splicing can be exploited as an on/off switch for the expression of intron containing genes and provides opportunities for regulation of plastid gene expression by environmental and developmental cues. The machinery responsible for plastid splicing is encoded primarily by nuclear genes, with almost all characterized plastid splicing factors being – as expected – essential for chloroplast development. The growing number of such factors identified in recent years suggests a complexity that was not anticipated given the resemblance of plastid introns to "self-splicing" introns found in other organisms.

2.1 Intron classes and splicing mechanisms

Chloroplast introns are classified as either group I, group II, or group III introns by virtue of conserved features of primary sequence and predicted secondary structure (Michel and Dujon 1983; Michel et al. 1989; Copertino and Hallick 1993; Lambowitz et al. 1999; Bonen and Vogel 2001). Group I and group II introns are distributed broadly among prokaryotes and prokaryote-derived organelles, but they are particularly prevalent in the mitochondria and chloroplasts of plants and algae. Both group I and group II introns are considered to be ribozymes because some introns in each class exhibit self-splicing activity *in vitro.* Despite this similarity, group I and group II introns differ fundamentally in structure and catalytic mechanism. Group I introns are relatively small and uniform in structure, with two central helical domains that are stabilized by peripheral domains (Michel and Westhof 1990; Lehnert et al. 1996; Golden et al. 1998; Westhof 2002). Group II introns are larger and more structurally variable, consisting of six helical domains emanating from a central core, and inter-domain tertiary contacts that create a compact catalytic center (Michel et al. 1989; Michel and Ferat 1995; Qin and Pyle 1998). This canonical group II intron structure is often highly modified; for example, the chloroplast *psa*A mRNA in *C. reinhardtii* (Kuck et al. 1987; Choquet et al. 1988; Herrin and Schmidt 1988) and the land plant chloroplast *rps*12 mRNA (Fromm et al. 1986; Zaita et al. 1987) are transcribed in pieces that must then be "*trans*-spliced" (reviewed in Bonen and Vogel 2001). During *trans*-splicing, intron fragments are believed to assemble via RNA-RNA interactions to recreate an intact group II intron structure. Group III introns are believed to be still more highly degenerate group II introns, and are a specialized case in that they have been found only in Euglenoids (Christopher and Hallick 1989).

The chemistry of group I and group II intron splicing has been elucidated primarily from the study of self-splicing introns from fungal mitochondria (reviewed in Qin and Pyle 1998; Lambowitz et al. 1999; Bonen and Vogel 2001). Both intron classes splice via two consecutive *trans*-esterifications involving first the 5' and then the 3' splice junction, but the reactions otherwise differ. The splicing of group I introns is initiated by an exogenous guanosine that attacks the 5' splice junction; the liberated 3' hydroxyl group then attacks the 3' splice junction, resulting in exon ligation and intron release. By contrast, group II splicing generally initiates when a 2' hydroxyl group of a "bulged" adenosine residue in the domain 6 helix attacks the 5' splice junction. This yields a branched "lariat" structure in which the 2'hydroxyl group is covalently attached to the 5' phosphate at the 5' end of the intron. Splicing is completed during a second step during which the liberated 3'-hydroxyl group at the 5' splice junction attacks the 3' splice junction, resulting in exon ligation and release of an excised intron lariat. These canonical splicing mechanisms likely apply to most plastid introns, as excised group II intron lariats have been identified for many chloroplast introns (Kim et al. 1993; Vogel and Borner 2002) and mutagenesis of predicted functional elements of a chloroplast group I intron disrupted splicing *in vivo* and *in vitro* (Lee and Herrin 2003).

An alternative pathway for group II splicing *in vitro* is initiated by water or hydroxide rather than by a bulged intron adenosine (Daniels et al. 1996). In fact, a bulged adenosine is missing from domain six in land plant chloroplast *trnV-UAC* introns. The excised *trnV-UAC* intron was detected as a linear molecule but not as a lariat in barley chloroplasts (Vogel and Borner 2002), supporting the notion that this intron is, indeed, spliced via a hydrolytic pathway *in vivo*. Whether the use of this alternative pathway has any physiological significance is unclear.

Similarities between the chemistry of group II splicing and spliceosome-mediated splicing in the nucleus, together with structural similarities between specific snRNAs and specific group II intron domains have led to the intriguing speculation that spliceosomal snRNAs might be derived domains of ancestral self-splicing group II introns (Cech 1986; Hetzer et al. 1997; Shukla and Padgett 2002; Villa et al. 2002; Sashital et al. 2004). The evolutionary lability of group II introns in plant organelles, as exemplified by the *trans*-spliced introns in chloroplasts, lends credence to this idea. If true, endosymbiotic organelles could have been the initial donors of a proto-intron, thereby shaping eukaryotic genomes in a most fundamental way.

2.2 Intron distribution

Bryophytes, gymnosperms, angiosperms and their closest algal relatives (members of the charophyta) share a basic set of chloroplast introns, consisting of one group I intron and ~20 group II introns. 20 out of the 21 plastid introns represented in land plants have been detected in at least one charophyte lineage, indicating that these introns were acquired prior to the emergence of land plants; only the *clpP-2* group II intron was incorporated later, during the transition to land plants (Wakasugi et al. 2001; Turmel et al. 2002; Kugita et al. 2003a; Sugiura et al. 2003; Turmel et al. 2006). Thus, plastid introns were acquired early during plant evolution and are among the most stable features of the chloroplast genome. Land plants and charophyte algae (together called the streptophyta) contain a single group I intron, in the *trnL-UAA* gene. This intron is considered to be the most ancient of all plastid introns as it is represented in land plants, in both charophyte and chlorophyte green algae, in red algae and even in cyanobacteria (Xu et al. 1990; Simon et al. 2003).

Differences in plastid intron content among land plant species reflect lineage specific intron loss via either the complete loss of intron-containing genes (e.g. loss of the *ndhA* and *ndhB* genes in black pine chloroplasts, Wakasugi et al. 1994), or intron loss with retention of the host gene (e.g. introns disrupt the *clpP* and *rpoC* genes in dicots and ancestral embryophytes but not in monocot grasses). Thus, maize and rice chloroplasts have seventeen group II introns, whereas *Arabidopsis* and tobacco chloroplasts have twenty. The parasitic plant *Epifagus virginiana* provides an extreme case of intron loss, in that it retains only six chloroplast introns (Wolfe et al. 1992; Ems et al. 1995).

Group II introns are absent from the chloroplasts of the most basal species within the streptophyta, like the charophyte algae *Mesostigma viride* and *Chlorokybus atmophyticus* (Lemieux et al. 2000; Turmel et al. 2006), whereas the overall plastid gene organization in these species is highly conserved with land plants. This supports the idea that the acquisition of chloroplast group II introns within and outside the streptophyta were independent events. Accordingly, the chloroplasts of the chlorophyte *C. reinhardtii* harbors five group I introns and only two group II introns, none of which are found in land plants (Maul et al. 2002). *Euglena gracilis*, a photosynthetic protist, houses the most intron rich chloroplast genome described to date, with at least 155 introns; these introns fall into the group II and group III classes (Hallick et al. 1993).

2.3 Proteins involved in the splicing of chloroplast introns

2.3.1 Proteins are required for chloroplast intron splicing

Group I and group II introns are classified as ribozymes because representatives of both intron classes have been shown to self-splice *in vitro* (reviewed in Lambowitz et al. 1999). However, many introns with the characteristic features of group I or group II introns cannot be coerced to self-splice and, in fact, not one of the ~40 introns in the organelles of vascular plants has been reported to self-splice *in vitro*. Only two examples of self-splicing group II introns in chloroplasts have been described, one of them in a psychrophilic *Chlamydomonas* species (Odom et al. 2004), the other in *Euglena myxocylindracea* (Sheveleva and Hallick 2004). This latter intron is, however, an evolutionary oddity because it likely represents a recent horizontal transfer from a cyanobacterial donor. In *Chlamydomonas* and other algae, several plastid group I introns have been demonstrated to self-splice (Herrin et al. 1990, 1991; Deshpande et al. 1997; Kapoor et al. 1997; Simon et al. 2003), while the group I intron in higher plant *trnL* genes does not (Simon et al. 2003). Even where self-splicing has been detected, the reactions generally require non-physiological salt and temperature conditions. Moreover, a self-splicing group I intron from *C. reinhardtii* chloroplasts was more tolerant of mutations in core elements when expressed *in vivo* than during self-splicing *in vitro* (Lee and Herrin 2003). Together, these data strongly suggest that accessory factors facilitate the splicing of most or all group I and group II introns *in vivo*. Indeed, genetic data summarized below have provided evidence for the involvement of proteins in the splicing of almost all of the group II introns represented in the chloroplasts of vascular plants and *C. reinhardtii*.

Proteins involved in group I and group II intron splicing fall into two general classes: conserved intron-encoded "maturase" proteins and diverse "host-encoded" factors (reviewed in Lambowitz et al. 1999). The majority of splicing factors in chloroplasts fall into this second category.

2.3.2 Intron-encoded maturases in chloroplasts

Group I and group II intron maturases have been studied primarily in fungal mitochondria and bacteria (reviewed in Lambowitz et al. 1999). Group I maturases are related to the LAGLIDADG class of homing endonucleases; some group I maturases promote both intron homing and splicing, whereas others have lost their DNA endonuclease function and are now specialized splicing factors. The single group I intron in land plant chloroplasts lacks a maturase open reading frame. However, three group I introns in *C. reinhardtii* chloroplasts encode maturase-like proteins; these have been shown to promote insertion of their host intron into intronless alleles (Durrenberger and Rochaix 1991; Holloway et al. 1999; Odom et al. 2001) but deletion of these open reading frames did not result in splicing defects (Thompson and Herrin 1991; Johanningmeier and Heiss 1993). Therefore, the available data suggest that these maturase-like proteins do not function in splicing.

Group II intron maturases are characterized by reverse-transcriptase and endonuclease domains involved in intron mobility, and a "domain X", which is implicated in RNA binding and splicing (reviewed in Lambowitz et al. 1999). The *C. reinhardtii* chloroplast genome lacks group II maturase open reading frames, whereas a single open reading frame related to group II maturases is encoded in land plant chloroplasts genomes; this gene is called *matK* and resides in the *trnK* intron (Neuhaus and Link 1987). MatK is a degenerate maturase-like protein, consisting of domain X fused to a remnant of the reverse transcriptase domain. Several lines of evidence suggest that MatK is involved not only in the splicing of its host *trnK* intron, but also in the splicing of other group II introns. First, MatK binds RNA *in vitro* (Liere and Link 1995) and the sequence encoding MatK is subject to an RNA editing event that increases its conservation with functional maturases (Vogel et al. 1997). Furthermore, the absence of MatK in maize and barley mutants lacking plastid ribosomes correlates with the failure to splice the *trn*K intron (Vogel et al. 1997) as well as several other group II introns (Jenkins et al. 1997; Vogel et al. 1999). Although the pleiotropic nature of the mutants used in these studies precludes firm conclusions about the role of MatK in splicing, these findings raised the possibility that MatK may facilitate the splicing of multiple introns, unlike canonical group II maturases which act specifically on the intron in which they are encoded (Lambowitz et al. 1999). Additional evidence that MatK promotes the splicing of multiple plastid introns arose from the sequence of the plastid genome of the non-photosynthetic angiosperm *Epifagus virginiana*. The Epifagus plastid genome lacks *trnK* but retains a stand-alone *matK* gene; it was proposed that retention of *matK* reflects a role for MatK in the splicing of one or more of the six group II introns retained in the *Epifagus* chloroplast genome, all of which are accurately spliced *in vivo* (Wolfe et al. 1992; Ems et al. 1995). Still, proof that MatK promotes splicing is lacking. Initial attempts to delete *matK* from the chloroplast genome in tobacco resulted only in heteroplastomic plants (R. Maier, personal communication); this is consistent with a role for *matK* in splicing *trnK* and/or other essential plastid RNAs that contain group II introns. Biochemical approaches and the analysis of hypomorphic, non-lethal alleles of *matK* may help to clarify this issue.

2.3.3 Nucleus-encoded splicing factors in chloroplasts

Numerous nucleus-encoded proteins involved in the splicing of chloroplast introns in both vascular plants and algae have been identified in recent years, primarily through genetic screens for nuclear mutations that cause defects in chloroplast gene expression. In *C. reinhardtii*, several splicing factors involved in the maturation of the *psaA* mRNA have been described. Maturation of this mRNA is particularly complex, as it involves the *trans*-splicing of two group II introns: intron 2 is transcribed in two segments together with the flanking exons, whereas intron 1 consists of three pieces: 5' and 3' intron fragments that are cotranscribed with flanking exons and an internal intron fragment that is independently transcribed from a chloroplast locus called *tscA* (Kuck et al. 1987; Choquet et al. 1988; Herrin and Schmidt 1988; Roitgrund and Mets 1990; Goldschmidt-Clermont et al. 1991). The *tscA* RNA is proposed to bridge the 5' and 3' fragments of intron 1, but domain 1 of this mosaic intron appears to lack critical elements, suggesting that an additional intron fragment remains to be discovered (Turmel et al. 1995). As might be expected, a large number of accessory factors are required to assemble and splice the two *psaA* introns. In fact, mutations that disrupt this process define at least fourteen nuclear genes (Goldschmidt-Clermont et al. 1990); this gene set includes genes that function directly in splicing, as well as genes that affect splicing indirectly by promoting the maturation of the *tscA* RNA. One gene in the latter class, *Rat1,* has been cloned. *Rat1* codes for a protein with a domain that is related to NAD+-binding domains from eukaryotic organisms, and that can interact with the *tscA* RNA in a yeast 3-hybrid assay (Balczun et al. 2005).

The molecular cloning of three genes that seem likely to function directly in the *trans*-splicing of the *C. reinhardtii psaA* mRNA has been reported: *Raa2,* which is required for the *trans*-splicing of intron 2, *Raa3,* which is required for the *trans*-splicing of intron 1, and *Raa1,* which is required for the *trans*-splicing of both introns (Perron et al. 1999; Rivier et al. 2001; Merendino et al. 2006). Raa3 exhibits limited similarity to pyridoxamine 5'-phosphate oxidases and is found in a large complex in the chloroplast stroma, together with the *tscA* and *psaA* exon 1 precursor RNAs (Rivier et al. 2001). In contrast, Raa1 and Raa2 are associated with a chloroplast membrane fraction and are found, at least in part, in a complex with one another (Perron et al. 1999, 2004; Merendino et al. 2006). Raa2 resembles pseudouridine synthase enzymes; however, mutagenesis of amino acids that are essential for the catalytic activity of related bacterial enzymes did not disrupt *psaA*

Table 1. Nuclear-encoded proteins involved in plastid RNA splicing

Protein	Target Intron	Protein Class	Species	References
Raa1	psaA introns 1 and 2		*C. reinhardtii*	(Merendino et al. 2006)
Raa2	psaA intron 1	Pseudouridine synthase	*C. reinhardtii*	(Perron et al. 1999)
Raa3	psaA intron 2		*C. reinhardtii*	(Rivier et al. 2001)
Rat1	psaA intron 1 (tscA)	NAD+ binding	*C. reinhardtii*	(Balczun et al. 2005)
CRS1	atpF	CRM domain	*Z. mays* *A. thaliana*	(Jenkins et al. 1997; Till et al. 2001; Ostersetzer et al. 2005; Asakura and Barkan 2006)
CAF1	petD, trnG, rps16, rpl16, ycf3-intron 1,rpoC1*, clpP-intron 1*	CRM domain	*Z. mays* *A. thaliana*	(Ostheimer et al. 2003; Asakura and Barkan 2006)
CAF2	rps12- intron 1, petB, ndhB, ndhA, ycf3- intron 1	CRM domain	*Z. mays* *A. thaliana*	(Ostheimer et al. 2003; Asakura and Barkan 2006)
CRS2	All CAF1- and CAF2-dependent introns	Peptidyl-tRNA hydrolase	*Z. mays*	(Jenkins et al. 1997; Vogel et al. 1999; Jenkins and Barkan 2001)
PPR4	rps12-intron 1	PPR and RRM	*Z. mays*	(Schmitz-Linneweber et al. 2006)
HCF152	petB	PPR	*A. thaliana*	(Meierhoff et al. 2003; Nakamura et al. 2003)

* Introns present in *Arabidopsis* but not in maize

splicing *in vivo*, suggesting that pseudouridine synthase activity is not relevant to Raa2's role in splicing (Perron et al. 1999). Raa1 encodes a novel protein that includes repeated motifs that are reminiscent of tetratricopeptide (TPR) and pentatricopeptide (PPR) motifs (Merendino et al. 2006); it was speculated that these repeats might form an RNA binding surface analogous to the surface proposed for PPR proteins (Small and Peeters 2000); in fact, Raa1 resides in two high molecular weight complexes in chloroplasts, one of which contains RNA (Merendino et al. 2006). Mutational studies revealed that Raa1's C-terminal domain functions in the processing of the *tscA* RNA and the splicing of *psaA* intron 1, whereas the

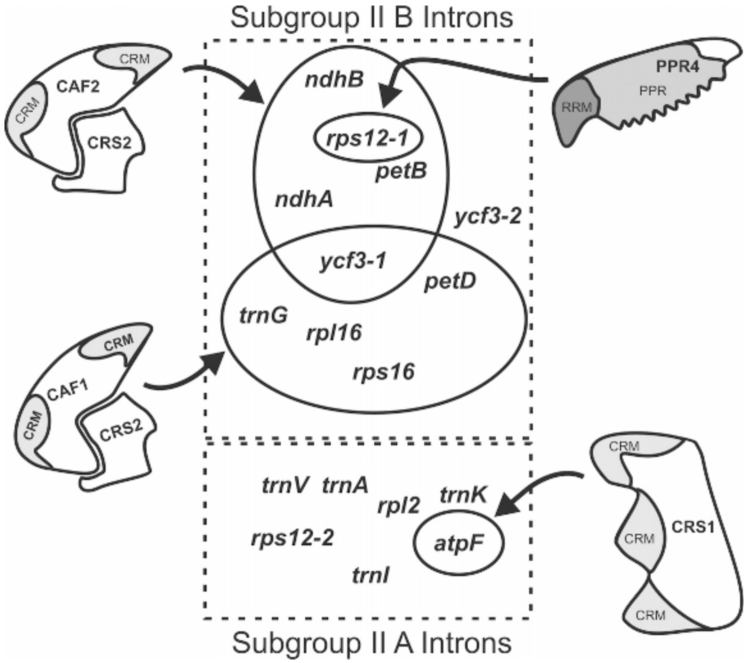


Fig. 1. Nucleus-encoded chloroplast splicing factors and their intron targets in maize. The intron targets indicated for each protein fail to splice in the corresponding mutant background and coimmunoprecipitate with the corresponding protein. CAF1, CAF2, and CRS1 are members of the CRM domain protein family (Barkan et al. 2007), CRS2 is a peptidyl-tRNA hydrolase homolog (Jenkins et al. 2001), and PPR4 contains a PPR tract and an RRM domain (Schmitz-Linneweber et al. 2006). All seven subgroup IIA introns fail to splice in mutant plastids lacking ribosomes, implicating a plastid translation product in their splicing. Results are summarized from Jenkins et al. 1997, Ostheimer et al. 2003, Vogel et al. 1999, Schmitz-Linneweber et al. 2005 and 2006. The functions of the *Arabidopsis* CAF1, CAF2, and CRS1 orthologs are conserved with those in maize, except that AtCAF1 promotes the splicing of two additional chloroplast introns that are not found in maize (*rpoC1* and *clpP*-intron 1; Asakura et al. 2006).

central domain mediates splicing of the second intron. Thus, Raa1 may serve to coordinate the two *trans*-splicing events during *psaA* maturation. This coordination may involve transient association between the predominantly stromal Raa3-containing complex and the predominantly membrane-bound Raa1/Raa2 complexes.

Analogous approaches have been used to identify nucleus-encoded proteins involved in the splicing of chloroplast introns in land plants. In maize, five proteins involved in the splicing of various subsets of its 17 chloroplast group II introns have been reported: CRS1, CRS2, CAF1, CAF2, and PPR4. For each of these proteins, splicing defects accompanying loss-of-function mutations have identified its intron targets, and RNA coimmunoprecipitation experiments have shown it to be associated *in vivo* with the corresponding intron RNAs (Jenkins et al. 1997; Jenkins and Barkan 2001; Till et al. 2001; Ostheimer et al. 2003; Schmitz-Linneweber et al. 2005b, 2006). Together, these results provide strong evidence for a direct role in splicing. These proteins are found in at least three distinct ribonucleoprotein complexes, all in the chloroplast stroma. CRS2 functions in complexes that contain either CAF1 or CAF2 to promote the splicing of nine introns, with CAF1 and CAF2 each required for the splicing of an overlapping subset of the CRS2-dependent introns (Table 1, Fig. 1). The CRS2-CAF complexes are bound to intron RNAs in the stroma, in complexes of ~500-600 kDa (Jenkins et al. 1997; Jenkins and Barkan 2001; Ostheimer et al. 2003; Schmitz-Linneweber et al. 2005b). CRS1 is required solely for the splicing of the *atpF* intron and is found in a distinct high molecular weight (~600-700 kDa) ribonucleoprotein complex that includes *atpF* intron RNA (Jenkins et al. 1997; Till et al. 2001; Ostheimer et al. 2003). PPR4 is required solely for the *trans*-splicing of the first intron of *rps12* and resides in stromal complexes that are heterogeneous in size, and that include both fragments of the split *rps12* intron (Schmitz-Linneweber et al. 2006). As noted above, a plastid translation product, possibly MatK, is required for the splicing of several chloroplast introns as well (Jenkins et al. 1997; Vogel et al. 1999). Taken together, the genetic data show that sixteen of the seventeen group II introns in maize chloroplasts rely on proteins for their splicing *in vivo* (Fig. 1). The splicing of the second intron in *ycf3* is not disrupted in any of the mutant backgrounds analyzed to date and is the only candidate for a truly self-splicing group II intron in the maize chloroplast genome.

The chloroplast splicing factors discovered in maize are unrelated to those identified in *C. reinhardtii*, which perhaps is not surprising, given the independent origin of chloroplast introns in land plants (e.g. maize) and chlorophyte algae (e.g. *C. reinhardtii*). CRS2 has strong sequence and structural similarity to bacterial peptidyl-tRNA hydrolases, but seems to lack peptidyl-tRNA hydrolase activity (Jenkins and Barkan 2001; Ostheimer et al. 2005). CRS1, CAF1, and CAF2 are related to one another in that they harbor several copies of the same conserved domain, which is represented as a stand-alone ORF in prokaryotes (Till et al. 2001; Ostheimer et al. 2003). The *E. coli* representative of this domain family, YhbY, is bound *in vivo* to 50S ribosomal subunit precursors and likely plays a role in ribosome maturation (Barkan et al. 2007); thus, the YhbY-like domain in the chloroplast splicing factors was named the chloroplast RNA splicing and ribosome maturation (CRM) domain (Ostheimer et al. 2003; Barkan et al. 2007). Structural and biochemical data show that CRM domains function as RNA binding domains: the crystal structure of YhbY shows structural similarity with a class of RNA binding proteins that includes IF3 (Ostheimer et al. 2002), and an isolated CRM domain from CRS1 binds RNA with high affinity *in vitro* (Barkan et al. 2007).

CRM domains are found in a protein family in vascular plants comprising 16 members in *Arabidopsis* and 14 members in rice (Barkan et al. 2007). A reverse-genetic approach in *Arabidopsis* showed that the splicing functions and intron specificities of the CRS1, CAF1, and CAF2 members of the CRM family are conserved between maize and *Arabidopsis*, indicating that these proteins were recruited to promote the splicing of plastid group II introns prior to the divergence of monocot and dicot plants (Asakura and Barkan 2006). The *Arabidopsis* CAF1 ortholog has additional functions in that it promotes the splicing of introns in *rpoC1* and *clpP*, which are found in *Arabidopsis* but not in maize (Asakura and Barkan 2006). Given that all three characterized members of the plant CRM family function in chloroplast group II splicing, it seems likely that additional group II intron splicing factors remain to be discovered among the uncharacterized CRM proteins. In fact, a CRS1 paralog has been shown to be bound to several group II introns in maize chloroplasts, and to be required for the splicing of the corresponding introns in *Arabidopsis* (Y. Asakura and A. Barkan, manuscript in preparation).

The PPR protein family, like the CRM family, is largely specific to plants and includes members that function in the splicing of chloroplast introns. PPR proteins are defined by tandem repeats of a degenerate 35 amino acid motif that is related to the TPR motif. The repeat tracts have been proposed to form an RNA-binding surface that is structurally similar to the protein-binding surface described for TPR tracts (Small and Peeters 2000). The maize protein PPR4 contains both a PPR tract and an RRM motif, and is required for the *trans*-splicing of the first intron in the chloroplast *rps12* mRNA, to which it is bound *in vivo* (Schmitz-Linneweber et al. 2006). *Arabidopsis* HCF152, another PPR protein, is required for the accumulation of spliced chloroplast *petB* RNA (Meierhoff et al. 2003) and binds *in vitro* to the *petB* precursor transcript (Nakamura et al. 2003); excised *petB* intron accumulates normally in *hcf152* mutants, however, suggesting that HCF152 may function to stabilize spliced *petB* mRNA rather than to promote splicing.

The nucleus-encoded chloroplast splicing factors described thus far are diverse in sequence and evolutionary origin, but a common theme is their derivation from RNA binding proteins that evolved in other contexts. For example, Raa2 was derived from a pseudouridine synthase (Perron et al. 1999), CRS2 was derived from a peptidyl-tRNA hydrolase (Jenkins and Barkan 2001), and CRS1, CAF1, and CAF2 were derived by duplication and diversification of a pre-ribosome binding protein (Barkan et al. 2007). This situation is mirrored in fungi, where derived tRNA synthetases promote the splicing of both group I and group II introns (reviewed in Lambowitz et al. 1999). The differences between these splicing factors and their ancestors can elucidate features responsible for their gain of intron splicing functions. For example, CRS2 maintains a three-dimensional structure that is highly similar to that of its peptidyl-tRNA hydrolase ancestor, but several amino acid substitutions result in a CRS2-specific hydrophobic surface that allows CRS2 to bind to its CAF1 and CAF2 partners (Ostheimer et al. 2005). Conversely, CAF1 and CAF2 acquired the corresponding CRS2-interaction motif: an amphipathic helix appended to their CRM domains that is lacking in their most closely-related paralogs (Ostheimer et al. 2006). These examples highlight how proteins with novel functions can emerge through minor evolutionary tinkering.

2.3.4 Biochemical functions of chloroplast splicing factors

Despite recent progress in identifying chloroplast splicing factors, little is known about the mechanisms by which they promote splicing. It is generally assumed that the catalytic activity of group I and group II introns is intrinsic to the intron RNAs, and that proteins facilitate their splicing by enhancing the productive folding of the introns into their catalytically-active structure. The folding of group I and group II introns, like that of other large and highly structured RNAs, is problematic because numerous non-native conformations are similar in stability to the active structures, so the RNAs can easily be trapped in inactive conformations (reviewed in Herschlag 1995; Weeks 1997). In addition, tertiary interactions that establish the three-dimensional architecture of the intron can be weak (Swisher et al. 2002). Proteins could potentially guide intron folding via high-affinity, sequence-specific interactions that stabilize an otherwise transient tertiary interaction, or that preclude competing non-productive folding pathways. Alternatively, proteins could act as "RNA chaperones" to resolve misfolded RNA structures through low-affinity non-specific interactions with unstructured RNA, or via an ATP-dependent helicase activity (Herschlag 1995; Lorsch 2002; Halls et al. 2007). The handful of group I and group II splicing factors that have been studied in detail (all from non-plant systems) act by promoting intron folding (reviewed in Lambowitz et al. 1999; Lambowitz and Zimmerly 2004), but it is likely that study of the diverse introns found in plant organelles will reveal additional mechanisms. For example, in the special case of *trans*-spliced introns, proteins such as PPR4, Raa1, Raa2, and Raa3 might assist in the assembly of intron fragments.

Among chloroplast splicing factors, details of protein-intron interactions have been reported only for the CRM-domain protein CRS1. CRS1 appears to function via the first of the general mechanisms outlined above, as it binds *in vitro* with high affinity and specificity to specific sequences in domains 1 and 4 of its *atpF* intron substrate (Ostersetzer et al. 2005). The results of hydroxyl-radical footprinting suggested that CRS1 binding promotes the internalization of intron elements that are expected to be at the core of the functional ribozyme. Thus, by making high-affinity contacts with two peripheral intron segments, CRS1 seems to act like a molecular scaffold to enhance the productive folding of internal intron segments (Ostersetzer et al. 2005).

In contrast to CRS1, CRS2 does not bind with high affinity to its target introns *in vitro* (Barkan lab, unpublished observations); therefore, it seems likely that CRS2 is recruited to specific introns via its interactions with its CRM-domain partners CAF1 and CAF2. A hypothesis for CRS2's role in splicing was suggested by the structure and sequence of its derived peptidyl-tRNA hydrolase active site, which is highly conserved in CRS2 despite the fact that CRS2 did not exhibit peptidyl-tRNA hydrolase activity when expressed in *E. coli* (Jenkins and Barkan 2001; Ostheimer et al. 2005). These observations suggest the intriguing possibility that the ancestral active site may have been subtly modified in CRS2 so that it now contributes to a chemical step in splicing.

2.4 The regulation of chloroplast RNA splicing

RNA splicing is essential for the expression of intron-containing genes, and is therefore a potential regulatory step in chloroplast gene expression. In fact, unspliced chloroplast transcripts typically accumulate to high levels, so changes in splicing efficiency are likely to be reflected by changes in the abundance of mature transcripts. Tissue-dependent changes in the ratio of spliced to unspliced chloroplast RNAs have been described for the maize *atp*F, *pet*D*, pet*B*, rpl*16, and *ycf3* introns (Barkan 1989; McCullough et al. 1992), and for the mustard *trn*G intron (Liere and Link 1994). In each of these cases, a higher proportion of transcripts is spliced in mature chloroplasts than in immature chloroplasts or in non-photosynthetic plastid forms, consistent with the possibility that an increase in splicing rates early in chloroplast development contributes to the burst in synthesis of chloroplast-encoded subunits of the photosynthetic apparatus. Although light has no apparent effect on the splicing of several chloroplast introns in vascular plants (Barkan 1989; Liere and Link 1994), light does activate the splicing of the group I introns in the *C. reinhardtii* chloroplast *psb*A gene (Deshpande et al. 1997).

These observations suggest that splicing can be developmentally regulated in plants and light-regulated in *C. reinhardtii*. Still, varying ratios of spliced versus unspliced RNAs do not prove that the rate of splicing is regulated, as this could also result from changes in the stability of the unspliced precursor with respect to that of its spliced product. Even if plastid splicing rates do change, these changes will be regulatory only if the level of spliced mRNA limits the ultimate accumulation of the protein product. Some chloroplast mRNAs are in excess of the amount needed for maximal translation in *C. reinhardtii* (Eberhard et al. 2002), so small decreases in the synthesis of these mRNAs are not anticipated to impact the level of their gene product. Nonetheless, a mutation in a group I intron in the *C. reinhardtii psbA* pre-mRNA caused a twofold reduction in both the level of spliced mRNA and the rate of PsbA protein synthesis (Lee and Herrin 2003), indicating that, at least for this mRNA, small changes in splicing efficiency effectively change the rate of synthesis of the corresponding protein.

A variety of mechanisms could potentially regulate splicing in chloroplasts. One obvious possibility is that the synthesis or activity of nucleus-encoded chloroplast splicing factors is regulated, which in turn, regulates the splicing of chloroplast introns. Unfortunately, few studies have attempted to correlate changes in the abundance of chloroplast splicing factors with changes in the splicing of their substrate RNAs; in fact, only for CRS1 has such a correlation been demonstrated (Till et al. 2001). A protein-independent mechanism for splicing regulation could involve developmentally-regulated changes in stromal [Mg^{2+}], as the folding and catalysis of group I and group II introns is dependent on Mg^{2+} (Pyle 2002) and the concentration of free Mg^{2+} rises during chloroplast maturation in spinach (Horlitz and Klaff 2000). It is also plausible that chloroplast splicing in vascular plants could change during development as a consequence of the developmental switch in the plastid transcription machinery. A nucleus-encoded phage-like polymerase (NEP) predominates early in chloroplast development, whereas a chloroplast-encoded bacterial-like RNA polymerase (PEP) predominates in mature chloroplasts (reviewed in Weihe 2004). Based on the properties of the phage and bacterial polymerases to which these enzymes are related (Iost et al. 1992), it is likely that NEP elongates more quickly than PEP. A more rapid transcription elongation rate might hinder the productive folding of chloroplast introns by reducing the length of the kinetic window during which non-native RNA partners are excluded from interaction with nascent intron segments. This general model could be tested by comparing splicing efficiencies in engineered tobacco chloroplasts that express the same intron-containing gene driven by either a NEP or PEP promoter.

2.5 Perspective

The findings summarized here raise numerous interesting questions: By what mechanisms do splicing factors promote the activity of chloroplast introns? How is chloroplast splicing coordinated with other steps in chloroplast gene expression and assembly of the photosynthetic apparatus? Is the rate of splicing in chloroplasts subject to regulation, and if so, how is this regulation accomplished? Did the "need" to promote the splicing of intrinsically poor chloroplast introns spur the evolution of plant-specific protein families such as the CRM and PPR families? What is the nature of the coevolutionary processes through which the degeneration of "self-splicing" group I and group II introns has been compensated by the recruitment and modification of pre-existing proteins to participate in splicing?

A thorough understanding of these issues cannot emerge without a more complete catalog of the proteins involved in the splicing of chloroplast introns. Results to date suggest a complexity that was not anticipated based on studies of protein-facilitated splicing in non-plant systems, where a single protein has, in several instances, been shown to be sufficient to reconstitute protein-facilitated splicing *in vitro* (reviewed in Lambowitz et al. 1999; Lambowitz and Zimmerly 2004). Reconstitution of protein-facilitated splicing of chloroplast introns has not been reported, suggesting that essential factors remain to be identified. Indeed, the large size of the particles harboring chloroplast intron RNAs and splicing factors *in vivo* cannot be accounted for by the identified components. Moreover, genetic screens for chloroplast splicing factors in land plants and *C. reinhardtii* are not yet saturating, and the genes underlying several known splicing mutants in maize and *Chlamydomonas* have not been identified (Goldschmidt-Clermont et al. 1990, Barkan lab, unpublished). Candidate gene approaches can be anticipated to play an increasingly important role in the efforts to identify more splicing factors. Candidates for reverse genetic analyses include nucleus-encoded group II maturase homologs in land plants (Mohr and Lambowitz 2003), paralogs of the plant CRM-domain splicing factors, and members of the PPR family: with more than 450 PPR proteins encoded in the genomes of vascular plants (Lurin et al. 2004), the PPR family constitutes a large pool of potential chloroplast splicing factors. Biochemical methods such as affinity purification of proteins that associate with known splicing factors and activity-based protein purifications can complement these efforts; indeed, two proteins were purified from *C. reinhardtii* chloroplasts by virtue of their ability to bind *in vitro* to chloroplast group II intron RNAs (Balczun et al. 2006; Glanz et al. 2006).

Studies addressing evolutionary questions are also limited by the restricted knowledge of organellar splicing machineries. It can be anticipated that the functions of splicing factors identified in one land plant species will generally be conserved in other land plants, as has been demonstrated for maize and *Arabidopsis* CRS1, CAF1, and CAF2 (Asakura and Barkan 2006). However, the more interesting questions concern the evolution of these proteins: does the emergence of specific splicing factors coincide with the appearance of the chloroplast genome organization that is characteristic of land plants, and can factors present in *Chlamydomonas* still be found in basal taxa of land plants? The availability of nuclear genome sequences of various "lower" plants will be necessary to address these questions.

Finally, to understand the role of regulated splicing in chloroplast function, it will be necessary to more thoroughly catalog changes in chloroplast splicing under various conditions, to correlate the levels of the known splicing factors with these changes, and to generate engineered organisms in which the abundance of specific splicing factors can be manipulated such that their effects on protein output can be assessed. Recent advances in the tools available for genetic and genomic analyses in chloroplast-bearing organisms should enhance progress in understanding these issues during the coming years.

3 Plastid RNA editing

RNA splicing is embedded in a series of additional RNA processing events, among them RNA editing - the modification of single ribonucleotides such that the RNA sequence does not match that of its DNA template. Indeed, a link between splicing and editing has been demonstrated for an exonic nucleotide in *ndhA* close by the 3' intron/exon border, such that only spliced mRNAs are edited (Schmitz-Linneweber et al. 2001). However, few chloroplast introns have been analyzed for editing, and it is unclear whether intron-internal editing events are instrumental in the splicing of any chloroplast introns (e.g. Bonen and Vogel 2001; Vogel and Borner 2002; Kugita et al. 2003b). Much more information is available regarding the impact of RNA editing on exonic sequences.

After the initial discovery of RNA editing in trypanosome mitochondria (Benne et al. 1986), various examples of RNA editing were described in organisms from diverse taxa (Gott and Emeson 2000). These encompass a variety of alterations of RNA primary sequence that arise from base modifications, nucleotide insertions or deletions, and nucleotide replacements. Many of the editing processes discovered to date employ widely different mechanisms and are therefore believed to be polyphyletic (Smith et al. 1997; Gott and Emeson 2000).

In chloroplasts, RNA editing is restricted to nucleotide conversions (for recent reviews, see Bock 2000; Wakasugi et al. 2001; Shikanai 2006). Only changes from C to U or – less frequently – from U to C have been observed so far. This type of RNA editing usually affects the coding potential of the mRNA. Like any other RNA editing system, plastid conversional editing depends on *cis*-acting sequences that determine the base to be edited and *trans*-acting factors that carry out site recognition and catalysis. Since its discovery, substantial progress has been made on understanding the *cis*-elements. In contrast, despite 15 years of research, very little is known about the executing machinery.

3.1 Editing sites impact protein function

Overwhelmingly, RNA editing restores evolutionarily conserved codons and thus conserved amino acids (Hirose et al. 1996; Inada et al. 2004; Tillich et al. 2005). Moreover, the most commonly observed codon conversions lead to amino acid substitutions that differ pronouncedly in their physico-chemical properties. In several instances it was found that if editing does not occur, the affected protein is severely impaired or altogether non-functional. For example, substitution of the unedited spinach-specific *psbF* editing site for the corresponding sequence in the *psbF* gene of tobacco led to tobacco plants in which the spinach editing site remained unprocessed and that had compromised photosynthesis (Bock et al. 1994). Presumably the aberrant PsbF protein encoded by this engineered gene led to reduced activity of photosystem II, of which PsbF is a subunit (Bock et al. 1994). Analogously, the introduction of the non-edited form of maize *petB* into the chloroplast genome of *Chlamydomonas*, which shows no editing at all, led to strains that were non-phototrophic, consistent with a lack of cytochrome b_6f activity (Zito et al. 1997). The mutant phenotype was due to defective assembly of cytochrome b_6f complexes, of which PetB is a subunit, confirming that the edited codon is essential for the functional interactions of PetB with that complex (Zito et al. 1997). Also, the carboxyltransferase subunit of the acetyl-CoA carboxylase showed no activity *in vitro* when translated from a message containing an unprocessed editing site from pea (Sasaki et al. 2001). In tobacco, mutation of an edited serine codon to a tryptophan codon in the plastid *atpA* gene led to albino plants, suggesting that this codon is essential for the function of the encoded alpha subunit of the plastid ATPase (Schmitz-Linneweber et al. 2005a). RNA editing sometimes is also necessary to create an initiation codon for translation; in such cases, it seems self-evident that the editing event plays an essential role in translation. This expectation was confirmed for the *ndhD* transcript in a tobacco *in vitro* translation system: only the edited version of the *ndhD* transcript gave rise to NdhD protein (Hirose and Sugiura 1997). Loss of editing at this site obliterated the function of the NDH complex, of which NdhD is a subunit (Okuda et al. 2006).

In summary, RNA editing is crucial for the proper expression or function of the encoded protein in every case in which this has been analyzed. This implies that edited codons generally code for amino acids that are critical for protein function.

3.2 Mechanism of RNA editing

3.2.1 Biochemistry

Initially, RNA editing was investigated by two methods: 1) direct sequencing of RNA and 2) sequencing of cloned or PCR-amplified cDNA. Thus, whether the base resulting from editing is a U or a modified C that reverse transcriptase recognizes as a U, was unclear. More recently, however, *in vitro* editing techniques and single strand conformation polymorphism assays allowed the unequivocal demonstration that uridine bases are the product of editing (Fuchs et al. 2001; Hirose and Sugiura 2001). The next question, then, was how the U is produced from the genomically encoded C.

Three reactions could in principle underlie C-to-U and U-to-C conversions: *trans*-amination, base-exchange (transglycosylation), or nucleotide replacement. In plastids, biochemical data on enzymatic aspects of RNA editing are scarce, but in plant mitochondria, which are believed to have a phylogenetically related RNA-editing system (Maier et al. 1996), it seems that the N-glycosidic bond between the ribose and the pyrimidine base remains intact (Yu and Schuster 1995). Also, in both organelles, the sugar-phosphate backbone remains untouched by editing (Rajasekhar and Mulligan 1993; Hirose and Sugiura 2001). This clearly speaks against a nucleotide excision mechanism and has led to a search for cytidine deaminases or transaminases - that is, enzymes that modify the bases while leaving the RNA backbone intact- as editing enzymes.

Although *cis*-acting sequence requirements are defined for several editing sites (see Section 3.4), it is unclear whether these sequences are presented as single-stranded RNA, part of a stem-loop or as double-stranded RNA. This question is of particular interest in light of the editing system of trypanosome mitochondria, which uses complementary guide RNAs to direct editing events. Experiments involving antisense RNAs to the tobacco chloroplast editing site *rpoB*-2 (Hegeman et al. 2005a) suggested that the edited site itself must be single-stranded whereas the adjacent *cis*-element can be either single-stranded or double-stranded. Potential guide RNAs and complementary sections inside the same transcript were computationally predicted for tobacco and hornwort chloroplasts (Bock and Maliga 1995; Yoshinaga et al. 1997), but mutation of a putative guide RNA did not inhibit RNA editing (Bock and Maliga 1995).

In *in vitro* editing systems, processing of artificial templates is highly dependent on the magnesium concentration and on the presence of hydrolysable NTP (Hirose and Sugiura 2001; Hegeman et al. 2005b; Nakajima and Mulligan 2005). This reliance on an external energy source sets plastid RNA editing apart from other C-to-U editing systems like the mammalian APOBEC system, which functions *in vitro* without added nucleotides (Driscoll et al. 1989), and may point to the involvement of an ATP-dependent RNA helicase in plastid RNA editing (Nakajima and Mulligan 2005). Both APOBEC and chloroplast editing are dependent on free zinc (Navaratnam et al. 1993; Bhattacharya et al. 1994; Hegeman et al. 2005b). Whether this means that the chloroplast editase, like the APOBEC enzyme, is a zinc-dependent cytidine-deaminase, remains to be established.

3.2.2 Kinetics

It is unclear at what point during transcript maturation RNA editing occurs or whether this is uniform among different edited transcripts. Potentially, RNA editing could be co-transcriptional, either via the incorporation of U instead of C by

RNA polymerase, or by the action of a cytidine deaminase in close contact with the nascent transcript. Alternatively, editing might occur on transcripts that have been already released from the polymerase. No definite answer regarding this question can be given at the moment, but it is clear that editing is highly efficient as most sites are fully or almost completely edited (e.g. Maier et al. 1995; Hirose et al. 1999; Peeters and Hanson 2002). There are several exceptional editing sites, however, for which a large pool of unedited RNAs accumulates (Hirose et al. 1999; Peeters and Hanson 2002; Schmitz-Linneweber et al. 2002; Inada et al. 2004; Tillich et al. 2005). Taken together, this shows that for most sites, the capacity of the editing machinery is sufficient to cope with template abundance. Whether this means that nascent transcripts or fully transcribed, released transcripts are the substrates for editing remains an open question.

Another interesting question concerns how editing relates to other processing events like RNA splicing or endonucleolytic cleavage. Some editing sites seem entirely independent of other processing steps. For instance, the *petB* and *ycf3* transcripts were fully edited regardless of whether they were spliced or still part of a polycistronic precursor (Freyer et al. 1993; Ruf et al. 1994). Other sites show a strong or even obligate link to another processing step occurring on the same precursor. This is true for editing of the *rpl2* initiation codon, which is complete in mature, spliced RNA molecules but which is rare in uncleaved and unspliced *rpl2-rpl23* precursor molecules (Freyer et al. 1993). Similarly, unspliced *ndhA* mRNAs in spinach were not edited at all (Schmitz-Linneweber et al. 2001). For *ndhD* in *Allium porrum*, RNA editing is linked to intercistronic cleavage between *psaC* and downstream *ndhD* (Del Campo et al. 2002). The translational status of an mRNA can influence editing as well: heat induced reduction of plastid translation or mutational loss of plastid ribosomes leads to a reduction in editing efficiency at specific sites (Zeltz et al. 1993; Hess et al. 1994; Karcher and Bock 1998; Nakajima and Mulligan 2001; Karcher and Bock 2002b; Halter et al. 2004). It is unclear, however, whether plastid translation acts indirectly via (i) the synthesis of a translation product that functions in editing, (ii) the co-translational recruitment of editing factors, or (iii) a change in transcript abundance via alteration of the PEP/NEP ratio, or acts directly by influencing the accessibility of the editing site.

In summary, the timing of RNA editing events relative to other RNA maturation steps is specific to each site. While at least for some sites, a link between editing and translation, splicing and/or endonucleolytic cleavage has been established, editing of other sites appears to be indifferent to the processing state of the RNA. Data on how links between editing and other steps in gene expression might be reflected by commonalities among the different processing machineries is lacking.

Table 2. Requirements of *cis*-sequences for editing sites *in vitro* and *in vivo*.

site [a]	*in vivo* (v); *in vitro* (r)	species	cis-element for RNA editing [b]	Reference
ndhB-156	r	tobacco	-10 to -5	(Sasaki et al. 2006)
ndhB-246	r	tobacco	-22 to +9	(Hirose and Sugiura 2001)
ndhB-246	v	tobacco	-12 to -2	(Bock et al. 1997)
ndhB-249	v	tobacco	-21 to -11	(Bock et al. 1997)
ndhB-277	v	tobacco	-42 to +48	(Bock et al. 1996)
ndhB-279	v	tobacco	-48 to +42	(Bock et al. 1996)
ndhF-97	r	tobacco	-15 to +5 plus -40 to -35	(Sasaki et al. 2006)
petB-204	r	pea	-20 to -1	(Miyamoto et al. 2002; Nakajima and Mulligan 2005)
petB-204	r	tobacco	-20 to -1	(Miyamoto et al. 2002)
psbE-72	r	pea	-15 to -1	(Miyamoto et al. 2002)
psbE-72	r	tobacco	-15 to -1	(Miyamoto et al. 2002)
psbE-72	r	*Arabidopsis*	-13 to +15	(Hegeman et al. 2005b;
psbL-1	v	tobacco	-16 to +5	(Chaudhuri and Maliga 1996)
psbL-1	r	tobacco	-22 to +9	(Hirose and Sugiura 2001)
rpoB-158	v	tobacco	-20 to +6	(Reed et al. 2001b)
rpoB-158	r	tobacco	-27 to +11	(Hayes et al. 2006)
rpoB-158	r	tobacco [c]	-31 to +61	(Hayes et al. 2006)
rpoB-158	r	maize	-27 to +11	(Hayes et al. 2006)

[a] Numbers refer to the codon affected by RNA editing in tobacco (not necessarily the same in the other species listed here)
[b] Not all listed elements have been mapped down to the minimal *cis*-sequence required for editing; in most cases, the indicated sequence range defines a core element sufficient for editing, which is not to say that all nucleotides of an element are also essential for editing, and which also does not exclude that longer templates lead to higher editing efficiencies
[c] Template from maize

3.3 *cis*-elements involved in plastid RNA editing

Efforts to identify *cis*-elements for RNA editing started from the hypothesis that sequences surrounding the nucleotide to be edited participate in its recognition by *trans*-factors. For example, position -1 is likely to be critical for the editing of most mRNAs, since 29 of 31 tobacco editing sites include pyrimidines at this position (Maier et al. 1992a, 1992b; Hirose et al. 1999). Moreover, editing of *ndhB* mRNAs (site V) was impaired if the U at position -1 was converted to a G (Bock et al. 1996), confirming that bases adjacent to editing sites do play a role in the reaction. Several studies demonstrated that, in addition to the upstream nucleotide a minimum sequence context is necessary and sufficient to direct editing (summarized in Table 2). The early studies of this nature involved laborious *in vivo* experiments, and revealed that the recognition of most editing sites relies on short sequences immediately upstream of the edited site, most of them less than 20 nt long (Chaudhuri et al. 1995; Bock et al. 1996; Chaudhuri and Maliga 1996; Reed et al. 2001b). No consensus sequence could be identified for these sites or any other sequences 5' to editing sites, nor could a consensus secondary structure be identified. Recently however, inter-site homologies were found in the 15 nt upstream of editing sites when all editing sites of *A. capsillus-veneris* and *A. formosae* were compared (Tillich et al. 2006a). These homologies do not allow generation of a consensus for all sites but rather point to small clusters of similar sites, at least in angiosperms (Chateigner-Boutin and Hanson 2002; Chateigner-Boutin and Hanson 2003; Tillich et al. 2005). Indirect evidence suggests that the members of each cluster of related *cis*-sequences are recognized by the same *trans*-factor (see Section 3.4).

Recently, *in vitro* editing systems have become available for four species: tobacco (Hirose and Sugiura 2001), pea (Miyamoto et al. 2002; Nakajima and Mulligan 2005), maize (Hayes et al. 2006), and *Arabidopsis* (Hegeman et al. 2005b). They have been used to dissect *cis*-elements at higher resolution. These studies confirmed the predominant role of 5' sequences over 3' sequences for determining editing efficiency, and showed further that nucleotides inside the *cis*-element do not contribute equally to editing. In particular, the nucleotides immediately preceding the editing site (one to four depending on the specific site) and the editing site itself are not essential for binding of the *trans*-factor(s), although they are required for the reaction itself (Miyamoto et al. 2002). Closer inspection of the proximal bases for the two editing sites in *psbL* and *petB* revealed that the sequence of these elements is recognized in a highly specific manner (Miyamoto et al. 2004). Thus, both binding of the site as well as catalysis after binding require sequence-specific interactions.

In addition to sites that require a short sequence element immediately upstream of the edited C, there are also reports of more complex *cis*-elements. For instance, the *cis*-element of the editing site in the tobacco *ndhF* mRNA is bipartite, with essential elements spaced by 19 nt (Sasaki et al. 2006). For other sites, editing efficiency increases with longer templates, although additional elements outside the usual -20 to +6 core are not essential (Hayes et al. 2006). Rarely, though, more distant putative elements can be essential as suggested by editing site *ndhB*-2 and -3, which were not edited *in vivo* despite 42 nt of both 5' and 3' adjacent sequences (Bock et al. 1996). In this context it is interesting that the more distant context of an editing site can determine how critical point mutations in the core element are: a point mutation 20 nt upstream of editing site *rpoB*-158 abolished editing in a construct stretching from -27 to + 6, but had little effect in a construct only little longer (-31 to +22, Hayes et al. 2006). This suggests that editing sites with short essential *cis*-elements have additional, non-essential elements farther away from the editing site that can compensate for mutations in the core elements. In fact, most editing sites are poorly edited *in vitro*, not reaching efficiencies greater than 10% (Sasaki et al. 2006) despite the high editing efficiency *in vivo*. This is also true for most studies in which short sequences around editing sites were introduced into chloroplasts by biolistic transformation. In these experiments, editing of the short transgenes was low (Bock et al. 1996; Reed et al. 2001a). Whether this low editing efficiency is solely due to the overexpression of introduced editing sites, which overburdened the editing apparatus, or whether important distal sequence elements were lacking in these constructs remains to be determined.

In summary, essential elements for RNA editing are mostly situated immediately 5' to the edited site, but this does not exclude the possibility that additional elements contribute to editing efficiency. Given that the translational and processing status of the edited message contributes to editing efficiency (see Section 3.2), it seems likely that further sequence elements will play into determining editing efficiency.

3.4 *trans*-factors involved in plastid RNA editing

The finding that there is no clear consensus for editing site recognition led to the proposal that each site is served by its own specific factor, presumably an RNA binding protein. This was supported by titration studies, in which overexpression of an introduced site leads to a reduction in editing of the endogenous site, but not of any other site examined (Chaudhuri et al. 1995). Later, this conclusion was modified due to the finding that smaller clusters of related editing sites may exist and that titration of one factor could affect several related sites (Chateigner-Boutin and Hanson 2002, 2003). However, given the small size of these clusters (usually two to three sites), a substantial set of factors would still be needed to serve all sites. The nature of these factors has long been elusive. Experiments transferring plastids between different species demonstrated that at least some of these specificity factors are nuclear-encoded (Bock and Koop 1997). This finding was not entirely unexpected as the well-annotated and small plastid chromosome was unlikely to code for dozens of hitherto unidentified editing factors. Still, small RNAs functioning as editing specificity factors might be hidden in the chloroplast genome. In trypanosome mitochondria, small so-called guide RNAs (gRNAs) form Watson-Crick base pairs with pre-mRNAs thereby specifying RNA editing sites. To assess the involvement of gRNAs in chloroplast RNA editing, tobacco *in vitro* editing extracts were treated with RNAse. This did not abolish editing activity of the treated extracts, which suggests that editing factors are not ribonucleic acids but rather of a proteinacious nature (Hirose and Sugiura 2001). Attempts to identify putative guide RNAs by crosslink strategies were also unsuccessful (Hirose and Sugiura 2001), which further strengthens the supposition that it is proteins rather than RNAs that do the main job in plastid RNA editing (see Section 3.1).

A first major advance in identifying *trans*-factors for RNA editing came from studies on proteins bound in the sequence environment of editing sites. Using a stromal extract competent for *in vitro* editing, Sugiura and colleagues were able to UV-crosslink several proteins to short bait-RNAs containing editing sites (Hirose and Sugiura 2001). First, they found the highly abundant cpRNP proteins, which contain two RNA recognition motifs (RRMs) and an additional acidic domain.

Table 3. Potential and confirmed editing factors

factor	target site [a]	species	references
cp31	psbL-1, ndhB-246	tobacco	(Hirose and Sugiura 2001)
CRR4	ndhD-1	*Arabidopsis*	(Kotera et al. 2005)
p25	psbL-1	tobacco	(Hirose and Sugiura 2001)
p70	psbE-72	tobacco/pea	(Hirose and Sugiura 2001; Miyamoto et al. 2002)
p56	petB-204	tobacco/not in pea	(Hirose and Sugiura 2001; Miyamoto et al. 2002)

[a] Numbers refer to the codon affected by RNA editing

These proteins were bound to all editing sites they provided as targets suggesting that binding was nonspecific and had nothing to do with editing. Surprisingly however, after depleting their *in vitro* editing extract of one of these RRM proteins, cp31, they did observe inhibition of editing at the two sites tested (Hirose and Sugiura 2001). Depletion of other cpRNP proteins, some of them closely related to cp31, did not lead to this effect. In complementation studies, they could show that the acidic domain of cp31 is necessary for editing. In conclusion, cp31 appears to be a general editing factor, probably acting via its acidic domain.

In addition to cpRNPs, factors specific to selected editing sites were identified by UV crosslinking (Table 3). In tobacco, editing sites *psbL*, *psbE*, and *petB* were associated with proteins of 25, 56, and 70 kD, respectively. All three proteins could be titrated off the bait with a sequence-specific competitor, but not with unrelated sequences (Hirose and Sugiura 2001). Similarly, in pea, the *petB* editing site was also specifically associated with a 70 kD protein (Miyamoto et al. 2002). No sequence information for any of these proteins has been reported.

A breakthrough in the search for specificity factors involved in chloroplast RNA editing came from researchers originally interested in other features of chloroplast biogenesis. Shikanai and colleagues were studying the chloroplast NADH dehydrogenase (NDH) complex and isolated mutants affected in the activity of this complex. In an elegant forward genetic screen in *Arabidopsis*, they isolated numerous nuclear mutations that caused the loss of NDH complex activity (Hashimoto et al. 2003). Most of the subunits of the NDH complex are encoded on the chloroplast chromosome by *ndh* genes A through K, which contain several editing sites. One of the mutants isolated, *chlororespiratory reduction 4* (*crr4*), exhibited an editing defect of the *ndhD* start codon, while no other *ndh* editing site was affected (Kotera et al. 2005). Transcript patterns for the *ndh* genes in *crr4* mutants did not deviate from wild type, indicating that the encountered editing defect is likely not a secondary effect of the *crr4* mutation. Later, the authors provided *in vitro* evidence for a specific, direct interaction of CRR4 with the *ndhD* editing site (Okuda et al. 2006). Positional cloning revealed that the *crr4* gene encodes a member of the pentatricopeptide repeat (PPR) protein family (Kotera et al. 2005).

Intriguingly, PPR proteins had long been candidates for editing factors (Small and Peeters 2000; Lurin et al. 2004). These proteins are defined by the PPR motif (Small and Peeters 2000), which is discussed above in the context of chloroplast

RNA splicing. PPR family members have been found in diverse eukaryotic species with usually only a handful of genes per genome, but in embryophytes, the PPR lineage has greatly expanded, with over 450 members in *Arabidopsis* and rice. Most PPR proteins are predicted to be targeted to either mitochondria or chloroplasts, and a string of recent genetic studies suggests that they are generally involved in various aspects of organellar RNA metabolism (e.g. PPR4 and HCF152 in Section 2.3.3 above; reviewed by Small and Peeters 2000; Lurin et al. 2004). The common functions in RNA metabolism for many PPR proteins suggested that PPR proteins bind RNA, but for only a few plant PPR proteins has RNA association *in vivo* (Schmitz-Linneweber et al. 2005b, 2006) or RNA binding *in vitro* been demonstrated (Lahmy et al. 2000; Nakamura et al. 2003; Lurin et al. 2004). The editing factor CRR4 is one of these few: recombinant CRR4 binds to a short segment (-25 to +10) surrounding the *ndhD* start codon editing site in a sequence specific manner and with high affinity (Okuda et al. 2006). This suggests that CRR4 indeed is the factor conferring sequence specificity to this particular editing reaction. In tobacco, *ndhD*-1 has been clustered with two other editing sites, *rpoB*-3 and *ndhF*-2 (Chateigner-Boutin and Hanson 2002). Overexpression of *ndhF*-2 leads to a reduction in editing of the two related sites, suggesting that they share the same specificity factor (Chateigner-Boutin and Hanson 2002). Sites *ndhF*-2 and *ndhD*-1 are also present in *Arabidopsis*, while *rpoB*-3 has been lost. Although CRR4 is not essential for editing *ndhF*-2, it would be still interesting to test whether it binds to this site. In general, it remains an exciting prospect to test other PPR proteins that are evolutionarily or structurally related to CRR4 for a potential role in editing of other sites.

3.5 Models for the editosome

There are two competing models for the machinery responsible for editing site recognition and catalysis. Both models propose a host of specificity factors akin to CRR4 that dock to target *cis*-elements in a highly specific manner. The PPR family of RNA binding proteins is large enough to fill this job easily, but it is too early to exclude roles for other types of RNA binding proteins. The second pressing question is how catalysis occurs; this aspect is addressed differently by the two models.

The original model for the RNA editing apparatus proposed that site recognition factors like PPRs are only a platform for a common factor with enzymatic activity that serves all sites (Fig. 2). Such an activity has not been isolated so far, maybe because a knockout of a general editing enzyme would be gametophyte or embryo-lethal. Of course, cytidine deaminases have been on top of the candidate list for such a general editase, but the few studies on these enzymes did not find any evidence for their involvement in editing (Faivre-Nitschke et al. 1999). Another finding supporting the two-factor-model is that plastid-localized PPRs have been shown to reside in large ribonucleoprotein complexes, presumably together

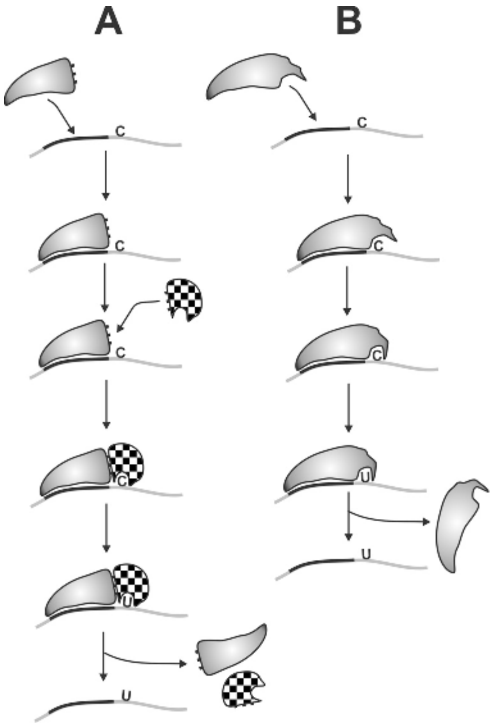


Fig. 2. Two models for the plastid editing machinery. A) Two-component model; a site specific factor (grey) recognizes a *cis*-element (black) upstream of an editing site (C). It forms a platform for an additional factor, the editase (checker pattern) that possesses an activity for converting Cs to Us, but is not necessarily an RNA-binding protein. After catalysis, factors might dissociate from the RNA. B) One-component model; the site-specific factor makes contacts with the *cis*-element, but also directly interacts with the editing site and is sufficient or at least required for catalysis.

with their RNA targets and additional proteins (Meierhoff et al. 2003; Williams and Barkan 2003; Schmitz-Linneweber et al. 2005b). Whether these additional proteins are important for editing is one of the more exciting questions in the field.

The second model for catalysis by the editosome was put forward by Sugiura's group after their finding that the specificity factor p56 makes contacts not only with the upstream *cis*-element, but also with the editing site and the adjacent nucleotides *in vitro*, although these latter interactions are comparatively weak (Hirose et al. 2004; Miyamoto et al. 2004). Consistent with this result, CRR4 also binds with a slight preference to non-edited rather than to pre-edited *ndhD* RNA (Okuda et al. 2006). It seems therefore possible that specificity factors work in a two-step mode, making first solid contact with upstream *cis*-elements and then in a second step also interact with the editing site itself to permit catalysis (Fig. 2). If this model is correct, specificity factors may have different functional protein domains, those for RNA binding and others for catalysis. In fact, many PPR proteins have additional protein domains other than the PPR tract itself. A large subgroup of more than 87 proteins in *Arabidopsis* possesses for example a so-called DYW domain, to which no function has yet been assigned (Lurin et al. 2004). Other domains without known function are present in many PPRs as well. Potentially, these domains could carry out editing catalysis. In the end, both models may be correct: considering the number of sites to be served and the differences in *cis*-elements, there might well be different solutions for the recognition and catalysis of different sites.

3.6 Function and evolution of plastid RNA editing

3.6.1 Evolution of editing sites

Chloroplast RNA editing is widespread in land plants. Of the taxa studied so far, only the marchantiid liverworts do not seem to have RNA editing (Freyer et al. 1997; Duff and Moore 2005). Members of other ancient embryophyte taxa like *Adiantum capillus-veneris* of the ferns (Wolf et al. 2004), *Physcomitrella patens*, and *Takakia lepidozioides* of the mosses (Miyata et al. 2002; Sugita et al. 2006), or *Anthoceros formosae* and other hornworts (Yoshinaga et al. 1996; Yoshinaga et al. 1997; Kugita et al. 2003b; Duff and Moore 2005) each display chloroplast RNA editing. For example, 509 C-to-U and 433 U-to-C editing sites were found in the chloroplast of *A. formosae* (Kugita et al. 2003b). By contrast, spermatophytes exhibit only about 30 C-to-U editing events and no U-to-C editing (Maier et al. 1996; Tsudzuki et al. 2001). Some editing sites are conserved even between vastly divergent embryophyte taxa like ferns and dicots, but most editing sites are restricted to a more narrow taxonomic range (Tillich et al. 2006a). Even between species of the same genus, differences in editing sites were observed (Sasaki et al. 2003). This led to the conclusion that editing sites evolve rapidly (Freyer et al. 1997; Schmitz-Linneweber et al. 2002; Fiebig et al. 2004), at rates similar to those of synonymous codon positions (Shields and Wolfe 1997). This also means that no stabilizing selection acts on editing sites; that is, whether C-to-T editing occurs or whether a T is already encoded on the genomic level does not seem to influence chloroplast function. This apparent futility of RNA editing is reflected in the absence of any data that would support a regulatory role of RNA editing. Most sites are edited at high efficiencies in various tissues. Fluctuations in the ratio between edited and unedited messages over time and space or in response to environmental clues – a prerequisite for regulation – have only been rarely observed (Bock et al. 1993; Ruf and Kössel 1997; Hirose et al. 1999; Karcher and Bock 2002b, 2002a; Miyata and Sugita 2004). Even if quantitative changes in editing efficiency do occur, such effects are expected to be superceded by the much larger variations in abundance of the respective transcripts (Peeters and Hanson 2002). Thus, it is not surprising that the functional significance of quantitative differences in editing efficiency has in no case been established. Nor has the restoration of cryptic translational start codons by editing been shown to impact regulation of protein synthesis (Hirose and Sugiura 1997).

In summary, it is not regulation but simply the generation of conserved codons that makes RNA editing important for chloroplast gene expression. In fact, in the one case where the C of an editing site has been replaced by a T on the genomic level, no deviant phenotype was observed (Schmitz-Linneweber et al. 2005a). This raises the obvious question of why edited Cs are not ultimately substituted by Ts in the DNA. A potential answer is that these edited Cs are simply very stable in evolutionary terms. As a matter of fact, the plastid chromosome in its entirety is evolving rather sluggishly, with a mutation rate lower than that encountered in the nucleus (Palmer 1990; Lynch et al. 2006). In addition, certain sites are less likely to be mutated than others, depending on the identity of the immediate neighboring bases. For Cs in spermatophyte organelle DNA, the context with the lowest C-to-T transition rate is a preceding T and a trailing A: tCa (Morton et al. 1997, 2003). Intriguingly, there is a striking bias towards such a tCa context around editing sites (Tillich et al. 2006a). Apparently, editing sites occur mainly in places where regular point mutations are rare and it might be faster (in evolutionary terms) to evolve a *trans*-acting factor in the nucleus that disposes of an unwanted C at the RNA level. Hence, RNA editing would be compensating for a lack of variation at certain genomic sites, providing an alternative to regular point mutations (Tillich et al. 2006b). It has been calculated that this can only occur in genomes that are slowly evolving, because otherwise, the disadvantage of maintaining *cis*-sequences that are prone to mutation defects would be too great (Lynch et al. 2006). This is consistent with the fact that more rapidly evolving genomes like those of animal mitochondria do not support RNA editing.

3.6.2 Evolution of trans-factors

The paucity of data on *trans*-factors for RNA editing precludes any detailed delineation of trends in *trans*-factor evolution. Still, indirect data on the presence of editing activities in heterologous experiments allow some general conclusions on the evolution of nuclear-encoded *trans*-factors.

Editing sites have been artificially transferred between species by basically two methods: introduction via particle gun transformation or transfer of whole plastid genomes via cybridization. Here, only sites not present in the recipient's plastid genome, so-called foreign or heterologous sites, are considered. The first foreign sites introduced into tobacco were maize site *rpoB*-4 and spinach site *psbF*-1, neither of which was edited (Bock et al. 1994; Reed and Hanson 1997). Similarly, four sites introduced by cybridization in tobacco and *Atropa belladonna* remained unedited in the genomic background of the host species (Schmitz-Linneweber et al. 2005a). In addition, no editing of a tobacco-specific editing site was found in a pea *in vitro* editing system (Miyamoto et al. 2002). This was taken as evidence that the cognate editing factors are evolving rapidly and seemed in accordance with the rapid evolution of editing sites themselves. The picture became more complicated when four examples for the processing of heterologous editing sites were reported (Schmitz-Linneweber et al. 2001, 2005a; Tillich et al. 2006b). These unexpected findings demonstrated that there is a subgroup of editing factors that are conserved between plant taxa independently of their target sites. At the moment, only speculative answers exist as to the reason for their survival despite the absence of their cognate target site. Possibly, such factors are retained because they edit additional sites in plastid transcriptomes as part of a related cluster of sites. Alternatively, these evolutionarily stable factors have additional functions unrelated to editing that provide a selective force for keeping them.

3.7 Perspectives

In comparison to what is known on RNA editing phenomena in humans and trypanosomes, research on plastid RNA editing is lagging far behind. A particularly serious gap is our lack of data on the editing apparatus, the plastid editosome. What are the factors, what is their chemistry, where did this machinery come from and how did it evolve? All these questions remain unanswered despite 15 years of research since the discovery of RNA editing in plastids (Hoch et al. 1991). The recent cloning of the first specificity factor for plastid RNA editing, a PPR protein, may mark the beginning of a rapid elucidation of the machinery behind this enigmatic processing step in the life of chloroplast RNAs.

Acknowledgement

We thank Michi Tillich for critical reading of the manuscript. Work in the authors' laboratories is funded by grants from the National Science Foundation and USDA (AB) and the DFG Emmy-Noether program (CSL).

References

Asakura Y, Barkan A (2006) *Arabidopsis* orthologs of maize chloroplast splicing factors promote splicing of orthologous and species-specific group II introns. Plant Physiol 142:1656-1663

Balczun C, Bunse A, Hahn D, Bennoun P, Nickelsen J, Kuck U (2005) Two adjacent nuclear genes are required for functional complementation of a chloroplast *trans*-splicing mutant from *Chlamydomonas reinhardtii*. Plant J 43:636-648

Balczun C, Bunse A, Schwarz C, Piotrowski M, Kuck U (2006) Chloroplast heat shock protein Cpn60 from *Chlamydomonas reinhardtii* exhibits a novel function as a group II intron-specific RNA-binding protein. FEBS Lett 580:4527-4532

Barkan A (1989) Tissue-dependent plastid RNA splicing in maize: transcripts from four plastid genes are predominantly unspliced in leaf meristems and roots. Plant Cell 1:437-445

Barkan A, Klipcan L, Ostersetzer O, Kawamura T, Asakura Y, Watkins KP (2007) The CRM domain: An RNA binding module derived from an ancient ribosome-associated protein. RNA 13:55-64

Benne R, Van den Burg J, Brakenhoff JP, Sloof P, Van Boom JH, Tromp MC (1986) Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. Cell 46:819-826

Bhattacharya S, Navaratnam N, Morrison JR, Scott J, Taylor WR (1994) Cytosine nucleoside/nucleotide deaminases and apolipoprotein B mRNA editing. Trends Biochem Sci 19:105-106

Bock R (2000) Sense from nonsense: how the genetic information of chloroplasts is altered by RNA editing. Biochimie 82:549-557

Bock R, Hagemann R, Kossel H, Kudla J (1993) Tissue- and stage-specific modulation of RNA editing of the *psbF* and *psbL* transcript from spinach plastids--a new regulatory mechanism? Mol Gen Genet 240:238-244

Bock R, Hermann M, Fuchs M (1997) Identification of critical nucleotide positions for plastid RNA editing site recognition. RNA 3:1194-1200

Bock R, Hermann M, Kossel H (1996) *In vivo* dissection of *cis*-acting determinants for plastid RNA editing. EMBO J 15:5052-5059

Bock R, Koop HU (1997) Extraplastidic site-specific factors mediate RNA editing in chloroplasts. EMBO J 16:3282-3288

Bock R, Kossel H, Maliga P (1994) Introduction of a heterologous editing site into the tobacco plastid genome: the lack of RNA editing leads to a mutant phenotype. EMBO J 13:4623-4628

Bock R, Maliga P (1995) *In vivo* testing of a tobacco plastid DNA segment for guide RNA function in *psbL* editing. Mol Gen Genet 247:439-443

Bonen L, Vogel J (2001) The ins and outs of group II introns. Trends Genet 17:322-331

Cech TR (1986) The generality of self-splicing RNA: relationship to nuclear mRNA splicing. Cell 44:207-210

Chateigner-Boutin AL, Hanson MR (2002) Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared *cis* elements. Mol Cell Biol 22:8448-8456

Chateigner-Boutin AL, Hanson MR (2003) Developmental co-variation of RNA editing extent of plastid editing sites exhibiting similar *cis*-elements. Nucleic Acids Res 31:2586-2594

Chaudhuri S, Carrer H, Maliga P (1995) Site-specific factor involved in the editing of the *psbL* mRNA in tobacco plastids. EMBO J 14:2951-2957

Chaudhuri S, Maliga P (1996) Sequences directing C to U editing of the plastid *psbL* mRNA are located within a 22 nucleotide segment spanning the editing site. EMBO J 15:5958-5964

Choquet Y, Goldschmidt-Clermont M, Girard-Bascou J, Kuck U, Bennoun P, Rochaix JD (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 DA, Hallick RB (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

Copertino DW, Hallick RB (1993) Group II and group III introns of twintrons: potential relationships with nuclear pre-mRNA introns. Trends Biochem Sci 18:467-471

Daniels DL, Michels WJ Jr, Pyle AM (1996) Two competing pathways for self-splicing by group II introns: a quantitative analysis of *in vitro* reaction rates and products. J Mol Biol 256:31-49

Del Campo EM, Sabater B, Martin M (2002) Post-transcriptional control of chloroplast gene expression: Accumulation of stable *psaC* mRNA is due to downstream RNA cleavages in *ndhD* gene. J Biol Chem 277:36457-36464

Deshpande NN, Bao Y, Herrin DL (1997) Evidence for light/redox-regulated splicing of *psbA* pre-RNAs in *Chlamydomonas* chloroplasts. RNA 3:37-48

Driscoll DM, Wynne JK, Wallis SC, Scott J (1989) An *in vitro* system for the editing of apolipoprotein B mRNA. Cell 58:519-525

Duff RJ, Moore FB (2005) Pervasive RNA editing among hornwort rbcL transcripts except *Leiosporoceros*. J Mol Evol 61:571-578

Durrenberger F, Rochaix JD (1991) Chloroplast ribosomal intron of *Chlamydomonas reinhardtii*: *in vitro* self-splicing, DNA endonuclease activity and *in vivo* mobility. EMBO J 10:3495-3501

Eberhard S, Drapier D, Wollman F (2002) Searching limiting steps in the expression of chloroplast-encoded proteins: relations between gene copy number, transcription, transcript abundance and translation rate in the chloroplast of *Chlamydomonas reinhardtii*. Plant J 31:149-160

Ems SC, Morden CW, Dixon CK, Wolfe KH, dePamphilis CW, Palmer JD (1995) Transcription, splicing and editing of plastid RNAs in the nonphotosynthetic plant *Epifagus virginiana*. Plant Mol Biol 29:721-733

Faivre-Nitschke SE, Grienenberger JM, Gualberto JM (1999) A prokaryotic-type cytidine deaminase from *Arabidopsis thaliana* gene expression and functional characterization. Eur J Biochem 263:896-903

Fiebig A, Stegemann S, Bock R (2004) Rapid evolution of editing sites in a small non-essential plastid gene. Nucl Acids Res 7:3615-3622

Freyer R, Hoch B, Neckermann K, Maier RM, Kossel H (1993) RNA editing in maize chloroplasts is a processing step independent of splicing and cleavage to monocistronic mRNAs. Plant J 4:621-629

Freyer R, Kiefer-Meyer MC, Kossel H (1997) Occurrence of plastid RNA editing in all major lineages of land plants. Proc Natl Acad Sci USA 94:6285-6290

Fromm H, Edelman M, Koller B, Goloubinoff P, Galun E (1986) The enigma of the gene coding for ribosomal protein S12 in the chloroplast of *Nicotiana*. Nucleic Acids Res 14:883-898

Fuchs M, Maier RM, Zeltz P (2001) RNA editing in higher plant plastids: oligoribonucleotide SSCP analysis allows the proof of base conversion directly at the RNA level. Curr Genet 39:384-387

Glanz S, Bunse A, Wimbert A, Balczun C, Kuck U (2006) A nucleosome assembly protein-like polypeptide binds to chloroplast group II intron RNA in *Chlamydomonas reinhardtii*. Nucleic Acids Res 34:5337-5351

Golden BL, Gooding AR, Podell ER, Cech TR (1998) A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. Science 282:259-264

Goldschmidt-Clermont M, Choquet Y, Girard-Bascou J, Michel F, Schirmer-Rahire M, Rochaix JD (1991) A small chloroplast RNA may be required for *trans*-splicing in *Chlamydomonas reinhardtii*. Cell 65:135-143

Goldschmidt-Clermont M, Girard-Bascou J, Choquet Y, Rochaix JD (1990) *Trans*-splicing mutants of *Chlamydomonas reinhardtii*. Mol Gen Genet 223:417-425

Gott JM, Emeson RB (2000) Functions and mechanisms of RNA editing. Annu Rev Genet 34:499-531

Hallick RB, Hong L, Drager RG, Favreau MR, Monfort A, Orsat B, Spielmann A, Stutz E (1993) Complete sequence of *Euglena gracilis* chloroplast DNA. Nucleic Acids Res 21:3537-3544

Halls C, Mohr S, Del Campo M, Yang Q, Jankowsky E, Lambowitz AM (2007) Involvement of DEAD-box proteins in group I and group II intron splicing. Biochemical characterization of Mss116p, ATP hydrolysis-dependent and -independent mechanisms, and general RNA chaperone activity. J Mol Biol 365:835-855

Halter C, Peeters N, Hanson M (2004) RNA editing in ribosome-less plastids of iojap maize. Curr Genet 45:331-337

Hashimoto M, Endo T, Peltier G, Tasaka M, Shikanai T (2003) A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast *ndhB* in *Arabidopsis*. Plant J 36:541-549

Hayes ML, Reed ML, Hegeman CE, Hanson MR (2006) Sequence elements critical for efficient RNA editing of a tobacco chloroplast transcript *in vivo* and *in vitro*. Nucleic Acids Res 34:3742-3754

Hegeman CE, Halter CP, Owens TG, Hanson MR (2005a) Expression of complementary RNA from chloroplast transgenes affects editing efficiency of transgene and endogenous chloroplast transcripts. Nucleic Acids Res 33:1454-1464

Hegeman CE, Hayes ML, Hanson MR (2005b) Substrate and cofactor requirements for RNA editing of chloroplast transcripts in *Arabidopsis in vitro*. Plant J 42:124-132

Herrin DL, Bao Y, Thompson AJ, Chen Y-F (1991) Self-splicing of the *Chlamydomonas* chloroplast *psbA* introns. Plant Cell 3:10951107

Herrin DL, Chen Y-F, Schmidt GW (1990) RNA splicing in *Chlamydomonas* chloroplasts: self-splicing of 23S preRNA. J Biol Chem 265:21134-21140

Herrin DL, Schmidt GW (1988) *trans*-splicing of transcripts for the chloroplast psaA1 gene. *In vivo* requirement for nuclear gene products. J Biol Chem 263:14601-14604

Herschlag D (1995) RNA chaperones and the RNA folding problem. J Biol Chem 270:20871-20874

Hess WR, Hoch B, Zeltz P, Hubschmann T, Kossel H, Borner T (1994) Inefficient *rpl2* splicing in barley mutants with ribosome-deficient plastids. Plant Cell 6:1455-1465

Hetzer M, Wurzer G, Schweyen R, Mueller M (1997) *Trans*-activation of group II intron splicing by nuclear U5 snRNA. Nature 386:417-420

Hirose T, Fan H, Suzuki JY, Wakasugi T, Tsudzuki T, Kossel H, Sugiura M (1996) Occurrence of silent RNA editing in chloroplasts: its species specificity and the influence of environmental and developmental conditions. Plant Mol Biol 30:667-672

Hirose T, Kusumegi T, Tsudzuki T, Sugiura M (1999) RNA editing sites in tobacco chloroplast transcripts: editing as a possible regulator of chloroplast RNA polymerase activity. Mol Gen Genet 262:462-467

Hirose T, Miyamoto T, Obokata J, Sugiura M (2004) *In vitro* RNA editing systems from higher plant chloroplasts. Methods Mol Biol 265:333-344

Hirose T, Sugiura M (1997) Both RNA editing and RNA cleavage are required for translation of tobacco chloroplast *ndhD* mRNA: a possible regulatory mechanism for the expression of a chloroplast operon consisting of functionally unrelated genes. EMBO J 16:6804-6811

Hirose T, Sugiura M (2001) Involvement of a site-specific *trans*-acting factor and a common RNA-binding protein in the editing of chloroplast mRNAs: development of a chloroplast *in vitro* RNA editing system. EMBO J 20:1144-1152

Hoch B, Maier RM, Appel K, Igloi GL, Kossel H (1991) Editing of a chloroplast mRNA by creation of an initiation codon. Nature 353:178-180

Holloway SP, Deshpande NN, Herrin DL (1999) The catalytic group-I introns of the *psbA* gene of *Chlamydomonas reinhardtii* : core structures, ORFs and evolutionary implications. Curr Genet 36:69-78

Horlitz M, Klaff P (2000) Gene-specific *trans*-regulatory functions of magnesium for chloroplast mRNA stability in higher plants. J Biol Chem 275:35638-35645

Inada M, Sasaki T, Yukawa M, Tsudzuki T, Sugiura M (2004) A systematic search for RNA editing sites in pea chloroplasts: an editing event causes diversification from the evolutionarily conserved amino acid sequence. Plant Cell Physiol 45:1615-1622

Iost I, Guillerez J, Dreyfus M (1992) Bacteriophage T7 RNA polymerase travels far ahead of ribosomes *in vivo*. J Bacteriol 174:619-622

Jenkins BD, Barkan A (2001) Recruitment of a peptidyl-tRNA hydrolase as a facilitator of group II intron splicing in chloroplasts. EMBO J 20:872-879

Jenkins BD, Kulhanek DJ, Barkan A (1997) Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. Plant Cell 9:283-296

Johanningmeier U, Heiss S (1993) Construction of a *Chlamydomonas reinhardtii* mutant with an intronless *psbA* gene. Plant Mol Biol 22:91-99

Kapoor M, Nagai T, Wakasugi T, Yoshinaga K, Sugiura M (1997) Organization of chloroplast ribosomal RNA genes and *in vitro* self-splicing activity of the large subunit rRNA intron from the green alga *Chlorella vulgaris* C-27. Curr Genet 31:503-510

Karcher D, Bock R (1998) Site-selective inhibition of plastid RNA editing by heat shock and antibiotics: a role for plastid translation in RNA editing. Nucleic Acids Res 26:1185-1190

Karcher D, Bock R (2002a) The amino acid sequence of a plastid protein is developmentally regulated by RNA editing. J Biol Chem 277:5570-5574

Karcher D, Bock R (2002b) Temperature sensitivity of RNA editing and intron splicing reactions in the plastid *ndhB* transcript. Curr Genet 41:48-52

Kim M, Rapp JC, Mullet JE (1993) Direct evidence for selective modulation of *psbA*, *rpoA*, *rbcL*, and 16S RNA stability during barley chloroplast development. Plant Mol Biol 22:447-463

Kotera E, Tasaka M, Shikanai T (2005) A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. Nature 433:326-330

Kuck U, Choquet Y, Schneider M, Dron M, 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

Kugita M, Kaneko A, Yamamoto Y, Takeya Y, Matsumoto T, Yoshinaga K (2003a) The complete nucleotide sequence of the hornwort (*Anthoceros formosae*) chloroplast genome: insight into the earliest land plants. Nucleic Acids Res 31:716-721

Kugita M, Yamamoto Y, Fujikawa T, Matsumoto T, Yoshinaga K (2003b) RNA editing in hornwort chloroplasts makes more than half the genes functional. Nucleic Acids Res 31:2417-2423

Lahmy S, Barneche F, Derancourt J, Filipowicz W, Delseny M, Echeverria M (2000) A chloroplastic RNA-binding protein is a new member of the PPR family. FEBS Lett 480:255-260

Lambowitz AM, Caprara MG, Zimmerly S, Perlman PS (1999) Group I and group II ribozymes as RNPs: clues to the past and guides to the future. In: Gesteland RF, Cech

TR, Atkins JF (eds) The RNA World, 2nd edn Cold Spring Harbor: Cold Spring Harbor Laboratory Press, pp 451–485

Lambowitz AM, Zimmerly S (2004) Mobile group II introns. Annu Rev Genet 38:1-35

Lee J, Herrin DL (2003) Mutagenesis of a light-regulated *psbA* intron reveals the importance of efficient splicing for photosynthetic growth. Nucleic Acids Res 31:4361-4372

Lehnert V, Jaeger L, Michel F, Westhof E (1996) New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the *Tetrahymena thermophila* ribozyme. Chem Biol 3:993-1009

Lemieux C, Otis C, Turmel M (2000) Ancestral chloroplast genome in *Mesostigma viride* reveals an early branch of green plant evolution. Nature 403:649-652

Liere K, Link G (1994) Structure and expression characteristics of the chloroplast DNA region containing the split gene for tRNA-gly (UCC) from mustard. Curr Genet 26:557-563

Liere K, Link G (1995) RNA-binding activity of the *matK* protein encoded by the chloroplast *trnK* intron from mustard (*Sinapis alba* L). Nucleic Acids Res 23:917-921

Lorsch JR (2002) RNA chaperones exist and DEAD box proteins get a life. Cell 109:797-800

Lurin C, Andres C, Aubourg S, Bellaoui M, Bitton F, Bruyere C, Caboche M, Debast C, Gualberto J, Hoffmann B, Lecharny A, Le Ret M, Martin-Magniette ML, Mireau H, Peeters N, Renou JP, Szurek B, Taconnat L, Small I (2004) Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell 16:2089-2103

Lynch M, Koskella B, Schaack S (2006) Mutation pressure and the evolution of organelle genomic architecture. Science 311:1727-1730

Maier RM, Hoch B, Zeltz P, Kossel H (1992a) Internal editing of the maize chloroplast *ndhA* transcript restores codons for conserved amino acids. Plant Cell 4:609-616

Maier RM, Neckermann K, Hoch B, Akhmedov NB, Kossel H (1992b) Identification of editing positions in the *ndhB* transcript from maize chloroplasts reveals sequence similarities between editing sites of chloroplasts and plant mitochondria. Nucleic Acids Res 20:6189-6194

Maier RM, Neckermann K, Igloi GL, Kossel H (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. J Mol Biol 251:614-628

Maier RM, Zeltz P, Kossel H, Bonnard G, Gualberto JM, Grienenberger JM (1996) RNA editing in plant mitochondria and chloroplasts. Plant Mol Biol 32:343-365

Maul JE, Lilly JW, Cui L, dePamphilis CW, Miller W, Harris EH, Stern DB (2002) The *Chlamydomonas reinhardtii* plastid chromosome: islands of genes in a sea of repeats. Plant Cell 14:2659-2679

McCullough AJ, Kangasjarvi J, Gengenback BG, Jones RJ (1992) Plastid DNA in developing maize endosperm. Plant Physiol 100:958-964

Meierhoff K, Felder S, Nakamura T, Bechtold N, Schuster G (2003) HCF152, an *Arabidopsis* RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast *psbB-psbT-psbH-petB-petD* RNAs. Plant Cell 15:1480-1495

Merendino L, Perron K, Rahire M, Howald I, Rochaix JD, Goldschmidt-Clermont M (2006) A novel multifunctional factor involved in *trans*-splicing of chloroplast introns in *Chlamydomonas*. Nucleic Acids Res 34:262-274

Michel F, Dujon B (1983) Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast- and nuclear-encoded members. EMBO J 2:33-38

Michel F, Ferat JL (1995) Structure and activities of group II introns. Annu Rev Biochem 64:435-461

Michel F, Umesono K, Ozeki H (1989) Comparative and functional anatomy of group II catalytic introns--a review. Gene 82:5-30

Michel F, Westhof E (1990) Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. J Mol Biol 216:585-610

Miyamoto T, Obokata J, Sugiura M (2002) Recognition of RNA editing sites is directed by unique proteins in chloroplasts: biochemical identification of *cis*-acting elements and *trans*-acting factors involved in RNA editing in tobacco and pea chloroplasts. Mol Cell Biol 22:6726-6734

Miyamoto T, Obokata J, Sugiura M (2004) A site-specific factor interacts directly with its cognate RNA editing site in chloroplast transcripts. Proc Natl Acad Sci USA 101:48-52

Miyata Y, Sugita M (2004) Tissue- and stage-specific RNA editing of *rps14* transcripts in moss (*Physcomitrella patens*) chloroplasts. J Plant Physiol 161:113-115

Miyata Y, Sugiura C, Kobayashi Y, Hagiwara M, Sugita M (2002) Chloroplast ribosomal S14 protein transcript is edited to create a translation initiation codon in the moss *Physcomitrella patens*. Biochim Biophys Acta 1576:346-349

Mohr G, Lambowitz AM (2003) Putative proteins related to group II intron reverse transcriptase/maturases are encoded by nuclear genes in higher plants. Nucleic Acids Res 31:647-652

Morton BR (2003) The role of context-dependent mutations in generating compositional and codon usage bias in grass chloroplast DNA. J Mol Evol 56:616-629

Morton BR, Oberholzer VM, Clegg MT (1997) The influence of specific neighboring bases on substitution bias in noncoding regions of the plant chloroplast genome. J Mol Evol 45:227-231

Nakajima Y, Mulligan RM (2001) Heat stress results in incomplete C-to-U editing of maize chloroplast mRNAs and correlates with changes in chloroplast transcription rate. Curr Genet 40:209-213

Nakajima Y, Mulligan RM (2005) Nucleotide specificity of the RNA editing reaction in pea chloroplasts. J Plant Physiol 162:1347-1354

Nakamura T, Meierhoff K, Westhoff P, Schuster G (2003) RNA-binding properties of HCF152, an *Arabidopsis* PPR protein involved in the processing of chloroplast RNA. Eur J Biochem 270:4070-4081

Navaratnam N, Morrison JR, Bhattacharya S, Patel D, Funahashi T, Giannoni F, Teng BB, Davidson NO, Scott J (1993) The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. J Biol Chem 268:20709-20712

Neuhaus H, Link G (1987) The chloroplast tRNA(Lys)(UUU) gene from mustard (*Sinapis alba*) contains a class II intron potentially encoding for a maturase-related polypeptide. Curr Genet 11:251-257

Odom OW, Holloway SP, Deshpande NN, Lee J, Herrin DL (2001) Mobile self-splicing group I introns from the *psbA* gene of *Chlamydomonas reinhardtii*: highly efficient homing of an exogenous intron containing its own promoter. Mol Cell Biol 21:3472-3481

Odom OW, Shenkenberg DL, Garcia JA, Herrin DL (2004) A horizontally acquired group II intron in the chloroplast *psbA* gene of a psychrophilic *Chlamydomonas*: *in vitro* self-splicing and genetic evidence for maturase activity. RNA 10:1097-1107

Okuda K, Nakamura T, Sugita M, Shimizu T, Shikanai T (2006) A pentatricopeptide repeat protein is a site-recognition factor in chloroplast RNA editing. J Biol Chem 281:37661-37667

Ostersetzer O, Cooke AM, Watkins KP, Barkan A (2005) CRS1, a chloroplast group II intron splicing factor, promotes intron folding through specific interactions with two intron domains. Plant Cell 17:241-255

Ostheimer GJ, Barkan A, Matthews BW (2002) Crystal structure of *E. coli* YhbY: a representative of a novel class of RNA binding proteins. Structure (Camb) 10:1593-1601

Ostheimer GJ, Hadjivasiliou H, Kloer DP, Barkan A, Matthews BW (2005) Structural analysis of the group II intron splicing factor CRS2 yields insights into its protein and RNA interaction surfaces. J Mol Biol 345:51-68

Ostheimer GJ, Rojas M, Hadjivassiliou H, Barkan A (2006) Formation of the CRS2-CAF2 group II intron splicing complex is mediated by a 22-amino acid motif in the COOH-terminal region of CAF2. J Biol Chem 281:4732-4738

Ostheimer GJ, Williams-Carrier R, Belcher S, Osborne E, Gierke J, Barkan A (2003) Group II intron splicing factors derived by diversification of an ancient RNA-binding domain. EMBO J 22:3919-3929

Palmer JD (1990) Contrasting modes and tempos of genome evolution in land plant organelles. Trends Genet 6:115-120

Peeters NM, Hanson MR (2002) Transcript abundance supercedes editing efficiency as a factor in developmental variation of chloroplast gene expression. RNA 8:497-511

Perron K, Goldschmidt-Clermont M, Rochaix JD (1999) A factor related to pseudouridine synthases is required for chloroplast group II intron *trans*-splicing in *Chlamydomonas reinhardtii*. EMBO J 18:6481-6490

Perron K, Goldschmidt-Clermont M, Rochaix JD (2004) A multiprotein complex involved in chloroplast group II intron splicing. RNA 10:704-711

Pyle A (2002) Metal ions in the structure and function of RNA. J Biol Inorg Chem 7:679-690

Qin PZ, Pyle AM (1998) The architectural organization and mechanistic function of group II intron structural elements. Curr Opin Struct Biol 8:301-308

Rajasekhar VK, Mulligan RM (1993) RNA editing in plant mitochondria: [alpha]-phosphate is retained during C-to-U conversion in mRNAs. Plant Cell 5:1843-1852

Reed ML, Hanson MR (1997) A heterologous maize *rpoB* editing site is recognized by transgenic tobacco chloroplasts. Mol Cell Biol 17:6948-6952

Reed ML, Lyi SM, Hanson MR (2001a) Edited transcripts compete with unedited mRNAs for *trans*-acting editing factors in higher plant chloroplasts. Gene 272:165-171

Reed ML, Peeters NM, Hanson MR (2001b) A single alteration 20 nt 5' to an editing target inhibits chloroplast RNA editing *in vivo*. Nucleic Acids Res 29:1507-1513

Rivier C, Goldschmidt-Clermont M, Rochaix JD (2001) Identification of an RNA-protein complex involved in chloroplast group II intron *trans*-splicing in *Chlamydomonas reinhardtii*. EMBO J 20:1765-1773

Roitgrund C, Mets JL (1990) Localization of two novel chloroplast genome functions: *trans*-splicing of RNA and protochlorophyllide reduction. Curr Genet 17:147-153

Ruf S, Kössel H (1997) Tissue-specific and differential editing of the two *ycf3* editing sites in maize plastids. Curr Genet 32:19–23

Ruf S, Zeltz P, Kossel H (1994) Complete RNA editing of unspliced and dicistronic transcripts of the intron-containing reading frame IRF170 from maize chloroplasts. Proc Natl Acad Sci USA 91:2295-2299

Sasaki T, Yukawa Y, Miyamoto T, Obokata J, Sugiura M (2003) Identification of RNA editing sites in chloroplast transcripts from the maternal and paternal progenitors of tobacco (*Nicotiana tabacum*): comparative analysis shows the involvement of distinct *trans*-factors for *ndhB* editing. Mol Biol Evol 20:1028-1035

Sasaki T, Yukawa Y, Wakasugi T, Yamada K, Sugiura M (2006) A simple *in vitro* RNA editing assay for chloroplast transcripts using fluorescent dideoxynucleotides: distinct types of sequence elements required for editing of *ndh* transcripts. Plant J 47:802-810

Sasaki Y, Kozaki A, Ohmori A, Iguchi H, Nagano Y (2001) Chloroplast RNA editing required for functional acetyl-CoA carboxylase in plants. J Biol Chem 276:3937-3940

Sashital DG, Cornilescu G, McManus CJ, Brow DA, Butcher SE (2004) U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction. Nat Struct Mol Biol 11:1237-1242

Schmitz-Linneweber C, Kushnir S, Babiychuk E, Poltnigg P, Herrmann RG, Maier RM (2005a) Pigment deficiency in nightshade/tobacco cybrids is caused by the failure to edit the plastid ATPase alpha-subunit mRNA. Plant Cell 17:1815-1828

Schmitz-Linneweber C, Regel R, Du TG, Hupfer H, Herrmann RG, Maier RM (2002) The plastid chromosome of *Atropa belladonna* and its comparison with that of *Nicotiana tabacum*: the role of RNA editing in generating divergence in the process of plant speciation. Mol Biol Evol 19:1602-1612

Schmitz-Linneweber C, Tillich M, Herrmann RG, Maier RM (2001) Heterologous, splicing-dependent RNA editing in chloroplasts: allotetraploidy provides *trans*-factors. EMBO J 20:4874-4883

Schmitz-Linneweber C, Williams-Carrier R, Barkan A (2005b) RNA immunoprecipitation and microarray analysis show a chloroplast Pentatricopeptide repeat protein to be associated with the 5' region of mRNAs whose translation it activates. Plant Cell 17:2791-2804

Schmitz-Linneweber C, Williams-Carrier R, Williams P, Kroeger T, Vichas A, Barkan A (2006) A pentatricopeptide repeat protein binds to and facilitates the *trans*-splicing of the maize chloroplast *rps12* pre-mRNA. Plant Cell 18:2650-2663

Sheveleva EV, Hallick RB (2004) Recent horizontal intron transfer to a chloroplast genome. Nucleic Acids Res 32:803-810

Shields DC, Wolfe KH (1997) Accelerated evolution of sites undergoing mRNA editing in plant mitochondria and chloroplasts. Mol Biol Evol 14:344-349

Shikanai T (2006) RNA editing in plant organelles: machinery, physiological function and evolution. Cell Mol Life Sci 63:698-708

Shukla G, Padgett R (2002) A catalytically active group II intron domain 5 can function in the U12-dependent spliceosome. Mol Cell 9:1145-1150

Simon D, Fewer D, Friedl T, Bhattacharya D (2003) Phylogeny and self-splicing ability of the plastid tRNA-Leu group I Intron. J Mol Evol 57:710-720

Small I, Peeters N (2000) The PPR motif - a TPR-related motif prevalent in plant organellar proteins. Trends Biochem Sci 25:46-47

Smith H, Gott J, Hanson M (1997) A guide to RNA editing. RNA 3:1105-1123

Sugita M, Miyata Y, Maruyama K, Sugiura C, Arikawa T, Higuchi M (2006) Extensive RNA editing in transcripts from the PsbB operon and RpoA gene of plastids from the enigmatic moss *Takakia lepidozioides*. Biosci Biotechnol Biochem 70:2268-2274

Sugiura C, Kobayashi Y, Aoki S, Sugita C, Sugita M (2003) Complete chloroplast DNA sequence of the moss *Physcomitrella patens*: evidence for the loss and relocation of rpoA from the chloroplast to the nucleus. Nucleic Acids Res 31:5324-5331

Swisher J, Su L, Brenowitz M, Anderson V, Pyle A (2002) Productive folding to the native state by a group II intron ribozyme. J Mol Biol 315:297-310

Thompson AJ, Herrin DL (1991) *In vitro* self-splicing reactions of the chloroplast group I intron Cr.LSU from *Chlamydomonas reinhardtii* and *in vivo* manipulation via gene-replacement. Nucleic Acids Res 19:6611-6618

Till B, Schmitz-Linneweber C, Williams-Carrier R, Barkan A (2001) CRS1 is a novel group II intron splicing factor that was derived from a domain of ancient origin. RNA 7:1227-1238

Tillich M, Funk HT, Schmitz-Linneweber C, Poltnigg P, Sabater B, Martin M, Maier RM (2005) Editing of plastid RNA in *Arabidopsis thaliana* ecotypes. Plant J 43:708-715

Tillich M, Lehwark P, Morton BR, Maier UG (2006a) The evolution of chloroplast RNA editing. Mol Biol Evol 23:1912-1921

Tillich M, Poltnigg P, Kushnir S, Schmitz-Linneweber C (2006b) Maintenance of plastid RNA editing activities independently of their target sites. EMBO Rep 7:308-313

Tsudzuki T, Wakasugi T, Sugiura M (2001) Comparative analysis of RNA editing sites in higher plant chloroplasts. J Mol Evol 53:327-332

Turmel M, Choquet Y, Goldschmidt-Clermont M, Rochaix JD, Otis C, Lemieux C (1995) The *trans*-spliced intron 1 in the *psaA* gene of the *Chlamydomonas* chloroplast: a comparative analysis. Curr Genet 27:270-279

Turmel M, Otis C, Lemieux C (2002) The chloroplast and mitochondrial genome sequences of the charophyte *Chaetosphaeridium globosum*: insights into the timing of the events that restructured organelle DNAs within the green algal lineage that led to land plants. Proc Natl Acad Sci USA 99:11275-11280

Turmel M, Otis C, Lemieux C (2006) The chloroplast genome sequence of *Chara vulgaris* sheds new light into the closest green algal relatives of land plants. Mol Biol Evol 23:1324-1338

Villa T, Pleiss J, Guthrie C (2002) Spliceosomal snRNAs: Mg^{2+} - dependent chemistry at the catalytic core? Cell 109:149-152

Vogel J, Borner T (2002) Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing. EMBO J 21:3794-3803

Vogel J, Borner T, Hess WR (1999) Comparative analysis of splicing of the complete set of chloroplast group II introns in three higher plant mutants. Nucleic Acids Res 27:3866-3874

Vogel J, Hubschmann T, Borner T, Hess WR (1997) Splicing and intron-internal RNA editing of *trnK-matK* transcripts in barley plastids: support for MatK as an essential splice factor. J Mol Biol 270:179-187

Wakasugi T, Tsudzuki J, Ito S, Nakashima K, Tsudzuki T, Sugiura M (1994) Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. Proc Natl Acad Sci USA 91:9794-9798

Wakasugi T, Tsudzuki T, Sugiura M (2001) The genomics of land plant chloroplasts: Gene content and alteration of genomic information by RNA editing. Photosynthesis Res 70:107-118

Weeks KM (1997) Protein-facilitated RNA folding. Curr Opin Struct Biol 7:336-342

Weihe A (2004) The transcription of plant organelle genomes. In: Daniell H, Chase C (eds) Molecular biology and biotechnology of plant Organelles. Chloroplasts and mitochondria. Dordrecht: Springer, pp 213-237

Westhof E (2002) Group I introns and RNA folding. Biochem Soc Trans 30:1149-1152

Williams PM, Barkan A (2003) A chloroplast-localized PPR protein required for plastid ribosome accumulation. Plant J 36:675-686

Wolf PG, Rowe CA, Hasebe M (2004) High levels of RNA editing in a vascular plant chloroplast genome: analysis of transcripts from the fern *Adiantum capillus-veneris*. Gene 339:89-97

Wolfe KH, Morden CW, Palmer JD (1992) Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. Proc Natl Acad Sci USA 89:10648-10652

Xu MQ, Kathe SD, Goodrich-Blair H, Nierzwicki-Bauer SA, Shub DA (1990) Bacterial origin of a chloroplast intron: conserved self-splicing group I introns in cyanobacteria. Science 250:1566-1570

Yoshinaga K, Iinuma H, Masuzawa T, Uedal K (1996) Extensive RNA editing of U to C in addition to C to U substitution in the *rbcL* transcripts of hornwort chloroplasts and the origin of RNA editing in green plants. Nucleic Acids Res 24:1008-1014

Yoshinaga K, Kakehi T, Shima Y, Iinuma H, Masuzawa T, Ueno M (1997) Extensive RNA editing and possible double-stranded structures determining editing sites in the *atpB* transcripts of hornwort chloroplasts. Nucleic Acids Res 25:4830-4834

Yu W, Schuster W (1995) Evidence for a site-specific cytidine deamination reaction involved in C to U RNA editing of plant mitochondria. J Biol Chem 270:18227-18233

Zaita N, Torazawa K, Shinozaki K, 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

Zeltz P, Hess WR, Neckermann K, Borner T, Kossel H (1993) Editing of the chloroplast *rpoB* transcript is independent of chloroplast translation and shows different patterns in barley and maize. EMBO J 12:4291-4296

Zito F, Kuras R, Choquet Y, Kossel H, Wollman FA (1997) Mutations of cytochrome b6 in *Chlamydomonas reinhardtii* disclose the functional significance for a proline to leucine conversion by *petB* editing in maize and tobacco. Plant Mol Biol 33:79-86

Barkan, Alice

Institute of Molecular Biology, University of Oregon, Eugene, OR 97403, USA

Schmitz-Linneweber, Christian

Institute of Biology, Humboldt-University Berlin, Chausseestr. 117, 10115 Berlin, Germany

christian.schmitz-linneweber@rz.hu-berlin.de

Translation and translational regulation in chloroplasts

Hadas Peled-Zehavi and Avihai Danon

Abstract

The translation mechanism of chloroplast mRNAs originated as prokaryotic-type, but has since evolved considerably. Chloroplast translation became, in large part, uncoupled from transcription, and turned into a highly regulated process. Concomitantly, chloroplast ribosomes, general translation factors, and transcripts changed substantially from their prokaryotic counterparts. A multitude of nucleus encoded regulatory proteins evolved that interact in a specific manner with elements in mRNAs to allow translation regulation in response to environmental and developmental cues. In this chapter, we sum up the current knowledge regarding the translation machinery in the chloroplast using examples of mechanisms utilized for chloroplast translation regulation.

1 Introduction

Chloroplasts are derived from endosymbiosis of oxygenic photosynthetic eubacteria in a non-photosynthetic eukaryotic host (Gray 1993). Hence, the translation mechanism of chloroplast mRNAs originated as prokaryotic-type. As summarized hereinafter, accumulating evidence suggests that chloroplast translation has evolved considerably from its prokaryotic origin. The chloroplast ribosomes, the translation factors, and the transcripts resemble their prokaryotic counterparts but also contain many changes that most likely evolved to facilitate the particular requirements of chloroplast gene expression. A better understanding of the unique features of chloroplast translation is likely to uncover these special requisites.

Gene expression regulation might occur at different points along the linear path from gene to functional protein. In its evolution, chloroplast gene expression underwent a shift from the mostly (but not entirely, Gold 1988) transcriptional regulation observed in prokaryotes to primarily posttranscriptional-based regulation, including regulation of transcript stability, translation, protein turnover, and protein activity (for reviews see Mullet 1988; Gruissem and Tonkyn 1993; Mayfield et al. 1995; Danon 1997; Stern et al. 1997; Zerges 2000; Rochaix 2001; Choquet and Wollman 2002; Eberhard et al. 2002; Nickelsen 2003). The shift to posttranscriptional regulation is also reflected by the mechanism of translation. Whereas translation of nascent transcripts, i.e., cotranscriptional translation, pre-

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0234 / Published online: 13 June 2007

vails in bacteria, the translation of chloroplast mRNAs is typically uncoupled to transcription, and is self-regulated in response to environmental and developmental cues (reviewed in Gillham et al. 1994; Mayfield et al. 1995; Danon 1997; Zerges 2000; Eberhard et al. 2002). The finding that transcription and translation are commonly uncoupled indicates the presence of new inhibitory steps that disrupt the constitutive course of prokaryotic-type translation and, thereby, converting it into a regulated mechanism. As will be discussed below, the regulation of the translatability of chloroplast messages entails the concerted action of RNA structures and sequence motifs, mostly in the untranslated region (UTR) of the mRNA, and of nucleus-encoded transacting proteins.

It is important to note that the shift to posttranscriptional-based regulation occurred concomitantly to the most dramatic change of chloroplast genome evolution, the retaining of only about 60 to 200 genes out of the several thousands of its progenitor genome (Martin and Herrmann 1998). Thus, the chloroplast underwent a drastic reduction in the number of transcripts encoded by its own genome. In contrast, hundreds of nucleus-encoded proteins are expected to be transported into the chloroplast and to be involved in RNA-binding activities (Martin and Herrmann 1998; Lorkovic and Barta 2002, Plasmid Proteome Data Bank http://ppdb.tc.cornell.edu; Lurin et al. 2004; van Wijk 2004), suggesting a high ratio of interacting proteins per single chloroplast transcript. Notably, accumulative results of genetic analyses in *Chlamydomonas reinhardtii* (Kuchka et al. 1988, 1989; Rochaix et al. 1989; Drapier et al. 1992; Girard-Bascou et al. 1992; Zerges and Rochaix 1994; Yohn et al. 1996; Stampacchia et al. 1997; Zerges et al. 1997; Cahoon and Timko 2000; Rochaix 2001; Wostrikoff et al. 2001; Dauvillee et al. 2003) and *Arabidopsis thaliana* (Meurer et al. 1996, 1998; Felder et al. 2001; Lennartz et al. 2001; Plucken et al. 2002; Nakamura et al. 2003; Sane et al. 2005; Barneche et al. 2006; Lennartz et al. 2006) have identified nucleus-encoded gene products that are required for the posttranscriptional regulation of chloroplast mRNAs. Interestingly, many of these mutations were each specific to a unique chloroplast mRNA. Moreover, mutational analysis identified three nuclear genes that are required for the translation of a single chloroplast mRNA (Zerges and Rochaix 1994; Zerges et al. 1997). Hence, it is possible that the translation of the small number of chloroplast transcripts itself underwent a shift towards a more transcript specific-type regulation, and that each transcript might be regulated by one, or more, distinct nucleus-encoded proteins. The possibility that the regulatory mechanisms of translation have diversified during chloroplast evolution is further implicated by the non-conserved position of the Shine-Dalgarno (SD) ribosome binding site in chloroplast transcripts and the finding that existing SD sequences are not always necessary for translation initiation (Fargo et al. 1998; Sugiura et al. 1998).

This review aims at summarizing the present understanding of translation and translation regulation mechanisms in the chloroplast. As described below, the most recurrent theme seems to be that a multitude of different strategies were adopted for the regulation of translation of small number of chloroplast mRNAs. The different regulatory schemes use an abundance of unique nucleus-encoded factors acting together with structured and unstructured cis-elements located predominantly in the 5'UTR of the chloroplast mRNAs.

2 Chloroplast translation machinery

The translation machinery in the chloroplast generally resembles that found in prokaryotes; the chloroplast ribosomes are closely related to the eubacterial 70S-type ribosomes, chloroplast transcripts are not m^7G capped at their 5' end, and lack 3' poly(A) tails. Furthermore, the anti-Shine-Dalgarno (SD) sequences at the 3' ends of the 16S rRNAs of cyanobacteria and chloroplasts share high homology with the *E. coli* anti-SD sequence (Dron et al. 1982; Steege et al. 1982; Maidak et al. 1996). Yet, important differences, which will be presented in detail hereinafter, indicate that the translation and its regulation have evolved considerably.

The genes encoding the chloroplast translational machinery are distributed between the chloroplast and nuclear genomes. The rRNA and tRNA genes are located in the chloroplast genome, while the genes for tRNA synthetases, processing/modification enzymes and part of the ribosomal proteins are located in the nuclear genome. Proteomic studies identified all of the protein components of both the ribosomal 30S and 50S subunits in spinach and in the unicellular green alga *C. reinhardtii* (Yamaguchi and Subramanian 2000; Yamaguchi et al. 2000, 2002, 2003). Spinach plastid ribosome comprises 59 proteins (33 in 50S subunit and 25 in 30S subunit and a putative ribosome recycling factor in the 70S ribosome) of which 53 are *Escherichia coli* orthologues and six have no *E. coli* orthologues and are plastid-unique proteins. Two 50S subunit *E. coli* proteins have no orthologues in the spinach plastid. Similarly, the majority of the proteins that were identified in *C. reinhardtii* are *E. coli* orthologues. Only 20 proteins out of the 59 ribosomal proteins of spinach are encoded in the plastid genome, while the rest are encoded by the nuclear genome. Due to the plastid specific ribosomal proteins and to N- and C-terminal extensions added to some of the other ribosomal proteins, the protein mass of the plastid ribosome is bigger then *E. coli* ribosome in both spinach and *C. reinhardtii*, though the specifics differ between the two organisms. In contrast, only minor changes occur in chloroplast rRNA structure. It was proposed that the additional domains of 30S ribosomal proteins and the 30S plastid-specific proteins might be involved in the regulation of chloroplast translation by mediating the effect of nucleus-encoded factors and/or by assisting in positioning of mRNAs on the ribosome for translation initiation. One suggested role of the plastid-specific 50S ribosomal proteins might be associated with protein targeting to thylakoid membranes (Yamaguchi and Subramanian 2000; Yamaguchi et al. 2000; Manuell et al. 2004).

The 30S ternary complex with mRNA in prokaryotes includes three initiation factors, IF1, IF2, and IF3 (for review see Laursen et al. 2005). Homologues of IF2 have been identified in the chloroplast of *Euglena gracilis* (Ma and Spremulli 1990) and bean (Campos et al. 2001). IF3 homologue has been identified in *Euglena gracilis* (Wang and Spremulli 1991; Lin et al. 1996). Despite the homology,

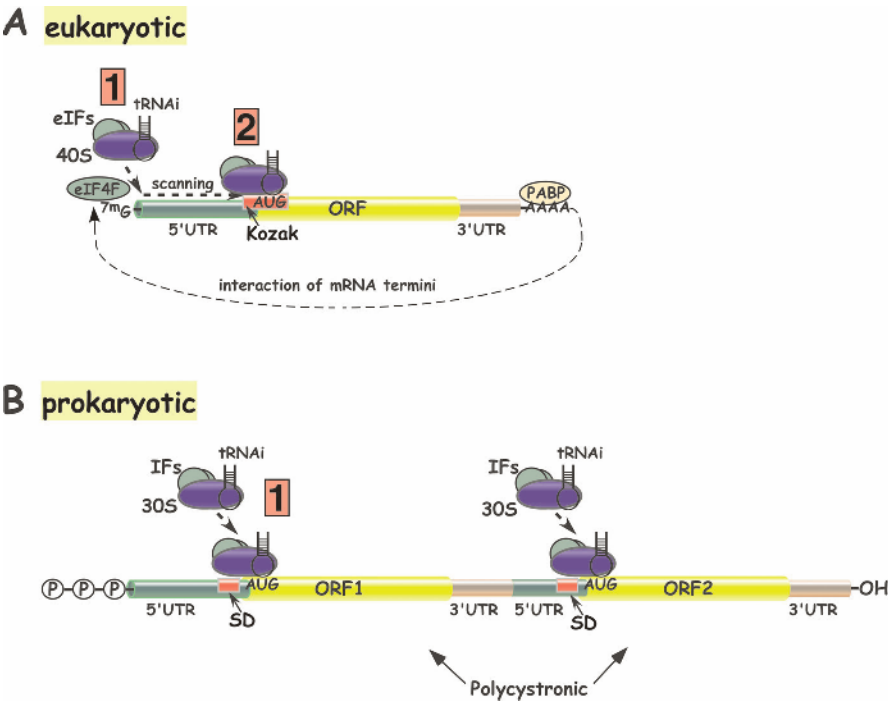


Fig. 1. Translation initiation in prokaryotes and eukaryotes. a. In eukaryotes translation initiation occurs through the scanning mechanism of initiation, which consists of two separate steps. First, the 40S small ribosomal subunit is loaded on the mRNA immediately downstream of the 5'-cap through an interaction between the cap-binding eIF4F complex and the 40S bound eIFs. The eIF4F complex also interacts with the poly(A) binding protein (PABP) bound to the 3' poly(A) tail, creating a 'closed loop' that promotes the recruitment of the 40S ribosomal subunit. In the second step of initiation, the 40S subunit, with the aid of an RNA helicase, which is also a component of the eIF4F complex, scans the RNA in the 5'-->3' direction for the first AUG codon that is embedded in the proper sequence context. b. In prokaryotes the initiation complex binds directly to the initiation codon. The binding of the 30S small ribosomal subunit to the mRNA is facilitated by base pairing between the SD ribosome binding site on the 5' UTR and sequences in the 3' end of the 16S rRNA of the 30S subunit, and this binding localizes the initiation complex to the correct initiation codon. This mechanism allows for the simultaneous translation of several ORFs in a polycistronic transcript. IFs, initiation factors; eIFs, eukaryotic initiation factors; SD, Shine-Dalgarno ribosome binding site.

the chloroplast initiation factors also seems to differ from their prokaryotic homologues. The algal IF3 contains NH_2- and COOH-terminal extensions that are not found in *E. coli* IF3. Sequences in these extensions reduce the activity of IF3 in promoting initiation complex formation with chloroplast mRNAs and 30S ribosomal subunits. It was proposed that these regions allow for a chloroplast-specific regulatory mechanism of initiation, and that their inhibitory effect might be alleviated in response to developmental or environmental conditions such as light (Yu and Spremulli 1998).

3 Mechanisms of translation initiation

Translation initiation of both prokaryotes and eukaryotes begins with two key steps: The *first step* involves the binding of the mRNA by the ribosomal small subunit, and the *second step* is the selection of the proper initiation codon. While in eukaryotic translation these are two well-separated steps, in prokaryotes the two steps are generally combined, such that the binding of the ribosomal small subunit to the mRNA concurrently positions it on top of the selected initiation codon. In eukaryotes, initiation of translation of the vast majority of mRNAs occurs through the following intermediate steps, each of which might be subjected to regulation, of the scanning mechanism (Fig. 1a); (i) The preinitiation complex, comprised of the 40S ribosomal subunit, $tRNA_i^{Met}$ and initiation factors, interacts with the cap-binding initiation factor eIF4F, and as a result binds the mRNA immediately downstream of the 5'-cap; (ii) After exchanging several translation initiation factors and acquiring helicase activity the preinitiation complex scans the mRNA in the 5'-->3' direction for the first AUG codon that is embedded in a consensus sequence that promotes initiation; (iii) On recognition of the proper AUG triplet, which in most cases is the closest to the mRNA cap, base-pairing with the $tRNA_i^{Met}$ anticodon takes place, triggering GTP hydrolysis, and release of several initiation factors. The efficiency of recognition is determined by the sequence surrounding the initiation codon. Energy is needed for the scanning process, and secondary structures in the 5'UTR, or the association of RNA-binding proteins can regulate the rate of translation initiation. A few mammalian cellular mRNAs and several RNA viruses utilize a different translation initiation mechanism that recruits the preinitiation complex to the mRNA by a structured stem-loop RNA motif called the internal ribosome entry element (IRES) in the 5' untranslated region (UTR) of the mRNA (for reviews see Jackson 2005; Kozak 2005).

In prokaryotes, the initiation complex does not bind to the 5' end of the mRNA but rather directly to the SD site, a purine rich sequence typically found approximately 5 to 9 bases upstream of the initiation codon (Fig. 1b). The recognition of and binding to the SD is facilitated by complementary base pairing between the SD and sequences in the 3' end of the 16S rRNA of the small ribosomal subunit. The binding of the initiation complex is further augmented by the interaction of the S1 protein of the 30S ribosomal small subunit with a pyrimidine rich sequence in the 5'UTR of the mRNA (Subramanian 1983; Boni et al. 1991; Sorensen et al. 1998; Sengupta et al. 2001). The binding of the initiation complex to the SD sequence localizes the initiation complex to the proper initiation codon. The *exact distance* between the SD sequence and the initiation codon is therefore *critical* for bona fide translation initiation. This process does not require energy, is usually independent of upstream sequences, and therefore allows for the simultaneous translation of several ORFs in a polycistronic transcript (Gold 1988). The combined steps of mRNA binding and initiation codon selection in prokaryotic-type translation, thereby, result in the typical cotranscriptional translation. In the rare cases of regulated translation, it occurs via the occlusion of the SD by either RNA structure or RNA-binding protein.

Alternative translation initiation pathways exit in prokaryotes as well, as leaderless mRNAs, lacking an SD sequence, are efficiently translated *in vitro* (Moll et al. 2002). The binding of leaderless mRNA to the ribosome differs from that of canonical mRNAs, but the exact mechanism is not clear yet (for review see Laursen 2005). The faithful translation of leaderless mRNAs in heterologous systems indicated that the ability to translate leaderless mRNAs might be an evolutionary conserved function of the translational apparatus (Moll et al. 2002).

The common uncoupling of transcription and translation in the chloroplast suggests that additional levels of regulation are needed to prevent constitutive translation initiation of mRNAs translation through SD-16S rRNA interactions. Indeed, chloroplast translation initiation seems to deviate from the "classical" prokaryotic translation initiation in several key issues, and there might be more than one way for initiating translation in the chloroplast. A key difference is the role of SD sequences in translation initiation in the chloroplast. Whereas, the exact distance between the SD sequence and the initiation codon is critical for bona fide translation initiation in prokaryotes, the precise role of the SD in chloroplast mRNAs is yet unclear. Though many of the plant chloroplast genes have SD-like sequences in their 5'UTR, the distance of this sequence from the initiation codon is often variable, and deviates from its conserved position in *E. coli*. In the tobacco chloroplast genome, 30 of 79 chloroplast protein-coding genes have no SD-like sequence located within 20 nt upstream from the initiation codon. The remaining 49 genes have SD-like sequences, but not necessarily at a conserved position, and overall two thirds of chloroplast mRNAs do not contain SD sequences at the correct position (Sugiura et al. 1998). Bearing in mind that in prokaryotes the distance between the SD sequence and the initiation codon is critical to the *positioning* of the small ribosomal subunit on top of the authentic initiation codon, the absence of SD sequence at the correct position in approximately two thirds of chloroplast mRNAs suggests that either the binding of mRNA by the ribosomal small subunit and selection of initiation codon are *two separated steps* in the translation initiation of these mRNAs, or that the binding and positioning is achieved by alternative factors, perhaps by regulatory nucleus-encoded proteins that interact with structural RNA elements. Hence, though for some chloroplast mRNAs translation initiation might occur similarly to prokaryotic-type translation initiation (for graphical illustration see Fig. 2a), translation initiation of the majority of mRNAs require additional elements to facilitate recognition of the proper initiation codon by the small ribosomal subunit and efficient translation, allowing for additional levels of regulation. Thus, alternative pathways for the identification of the correct initiation codon in the chloroplast may coexist (Fig. 2b, 2c).

This notion is reinforced by a series of experiments utilizing different species of plants and different methods (for summary see Table 1). First, chloroplast transformation in *C. reinhardtii* was used to look at mutants of chloroplast SD-like sequences. The results demonstrated differing levels of effect on translation

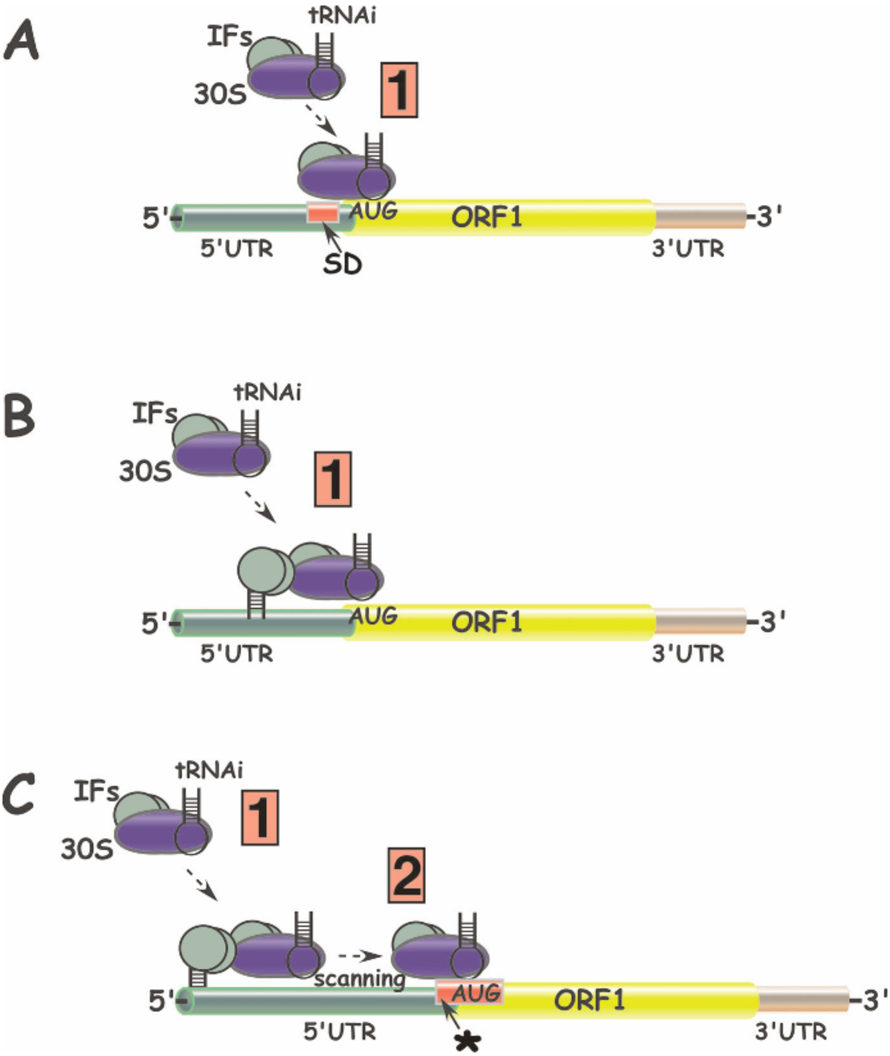


Fig. 2. Models of chloroplast initiation of translation. a. For some mRNAs, translation initiation might occur in the "classical" prokaryotic mechanism. The SD sequence is situated in a conserved position on the 5' UTR, and can recruit the 30S small ribosomal subunit to the correct initiation codon. b. In 5' UTR that lack a conserved ribosome binding site, secondary structure elements within the 5'UTR aided by transacting factors can correctly position the initiation complex to the initiation codon. **c**. Alternatively, cis- and trans-acting elements might bind the initiation complex upstream of the initiation codon, and a mechanism reminiscent of the eukaryotic scanning mechanism might be needed to direct the initiation complex to the correct initiation codon. Notably, a scanning mechanism is likely to require a signature sequence, functionally analogous to the eukaryotic Kozak consensus sequence, stimulating the recognition of the authentic initiation codon. IFs, initiation factors; SD, Shine-Dalgarno ribosome binding site; ∗ initiation codon recognition signature sequence.

Table 1. SD and regulatory cis-elements in the 5'UTR of chloroplast mRNAs

Species	essential SD	nonessential/no SD	alternative cis-elements
Algae			
C. reinhardtii	*psbA*[a], *psbD*[b]	*petD*[c], *atpB*[d], *atpE*[d], *rps4*[d], *rps7*[d]	*petD*[c,e], *atpB*[f], *rps7*[f,g], *psaB*[h], *psbC*[i,j], *psbA*[a], *psbD*[b]
E. gracilis	*atpH*[k]	*rbcL*[l]	*rbcL*[l]
Higher plants			
N. tabacum	*atpE*[m], *rbcL*[m], *rps14*[n]	*rps12*[m], *petB*[m], *psbA*[o], *atpB*[p]	*psbA*[o], *atpB*[p]
Z. mays			*petA*[q], *psaC*[q],
S. oleracea			*psbA*[r], *atpI*[s]

[a]Mayfield et al. 1994, [b]Nickelsen 1999, [c] Sakamoto et al. 1994, [d]Fargo et al. 1998, [e]Higgs et al. 1999, [f]Hauser et al. 1996, [g]Fargo et al. 1999, [h]Stampacchia et al. 1997, [i]Rochaix et al. 1989, [j]Zerges et al. 1997, [k]Betts and Spremulli 1994, [l]Koo and Spremulli 1994, [m] Hirose and Sugiura 2004a, [n] Hirose et al. 1998, [o] Hirose and Sugiura 1996, [p] Hirose and Sugiura 2004b, [q] Schmitz-Linneweber et al. 2005, [r] Boni et al. 1991, [s]Merhige et al. 2005.

depending on the mRNA. Deletion of the SD-like sequence from *psbA* 5'UTR abolished translation, and reduced the level of mRNA (Mayfield et al. 1994). On the other hand, replacement mutations of the SD-like sequences in the 5'UTR of five mRNAs, *petD, atpB, atpE, rps4*, and *rps7* had little or no effect on their translation *in vivo* (Sakamoto et al. 1994; Fargo et al. 1998). In another experiment, replacement mutagenesis of the SD-like sequence in the *psbD* 5'UTR reduced synthesis of the polypeptide to 25% of the wild type level (Nickelsen et al. 1999). Hence, the function of the SD sequence appears to be dependent on the identity of the mRNA.

In a different approach, the 'toe-printing' method was used to examine the role of SD-like sequences in the 5'UTRs of chloroplast mRNAs in *E. gracilis*. In the 'toe-printing' method the position of an initiation complex on the mRNA is determined by its hindering effect of primer extension reaction. Even though mutations of the SD-like sequence of the *E. gracilis* chloroplast *atpH* mRNA resulted in two to fivefold reductions in the efficiency of initiation complex formation, the *rbcL* mRNA was found to be translated independently of the SD like sequence (Betts and Spremulli 1994; Koo and Spremulli 1994).

Biochemical assays utilizing an *in vitro* translation system from tobacco chloroplast proved useful in looking at different regulatory elements of translation initiation (Hirose and Sugiura 1996). Using this system to look at SD-like sequences in tobacco chloroplast mRNAs it was demonstrated that the position of the SD sequence relative to the initiation codon determines its necessity for translation initiation (Hirose et al. 1998; Hirose and Sugiura 2004a). The *atpE*, *rps14*, and *rbcL* mRNAs have SD-like sequences at a position similar to the conserved SD region in *E. coli*, and these sequences were found to be essential for translation. On the other hand, SD-like sequences in the *rps12* mRNA and in *petB* mRNA are located far from and too close to the initiation codon, respectively, and these sequences are not essential for translation of the corresponding message. The same *in vitro* translation system yielded two very informative observations. First, it was shown that although the tobacco *rps2* mRNA possesses an SD-like sequence at a proper position from the initiation codon, this sequence functions as a negative regulatory element for translation (Plader and Sugiura 2003), suggesting, at least for *rps2* mRNA, a deviation from the prokaryotic function of the SD sequence. Second, an unstructured sequence containing the initiation codon in the proper sequence context was shown to be required for the translation of the tobacco *atpB* mRNA, which does not contain an SD-like sequence (Hirose and Sugiura 2004b), suggesting an additional divergent mechanism for translation initiation. In both cases, trans-acting factors were implied in translation regulation.

Analysis of the binding of *E. coli* 30S ribosomal subunits to barley chloroplast mRNAs *in vitro* has shown that it varies among different messages depending on the existence of a conserved SD sequence. In a message containing an SD-like sequence located in close proximity to the initiation codon, the *E. coli* ribosomal subunits associated with the same region as chloroplast ribosomes. Conversely, in a message that does not contain an SD-like sequence located in close proximity to the initiation codon, their patterns of association differed (Kim and Mullet 1994). These results support the notion that mRNA binding by plastid ribosomes seems to have evolved distinct features, and that it may require interactions with trans-acting proteins that are unique to plastid ribosomes.

It is interesting to note that the SD sequence is necessary for the translation of the *psbA* mRNA in *C. reinhardtii*, but is not necessary for its translation in tobacco (Mayfield et al. 1994; Hirose and Sugiura 1996). Similarly, the SD sequence is necessary for the translation of the *rbcL* mRNA in tobacco (Hirose and Sugiura 2004a), but not in *E. gracilis* (Koo and Spremulli 1994). Although the differences might stem from the different assays used to dissect the importance of the SD sequence, the data suggest that the role of the SD sequence in the translation of a specific protein is not necessarily conserved across species.

The role of the prokaryotic S1 protein in binding of ribosomes to mRNAs may suggest a similar function in the chloroplast, especially in messages lacking SD in the correct position. Chloroplast homologs of bacterial S1 were identified in cyanobacteria (Sugita et al. 1995), spinach (Franzetti et al. 1992; Shteiman-Kotler and Schuster 2000), and *C. reinhardtii* (Merendino et al. 2003). The chloroplast S1 protein is a nuclear-encoded protein and is much shorter than the bacterial protein. Different RNA-binding specificities were reported for the chloroplast S1 protein with preference to AU-rich RNA sequences that are common in the 5'UTR of chloroplast genes (Franzetti et al. 1992; Alexander et al. 1998; Shteiman-Kotler and Schuster 2000; Merendino et al. 2003). Thus, further work is needed to determine the authentic binding site of the chloroplast S1 and its possible contribution to the positioning of the ribosome in translation initiation.

Extended interactions between the mRNA and the initiator tRNA might also contribute to the efficiency of translation initiation. In prokaryotes, a uridine at position –1 upstream of the initiation codon was proposed to allow a fourth base pair with the adenine immediately downstream of the initiator tRNA anticodon. Similarly to *E. coli*, *C. reinhardtii* chloroplast genes preferentially have a U at the -1 position. Indeed, *in vitro* and *in vivo* experiments support a 5' extended codon– anticodon interaction in *C. reinhardtii petA* mRNA translation initiation (Esposito et al. 2001, 2003). It will be interesting to check the relative importance of this extended base-pairing in mRNAs that lack a conserved SD sequence.

Taken together, the above results strongly suggest that: 1) consistently with the expanded role of translational control in the chloroplast, the mechanism of translation initiation of a large portion of chloroplast mRNAs deviates from the classical prokaryotic mechanism of translation initiation; 2) alternative divergent translation initiation pathways exist; 3) trans-acting factors are probably involved in the translational control.

4 Translation initiation regulation – intricate interplay between cis- and trans-acting elements

The findings that SD-like sequences are not always necessary for translation initiation in the chloroplast suggest that the binding and/or the positioning of the initiation complex along the mRNA is mediated by alternative cis-element and trans-acting factors. Though detailed data on the mechanisms that allow for efficient translation initiation in these alternative pathways is still missing, several repeating themes seem to emerge from existing information, and those will be summarized below.

4.1 Cis-elements in chloroplast 5'UTRs

A hallmark of regulated translation in the chloroplast seems to be an abundance of cis-acting elements in the 5'UTRs of chloroplast mRNAs. The mechanisms that are involved in translation attenuation by these cis-elements vary as well. The chloroplast 5'UTR cis-elements can presumably participate in translation initiation pathways in mRNAs that do not utilize a SD sequence. Alternatively, they might complement the SD sequence, or confer specific regulation of translation in response to environmental or developmental cues. In support of this notion, a considerable portion of the cis-elements, that were found up to date in the 5'UTRs of chloroplast mRNAs, appear to promote translation. Examples include elements in the 5'UTRs of the mRNAs of *psbC, petD, rps7, psbD, psbA, psaB*, and *atpB* in *C. reinhardtii* and in several mRNAs from higher plants (for review see Danon 1997; Zerges 2000, and Table 1 and below). There are at least two examples, the *psbD* and *psbA* mRNAs of *C. reinhardtii*, in which both an SD-like sequence and additional cis-elements in the 5' UTR appear to be required for efficient translation initiation, though the mechanisms particulars seem to differ (Mayfield et al. 1994; Nickelsen et al. 1999). The prevalence of positive regulation of translation is higher in chloroplasts relative to prokaryotes, where secondary structure of cis-elements or protein binding to the mRNA, usually repress translation by blocking access to the initiation site (Gold 1988, Kozak 2005). On the other hand, there are examples of translation regulation that are reminiscent of prokaryotic type translation regulation. The Control by Epistasy of Synthesis (CES) mechanism (Choquet and Vallon 2000, see below) is similar to negative feedback loops found in prokaryotes (Kozak 2005). Another example is the *C. reinhardtii psbD* gene, encoding the D2 protein of photosystem II, which exhibits both positive and negative translation regulation by cis-elements. One *cis*-acting element comprises a stretch of multiple U residues whose deletion completely abolishes the synthesis of the *psbD* gene product (Nickelsen et al. 1999). A second negative regulator element is a double-stranded RNA region encompassing the initiation codon, whose conformation needs to be changed before translation initiation (Klinkert et al. 2006). An interesting example of translation regulation by cis-elements, that demonstrate the divergence of chloroplast translation mechanisms from the “classical” prokaryotic type translation initiation, is the translation regulation of the chloroplast *rps2* gene. The tobacco *rps2* mRNA, which encodes ribosomal protein S2 of the 30S ribosomal subunit, has an SD-like sequence at a proper position relative to the initiation codon. Hence, it was expected that this sequence would play important role in translation initiation of the *rps2* mRNA. Unexpectedly, using *in vitro* translation assay the SD-like sequence of tobacco *rps2* mRNA was found to act as a negative regulator of translation. A trans-acting factor was implicated in the process (Plader and Sugiura 2003).

4.2 Structural elements in 5’UTRs

Stem-loop structures in the 5’UTRs of chloroplast mRNA seem to play an important role in the regulation of translation initiation. One example is the translation of the *C. reinhardtii psbC* mRNA, encoding the 51-kDa chlorophyll-binding PSII reaction center subunit P6, which requires the central 100 nt of its 547 nt 5’UTR (Zerges et al. 1997). This region has the potential to form a stable stem-loop secondary structure. Two bulges in the stem, caused by two sites of noncomplementarity between the strands, are essential for translation. Stem-loop mutations that increased the structure stability resulted in inhibited translation of *psbC* mRNA *in vivo*, whereas point mutation that weakened the structure suppressed the effect of nuclear mutation (Rochaix et al. 1989).

Stem-loop structures were shown as well to be critical to the translation initiation of *C. reinhardtii psbA* (Mayfield et al. 1994, see below) and *petD* mRNAs. Site-directed and linker-scanning mutagenesis identified three distinct elements within the 5’UTR of *petD* mRNA (encoding subunit IV of the cytochrome b6/complex) (Sakamoto et al. 1994; Higgs et al. 1999). Element I appears to form a small stem-loop and is located at the 5’ end of the mRNA. It is required for both stability and translation of the mRNA, and may interact with a protein factor to block 5’ to 3’ exoribonucleolytic degradation of the mRNA (Higgs et al. 1999). The two other elements, II and III, are required for translation, but not mRNA stability. Element II is an unstructured region of 16 nt located in the center of the

UTR and appears to bind proteins that protect it from dimethyl sulfate modification. Element III spans a region of 14 nt close to the AUG initiation codon. This sequence appears to form a stem-loop *in vivo* (Higgs et al. 1999).

Comparison of the 5'UTR sequences of orthologous *petD* mRNAs among four *Chlamydomonas* species demonstrated that although the overall sequence conservation across these species is low, the sequences of the three regulatory elements present in the 5'UTR of the *petD* mRNA and their relative positions appear partially conserved (Kramzar et al. 2006). Functionality of the divergent 5'UTRs was tested in *C. reinhardtii* chloroplasts using reporter genes. Only the nearly identical *C. incerta petD* 5'UTR retained its translational control in *C. reinhardtii* chloroplasts (Kramzar et al. 2006).

Thus, the work on both the *psbC* and *petD* mRNAs suggests that the regulatory interactions between 5'UTR elements and nucleus-encoded factors are highly specific and very sensitive to minor sequence changes (Rochaix et al. 1989; Kramzar et al. 2006).

An interesting case is the translation of *rps7* mRNA encoding the chloroplast ribosomal protein S7. Several mutations isolated in the 5'UTR of the chloroplast *rps7* gene in *C. reinhardtii* reduce expression of reporter genes. These mutations altered the predicted secondary structure of the 5'UTR by weakening the stability of stem structures. Second site mutations that restored the predicted secondary structure suppressed the loss of reporter activity caused by the original mutations, suggesting that a *stable* RNA structure is required for translation (Fargo et al. 1999). The translational negative mutations failed to bind a 20 kDa protein that turned out to be S7 itself (Fargo et al. 2001).

Translation initiation that involves stem-loop 5'UTR elements that interact with protein factors is reminiscent of the translation initiation mechanism demonstrated in a few mammalian cellular mRNAs and several RNA viruses that recruit the 40S ribosomal subunit to the mRNA by a structured stem-loop RNA motif in the 5'UTR called the internal ribosome entry element (IRES) (Jackson 2005).

4.3 General and specific translation factors

A multitude of genetic and biochemical approaches were used to identify nuclear genes that participate in chloroplast gene expression (reviewed in Barkan and Goldschmidt-Clermont 2000). Very little information emerged concerning general factors that promote translation of multiple mRNAs. Using UV crosslinking, at least seven binding proteins of 81, 62, 56, 47, 38, 36, and 15 kDa were detected that bind several different *C. reinhardtii* chloroplast mRNAs. The 81, 47, and 38 kDa proteins were shown to associate with all tested 5'UTRs (Hauser et al. 1996). The identity and function of the different proteins is not yet clear. The level of the 36 kDa protein was diminished in cells that preferentially translate chloroplast-encoded ribosomal proteins, suggesting that it may be required for translation of a class of proteins encoding photosynthetic proteins (Hauser et al. 1996). In another set of experiment the S7 ribosomal protein was shown to bind several different mRNAs including the *rps12, rbcL, atpB, and psbA* mRNAs, raising the intriguing hypothesis that S7 might have a role in the translation initiation of a subset of chloroplast mRNAs (Fargo et al. 2001). Competition assays in spinach demonstrated that four ATP synthase 5'UTRs were able to compete with each other for binding by proteins in a chloroplast extract. Thus, at least some of the binding proteins recognized all four of those 5'UTRs (Hotchkiss and Hollingsworth 1999). Furthermore, competition-binding assays between an ATP synthase 5'UTR and 5'UTRs from several other chloroplast genes revealed that the ATP synthase-binding proteins can bind the majority of the 5'UTRs examined (Robida et al. 2002). Though the function of these binding proteins is not known, these findings suggest that some RNA-binding proteins have a more general role in the regulation of either mRNA stability or translation.

In contrast to the paucity of data regarding general translation initiation factors, a growing body of data supports the importance of mRNA specific protein factors. Thus, a large number of nucleus-encoded proteins were found, each needed for the translation of only one or few chloroplast mRNAs. This became evident at first with the discovery of several nuclear mutations in *C. reinhardtii* that cause reduction or elimination of translation of specific chloroplast mRNAs. Examples include nuclear mutations that disrupt the translation of the *psbC* (Rochaix et al. 1989; Zerges and Rochaix 1994; Zerges et al. 1997), *psbA* (Girard-Bascou et al. 1992; Yohn et al. 1996), *psaB* (Stampacchia et al. 1997; Dauvillee et al. 2003), *atpA* (Drapier et al. 1992), *petA* (Wostrikoff et al. 2001), and *psbD* mRNAs (Kuchka et al. 1988, 1989). Though less common, there are also examples of nuclear mutations in higher plants that cause decrease in synthesis of specific chloroplast proteins. In maize, mutants of the nuclear gene *crp1* are defective in the translation of the chloroplast *petA* and *petD* mRNAs, and also fail to process a monocistronic *petD* mRNA from its polycistronic precursor (Barkan et al. 1994). The maize nuclear *atp1* gene is required for translation of the *atpB* mRNA (McCormac and Barkan 1999).

4.4 Multiple proteins interact with single mRNA

As exemplified in the translation of *psbA* mRNA, that is discussed in detail further on, another emerging theme in the translation regulation of chloroplast mRNAs is the involvement of several protein factors in the translation regulation of a single mRNA. UV crosslinking experiments identified at least seven proteins that bind to several different *C. reinhardtii* chloroplast mRNAs (Hauser et al. 1996), and five RNA binding proteins ranging from 16 to 80 kDa were shown to bind to the *rps7* 5'UTR (Fargo et al. 2001). The different protein factors can interact with each other and the 5'UTR or can function independently of each other. For example, the products of three nuclear loci were shown to interact with the *psbC* 5'UTR of *C. reinhardtii*. Two of them, Tbc1 and Tbc3, interact with each other and sequence elements in the 5'UTR to activate translation initiation at the GUG initiation codon of the mRNA (Zerges et al. 1997). Another nuclear gene, Tbc2, appears to function in *psbC* translation independently of Tbc1 and Tbc3 (Zerges et al. 1997, 2003).

The 5'UTR of the chloroplast *psbD* gene of *C. reinhardtii* encoding the D2 protein of photosystem II contains several distinct RNA elements, which are involved in the translational control of its expression. One of these elements is an SD-like sequence. A second is a stretch of eleven consecutive U residues, interrupted by a single A residue. Deletion of this sequence abolishes translation of the *psbD* mRNA (Nickelsen et al. 1999). A 40 kDa RNA binding protein (RBP40) was shown to interact specifically with the U-rich element, and is needed for translation of the *psbD* mRNA. Furthermore, interaction of RBP40 with the *psbD* 5'UTR was found to be dependent on the Nac2 factor, which is required for the stabilization of the *psbD* mRNA (Nickelsen et al. 1994; Boudreau et al. 2000; Ossenbuhl and Nickelsen 2000).

The involvement of multiple proteins in the translation regulation of a single mRNA was demonstrated in higher plants as well. In spinach chloroplasts, two conserved regions in the 5'UTR of *atpI* mRNA were shown to bind at least two different proteins, though the exact function of the proteins remains to be clarified (Merhige et al. 2005).

The findings of both general and transcript specific RNA-binding proteins may implicate a more eukaryotic type regulation. The emergence of systems biology has effectively demonstrated that RNA-binding proteins that regulate eukaryotic gene expression tend to bind specific mRNA subpopulations ranging from tens to hundreds different mRNAs (Hieronymus and Silver 2004; Keene and Lager 2005), and that the final level of synthesized protein is influenced by the combinatorial effect of several regulatory circuits at the same level. Gene expression regulation in the chloroplast might represent a similar but much smaller regulatory network. Nucleus-encoded RNA-binding proteins might bind specifically only one or a few chloroplast mRNAs, and the final protein level might depend on the combined effect of several smaller regulatory circuits on any specific mRNA.

5 Translation regulation examples

The complex array of general and mRNA specific cis- and trans-regulatory elements creates a network that allows for the dynamic and coordinated chloroplast translation regulation necessary to respond to developmental and environmental cues. There are only a few cases for which detailed information regarding this intricate regulation exists. Some examples are given below.

5.1 Translation regulation of D1 synthesis

Photodamage to the D1 protein of photosystem II necessitates rapid turnover and replacement with newly synthesized D1 for continuation of efficient photosynthesis. Light induces a 50 to 100-fold enhancement of synthesis of D1 without an equivalent increase in *psbA* mRNA levels in higher plants and algae cells, suggesting that translation is the regulated step (Fromm et al. 1985; Klein et al. 1988;

Malnoe et al. 1988; Krupinska and Apel 1989). The mechanism of light-regulated translation of *C. reinhardtii psbA* mRNA, encoding the D1 protein, was thoroughly studied over the last few years, and therefore provides a good case study to look at the way the repeating themes of translation regulation in the chloroplast converge to create an orchestrated response to changing environmental conditions.

The 5'UTR of *C. reinhardtii psbA* gene contains a stem-loop element immediately upstream of a putative SD sequence. The SD sequence is located 27 nucleotides upstream of the initiation codon, i.e., in a non-conserved position. Both elements play a role in protein synthesis. Deletion of the SD-like sequence abolished translation, and reduced the level of mRNA, while site-directed mutations that disrupt the stem-loop element reduce D1 protein synthesis without affecting *psbA* mRNA accumulation (Mayfield et al. 1994).

A set of mRNA binding proteins which bind the *psbA* 5'UTR with high affinity and specificity has been identified and purified from *C. reinhardtii* cells by RNA-affinity chromatography, capable of isolating both proteins that bind directly to the RNA and proteins that are associated through protein-protein interactions (Danon and Mayfield 1991). *psbA* 5'UTR-binding proteins are composed of four proteins: RB38, RB47, RB55, and RB60. These form a complex (*psbA* 5'PC), which binds the mRNA through the RB47 protein. The level of binding of *psbA* 5'PC to the mRNA parallels the level of *psbA* mRNA translation and association with polyribosomes in light- and dark-grown wild type *C. reinhardtii* and in several mutants lacking translation of *psbA* mRNA (Danon and Mayfield 1991; Trebitsh and Danon 2001; Zou et al. 2003). This suggests that light regulates polyribosome association and translation of *psbA* mRNA by modulating the binding of *psbA* 5'PC to the 5'UTR.

In contrast to most of the nucleus-encoded translation regulators of chloroplast mRNAs, two of the four proteins that constitutes the regulatory *psbA* 5'PC, namely RB47 and RB60, have been cloned and characterized. Both proteins are nucleus-encoded proteins that are targeted to the chloroplast of *C. reinhardtii* (Yohn et al. 1998b; Trebitsh et al. 2001), where they associate with both the full length 5'UTR of *psbA* mRNA and its mature processed form (Danon and Mayfield 1991; Bruick and Mayfield 1998) and regulate the expression of the message (Yohn et al. 1996, 1998a; Trebitsh et al. 2000; Trebitsh and Danon 2001). RB47 is a member of the eukaryotic poly(A)-binding protein (PABP) family, and like all members of the family, contains four conserved RNA recognition motifs (RRMs) (Yohn et al. 1998a). PABPs are involved in polyadenylation of mRNA, but also in different aspects of translation initiation and termination, and mRNA decay (Mangus et al. 2003). In *Chlamydomonas*, the cytoplasmic PABP, a 69 kDa polypeptide, is imported from the cytoplasm into the chloroplast, where it is processed to the 47 kDa form (Yohn et al. 1998a). The RB60 protein shows high homology to protein disulfide isomerase (PDI) (Trebitsh et al. 2001), an oxidoreductase that was identified first as a highly abundant, essential protein in the lumen of the ER where it catalyzes the formation, reduction and isomerization of disulfide bridges of nascent proteins during their folding in the ER. RB47 binds directly to the mRNA (Danon and Mayfield 1991), whereas, RB60 is thought to bind to RB47 and to modulate its activity, probably by oxidoreducing specific thiol groups of

RB47 (Danon and Mayfield 1994b; Kim and Mayfield 1997; Fong et al. 2000; Trebitsh et al. 2000; Alergand et al. 2006).

Two complementary regulatory mechanisms have been proposed for RB60 control of RB47 activity. In the first, the binding of *psbA* mRNA is regulated by the reduction and oxidation of disulfide groups in RB60 (Danon and Mayfield 1994b; Trebitsh et al. 2000). Because the pool of RB60-thiols in the chloroplast becomes proportionally reduced with increasing light intensity it was suggested that the purpose of this regulatory mechanism is to modulate *psbA* mRNA translation in parallel to incident light (Trebitsh et al. 2000). In the second mechanism, ADP-dependent phosphorylation of RB60 inactivates the binding to *psbA* mRNA. As the inactivation by phosphorylation of RB60 requires high ADP concentrations, normally attained only in chloroplasts in the dark, the role of this mechanism is thought to diminish *psbA* mRNA translation in darkness (Danon and Mayfield 1994a). The mechanism by which the phosphorylation, or redox state, of RB60 activates or inactivates the translation of the *psbA* mRNA is still unknown.

Recently, a new player in the translation regulation of *psbA* mRNA was cloned and characterized. Tba1 is a novel protein, whose expression is needed for *psbA* mRNA/ribosome association and D1 translation. Tba1 is also needed for RB47 RNA binding activity, but its exact role in the mechanism described above is still unknown (Somanchi et al. 2005).

Whether a similar mechanism of light-regulated translation of *psbA* mRNA exists in higher plants is unclear to date. Yet, light-regulated translation in higher plants exhibits several similar characteristics to the mechanism identified in *C. reinhardtii* and some differences as well. In *A. thaliana,* two proteins of 43- and 30-kDa were shown to bind the *psbA* 5'UTR. Oxidizing conditions abolished the association of the proteins with the 5'UTR, while RNA-binding activity was recovered upon incubation with a reductant. Thus, it was hypothesized that similarly to *C. reinhardtii,* redox-dependent interactions play a role in the posttranscriptional regulation of *psbA* gene expression in *A. thaliana* (Shen et al. 2001).

Heterologous genes fused to tobacco *psbA* 5’UTR are enhanced by light, suggesting that similarly to *C. reinhardtii*, initiation of D1 translation in tobacco plastids is controlled via the *psbA* 5'UTR (Staub and Maliga 1994). But, unlike *C. reinhardtii psbA* mRNA, the SD-like sequence in the 5’UTR of tobacco *psbA* mRNA has little influence on translation. Translation requires three other elements within the 5’UTR. Two of them are complementary to the 3'-terminus of chloroplast 16S rRNA (termed RBS1 and RBS2) and the other is an AU-rich sequence located between RBS1 and RBS2 and is termed the AU box (Hirose and Sugiura 1996). The AU box was shown to be recognized by a protein factor(s) and a model was proposed for the initiation of *psbA* translation whereby RBS1 and RBS2 bind cooperatively to the 3'-end of 16S rRNA resulting in looping out of the AU box, which facilitates the interaction of a transacting factor(s) (Hirose and Sugiura 1996).

In spinach, a 43 kDa protein homologous to the *E. coli* ribosomal S1 protein has been shown to bind an element in the 5’UTR of *psbA* mRNA that comprises an SD-like sequence. Binding activity of this protein can be detected only after plants have been illuminated (Alexander et al. 1998). There is evidence suggesting that *E. coli* ribosomal S1 protein can mediate association of mRNA with the 30S ribosomal subunit by binding pyrimidine rich sequences upstream of the SD sequence in the mRNA (Boni et al. 1991). Whether or not the 43 KDa protein might play a similar role in spinach is unclear.

5.2 Negative feedback loops: assembly-controlled regulation of translation

A second well-studied example of cis-acting elements in the 5'UTR of mRNA, which regulate translation initiation through protein binding, is the CES mechanism. The four major multimeric complexes in the thylakoid membrane (PSI, PSII, ATP synthase and cytochrome b_6f) comprise subunits encoded by the chloroplast genome side by side with nucleus-encoded subunits. Thus, a regulated coordinated expression of proteins from the two genomes is essential for an energy efficient and functional assembly of the complexes. This is achieved by post-translational degradation of unassembled subunits (for review see Choquet and Vallon 2000), but also by an assembly regulated translation of some chloroplast-encoded proteins, a phenomenon called CES. The CES process was first studied in the cytochorome b_6f complex in *C. reinhardtii* (Choquet and Wollman 2002). Cytochrome f shows a reduced rate of synthesis in the absence of its assembly partners, cytochrome b_6 or subunit IV, but there is no change in the stability of the protein that is synthesized (Kuras and Wollman 1994). This assembly-dependant regulation of cytochrome f synthesis stems from autoregulation of translation initiation of its own *petA* mRNA. Two components which are required for the reduced translation initiation were identified: the 5'UTR of *petA* mRNA which is sufficient to confer the CES behavior to a reporter gene (Choquet et al. 1998), and a repressor motif on the C-terminal of the unassembled cytochrome f protein that is able to inhibit further translation of its own mRNA (Kuras and Wollman 1994; Choquet et al. 2001, 2003). As there is no evidence for direct binding of cytochrome f to its mRNA, the involvement of a ternary effector was suggested (Choquet et al. 2003). Thus, a negative feedback mechanism insures translation arrest following accumulation of unassembled cytochrome f.

In *C. reinhardtii*, similar negative feedback loops, controlling translation via the 5'UTRs of relevant mRNAs, were shown to exist for the three other multi-complexes in the thylakoid membrane, PSI, PSII, and ATP synthase (Choquet and Wollman 2002; Wostrikoff et al. 2004; Minai et al. 2006). Furthermore, though the molecular details are different, assembly controlled translation of cytochrome b_6f exists also in tobacco (Monde et al. 2000b), and there is evidence for a similar regulation of PSII components in barley (Gamble and Mullet 1989; Kim et al. 1994b).

Ribulose1,5-bisphosphate carboxylase/oxygenase (Rubisco), the enzyme responsible for CO_2 fixation during photosynthesis, is another example of a multimeric complex containing subunits encoded by the chloroplast, side by side with nucleus-encoded subunits. It is composed of a nucleus-encoded small subunit and a chloroplast encoded large subunit. It was shown that when accumulation of the large subunit is limiting (as in some *rbcL* nonsense and missense mutants), the small subunit levels are adjusted to those of the large subunit at the level of protein degradation (for review see Rodermel 1999). On the other hand, limiting the amounts of small subunits by expression of an *rbcS* anti-sense RNA in tobacco (Rodermel et al. 1988, 1996) or deletion of the gene by insertional mutagenesis in *C. reinhardtii* (Khrebtukova and Spreitzer 1996) resulted in decrease in the translation of the chloroplast *rbcL* mRNA. Thus, it was suggested that the level of small subunits regulates large subunits accumulation at the level of *rbcL* translation. The decrease in large subunit translation might be mediated by inhibition of translation by free large subunits, similarly to CES, or by lack of positive regulation by small subunits. A recent work in tobacco suggests that indeed the underlying mechanism is CES-like, and unassembled large subunit autoregulates its own translation (Wostrikoff 2007), though the identity of the cis-elements involved are not yet known. Interestingly, it was suggested that the light-induced oxidative stress inhibition of Rubisco large subunit translation is caused by structural changes that result in exposure of an RNA recognition motif (RRM) at the N-terminal of the large subunit. It was suggested that the exposed RRM will then bind any RNA in its vicinity including its own transcript, resulting in the translational arrest of the large subunit (Cohen et al. 2005).

6 Regulation of translation elongation

Most of the aforementioned discussion has focused on translation regulation at the level of initiation, which is most commonly the rate-limiting step in translation, and has gained most of the attention of researchers. Yet, evidence for regulation at the level of elongation was found as well. Conceivably, regulation of translation elongation might be beneficial for processes such as cotranslational membrane insertion or assembly.

The activities of the *E. gracilis* chloroplast elongation factors EF-Tu, EF-G, and EF-Ts, as well as the activities of the pea chloroplast EF-G and EF-Tu were shown to be regulated by light (Breitenberger et al. 1979; Fox et al. 1980; Sreedharan et al. 1985; Akkaya and Breitenberger 1992; Singh et al. 2004), suggesting a possible role for these factors in regulation of translation in response to environmental cues. Furthermore, translation elongation of the *psbA* mRNA was demonstrated to be light-regulated in higher plants (Kim et al. 1991; Taniguchi et al. 1993; Kim et al. 1994a; Edhofer et al. 1998; Muhlbauer and Eichacker 1998). Toe print analysis in barley showed that ribosomes indeed pause at distinct sites during the elongation phase of *psbA* mRNA translation (Kim et al. 1991, 1994b; Kim and Mullet 1994). Taken together with extensive work demonstrating that chlorophyll stimulates the accumulation of Dl and other chlorophyll proteins by increasing chlorophyll apoprotein stability, it was suggested that ribosome pausing during elongation improves the efficiency of Dl synthesis by providing additional time for nascent chains to bind cofactors such as chlorophyll prior to polypeptide release from the ribosomes (Klein et al. 1988; Mullet et al. 1990; Kim et al. 1994a; Kim and Mullet 1994; Zhang et al. 1999, 2000).

Regulation of elongation might occur in additional chloroplast mRNAs. For example, ribosome pausing was also suggested to play a role in the expression of the large ATP synthase gene cluster in spinach chloroplasts (Stollar et al. 1994). Additionally, it was shown that translation initiation complexes for *rbcL* mRNA (encoding the large subunit of Rubisco) are normally formed in the dark, but the elongating step right after formation of translation initiation complexes might be blocked. The release of this translational elongation block upon illumination may contribute to light-activated translation of the *rbcL* mRNA (Kim and Mullet 2003).

Recently, it was reported that that the *C. reinhardtii* plastid-specific ribosomal protein PSRP-7, which contains two S1 domains, is encoded by a gene whose complete ORF codes for a 110 kDa polyprotein that also contains two EF-Ts domains on its carboxy end. The 110 kDa protein containing the S1 domains and the EF-Ts was identified in cell extracts, as well as proteins containing only the S1 or the EF-Ts domains. It was suggested that the structure of this gene implicates coordinated expression of the S1 like protein and EF-Ts, but that the stable expression of the full 110 kDa protein implies that this protein plays a novel, yet unknown, role in translation (Beligni et al. 2004).

7 Interactions of 5' and 3' ends of chloroplast mRNA

In eukaryotes, interactions between the two termini of cytoplasmic mRNAs stimulate the initiation of translation. The poly(A) binding protein (PABP) bound to the 3' poly(A) tail interacts with initiation factors bound to the 5'UTR, thus creating a 'closed loop' that promotes the recruitment of the 40S ribosomal subunit. It is generally thought that the 'closed loop' role is a quality control mechanism to promote translation of full-length mRNAs rather than truncated forms (Gallie 1998). Translatable chloroplast mRNAs do not contain poly(A) tails. Most of them, similarly to prokaryotic mRNAs, contain an AU-rich 3'UTR with a terminal inverted repeat. The 3'UTR inverted repeat has been shown to play a role in the processing and stabilization of the mRNA (for review see Monde et al. 2000a). Examples of modulation of translation initiation by interactions between the two termini of mRNA in prokaryotes (Lindahl and Hinnebusch 1992; Franch and Gerdes 1996; Voorma 1996) raise the possibility that such interactions might also exist in chloroplast mRNAs and influence their expression. Indeed, there are several reports that support a role for the 3'UTR in translation initiation of several mRNAs. Correct processing of the 3'UTR was suggested to be required for high levels of translation initiation and polysomal association in *C. reinhardtii* cells (Rott et al. 1998). Recent results from tobacco transformants in which the influence of the *psbA* UTRs on translation of a reporter gene were studied indicated that including the *psbA* 3'UTR resulted in a three to fourfold enhancement of translation (Eibl et al. 1999). Furthermore, though high affinity binding of regulator proteins to *C. reinhardtii psbA* mRNA is primarily via its 5'UTR, the 3'UTR was shown to increase the affinity of binding of the 5'UTR-binding protein complex (Katz and Danon 2002). In another study, deletion of the inverted repeat of the 3'UTR of tobacco *petD* mRNA led to a reduction in *petD* expression beyond that expected by the decrease in mRNA accumulation alone, indicating that the 3'UTR might also contribute to efficient translation (Monde et al. 2000a). Further research is needed to establish the generality of this phenomena and its importance for translation efficiency.

8 Subchloroplast location of translation

The chloroplast consists of several different subcompartments (such as the soluble stroma, thylakoids, and the chloroplast envelope), each requiring its own set of proteins as well as other molecules such as pigments, cofactors, and lipids. Furthermore, the assembly of functional complexes within the different subcompartments requires the coordinated assembly of components synthesized within the chloroplasts and components imported from the cytosol. Thus, the location of protein translation within the chloroplast is not self-evident, and several subcompartments were suggested to be involved in the process. Early sedimentation studies in extracts of *C. reinhardtii* demonstrated that a significant percentage of polyribosomes are attached to thylakoid membranes, and this attachment is light dependant (Chua et al. 1973; Margulies and Michaels 1974, 1975; Chua et al. 1976; Bolli et al. 1981). Furthermore, many thylakoid proteins were shown to be synthesized on thylakoid-attached polyribosomes of *C. reinhardtii* and higher plants (Margulies 1983; Minami and Watanabe 1984; Bhaya and Jagendorf 1985; Margulies et al. 1987; Breidenbach et al. 1988; Klein et al. 1988; Shinohara et al. 1988). It was thus suggested that an evolved function of the plastid-specific 50S ribosomal proteins might be associated with protein targeting to thylakoid membranes (Yamaguchi and Subramanian 2000; Yamaguchi et al. 2000; Manuell et al. 2004). Recently, proteomic analysis of *A. thaliana* identified components of the translation machinery in the chloroplast thylakoid membranes thus supporting the notion that thylakoid membranes play a role in chloroplast translation (Friso et al. 2004). Translation of thylakoid proteins on thylakoid-bound polyribosomes makes sense as the proper assembly of thylakoid complexes necessitates coordinated and stepwise assembly of the different components of the complex. Indeed, D1 is synthesized on membrane-bound polyribosomes, and assembled cotranslationally into the membrane (for review see Zhang and Aro 2002). The picture is complicated by the finding that the stromal large subunit of Rubisco is also synthesized on membrane-associated polyribosomes (Hattori and Margulies 1986; Breidenbach et al. 1988; Klein et al. 1988). It was suggested that the translation of the large subunit of Rubisco by thylakoid-attached polyribosomes allows for regulation of translation by the photosynthetic proton gradient without the need for signal transduction to stromal ribosomes (Muhlbauer and Eichacker 1999).

There is also suggestive evidence that translation is associated with the inner membrane of chloroplast envelope (for review see Sato et al. 1999). Furthermore, there is data that imply that thylakoid proteins might be translated in the chloroplast envelope. A set of RNA-binding proteins, including RB47, which was reported to be a specific activator of *psbA* mRNA translation, were found to be associated with chloroplast membranes whose buoyant density and acyl lipid composition imply that their origin is the inner chloroplast envelope membrane. These membranes were found to be associated with thylakoid membranes (Zerges and Rochaix 1998). An earlier report also found polyribosomes attached to a membrane fraction, which differed from thylakoid membranes in polypeptide composition and the amount of chlorophyll it contained (Margulies and Weistrop 1980). It is not entirely clear whether these membranes are a subfraction of the thylakoids or inner envelope membrane or a previously uncharacterized intra-chloroplast compartment.

9 Concluding remarks

An increasing body of evidence suggests that chloroplast translation has evolved considerably from its prokaryotic origin. Chloroplast translation became, in large part, uncoupled from transcription, and turned into a highly regulated process. Concomitantly, chloroplast ribosomes, general translation factors and transcripts changed substantially from their prokaryotic counterparts. Accumulating evidence based on genetic, biochemical and proteomic approaches imply that a plethora of nucleus-encoded regulatory proteins that interact in a specific manner with structured and unstructured cis-elements located predominantly in the 5'UTR of chloroplast transcripts have evolved. The dramatic reduction in the number of chloroplast genes, and the expanded number of nucleus-encoded RNA-binding proteins, indicate that the translation of the small number of chloroplast transcripts underwent a shift towards a more transcript specific-type regulation.

While the importance of the chloroplast unique trans-acting proteins and 5'UTR elements to the regulation of translation has been demonstrated repeatedly, a most intriguing question is yet to be resolved; how the initiation complex is positioned onto the bona fide initiation codon? The variable location of transcript-unique translational cis-elements in the 5'UTR relative to the initiation codon, some located far upstream, indicate that the mechanism of positioning diversified from the prokaryotic type. Two likely scenarios might be envisioned; 1) the gap between the binding site of the initiation complex and the initiation codon is bridged by structural elements in the 5'UTR (Fig. 2b); 2) helicase activity associated with the initiation complex promotes, in a similar fashion to eukaryotic translation, scanning for the initiation codon (Fig. 2c). Such a mechanism is likely to require a signature sequence that will enhance the binding of the 30S subunit to the correct initiation codon, similarly to the function of the Kozak consensus sequence in eukaryotes.

Is there a biological gain in the convoluted evolution of chloroplast gene expression or is it a mere outcome of random selection? Interestingly, the mitochondrion, the other endosymbiotic prokaryotic-like organelle, shows many evolutionary parallelism in its evolution to the chloroplast including the transfer of most of its self-encoded functions to the nucleus and an increase in translationally regulated gene expression (Fox 1996). The similar evolution of chloroplast and mitochondrion gene expression indicates a high selection pressure for this type of regulation. The biological advantage of this type of system organization is yet unclear, but its clarification is important to our understanding of the primary principles governing organellar functions.

What might be the special requirements in the chloroplast that made translational control a favored mechanism for regulating gene expression? Examining the type of genes that were retained in the chloroplast and mitochondrion genomes might suggest a possible explanation. In addition to components of gene expression system, i.e., tRNAs, ribosomal RNAs, and proteins, most of the retained organellar genes encode proteins involved in electron transport and energy coupling. Thus, perhaps to counteract the potentially harmful side effects of electron transport chain reactions, structural proteins that maintain redox balance within bioenergetic membranes must be synthesized when and where they are needed (Race et al. 1999). The requirement for dynamic and tight regulation is further accentuated for photosynthetic gene expression in the chloroplast as rapid adjustment is critical to ensure efficient energy production and prevention of deleterious side effects in response to changing light intensity and availability. Translational regulation of gene expression allows for rapid on and off adjustment of rates of protein synthesis from an existing pool of transcripts. In contrast, transcriptional regulation is relatively slow to induce protein synthesis and has to be accompanied by mRNA instability to enable turning off translation in a short time. Why such a complex network of RNA-binding proteins and RNA cis-elements is required for the regulation of chloroplast gene expression? It is possible that the small number of pivotal chloroplast genes is subject to multiple regulatory circuits, including the coordination with nucleus expression, developmental regulation of plastid-type specific expression, metabolism switches between light and dark, and adjustments to changes in light intensity.

Acknowledgement

Supported by grants from the Israeli Science Foundation, the Minerva Foundation, and by the Charles W. and Tillie K. Lubin Center for Plant Biotechnology at the Weizmann Institute of Science. AD is incumbent of The Henry and Bertha Benson Chair, Weizmann Institute of Science.

References

Akkaya MS, Breitenberger CA (1992) Light regulation of protein synthesis factor EF-G in pea chloroplasts. Plant Mol Biol 20:791-800

Alergand T, Peled-Zehavi H, Katz Y, Danon A (2006) The chloroplast protein disulfide isomerase RB60 reacts with a regulatory disulfide of the RNA-binding protein RB47. Plant Cell Physiol 47:540-548

Alexander C, Faber N, Klaff P (1998) Characterization of protein-binding to the spinach chloroplast psbA mRNA 5' untranslated region. Nucleic Acids Res 26:2265-2272

Barkan A, Goldschmidt-Clermont M (2000) Participation of nuclear genes in chloroplast gene expression. Biochimie 82:559-572

Barkan A, Walker M, Nolasco M, Johnson D (1994) A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. EMBO J 13:3170-3181

Barneche F, Winter V, Crevecoeur M, Rochaix JD (2006) ATAB2 is a novel factor in the signalling pathway of light-controlled synthesis of photosystem proteins. EMBO J 25:5907-5918

Beligni MV, Yamaguchi K, Mayfield SP (2004) Chloroplast elongation factor ts pro-protein is an evolutionarily conserved fusion with the s1 domain-containing plastid-specific ribosomal protein-7. Plant Cell 16:3357-3369

Betts L, Spremulli LL (1994) Analysis of the role of the Shine-Dalgarno sequence and mRNA secondary structure on the efficiency of translational initiation in the *Euglena gracilis* chloroplast atpH mRNA. J Biol Chem 269:26456-26463

Bhaya D, Jagendorf AT (1985) Synthesis of the alpha and beta subunits of coupling factor 1 by polysomes from pea chloroplasts. Arch Biochem Biophys 237:217-223

Bolli R, Mendiola-Morgenthaler L, Boschetti A (1981) Isolation and characterization of polysomes from thylakoid membranes of *Chlamydomonas reinhardii*. Biochim Biophys Acta 653:276-287

Boni IV, Isaeva DM, Musychenko ML, Tzareva NV (1991) Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. Nucleic Acids Res 19:155-162

Boudreau E, Nickelsen J, Lemaire SD, Ossenbuhl F, Rochaix JD (2000) The Nac2 gene of *Chlamydomonas* encodes a chloroplast TPR-like protein involved in psbD mRNA stability. EMBO J 19:3366-3376

Breidenbach E, Jenni E, Boschetti A (1988) Synthesis of two proteins in chloroplasts and mRNA distribution between thylakoids and stroma during the cell cycle of *Chlamydomonas reinhardii*. Eur J Biochem 177:225-232

Breitenberger CA, Graves MC, Spremulli LL (1979) Evidence for the nuclear location of the gene for chloroplast elongation factor G. Arch Biochem Biophys 194:265-270

Bruick RK, Mayfield SP (1998) Processing of the psbA 5' untranslated region in *Chlamydomonas reinhardtii* depends upon factors mediating ribosome association. J Cell Biol 143:1145-1153

Cahoon AB, Timko MP (2000) yellow-in-the-dark mutants of chlamydomonas lack the CHLL subunit of light-independent protochlorophyllide reductase. Plant Cell 12:559-568

Campos F, Garcia-Gomez BI, Solorzano RM, Salazar E, Estevez J, Leon P, Alvarez-Buylla ER, Covarrubias AA (2001) A cDNA for nuclear-encoded chloroplast translational initiation factor 2 from a higher plant is able to complement an infB *Escherichia coli* null mutant. J Biol Chem 276:28388-28394

Choquet Y, Stern DB, Wostrikoff K, Kuras R, Girard-Bascou J, Wollman FA (1998) Translation of cytochrome f is autoregulated through the 5' untranslated region of petA mRNA in *Chlamydomonas* chloroplasts. Proc Natl Acad Sci USA 95:4380-4385

Choquet Y, Vallon O (2000) Synthesis, assembly and degradation of thylakoid membrane proteins. Biochimie 82:615-634

Choquet Y, Wollman FA (2002) Translational regulations as specific traits of chloroplast gene expression. FEBS Lett 529:39-42

Choquet Y, Wostrikoff K, Rimbault B, Zito F, Girard-Bascou J, Drapier D, Wollman FA (2001) Assembly-controlled regulation of chloroplast gene translation. Biochem Soc Trans 29:421-426

Choquet Y, Zito F, Wostrikoff K, Wollman FA (2003) Cytochrome f translation in *Chlamydomonas* chloroplast is autoregulated by its carboxyl-terminal domain. Plant Cell 15:1443-1454

Chua NH, Blobel G, Siekevitz P, Palade GE (1973) Attachment of chloroplast polysomes to thylakoid membranes in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA 70:1554-1558

Chua NH, Blobel G, Siekevitz P, Palade GE (1976) Periodic variations in the ratio of free to thylakoid-bound chloroplast ribosomes during the cell cycle of *Chlamydomonas reinhardtii*. J Cell Biol 71:497-514

Cohen I, Knopf JA, Irihimovitch V, Shapira M (2005) A proposed mechanism for the inhibitory effects of oxidative stress on Rubisco assembly and its subunit expression. Plant Physiol 137:738-746

Danon A (1997) Translational regulation in the chloroplast. Plant Physiol 115:1293-1298

Danon A, Mayfield SP (1991) Light regulated translational activators: identification of chloroplast gene specific mRNA binding proteins. EMBO J 10:3993-4001

Danon A, Mayfield SP (1994a) ADP-dependent phosphorylation regulates RNA-binding *in vitro*: implications in light-modulated translation. EMBO J 13:2227-2235

Danon A, Mayfield SP (1994b) Light-regulated translation of chloroplast messenger RNAs through redox potential. Science 266:1717-1719

Dauvillee D, Stampacchia O, Girard-Bascou J, Rochaix JD (2003) Tab2 is a novel conserved RNA binding protein required for translation of the chloroplast psaB mRNA. EMBO J 22:6378-6388

Drapier D, Girard-Bascou J, Wollman FA (1992) Evidence for nuclear control of the expression of the atpA and atpB chloroplast genes in *Chlamydomonas*. Plant Cell 4:283-295

Dron M, Rahire M, Rochaix JD (1982) Sequence of the chloroplast 16S rRNA gene and its surrounding regions of *Chlamydomonas reinhardii*. Nucleic Acids Res 10:7609-7620

Eberhard S, Drapier D, Wollman FA (2002) Searching limiting steps in the expression of chloroplast-encoded proteins: relations between gene copy number, transcription, transcript abundance and translation rate in the chloroplast of *Chlamydomonas reinhardtii*. Plant J 31:149-160

Edhofer I, Muhlbauer SK, Eichacker LA (1998) Light regulates the rate of translation elongation of chloroplast reaction center protein D1. Eur J Biochem 257:78-84

Eibl C, Zou Z, Beck a, Kim M, Mullet J, Koop HU (1999) *In vivo* analysis of plastid psbA, rbcL and rpl32 UTR elements by chloroplast transformation: tobacco plastid gene expression is controlled by modulation of transcript levels and translation efficiency. Plant J 19:333-345

Esposito D, Fey JP, Eberhard S, Hicks AJ, Stern DB (2003) *In vivo* evidence for the prokaryotic model of extended codon-anticodon interaction in translation initiation. EMBO J 22:651-656

Esposito D, Hicks AJ, Stern DB (2001) A role for initiation codon context in chloroplast translation. Plant Cell 13:2373-2384

Fargo DC, Boynton JE, Gillham NW (1999) Mutations altering the predicted secondary structure of a chloroplast 5' untranslated region affect its physical and biochemical properties as well as its ability to promote translation of reporter mRNAs both in the *Chlamydomonas reinhardtii* chloroplast and in *Escherichia coli*. Mol Cell Biol 19:6980-6990

Fargo DC, Boynton JE, Gillham NW (2001) Chloroplast ribosomal protein S7 of *Chlamydomonas* binds to chloroplast mRNA leader sequences and may be involved in translation initiation. Plant Cell 13:207-218

Fargo DC, Zhang M, Gillham NW, Boynton JE (1998) Shine-Dalgarno-like sequences are not required for translation of chloroplast mRNAs in *Chlamydomonas reinhardtii* chloroplasts or in *Escherichia coli*. Mol Gen Genet 257:271-282

Felder S, Meierhoff K, Sane AP, Meurer J, Driemel C, Plucken H, Klaff P, Stein B, Bechtold N, Westhoff P (2001) The nucleus-encoded HCF107 gene of *Arabidopsis* provides a link between intercistronic RNA processing and the accumulation of translation-competent psbH transcripts in chloroplasts. Plant Cell 13:2127-2141

Fong CL, Lentz A, Mayfield SP (2000) Disulfide bond formation between RNA binding domains is used to regulate mRNA binding activity of the chloroplast poly(A)-binding protein. J Biol Chem 275:8275-8278

Fox L, Erion J, Tarnowski J, Spremulli L, Brot N, Weissbach H (1980) *Euglena gracilis* chloroplast EF-Ts. Evidence that it is a nuclear-coded gene product. J Biol Chem 255:6018-6019

Fox T (1996) Translational control of endogenous and recoded nuclear genes in yeast mitochondria: regulation and membrane targeting. Experientia 52:1130-1135

Franch T, Gerdes K (1996) Programmed cell death in bacteria: translational repression by mRNA end-pairing. Mol Microbiol 21:1049-1060

Franzetti B, Carol P, Mache R (1992) Characterization and RNA-binding properties of a chloroplast S1-like ribosomal protein. J Biol Chem 267:19075-19081

Friso G, Giacomelli L, Ytterberg AJ, Peltier JB, Rudella A, Sun Q, Wijk KJ (2004) In-depth analysis of the thylakoid membrane proteome of *Arabidopsis thaliana* chloroplasts: new proteins, new functions, and a plastid proteome database. Plant Cell 16:478-499

Fromm H, Devic M, Fluhr R, Edelman M (1985) Control of psbA gene expression: in mature *Spirodela* chloroplasts light regulation of 32-kd protein synthesis is independent of transcript level. EMBO J 4:291-295

Gallie DR (1998) A tale of two termini: a functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation. Gene 216:1-11

Gamble PE, Mullet JE (1989) Translation and stability of proteins encoded by the plastid psbA and psbB genes are regulated by a nuclear gene during light-induced chloroplast development in barley. J Biol Chem 264:7236-7243

Gillham NW, Boynton JE, Hauser CR (1994) Translational regulation of gene expression in chloroplasts and mitochondria. Annu Rev Genet 28:71-93

Girard-Bascou J, Pierre Y, Drapier D (1992) A nuclear mutation affects the synthesis of the chloroplast psbA gene production *Chlamydomonas reinhardtii*. Curr Genet 22:47-52

Gold L (1988) Posttranscriptional regulatory mechanisms in *Escherichia coli*. Annu Rev Biochem 57:199-233

Gray MW (1993) Origin and evolution of organelle genomes. Curr Opin Genet Dev 3:884-890

Gruissem W, Tonkyn JC (1993) Control mechanisms of plastid gene-expression. Crit Rev Plant Sci 12:19-55

Hattori T, Margulies MM (1986) Synthesis of large subunit of ribulosebisphosphate carboxylase by thylakoid-bound polyribosomes from spinach chloroplasts. Arch Biochem Biophys 244:630-640

Hauser CR, Gillham NW, Boynton JE (1996) Translational regulation of chloroplast genes. Proteins binding to the 5'-untranslated regions of chloroplast mRNAs in *Chlamydomonas reinhardtii*. J Biol Chem 271:1486-1497

Hieronymus H, Silver PA (2004) A systems view of mRNP biology. Genes Dev 18:2845-2860

Higgs DC, Shapiro RS, Kindle KL, Stern DB (1999) Small cis-acting sequences that specify secondary structures in a chloroplast mRNA are essential for RNA stability and translation. Mol Cell Biol 19:8479-8491

Hirose T, Kusumegi T, Sugiura M (1998) Translation of tobacco chloroplast rps14 mRNA depends on a Shine-Dalgarno-like sequence in the 5'-untranslated region but not on internal RNA editing in the coding region. FEBS Lett 430:257-260

Hirose T, Sugiura M (1996) Cis-acting elements and trans-acting factors for accurate translation of chloroplast psbA mRNAs: development of an *in vitro* translation system from tobacco chloroplasts. EMBO J 15:1687-1695

Hirose T, Sugiura M (2004a) Functional Shine-Dalgarno-like sequences for translational initiation of chloroplast mRNAs. Plant Cell Physiol 45:114-117

Hirose T, Sugiura M (2004b) Multiple elements required for translation of plastid atpB mRNA lacking the Shine-Dalgarno sequence. Nucleic Acids Res 32:3503-3510

Hotchkiss TL, Hollingsworth MJ (1999) ATP synthase 5' untranslated regions are specifically bound by chloroplast polypeptides. Curr Genet 35:512-520

Jackson RJ (2005) Alternative mechanisms of initiating translation of mammalian mRNAs. Biochem Soc Trans 33:1231-1241

Katz YS, Danon A (2002) The 3'-untranslated region of chloroplast psbA mRNA stabilizes binding of regulatory proteins to the leader of the message. J Biol Chem 277:18665-18669

Keene JD, Lager PJ (2005) Post-transcriptional operons and regulons co-ordinating gene expression. Chromosome Res 13:327-337

Khrebtukova I, Spreitzer RJ (1996) Elimination of the *Chlamydomonas* gene family that encodes the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. Proc Natl Acad Sci USA 93:13689-13693

Kim J, Klein PG, Mullet JE (1991) Ribosomes pause at specific sites during synthesis of membrane-bound chloroplast reaction center protein D1. J Biol Chem 266:14931-14938

Kim J, Klein PG, Mullet JE (1994a) Synthesis and turnover of photosystem II reaction center protein D1. Ribosome pausing increases during chloroplast development. J Biol Chem 269:17918-17923

Kim J, Klein PG, Mullet JE (1994b) Vir-115 gene product is required to stabilize D1 translation intermediates in chloroplasts. Plant Mol Biol 25:459-467

Kim J, Mayfield SP (1997) Protein disulfide isomerase as a regulator of chloroplast translational activation. Science 278:1954-1957

Kim J, Mullet JE (1994) Ribosome-binding sites on chloroplast rbcL and psbA mRNAs and light-induced initiation of D1 translation. Plant Mol Biol 25:437-448

Kim J, Mullet JE (2003) A mechanism for light-induced translation of the rbcL mRNA encoding the large subunit of ribulose-1,5-bisphosphate carboxylase in barley chloroplasts. Plant Cell Physiol 44:491-499

Klein RR, Mason HS, Mullet JE (1988) Light-regulated translation of chloroplast proteins. I. Transcripts of psaA-psaB, psbA, and rbcL are associated with polysomes in dark-grown and illuminated barley seedlings. J Cell Biol 106:289-301

Klinkert B, Elles I, Nickelsen J (2006) Translation of chloroplast psbD mRNA in *Chlamydomonas* is controlled by a secondary RNA structure blocking the AUG start codon. Nucleic Acids Res 34:386-394

Koo JS, Spremulli LL (1994) Analysis of the translational initiation region on the *Euglena gracilis* chloroplast ribulose-bisphosphate carboxylase/oxygenase (rbcL) messenger RNA. J Biol Chem 269:7494-7500

Kozak M (2005) Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene 361:13-37

Kramzar LM, Mueller T, Erickson B, Higgs DC (2006) Regulatory sequences of orthologous petD chloroplast mRNAs are highly specific among *Chlamydomonas* species. Plant Mol Biol 60:405-422

Krupinska k, Apel k (1989) Light-induced transformation of etioplasts to chloroplasts of barley without transcriptional control of plastid gene expression. Mol Gen Genet 219:467-473

Kuchka MR, Goldschmidt-Clermont M, van Dillewijn J, Rochaix JD (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

Kuchka MR, Mayfield SP, Rochaix JD (1988) Nuclear mutations specifically affect the synthesis and/or degradation of the chloroplast-encoded D2 polypeptide of photosystem II in *Chlamydomonas reinhardtii*. EMBO J 7:319-324

Kuras R, Wollman FA (1994) The assembly of cytochrome b6/f complexes: an approach using genetic transformation of the green alga *Chlamydomonas reinhardtii*. EMBO J 13:1019-1027

Laursen BS, Sorensen HP, Mortensen KK, Sperling-Petersen HU (2005) Initiation of protein synthesis in bacteria. Microbiol Mol Biol Rev 69:101-123

Lennartz K, Bossmann S, Westhoff P, Bechtold N, Meierhoff K (2006) HCF153, a novel nuclear-encoded factor necessary during a post-translational step in biogenesis of the cytochrome bf complex. Plant J 45:101-112

Lennartz K, Plucken H, Seidler A, Westhoff P, Bechtold N, Meierhoff K (2001) HCF164 encodes a thioredoxin-like protein involved in the biogenesis of the cytochrome b(6)f complex in *Arabidopsis*. Plant Cell 13:2539-2551

Lin Q, Yu NJ, Spremulli LL (1996) Expression and functional analysis of *Euglena Gracilis* chloroplast initiation factor 3. Plant Mol Biol 32:937-945

Lindahl L, Hinnebusch A (1992) Diversity of mechanisms in the regulation of translation in prokaryotes and lower eukaryotes. Curr Opin Genet Dev 2:720-726

Lorkovic ZJ, Barta A (2002) Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant *Arabidopsis thaliana*. Nucleic Acids Res 30:623-635

Lurin C, Andres C, Aubourg S, Bellaoui M, Bitton F, Bruyere C, Caboche M, Debast C, Gualberto J, Hoffmann B, Lecharny A, Le Ret M, Martin-Magniette ML, Mireau H, Peeters N, Renou JP, Szurek B, Taconnat L, Small I (2004) Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell 16:2089-2103

Ma L, Spremulli LL (1990) Identification and characterization of large, complex forms of chloroplast translational initiation factor 2 from *Euglena gracilis*. J Biol Chem 265:13560-13565

Maidak BL, Olsen GJ, Larsen N, Overbeek R, McCaughey MJ, Woese CR (1996) The Ribosomal Database Project (RDP). Nucleic Acids Res 24:82-85

Malnoe P, Mayfield SP, Rochaix JD (1988) Comparative analysis of the biogenesis of photosystem II in the wild-type and Y-1 mutant of *Chlamydomonas reinhardtii*. J Cell Biol 106:609-616

Mangus DA, Evans MC, Jacobson A (2003) Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. Genome Biol 4:223

Manuell A, Beligni MV, Yamaguchi K, Mayfield SP (2004) Regulation of chloroplast translation: interactions of RNA elements, RNA-binding proteins and the plastid ribosome. Biochem Soc Trans 32:601-605

Margulies MM (1983) Synthesis of photosynthetic membrane proteins directed by RNA from rough thylakoids of *Chlamydomonas reinhardtii*. Eur J Biochem 137:241-248

Margulies MM, Michaels A (1974) Ribosomes bound to chloroplast membranes in *Chlamydomonas reinhardtii*. J Cell Biol 60:65-77

Margulies MM, Michaels A (1975) Free and membrane-bound chloroplast polyribosomes *Chlamydomonas reinhardtii*. Biochim Biophys Acta 402:297-308

Margulies MM, Tiffany HL, Hattori T (1987) Photosystem I reaction center polypeptides of spinach are synthesized on thylakoid-bound ribosomes. Arch Biochem Biophys 254:454-461

Margulies MM, Weistrop JS (1980) Sub-thylakoid fractions containing ribosomes. Biochim Biophys Acta 606:20-33

Martin W, Herrmann RG (1998) Gene transfer from organelles to the nucleus: how much, what happens, and Why? Plant Physiol 118:9-17

Mayfield SP, Cohen A, Danon A, Yohn CB (1994) Translation of the psbA mRNA of *Chlamydomonas reinhardtii* requires a structured RNA element contained within the 5' untranslated region. J Cell Biol 127:1537-1545

Mayfield SP, Yohn CB, Cohen A, Danon A (1995) Regulation of chloroplast gene expression. Annu Rev Plant Physiol Plant Mol Biol 46:147-166

McCormac DJ, Barkan A (1999) A nuclear gene in maize required for the translation of the chloroplast atpB/E mRNA. Plant Cell 11:1709-1716

Merendino L, Falciatore A, Rochaix JD (2003) Expression and RNA binding properties of the chloroplast ribosomal protein S1 from *Chlamydomonas reinhardtii*. Plant Mol Biol 53:371-382

Merhige PM, Both-Kim D, Robida MD, Hollingsworth MJ (2005) RNA-protein complexes that form in the spinach chloroplast atpI 5' untranslated region can be divided into two subcomplexes, each comprised of unique cis-elements and trans-factors. Curr Genet 48:256-264

Meurer J, Grevelding C, Westhoff P, Reiss B (1998) The PAC protein affects the maturation of specific chloroplast mRNAs in *Arabidopsis thaliana*. Mol Gen Genet 258:342-351

Meurer J, Meierhoff K, Westhoff P (1996) Isolation of high-chlorophyll-fluorescence mutants of *Arabidopsis thaliana* and their characterisation by spectroscopy, immunoblotting and northern hybridisation. Planta 198:385-396

Minai L, Wostrikoff K, Wollman FA, Choquet Y (2006) Chloroplast biogenesis of photosystem II cores involves a series of assembly-controlled steps that regulate translation. Plant Cell 18:159-175

Minami E, Watanabe A (1984) Thylakoid membranes: the translational site of chloroplast DNA-regulated thylakoid polypeptides. Arch Biochem Biophys 235:562-570

Moll I, Grill S, Gualerzi CO, Blasi U (2002) Leaderless mRNAs in bacteria: surprises in ribosomal recruitment and translational control. Mol Microbiol 43:239-246

Monde RA, Greene JC, Stern DB (2000a) The sequence and secondary structure of the 3'-UTR affect 3'-end maturation, RNA accumulation, and translation in tobacco chloroplasts. Plant Mol Biol 44:529-542

Monde RA, Zito F, Olive J, Wollman FA, Stern DB (2000b) Post-transcriptional defects in tobacco chloroplast mutants lacking the cytochrome b6/f complex. Plant J 21:61-72

Muhlbauer SK, Eichacker LA (1998) Light-dependent formation of the photosynthetic proton gradient regulates translation elongation in chloroplasts. J Biol Chem 273:20935-20940

Muhlbauer SK, Eichacker LA (1999) The stromal protein large subunit of ribulose-1,5-bisphosphate carboxylase is translated by membrane-bound ribosomes. Eur J Biochem 261:784-788

Mullet JE (1988) Chloroplast development and gene-expression. Annu Rev Plant Physiol Plant Mol Biol 39:475-502

Mullet JE, Klein PG, Klein RR (1990) Chlorophyll regulates accumulation of the plastid-encoded chlorophyll apoproteins CP43 and D1 by increasing apoprotein stability. Proc Natl Acad Sci USA 87:4038-4042

Nakamura T, Meierhoff K, Westhoff P, Schuster G (2003) RNA-binding properties of HCF152, an *Arabidopsis* PPR protein involved in the processing of chloroplast RNA. Eur J Biochem 270:4070-4081

Nickelsen J (1999) Transcripts containing the 5' untranslated regions of the plastid genes psbA and psbB from higher plants are unstable in *Chlamydomonas reinhardtii* chloroplasts. Mol Gen Genet 262:768-771

Nickelsen J (2003) Chloroplast RNA-binding proteins. Curr Genet 43:392-399

Nickelsen J, Fleischmann M, Boudreau E, Rahire M, Rochaix JD (1999) Identification of cis-acting RNA leader elements required for chloroplast psbD gene expression in *Chlamydomonas*. Plant Cell 11:957-970

Nickelsen J, van Dillewijn J, Rahire M, Rochaix JD (1994) Determinants for stability of the chloroplast psbD RNA are located within its short leader region in *Chlamydomonas reinhardtii*. EMBO J 13:3182-3191

Ossenbuhl F, Nickelsen J (2000) cis- and trans-Acting determinants for translation of psbD mRNA in *Chlamydomonas reinhardtii*. Mol Cell Biol 20:8134-8142

Plader W, Sugiura M (2003) The Shine-Dalgarno-like sequence is a negative regulatory element for translation of tobacco chloroplast rps2 mRNA: an additional mechanism for translational control in chloroplasts. Plant J 34:377-382

Plucken H, Muller B, Grohmann D, Westhoff P, Eichacker LA (2002) The HCF136 protein is essential for assembly of the photosystem II reaction center in *Arabidopsis thaliana*. FEBS Lett 532:85-90

Race HL, Herrmann RG, Martin W (1999) Why have organelles retained genomes? Trends Genet 15:364-370

Robida MD, Merhige PM, Hollingsworth MJ (2002) Proteins are shared among RNA-protein complexes that form in the 5' untranslated regions of spinach chloroplast mRNAs. Curr Genet 41:53-62

Rochaix JD (2001) Posttranscriptional control of chloroplast gene expression. From RNA to photosynthetic complex. Plant Physiol 125:142-144

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

Rodermel S (1999) Subunit control of Rubisco biosynthesis – a relic of an endosymbiotic past? Photosynthesis Res 59:105–123

Rodermel S, Haley J, Jiang CZ, Tsai CH, Bogorad L (1996) A mechanism for intergenomic integration: abundance of ribulose bisphosphate carboxylase small-subunit protein influences the translation of the large-subunit mRNA. Proc Natl Acad Sci USA 93:3881-3885

Rodermel SR, Abbott MS, Bogorad L (1988) Nuclear-organelle interactions: nuclear antisense gene inhibits ribulose bisphosphate carboxylase enzyme levels in transformed tobacco plants. Cell 55:673-681

Rott R, Levy H, Drager RG, Stern DB, Schuster G (1998) 3'-Processed mRNA is preferentially translated in *Chlamydomonas reinhardtii* chloroplasts. Mol Cell Biol 18:4605-4611

Sakamoto W, Chen X, Kindle KL, Stern DB (1994) Function of the *Chlamydomonas reinhardtii* petd 5' untranslated region in regulating the accumulation of subunit IV of the cytochrome b6/f complex. Plant J 6:503-512

Sane AP, Stein B, Westhoff P (2005) The nuclear gene HCF107 encodes a membrane-associated R-TPR (RNA tetratricopeptide repeat)-containing protein involved in expression of the plastidial psbH gene in *Arabidopsis*. Plant J 42:720-730

Sato N, Rolland N, Block MA, Joyard J (1999) Do plastid envelope membranes play a role in the expression of the plastid genome? Biochimie 81:619-629

Schmitz-Linneweber C, Williams-Carrier R, Barkan A (2005) RNA immunoprecipitation and microarray analysis show a chloroplast pentatricopeptide repeat protein to be associated with the 5' region of mRNAs whose translation it activates. Plant Cell 17:2791-2804

Sengupta J, Agrawal RK, Frank J (2001) Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. Proc Natl Acad Sci USA 98:11991-11996

Shen Y, Danon A, Christopher DA (2001) RNA binding-proteins interact specifically with the *Arabidopsis* chloroplast psbA mRNA 5' untranslated region in a redox-dependent manner. Plant Cell Physiol 42:1071-1078

Shinohara K, Minami E, Watanabe A (1988) Synthesis and assembly of H+-ATPase complex by isolated "rough" thylakoids. Arch Biochem Biophys 260:452-460

Shteiman-Kotler A, Schuster G (2000) RNA-binding characteristics of the chloroplast S1-like ribosomal protein CS1. Nucleic Acids Res 28:3310-3315

Singh BN, Mishra RN, Agarwal PK, Goswami M, Nair S, Sopory SK, Reddy MK (2004) A pea chloroplast translation elongation factor that is regulated by abiotic factors. Biochem Biophys Res Commun 320:523-530

Somanchi A, Barnes D, Mayfield SP (2005) A nuclear gene of *Chlamydomonas reinhardtii*, Tba1, encodes a putative oxidoreductase required for translation of the chloroplast psbA mRNA. Plant J 42:341-352

Sorensen MA, Fricke J, Pedersen S (1998) Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in *Escherichia coli in vivo*. J Mol Biol 280:561-569

Sreedharan SP, Beck CM, Spremulli LL (1985) *Euglena gracilis* chloroplast elongation factor Tu. Purification and initial characterization. J Biol Chem 260:3126-3131

Stampacchia O, Girard-Bascou J, Zanasco JL, Zerges W, Bennoun P, Rochaix JD (1997) A nuclear-encoded function essential for translation of the chloroplast psaB mRNA in *Chlamydomonas*. Plant Cell 9:773-782

Staub JM, Maliga P (1994) Translation of psbA mRNA is regulated by light via the 5'-untranslated region in tobacco plastids. Plant J 6:547-553

Steege DA, Graves MC, Spremulli LL (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

Stern DB, Higgs DC, Yang JJ (1997) Transcription and translation in chloroplasts. Trends in Plant Science 2:308-315

Stollar NE, Kim JK, Hollingsworth MJ (1994) Ribosomes pause during the expression of the large ATP synthase gene cluster in spinach chloroplasts. Plant Physiol 105:1167-1177

Subramanian AR (1983) Structure and functions of ribosomal protein S1. Prog Nucleic Acid Res Mol Biol 28:101-142

Sugita M, Sugita C, Sugiura M (1995) Structure and expression of the gene encoding ribosomal protein S1 from the cyanobacterium *Synechococcus* sp. strain PCC 6301: striking sequence similarity to the chloroplast ribosomal protein CS1. Mol Gen Genet 246:142-147

Sugiura M, Hirose T, Sugita M (1998) Evolution and mechanism of translation in chloroplasts. Annu Rev Genet 32:437-459

Taniguchi M, Kuroda H, Satoh K (1993) ATP-dependent protein synthesis in isolated pea chloroplasts. Evidence for accumulation of a translation intermediate of the D1 protein. FEBS Lett 317:57-61

Trebitsh T, Danon A (2001) Translation of chloroplast psbA mRNA is regulated by signals initiated by both photosystems II and I. Proc Natl Acad Sci USA 98:12289-12294

Trebitsh T, Levitan A, Sofer A, Danon A (2000) Translation of chloroplast psbA mRNA is modulated in the light by counteracting oxidizing and reducing activities. Mol Cell Biol 20:1116-1123

Trebitsh T, Meiri E, Ostersetzer O, Adam Z, Danon A (2001) The protein disulfide isomerase-like RB60 is partitioned between stroma and thylakoids in *Chlamydomonas reinhardtii* chloroplasts. J Biol Chem 276:4564-4569

van Wijk KJ (2004) Plastid proteomics. Plant Physiol Biochem 42:963-977

Voorma HO (1996) Control of translation initiation in prokaryotes In: Hershey JWB, Mathews MB, Sonenberg N (eds) Translational control. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp 759-777

Wang CC, Spremulli LL (1991) Chloroplast translational initiation factor 3. Purification and characterization of multiple forms from *Euglena gracilis*. J Biol Chem 266:17079-17083

Wostrikoff K, Choquet Y, Wollman FA, Girard-Bascou J (2001) TCA1, a single nuclear-encoded translational activator specific for petA mRNA in *Chlamydomonas reinhardtii* chloroplast. Genetics 159:119-132

Wostrikoff K, Girard-Bascou J, Wollman FA, Choquet Y (2004) Biogenesis of PSI involves a cascade of translational autoregulation in the chloroplast of *Chlamydomonas*. EMBO J 23:2696-2705

Yamaguchi K, Beligni MV, Prieto S, Haynes PA, McDonald WH, Yates JR 3rd, Mayfield SP (2003) Proteomic characterization of the *Chlamydomonas reinhardtii* chloroplast ribosome. Identification of proteins unique to th e70 S ribosome. J Biol Chem 278:33774-33785

Yamaguchi K, Prieto S, Beligni MV, Haynes PA, McDonald WH, Yates JR 3rd, Mayfield SP (2002) Proteomic characterization of the small subunit of *Chlamydomonas reinhardtii* chloroplast ribosome: identification of a novel S1 domain-containing protein and unusually large orthologs of bacterial S2, S3, and S5. Plant Cell 14:2957-2974

Yamaguchi K, Subramanian AR (2000) The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast). J Biol Chem 275:28466-28482

Yamaguchi K, von Knoblauch K, Subramanian AR (2000) The plastid ribosomal proteins. Identification of all the proteins in the 30 S subunit of an organelle ribosome (chloroplast). J Biol Chem 275:28455-28465

Yohn CB, Cohen A, Danon A, Mayfield SP (1996) Altered mRNA binding activity and decreased translational initiation in a nuclear mutant lacking translation of the chloroplast psbA mRNA. Mol Cell Biol 16:3560-3566

Yohn CB, Cohen A, Danon A, Mayfield SP (1998a) A poly(A) binding protein functions in the chloroplast as a message-specific translation factor. Proc Natl Acad Sci USA 95:2238-2243

Yohn CB, Cohen A, Rosch C, Kuchka MR, Mayfield SP (1998b) Translation of the chloroplast psbA mRNA requires the nuclear-encoded poly(A)-binding protein, RB47. J Cell Biol 142:435-442

Yu NJ, Spremulli LL (1998) Regulation of the activity of chloroplast translational initiation factor 3 by NH2- and COOH-terminal extensions. J Biol Chem 273:3871-3877

Zerges W (2000) Translation in chloroplasts. Biochimie 82:583-601

Zerges W, Auchincloss AH, Rochaix JD (2003) Multiple translational control sequences in the 5' leader of the chloroplast psbC mRNA interact with nuclear gene products in *Chlamydomonas reinhardtii*. Genetics 163:895-904

Zerges W, Girard-Bascou J, Rochaix JD (1997) Translation of the chloroplast psbC mRNA is controlled by interactions between its 5' leader and the nuclear loci TBC1 and TBC3 in *Chlamydomonas reinhardtii*. Mol Cell Biol 17:3440-3448

Zerges W, Rochaix JD (1994) The 5' leader of a chloroplast mRNA mediates the translational requirements for two nucleus-encoded functions in *Chlamydomonas reinhardtii*. Mol Cell Biol 14:5268-5277

Zerges W, Rochaix JD (1998) Low density membranes are associated with RNA-binding proteins and thylakoids in the chloroplast of *Chlamydomonas reinhardtii*. J Cell Biol 140:101-110

Zhang L, Aro EM (2002) Synthesis, membrane insertion and assembly of the chloroplast-encoded D1 protein into photosystem II. FEBS Lett 512:13-18
Zhang L, Paakkarinen V, van Wijk KJ, Aro EM (1999) Co-translational assembly of the D1 protein into photosystem II. J Biol Chem 274:16062-16067
Zhang L, Paakkarinen V, van Wijk KJ, Aro EM (2000) Biogenesis of the chloroplast-encoded D1 protein: regulation of translation elongation, insertion, and assembly into photosystem II. Plant Cell 12:1769-1782
Zou Z, Eibl C, Koop HU (2003) The stem-loop region of the tobacco psbA 5'UTR is an important determinant of mRNA stability and translation efficiency. Mol Genet Genomics 269:340-349

Danon, Avihai
Department of Plant Sciences, Weizmann Institute of Science, Rehovot 76100, Israel
avihai.danon@weizmann.ac.il

Peled-Zehavi, Hadas
Department of Plant Sciences, Weizmann Institute of Science, Rehovot 76100, Israel

Assembly of protein complexes in plastids

Eira Kanervo, Marjaana Suorsa, and Eva-Mari Aro

Abstract

Photosynthetic multiprotein complexes in plants and cyanobacteria are mainly responsible for the function of the oxygenic photosynthesis. Great progress has recently been made in resolving the structures of these complexes, most of which are now known at 2 to 4Å resolution. Compared to these achievements, amazingly little is known about the biogenesis, maintenance, and stability of these macromolecular photosynthetic complexes. So far, the sequential assembly of the Photosystem II subunits is best characterized, yet the ligation of redox co-factors and other pigments still remain only poorly understood. There seems to be a general CES control of translation of the key chloroplast-encoded subunits of all thylakoid protein complexes thus ensuring a coordinated synthesis and assembly of the chloroplast- and nucleus-encoded subunits. Some light has also recently been shed on the function and abundance of the auxiliary proteins necessary for the translocation of the nucleus-encoded proteins into chloroplasts and for facilitating the assembly processes of the macromolecular photosynthetic protein complexes.

1 Introduction

Photosynthetic membrane protein complexes comprise the marvelous machinery that provides energy for all living forms on Earth. During the past few years the knowledge on the 3D structure of the photosynthetic protein complexes has advanced tremendously. Photosystem II (PSII) and Photosystem I (PSI) structures have been resolved to nearly atomic resolution, and in addition to the protein subunits, the coordinates for different ligands and co-factors are largely known (for PSII, see Zouni et al. 2001; Ferreira et al. 2004; Loll et al. 2005; for PSI see Ben-Shem et al. 2003). Similarly, the structure of the cytochrome b_6f (Cyt b_6f) complex was resolved both from the green alga *Chlamydomonas reinhardtii* (Stroebel et al. 2003) and cyanobacteria (Kurisu et al. 2003) as well as the structure of the ATP synthase (Seelert et al. 2000). It is a big challenge now to clarify the mechanisms, how these multisubunit complexes are properly assembled in the thylakoid membrane. Most of the structural subunits of these multiprotein complexes have been identified by now, yet it is not exceptional that new subunits are still recognized (e.g. Khrouchtchova et al. 2005). However, the assembly mechanisms of the subunits into the multiprotein complexes are far from being resolved, and we are only in the very beginning of understanding of the number and the functions of the

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0236 / Published online: 4 July 2007

auxiliary proteins that finally guarantee the proper synthesis, assembly, and stability of the protein subunits. Furthermore, the mechanisms of the ligation of different co-factors to the multiprotein complexes still largely remain to be elucidated.

In general, the synthesis and assembly of the main photosynthetic multiprotein complexes in plastids require concerted interactions between the nucleo-cytosolic and plastid genetic systems. These interactions occur mostly at the translational and posttranslational levels and are controlled by the nuclear-encoded regulatory factors (Rochaix 1996; Wollman et al. 1999). A common feature in the assembly of the photosynthetic multiprotein complexes in chloroplasts seems to be an assembly-dependent autoregulation of translation of the central chloroplast-encoded subunits that makes the core of all thylakoid protein complexes. This phenomenon has been defined as a CES process - Control by Epistasy of Synthesis (Wollman et al. 1999; Choquet and Vallon 2000; Wostrikoff et al. 2004; Minai et al. 2006). In the CES process, the translation of a CES protein is dependent on the presence of a specific dominant protein (or protein sub-complex) thus representing a protein-assembly-mediated autoregulation of translation. Furthermore, a strict quality control is constantly operating in chloroplasts to guarantee a proper stoichiometry of the protein subunits for efficient assembly of the complexes (Yamamoto 2001).

Besides structural subunits, the photosynthetic membrane protein complexes PSI, PSII, and the Cyt b_6f complex contain pigments and co-factors that need to be ligated to the proteins during or after the assembly of the complexes, or even concomitantly with the translation process. Assembly of the protein subunits occurs sequentially and may require the interaction with several soluble or membrane-bound chaperones, or assembly factors.

The distinct assembly steps of the multiprotein complexes in plastids are not properly known yet. PSII and the Cyt b_6f complex are by far the protein complexes whose biogenesis and assembly have been studied in greatest detail, particularly of PSII, since it is the major target for the photo-destructive processes. Experimental evidence exists on several individual assembly steps of PSII with distinct sub-complexes, the processes, which will be reviewed in the following. In addition, the assembly of PSI and Cyt b_6f are shortly reviewed, as well as the assembly of Rubisco and ferredoxin:thioredoxin reductase (FTR) as examples of the soluble plastid complexes. Focus will be put especially on the assembly of the protein subunits, since our knowledge on the mechanisms of co-factor ligation into the complexes still remains poor. Furthermore, a short survey is presented on post-translational modifications of plastid proteins that affect the turnover and assembly/disassembly of the protein subunits in the thylakoid membrane complexes.

2 Assembly of the protein complexes

2.1 Assembly of PSII

The PSII complex contains 29 different subunits, from which 15 are plastid-encoded (PsbA-PsbF, PsbH-PsbN, PsbTc, and PsbZ), the rest of them being nucleus-encoded (for a review see van Wijk 2001; Rochaix 2006). The nucleus-encoded PSII proteins are synthesized on cytoplasmic ribosomes as precursor proteins, which contain an N-terminal transit peptide for plastid targeting and for determining the destination of the protein inside the chloroplast. The nucleus-encoded PSII proteins include proteins, such as PsbR, PsbW, and PsbY, three oxygen- evolving complex (OEC) proteins (PsbO, PsbP, and PsbQ), six Lhcb proteins (Lhcb1-6), and PsbS, which also belongs to the Lhcb family of proteins. The plastid-encoded PSII proteins are mainly integral membrane proteins, which are synthesized on thylakoid-bound ribosomes. Of these proteins particularly the reaction center protein D1 has been shown to be co-translationally inserted into the thylakoid membrane. Likewise, the assembly of the D1 protein to PSII during the repair process of photodamaged PSII centers was shown to occur co-translationally (Klein et al. 1988; Keegstra and Cline 1999; Zhang et al. 1999).

From the methodological point of view, the characterization of gene interruption or knockout mutants and the studies on plastid development from etioplast to chloroplast have been used to get insights into the assembly order of the protein subunits to PSII. Furthermore, isolated, intact chloroplasts have been subjected to the approaches, such as pulse and chase experiments followed by subfractionation of the various PSII subassemblies by sucrose density centrifugation (van Wijk et al. 1995; Müller and Eichacker 1999; Zhang et al. 1999). However, these latter experiments could only reveal the assembly of the major chloroplast-encoded PSII proteins D1, D2, CP43, and CP47, but failed to reveal the synthesis and assembly of the low-molecular-mass (LMM) subunits and the nucleus-encoded subunits.

A more thorough insight into the assembly steps of PSII proteins was received recently using different chromatographic methods and the two-dimensional blue native (BN)/SDS-PAGE system for separation of both the *in vitro* and *in vivo* labeled and assembled thylakoid proteins and protein complexes (e.g. Rokka et al. 2005; Nowaczyk et al. 2006). For resolving the mechanisms of assembly-dependent autoregulation of translation, an approach of chimeric gene constructs and their expression under the control of 5′UTRs of the genes of interest have been employed (Minai et al. 2006). Moreover, research on the light-induced turnover of the PSII complex has provided information that has also been applied for the research on the assembly process of a new PSII center.

2.1.1. Assembly of the PSII core monomers and dimers

The prerequisite for PSII assembly is the presence of α- and β- subunits of cytochrome b_{559} (Cyt b_{559}), which accumulate in the thylakoid membrane even in the absence of other PSII subunits (Morais et al. 1998; Müller and Eichacker 1999). Cyt b_{559} interacts with the D2 protein to form an initial complex that further serves as a receptor for the co-translational assembly of the D1 protein (Komenda et al. 2004) (Fig. 1). Indeed, evidence was recently provided indicating that the translation of D1 (a CES subunit) is strongly decreased in the absence of D2 (Minai et al. 2006). Also light is required for an efficient translation elongation and accumulation of the D1 protein, most probably due to the requirement of light for the synthesis of the pigment and other co-factor molecules that are ligated to the PSII

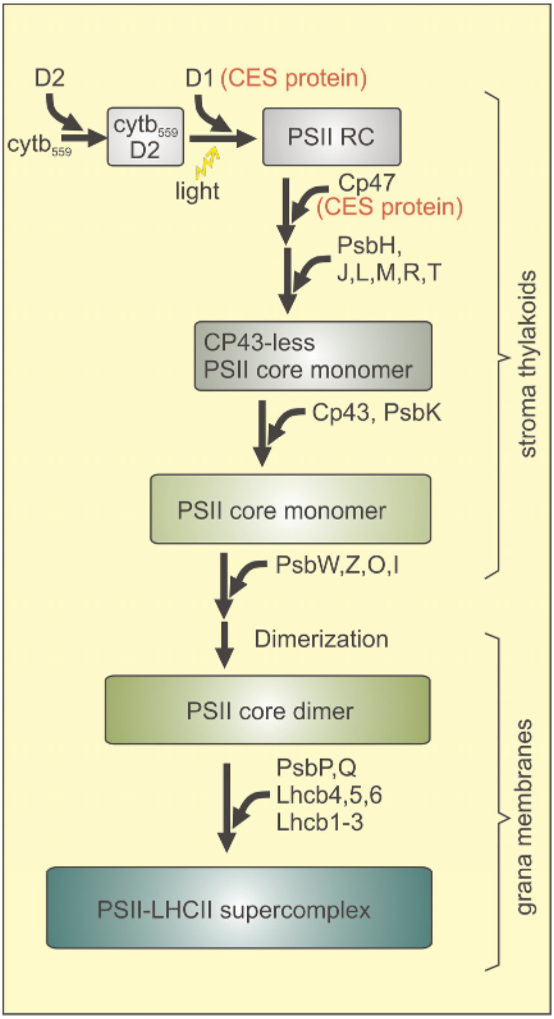


Fig. 1. Hypothetical scheme of the sequential protein assembly during biogenesis of the PSII complexes. Note that a hierarchical CES control is functioning for the translation of the D1 and CP47 proteins. So far, no data has been published on the assembly order of the PsbN and PsbY proteins, or PsbS.

complex concomitantly with the assembly process (Kim et al.1991; van Wijk and Eichacker 1996; Edhofer et al. 1998). Rapidly after the termination of translation, the D1 protein undergoes C-terminal processing (Diner et al. 1988; Bowyer et al. 1992) by the carboxyl-terminal processing protease CtpA that functions exclusively in the processing of the D1 protein in the thylakoid lumen (Anbudurai et al. 1994; Oelmüller et al. 1996).

The assembly of the D1 protein to the Cyt b_{559}/D2 subcomplex is followed by the association of the CP47 protein (Sharma et al. 1997; Müller and Eichacker 1999; Tsiotis et al. 1999; Zhang et al. 1999; Szabò et al. 2001; Rokka et al. 2005). In fact, the presence of D1 is a prerequisite for the high-level translation of the core antenna subunit CP47. The biogenesis of PSII thus involves a CES cascade where translation of D1 is dependent on the presence of D2 and the translation of CP47, in turn, is dependent on the presence of D1 (Minai et al. 2006). After the assembly of CP47, the LMM subunits PsbH, PsbL, PsbM, PsbTc, PsbR, and also PsbJ associate with the growing PSII subcomplex (Suorsa et al. 2004; Rokka et al. 2005). These LMM subunits are thought to stabilize the D1/D2/Cyt b_{559}/CP47 subassembly of PSII. The subunits PsbL, PsbM, and PsbT are located in the monomer-monomer interphase (Loll et al. 2005) and therefore are also crucial for the dimerization of PSII, whereas PsbR and PsbJ have been shown to be essential for the stable assembly of the OEC (see below). In cyanobacteria, the PsbH protein was found to be associated with CP47 and to be important for the prompt incorporation of the newly-synthesized D1 protein to the PSII complex (Komenda et al. 2005).

The subsequent assembly steps involve the association of the core antenna protein CP43 and the LMM subunit PsbK (Suorsa et al. 2004; Rokka et al. 2005) (Fig. 1). PsbK is tightly bound to CP43, and the assembly of PsbK was shown to occur only in the presence of CP43 (Sugimoto and Takahashi 2003). Furthermore, PsbK was postulated to be required for the PSII core dimerization (Zheleva et al. 1998). Due to the location of PsbK in the periphery of the dimer complex (Loll et al. 2005), this interpretation, however, needs further examination. The PsbI subunit, which was earlier reported to be a component of the PSII reaction center complex, was recently assigned to have a role in the PSII dimerization (Schwenkert et al. 2006). It is natural that the LMM subunits in the monomer/monomer interphase are crucial for dimerization. Such a role was recently experimentally proven for PsbL, which was likewise shown to be required for the stable association of CP43 (Suorsa et al. 2004).

The PsbZ protein, as well as the nuclear-encoded PsbW, are probably the last subunits that assemble to the PSII core and thereby facilitate the assembly of the minor LHCII proteins (Swiatek et al. 2001), which, in turn, are required for binding of the trimers of the light-harvesting antenna complex (LHCII) to the PSII core dimer (Shi et al. 2000; Rokka et al. 2005). Furthermore, chlorophyll (Chl) *a* synthesis enhances the accumulation and stability of monomers and, subsequently, the dimerization of the PSII core monomers (Müller and Eichacker 1999).

Light and Chl biosynthesis are essential for the synthesis and stability of the core antenna proteins CP43 and CP47 (van Wijk and Eichacker 1996). In addition to light and the availability of chlorophyll and various assembly partners, the regulation of synthesis and assembly of the PSII complex involves the presence of a variety of other factors, such as pheophytin, β-carotene, Fe, Mn, and plastoquinone. However, pigments are not always needed for the core protein accumulation; for example, in etioplasts, isolated from dark-grown barley seedlings, a PSII pre-complex has been found to exist, consisting of Cyt b_{559}, the D2 protein and the precursor form of the D1 protein (Müller and Eichacker 1999).

Only a few assembly factors important for the biogenesis of the PSII core complex have been identified so far (Table 1). A molecular chaperone HCF136 in the thylakoid lumen was first characterized with a selective role in the assembly of only the PSII reaction center complex (Meurer et al. 1998; Plucken et al. 2002) (Table 1), yet the specific molecular interactions in assisting the assembly process still remain unknown. More recently, an LPA1 (low PSII accumulation1) protein was identified in *Arabidopsis* and shown to be an integral membrane chaperone essential for the maintenance and assembly of the PSII core complex, probably through a direct interaction with the PSII reaction center protein D1 (Peng et al. 2006). Yet another membrane-localized protein HCF107, a component of a multi-subunit complex, has been shown to be crucial for the assembly of PSII, affecting particularly the expression of the *psbH* gene in *Arabidopsis* (Sane et al. 2005).

HCF107, HCF136, and LPA1 probably represent only the first examples of the growing number of assembly factors, or molecular chaperones, facilitating the biogenesis of the PSII core complexes. Additional assembly factors are likely to be discovered in the near future by advanced proteomic and reverse genetics approaches. For example, some member(s) of the Alb3 family are likely to have such functions (see Section 3.2.)

2.1.2 Assembly of the proteins of the oxygen evolving complex

The OEC complex of higher plant PSII contains three extrinsic, nucleus-encoded subunits, PsbO (OEC33), PsbP (OEC23), and PsbQ (OEC16). The OEC complex is attached to the lumenal side of PSII and protects the $CaMn_4$ cluster bound to the D1 and CP43 proteins (Ferreira et al. 2004). Contrary to the membrane-embedded PSII core polypeptides that are subject to rapid proteolytic degradation when not assembled, a pool of free, assembly-competent OEC proteins has been shown to exist in the thylakoid lumen (Hashimoto et al. 1996, 1997).

One clear requirement for the assembly of OEC to the luminal side of PSII is the C-terminal processing of the precursor D1 protein. It has been demonstrated that the CP43 protein is stably assembled only if the D1 protein has undergone maturation via C-terminal processing (Zhang et al. 2000). Coordination of the C-terminal processing and the assembly of CP43 may thus be essential for stable ligation of the $CaMn_4$-cluster to the PSII core (Roose and Pakrasi 2004) and for the subsequent photoactivation of the oxygen evolving complex.

Assembly of the PsbO protein of OEC to the PSII core complex occurs in the stroma-exposed thylakoid membranes whereas the PsbP and PsbQ proteins have been found to associate with PSII in the grana thylakoids (Hashimoto et al. 1997). It was believed for a long time that the PsbO protein is the only OEC protein that directly binds to the PSII core on the lumenal side of the thylakoid membrane, and thereby provides a docking site for PsbP, which in turn binds the PsbQ protein (e.g. Miyao and Murata 1989). This model has, however, been recently challenged and evidence is accumulating supporting the concept that either all the three OEC proteins are independently bound to PSII, or only PsbO and PsbP are independently bound to PSII and one, or both, of them provide a docking site for PsbQ.

Table 1. Assembly factors, chaperones and translocator components involved in the assembly of thylakoid protein complexes.

Factor or chaperone (or nuclear locus)	function	reference
Assembly factors, chaperones		
HCF136	PSII assembly	Meurer et al. 1998; Plucken et al. 2002
HCF107	PSII assembly	Sane et al. 2005
LPA1	PSII assembly	Peng et al. 2006
TLP40	PSII assembly	Fulgosi et al. 1998
Alb3.1	PSII assembly	Ossenbühl et al. 2004
Alb3.2	PSII and PSI assembly	Gerdes et al. 2006; Göhre et al. 2006
Alb4	PSII assembly	Gerdes et al. 2006; Göhre et al. 2006
Slr1471p	PSII assembly (cyano)	Ossenbühl et al. 2006
Psb27	PSII/OEC assembly	Chen et al. 2006; Nowaczyk et al. 2006
Psb29	PSII/OEC assembly	Keren et al. 2005
Ycf3, Ycf4	PSI assembly	Boudreau et al. 1997; Naver et al. 2001
Ycf37	PSI assembly (cyano)	Wilde et al. 2001
BtpA	PSI assembly (cyano)	Bartsevich and Pakrasi 1997 Zak and Pakrasi 2000
PYG7	PSI assembly	Stöckel et al. 2006
HCF145	PSI assembly	Lezhneva and Meurer 2004
HCF101	PSI and FTR assembly	Stöckel and Oelmüller 2004
APO1	PSI and FTR assembly	Amann et al. 2004
RubA	PSI assembly (cyano)	Shen et al. 2002a, 2002b
HCF164	Cyt b_6 f assembly	Lennartz et al. 2001
HCF153	Cyt b_6 f assembly	Lennartz et al. 2006
CCDA	Cyt b_6 f assembly	Page et al. 2004
CCSA (Ycf5)	Cyt b_6 assembly	Xie and Merchant 1996; Hamel et al. 2003
CCS1-4	Cyt b_6 f assembly	Inoue et al. 1997; van Wijk 2001
CCB1-4	Cyt b_6 f assembly	Kuras et al. 1997; van Wijk 2001
HSP70	Cyt b_6 f assembly	Madueno et al. 1993
HSP70B	PSII stability and turnover; Rubisco assembly	Schroda et al. 2001; Yokthongwattana et al. 2001; Brutnell et al. 1999
BSD2	Rubisco assembly	Brutnell et al. 1999
Hsp100/ClpC1	PSI, PSII biogenesis	Sjögren et al. 2004
DnaJ	Rubisco assembly	Hartl 1996; Schlicher and Soll 1997
GrpE	Rubisco assembly	Hartl 1996; Schlicher and Soll 1997

Factor or chaperone (or nuclear locus)	**function**	**reference**
Cpn60	Rubisco and Cyt $b_6 f$ assembly	Gatenby and Ellis 1990; Madueno et al. 1993
Cpn21	Rubisco assembly	Gatenby and Ellis 1990; Madueno et al. 1993
cpSRP54	Lhcb-protein assembly	Tu et al. 1999; Woolhead et al. 2001
cpSRP43	Lhcb-protein assembly	Tu et al. 1999; Woolhead et al. 2001
Translocator components		
Hcf106	TAT-translocation	Settles et al. 1997; Mori et al. 2001
Tha4	TAT-translocation	Mori et al. 2001
cpTatC	TAT-translocation	Mori et al. 2001
cpSecY; SecE, SecA	Plastocyanin, PsbO	Schuenemann et al. 1999

cyano = cyanobacteria

The PsbO protein attaches to the lumenal loops of the D2 and CP47 core proteins (Nield et al. 2000) and also requires the presence of CP43 for the stable assembly (Suorsa et al. 2004). For PsbP association, it was recently shown using reverse genetics approaches that the presence of the LMM protein PsbJ is an absolute requirement (Hager et al. 2002; Suorsa et al. 2004). This requirement, however, may be only indirect and result from the fact that another PSII protein, PsbR, is also missing from the PsbJ mutant thylakoids (Suorsa et al. 2006). Indeed, the PsbR protein was shown to be important for the structure and function of the OEC complex. It was demonstrated that the absence of PsbR results in a reduction of the PsbP and PsbQ proteins as well as a reduction in the light-saturated rate of oxygen evolution (Suorsa et al. 2006; Allahverdiyeva et al. 2007). These results provide evidence that PsbR is an important component in the PSII core complex, especially for the stable assembly of the PsbP protein. The third OEC protein, PsbQ, was found to be completely missing from a tobacco mutant lacking the PsbP protein (Ifuku et al. 2005) suggesting that the PsbP protein provides a docking site for the PsbQ protein (for further discussion see Suorsa and Aro, 2007). Two other PSII proteins, encoded by a single nuclear gene *psbY* (Gau et al. 1998), are also important for water oxidation (Neufeld et al. 2004) and possibly play a similar role as PsbR by stabilizing the association of the OEC proteins to the PSII core dimer.

Recent proteomic studies have revealed the existence of novel proteins in substoichiometric amounts in various purified PSII preparations (Kashino et al. 2002). Of these proteins, Psb29 was shown important for the assembly of PSII (Keren et al. 2005) but Psb27 was particularly assigned a role in the assembly of the OEC proteins to the PSII core (Roose and Pakrasi 2004). Psb27 protein seems to bind to the PSII core monomer prior to the assembly of the OEC proteins (Nowaczyk et al. 2006). Studies with the Psb27 mutant also revealed an impaired repair of the PSII centers after photoinhibition, providing evidence that the Psb27 protein possibly facilitates the assembly of OEC to the PSII core (Chen et al. 2006).

2.1.3 Assembly of the PSII-LHCII supercomplexes

The functional PSII complexes of higher plants exist as PSII-LHCII (light-harvesting chlorophyll-protein complex II) supercomplexes in the grana appressions. Of the LMM proteins of PSII, particularly the PsbZ (and PsbW) protein has been reported to be essential for the stable assembly of the PSII-LHCII supercomplexes (Swiatek et al. 2001; Rokka et al. 2005). This chloroplast-encoded protein is located in the periphery of the PSII core dimer, in a close vicinity to CP43. Overlay of the X-ray structures of spinach LHCII and the cyanobacterial PSII core onto the projection map of the cryo-EM 3D structure of the isolated PSII-LHCII supercomplexes of spinach revealed a close vicinity of PsbZ to CP26 (Lhcb5) (Loll et al. 2005). It remains to be elucidated whether the nucleus-encoded PsbW protein is located in the similar vicinity to the CP47 and CP29 (Lhcb4) proteins. Upon formation of the PSII-LHCII supercomplex, the CP29 and CP26 proteins attach the LHCII trimers, consisting of the Lhcb1 and Lhcb2 proteins, to the core dimers (Boekema et al. 1999). Furthermore, CP24 (Lhcb6) together with CP29 and CP26 most probably bind additional trimers (composed of Lhcb1-3) in the periphery of the PSII-LHCII supercomplex. The LHCII trimers are bound to the PSII dimer either strongly (S), moderately (M), or loosely (L) (Dekker and Boekema 2005). Recently, it was shown that the CP24 (Lhcb6)-deficient plants displayed a major change in the macro-organization of the PSII-LHCII supercomplexes in the grana (Kovacs et al. 2006). It was concluded that CP24 provides the linker for association of the M-trimer into the PSII complex, thereby allowing a specific macro-organization necessary for optimal function of PSII.

It is intriguing to note that the OEC proteins possibly also have specific roles in the structural integrity of the PSII-LHCII supercomplexes and their macro-organization in the grana (Dekker and Boekema 2005). Electron microscopy and single particle analysis have revealed that the PSII-LHCII supercomplexes lacking the OEC proteins differ from the native PSII supercomplexes (Boekema et al. 2000). It was concluded that the OEC proteins are needed to keep the CP29 and S-LHCII trimers at a correct distance from the PSII core in order to optimize the migration of excitation energy to the PSII core.

The Alb3 protein has been assigned an important role in the membrane insertion and assembly of the Lhcb proteins (see below more about Alb3). So far, however, no specific assembly factors have been detected to be involved in the association of the light-harvesting apparatus to the PSII core dimer, i.e., in the formation of the PSII-LHCII supercomplexes and their macro-organizations in the grana.

2.1.4 Reassembly of the PSII complexes during the photoinhibition repair cycle

The PSII complex performs a unique task in splitting water molecules to oxygen and hydrogen (protons). Such oxidizing electron transfer reactions of PSII in an atmosphere containing oxygen readily result in the formation of highly reactive radicals that are potentially harmful to the proteins and induce imbalance during the linear electron transfer process. Situation like this may lead to photoinactivation and photodamage of PSII, when the PSII complex is unable to transfer electrons and split water molecules. A constant repair of the photodamaged PSII complexes is required for the maintenance of a sufficient level of active PSII complexes for photosynthesis. The efficiency of repair is dependent on the environmental conditions, stress factors such as high light or low temperature, impairing the efficiency of the repair process. As far as the repair process is in balance with the rate of photodamage, nonfunctional PSII complexes do not accumulate and a measurable decrease in the rate of total photosynthesis is not detected. An extensive literature has been published on the mechanisms of the PSII photoinactivation and damage to the D1 protein (for a review see Melis 1999; Prasil et al. 1992; Aro et al. 1993; Chow and Aro 2005) and therefore these subjects are not considered here in more detail.

At the protein level, most often only the D1 protein is the target for the light-induced damage, but occasionally also the D2 and PsbH proteins become damaged and require replacement during the repair cycle (Schuster et al.1988; Bergantino et al. 2003; Rokka et al. 2005). However, only the repair steps concerning the replacement of the D1 protein are considered here. It is also worth noting that the CES process seems to play no role in the recovery from photoinhibition (Minai et al. 2006).

In the beginning of the repair, the LHCII antenna dissociates from the dimer and monomerization of PSII occurs. Damaged PSII monomers then migrate from the grana to the stroma-exposed membranes, where a contact with the components required in degradation and synthesis of the D1 protein are available. OEC dissociates from PSII and a partial disassembly of the PSII core proteins takes place. The stages from photodamage to degradation of the D1 protein are regulated by phosphorylation-dephosphorylation events of the core proteins (Koivuniemi et al. 1995; Rintamäki et al. 1996) (see below). The D1 protein is degraded proteolytically, proteases from the DegP and FtsH families known to act on the process (for a review see, for example, Adam et al. 2005; Sakamoto 2006). The closest partner of the D1 protein, D2, remains most often intact in the repair process. In the synthesis of the new D1 protein, the nascent D1 protein is co-translationally inserted into the thylakoid membrane where the D2 and Cyt b_{559} act as the first assembly partners. In fact, it was demonstrated that not only the insertion into the membrane but also the assembly of the D1 protein into the PSII complex, composed of Cyt b_{559}, D2 and possibly also of CP47 and several LMM subunits, occur co-translationally during the repair process (Zhang et al. 1999, 2000; Rokka et al. 2005). Re-synthesis of the assembly partner subunits is not needed, since they are already present in the existing PSII centers under repair.

After maturation of the D1 protein, the reassembly of the internal core antenna protein CP43 occurs. CP43, residing next to the D1 protein, is always dissociated from PSII upon the repair process. Before the OEC proteins can re-associate, also most of the LMM subunits have to be assembled to the PSII complex. Finally, the properly assembled, repaired PSII monomer migrates back to the grana thylakoids, where PSII core dimerization and reactivation, with the association of the LHCII antenna proteins, take place. These last assembly steps thus accomplish the PSII photoinhibition repair cycle providing active PSII-LHCII supercomplexes for photosynthesis.

2.2 Assembly of the PSI complex

The assembly of the PSI complex is rather poorly known due to a difficulty in isolation of various PSI subcomplexes and also to a very slow turnover rate of the PSI complexes, which results in a technical difficulty to accumulate radiolabeled amino acids into newly synthesized PSI subunits. In higher plants, the PSI core complex is composed of 14 subunits (PsaA to PsaL, PsaN, and PsaO), of which PsaA, PsaB, PsaC, PsaI, and PsaJ are plastid-encoded (Jensen et al. 2003; Ben-Shem et al. 2003). A novel subunit of PSI, the previously found phosphoprotein TMP14 (Hansson and Vener 2003) was recently identified in *Arabidopsis* (Khrouchtchova et al. 2005). This protein, designated as PSI-P, was suggested to locate in the proximity of PsaL, PsaH, and PsaO subunits, on the opposite side to the location of the LHCI antenna. Furthermore, the PSI-G subunit has been found to be bound to PsaB and to be in contact with Lhca1 (Zygadlo et al. 2006). The PSI peripheral antenna is arranged around one side of the PSI core and is composed of four different nuclear-encoded Lhca polypeptides (Lhca1-4) in higher plants. In addition, the fifth Lhca protein, which shows a different mode of regulation as compared to the other Lhca proteins, and which is present at sub-stoichiometric amounts under standard conditions, has recently been characterized (Ganeteg et al. 2004).

A key step in the assembly of the PSI complex is the coordinate synthesis and assembly of its two chloroplast-encoded core polypeptides, PsaB and PsaA, that form, together with ca 100 Chl *a* molecules and several redox ligands, the main part of the reaction center complex (for a review see Rochaix 2006). In *Chlamydomonas*, the accumulation of PsaB was shown to be required for synthesis of the PsaA subunit that, in turn, is needed for synthesis of the PsaC subunit (Wostrikoff et al. 2004) on the stromal side of the membrane. All these three subunits, PsaA, PsaB, and PsaC, are required for stable accumulation of the PSI core complex. The rate of production of PsaB is the controlling stage in order to determine the stoichiometric expression of all subunits of the PSI core complex. There is thus a clear CES hierarchy in the sequence of polypeptide assembly during PSI biogenesis (Wostrikoff et al. 2004). Whether the other chloroplast-encoded PSI subunits PsaI and PsaJ are also CES proteins remains unknown. PsaC then coordinates the stable assembly of PsaD and PsaE, both on the stromal side of PSI (Yu et al. 1995).

Some assembly factors have been assigned a role particularly in the biogenesis of the PSI complexes (Table 1). These include the plastid-encoded Ycf3 and Ycf4 factors (Boudreau et al. 1997). Ycf3 has been found to interact directly with PsaA and PsaD, but not with the subunits of other photosynthetic complexes (Naver et al. 2001). When Ycf3 and Ycf4 were missing in the deletion mutans of *Chlamydomonas*, no stable assembly of PSI occurred, even though the PsaA, PsaB, and PsaC transcripts accumulated (Boudreau et al. 1997). In cyanobacteria, the lack of the Ycf37 protein caused a decrease in photosynthetic activity and lowered levels of the PSI complexes, yet the mutant cells were capable of photoautorophic growth (Wilde et al. 2001). Recently, the role of a higher plant homolog for Ycf37, PYG7 was characterized in *Arabidopsis* (Stöckel et al. 2006). The plants lacking PYG7 were unable for photoautorophic growth and did not accumulate PSI complexes. However, the PSI subunits were synthesized in the mutants, indicating that the lack of the PSI complexes is due to accelerated degradation of the unassembled subunits (Stöckel et al. 2006). The lack of the HCF145 protein, on the other hand, caused dramatically decreased amounts of the PSI subunits as well, but the protein was shown to function at the mRNA level, by stabilizing the *psaA-psaB-rps14* operon (Lezhneva and Meurer 2004). In cyanobacteria, the BtpA protein has been shown to posttranscriptionally affect the accumulation of PSI (Bartsevich and Pakrasi 1997), especially under low temperature (Zak and Pakrasi 2000).

The correct assembly of the iron-sulphur clusters has been found to be essential for the accumulation of the PSI and Cyt b_6f complexes, and some proteins needed for the (general) assembly of Fe-S clusters have already been identified (Touraine et al. 2004; Yabe et al. 2004). PSI has three iron sulphur centers of type [4Fe-4S], one of which (F_x) is associated with the PsaA/B heterodimer and the two others (F_A and F_B) with PsaC. The evolutionarily conserved HCF101 protein, found to be essential for the accumulation of PSI (Stöckel and Oelmüller 2004), has been shown to function particularly in the assembly of the [4Fe-4S] clusters (Lezhneva et al. 2004). Also the APO1 protein, which is specific for vascular plants, is needed for accumulation of PSI via assembly of the [4Fe-4S] clusters (Amann et al. 2004). APO1-mediated function, however, occurs at a different stage or through a different mechanism than that os HCF101, since the phenotypes, some functional characteristics, chloroplast ultrastructure and the levels of the PSI antenna proteins differ between the *apo1* and *hcf101* mutants (Lezhneva et al. 2004; Amann et al. 2004). Nevertheless, the role of both HCF101 and APO1 in the assembly of PSI is specific for the [4Fe-4S] clusters, since both the *hcf101* and *apo1* mutants also exhibited lowered levels of the ferrodoxin-thioredoxin reductase containing [4Fe-4S] clusters (Amann et al. 2004; Lezhneva et al. 2004). Moreover, the *apo1* mutant also had reduced amounts of the NAD(P)H dehydrogenase (NDH) complexes, which likewise harbor[4Fe-4S] clusters (Amann et al. 2004). The specificity of HCF101 and APO1 for [4Fe-4S] clusters is corroborated by the fact that ferredoxin, which contains a [2Fe-2S] cluster, was present at normal levels in both the *hcf101* and *apo1* mutants (Amann et al. 2004; Lezhneva et al. 2004). In cyanobacteria, a rubredoxin protein RubA has been shown to be needed for the assembly of the F_x [4Fe-4S] cluster (Shen et al. 2002a), and the *rubA* inactivation mutant had significantly lower amounts of PSI, and was not capable of photoautotrophic growth (Shen et al. 2002b).

2.3 Assembly of the Cyt b_6f complex

The Cyt b_6f complex is a dimer, with one monomer composed of eight subunits, from which six subunits are plastid-encoded (PetA, Pet B, Pet D, PetG, PetL, and PetN) and two nuclear-encoded (PetC and PetM). The three-dimensional structure of the Cyt b_6f complex was resolved recently both from cyanobacteria (Kurisu et al. 2003) and *Chlamydomonas* (Stroebel et al. 2003). The Cyt b_6f complex is also the thylakoid protein complex, in which the CES control of the synthesis of the chloroplast-encoded proteins was first demonstrated (Choquet et al. 1998), yet the precise molecular mechanisms of the CES processes in chloroplasts remain to be elucidated.

Cytochrome f (Cyt *f*, PetA) is a CES protein because its rate of synthesis is regulated by the availability of its assembly partners, which are the chloroplast-encoded cytochrome b_6 (PetB) and the subunit IV (SU IV, PetD). In the absence of these assembly partners (or dominant subunits, Cyt b_6 and SU IV), the synthesis of Cyt *f* decreases tenfold (Kuras and Wollman 1994). The C-terminal region of Cyt *f* is important for the assembly into the complex (Mould et al. 2001). More recently, it was shown that Cyt *f* translation is autoregulated by its C-terminal domain and that this CES process for Cyt *f* expression most likely requires an interaction with the membrane-bound translational activator (Choquet et al. 2003).

One of the major Cyt b_6f subunits, the nucleus-encoded Rieske iron-sulphur protein (PetC) is synthesized in cytosol as a 26 kDa precursor and subsequently transported to the plastid. It is processed in the stroma to the mature 20 kDa protein, found to be associated with the chaperones Cpn60 and Hsp70 in the stroma and targeted to the thylakoid membrane where it is assembled into the Cyt b_6f complex (Madueno et al. 1993). For the assembly, it has been found that the presence of the Rieske [2Fe-2S] cluster, the glycine-rich region or the conserved C-terminal region is not required as a prerequisite (Kapazoglou et al. 2000). Interestingly, the Rieske protein has also been assigned a role in the assembly-mediated control of the Cyt *f* synthesis, though the effect was lower than that observed in the absence of Cyt*b* and SU IV (de Vitry et al. 2004).

The function of the small subunits PetG, PetL (ycf7), PetM, and PetN (ycf6) of the Cyt b_6f complex is not yet known properly. However, it has been demonstrated in cyanobacteria that inactivation of the *petM* gene did not affect the activity of the Cyt b_6f complex itself, but instead affected the stoichiometry of other protein complexes, suggesting that specific regulatory processes are mediated by the Cyt b_6f complex (Schneider et al. 2001). In the tobacco knockout mutant for the *petN* gene, on the contrary, the Cyt b_6f complex was totally absent, resulting in interruption in the electron transfer from PSII to PSI, these two latter complexes being, however, intact and physiologically active (Hager et al. 1999).

At least two auxiliary proteins, HCF164 (Lennartz et al. 2001) and HCF153 (Lennartz et al. 2006) have been identified that specifically regulate the accumulation of the Cyt b_6f complexes in the thylakoid membrane (Table 1). Both proteins have been found to be tightly associated with the thylakoid membrane. HCF164 is a thioredoxin-like protein and was recently shown to be able to mediate reducing equivalents across the thylakoid membrane (Motohashi and Hisabori 2006).

Among the identified target proteins for HCF164 were Cyt *f* and the Rieske protein, indicating that the interaction between HCF164, Cyt *f* and the Rieske protein might be an important prerequisite for the assembly of the Cyt b_6f complex (Motohashi and Hisabori 2006). Moreover, the CCDA protein, which is a homolog for prokaryotic thiol disulfide transporter, might be a component of the HCF164-dependent transthylakoid thioreduction pathway, and the lack of the CCDA protein caused defects in the accumulation of Cyt b_6f, and resulted in impaired photosynthesis (Page et al. 2004). The plastid-encoded CCSA protein (Xie and Merchant 1996) and the nuclear-encoded CCS1-4 proteins are needed for the *c*-heme attachment (Inoue et al. 1997; Hamel et al. 2003). In addition, the nuclear-encoded CCB1-4 proteins are specific for binding heme to Cyt b_6 (Kuras et al. 1997). For a review of the CCS and CCB proteins, see van Wijk (2001).

2.4 Assembly of soluble complexes

Increasing amount of research has recently been focused on the assembly of the thylakoid-membrane-embedded protein complexes (with the NDH complex as an exception) whereas the knowledge concerning the assembly of the chloroplast soluble complexes has not much advanced during the past few years. Here we briefly summarize the assembly processes of two stromal protein complexes, Rubisco and FTR.

2.4.1 Rubisco

In higher plants and green algae, Rubisco holoenzyme exists as a 600 kDa soluble complex of the L8S8 form. It thus consists of eight large subunits (LSU) of 55 kDa encoded by the plastome *rbcL* gene and eight small subunits (SSU) of 15-18 kDa encoded by the *rbcS* gene in the nucleus (Spreitzer 1993). Also in red algae, Rubisco is of the L8S8 form, but both subunits are plastome-encoded.

During the assembly of LSU chains, the DnaK/DnaJ/GrpE chaperone complex has been found to associate to the chains in order to maintain them in an unfolded state (Hartl 1996). Also the BSD2 protein, having homology with the DnaJ proteins, has been suggested to prevent the aggregation of the nascent LSU chains (Brutnell et al. 1999). The SSU precursors are processed during their entry into the plastid and are subsequently assembled. The Cpn60 and Cpn21 chaperonins assist in the assembly of the L8S8 holoenzyme (reviewed in Gatenby and Ellis 1990; Gutteridge and Gatenby 1995). The SSU assembly stabilizes the holoenzyme complex generating a fully active enzyme complex. In particular, the highly conserved tyrosine residues at the beta A-beta B loop of the SSU were recently identified to play a stabilizing role for the holoenzyme (Esquivel et al. 2006). SSU assembly controls LSU expression, but SSU does not have a direct effect on LSU translation. If the SSU expression is inhibited (antisense silencing in tobacco), Rubisco assembly is prevented and LSU synthesis is reduced (Rodermel et al. 1996). The assembly of Rubisco has been shown to be sensitive to oxidative stress, and it was recently proposed that during oxidative stress, the RNA recognition motif in the N-terminus of the LSU becomes exposed and binds any RNA molecule, which causes blocking of the translation and degradation of the unpaired SSU (Cohen et al. 2005, 2006). Thus, in the absence of one subunit in the complex, synthesis of another subunit decreases that has also been detected in the assembly of other photosynthetic complexes in chloroplasts (Minai et al. 2006).

2.4.2 Ferredoxin:thioredoxin reductase

The stromal FTR is a heterodimer protein of 26 kDa, consisting of the catalytical β subunit with a [4Fe-4S] cluster and a variable α subunit. The primary structure of the catalytical subunit is highly conserved between different species, whereas the variable subunit of higher plants has a N-terminal tail. The catalytical β subunit stabilizes the α subunit, since the [4Fe-4S] cluster has been shown to be important for the stability of FTR (Manieri et al. 2003). Thus, the nuclear-encoded proteins HCF101 (Lezhneva et al. 2004) and APO1 (Amann et al. 2004), essential for the assembly of the PSI [4Fe-4S] clusters, have been shown to be needed for the accumulation of FTR subunits as well.

3 Insertion of proteins to the thylakoid membrane - thylakoid translocase complexes and chaperones

3.1 Thylakoid translocases

Nucleus-encoded thylakoid proteins, first translocated to the chloroplast stroma via the envelope membrane, are generally dependent on thylakoid protein complexes, the translocases, to find their final location. They can be inserted into the thylakoid membrane or translocated to the lumen by three distinct pathways that have bacterial homologues: the SRP (signal recognition particle), the Tat (twin-arginine translocase) and the Sec (secretory) pathways. In addition, a fourth pathway exists that is considered to be ‘spontaneous’. The protein composition of these translocases has been partially resolved, but very little is known about the assembly processes of the translocases themselves.

The SRP and Sec pathways translocate proteins in their unfolded state and require the activity of soluble chaperones (Mori and Cline 2001), while the Tat pathway has the rare ability to translocate proteins in their fully folded state (Clark and Theg 1997). Proteins using the SRP pathway have a single pre-sequence, which is cleaved off after the envelope translocation, while proteins using the Tat and Sec routes have bipartite pre-sequences for translocation of proteins to the thylakoid lumen. There are also differences in the energetic requirements of protein translocation between the three routes: the Sec and SRP pathways require hydrolysis of nucleoside triphosphates, ATP and GTP, respectively, even though a proton motive force may also be involved (Mant et al. 1995; Kouranov and Schnell 1996).

The chloroplast SRP is a trimer consisting of two subunits of cpSRP43 and one cpSRP54 subunit (Li et al. 1995; Tu et al. 1999). The specific substrates for SRP pathway are the Lhcb proteins, especially the Lhcb4.1 and Lhcb5 proteins have been investigated in detail (Cline 1986; Woolhead et al. 2001). The integration of an Lhcb protein into the thylakoid membrane occurs in two steps: the Lhcb protein interacts first with cpSRP to form a soluble targeting intermediate, called the transit complex, and subsequently integrates into the thylakoid membrane in the presense of GTP and FtsY (Tu et al. 1999). Furthermore, insertion of the Lhcb protein into the thylakoid membrane is known to require an additional component, Alb3 (see also below), a protein that belongs to the Oxa1-YidC family (Moore et al. 2000; Woolhead et al. 2001).

The Tat-pathway is the major route for protein export in prokaryotes, also participating in translocation of proteins to plastids (Finazzi et al. 2003). A substrate protein for the Tat-pathway contains a characteristic, conserved twin-arginine motif situated upstream of a hydrophobic stretch in the pre-sequence. The complete structure of the Tat-translocation channel is not resolved yet, but three proteins, Hcf106, Tha4, and cpTatC, have been identified as the primary components of the Tat-pathway (Settles et al. 1997; Mori et al. 2001). Such proteins as PsaN, PsbP, and PsbQ have been reported to use the Tat-pathway in their translocation (Nielsen et al. 1994; Clark and Theg 1997). The Tat-translocation has also been found to be dependent on the ΔpH across the thylakoid membrane, but this has recently been questioned by showing that the transport of the Tat-pathway substrates can take place *in vivo* in the absence of ΔpH (Finazzi et al. 2003).

The Sec-pathway translocates proteins such as plastocyanin and PsbO across the membrane to the thylakoid lumen. Components of the Sec-pathway include the membrane-bound SecY and SecE proteins, as well as the soluble stromal protein SecA (Shuenemann et al. 1999). By analogy to the bacterial Sec-pathway, it is assumed that SecA interacts with a precursor protein in the stroma and subsequently inserts itself into the membrane. SecY and SecE, in turn, form the translocation channel, maybe with some so far unidentified protein(s).

Many thylakoid proteins insert spontaneously to the membrane, without any aid of stromal components, nucleoside triphosphates, SRP, Alb3, or SecA. These include the photosynthetic reaction center proteins PsbW, PsbY, and PsaK, as well as SecE (Mant et al. 2001; Steiner et al. 2002).

The insertion mechanisms of the chloroplast-encoded proteins to the thylakoid membrane have not been thoroughly investigated. However, there is emerging evidence that the chloroplast-encoded proteins, usually synthesized on thylakoid-bound ribosomes, also use the thylakoid translocases, like SecY (Zhang et al. 2000). Alb3 interactions with the PSI and PSII reaction center proteins (Göhre et al. 2006) also propose the role of Alb3 protein in the folding and translocation of chloroplast-encoded proteins.

3.2 Chaperones

Besides the assembly factors discussed above in the context of the assembly of specific thylakoid protein complexes, several other assembly factors or molecular chaperones have been identified in chloroplasts. These chaperones include chloroplast-envelope-associated and stromal members of the Hsp70 family (for review see Jackson-Constan et al. 2001; van Wijk 2001; Schroda 2004). In addition to the general role of Hsp70 in refolding denatured proteins, some specialized functions have also been found for this chaperone. In *Chlamydomonas* it was shown that HSP70B may protect PSII under light stress and/or stabilize photodamaged PSII to allow for a coordinated repair (Schroda et al. 2001). Furthermore, in *Dunaliella salina* it was detected that a PSII repair intermediate indeed contained the HSP70B protein (Yokthongwattana et al. 2001). Moreover, folding of Rubisco by the stromal Hsp70 was shown to be assisted by the BSD2 protein, which has a high sequence similarity to the Zn-finger domain of DnaJ proteins (Brutnell et al. 1999). DnaJ (and also GrpE) proteins function as co-chaperonins in the prokaryotic Hsp70 system (Schlicher and Soll 1997). In addition, the members of the Hsp100/Clp chaperone family participate in specific functions in chloroplasts. In *Arabidopsis clpC1* mutant line lacking approximately 65% of the total Hsp100/ClpC protein, growth retardation, impaired photosynthetic capacity and reduced amounts of PSI and PSII were found, indicating that ClpC1 is essential for the normal function of the photosynthetic machinery (Sjögren et al. 2004) (For a review concerning the recent advances in the study of the Clp proteins, see Adam et al. 2006).

Also the thylakoid lumen contains a separate set of molecular chaperones, such as cpn60, cpn10, and hsc70 proteins (Schlicher and Soll 1996). Another lumenal protein TLP40 is a cyclophilin-type PPIase that is assumed to catalyze the folding of proteins newly inserted in the thylakoid membrane, or translocated into the thylakoid lumen (Fulgosi et al. 1998). This protein also functions as a phosphatase inhibitor (Vener et al. 1999). Recent characterization of the TLP40 knockout mutants has revealed that the TLP40 protein is crucial in the growth and development of *Arabidopsis* plants thus indicating its crucial importance for the biogenesis and assembly of the thylakoid protein complexes (Khrouchtchova et al. manuscript in preparation).

The Alb3 protein located in the thylakoid membrane is a member of the YidC/Oxa1/Alb3 membrane protein family, whose members are multifunctional mediators of membrane protein integration, folding and assembly into larger complexes. Their evolutionary conserved and physiologically important roles are generally linked to the assembly of the major energy-transducing membrane protein complexes (van der Laan et al. 2005). In chloroplasts, Alb3 (Alb3.1) is an important component of the thylakoid SRP pathway import complex, which is, however, not the only function of chloroplast Alb proteins in the insertion of proteins to the thylakoid membrane. Indeed, Alb3 is involved in the membrane insertion and assembly of both the nucleus- and plastid-encoded subunits of various photosynthetic membrane protein complexes (Ossenbühl et al. 2004). In *Arabidopsis*, loss of Alb3 results in an albino phenotype and a reduction in the amount of thylakoid membranes (Sundberg et al. 1997). Although the major function of Alb3 (Alb3.1) seems to be to assist the integration and assembly of the Lhcb proteins, other members of the Alb family, Alb3.2 and Alb4, have recently been reported to also participate in the assembly of thylakoid proteins (Göhre et al. 2006; Gerdes et al. 2006). Alb3.2 was found in a large thylakoid protein complex and showed interaction with Alb3.1 and the reaction center proteins of PSI and PSII (Göhre et al. 2006). Moreover, downregulation of Alb3.1 resulted in concomitant decrease in the number of PSII and PSI reaction centers suggesting a fundamental role of Alb3.2 in the assembly of these complexes. More support for the involvement of Alb proteins in PSII biogenesis and turnover come from experiments with cyanobacterial cells where an Alb3 homolog Slr1471p was shown to directly interact with the precursor-D1 protein and facilitate the proper repair of the PSII centers (Ossenbühl et al. 2006).

4 Posttranslational modifications of chloroplast proteins

Chloroplast proteins are prone to several modifications, which occur either after nucleus-encoded proteins have been imported into chloroplasts, or upon or after protein translation in chloroplasts. The most important irreversible modifications are the N-terminal deformylation, removal of N-terminal methionine, and internal processing, whereas protein phosphorylation represents the most common reversible posttranslational modification of chloroplast proteins. Other modifications include the reversible addition and removal of functional groups by glycosylation, acylation, and nitration resulting in structural changes in proteins. Posttranslational modifications of proteins are important regulators that enhance and increase protein complexity and dynamics. They are covalent processes that change the primary structure of proteins in a sequence-specific manner. In the following, we shortly summarize the recent advances in the fields concerning N-terminal methionine excision and thylakoid protein phosphorylation in plastids. In addition, the reader is referred to the recent reviews on studies of posttranslational modifications in plants (Peck 2006; Kwon et al. 2006; Rossignol 2006; de la Fuente van Bentem et al. 2006). For imported proteins, the cleavage of the transit peptide occurs in one or two phases, depending on the final destination of the protein in chloroplast (Mori and Cline 2001) as discussed above (Section 3.1.).

4.1 N-terminal methionine excision

Although Met is the first amino acid of the newly synthesized proteins, it is usually removed from mature proteins in a process called N-terminal Met excision (NME). NME is an irreversible co-translational mechanism, completed before the nascent polypeptide chains are fully synthesized (Arfin and Bradshaw 1988). NME is best documented in plastids where the N-termini of most of the proteins encoded by the chloroplast genome have been determined (Giglione et al. 2004).

Two enzymes of sequential action are needed for NME: 1) peptidyl deformylase (PDF), which specifically removes the N-formyl group present in all nascent polypeptides synthesized in eubacteria and organelles and 2) methionine aminopeptidase (MAP), which removes the methionine specifically in all organisms (Giglione et al. 2004).

Whether the N-formyl group only, or the entire N-formylMet group, is cleaved or retained, depends mostly on the nature and bulkiness of the side chains of the second amino acid (Frottin et al. 2006). In the proteome of chloroplast-encoded proteins, however, all different possibilities exist. The excision of the N-formylMet is the most common one, this group including, among others, the reaction center proteins D1 and D2 of PSII. Additionally, a more extensive cleavage than only the N-formylMet occurs in some chloroplast proteins including RbcL, AtpI, PetA, PscC, and PsbK (Giglione et al. 2004).

In attempts to find the physiological role for NME in chloroplasts, Meinnel and colleagues (Giglione et al. 2003) tested the hypothesis whether MNE is determining the protein half-life. To this end, a specific inhibitor of PDF, actinonin, was used and found to cause a progressive loss of photosynthetic activity both in *Arabidopsis* and *Chlamydomonas* due to the destabilization of the PSII core proteins, particularly the D2 protein. Since the function of PDF is a prerequisite for MAP function, it is likely that methionine at the N-terminus of some proteins, like the D2 protein, possibly acts as a destabilizing residue. Thus, it was concluded that NME is essential for biogenesis of PSII primarily by stabilizing the D2 subunit. This conclusion is corroborated by the fact that the disruption of *PDF1B* (a gene encoding the chloroplast targeted PDF) in *Arabidopsis* led to an albino phenotype (Giglione et al. 2003). However, several proteins of various thylakoid complexes are substrates of PDF, yet the stability of only PSII and its D2 protein were primarily affected in the presence of actinonin. Therefore, the detailed mechanisms of NME in regulation of the life span of chloroplast proteins and thereby the assembly of the chloroplast protein complexes remains to be established.

4.2 Protein phosphorylation

A dynamic light- and redox-controlled protein phosphorylation system has evolved in the thylakoid membranes of chloroplasts for regulation of photosynthesis and the dynamics of the photosynthetic protein complexes (Bennett 1977, 1991; Allen 1992; Vener et al. 1998, 2007). The reversible phosphorylation concerns given amino acid residues, most commonly the tyrosine residue on the stromal side of the thylakoid membrane.

A number of PSII proteins are reversibly phosphorylated in the thylakoid membrane. Thylakoid-bound kinases are responsible for protein phosphorylation, for which several regulatory patterns have been described (Pursiheimo et al. 2003). Protein dephosphorylation, in turn, is catalyzed by the chloroplast phosphatases, being either thylakoid-bound or soluble ones (Bennett 1991). Furthermore, modulation of the thylakoid protein phosphorylation involves the thiol redox state (Rintamäki et al. 2000) and the light-induced conformational changes in the substrate proteins (Zer et al. 1999; Jeschke et al. 2005). Thylakoid phosphoproteins include the D1, D2, CP43, and PsbH proteins of the PSII core (Bennett 1991; Vener et al. 2001; Andreuzzi et al. 2005), the Lhcb1, Lhcb2, and Lhcb4 proteins of the light-harvesting II antenna (Bennett 1991; Bergantino et al. 1995; Vener et al. 2001; Turkina et al. 2004; Tikkanen et al. 2006) as well as the PsaD protein of PSI (Hansson and Vener 2003), 9 kDa soluble phosphoprotein (TSP9) (Carlberg et al. 2003) and TMP14, the latter demonstrated recently to be a novel subunit of PSI (Khrouchtchova et al. 2005). In addition, two phosphorylation sites (Thr-2 and Ser-3) were detected recently in the Rieske Fe-S protein (PetC) of the Cyt $b_6 f$ complex in spinach, and three new threonine phosphorylation sites in the CP43 protein (Rinalducci et al. 2005).

The role of reversible phosophorylation of the above-mentioned photosynthetic proteins is not completely understood, but it has been shown to be involved in several aspects of the dynamics of photosynthetic membrane protein complexes, especially as a response to environmental cues. Light induces reversible phosphorylation of a number of PSII core proteins and of the LHCII antenna proteins Lhcb1, Lhcb2, and Lhcb 4 (Bennett 1991) via activation of the redox-dependent protein kinases, the identity of which is not yet fully elucidated.

5 Concluding remarks

Elucidation of the mechanisms, pathways, and auxiliary components involved in the synthesis, assembly, stability, and dynamics of the photosynthetic membrane protein complexes is still in its infancy. One pertinent task is to increase our understanding about the protein networks involved in auxiliary functions in guiding the assembly of the individual protein subunits to macromolecular photosynthetic complexes. Moreover, the biosynthesis and regulation of the ligation of various redox co-factors to the bioenergetic membrane protein complexes awaits extensive investigation. Table 1 summarizes our present knowledge of the assembly factors and chaperones involved in the biosynthesis of plastid protein complexes. We are now in an urgent need to get a systems biology view on the biogenesis of the photosynthetic energy providing pigment protein complexes. This will greatly facilitate, for example, the future plans to construct artificial cell factories for clean solar energy production.

Acknowledgements

The research in the author's laboratory has been supported by the Academy of Finland and Finish Ministry of Agriculture. Mr. Kurt Ståhle is thanked for his help in preparing the figures.

References

Adam Z, Rudella A, van Wijk KJ (2006) Recent advances in the study of Clp, FtsH and other proteases located in chloroplasts. Curr Opin Plant Biol 9:234-240

Adam Z, Zaltsman A, Sinvany-Villalobo G, Sakamoto W (2005) FtsH proteases in chloroplasts and cyanobacteria. Physiol Plant 123:386-390

Allahverdiyeva Y, Mamedov F, Suorsa M, Styring S, Vass I, Aro E-M (2007) Insights into the function of PsbR protein in *Arabidopsis thaliana*. Biochem Biophys Acta: in press

Allen JF (1992) Protein phosphorylation in regulation of photosynthesis. Biochim Biophys Acta 1098:275-335

Amann K, Lezhneva L, Wanner G, Herrmann RG, Meurer J (2004) ACCUMULATION OF PHOTOSYSTEM ONE1, a member of a novel gene family, is required for accumulation of [4Fe-4S] cluster-containing chloroplast complexes and antenna proteins. Plant Cell 16:3084-3097

Anbudurai PR, Mor TS, Ohad I, Shestakov SV, Pakrasi HB (1994) The *ctpA* gene encodes the C-terminal processing protease for the D1 protein of the Photosystem II reaction center complex. Proc Natl Acad Sci USA 91:8082-8086

Andreuzzi F, Barbato R, Picollo C, Segalla A (2005) Isolation of phosphorylated and dephosphorylated forms of the CP43 internal antenna of photosystem II in *Hordeum vulgare* L. J Exp Bot 56:1239-1244

Arfin SM, Bradshaw RA (1988) Cotranslational processing and protein turnover in eukaryotic cells. Biochemistry 27:7979-7984

Aro EM, Virgin I, Andersson B (1993) Photoinhibition of photosystem II: inactivation, protein damage and turnover. Biochim Biophys Acta 1143:113-134

Bartsevich VV, Pakrasi HB (1997) Molecular identification of a novel protein that regulates biogenesis of photosystem I, a membrane protein complex. J Biol Chem 272:6382-6387

Bennett J (1977) Phosphorylation of chloroplast membrane polypeptides. Nature 269:344-346

Bennett J (1991) Protein phosphorylation in green plant chloroplasts. Annu Rev Plant Physiol Plant Mol Biol 42:281-311

Ben-Shem A, Frolow F, Nelson N (2003) Crystal structure of plant photosystem I. Nature 426:630-635

Bergantino E, Dainese P, Cerovic Z, Sechi S, Bassi R (1995) A post-translational modification of the photosystem II subunit CP29 protects maize from cold stress. J Biol Chem 270:8474-8481

Bergantino E, Brunetta A, Touloupakis E, Segalla A, Szabò I, Giacometti GM (2003) Role of the PSII-H subunit in photo-protection. Novel aspects of D1 turnover in *Synechocystis* 6803. J Biol Chem 278:41820-41829

Boekema EJ, van Roon H, van Breemen JF, Dekker JP (1999) Supramolecular organization of photosystem II and its light-harvesting antenna in partially solubilized photosystem II membranes. Eur J Biochem 266:444-452

Boekema EJ, van Breemen JFL, van Roon H, Dekker JP (2000) Conformational changes in photosystem II supercomplexes upon removal of extrinsic subunits. Biochemistry 39:12907-12915

Boudreau E, Takahashi Y, Lemieux C, Turmel M, Rochaix JD (1997) The chloroplast *ycf3* and *ycf4* open reading frames of *Chlamydomonas reinhardtii* are required for the accumulation of the photosystem I complex. EMBO J 16:6095-6104

Bowyer JR, Packer JCL, McCormack BA, Whitelegge JP, Bobinson C, Taylor M (1992) Carboxyl-terminal processing of the D1 protein and photoactivation of water splitting in Photosystem II. J Biol Chem 267:5424-5433

Brutnell TP, Sawers RJH, Mant A, Langdale JA (1999) BUNDLE SHEATH DEFECTIVE2, a novel protein required for post-translational regulation of the *rbcL* gene of maize. Plant Cell 11:849-864

Carlberg I, Hansson M, Kieselbach T, Schröder WP, Andersson B, Vener AV (2003) A novel plant protein undergoing light-induced phosphorylation and release from the photosynthetic thylakoid membranes. Proc Natl Acad Sci USA 100:757-762

Chen H, Zhang D, Guo J, Wu H, Jin M, Lu Q, Lu C, Zhang LX (2006) A Psb27 homologue in *Arabidopsis thaliana* is required for efficient repair of photodamaged photosystem II. Plant Mol Biol 61:567-575

Choquet Y, Vallon O (2000) Synthesis, assembly and degradation of thylakoid membrane proteins. Biochimie 82:615-634

Choquet Y, Stern DB, Wostrikoff K, Kuras R, Girard-Bascou J, Wollman FA (1998) Translation of cytochrome *f* is autoregulated through the 5' untranslated region of *petA* mRNA in *Chlamydomonas* chloroplasts. Proc Natl Acad Sci USA 95:4380-4385

Choquet Y, Zito F, Wostrikoff K, Wollman FA (2003) Cytochrome *f* translation in *Chlamydomonas* chloroplasts is autoregulated by its carboxyl-terminal domain. Plant Cell 15:1443-1454

Chow WS, Aro EM (2005) Photoinactivation and mechanisms of recovery. In: Wydrzynski TJ and Satoh K (eds) Photosystem II: The Light-Driven Water: Plastoquinone Oxidoreductase. Advances in Photosynthesis and Respiration, vol 22. Dordrecht: Springer, pp 627-648

Clark SA, Theg SM (1997) A folded protein can be transported across the chloroplast envelope and thylakoid membranes. Mol Biol Cell 8:923-934

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

Cohen I, Knopf JA, Irihimovitch V, Shapira M (2005) A proposed mechanism for the inhibitory effect of oxidative stress on Rubisco assembly and its subunit expression. Plant Physiol 137:738-746

Cohen I, Sapir Y, Shapira M (2006) A conserved mechanism controls translation of Rubisco large subunit in different photosynthetic organisms. Plant Physiol 141:1089-1097

Dekker JP, Boekema EJ (2005) Supramolecular organization of thylakoid membrane proteins in green plants. Biochim Biophys Acta 1706:12-39

De la Fuente van Bentem S, Roitinger E, Anrather D, Csaszar E, Hirt H (2006) Phosphoproteomics as a tool to unravel plant regulatory mechanisms. Physiol Plant 126:110-119

De Vitry C, Desbois A, Redeker V, Zito F, Wollmann FA (2004) Biochemical and spectroscopic characterization of the covalent binding of heme to cytochrome b_6. Biochemistry 43:3956-3968

Diner BA, Ries DF, Cohen BN, Metz JG (1988) COOH-terminal processing of polypeptide D1 of the Photosystem II reaction center of Scenedesmus obliquus is necessary for the assembly of oxygen evolving complex. J Biol Chem 263:8972-8980

Edhofer I, Mühlbauer SK, Eichacker LA (1998) Light regulates the rate of translation elongation of chloroplast reaction center protein D1. Eur J Biochem 257:78-84

Esquivel MG, Pinto TS, Marin-Navarro J, Moreno J (2006) Substitution of tyrosine residues at the aromatic cluster around the beta A-beta B loop of rubisco small subunit affects the structural stability of the enzyme and the in vivo degradation under stress conditions. Biochemistry 45:5745-5753

Ferreira KN, Iverson TM, Maghlaoui K, Barber J, Iwata S (2004) Architecture of the photosynthetic oxygen-evolving center. Science 303:1831-1838

Finazzi G, Chasen C, Wollman FA, de Vitry C (2003) Thylakoid targeting of Tat passenger proteins shows no ΔpH dependence *in vivo*. EMBO J 22:807-815

Frottin F, Martinez A, Peynot P, Mitra S, Holz RC, Giglione C, Meinnel T (2006) The proteomics of N-terminal methionine cleavage. Molec Cellul Proteom 5:2336-2349

Fulgosi H, Vener AV, Altschmied L, Herrmann RG, Andersson B (1998) A novel multifunctional chloroplast protein: identification of a 40 kDa immunophilin-like protein located in the thylakoid lumen. EMBO J 17:1577-1587

Ganeteg U, Klimmek F, Jansson S (2004) Lhca5 – an LHC-type protein associated with photosystem I. Plant Mol Biol 54:641-651

Gatenby AA, Ellis RJ (1990) Chaperone function: the assembly of ribulose bisphosphate carboxylase-oxygenase. Annu Rev Cell Biol 6:125-149

Gau AE, Thole HH, Sokolenko A, Altschmied L, Herrmann RG (1998) PsbY, a novel manganese-binding, low-molecular-mass protein associated with photosystem II. Mol Gen Genet 260:56-68

Gerdes L, Bals T, Klostermann E, Karl M, Philippar K, Hunken M, Soll J, Schunemann D (2006) A second thylakoid membrane-localized Alb3/Oxa1/YidC homologue is involved in proper chloroplast biogenesis in *Arabidopsis thaliana*. J Biol Chem 281:16632-16642

Giglione C, Vallon O, Meinnel T (2003) Control of protein life-span by N-terminal methionine excision. EMBO J 22:13-23

Gioglione C, Boularot A, Meinnel T (2004) Protein N-terminal methionine excision. Cellul Molec Life Sci 61:1455-1474

Gutteridge S, Gatenby AA (1995) Rubisco synthesis, assembly, mechanism and regulation. Plant Cell 7:809-819

Göhre V, Ossenbühl F, Crèvecoeur M, Eichacker LA, Rochaix JD (2006) One of the two Alb3 proteins is essential for the assembly of the photosystems and for cell survival in *Chlamydomonas*. Plant Cell 18:1454-1466

Hager M, Biehler K, Illerhaus J, Ruf S, Bock R (1999) Targeted inactivation of the smallest plastid genome-encoded open reading frame reveals a novel and essential subunit of the cytochrome *b6f* complex. EMBO J 18:5834-5842

Hager M, Hermann M, Biehler K, Krieger-Liszkay A, Bock R (2002) Lack of the small plastid-encoded PsbJ polypeptide results in a defective water-splitting apparatus of photosystem II, reduced photosystem I levels and hypersensitivity to light. J Biol Chem 277:14031-14039

Hamel PP, Dreyfuss BW, Xie Z, Gabilly T, Merchant S (2003) Essential histidine and tryptophan residues in CcsA, a system II polytopic cytochrome c biogenesis protein. J Biol Chem 278:2593-2603

Hansson M, Vener AV (2003) Identification of three previously unknown *in vivo* phosphorylation sites in thylakoid membranes of *Arabidopsis thaliana*. Mol Cellul Proteom 2:550-559

Hartl FU (1996) Molecular chaperones in cellular protein folding. Nature 381:571-580

Hashimoto A, Yamamoto Y, Theg SM (1996) Unassembled subunits of the photosynthetic oxygen-evolving complex present in the thylakoid lumen are long-lived and assembly-competent. FEBS Lett 391:29-34

Hashimoto A, Ettinger WF, Yamamoto Y, Theg SM (1997) Assembly of newly imported oxygen-evolving complex subunits in isolated chloroplasts: sites of assembly and mechanism of binding. Plant Cell 9:441-452

Ifuku K, Yamamoto Y, Ono T, Ishihara S, Sato F (2005) PsbP protein, but not PsbQ protein, is essential for the regulation and stabilization of photosystem II in higher plants. Plant Physiol 139:1175-1184

Inoue K, Dreyfuss BW, Kindle KL, Stern DB, Merchant S, Sodeinde OA (1997) *Ccs1*, a nuclear gene required for the post-translational assembly of chloroplast *c*-type cytochromes. J Biol Chem 272:31747-31754

Jackson-Constan D, Akita M, Keegstra K (2001) Molecular chaperones involved in chloroplast protein import. Biochim Biophys Acta 1541:102-113

Jensen PE, Haldrup A, Rosgaard L, Scheller HV (2003) Molecular dissection of photosystem I in higher plants: topology, structure and function. Physiol Plant 119:313-321

Jeschke G, Bender A, Schweikardt T, Panek G, Decker H, Paulsen H (2005) Localization of the N-terminal domain in light-harvesting chlorophyll *a/b* protein by EPR measurements. J Biol Chem 280:18623-18630

Kapazoglou A, Mould RM, Gray JC (2000) Assembly of the Rieske iron-sulphur protein into the cytochrome *bf* complex in thylakoid membranes of isolated pea chloroplasts. Eur J Biochem 267:352-360

Kashino Y, Lauber WM, Carrol JA, Wang Q, Whitmarsh J, Satoh K, Pakrasi HB (2002) Proteomic analysis of highly active photosystem II preparation from the cyanobacterium *Synechocystis* sp. PCC 6803 reveals the presence of novel polypeptides. Biochemistry 41:8004-8012

Keegstra K, Cline K (1999) Protein import and routing systems of chloroplasts. Plant Cell 11:289-301

Keren N, Ohkawa H, Welsh EA, Liberton M, Pakrasi HB (2005) Psb29, a conserved 22-kD protein, functions in the biogenesis of photosystem II complexes in *Synechocystis* and *Arabidopsis*. Plant Cell 17:2768-2781

Khrouchtchova A, Hansson M, Paakkarinen V, Vainonen JP, Zhang SP, Jensen PE, Scheller HV, Vener AV, Aro EM, Haldrup A (2005) A previously found thylakoid membrane protein of 14 kDa (TMP14) is a novel subunit of plant photosystem I and is designated PSI-P. FEBS Lett 579:4808-4812

Kim J, Klein PG, Mullet JE (1991) Ribosomes pause at specific sites during synthesis of membrane-bound chloroplast reaction center protein D1. J Biol Chem 266:14931-14938

Klein RR, Mason HS, Mullet JE (1988) Light-regulated translation of chloroplast proteins. I. Transcripts of psaA, psaB, psbA, and rbcL are associated with polysomes in dark-grown and illuminated barley seedlings. J Cell Biol 106:289-301

Koivuniemi A, Aro EM, Andersson B (1995) Degradation of the D1 and D2 proteins of photosystem II in higher plants is regulated by reversible phosphorylation. Biochemistry 34:16022-16029

Komenda J, Reisinger V, Muller BC, Dobakova M, Granvogl B, Eichacker LA (2004) Accumulation of the D2 protein is a key regulatory step for assembly of the photosystem II reaction center complex in *Synechocystis* PCC 6803. J Biol Chem 279:48620-48629

Komenda J, Tichy M, Eichacker LA (2005) The PsbH protein is associated with the inner antenna CP47 and facilitates D1 processing and incorporation into PSII in the cyanobacterium *Synechocystis* PCC 6803. Plant Cell Physiol 46:1477-1483

Kouranov A, Schnell DJ (1996) Protein translocation at the envelope and thylakoid membranes of chloroplasts. J Biol Chem 271:31009-31012

Kovács L, Damkjær J, Kereïche S, Ilioaia C, Ruban AV, Boekema EJ, Jansson S, Horton P (2006) Lack of the light-harvesting complex CP24 affects the structure and function of the grana membranes of higher plant chloroplasts. Plant Cell 18:3106-3120

Kuras R, Wollman FA (1994) The assembly of cytochrome $b_6 f$ complexes: An approach using genetic transformation of the green alga *Chlamydomonas reinhardtii*. EMBO J 13:1019-1027

Kuras R, de Vitry C, Choquet Y, Girard-Bascou J, Culler D, Buschlein S, Merchant S, Wollman FA (1997) Molecular genetic identification of a pathway for heme binding to cytochrome b6. J Biol Chem 272:32427-32435

Kurisu G, Zhang HM, Smith JL, Cramer WA (2003) Structure of the cytochrome $b_6 f$ complex of oxygenic photosynthesis: tuning the cavity. Science 302:1009-1014

Kwon SJ, Choi EY, Choi YJ, Ahn JH, Park OK (2006) Proteomics studies of post-translational modifications in plants. J Exp Bot 57:1547-1551

Lennartz K, Plucken H, Seidler A, Westhoff P, Bechtold N, Meierhoff K (2001) HCF164 encodes a thioredoxin-like protein involved in the biogenesis of the cytochrome b(6)f complex in *Arabidopsis*. Plant Cell 13:2539-2551

Lennartz K, Bossmann S, Westhoff P, Bechtold N, Meierhoff K (2006) HCF153, a novel nuclear-encoded factor necessary during a post-translational step in biogenesis of the cytochrome b(6)f complex. Plant J 45:101-112

Lezhneva L, Meurer J (2004) The nuclear factor HCF145 affects chloroplast psaA-psaB-rps14 transcript abundance in *Arabidopsis thaliana*. Plant J 38:740-753

Lezhneva L, Amann K, Meurer J (2004) The universally conserved HCF101 protein is involved in assembly of [4Fe-4S]-cluster-containing complexes in *Arabidopsis thaliana* chloroplasts. Plant J 37:174-185

Li X, Henry R, Yuan J, Cline K, Hoffman NE (1995) A chloroplast homologue of the signal recognition particle subunit SRP54 is involved in the posttranslational integration of a protein into thylakoid membranes. Proc Natl Acad Sci USA 92:3789-3793

Loll B, Kern J, Saenger W, Zouni A, Biesiadka J (2005) Towards complete cofactor arrangement in the 3.0 Å resolution structure of photosystem II. Nature 438:1040-1044

Madueno F, Napier JA, Gray JC (1993) Newly imported Rieske iron-sulfur protein associates with both Cpn60 and Hsp70 in the chloroplast stroma. Plant Cell 5:1865-1876

Manieri W, Franchini L, Raeber L, Dai S, Stritt-Etter AL, Shurmann P (2003) N-terminal truncation of the variable subunit stabilizes spinach ferredoxin:thioredoxin reductase. FEBS Lett 549:167-170

Mant A, Schmidt L, Herrmann RG, Robinson C, Klosgen RB (1995) Sec-dependent thylakoid protein translocation. ΔpH requirement is dictated by passenger protein and ATP concentration. J Biol Chem 270:23275-23281

Mant A, Woolhead CA, Moore M, Henry R, Robinson C (2001) Insertion of PsaK into the thylakoid membrane in a "Horseshoe" conformation occurs in the absence of signal recognition particle, nucleoside triphosphates, or functional albino3. J Biol Chem 276:36200-36206

Melis A (1999) Photosystem-II damage and repair cycle in chloroplasts: what modulates the rate of photodamage *in vivo*? Trends Plant Sci 4:130-135

Meurer J, Plucken H, Kowallik KV, Westhoff P (1998) A nuclear-encoded protein of prokaryotic origin is essential for the stability of Photosystem II in *Arabidopsis thaliana*. EMBO J 17:5286-5297

Minai L, Wostrikoff K, Wollman FA Choquet Y (2006) Chloroplast biogenesis of Photosystem II cores involves a series of assembly-controlled steps that regulate translation. Plant Cell 18:159-175

Miyao M, Murata N (1989) The mode of binding of three extrinsic proteins of 33 kDa, 23 kDa and 18 kDa in the photosystem II complex of spinach. Biochim Biophys Acta 977:315-321

Moore M, Harrison MS, Peterson EC, Henry R (2000) Chloroplast Oxa1p homolog albino3 is required for post-translational integration of the light harvesting chlorophyll-binding protein into thylakoid membranes. J Biol Chem 275:1529-1532

Morais F, Barber J, Nixon PJ (1998) The chloroplast-encoded alpha subunit of cytochrome *b*559 is required for assembly of the photosystem two complex in both the light and the dark in *Chlamydomonas reinhardtii*. J Biol Chem 273:29315-29320

Mori H, Cline K (2001) Post-translational protein translocation into thylakoids by the Sec and ΔpH-dependent pathways. Biochim Biophys Acta 1541:80-90

Mori H, Summer EJ, Cline K (2001) Chloroplast TatC plays a direct role in thylakoid (Delta) pH-dependent protein transport. FEBS Lett 501:65-68

Motohashi K, Hisabori T (2006) HCF164 receives reducing equivalents from stromal thioredoxin across the thylakoid membrane and mediates reduction of target protein in the thylakoid lumen. J Biol Chem 281:35039-35047

Mould RM, Kapazoglou A, Gray JC (2001) Assembly of cytochrome f into the cytochrome bf complex in isolated pea chloroplasts. Eur J Biochem 268:792-799

Müller B, Eichacker LA (1999) Assembly of the D1 precursor in monomeric photosystem II reaction center precomplexes precedes chlorophyll a-triggered accumulation of reaction center II in barley etioplasts. Plant Cell 11:2365-2377

Naver H, Boudreau E, Rochaix JD (2001) Functional studies of Ycf3: its role in assembly of photosystem I and interactions with some of its subunits. Plant Cell 13:2731-2745

Neufeld S, Zinchenko V, Stephan DP, Bader KP, Pistorius EK (2004) On the functional significance of the polypeptide PsbY for photosynthetic water oxidation in the cyanobacterium *Synechocystis* sp strain PCC 6803. Mol Genet Genom 271:458-467

Nield J, Kruse O, Ruprecht J, da Fonseca P, Büchel C, Barber J (2000) Three-dimensional structure of *Chlamydomonas reinhardtii* and *Synechococcus elongatus* photosystem II complexes allows for comparison of their oxygen-evolving complex organization. J Biol Chem 275:27940-27946

Nielsen VS, Mant A, Knoetzel J, Moller BL, Robinson C (1994) Import of barley photosystem I subunit N into the thylakoid lumen is mediated by a bipartite presequence lacking an intermediate processing site. Role of delta pH in translocation across the thylakoid membrane. J Biol Chem 269:3762-3766

Nowaczyk M, Hebeler R, Schlodder E, Meyer H, Warscheid B, Rögner M (2006) Psb27, a cyanobacterial lipoprotein, is involved in the repair cycle of photosystem II. Plant Cell 18:3121-3131

Oelmüller R, Herrmann RG, Pakrasi HB (1996) Molecular studies of CtpA, the carboxyl-terminal processing protease for the D1 protein of the Photosystem II reaction center in higher plants. J Biol Chem 271:21848-21852

Ossenbühl F, Gohre V, Meurer J, Liszkay-Krieger A, Rochaix JD, Eichacker LA (2004) Efficient assembly of photosystem II in *Chlamydomonas reinhardtii* requires Alb3.1p, a homolog of *Arabidopsis* ALBINO3. Plant Cell 16:1790-1800

Ossenbühl F, Inaba-Sulpice M, Meurer J, Soll J, Eichacker LA (2006) The *Synechocystis* sp PCC 6803 Oxa1 homolog is essential for membrane integration of reaction center precursor protein pD1. Plant Cell 18:2236-2246

Page MLD, Hamel PP, Gabilly ST, Zegzouti H, Perea JV, Alonso JM, Ecker JR, Theg SM, Christensen SK, Merchant S (2006) A homolog of prokaryotic thiol disulfide transporter CcdA is required for the assembly of the Cytochrome b_6f complex in *Arabidopsis* chloroplasts. J Biol Chem 279:32474-32482

Peck SC (2006) Phosphoproteomics in *Arabidopsis*: moving from empirical to predictive science. J Exp Bot 57:1523-1527

Peng LW, Ma JF, Chi W, Guo JK, Zhu SY, Lu QT, Lu CM, Zhang LX (2006) LOW PSII ACCUMULATION1 is involved in efficient assembly of photosystem II in *Arabidopsis thaliana*. Plant Cell 18:955-969

Plucken H, Muller B, Grohmann D, Westhoff P, Eichacker LA (2002) The HCF136 protein is essential for assembly of the photosystem II reaction center in *Arabidopsis thaliana*. FEBS Lett 532:85-90

Prasil O, Adir N, Ohad I (1992) Dynamics of photosystem II. Mechanism of photoinhibition and recovery process. In: Barber J (ed) Topics in Photosynthesis. Vol II. The Netherlands: Elsevier, pp 293-348

Pursiheimo S, Martinsuo P, Rintamäki E, Aro EM (2003) Photosystem II protein phosphorylation follows four distinctly different regulatory patterns induced by environmental cues. Plant Cell Environ 26:1995-2993

Rinalducci S, Larsen MR, Mohammed S, Zolla L (2005) Novel phosphorylation site identification in spinach stroma membranes by titanium dioxide microcolumns and tandem mass spectrometry. J Proteome Res 5:973-982

Rintamäki E, Kettunen R, Aro EM (1996) Differential D1 dephosphorylation in functional and photodamaged photosystem II centres. Dephosphorylation is a prerequisite for degradation of damaged D1*. J Biol Chem 271:14870-14875

Rintamäki E, Martinsuo P, Pursiheimo S, Aro EM (2000) Cooperative regulation of light-harvesting complex II phosphorylation via the plastoquinol and the ferredoxin-thioredoxin system in chloroplasts. Proc Natl Acad Sci USA 97:11644-11649

Rochaix JD (1996) Post-transcriptional regulation of chloroplast gene expression in *Chlamydomonas reinhardtii*. Plant Mol Biol 32:209-221

Rochaix JD (2006) The role of nucleus- and chloroplast-encoded factors in the synthesis of the photosynthetic apparatus. In: Wise RR, Hoober JK (eds) The Structure and Function of Plastids. The Netherlands: Springer, pp 145-165

Rodermel S, Haley J, Jiang CZ, Tsai CH, Bogorad L (1996) A mechanism for intergenomic integration: abundance of ribulose bisphosphate carboxylase small-subunit protein influences the translation of the large-subunit mRNA. Proc Natl Acad Sci USA 93:3881-3885

Rokka A, Suorsa M, Saleem A, Battchikova N, Aro EM (2005) Synthesis and assembly of thylakoid protein complexes: multiple assembly steps of photosystem II. Biochem J 388:159-168

Roose JL, Pakrasi HB (2004) Evidence that D1 processing is required for manganese binding and extrinsic protein assembly into photosystem II. J Biol Chem 279:45417-45422

Rossignol M (2006) Proteomic analysis of phosphorylated proteins. Curr Opin Plant Biol 9:538-543
Sakamoto W (2006) Protein degradation machineries in plastids. Annu Rev Plant Biol 57:599-621
Sane AP, Stein B, Westhoff P (2005) The nuclear gene HCF107 encodes a membrane-associated R-TPR (RNA tetratricopeptide repeat)-containing protein involved in expression of the plastidial psbH gene in *Arabidopsis*. Plant J 42:720-730
Schlicher T, Soll J (1996) Molecular chaperones are present in the thylakoid lumen of pea chloroplasts. FEBS Lett 379:302-304
Schlicher T, Soll J (1997) Chloroplastic isoforms of DnaJ and GrpE in pea. Plant Mol Biol 33:181-185
Schneider D, Berry S, Rich P, Seidler A, Rögner M (2001) A regulatory role of the PetM subunit in a cyanobacterial cytochrome *b6f* complex. J Biol Chem 276:16780-16785
Schroda M (2004) The *Chlamydomonas* genome reveals it secrets: chaperone genes and the potential roles of their gene products in the chloroplast. Photosynth Res 82:221-240
Schroda M, Kropat J, Oster U, Rüdiger W, Vallon O, Wollman FA, Beck CF (2001) A role for molecular chaperones in assembly and repair of Photosystem II. Biochem Soc Trans 29:413-418
Schuenemann D, Amin P, Hartmann E, Hoffman NE (1999) Chloroplast SecY is complexed to SecE and involved in the translocation of the 33-kDa but not the 23-kDa subunit of the oxygen-evolving complex. J Biol Chem 274:12177-12182
Schuster G, Timberg R, Ohad I (1988) Turnover of thylakoid photosystem II proteins during photoinhibition of *Chlamydomonas reinhardtii*. Eur J Biochem 177:403-410
Schwenkert S, Umate P, Dal Bosco C, Volz S, Mlcochová L, Zoryan M, Eichacker LA, Ohad I, Herrmann RG, Meurer J (2006) PsbI affects the stability, function, and phosphorylation patterns of photosystem II assemblies in tobacco. J Biol Chem 281:34227-34238
Seelert H, Poetsch A, Dencher N, Engel A, Stahlberg H, Muller DJ (2000) Proton-powered turbine of a plant motor. Nature 405:418-419
Settles AM, Yonetani A, Baron A, Bush DR, Cline K, Martienssen R (1997) Sec-independent protein translocation by the maize Hcf106 protein. Science 278:1467-1470
Sharma J, Panico M, Barber J, Morris HR (1997) Characterization of the low molecular weight photosystem II reaction center subunits and their light-induced modifications by mass spectrometry. J Biol Chem 272:3935-3943
Shen G, Antonkine ML, van der Est A, Vassiliev IR, Brettel K, Bittl R, Zech SG, Zhao J, Stehlik D, Bryant DA, Golbeck JH (2002a) Assembly of photosystem I. II. Rubredoxin is required for the in vivo assembly of F_x in *Synechococcus* sp. PCC 7002 as shown by optical and EPR spectroscopy. J Biol Chem 277:20355-20366
Shen G, Zhao J, Reimer SK, Antonkine ML, Cai Q, Weiland SM, Goldbeck JH, Bryant DA (2002b) Assembly of Photosystem I. I Inactivation of the *rubA* gene encoding a membrane-associated rubredoxin in the cyanobacterium *Synechococcus* sp. PCC 7002 causes a loss of photosystem I activity. J Biol Chem 277:20343-20354
Shi LX, Lorcovic ZJ, Oelmüller R, Schröder WP (2000) The low molecular mass PsbW protein is involved in the stabilization of the dimeric photosystem II complex in *Arabidopsis thaliana*. J Biol Chem 275:37945-37950
Sjögren LLE, MacDonald TM, Sutinen S, Clarke AK (2004) Inactivation of the clpC1 gene encoding a chloroplast Hsp100 molecular chaperone causes growth retardion, leaf

chlorosis, lower photosynthetic activity, and a specific reduction in photosystem content. Plant Physiol 136:4114-4126

Spreitzer RJ (1993) Genetic dissecting of Rubisco structure and function. Annu Rev Plant Physiol Plant Mol Biol 44:1-49

Steiner JM, Kocher T, Nagy C, Löffelhardt W (2002) Chloroplast SecE: evidence for spontaneous insertion into the thylakoid membrane. Biochem Biophys Res Commun 293:747-752

Stroebel D, Choquet Y, Popot JL, Picot D (2003) An atypical haem in the cytochrome b_6f complex. Nature 426:413-418

Stöckel J, Oelmüller R (2004) A novel protein for photosystem I biogenesis. J Biol Chem 279:10243-10251

Stöckel J, Bennewitz S, Oelmüller R (2006) The evolutionarily conserved tetratrico peptide repeat protein pale yellow green7 is required for photosystem I accumulation in *Arabidopsis* and copurifies with the complex. Plant Physiol 141:870-878

Sugimoto I, Takahashi Y (2003) Evidence that the PsbK polypeptide is associated with the photosystem II core antenna complex CP43. J Biol Chem 278:45004-45010

Sundberg E, Slagter JG, Fridborg I, Cleary SP, Robinson C, Coupland G (1997) ALBINO3, an *Arabidopsis* nuclear gene essential for chloroplast differentiation, encodes a chloroplast protein that shows homology to proteins present in bacterial membranes and yeast mitochondria. Plant Cell 9:717-730

Suorsa M, Aro EM (2007) Expression, assembly and auxiliary functions of photosystem II oxygen evolving proteins in higher plants. Photosynth Res: in press

Suorsa M, Regel R, Paakkarinen V, Battchikova N, Herrmann RG, Aro EM (2004) Protein assembly of photosystem II and accumulation of subcomplexes in the absence of low molecular mass subunits PsbL and PsbI. Eur J Biochem 271:96-107

Suorsa M, Sirpiö S, Allahverdiyeva Y, Paakkarinen V, Mamedov F, Styring S, Aro EM (2006) PsbR, a missing link in the assembly of the oxygen-evolving complex of plant photosystem II. J Biol Chem 281:145-150

Szabò I, Seraglia R, Rigoni F, Traidi P, Giacometti GM (2001) Determination of photosystem II subunits by matrix-assisted laser desorption/ionization mass spectrometry. J Biol Chem 276:13784-13790

Tikkanen M, Piippo M, Suorsa M, Sirpio S, Mulo P, Vainonen J, Vener AV, Allahverdiyeva Y, Aro EM (2006) State transitions revisited – a buffering system for dynamic low light acclimation of *Arabidopsis.* Plant Mol Biol 62:779-793

Touraine B, Boutin JP, Marion-Poll A, Briat JF, Peltier G, Lobreaux S (2004) Nfu2: a scaffold protein required for [4Fe-4S] and ferredoxin iron-sulphur cluster assembly in *Arabidopsis* chloroplasts. Plant J 40:101-111

Tsiotis G, Psylinakis M, Woplensinger B, Lustig A, Engel A, Ghanotakis D (1999) Investigation of the structure of spinach photosystem II reaction center complex. Eur J Biochem 259:320-324

Tu CJ, Schuenemann D, Hoffmann NE (1999) Chloroplast FtsY, chloroplast signal recognition particle, and GTP are required to reconstitute the soluble phase of light-harvesting chlorophyll protein transport into thylakoid membrane. J Biol Chem 274:27219-27224

Turkina M, Villarejo A, Vener AV (2004) The transit peptide of CP29 thylakoid protein is not removed but undergoes acetylation and phosphorylation. FEBS Lett 564:104-108

Van der Laan M, Nouwen NP, Driessen AJM (2005) YidC – an evolutionary conserved device for the assembly of energy-transducing membrane protein complexes. Curr Opin Microbiol 8:182-187

Van Wijk KJ (2001) Proteins involved in biogenesis of the thylakoid membrane. In: Aro EM, Andersson B (eds) Regulation of Photosynthesis. The Netherlands: Kluwer Academic Publishers, pp 153-175

Van Wijk KJ, Eichacker L (1996) Light is required for efficient translation elongation and subsequent integration of the D1-protein into Photosystem II. FEBS Lett 388:89-93

Van Wijk KJ, Bingsmark S, Aro EM, Andersson B (1995) *In vitro* synthesis and assembly of photosystem II core proteins. The D1 protein can be incorporated into photosystem II in isolated chloroplasts and thylakoids. J Biol Chem 270:25685-25695

Vener AV (2007) Environmentally modulated phosphorylation and dynamics of proteins in photosynthetic membranes. Biochem Biophys Acta: in press

Vener AV, Ohad I, Andersson B (1998) Protein phosphorylation and redox sensing in chloroplast thylakoids. Curr Opin Plant Biol 1:217-223

Vener AV, Rokka A, Fulgosi H, Andersson B, Herrmann RG (1999) A cyclophilin-regulated PP2A-like protein phosphatase in thylakoid membrane of plant chloroplasts. Biochemistry 38:14955-14965

Vener AV, Harms A, Sussman MR, Vierstra RD (2001) Mass spectrometric resolution of reversible protein phosphorylation in photosynthetic membranes of *Arabidopsis thaliana*. J Biol Chem 276:6959-6966

Wilde A, Lünser K, Ossenbühl F, Nickelsen J, Börner T (2001) Characterization of the cyanobacterial *ycf37*: mutation decreases the photosystem I content. Biochem J 357:211-216

Wollman FA, Minai L, Nechushtai R (1999) The biogenesis and assembly of photosynthetic proteins in thylakoid membranes. Biochim Biophys Acta 1411:21-85

Woolhead CA, Thompson SJ, Moore M, Tissier C, Mant A, Rodger A, Henry R, Robinson C (2001) Distinct Albino3-dependent and –independent pathways for thylakoid membrane protein insertion. J Biol Chem 276:40841-40846

Wostrikoff K, Girard-Bascou J, Wollman FA, Choquet Y (2004) Biogenesis of PSI involves a cascade of translational autoregulation in the chloroplast of *Chlamydomonas*. EMBO J 23:2696-2705

Xie ZY, Merchant S (1996) The plastid-encoded ccsA gene is required for heme attachment to chloroplast c-type cytochromes. J Biol Chem 271:4632-4639

Yabe T, Morimoto K, Kikuchi S, Nishio K, Terashima I, Nakai M (2004) The *Arabidopsis* chloroplastic NifU-like protein CnfU, which can act as an iron-sulphur cluster scaffold protein, is required for biogenesis of ferrodoxin and photosystem I. Plant Cell 16:993-1007

Yamamoto Y (2001) Quality control of photosystem II. Plant Cell Physiol 42:121-128

Yokthongwattana K, Chrost B, Behrman S, Casper-Lindley C, Melis A (2001) Photosystem II damage and repair cycle in the green alga *Dunaliella salina*: involvement of a chloroplast-localized HSP70. Plant Cell Physiol 42:1389-1397

Yu J, Smart LB, Jung YS, Golbeck J, McIntosh L (1995) Absence of PsaC subunit allows assembly of photosystem I core but prevents the binding of PsaD and PsaE in *Synechocystis* sp. PCC6803. Plant Mol Biol 29:331-342

Zak E, Pakrasi HB (2000) The BtpA protein stabilizes the reaction center protein of photosystem I in the cyanobacterium *Synechocystis* sp. PCC 6803 at low temperature. Plant Physiol 123:215-222

Zer H, Vink M, Keren N, Dilly-Hartwig HG, Paulsen H, Herrmann RG, Andersson B, Ohad I (1999) Regulation of thylakoid protein phosphorylation at the substrate level: reversible light-induced conformational changes expose the phosphorylation site of the light-harvesting complex II. Proc Natl Acad Sci USA 96:8277-8282

Zhang L, Paakkarinen V, van Wijk KJ, Aro EM (1999) Co-translational assembly of the D1 protein into photosystem II. J Biol Chem 274: 16062-16067

Zhang L, Paakkarinen V, van Wijk KJ, Aro EM (2000) Biogenesis of the chloroplast-encoded D1 protein: regulation of translation elongation, insertion, and assembly into photosystem II. Plant Cell 12:1769-1782

Zheleva D, Sharma J, Panico M, Morris HR, Barber J (1998) Isolation and characterization of monomeric and dimeric CP47-reaction center photosystem II complexes. J Biol Chem 273:16122-16127

Zouni A, Witt HT, Kern J, Fromme P, Krauss N, Saenger W, Orth P (2001) Crystal structure of photosystem II from *Synechococcus* elongatus at 3.8 Å resolution. Nature (London) 409:739-743

Zygadlo A, Robinson C, Scheller HV, Mant A, Jensen PE (2006) The properties of the positively charged loop region in PSI-G are essential for its “spontaneous” insertion into thylakoids and rapid assembly into the photosystem I complex. J Biol Chem 281:10548-10554

Aro, Eva-Mari
Department of Biology, University of Turku, FIN-20014 Turku, Finland
evaaro@utu.fi

Kanervo, Eira
Department of Biology, University of Turku, FIN-20014 Turku, Finland

Suorsa, Marjaana
Department of Biology, University of Turku, FIN-20014 Turku, Finland

Protein stability and degradation in plastids

Zach Adam

Abstract

Steady-state levels of chloroplast proteins rely on the balance between synthesis and degradation rates. Thus, the importance of protein-degradation processes in shaping the chloroplast proteome, and hence proper organellar functioning, cannot be overestimated. Chloroplast proteases and peptidases participate in chloroplast biogenesis through maturation or activation of pre-proteins, adaptation to changing environmental conditions through degradation of certain proteins, and maintenance of protein quality through degradation of unassembled or damaged proteins. These activities are mediated by ATP-dependent and -independent proteases, many of which are encoded by multigene families. Newly imported proteins are processed by stroma- and thylakoid-localized peptidases that remove signal sequences, which are then further degraded. The multisubunit ATP-dependent Clp and FtsH complexes degrade housekeeping and oxidatively damaged proteins in the stroma and thylakoid membranes, respectively. A number of other chloroplast proteases have been identified, but their function and substrates are still unknown, as are the nature of degradation signals and determinants of protein instability. Future research is expected to focus on these questions.

1 Introduction

The chloroplast proteome comprises more than 2000 nuclear- and chloroplast-encoded proteins. Steady-state levels of these proteins are determined by the balance between transcription and translation rates on the one hand, and degradation rates on the other. Thus, the importance of protein-degradation processes in shaping the chloroplast proteome, and hence proper functioning of the organelle, cannot be overestimated. Proteolytic activities, determined as cleavage of peptide bonds, are carried out by proteases or peptidases, which differ in a number of aspects. Some activities are limited to the hydrolysis of a single bond in a given substrate, whereas others function processively. Products of such activities can be either free amino acids or peptides of different lengths, from di- and tri-peptides to much longer ones. The hydrolysis itself can be catalyzed by different mechanisms, depending on the chemistry of the active site, giving rise to the categorization of proteases into seven different families based on the catalytic centers: serine, cysteine, aspartic, metalloproteases, threonine, glutamic, and peptidases of unknown catalytic mechanisms. Although cleavage of a peptide bond does not require

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0227 / Published online: 22 May 2007

metabolic energy, some proteases couple the hydrolysis of ATP to the unfolding of their substrates as a prerequisite for the actual cleavage of peptide bonds. The *in vivo* contexts of proteolytic reactions are also highly variable: maturation or activation of pre-proteins require either N- or C-terminal processing by specific peptidases; proteolytic enzymes participate in some cases of signal transduction by releasing factors from membranes into the soluble phase; rapid turnover rates of certain regulatory proteins allow their function as 'timing proteins' in the control of gene expression; protein quality control is maintained by the degradation of unassembled or damaged proteins. Thus, proteolytic processes are intimately involved in almost every aspect of the cell's life cycle. Organelles such as chloroplasts are no exception. Although examples have been documented for the involvement of only some of the above proteolytic processes in chloroplasts, it is already clear that proteases play an essential role in this organelle's biogenesis and function.

Looking back 25 years or so, research in the field of chloroplast proteolysis can be roughly divided into three periods. During the 1980s and early 1990s, a number of proteolytic processes were documented and characterized. However, attempts to identify the proteases involved in these processes, primarily through biochemical approaches, were largely unsuccessful. In the mid-1990s, the identities of the chloroplast proteases began to be revealed. These all turned out to be homologues of known bacterial proteases. Completion of the *Arabidopsis* genome project enabled comprehensive homology searches, and in conjunction with the use of programs for predicting the intracellular location of proteins, a list of putative components of the proteolytic machinery of chloroplasts was compiled (Sokolenko et al. 2002). Research in the field in recent years has been characterized by attempts to link identified proteases with the previously described proteolytic processes, and to reveal their physiological roles, primarily through a reverse-genetics approach.

This chapter reviews the different components of the chloroplast proteolytic machinery, the different proteolytic processes delineated to date in chloroplasts, and the limited information on determinants of protein stability and instability in chloroplasts. Where possible, proteolytic enzymes will be referred to according to their names and classification in the peptidase database MEROPS (Rawlings et al. 2006) (http://merops.sanger.ac.uk/index.htm) and its corresponding handbook (Barrett et al. 2004).

2 Major chloroplast proteases

Given the prokaryotic evolutionary origin of chloroplasts, it is not surprising that all chloroplast proteases are homologues of known bacterial ones. In fact, this relationship facilitated the initial identification of some chloroplast proteases. Proteases involved in intracellular proteolysis in any biological system can be categorized, based on their energy requirement, into ATP-dependent and -independent ones. Hydrolysis of a peptide bond does not require metabolic energy. Thus, the

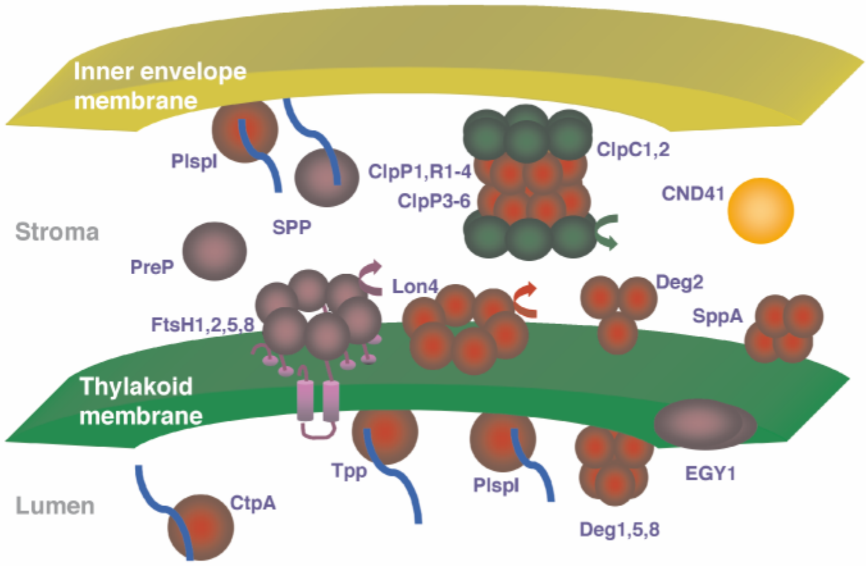


Fig. 1. Distribution and characteristics of chloroplast proteases. Serine proteases are depicted in red, metalloproteases are in purple, and the one aspartic protease is in orange. Arrows indicate ATPases that are found either together with the protease domain on the same polypeptide (FtsH and Lon proteases) or on a separate polypeptide (Clp protease). Ribbons attached to peptidases indicate precursor proteins processed by them.

requirement for ATP in certain enzymes is limited to unfolding the substrate and feeding it into a catalytic chamber, which is secluded from the cellular environment, a paradigm that led to classifying these enzymes as self-compartmentalizing proteases (Baumeister et al. 1998). Similar to all bacteria, chloroplasts contain both ATP-dependent and independent proteases. However, whereas *Escherichia coli* and most other bacteria contain single genes encoding these enzymes, higher plants have evolved multiple genes for most of them (Adam et al. 2001). These enzymes are described below.

2.1 Clp protease

Clp protease in *E. coli* is a multisubunit complex, composed of two main components, proteolytic, and regulatory (for review, see Sauer et al. 2004). The proteolytic chamber is made up of two heptameric rings of the serine peptidase ClpP. Together, they form a barrel-like structure with a narrow inlet and an internal cavity where the active-site subunits, composed of the catalytic triad of Ser-His-Asp, are located. The openings of the ClpP subcomplex are capped by hexameric rings of specific ATP-dependent chaperones of the AAA^+ superfamily (Neuwald et al. 1999), either ClpA or ClpX, which recognize potential substrates, unfold them, and feed them into the catalytic chamber. ClpAP and ClpXP specifically degrade different regulatory proteins, and participate in protein quality control by degrading aggregated, misfolded and otherwise abnormal proteins (Sauer et al. 2004).

Chloroplast Clp protease is much more complex (for recent reviews, see Clarke et al. 2005; Adam et al. 2006). ClpP in *Arabidopsis* (peptidase S14.002) is encoded by six different genes, giving rise to proteins of 20 to 29 kDa, five of which are targeted to chloroplasts. Only one of these, ClpP2, is targeted to mitochondria where, together with ClpX, they form the mitochondrial Clp complex (Halperin et al. 2001b; Peltier et al. 2004). One of the ClpPs, ClpP1, is the only component of the chloroplast proteolytic machinery that is encoded in the organelle's genome. The *Arabidopsis* nuclear genome encodes four ClpP-like proteins, designated ClpR. These are similar in size and sequence to ClpP and located exclusively in chloroplasts, but they lack the conserved residues of the catalytic triad, and thus are not expected to perform a proteolytic function. The ClpP cognate chaperones in *Arabidopsis* include two copies of ClpC, the plant homologue of ClpA, and another related protein designated ClpD, all located in the chloroplast, and three ClpX proteins that are located in the mitochondria. Expression of all of these, with the exception of ClpD, appears to be constitutive under different short- and long-term stress conditions (Zheng et al. 2002). Additional Clp proteins include two copies of ClpS and one of ClpT. ClpS is unique to land plants, being absent from algae and cyanobacteria, and shares homology with the N terminus of ClpC (Peltier et al. 2001). ClpT is homologous to the *E. coli* ClpS (which shares no homology with the *Arabidopsis* ClpS), a substrate modulator of the bacterial ClpAP complex (Dougan et al. 2002), which is essential for the operation of the N-end rule pathway (see Section 6) in bacteria (Erbse et al. 2006).

Native isoelectric focusing followed by mass spectrometry revealed that the core of the chloroplast Clp protease is a complex of 325 to 350 kDa, composed of one to three copies of ClpP (ClpP1, ClpP3-ClpP6), four copies of ClpR (ClpR1-ClpR4), and one copy of ClpS (Peltier et al. 2001). Interestingly, the same core Clp complex is found in the stroma of chloroplasts and non-green plastids from roots and flowers (Peltier et al. 2004). More recent work, using native polyacrylamide gel electrophoresis followed by immunoblot analysis with specific antibodies for each of the Clp isomers, has shed more light on the structure of the core Clp complex. Two sub-core complexes were observed, probably corresponding to the two different rings. Whereas a 335-kDa core contained all chloroplastic ClpP and ClpR subunits, two smaller sub-complexes had different compositions: a 230-kDa complex contained ClpP1 and ClpR1-ClpR4, and a 180-kDa complex contained ClpP3-ClpP6 (Sjogren et al. 2006). How this asymmetrical distribution of subunits between the different rings affects the function of the Clp core remains to be determined.

Knockout and downregulation of Clp genes revealed some of the functions of Clp protease *in vivo*. Disruption of the chloroplast ClpP1 gene in tobacco resulted in loss of shoot development (Shikanai et al. 2001; Kuroda and Maliga 2003). Since there is still no reliable chloroplast transformation system for *Arabidopsis*, it is not known whether inactivation of the ClpP1 gene in this species would be as detrimental as in tobacco. Inactivation or downregulation of several ClpP, ClpR and ClpC genes in *Arabidopsis* led to phenotypes of variable severity. Viable ClpP4 and ClpP6 knockout mutants could not be obtained, but repression of their expression by antisense constructs resulted in slow growth and a variegated 'yellow-heart' phenotype (Sjogren et al. 2006; Zheng et al. 2006). Yellow variegated leaves were also observed in rice as a result of disrupting the ClpP5 gene (Tsugane et al. 2006). Mutations leading to loss of ClpR1 (Koussevitzky et al. 2007) or a lower level of ClpR2 (Rudella et al. 2006) also resulted in a slow-growing, pale green phenotype. These results suggest that ClpR, although lacking a proteolytic site, is important for stabilizing the structure and/or regulating the function of the chloroplast Clp protease. However, details of ClpR's involvement are unknown.

Mutations in the regulatory ATPase have somewhat less severe consequences than mutations in ClpPs. ClpC1 mutants can grow autotrophically, but they are small and pale relative to wild type plants (Constan et al. 2004; Sjogren et al. 2004; Kovacheva et al. 2005). In contrast, a ClpC2 mutant is indistinguishable from the wild type (Park and Rodermel 2004), suggesting that the two copies of ClpC are redundant. It is not known why these two mutants have different phenotypes, but it might be due to different levels of accumulation of these two isomers, such that loss of the more abundant one has more severe effects.

Insights into the structure of the Clp core complex were also obtained from these mutants. The T-DNA line of ClpR2 was not a complete knockout. Instead, it contained lower levels of the ClpR2 transcript and protein (Rudella et al. 2006). Interestingly, this was accompanied by a decrease in the level of all other Clp core-complex subunits, demonstrating that they are all essential for the assembly and stability of the complex. Furthermore, the same analysis that suggested that the two heptameric rings have different compositions (Sjogren et al. 2006) was performed on the ClpP6 mutant; in the absence of ClpP6, the other components of the ring (ClpP3-ClpP5) did not accumulate, whereas components of the other ring (ClpP1 and ClpR1-ClpR4) did. However, only small amounts of the existing rings dimerized (Sjogren et al. 2006). Thus, it appears that each of the rings is stabilized only if it contains the full complement of its components, and that most, if not all of these components are not redundant.

2.2 FtsH protease

The *E. coli* FtsH is a membrane-bound ATP-dependent metalloprotease (for a recent review, see Ito and Akiyama 2005). Of all the ATP-dependent proteases in this organism, FtsH is the only essential one. Unlike Clp protease, its proteolytic and ATPase domains are found on the same polypeptide and not on separate ones. The N terminus of the protein contains two trans-membrane helices, which anchor the protein to the plasma membrane. This region is followed by the ATPase domain, which relates this protein to the AAA^+ superfamily (Neuwald et al. 1999). The proteolytic domain of the protein is found in the C terminus of the protein, and it contains the zinc-binding motif His-Glu-X-X-His, which serves as the catalytic site of the protease. Similar to other ATP-dependent proteases, FtsH forms a hexameric ring-like structure, in which access to the proteolytic site is controlled by the ATPase domain. Details of these structural features were recently revealed when the three-dimensional structure of bacterial FtsH was determined (Bieniossek et al. 2006; Suno et al. 2006).

The FtsH gene family in *Arabidopsis* contains twelve members (for recent reviews, see Adam et al. 2005, 2006; Sakamoto 2006). Products of three of these (FtsH3, FtsH4, and FtsH10) are targeted to the mitochondria whereas the other nine (FtsH1, FtsH2, FtsH5-FtsH9, FtsH11, and FtsH12) are targeted to the chloroplasts, as revealed by transient-expression assays with GFP fusions (Sakamoto et al. 2003). Mass spectrometry analyses confirmed the presence of FtsH1, FtsH2, FtsH5 and FtsH8 in chloroplasts (Friso et al. 2004; Sinvany-Villalobo et al. 2004; Yu et al. 2004). Immunoblot analysis of isolated organelles suggested that whereas FtsH4 is located exclusively in the mitochondria, FtsH11 is dually targeted to both the mitochondria and chloroplasts (Urantowka et al. 2005). The chloroplast-targeted FtsH1 and FtsH5, FtsH2 and FtsH8, and FtsH7 and FtsH9 comprise three pairs of duplicated genes (see phylogenetic trees in Sakamoto et al. 2003; Yu et al. 2004; Adam et al. 2005). Of the four proteins that were indeed identified in chloroplasts, FtsH2 is the most abundant, followed by FtsH5, FtsH8 and FtsH1, in decreasing order of abundance (Sinvany-Villalobo et al. 2004). The differential abundance of these four FtsHs is positively correlated with the severity of phenotypes associated with mutations in the corresponding genes. FtsH2 mutants have variegated leaves, containing distinct green and yellow/white sectors (Chen et al. 2000; Takechi et al. 2000). Mutants in FtsH5 have only slightly variegated leaves (Sakamoto et al. 2002), whereas mutants in FtsH1 and FtsH8 are indistinguishable from wild type plants (Sakamoto et al. 2003). These mutant phenotypes suggest that FtsH might be involved in chloroplast biogenesis.

The size of the chloroplast FtsH monomer (peptidase M41.005) is ~74 kDa. It is located in the thylakoid membrane with its ATPase and proteolytic domains facing the stroma (Lindahl et al. 1996). It forms a complex of 400 to 450 kDa, which is probably a hexamer (Sakamoto et al. 2003; Yu et al. 2004). Several lines of evidence suggest that FtsH complexes are heteromeric: FtsH2 and FtsH5, identified by either specific antibodies or mass spectrometry, co-migrate on native gels, sucrose gradients and size-exclusion chromatography. Moreover, in mutants lacking one of these proteins, the level of the other is also reduced (Sakamoto et al. 2003; Yu et al. 2004). Although an authentic native FtsH complex has not yet been purified to homogeneity, insights into its composition can be obtained from overexpression experiments and analysis of single and double knockout mutants. Overexpression of FtsH8 compensates for the loss of its duplicated gene FtsH2 (Yu et al. 2004), and FtsH1 can compensate for the loss of its close homologue FtsH5 (Yu et al. 2005). However, attempts to restore the wild type phenotype by overexpressing FtsH5 in the FtsH2-mutant background were unsuccessful. Furthermore, double mutants of duplicated genes, either FtsH1 and FtsH5, or FtsH2 and FtsH8, were completely albino, and could grow only on agar plates supplemented with sucrose. In each of these double mutants, the presumably remaining FtsHs did not accumulate (Zaltsman et al. 2005b). Taken together, these results suggested that the chloroplast FtsH complex is a hetero-oligomer composed of two types of subunits, each encoded by duplicated genes. Whereas subunits within a type are redundant, the presence of subunits from both types is essential for accumulation of the complex (Adam et al. 2005, 2006; Zaltsman et al. 2005b). Moreover, previous quantification of the different isomers (Sinvany-Villalobo et al. 2004) now suggests that the FtsH hexamer is composed of two subunits of ‘type A’—FtsH1 and/or FtsH5, and four subunits of ‘type B’—FtsH2 and/or FtsH8 (Zaltsman et al. 2005b; Adam et al. 2006). Why two types of subunits are needed for accumulation of the complex is not clear, but these conclusions, based primarily on genetic analyses, will have to be confirmed by a biochemical approach.

The FtsH protein does not accumulate in etiolated seedlings (Lindahl et al. 1996). Expression studies on the different FtsH genes demonstrated an increase in all of their transcript levels in response to short-term (2.5 h) exposure to high light intensity. Temperature shifts, to either high or low temperature, had almost no effect on FtsH transcript level (Sinvany-Villalobo et al. 2004). Interestingly, exposure to high light resulted in a transient decrease in the level of the FtsH protein itself (Zaltsman et al. 2005a). Thus, it is possible that the increase in FtsH transcript level in response to high light only compensates for the temporary loss of FtsH protein induced by this treatment, enabling its restoration to normal levels. Consistent with this view are recent findings from a proteomic analysis of the response to high light, where no increase in the level of chloroplast proteases or chaperones, including FtsH, was observed (Giacomelli et al. 2006). Differential spatial expression of different FtsHs can also be ruled out, as GUS-fusion experiments revealed similar patterns for FtsH1, FtsH2, FtsH5, and FtsH8 (Yu et al. 2004, 2005). However, FtsH transcript and protein levels appear to increase during the initial stages of *Arabidopsis* seedling development (Zaltsman et al. 2005a).

2.3 Lon protease

Another ATP-dependent protease in *E. coli* is Lon. Similar to FtsH protease, its ATPase and proteolytic domains are found on the same polypeptide. It is a hexameric serine protease that uses a Ser-Lys dyad in its active site (Botos et al. 2004), which is required for the degradation of abnormal as well as several short-lived regulatory proteins (for review, see Gottesman 1996). Plant homologues of Lon protease (peptidase S16.003) have been identified in mitochondria (Barakat et al. 1998; Sarria et al. 1998), but apparently they are also found in chloroplasts. Transient-expression assay of GFP fusions revealed that of the four genes found in *Arabidopsis*, products of Lon1 and Lon2 are targeted to the mitochondria and peroxisomes, respectively, whereas Lon4 is dually targeted to both the mitochondria and chloroplasts (Ostersetzer et al. 2007). Proteomic analysis of mitochondria revealed the presence of Lon3 in this organelle (Heazlewood et al. 2004; Heazlewood and Millar 2005). Moreover, immunoblot analysis of purified chloroplasts with an antibody against Lon1 revealed a cross-reacting protein of the correct size that was tightly associated with thylakoid membranes facing the stroma (Ostersetzer et al. 2007). This association is consistent with the previous localization of plant Lon to the inner membrane of the mitochondria (Sarria et al. 1998), and the archaeal Lon to the plasma membrane (Fukui et al. 2002). Expression of Lon4 appears to be constitutive, as its transcript level did not change upon exposure to high light, or low or high temperatures (Sinvany-Villalobo et al. 2004). The oligomeric structure of the plant Lon protease, in chloroplasts or mitochondria, is not known.

2.4 Deg protease

The *E. coli* DegP (also known as HtrA) is an ATP-independent serine protease, peripherally attached to the periplasmic side of the plasma membrane, which is essential for survival at elevated temperatures (for review, see Clausen et al. 2002). DegP forms a hexameric complex, made up of two staggered trimers. Its monomer size is 48 kDa, composed of two distinct domains: the proteolytic domain, with a typical catalytic triad of Ser-Asp-His, is found at the N terminus. Two PDZ domains in tandem, implicated in protein-protein interactions, are located at the C terminus. In addition to its proteolytic activity, DegP demonstrates chaperone activity. Whereas the chaperone activity dominates at low temperatures below 22^{o}C, the proteolytic activity is manifested at elevated temperatures (Spiess et al. 1999). This transition between the two activities can be explained by the structure of the protein. At normal temperatures, the active site of the protease is blocked by segments of the protein itself. At elevated temperatures, a conformational change is induced, which makes the active site accessible to substrates (Krojer et al. 2002).

The *Arabidopsis* genome contains 16 genes homologous to DegP. These have been recently renamed Deg proteases (for a recent review, see Huesgen et al. 2005). Products of four of these have been identified in chloroplasts. Deg2 is found peripherally attached to the stromal side of the thylakoid membrane (Haussuhl et al. 2001), whereas Deg1, Deg5, and Deg8 are found in the lumen (Itzhaki et al. 1998; Peltier et al. 2002; Schubert et al. 2002). Size-exclusion chromatography demonstrated that recombinant Deg1 can form hexamers (Chassin et al. 2002). Nevertheless, the oligomeric structure of native Deg1 has yet to be determined. The presence of Deg1, Deg5, and Deg8 in the same compartment suggests their interaction, but this possibility still needs to be tested experimentally. Expression analysis has shown a fourfold increase in the transcript level of Deg1 and Deg8 upon exposure of *Arabidopsis* plants to a short-term high-light treatment, whereas temperature shifts had no effect (Sinvany-Villalobo et al. 2004). However, it is not known whether this increase is accompanied by a parallel increase in protein level.

2.5 Intramembrane proteases

Intramembrane proteolysis refers to a relatively recently discovered phenomenon, the cleavage of peptide bonds within trans-membranes helices. Such cleavage events are catalyzed by four groups of proteases: S2P, Presenilin, SPP, and

Rhomboid (for reviews, see Weihofen and Martoglio 2003; Wolfe and Kopan 2004). Two recent studies identified homologues of S2P in chloroplasts. A genetic screen for *Arabidopsis* mutants deficient in both chlorophyll accumulation and ethylene-induced gravitropism revealed EGY1, a 59-kDa membrane-bound metalloprotease that is located in the chloroplast (Chen et al. 2005). Although the intraorganellar location of EGY1 was not determined, mutant plants had reduced levels of granal stacks and light-harvesting complex (LHC) proteins, suggesting that this protease is required for proper chloroplast development. Another protease related to S2P, designated AraSP, was localized to the chloroplast inner-envelope membrane (Bolter et al. 2006). Antisense and T-DNA insertion lines of this protease demonstrated severely impaired chloroplast biogenesis. However, how these proteases are involved in chloroplast biogenesis is not clear. As for other intramembrane proteases, although homologous genes are found in plant genomes, and products of some of these are predicted to reside in chloroplasts, no reports on these have appeared yet.

3 Proteolytic processes in chloroplasts and the enzymes involved

3.1 Maturation of pre-proteins

Similar to all proteins synthesized in prokaryotic organisms, the 80 or so proteins synthesized within the chloroplast contain an N-formyl Met residue at their N terminus. Most of these proteins undergo maturation that involves two hydrolytic reactions: the N-formyl group is removed by peptide deformylase (PDF), and in most cases, this is followed by the activity of methionine aminopeptidase (MAP) (peptidase M24.001), which removes the N-terminal Met residue (Giglione and Meinnel 2001; Giglione et al. 2003). Thus, the activity of MAP has implications for the identity of the N-terminal residue of proteins encoded and synthesized in chloroplasts, and hence, might affect their stability through the N-end rule pathway (see Section 6).

Most chloroplast proteins are encoded in the nucleus and imported post-translationally into the organelle (see Chapter 11). A key feature in their targeting and import is their N-terminal transit peptide, which is cleaved off in the stroma during or shortly after import. This cleavage event is catalyzed by the stromal processing peptidase (SPP) (peptidase M16.004). SPP is a metalloprotease of ~140 kDa, containing the inverted zinc-binding motif HXXEH at its catalytic site (for a recent review, see Richter and Lamppa 2005). Recombinant SPP was shown to be able to cleave a wide range of chloroplast precursor proteins *in vitro* (Richter and Lamppa 1998), suggesting that SPP is the only enzyme responsible for this processing step. This contention was supported by *in vivo* studies in which expression of the single *Arabidopsis* SPP gene was downregulated by an antisense construct, resulting in lethal seedlings (Zhong et al. 2003).

After cleavage of the transit peptide, it remains bound to SPP. Its release from the enzyme is mediated by another cleavage event, carried out by SPP, which results in the release of two subfragments into the stroma. These subfragments are then further degraded by a soluble metalloprotease (Richter and Lamppa 1999). More recent work has suggested that degradation of the transit peptide is catalyzed by a 110-kDa metalloprotease designated pre-sequence protease (PreP) (peptidase M16.012; for a recent review, see Glaser et al. 2006). *Arabidopsis* contains two such genes, the products of which are dually targeted to both the chloroplasts and mitochondria, where they perform a similar function—degradation of the respective signal peptides (Bhushan et al. 2003). The crystal structure of this enzyme was recently determined, demonstrating that the active-site pocket is composed not only of the inverted zinc-binding motif HXXEH found in the N terminus, but also of more remote Arg and Tyr residues from the C terminus of the enzyme (Johnson et al. 2006). Identification and characterization of PreP as the protease-degrading signal peptide (Glaser et al. 2006) is consistent with the previous suggestion that a metalloprotease catalyzes this step (Richter and Lamppa 1999). However, since degradation of transit peptides by SPP and PreP has been studied mostly *in vitro*, it is not clear whether formation of two subfragments of the transit peptide by SPP is a prerequisite for further degradation by PreP. In any case, since PreP cleaves peptides in the range of 10 to 65 residues, but not shorter ones (Stahl et al. 2002, 2005), products of PreP activity must be further degraded by other as yet unidentified peptidases.

Nuclear-encoded proteins targeted to the thylakoid lumen are synthesized with a bipartite transit peptide. They are first processed by SPP in the stroma to yield an intermediate form, which is then translocated across the thylakoid membrane (see Chapter 11). This translocation step is followed by cleavage of the thylakoid-targeting sequence by the thylakoidal processing peptidase (TPP) (peptidase S26.008), a homologue of type I signal peptidase (for review, see Paetzel et al. 2002). TPP is an ~30-kDa membrane-bound serine protease that uses a Ser-Lys dyad for catalysis (Chaal et al. 1998). However, it is not known whether TPP further degrades the cleaved thylakoid-targeting sequence, or if complete degradation to free amino acids is catalyzed by other lumenal peptidases. Chloroplasts contain another homologue of this peptidase, designated plastidic SPase I (PlsP1) (Inoue et al. 2005). This peptidase was recently implicated in the maturation of Toc75, a component of the protein translocation machinery at the outer envelope membrane, which undergoes multiple maturation steps. Toc75 is first processed by SPP in the stroma, and then, after being targeted to the outer envelope membrane, cleaved again to yield the mature protein (Tranel and Keegstra 1996; Inoue and Keegstra 2003). This second step in the maturation of Toc75 is catalyzed by PlsP1 (Inoue et al. 2005). Interestingly, this peptidase has been localized to the envelope as well as to thylakoid membranes. Consistent with this dual localization was the observation that a PlsP1 knockout plant demonstrates accumulation of unprocessed forms of both Toc75 and the lumenal protein OE33 (Inoue et al. 2005).

Precursor processing is not limited to nuclear-encoded chloroplast proteins. The best characterized example of processing of a chloroplast-encoded protein is the D1 protein of photosystem II (PSII) reaction center. D1 is synthesized with a C-terminal extension of unknown function. However, it is well established that this extension needs to be removed before the oxygen-evolving complex can assemble together with the core PSII complex (e.g. Taguchi et al. 1995). This maturation step is carried out by the C-terminal protease-2 (CtpA) (peptidase S41.002). CtpA is a 45-kDa soluble serine protease, which uses a catalytic Ser-Lys dyad located in the lumen (Inagaki et al. 1996; Yamamoto et al. 2001). To date, there is no evidence for its involvement in any proteolytic process beyond maturation of the D1 protein. Another chloroplast-encoded protein that is processed in the thylakoid lumen is cytochrome *f*. This protein is synthesized with an N-terminal extension. The function of this extension and the peptidase involved are unknown.

3.2 Adaptation to changing light intensities

Although light is essential to plants, it may also have detrimental effects on them, a phenomenon known as ‘photoinhibition’ (Barber and Andersson 1992). Plants have evolved a number of strategies to avoid the harmful effects of light on the photosynthetic machinery, and one of them involves proteolysis. Long-term adaptation to an increase in light intensity is accompanied by a decrease in the antenna size of PSII, leading to a decrease in the amount of excitation energy being funneled to the reaction center. This modulation of antenna size is achieved by proteolytic degradation of a subset of LHCII subunits (Lindahl et al. 1995; Yang et al. 1998). Several biochemical attempts to identify the protease involved in this process have been unsuccessful, but a recent report on a specific *Arabidopsis* mutant suggested the involvement of the FtsH6 protease (Zelisko et al. 2005). However, low sensitivity of the degradation assay and high variability within and between experiments suggested that further experimentation was needed before a firm conclusion could be made. Another proposed candidate for this activity is SppA, a homologue of the *E. coli* SppA protease (peptidase S49.001), which functions as a signal-peptide peptidase in bacteria. SppA is a thylakoid-membrane-bound, light-induced serine protease—characteristics which are consistent with its speculated role in LHCII degradation (Lensch et al. 2001). Interestingly, this is the only thylakoid protease whose level increased in response to high light (Giacomelli et al. 2006). Nevertheless, direct experimental support for the involvement of SppA in this process is still lacking.

The transition from high to low light is also accompanied by protein degradation. The best known example in this context is the ‘early light-inducible protein’ (ELIP). This protein, which is structurally related to LHCs, is rapidly degraded upon such a transition (Adamska et al. 1993). Despite its initial characterization (Adamska et al. 1996), the involved protease has not yet been identified.

3.3 Protein quality control

Accumulation of all major photosynthetic complexes requires coordination between the chloroplast and nuclear genomes. Although advances have been made in recent years in understanding how these genomes communicate with each other (see Chapter by Bräutigam et al.), little is known about the mechanisms involved in regulating the correct stoichiometry between the different subunits of a given complex. In this context, it is assumed that fine-tuning of their levels is achieved by proteolytic degradation of super-stoichiometric subunits. First support for this assumption comes from a work published more than 20 years ago. Inhibition of protein synthesis in the chloroplast of *Chlamydomonas*, including that of the large subunit of Ribulose 1,5-bisphosphate carboxylase/oxygenase, resulted in degradation of the nuclear-encoded small subunit within the chloroplast (Schmidt and Mishkind 1983). These results suggested that unassembled subunits of a multiprotein complex are rapidly degraded. Indeed, similar observations have been made in other photosynthetic complexes as well. For instance, in *Chlamydomonas,* when cytochrome b_6, subunit IV and the Rieske protein of the cytochrome b_6-f complex cannot assemble with cytochrome *f,* they are rapidly degraded (Kuras and Wollman 1994). Similarly, a point mutation in the Rieske protein led to a significant decrease in its level, as well as the levels of other subunits of the cytochrome b_6-f complex. Crossing this mutant with one containing reduced levels of ClpP1 resulted in stabilization of these proteins, suggesting a role for Clp protease in the degradation of some unassembled proteins (Majeran et al. 2000).

In vitro studies have hinted at a role for Clp protease in the degradation of unassembled or abnormal proteins in the stroma as well. Mistargeting of the lumenal protein OE33 to the stroma resulted in its rapid degradation, with characteristics reminiscent of those of Clp protease (Halperin and Adam 1996; Halperin et al. 2001a). A similar function may be fulfilled by other proteases as well. Experiments with wild type or mutant forms of the Rieske protein have demonstrated that molecules that fail to translocate across the thylakoid membrane are rapidly degraded by a membrane-bound metalloprotease. This *in vitro* degradation reaction could be inhibited by increasing amounts of antibody against native FtsH (Ostersetzer and Adam 1997), suggesting that FtsH may be involved in protein quality control in the chloroplast as well.

Many mutants have been shown to contain lower levels of subunits in a complex when one other subunit is missing. This observation is often interpreted as degradation of the not-fully-assembled complex. However, such conclusions should be viewed with caution when no direct evidence for degradation is provided, for instance, through pulse-chase experiments. In some cases, translation of a complex's subunits is regulated by one component of the complex (a regulatory mechanism known as 'control by epistasy', see Minai et al. 2006 and references therein). Thus, lower levels of the subunits in a complex may result from reduced rates of translation in the chloroplast and not only from degradation of unassembled ones.

It has long been known that chloroplast proteins are unstable without their cofactors. For example, in the absence of chlorophyll, due to either inhibition of synthesis or mutation, chlorophyll-binding proteins are rapidly degraded (e.g. Kim et al. 1994 and references therein). Similarly, the lack of a single copper ion is sufficient to destabilize the electron carrier plastocyanin (Li and Merchant 1995). These observations suggest that minor structural changes, induced by a lack of minor components of a protein, may render it susceptible to proteolysis. Nevertheless, although the above examples have long been known, the proteases involved in degrading these substrates remain a mystery.

3.4 Oxidatively damaged proteins

Not all light energy absorbed by the photosynthetic antenna is converted to chemical energy. Depending on environmental conditions, free radicals are generated in chloroplasts, and despite the presence of free-radical scavengers, chloroplast proteins are highly prone to oxidation processes, which may impair their structure and function. The best characterized oxidatively damaged protein in the chloroplast is the D1 protein of the PSII reaction center. Oxidative damage leads to its inactivation, and hence, to photoinhibition (for reviews, see Andersson and Aro 2001; Yamamoto 2001). A prerequisite for the repair of photoinhibited PSII is degradation of the D1 protein, and numerous attempts have been made to identify the protease(s) involved. Biochemical approaches have been largely unsuccessful. However, the identification of chloroplast proteases and availability of protease mutants have enabled testing the possible involvement of specific proteases in D1 degradation.

Attempts to test the potential involvement of FtsH in the degradation of D1 were first made using recombinant GST-FtsH1 (Lindahl et al. 2000). These experiments demonstrated weak albeit significant activity of the recombinant enzyme against the initial 23-kDa degradation product of the D1 protein, but not against the full-length protein. The weak activity and limited specificity may result from the homomeric nature of the recombinant enzyme used *in vitro*, as opposed to the heteromeric nature of the enzyme found *in vivo*, as described in Section 2.2. Variegated mutants of FtsH2 and FtsH5 provided an opportunity to test the possible involvement of FtsH protease in the repair cycle of PSII *in vivo* as well. These mutants demonstrated increased sensitivity to photoinhibition compared to the wild type, as revealed by PSII-activity measurements (Bailey et al. 2002; Sakamoto et al. 2002). Consistent with this proposed role was the finding that the D1 protein is stabilized, probably in an inactive form, in FtsH mutant plants after exposing them to photoinhibitory illumination (Bailey et al. 2002).

Deg2, associated with the stromal side of the thylakoid membrane, has also been implicated in D1 degradation. An *in vitro* study demonstrated that recombinant Deg2 could cleave the D1 protein at its stromal loop connecting the fourth and fifth trans-membrane helices, yielding an N-terminal 23-kDa product and a C-terminal 10-kDa product, suggesting that this protease participates in the initial stages of D1 degradation (Haussuhl et al. 2001). However, a recent *in vivo* study with mutants lacking Deg2 demonstrated a rate of D1 degradation under light stress that was comparable to the wild type, suggesting that Deg2 is not essential for D1 degradation (Huesgen et al. 2006).

A recent study suggested that Deg1, located on the luminal side of the thylakoid membrane, might also be involved in D1 degradation (Kapri-Pardes et al. 2007). Transgenic plants expressing a RNAi construct accumulated less Deg1, were smaller than wild type and more sensitive to photoinhibition. These plants accumulated more of the D1 protein, probably in an inactive form, but less of 16- and 5.2-kD degradation products. Moreover, addition of recombinant Deg1 to inside-out thylakoid membranes could induce the formation of the 5.2-kD D1 C-terminal fragment *in vitro*. Taken together, these results suggest that Deg1 cooperate with proteases found on the stromal side of the membrane in the degradation of D1 protein during repair from photoinhibition (Kapri-Pardes et al. 2007).

A *Chlamydomonas* ATP synthase mutant has also been shown to lose PSII upon exposure to light. Crossing this mutant with a strain containing lower levels of ClpP resulted in stabilization of several PSII subunits, including the D2 protein, CP43 and CP47 (Majeran et al. 2001). It is not known whether this degradation process is identical to the one occurring in response to exposure to photoinhibitory conditions, or even whether the effect of Clp protease is direct or not. Nevertheless, these results suggest the involvement of soluble Clp protease in the degradation of membrane substrates as well.

4 Other functions

4.1 Nutrient stress and senescence

Nutrient stresses and senescence are both characterized by the need to remobilize internal cellular resources, some of which can be provided by the building blocks of existing proteins. Thus, massive protein degradation is expected to accompany the plants' attempts to deal with nutrient stress or their final developmental stage, senescence. However, only little is known about the involvement of specific proteases in these processes. Downregulation of ClpP1 in *Chlamydomonas* suggests involvement of the Clp complex in the degradation of thylakoid membrane proteins upon exposure to nutrient stress (Majeran et al. 2000). Nitrogen starvation results in degradation of subunits of the cytochrome b_6-f complex. However, in cells containing reduced levels of ClpP1, this degradation process is retarded, suggesting that Clp protease may be involved in the adaptation to nitrogen starvation via the degradation of existing abundant proteins.

Protein degradation in senescing leaves followed by nitrogen mobilization to younger ones is a well-documented phenomenon (Hortensteiner and Feller 2002). To date, one specific protease has been linked to the degradation of the most abundant protein in chloroplasts, Rubisco. CND41 (peptidase A01.050) is an aspartic protease of 41 kDa that is associated with chloroplast nucleoids and exhibits proteolytic activity against denatured Rubisco, among others (Murakami et al. 2000). Moreover, antisense plants demonstrated delayed senescence, along with stabilization of Rubisco as well as other chloroplast proteins (Kato et al. 2004). Interestingly, it was recently found that CND41 itself must undergo a protelytic processing step for its activation (Kato et al. 2005). The significance of CND41 binding to DNA is not yet known, but it could be a means of sequestering it from other chloroplast proteins. Degradation of plastid DNA during early stages of senescence may release CND41 to the stroma, allowing the initiation of massive protein degradation.

4.2 Thermotolerance

A recent study on a thermosensitive *Arabidopsis* mutant suggests the involvement of FtsH11 in thermotolerance (Chen et al. 2006). This mutant was more sensitive to moderate high temperature than the wild type, whereas the FtsH2 and FtsH5 mutants were not. As a result, the FtsH11 mutant had lower photosynthetic capability than the wild type after exposure to 30°C. Unlike the FtsH2 and FtsH5 mutants, the sensitivity of the FtsH11 mutant to high light was similar to that observed in wild type plants (Chen et al. 2006). These results suggest that the physiological functions of the FtsH2-FtsH5-FtsH8-FtsH1 complex and FtsH11 differ. In this context, it should be noted that FtsH11 is found in both chloroplasts and mitochondria (Urantowka et al. 2005), and it is impossible to conclude at this stage whether its contribution to thermotolerance is related to its activity in chloroplasts, mitochondria, or both.

5 Identification of specific substrates

The availability of specific protease mutants lends itself to the identification of their substrates, when these are unknown. Specific substrates of a protease are expected to be stabilized in a mutant background, and thus comparative proteomics has the potential to yield their unbiased identification. One successful utilization of this approach involved a comparison between stromal proteins in the wild type and ClpP6 mutant (Sjogren et al. 2006). Potential substrates of Clp protease found in this work included a nuclear exchange factor for the elongation factor Tu, the molecular chaperone HSP90, an RNA helicase, the folding catalyst PPIase, and the UPRT and NDP kinase proteins involved in nucleic acid synthesis. These results suggest that Clp substrates are more involved in chloroplast homeostasis than in metabolism (Sjogren et al. 2006). A similar approach should prove useful in the identification of substrates of other proteases as well.

6 Determinants of protein instability

Although progress has been made in identifying components of the chloroplast proteolytic machinery, and proteolytic processes have been documented, determinants of instability within the protein substrates themselves are still obscure. The N-end rule, discovered and characterized in eukaryotic cells, relates the half-life of a protein to the identity of its N-terminal residue. Proteins carrying a destabilizing residue at their N terminus are ubiquitinated and degraded by the 26S proteasome (Varshavsky 1992). The N-end rule was shown to operate in *E. coli* as well, where degradation of substrates is mediated by the ClpAP protease (Tobias et al. 1991). The *E. coli* ClpS adaptor protein has been described as a modulator of ClpAP activity (Dougan et al. 2002), and more recently, shown to be essential for the operation of the N-end rule pathway in bacteria (Erbse et al. 2006). As plastids are descendants of a prokaryotic progenitor and many of their characteristics, including their proteolytic machinery, are prokaryote-like, it is highly likely that an N-end rule-like mechanism governs protein stability in plastids as well. As described in section 2.1, the plastid homologue of the bacterial ClpAP is designated ClpC. Plants also have a homologue of ClpS, known as ClpT (Peltier et al. 2004). However, ClpT was not identified in proteomic studies of the Clp core complex (Adam et al. 2006), and thus is not expected to be a component of the Clp protease core. Nevertheless, the presence of chloroplast homologues to components of the bacterial N-end rule pathway suggests that this pathway governs protein stability/instability in this organelle as well, a notion that awaits experimental validation.

The small stable RNA A (SsrA) system in *E. coli* tags proteins translated from incomplete mRNAs for degradation (Karzai et al. 2000). The ssrA RNA is a small molecule that acts as both tRNA and mRNA. When a ribosome stalls on an incomplete mRNA, the ssrA molecule binds the ribosome, which then reads through to add an 11-amino-acid long tag to the protein (a process referred to as 'trans-translation'). This tag contains the small nonpolar amino acid sequence Leu-Ala-Ala at its C terminus, which is recognized in the cytoplasm by the ClpAP, ClpXP or FtsH proteases, or in the periplasm by the DegP protease. Related sequences that can also be recognized by proteases are Leu-Val-Ala, Ala-Ser-Val and Ala-Ala-Val. To date, there is no evidence for the presence of ssrA RNA in plastids of higher plants (Gueneau de Novoa and Williams 2004). However, the presence of homologues of the bacterial proteases suggests that plastid proteins with C termini homologous to the SsrA tag would be short-lived. Moreover, even in the absence of an SsrA trans-translation system, the identity of C-terminal residues may confer stability or instability on a protein. However, as with the influence of the N terminus, this has not been explored to date.

7 Future prospects

Plant sequence data accumulated over the past 15 years, and completion of the *Arabidopsis* and rice genome sequencing projects, suggest that the identity of most, if not all chloroplast proteases and peptidases is now known. The major challenge ahead is to assign a function to each of them, and relate the proteases and peptidases to known proteolytic processes. A striking feature of many of the chloroplast proteases is the relatively large number of genes encoding them, compared with their prokaryotic progenitors. It is now clear that at least the Clp and FtsH proteases are heteromeric complexes, with little redundancy between their components. Assuming that ClpRs are indeed proteolytically inactive, it will be important to establish whether they have only a structural role, or perhaps other functions, such as substrate binding or recognition. The apparent need for more than just one ClpR is also intriguing. Overexpression of different ClpRs in a specific ClpR mutant is expected to help resolve the question of their redundancy. A similar approach will be applicable to the study of the relations between the different ClpPs as well. Some of these questions are partially answered with respect to the FtsH protease. However, it is still important to understand why the FtsH complex requires two types of subunits for its accumulation. Relations between the three luminal Deg proteases also need to be sorted out. Do they form homo- or hetero-oligomeric complexes? Are they redundant or not?

Major insight into the functions of different chloroplast proteases has been gained using specific mutants. There are now a number of publicly available mutant collections, particularly T-DNA insertion lines, and well-established procedures, such as antisense and RNA interference, for downregulating the expression of specific genes or gene families. Thus, these will probably continue to serve as the main tools in deciphering the physiological functions of specific proteases. *In vitro* approaches will probably continue to complement these efforts. However, special attention should be paid to possible pleiotropic effects. Many mutants missing different chloroplast proteins demonstrate a similar phenotype: slow growth, reduced pigmentation, altered chloroplast morphology and reduced levels of thylakoids. Thus, efforts should be made to distinguish between these general effects and the specific function of a given protease leading to these effects. This requires more specific assays for specific proteolytic processes, better linkage to substrate proteins, and attempts to understand their involvement in a given physiological response.

Identification of specific substrates for each of the proteases will have to be accompanied by attempts to reveal recognition determinants. To date, an understanding of recognition mechanisms between chloroplast proteases and their substrates is almost totally lacking. This applies to both partners—the proteases and their substrates. Efforts will need to be made to identify subunits within a proteolytic complex, or domains within a given protease, that are responsible for substrate recognition and binding, and determinants on the substrates themselves that allow this recognition. All of these questions will keep the growing community of scientists interested in chloroplast proteases busy for years to come.

Acknowledgement

Work in the author's lab was supported by grants from The Israel Science Foundation, The U.S.-Israel Binational Science Foundation, The U.S.-Israel Binational Agricultural Research and Development Fund, and The Sixth Framework Program of the E.U.

References

Adam Z, Adamska I, Nakabayashi K, Ostersetzer O, Haussuhl K, Manuell A, Vallon O, Rodermel SR, Shinozaki K, Clarke AK (2001) Chloroplast and mitochondrial proteases in *Arabidopsis*. A proposed nomenclature. Plant Physiol 125:1912-1918

Adam Z, Zaltsman A, Sinvany-Villalobo G, Sakamoto W (2005) FtsH proteases in chloroplasts and cyanobacteria. Physiol Plant 123:386-390

Adam Z, Rudella A, van Wijk KJ (2006) Recent advances in the study of Clp, FtsH, and other proteases located in chloroplasts. Curr Opin Plant Biol 9:234-240

Adamska I, Kloppstech K, Ohad I (1993) Early light-inducible protein in pea is stable during light stress but is degraded during recovery at low light intensity. J Biol Chem 268:5438-5444

Adamska I, Lindahl M, Roobol-Boza M, Andersson B (1996) Degradation of the light-stress protein is mediated by an ATP-independent, serine-type protease under low-light conditions. Eur J Biochem 236:591-599

Andersson B, Aro E-M (2001) Photodamage and D1 protein turnover in photosystem II. In: Andersson B, Aro E-M (eds) Regulation of photosynthesis. Dordrecht/Boston/London: Kluwer Academic Publishers, pp 377-393

Bailey S, Thompson E, Nixon PJ, Horton P, Mullineaux CW, Robinson C, Mann NH (2002) A critical role for the Var2 FtsH homologue of *Arabidopsis thaliana* in the photosystem II repair cycle *in vivo*. J Biol Chem 277:2006-2011

Barakat S, Pearce DA, Sherman F, Rapp WD (1998) Maize contains a Lon protease gene that can partially complement a yeast pim1-deletion mutant. Plant Mol Biol 37:141-154

Barber J, Andersson B (1992) Too much of a good thing: light can be bad for photosynthesis. Trends Biochem Sci 17:61-66

Barrett AJ, Rawlings ND, Woessner JF (2004) Handbook of proteolytic enzymes, 2nd edn. Amsterdam: Elsevier Ltd

Baumeister W, Walz J, Zuhl F, Seemuller E (1998) The proteasome: paradigm of a self-compartmentalizing protease. Cell 92:367-380

Bhushan S, Lefebvre B, Stahl A, Wright SJ, Bruce BD, Boutry M, Glaser E (2003) Dual targeting and function of a protease in mitochondria and chloroplasts. EMBO Rep 4:1073-1078

Bieniossek C, Schalch T, Bumann M, Meister M, Meier R, Baumann U (2006) The molecular architecture of the metalloprotease FtsH. Proc Natl Acad Sci USA 103:3066-3071

Bolter B, Nada A, Fulgosi H, Soll J (2006) A chloroplastic inner envelope membrane protease is essential for plant development. FEBS Lett 580:789-794

Botos I, Melnikov EE, Cherry S, Khalatova AG, Rasulova FS, Tropea JE, Maurizi MR, Rotanova TV, Gustchina A, Wlodawer A (2004) Crystal structure of the AAA+ alpha domain of *E. coli* Lon protease at 1.9A resolution. J Struct Biol 146:113-122

Chaal BK, Mould RM, Barbrook AC, Gray JC, Howe CJ (1998) Characterization of a cDNA encoding the thylakoidal processing peptidase from *Arabidopsis thaliana*. Implications for the origin and catalytic mechanism of the enzyme. J Biol Chem 273:689-692

Chassin Y, Kapri-Pardes E, Sinvany G, Arad T, Adam Z (2002) Expression and characterization of the thylakoid lumen protease DegP1 from *Arabidopsis thaliana*. Plant Physiol 130:857-864

Chen G, Bi YR, Li N (2005) EGY1 encodes a membrane-associated and ATP-independent metalloprotease that is required for chloroplast development. Plant J 41:364-375

Chen J, Burke JJ, Velten J, Xin Z (2006) FtsH11 protease plays a critical role in *Arabidopsis* thermotolerance. Plant J 48:73-84

Chen M, Choi Y, Voytas DF, Rodermel S (2000) Mutations in the *Arabidopsis* VAR2 locus cause leaf variegation due to the loss of a chloroplast FtsH protease. Plant J 22:303-313

Clarke AK, MacDonald TM, Sjogren LLE (2005) The ATP-dependent Clp protease in chloroplasts of higher plants. Physiol Plant 123:406-412

Clausen T, Southan C, Ehrmann M (2002) The HtrA family of proteases: implications for protein composition and cell fate. Mol Cell 10:443-455

Constan D, Froehlich JE, Rangarajan S, Keegstra K (2004) A stromal Hsp100 protein is required for normal chloroplast development and function in *Arabidopsis*. Plant Physiol 136:3605-3615

Dougan DA, Reid BG, Horwich AL, Bukau B (2002) ClpS, a substrate modulator of the ClpAP machine. Mol Cell 9:673-683

Erbse A, Schmidt R, Bornemann T, Schneider-Mergener J, Mogk A, Zahn R, Dougan DA, Bukau B (2006) ClpS is an essential component of the N-end rule pathway in *Escherichia coli*. Nature 439:753-756

Friso G, Giacomelli L, Ytterberg AJ, Peltier JB, Rudella A, Sun Q, van Wijk KJ (2004) In-depth analysis of the thylakoid membrane proteome of *Arabidopsis thaliana* chloroplasts: new proteins, new functions, and a plastid proteome database. Plant Cell 16:478-499

Fukui T, Eguchi T, Atomi H, Imanaka T (2002) A membrane-bound archaeal Lon protease displays ATP-independent proteolytic activity towards unfolded proteins and ATP-dependent activity for folded proteins. J Bacteriol 184:3689-3698

Giacomelli L, Rudella A, van Wijk KJ (2006) High light response of the thylakoid proteome in *Arabidopsis* wild type and the ascorbate-deficient mutant vtc2-2. A comparative proteomics study. Plant Physiol 141:685-701

Giglione C, Meinnel T (2001) Organellar peptide deformylases: universality of the N-terminal methionine cleavage mechanism. Trends Plant Sci 6:566-572

Giglione C, Vallon O, Meinnel T (2003) Control of protein life-span by N-terminal methionine excision. EMBO J 22:13-23

Glaser E, Nilsson S, Bhushan S (2006) Two novel mitochondrial and chloroplastic targeting-peptide-degrading peptidasomes in *A. thaliana*, AtPreP1 and AtPreP2. Biol Chem 387:1441-1447

Gottesman S (1996) Proteases and their targets in *Escherichia coli*. Annu Rev Genet 30:465-506

Gueneau de Novoa P, Williams KP (2004) The tmRNA website: reductive evolution of tmRNA in plastids and other endosymbionts. Nucl Acids Res 32:D104-108

Halperin T, Adam Z (1996) Degradation of mistargeted OEE33 in the chloroplast stroma. Plant Mol Biol 30:925-933

Halperin T, Ostersetzer O, Adam Z (2001a) ATP-dependent association between subunits of Clp protease in pea chloroplasts. Planta 213:614-619

Halperin T, Zheng B, Itzhaki H, Clarke AK, Adam Z (2001b) Plant mitochondria contain proteolytic and regulatory subunits of the ATP-dependent Clp protease. Plant Mol Biol 45:461-468

Haussuhl K, Andersson B, Adamska I (2001) A chloroplast DegP2 protease performs the primary cleavage of the photodamaged D1 protein in plant photosystem II. EMBO J 20:713-722

Heazlewood JL, Tonti-Filippini JS, Gout AM, Day DA, Whelan J, Millar AH (2004) Experimental analysis of the *Arabidopsis* mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. Plant Cell 16:241-256

Heazlewood JL, Millar AH (2005) AMPDB: the *Arabidopsis* mitochondrial protein database. Nuc Acids Res 33:D605-610

Hortensteiner S, Feller U (2002) Nitrogen metabolism and remobilization during senescence. J Exp Bot 53:927-937

Huesgen PH, Schuhmann H, Adamska I (2005) The family of Deg proteases in cyanobacteria and chloroplasts of higher plants. Physiol Plant 123:413-420

Huesgen P, Schumann H, Adamska I (2006) Photodamaged D1 protein is degraded in *Arabidopsis* mutants lacking the Deg2 protease. FEBS Lett 580:6929-6932

Inagaki N, Yamamoto Y, Mori H, Satoh K (1996) Carboxyl-terminal processing protease for the D1 precursor protein: cloning and sequencing of the spinach cDNA. Plant Mol Biol 30:39-50

Inoue K, Keegstra K (2003) A polyglycine stretch is necessary for proper targeting of the protein translocation channel precursor to the outer envelope membrane of chloroplasts. Plant J 34:661-669

Inoue K, Baldwin AJ, Shipman RL, Matsui K, Theg SM, Ohme-Takagi M (2005) Complete maturation of the plastid protein translocation channel requires a type I signal peptidase. J Cell Biol 171:425-430

Ito K, Akiyama Y (2005) Cellular functions, mechanism of action, and regulation of FtsH protease. Annu Rev Microbiol 59:211-231

Itzhaki H, Naveh L, Lindahl M, Cook M, Adam Z (1998) Identification and characterization of DegP, a serine protease associated with the luminal side of the thylakoid membrane. J Biol Chem 273:7094-7098

Johnson KA, Bhushan S, Stahl A, Hallberg BM, Frohn A, Glaser E, Eneqvist T (2006) The closed structure of presequence protease PreP forms a unique 10,000 Angstroms3 chamber for proteolysis. EMBO J 25:1977-1986

Kapri-Pardes E, Naveh L, Adam Z (2007) The thylakoid lumen protease Deg1 is involved in the repair of photosystem II from photoinhibition in *Arabidospsis*. Plant Cell 19:1039-1047

Karzai AW, Roche ED, Sauer RT (2000) The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue. Nat Struct Biol 7:449-455

Kato Y, Murakami S, Yamamoto Y, Chatani H, Kondo Y, Nakano T, Yokota A, Sato F (2004) The DNA-binding protease, CND41, and the degradation of ribulose-1,5-bisphosphate carboxylase/oxygenase in senescent leaves of tobacco. Planta 220:97-104

Kato Y, Yamamoto Y, Murakami S, Sato F (2005) Post-translational regulation of CND41 protease activity in senescent tobacco leaves. Planta 222:643-651

Kim J, Klein PG, Mullet JE (1994) Vir-115 gene product is required to stabilize D1 translation intermediates in chloroplasts. Plant Mol Biol 25:459-467

Koussevitzky S, Stanne TM, Peto CA, Giap T, Sjogren LL, Zhao Y, Clarke AK, Chory J (2007) An *Arabidopsis thaliana* virescent mutant reveals a role for ClpR1 in plastid development. Plant Mol Biol 63:85-96

Kovacheva S, Bedard J, Patel R, Dudley P, Twell D, Rios G, Koncz C, Jarvis P (2005) *In vivo* studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. Plant J 41:412-428

Krojer T, Garrido-Franco M, Huber R, Ehrmann M, Clausen T (2002) Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. Nature 416:455-459

Kuras R, Wollman FA (1994) The assembly of cytochrome b6/f complexes: an approach using genetic transformation of the green alga *Chlamydomonas reinhardtii*. EMBO J 13:1019-1027

Kuroda H, Maliga P (2003) The plastid clpP1 protease gene is essential for plant development. Nature 425:86-89

Lensch M, Herrmann RG, Sokolenko A (2001) Identification and characterization of SppA, a novel light-inducible chloroplast protease complex associated with thylakoid membranes. J Biol Chem 276:33645-33651

Li HH, Merchant S (1995) Degradation of plastocyanin in copper-deficient *C. reinhardtii*—evidence for a protease-susceptible conformation of the apoprotein and regulated proteolysis. J Biol Chem 270:23504-23510

Lindahl M, Yang DH, Andersson B (1995) Regulatory proteolysis of the major light-harvesting chlorophyll a/b protein of photosystem II by a light-induced membrane-associated enzymic system. Eur J Biochem 231:503-509

Lindahl M, Tabak S, Cseke L, Pichersky E, Andersson B, Adam Z (1996) Identification, characterization, and molecular cloning of a homologue of the bacterial FtsH protease in chloroplasts of higher plants. J Biol Chem 271:29329-29334

Lindahl M, Spetea C, Hundal T, Oppenheim AB, Adam Z, Andersson B (2000) The thylakoid FtsH protease plays a role in the light-induced turnover of the photosystem II D1 protein. Plant Cell 12:419-431

Majeran W, Wollman F-A, Vallon O (2000) Evidence for a role of ClpP in the degradation of the chloroplast cytochrome b6f complex. Plant Cell 12:137-149

Majeran W, Olive J, Drapier D, Vallon O, Wollman FA (2001) The light sensitivity of ATP synthase mutants of *Chlamydomonas reinhardtii*. Plant Physiol 126:421-433

Minai L, Wostrikoff K, Wollman FA, Choquet Y (2006) Chloroplast biogenesis of photosystem II cores involves a series of assembly-controlled steps that regulate translation. Plant Cell 18:159-175

Murakami S, Kondo Y, Nakano T, Sato F (2000) Protease activity of CND41, a chloroplast nucleoid DNA-binding protein, isolated from cultured tobacco cells. FEBS Lett 468:15-18

Neuwald AF, Aravind L, Spouge JL, Koonin EV (1999) AAA+: a class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome Res 9:27-43

Ostersetzer O, Adam Z (1997) Light-stimulated degradation of an unassembled Rieske FeS protein by a thylakoid-bound protease: the possible role of the FtsH protease. Plant Cell 9:957-965

Ostersetzer O, Kato Y, Adam Z, Sakamoto W (2007) Multiple intracellular locations of Lon protease in Arabidopsis: Evidence for the localization of AtLon4 to chloroplasts. Plant Cell Physiol (in press)

Paetzel M, Karla A, Strynadka NC, Dalbey RE (2002) Signal peptidases. Chem Rev 102:4549-4580

Park S, Rodermel SR (2004) Mutations in ClpC2/Hsp100 suppress the requirement for FtsH in thylakoid membrane biogenesis. Proc Natl Acad Sci USA 101:12765-12770

Peltier J-B, Ytterberg J, Liberles DA, Roepstorff P, van Wijk KJ (2001) Identification of a 350 kDa ClpP protease complex with 10 different Clp isoforms in chloroplasts of *Arabidopsis thaliana*. J Biol Chem 276:16318-16327

Peltier J-B, Emanuelsson O, Kalume DE, Ytterberg J, Friso G, Rudella A, Liberles DA, Soderberg L, Roepstorff P, von Heijne G, van Wijk KJ (2002) Central functions of the lumenal and peripheral thylakoid proteome of *Arabidopsis* determined by experimentation and genome-wide prediction. Plant Cell 14:211-236

Peltier JB, Ripoll DR, Friso G, Rudella A, Cai Y, Ytterberg J, Giacomelli L, Pillardy J, van Wijk KJ (2004) Clp protease complexes from photosynthetic and non-photosynthetic plastids and mitochondria of plants, their predicted 3-D structures and functional implications. J Biol Chem 279:4768-4781

Rawlings ND, Morton FR, Barrett AJ (2006) MEROPS: the peptidase database. Nucl Acids Res 34:D270-D272

Richter S, Lamppa GK (1998) A chloroplast processing enzyme functions as the general stromal processing peptidase. Proc Natl Acad Sci USA 95:7463-7468

Richter S, Lamppa GK (1999) Stromal processing peptidase binds transit peptides and initiates their ATP-dependent turnover in chloroplasts. J Cell Biol 147:33-43

Richter S, Lamppa GK (2005) Function of the stromal processing peptidase in the chloroplast import pathway. Physiol Plant 123:362-368

Rudella A, Friso G, Alonso JM, Ecker JR, van Wijk KJ (2006) Downregulation of ClpR2 leads to reduced accumulation of the ClpPRS protease complex and defects in chloroplast biogenesis in *Arabidopsis*. Plant Cell 18:1704-1721

Sakamoto W (2006) Protein degradation machineries in plastids. Annu Rev Plant Biol 57:599-621

Sakamoto W, Tamura T, Hanba-Tomita Y, Sodmergen, Murata M (2002) The VAR1 locus of *Arabidopsis* encodes a chloroplastic FtsH and is responsible for leaf variegation in the mutant alleles. Genes Cells 7:769-780

Sakamoto W, Zaltsman A, Adam Z, Takahashi Y (2003) Coordinated regulation and complex formation of YELLOW VARIEGATED1 and YELLOW VARIEGATED2, chloroplastic FtsH metalloproteases involved in the repair cycle of photosystem II in *Arabidopsis* thylakoid membranes. Plant Cell 15:2843-2855

Sarria R, Lyznik A, Vallejos CE, Mackenzie SA (1998) A cytoplasmic male sterility-associated mitochondrial peptide in common bean is post-translationally regulated. Plant Cell 10:1217-1228

Sauer RT, Bolon DN, Burton BM, Burton RE, Flynn JM, Grant RA, Hersch GL, Joshi SA, Kenniston JA, Levchenko I, Neher SB, Oakes ES, Siddiqui SM, Wah DA, Baker TA (2004) Sculpting the proteome with AAA(+) proteases and disassembly machines. Cell 119:9-18

Schmidt GW, Mishkind ML (1983) Rapid degradation of unassembled ribulose 1,5-biphosphate carboxylase small subunit in chloroplasts. Proc Natl Acad Sci USA 80:2632-2636

Schubert M, Petersson UA, Haas BJ, Funk C, Schroder WP, Kieselbach T (2002) Proteome map of the chloroplast lumen of *Arabidopsis thaliana*. J Biol Chem 277:8354-8365

Shikanai T, Shimizu K, Ueda K, Nishimura Y, Kuroiwa T, Hashimoto T (2001) The chloroplast clpP gene, encoding a proteolytic subunit of ATP-dependent protease, is indispensable for chloroplast development in tobacco. Plant Cell Physiol 42:264-273

Sinvany-Villalobo G, Davydov O, Ben-Ari G, Zaltsman A, Raskind A, Adam Z (2004) Expression in multigene families. Analysis of chloroplast and mitochondrial proteases. Plant Physiol 135:1336-1345

Sjogren LL, MacDonald TM, Sutinen S, Clarke AK (2004) Inactivation of the clpC1 gene encoding a chloroplast Hsp100 molecular chaperone causes growth retardation, leaf chlorosis, lower photosynthetic activity, and a specific reduction in photosystem content. Plant Physiol 136:4114-4126

Sjogren LL, Stanne TM, Zheng B, Sutinen S, Clarke AK (2006) Structural and functional insights into the chloroplast ATP-dependent Clp protease in *Arabidopsis*. Plant Cell 18:2635-2649

Sokolenko A, Pojidaeva E, Zinchenko V, Panichkin V, Glaser VM, Herrmann RG, Shestakov SV (2002) The gene complement for proteolysis in the cyanobacterium *Synechocystis* sp. PCC 6803 and *Arabidopsis thaliana* chloroplasts. Curr Genet 41:291-310

Spiess C, Beil A, Ehrmann M (1999) A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. Cell 97:339-347

Stahl A, Moberg P, Ytterberg J, Panfilov O, Brockenhuus Von Lowenhielm H, Nilsson F, Glaser E (2002) Isolation and identification of a novel mitochondrial metalloprotease (PreP) that degrades targeting presequences in plants. J Biol Chem 277:41931-41939

Stahl A, Nilsson S, Lundberg P, Bhushan S, Biverstahl H, Moberg P, Morisset M, Vener A, Maler L, Langel U, Glaser E (2005) Two novel targeting peptide degrading proteases, PrePs, in mitochondria and chloroplasts, so similar and still different. J Mol Biol 349:847-860

Suno R, Niwa H, Tsuchiya D, Zhang X, Yoshida M, Morikawa K (2006) Structure of the whole cytosolic region of ATP-dependent protease FtsH. Mol Cell 22:575-585

Taguchi F, Yamamoto Y, Satoh K (1995) Recognition of the structure around the site of cleavage by the carboxyl-terminal processing protease for D1 precursor protein of the photosystem II reaction center. J Biol Chem 270:10711-10716

Takechi K, Sodmergen, Murata M, Motoyoshi F, Sakamoto W (2000) The YELLOW VARIEGATED (VAR2) locus encodes a homologue of FtsH, an ATP- dependent protease in *Arabidopsis*. Plant Cell Physiol 41:1334-1346

Tobias JW, Shrader TE, Rocap G, Varshavsky A (1991) The N-end rule in bacteria. Science 254:1374-1377

Tranel PJ, Keegstra K (1996) A novel, bipartite transit peptide targets OEP75 to the outer membrane of the chloroplastic envelope. Plant Cell 8:2093-2104

Tsugane K, Maekawa M, Takagi K, Takahara H, Qian Q, Eun CH, Iida S (2006) An active DNA transposon nDart causing leaf variegation and mutable dwarfism and its related elements in rice. Plant J 45:46-57

Urantowka A, Knorpp C, Olczak T, Kolodziejczak M, Janska H (2005) Plant mitochondria contain at least two i-AAA-like complexes. Plant Mol Biol 59:239-252

Varshavsky A (1992) The N-end rule. Cell 69:725-735

Weihofen A, Martoglio B (2003) Intramembrane-cleaving proteases: controlled liberation of proteins and bioactive peptides. Trends Cell Biol 13:71-78

Wolfe MS, Kopan R (2004) Intramembrane proteolysis: theme and variations. Science 305:1119-1123

Yamamoto Y (2001) Quality control of photosystem II. Plant Cell Physiol 42:121-128

Yamamoto Y, Inagaki N, Satoh K (2001) Overexpression and characterization of carboxyl-terminal processing protease for precursor D1 protein: regulation of enzyme-substrate interaction by molecular environments. J Biol Chem 276:7518-7525

Yang DH, Webster J, Adam Z, Lindahl M, Andersson B (1998) Induction of acclimative proteolysis of the light-harvesting chlorophyll a/b protein of photosystem II in response to elevated light intensities. Plant Physiol 118:827-834

Yu F, Park S, Rodermel SR (2004) The *Arabidopsis* FtsH metalloprotease gene family: interchangeability of subunits in chloroplast oligomeric complexes. Plant J 37:864-876

Yu F, Park S, Rodermel SR (2005) Functional redundancy of AtFtsH metalloproteases in thylakoid membrane complexes. Plant Physiol 138:1957-1966

Zaltsman A, Feder A, Adam Z (2005a) Developmental and light effects on the accumulation of FtsH protease in *Arabidopsis* chloroplasts—implications for thylakoid formation and photosystem II maintenance. Plant J 42:609-617

Zaltsman A, Ori N, Adam Z (2005b) Two types of FtsH protease subunits are required for chloroplast biogenesis and Photosystem II repair in *Arabidopsis*. Plant Cell 17:2782-2790

Zelisko A, Garcia-Lorenzo M, Jackowski G, Jansson S, Funk C (2005) AtFtsH6 is involved in the degradation of the light-harvesting complex II during high-light acclimation and senescence. Proc Natl Acad Sci USA 102:13699-13704

Zheng B, Halperin T, Hruskova-Heidingsfeldova O, Adam Z, Clarke AK (2002) Characterization of chloroplast Clp proteins in *Arabidopsis*: localization, tissue specificity and stress responses. Physiol Plant 114:92-101

Zheng B, Macdonald TM, Sutinen S, Hurry V, Clarke AK (2006) A nuclear-encoded ClpP subunit of the chloroplast ATP-dependent Clp protease is essential for early development in *Arabidopsis thaliana*. Planta 224:1103-1115

Zhong R, Wan J, Jin R, Lamppa G (2003) A pea antisense gene for the chloroplast stromal processing peptidase yields seedling lethals in *Arabidopsis*: survivors show defective GFP import *in vivo*. Plant J 34:802-812

Adam, Zach

The Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture, The Hebrew University, Rehovot 76100, Israel.
zach@agri.huji.ac.il

Protein import into plastids

Birgit Agne and Felix Kessler

Abstract

Most chloroplast proteins are synthesized as preproteins with N-terminal transit sequences and imported from the cytosol. Protein translocons at the outer (Toc) and inner membranes of the chloroplast (Tic) recognize the presequence and enable the transfer of the polypeptide across the dual membrane envelope. After the initial characterization of the translocon components, research focused on the import mechanisms and their regulation. Recent research taking advantage of the *Arabidopsis* system has demonstrated the existence of substrate specific import sub-pathways in the general import pathway. New discoveries indicate that glycosylated proteins may take an entirely different pathway via the endoplasmic reticulum and the Golgi apparatus. This review will discuss the known import components in the light of the exciting new discoveries.

1 Plastids

Plants are characterized by a family of double-membrane bound organelles called plastids. Plastids and mitochondria originate from endosymbiotic events. In the case of the plastid, an ancient photosynthetic cyanobacterium was engulfed by a eukaryotic host (McFadden 2001; Kutschera and Niklas 2005). During evolution the plastid retained its own genetic system but most of its genes were transferred to the host cell nucleus (Martin et al. 2002; Leister and Schneider 2003; Timmis et al. 2004). Moreover, the plastid evolved into a family of remarkably versatile organelles (Lopez-Juez and Pyke 2005). Functionally specialized plastid types having varying morphologies and physiological properties are controlled by the host tissue. All plastid types derive from an undifferentiated plastid called proplastid which is present in meristematic cells. Differentiation is achieved by the import of functionally specific sets of protein. The development and interconversion of plastids may also be influenced by environmental cues such as light in the case of chloroplasts.

The chloroplast, present in green aerial tissues, constitutes the site of photosynthesis and metabolic functions such as fatty acid and amino acid biosynthesis or nitrite and sulphate reduction. Chloroplasts contain a unique internal membrane system, the thylakoids harbouring the photosynthetic machinery.

Chloroplasts may differentiate into chromoplasts colouring fruits and flowers due to their accumulation of carotenoid compounds. Finally, amyloplasts, elaio-

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0237 / Published online: 4 July 2007

plasts and leucoplasts are unpigmented plastid types present in non-green tissues and generally specialize in the storage of compounds such as starch, lipid, and protein.

1.1 Plastid biogenesis

The biogenesis of the different plastid types is linked to the differentiation of the host cell and is evident in plastid type specific proteomes (Kleffmann et al. 2004; van Wijk 2004; Siddique et al. 2006). Plastids being semiautonomous are largely under genetic control by the host cell and most plastid proteins are encoded in the nucleus. Therefore plastids import proteins that are synthesized in the cytosol. This requires the existence of mechanisms that reliably, specifically and efficiently target and translocate proteins into plastids. In this review, we will give an overview of the main import pathway across the two envelope membranes of chloroplasts as well as a short discussion of alternative import pathways.

2 Chloroplast targeting signals

Most chloroplast targeted proteins are synthesized in the cytosol as precursor proteins with a cleavable, N-terminal targeting signal termed transit sequence (transit peptide). The transit sequence is recognized by chloroplast import receptors and enables the passage of the precursor protein through the import complexes at the outer and inner envelope membrane (Bruce 2000, 2001). Upon translocation, the transit peptide is removed by a stromal processing peptidase yielding the mature protein (Richter and Lamppa 2003). Proteins destined for the thylakoid membrane system of chloroplasts are often synthesized with bipartite signals consisting of an N-terminal transit peptide, for import into the chloroplast stroma, followed by an additional targeting sequence specifying either insertion into the thylakoid membrane or translocation into the thylakoid lumen. At least four pathways contribute to thylakoid targeting and are conserved from the prokaryotic ancestor of the chloroplast: 1) The SRP-pathway facilitates insertion of hydrophobic proteins into the thylakoid membrane. 2) The Sec-pathway promotes translocation into the thylakoid lumen. 3) The TAT-pathway allows for translocation of folded proteins associated with cofactors into the thylakoid lumen. 4) Finally, some proteins may integrate into the thylakoid membrane without assistance by other proteins or energy consumption via the spontaneous insertion pathway. The detailed description of thylakoid translocation and membrane insertion exceeds the scope of this paper and we recommend the lecture of one of the excellent reviews on the subject (Jarvis and Robinson 2004; Gutensohn et al. 2006).

2.1 Structure of transit peptides

Transit peptides typically have 20 to > 70 amino acids. They are rich in hydrophobic and hydroxylated residues and have few acidic amino acids resulting in a net positive charge. The entire transit peptide seems to be necessary for correct targeting (Bhushan et al. 2006). No conserved sequence motifs or secondary structure elements have been identified complicating the definition of common features for chloroplast targeting. In aqueous environments transit peptides have been shown to be largely unstructured and form a random coil (Bruce 2001). However transit peptides may undergo secondary structural changes upon interaction with lipids (Horniak et al. 1993) or with receptor components (Bedard and Jarvis 2005) possibly corresponding to molecular events taking place during the import process. Transit peptides interact with Hsp70 molecular chaperones (Ivey and Bruce 2000; Rial et al. 2000; Zhang and Glaser 2002). Binding to cytosolic Hsp70s most likely prevents precursor protein aggregation prior to import, whereas binding to Hsp70 in the course of import might facilitate the translocation process.

3 Energy requirements of *in vitro* chloroplast protein import

Biochemical experiments carried out with isolated chloroplasts from pea (*Pisum sativum*) revealed the energy-requirements of chloroplast protein import and resulted in a three step model of protein import (Fig. 1) (Schnell and Blobel 1993). In the first step the precursor protein is recognized at the outer envelope membrane.

3.1 Precursor protein recognition at the chloroplast surface

This initial binding of a precursor protein to receptor components at the chloroplast surface does not require energy and is reversible (Perry and Keegstra 1994). In the second step, irreversible binding of the precursor protein at the outer membrane occurs.

3.2 The early translocation intermediate

Irreversible binding of the precursor protein at the outer membrane requires GTP as well as low concentrations of ATP (≤ 100 μM) in the intermembrane space (Olsen and Keegstra 1992; Kessler et al. 1994; Young et al. 1999). The outer membrane-bound form has been termed the "early intermediate" (Schnell and Blobel 1993). At the early intermediate stage the precursor protein extends across the outer membrane and makes contact with protein components at the surface of the inner membrane (Wu et al. 1994; Ma et al. 1996). 100 μM ATP may promote

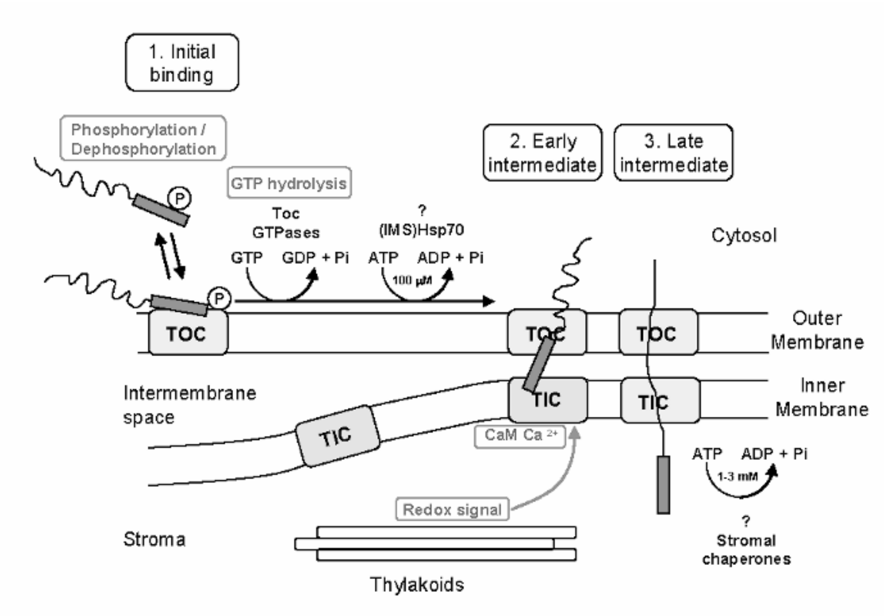


Fig. 1. Stages and regulation of Toc/Tic mediated chloroplast protein import. The three stages (1.-3.) were defined based on the energy requirements of *in vitro* import into isolated chloroplasts. 1. The initial binding of the precursor protein to the chloroplast surface does not require energy and is reversible. 2. For irreversible binding and partial translocation GTP and low amounts of ATP are required, probably used up by the Toc GTPases and intermembrane space-located Hsp70s, respectively. 3. Completion of translocation requires higher concentrations of ATP in the stroma, presumably consumed by stromal chaperones at the inner surface of the envelope. Phosphorylation/dephosphorylation of the transit-peptides and the Toc GTPases as well as GTP-hydrolysis regulate the Toc complex, whereas the Tic complex is regulated by redox-signals and calcium-calmodulin.

precursor protein binding to an intermembrane space Hsp70 protein preventing re-exit into the cytoplasm (Schnell et al. 1994; Becker et al. 2004b).

3.3 The late translocation intermediate

Completion of translocation requires higher concentrations of ATP (1-3 mM) in the stroma (Pain and Blobel 1987; Theg et al. 1989). The ATP in the stroma is presumably consumed by stromal chaperones such as Hsp60 (Cpn60), Hsp93 (ClpC) or of the Hsp70 family at the inner surface of the envelope (Kessler and Blobel 1996). Rapid chilling of the import reaction in the presence of 1-3 mM ATP results in the kinetic trapping of the precursor protein extending across both envelope membranes and engaging the translocation machineries at both the outer and inner envelope membranes. This trapped precursor constitutes the so-called “late intermediate” (Schnell and Blobel 1993).

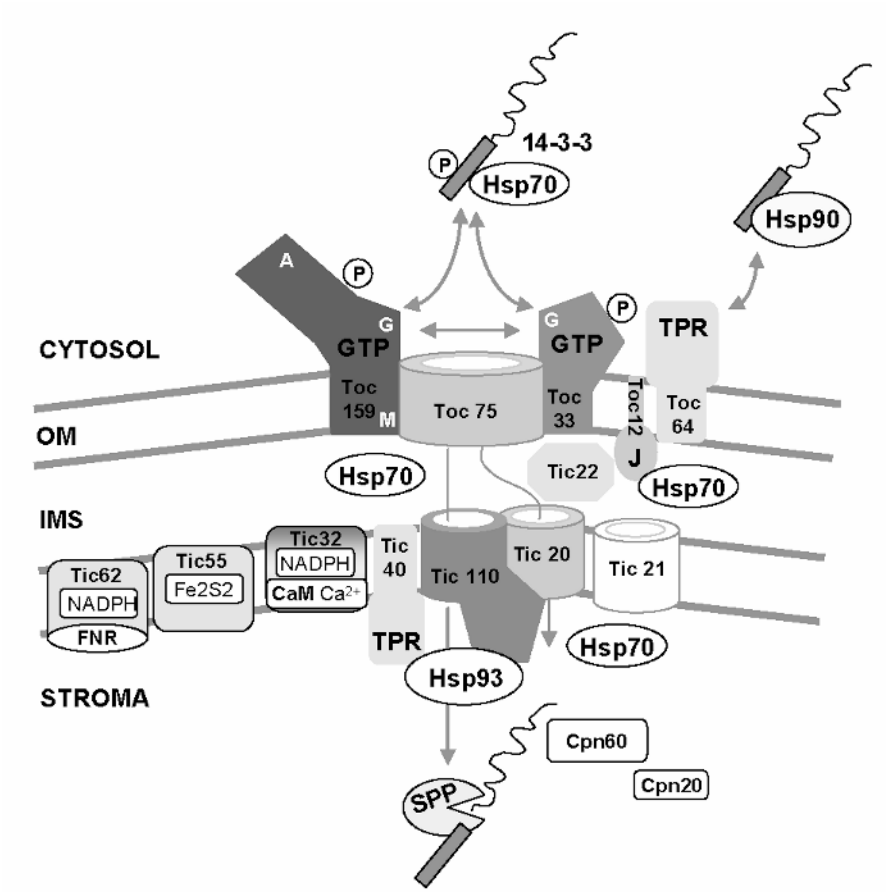
Fig. 2. overleaf. Overview of components of the Toc/Tic chloroplast import machinery of the model plant *Arabidopsis thaliana*. Phosphorylated or non-phosphorylated precursor proteins with an N-terminal transit-peptide are associated with molecular chaperones of the Hsp70 or Hsp90 family and 14-3-3 proteins. They are recognized at the chloroplast surface by the receptor GTPases of the Toc complex (translocase at the outer membrane of chloroplasts), Toc159 and Toc33. Hsp90 associated precursor proteins are recognized by the co-receptor Toc64 before being transferred to Toc33. The GTP-dependent interaction of Toc159 and Toc33 promotes the transfer of the precursor protein to the translocation channel Toc75. In the intermembrane space (IMS), the precursor protein interacts with Hsp70 recruited by the J-domain of the Toc component Toc12 and with Tic22. This interaction probably favour the tight association of the Toc and Tic complexes (translocase at the inner membrane of chloroplasts). Tic110 and/or Tic20 constitute the translocation channel at the inner membrane of the chloroplast. Tic110 has an additional role in the recruitment of chaperones (Hsp93 and cpn60) to the translocation machinery and is assisted by the co-chaperone Tic40. In the stroma the transit-peptide is cleaved off by stromal processing peptidase (SPP) and the folding of the mature protein is aided by the chaperonins cpn60 and cpn20. Tic62, Tic55, and Tic32 are redox-sensing Tic components having a regulatory function. As a calmodulin binding protein Tic32 is also involved in calcium-regulation of the import process. Like Tic20, Tic21 may take over a function as part of the inner membrane protein-conducting channel and replace Tic20 depending on the developmental stage.

4 Identification of components of the translocation machinery

Chloroplast envelope proteins participating in precursor protein import were first identified in the *in vitro* import system using pea chloroplasts. Different experimental approaches using translocation intermediates, chemical crosslinking and antibody inhibition, respectively, resulted in overlapping sets of candidate components (Fig. 2) (Hirsch et al. 1994; Kessler et al. 1994; Perry and Keegstra 1994; Schnell et al. 1994). Interestingly, three of the components now termed Toc159, Toc34, and Toc75 (Translocase at the outer membrane of the chloroplast followed by the molecular mass in kilodaltons) were associated with both the early intermediate and the late intermediate whereas another protein Tic110 (Translocase at the inner membrane of the chloroplast) was exclusively associated with the late intermediate (Schnell et al. 1994). These findings are consistent with the early intermediate extending solely across the outer membrane and the late intermediate extending across both membranes simultaneously engaging components at both envelope membranes.

4.1 Components of the Toc complex

To date five different components of the pea Toc complex have been described: two GTPases, Toc159 and Toc34, and a channel protein Toc75 together form a



stable core complex. In addition, a TPR-domain containing protein Toc64 (Sohrt and Soll 2000) and J-domain containing protein Toc12 (Becker et al. 2004b) have been identified as loosely associated components.

4.1.1 The Toc core complex

Reconstitution experiments into artificial lipid vesicles have demonstrated that the Toc core complex consisting of a 86 kDa proteolytic fragment of Toc159 (Toc86, see also 4.1.2), Toc34 and Toc75 translocates precursor proteins *in vitro* (Schleiff et al. 2003b). Analysis of the Toc core translocon purified from pea, revealed a molecular mass of ~ 500 kDa in size exclusion chromatography and a stoichiometry of Toc159:Toc34:Toc75 of 1:4:4 in this complex (Schleiff et al. 2003a). In single particle electron microscopy of the complex, 13 nm toroid-shaped particles were observed, which appeared in three-dimensional reconstitution as structures with four pores and a central finger domain (Schleiff et al. 2003a). In BN-PAGE the pea Toc complex migrates at 800-1000 kDa (Kikuchi et al. 2006; Chen and Li

2007). The molecular weight differences observed may be due to the techniques employed, partial degradation or the presence of Toc64, Toc12, precursor proteins or so far unidentified components in the higher molecular weight complex.

4.1.2 Toc159 and Toc34: chloroplast protein import receptors

Toc159 and Toc34 are protease sensitive proteins in isolated chloroplasts indicating exposure at the chloroplast surface (Hirsch et al. 1994; Kessler et al. 1994; Seedorf et al. 1995). In agreement with the finding numerous studies indicate a role in precursor protein recognition. The two proteins have homologous GTP-binding domains (G-domains). GTP-dependence of early intermediate formation implicates the two proteins in the early stages of protein import. Toc159 has an additional N-terminal acidic domain (A-domain) and is anchored in the outer chloroplast membrane by a C-terminal membrane anchoring domain (M-domain). Due to the protease sensitivity of the A-domain, pea Toc159 was initially identified as a smaller protein designated Toc86 comprising only the G- and M-domains (Bolter et al. 1998). In addition to the membrane-integrated form, a portion of cellular Toc159 was present as a soluble, cytoplasmic protein, hinting at the dynamics in Toc complex assembly (Hiltbrunner et al. 2001b; Bauer et al. 2002; Smith et al. 2002). Toc34 consists of the G-domain followed by a stretch of hydrophobic amino acids anchoring the protein in the outer membrane. The crystal structure of pea Toc34 showed GDP-bound dimers (Sun et al. 2002). Biochemical studies indicate that Toc34 not only forms homodimers but may also form heterodimers with Toc159 via their respective GTP-binding domains (Weibel et al. 2003).

4.1.3 Toc75: the protein conducting channel

Toc75 has similarity to bacterial solute channels and was therefore identified as a candidate for a protein conducting channel (Schnell et al. 1994; Gentle et al. 2005; Schleiff and Soll 2005). Electrophysiological studies on the reconstituted protein demonstrated ion conducting channel properties (Hinnah et al. 1997, 2002). During early intermediate formation the precursor protein is thought to insert across the Toc75 channel and make initial contact with Tic components at the outer surface of the inner membrane. Interestingly, Toc75 has sequence homology with Omp85 in bacteria and Tob55 in yeast which are proteins involved in the insertion of beta-barrel proteins in the bacterial and mitochondrial outer membranes respectively (Gentle et al. 2005; Paschen et al. 2005).

4.1.4 Toc64

Toc64 is loosely associated with the Toc-complex and therefore not considered part of its core (Sohrt and Soll 2000; Qbadou et al. 2006). Toc64 contains TPR repeats forming a putative docking site for a subset of precursor proteins bound to cytosolic Hsp90. The Toc64 TPR domain interacts with the receptor GTPase Toc34 initiating the transfer of the precursor protein to the Toc complex (Qbadou et al. 2006). Its function as an accessory import receptor for a subset of precursor proteins may be comparable to the function of Tom70 in mitochondrial protein import (Young et al. 2003). Toc64 is non-essential both in *Arabidopsis thaliana* and *Physcomitrella patens*. Moreover, Toc64 deficient *Physcomitrella patens* showed no chloroplast protein import deficiency. Therefore, PpToc64 has been renamed PpOEP64 (Hofmann and Theg 2005b).

4.1.5 Toc12

Toc12 is a J-domain containing protein facing the intermembrane space (Becker et al. 2004b). It stimulates ATP hydrolysis by DnaK *in vitro* and interacts directly with intermembrane space localized Hsp70 proteins. Toc12 may therefore function to anchor an intermembrane space Hsp70 which in turn retains precursor proteins in the translocation channel and prevents them from slipping back into the cytosol. The J-domain of Toc12 may be conformationally stabilized by an intramolecular disulfide bridge, which could be sensitive to the redox state of the chloroplast (Becker et al. 2004b; Bedard and Jarvis 2005).

4.2 Components of the Tic complex

A number of components of the inner membrane translocation machinery have been identified (Fig. 2). Unlike the Toc core complex, they do not appear to form a stable complex.

4.2.1 Tic22 and Tic20

Both Tic22 and Tic20 were not originally identified as components of the early or late translocation intermediates. However, chemical crosslinking at the early intermediate stage revealed the presence of the two inner membrane proteins and indicated that the precursor proteins had completely traversed the outer chloroplast membrane (Kouranov and Schnell 1997; Kouranov et al. 1998; Chen et al. 2002). Biochemical analysis demonstrated that Tic22 was an extrinsic protein exposed at the surface of the inner membrane suggesting a function as an inner membrane receptor or in the formation of contact sites between the import complexes (Kouranov and Schnell 1997; Kouranov et al. 1998).

In contrast, Tic20 was integral to the inner membrane suggesting a function as a component of the protein conducting channel at the inner membrane. Tic20 is predicted to span the membrane with four transmembrane α-helices (Kouranov et al. 1998; Chen et al. 2002) and shares some similarity with TIM channel proteins (Reumann and Keegstra 1999). A role as an inner membrane protein-conducting channel expressed in all tissues and likely to be required for early plastid development has been proposed (Chen et al. 2002; Teng et al. 2006).

4.2.2 Tic110

Tic110 was first described as a component of the late translocation intermediate (Schnell et al. 1994; Wu et al. 1994). The sequence contains two predicted hydrophobic alpha helices at the N-terminus and a large hydrophilic domain at the C-terminus (Kessler and Blobel 1996; Lubeck et al. 1996). Topology data suggest that the C-terminus extends into the stroma (Kessler and Blobel 1996; Jackson et al. 1998; Inaba et al. 2003). The C-terminus, in conjunction with Tic40 functions as a binding site for the chaperones ClpC and Hsp60 which are required for folding subsequent to import (Chou et al. 2003, 2006; Kovacheva et al. 2005). Via its transmembrane helices, Tic110 may participate in the formation of the translocation channel at the inner membrane (van den Wijngaard and Vredenberg 1999; Heins et al. 2002).

4.2.3 Tic40

Tic40 functions at the same late stage of import as Tic110 and Hsp93 (ClpC) and is recruited to Tic110 upon transit-peptide binding by Tic110. It is a membrane-anchored protein with a large hydrophilic domain protruding into the stroma. The C-terminal portion of Tic40 contains similarity to the co-chaperones Sti1p/Hop (Stahl et al. 1999; Chou et al. 2003). Recently, it was demonstrated that Tic40 indeed acts as a co-chaperone by stimulating ATP hydrolysis by Hsp93 and transit-peptide release from Tic110 (Chou et al. 2006).

4.2.4 Tic21

AtTic21 (also known as CIA5) was identified in a genetic screen scoring for mistargeting and accumulation of an antibiotic resistance marker in the cytosol of *Arabidopsis* (Sun et al. 2001; Teng et al. 2006). The phenotype of the knockout of the *Arabidopsis* gene At2g15290 annotated as atTic21 (CIA5) (Teng et al. 2006) was published by two independent research groups recently (Teng et al. 2006; Duy et al. 2007). Both groups found that At2g15290 mutant plants are seedling lethal on soil and chlorotic to albino when cultivated on media supplemented with sucrose. Furthermore, in accordance with each other, both studies identified the corresponding gene product as an integral inner envelope membrane protein of chloroplasts. However, different functions were attributed to the At2g15290 gene product. According to Teng et al. (Teng et al. 2006) At2g15290 (atTic21) functions as a part of the inner membrane protein-conducting channel, similar to Tic20 but at later stages of leaf development. Strong arguments for a function of At2g15290 in chloroplast protein import are the observed inner membrane import defect of chloroplasts isolated from plants expressing a mutated variant of the protein (*tic21*/*cia5* K112C) as well as the co-precipitation of the putative Tic21 with Toc and Tic components even in the absence of crosslinking agent (Teng et al. 2006). In contrast, Duy et al. (Duy et al. 2007) claimed At2g15290 to encode a permease that functions in iron transport across the inner envelope of chloroplasts and therefore annotated the gene as PIC1 for PERMEASE IN CHLOROPLASTS

1 (Duy et al. 2007). Indeed, At2g15290 shares sequence similarity with potential metal iron transporters from cyanobacteria (e.g. *Synechocystis* sll1656). In Affimetrix microarray analysis of the At2g15290 mutant changes in the expression of metal homeostasis-associated and a drastic downregulation of photosynthetic genes were observed (Duy et al. 2007). Ferritin was found to be upregulated and an accumulation of ferritin clusters in plastids was revealed by ultrastructural analysis. It is counterintuitive that the lack of a putative iron permease causes phytoferritin accumulation normally observed as consequence of iron-overload. Expression of the cDNA of At2g15290 as well as of its *Synechocystis* homolog sll1656 in a yeast mutant defective in low- and high-affinity Fe uptake partially restored its growth defect under iron-limited conditions (Duy et al. 2007) indicating their ability to transport iron. However, over-expression of *Synechocystis* sll1656 in the *Arabidopsis* At2g15290 knockout mutant did not result in complementation (Teng et al. 2006). Therefore At2g15290 is most likely not just an "ancient" permease but acquired additional essential functions during evolution. Additional work is clearly required to exclude a role of At2g15290 in chloroplast protein import.

4.2.5 Tic55

Tic55, Tic62, and Tic32 have been identified as a redox-sensing Tic components. Tic55 is an integral protein at the inner membrane of chloroplasts and comigrates with Tic110 and ClpC in BN-PAGE (Caliebe et al. 1997). Tic55 contains a Rieske-type iron-sulphur cluster and a mononuclear iron binding site and may therefore catalyse electron-transfer reactions. Tic55 has been suggested to act as regulatory component of the Tic complex involved in signal-transduction or redox-regulation during protein import (Soll 2002).

4.2.6 Tic62

Tic62 has a conserved NAD(P) binding site at its N-terminus and a binding site for ferredoxin-NAD(P) reductase (FNR) at its stroma-exposed C-terminus (Kuchler et al. 2002). The photosynthetic electron flux may regulate the import apparatus via FNR, transferring electrons from ferredoxin to the NAD(P) associated with Tic62.

4.2.7 Tic32

Tic32, with similarity to short-chain dehydrogenase/reductase (SDR), is an integral inner envelope protein and was shown to associate with Tic110 (Hormann et al. 2004). Tic32 is an essential protein, the *tic32 Arabidopsis* knockout mutant being embryo lethal (Hormann et al. 2004). Notably, Tic32 may function not only in redox but potentially also in calcium regulation of the protein import (see section calcium regulation) (Chigri et al. 2006). A striking feature of Tic32 is its lack of a cleavable transit-peptide and its import by a so far unknown pathway (Nada and Soll 2004) (see section alternative import pathways).

5 Regulation at the Toc and Tic complexes

The nature of some of the Tic and Toc proteins hint at their regulation. Analysis of GTP-binding proteins Toc34 and Toc159 in the Toc-complex has demonstrated their regulation by GTP-binding and hydrolysis. Phosphorylation/dephosphorylation of some precursors and Toc-components may provide an additional layer of regulation at the Toc-complex. At the Tic complex. Tic55, Tic62, and Tic32 suggest regulation by the redox state of the chloroplast and calcium signalling (Fig. 1).

5.1 GTP-regulated protein recognition at the Toc complex

GTP-binding and hydrolysis at the Toc receptor GTPases Toc159 and Toc34 most likely explain GTP-dependent precursor binding to the chloroplast surface as well as formation of the early import intermediate. Toc159 and Toc34 share some motifs involved in GTP-binding and hydrolysis with Ras-like GTPases (Kessler et al. 1994; Sun et al. 2002). However, the crystal structure of pea Toc34 (psToc34) revealed significant structural variations when compared to Ras-like GTPases. This suggests that the Toc-GTPases utilize a unique mechanism of GTP binding and hydrolysis (Sun et al. 2002) and therefore constitute a new class of GTPases. The observation of GDP-bound Toc34 homodimers as well as Toc159/Toc34 heterodimers suggests that the early stages of import involve GTPase-regulated interactions of the Toc-GTPases (Smith et al. 2002; Sun et al. 2002). In pea, Toc159 and Toc34 are the only Toc-GTPases known so far, but in other species small families have been identified: in *Arabidopsis thaliana* Toc159 has four homologues (atToc159, -132,-120 and -90) and Toc34 has two (atToc33 and Toc34). AtToc159 and atToc33 are considered the orthologs of pea Toc159 and Toc34, respectively.

5.1.1 Toc GTPase cycle

Small GTPases are known to behave like molecular switches cycling between "active" GTP-bound and "inactive" GDP bound states (Bourne et al. 1990). This normally involves interactions with regulatory proteins such as guanine nucleotide exchange factors (GEFs) (Cherfils and Chardin 1999) or GTPase activating proteins (GAPs) (Scheffzek and Ahmadian 2005). To date no GEFs for the Toc GTPases have been identified. But the low intrinsic GTP hydrolysis rate of Toc34 was stimulated by precursor proteins that may therefore serve as GAPs (Jelic et al. 2002; Becker et al. 2004a).

The crystal structure of Toc3, revealed a GDP-bound homodimer. The arrangement of the two Toc34 monomers suggested that one could act as a GAP for the other, by inserting a catalytic residue (arginine 133 in psToc34, arginine 130 in atToc33) into the active site of the other (Sun et al. 2002). However, studies diverge on the catalytic constants of the GTPase activities of the psToc34(R133A) and atToc33(R130A) mutants with regard to the wild type (Weibel et al. 2003; Reddick et al. 2007; Yeh et al. 2007). This has been attributed to the differing experimental conditions used by the different groups. As a consequence a clear verdict for or against the arginine finger theory is not yet possible.

In addition to its proposed function as an arginine finger, Arg133 of pea Toc34 (Arg130 in *Arabidopsis* Toc33) plays a key role in homodimer formation as well as heterodimer formation with Toc159. The mutated proteins psToc34(R133A) as well as atToc33(R130A) behave as monomeric proteins (Weibel et al. 2003; Reddick et al. 2007; Yeh et al. 2007) and are reduced in their ability to interact with Toc159 (Weibel et al. 2003). Recently, the crystal structure of atToc33(R130A) was published (Yeh et al. 2007). The atToc33(R130A) mutant indeed crystallized as a monomer its structure strongly resembling the monomer structure of psToc34 (Sun et al. 2002). The crystal structure of wild type atToc33, which would be valuable for the evaluation of the structural changes caused by the R130A mutation, is not available. It now appears clear that the dimerization between the Toc-GTPases and their respective interactions with precursor proteins are intimately entwined with GTP-binding and -hydrolysis. Many of the details including the existence of Toc34 homodimers *in vivo* and the nature of nucleotide exchange factors are still mysterious.

5.1.2 Interplay of the Toc GTPases

Substantiated by a series of studies it is well established that Toc159 and Toc34 bind to precursor proteins and dimerize via their GTP binding domains (Kessler and Schnell 2002, 2004). Regardless, the order of the GTP-regulated import events *in vivo* is not known with certainty and has been obscured by the unexpected complexity of the system.

One possibility is that the targeting of a cytosolic precursor protein to the Toc complex is coupled to the GTP-dependent association of the either soluble or integral membrane Toc159 receptor with the Toc-complex (Bauer et al. 2002; Smith et al. 2002; Bedard and Jarvis 2005). In this scenario, termed the targeting hypothesis, the precursor protein in complex with Toc159 is targeted to the Toc complex, where a homotypic interaction between Toc159 and Toc34 coupled to GTP hydrolysis initiates the association of Toc159 in the Toc complex and the transfer of the precursor into the translocation channel Toc75.

The observation that Toc159 GTPase mutants affected in GTP-binding and/or hydrolysis are mislocated to the cytosol and fail to functionally complement the import defect of the *toc159* null mutant *ppi2* (see section 6.1.1) are compatible with this hypothesis (Smith et al. 2002; Lee et al. 2003). The minimal requirements for Toc159 insertion were studied in an *in vitro* system with purified Toc core components reconstituted into artificial vesicles (Wallas et al. 2003). The study demonstrated that the insertion of Toc159 does not only require GTP hydrolysis at both GTPases as well as the presence of Toc75 but also involves a previously unknown interaction of the Toc159 membrane anchoring M-domain with the G-domain of atToc33 (Wallas et al. 2003). Upon insertion into the Toc complex the membrane anchoring domain appears to assume a substantial role in the import reaction itself. The fact that proteolytic removal of the cytosol-exposed A- and G-domain of Toc159 did not completely deactivate pre-protein import *in vitro* (Chen et al. 2000a) and that the M-domain alone partially complements the *ppi2* mutant (Lee et al. 2003), indicates that the M-domain is more than a passive membrane-anchor and may participate in translocation channel formation.

According to motor hypothesis likewise founded on a series of studies (Schleiff et al. 2003a, 2003b; Becker et al. 2004a), Toc159 functions only after membrane integration as a GTP-driven motor protein at the centre of the Toc complex. This motor hypothesis is based on an *in vitro* reconstitution experiment in which the carboxy-terminal Toc86 fragment of pea Toc159 alone was sufficient to mediate GTP-dependent translocation of a precursor across the Toc75 translocation channel (Schleiff et al. 2003b).

The motor hypothesis proposes that Toc34 acts as an initial receptor and Toc159 as a docking partner for recruiting Toc34. The transit-peptide of the precursor stimulates an interaction of two GTPases in their GTP-bound state. GTP-hydrolysis at Toc34 results in the transfer of the transit peptide to Toc159 and dissociation of Toc34 from the complex. Subsequent cycles of GTP-hydrolysis at Toc159 push the precursor protein across the translocation channel.

The two hypotheses have the GTP-regulated precursor recognition and the interaction of Toc159 and Toc34 in common. They differ mainly in the hierarchy of the two Toc GTPases, which has proven difficult to resolve. Further investigation will be required to determine the mechanistic details of the Toc GTPase cycle. For example, how the crucial GDP-GTP nucleotide exchange occurs at Toc159 and Toc34 is completely enigmatic so far.

5.2 Regulation by phosphorylation

Phosphorylation of the transit peptide of the small subunit of Rubisco in the cytosol has been demonstrated to influence the rate of its import *in vitro* (Waegemann and Soll 1996; May and Soll 2000). Transit peptide phosphorylation permits the binding of cytosolic 14-3-3 proteins and Hsp70 molecular chaperones. This complex of 14-3-3 proteins and Hsp70s has been designated the guidance complex and was shown to stimulate import three to fourfold (May and Soll 2000). Moreover, phosphorylation influences the recognition of the precursor proteins by the import receptors, as Toc34 binds with high affinity to phosphorylated precursors (Sveshnikova et al. 2000) whereas Toc159 recognizes only non-phosphorylated precursors (Becker et al. 2004a). Mutating the phosphorylation site of precursor proteins does not result in mistargeting *in vitro* and *in vivo* (Nakrieko et al. 2004). Thus, transit peptide phosphorylation is not essential for targeting specificity but influences import kinetics. A serine/threonine-specific protein kinase activity for transit peptides was found in pea leaf mesophyll cells and wheat germ lysate (Waegemann and Soll 1996; May and Soll 2000) and recently a family of chloroplast precursor protein kinases was purified from *Arabidopsis* (Martin et al. 2006). The latter consists of three cytosolic serine/threonine kinases (At2g17700, At4g35780, At4g38470). The three kinases utilize ATP to phosphorylate several chloroplast precursors but not a mitochondrial precursor protein *in vitro*.

Not only precursor proteins but also the Toc GTPases (Toc159 and Toc34 of pea) are subject to phosphorylation (Sveshnikova et al. 2000; Fulgosi and Soll 2002). Two outer envelope kinases phosphorylating the two receptors were partially purified from pea chloroplasts (Fulgosi and Soll 2002). Phosphorylation of pea Toc34 at serine 113 and atToc33 at serine 181 was demonstrated to inhibit precursor protein recognition and GTP binding *in vitro* (Sveshnikova et al. 2000; Jelic et al. 2002; Jelic et al. 2003). Mutations in atToc33 that prevent or mimic phosphorylation at serine 181 did not influence the function of the receptor *in vivo* (Aronsson et al. 2006). In summary, phosphorylation of the import receptor Toc33, similar to precursor protein phosphorylation, is not essential *in vivo*, but may influence the rate of import, which is consistent with a regulatory function.

5.3 Redox-regulation

The redox-state of the chloroplast depends on light and consequently photosynthetic electron transport. It has been suggested to have a regulatory influence on chloroplast protein import (Caliebe et al. 1997; Kuchler et al. 2002; Hormann et al. 2004). Under illumination, the reducing power could act on the redox-sensing components of the Tic complex Tic55, Tic62, and Tic32 as well as on the disulphide bridge stabilizing the J-domain of Toc12 and thereby modulate the import characteristics of the Toc and Tic complexes.

Support for a role of the chloroplast redox state in the regulation of chloroplast protein import came from studies on precursors that exhibit distinct import patterns in the chloroplast under light and dark conditions (Hirohashi et al. 2001). The precursors of maize FdIII (non-photosynthetic ferredoxin) and FNRII (ferredoxin-$NADP^+$ reductase II) accumulated in the intermembrane space of the chloroplast envelope membranes, whereas in the dark, the proteins were processed correctly. Furthermore import experiments with NAD-analogues indicated that the precursor of one isoform of *Arabidopsis* leaf specific FNR (pFNR-L1) is translocated preferentially at a high NAD(P)/NAD(P)H ratio, i.e., in the dark (Kuchler et al. 2002). As many precursor proteins are imported into chloroplasts equally well in the light as in the dark, the light-/redox-regulation is probably not a general regulatory element of the Toc and Tic complexes but rather specific for certain precursor proteins.

5.4 Calcium/calmodulin regulation

Recently, a new mode of regulation of chloroplast protein import by calcium has been proposed (Chigri et al. 2005, 2006). Ophiobolin A, a specific inhibitor of calmodulin, as well as two calcium ionophores inhibited the import of precursor proteins with N-terminal cleavable presequences into isolated pea chloroplasts. The calcium regulation seems to be restricted to the Toc/Tic import pathway, as

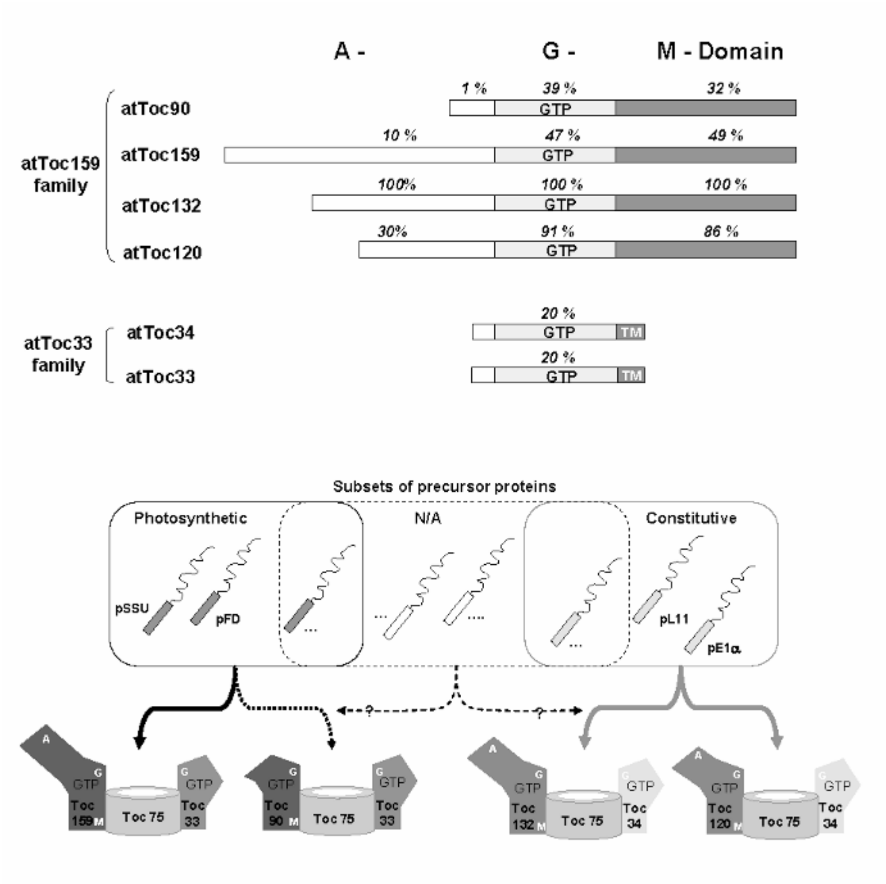


Fig. 3. The Toc GTPase families of *Arabidopsis thaliana*. The upper part shows a linear representation of the four Toc159 homologues (Toc90, Toc159, Toc132, and Toc120), and the two Toc34 homologues (Toc33 and Toc34) including the position of the acidic domains (A), GTPase domains (G) and membrane-anchoring domains (M). Toc33 and Toc34 consist only of a GTPase domain and a small trans-membrane anchor (TM). The values above the domains give the percentage of sequence identity between the domains relative to Toc132. The borders of the G-domains for this analysis were defined as follows: Toc90 aa 38-399, Toc159 aa 727-1092, Toc132 aa 455-805, Toc120 aa 339-687. Genetic and biochemical studies suggest that the members of the Toc GTPase families assemble into different Toc complexes with different substrate specificities (lower part). Toc159 associates preferentially with Toc33 to form the main import complex for photosynthetic proteins. Toc132 and Toc120 associate preferentially with Toc34 into import complexes more involved in the import of constitutive house-keeping proteins. The determinants in the transit-peptides for a specific import pathway have not yet been defined, therefore, a clear classification of precursor protein subsets is not possible. Toc90 may have a minor role in photosynthetic protein import. A preferential association with Toc33 has not been demonstrated.

the import of proteins that do not contain a presequence like psToc34 and psToc32 was not affected (Chigri et al. 2005). The site of ophiobolin A action could be located at the inner envelope. Intriguingly, Tic32, a component of the Tic complex, turned out to be the only calmodulin-binding protein at the inner envelope membrane (Chigri et al. 2006). As Tic32 also has NADPH-dependent dehydrogenase activity it could act as a transducer of both redox and calcium regulation of chloroplast protein import. It has been proposed that the association of Tic32 with other Tic components may be regulated in a NADPH and/or Ca^{2+}/calmodulin dependent manner (Chigri et al. 2006). A number of candidate chloroplast calmodulin and calmodulin-like sequences with potential N-terminal transit sequences have been retrieved from the databases (Chigri et al. 2005). Currently, the nature of the chloroplast calmodulin or calmodulin-like proteins involved in import regulation is not known.

6 Functional specialization in the general import pathway

In *Arabidopsis* the chloroplast protein import receptors pea Toc34 and Toc159 are encoded by small gene families of two (atToc33 and atToc34) and four genes (atToc159, atToc132, atToc120, atToc90), respectively (Fig. 3) (Jackson-Constan and Keegstra 2001a). All Toc34/Toc159 homologous in *Arabidopsis* share sequence similarity in their central GTP-binding domain and seem to derive from an ancient eukaryotic GTP-binding protein that evolved into the Toc GTPases and the AIG family of GTPases (avirulence induced gene) (Reuber and Ausubel 1996; Reumann et al. 2005). All Toc159 homologues have C-terminal membrane anchoring domains (M-domains) but differ significantly in their N-terminal parts. Toc90, the most ancient and distant family member lacks the N-terminal acidic domain (A-domain) and instead has only a short non-acidic N-terminal extension. The other members atToc159, atToc132, and atToc120 as a consequence of domain enlargement and introduction of negative charges have A-domains that vary greatly in length and sequence between the different isoforms. Phylogenetic analysis clearly revealed that atToc132 and atToc120 form a subgroup in the Toc159 family (Hiltbrunner et al. 2001a; Kubis et al. 2004). There is increasing evidence that the members of the Toc GTPase families represent functionally specialized import receptors assembling into Toc complexes of distinct composition. This model is supported by expression patterns of the isoforms and phenotypes of mutants for these receptors as well as complementation studies (Hiltbrunner et al. 2001a; Ivanova et al. 2004; Kessler and Schnell 2004; Kubis et al. 2004).

6.1 Plastid protein import mutants and phenotypes

Plastid protein import mutants exhibit a gradient of phenotypes ranging from embryo lethal to wild type (Table 1). The analysis of import mutants has provided important new insight into the import process and its role in plant development.

Table 1. Summary of *Arabidopsis* homologues of the pea chloroplast import machinery including single mutant phenotypes

Pea	*Arabidopsis* (ppi mutant) Designation	Proposed function	Mutant phenotype	Reference
Toc159	atToc159 (ppi2) At4g02510	Import receptor/motor protein	albino	(Bauer et al. 2002)
	atToc132 (toc132) At2g16640	Import receptor	none to pale-yellow green	(Ivanova et al. 2004; Kubis et al. 2004)
	atToc120 (toc120) At3g16620	Import receptor	none	(Ivanova et al. 2004; Kubis et al. 2004)
	atToc90 (ppi4) At5g20300	Import receptor	none	(Hiltbrunner et al. 2004; Kubis et al. 2004)
Toc34	atToc33 (ppi1) At1g02280	Import receptor	pale-green	(Jarvis et al. 1998)
	atToc34 (ppi3) At5g05000	Import receptor	none, reduced root length	(Gutensohn et al. 2000; Constan et al. 2004a)
Toc75	atToc75-III At3g46740	Translocation channel	embryo lethal	(Baldwin et al. 2005)
	atToc75-I At1g35860	no (pseudogene)	N/A	(Baldwin et al. 2005)
	atToc75-IV At4g09080	Translocation channel ?	abnormal etioplasts, de-etiolation defect	(Baldwin et al. 2005)
Toc64	atToc64-III At3g17970	Import co-receptor	none	(Qbadou et al. 2006)
	atToc64-I At1g08980	amidase 1 (AMI1) (no TPR domain)	N/A	(Pollmann et al. 2003; 2006)
	atToc64-V At5g09420	Mitochondrial receptor	N/A	(Chew et al. 2004)
Toc12	atToc12 ? At1g80920	dnaJ homolog	N/A	(Becker et al. 2004b; Becker 2005)
Tic110	atTic110 At1g06950	Translocation channel, chaperone recruitment	embryo lethal	(Inaba et al. 2005; Kovacheva et al. 2005)
Tic40	atTic40 At5g16620	co-chaperone	chlorotic	(Chou et al. 2003; Kovacheva et al. 2005)
Tic20	atTic20-I At1g04940	Translocation channel	seedling lethal, albino	(Chen et al. 2002; Teng et al. 2006)
	atTic20-IV At4g03320	N/A	N/A	(Jackson-Constan and Keegstra 2001a)
Tic62	atTic62 At3g18890	Redox-regulation	N/A	
Tic55	atTic55 At2g24820	Redox-regulation	N/A	

Pea	*Arabidopsis* (ppi mutant) Designation	Proposed function	Mutant phenotype	Reference
Tic32	atTic32 At4g23430	Redox, calcium regulation	embryo lethal	(Hormann et al. 2004)
Tic22	atTic22-IV At4g33350	Formation of contact sites	N/A	
	atTic22-III At3g23710	N/A	N/A	
Tic21	atTic21 (cia5) At2g15290	Translocation channel or PIC1 permease ?	albino, chlorotic, precursor accumulation	(Teng et al. 2006; Duy et al. 2007)
Hsp93	atHsp93-V (ClpC1) At5g50920	chaperone	Retarded growth chlorotic	(Sjogren et al. 2004; Constan et al. 2004b; Kovacheva et al. 2005)
	atHsp93-III (ClpC2) At3g48870	chaperone	not visible	(Constan et al. 2004b; Kovacheva et al. 2007)
SPP (CPE)	atCPE At5g42390	Transit peptide removal	(antisense) seedling-lethal	(Zhong et al. 2003)

Embryo lethality indicated that plastid protein import into plastids is absolutely required to establish essential, housekeeping biosynthetic pathways inside the organelle. Wild type, pale green and albino phenotypes suggest either accessory-functions or partial redundancy within small families of homologues. In the following, we discuss the phenotypes and expression patterns of the members of the Toc-GTPase family in detail as their analysis permitted the definition of their roles in protein import sub-pathways. These rely on components homologous and mechanisms similar to those of pea Toc159 and Toc34 while their substrates vary. For a complete summary of phenotypes of the known *Arabidopsis* chloroplast protein import components please refer to Table 1.

6.1.1 ppi2: the Toc159 knockout mutant

The *Arabidopsis* plastid protein import mutant 2 (*ppi2*), has a revealing albino phenotype due to a T-DNA insertion in the *atTOC159* gene (Bauer et al. 2000; Asano et al. 2004). The *ppi2* albino phenotype results in seedling lethality when plants are grown on soil. Similar to proplastids, *ppi2* plastids in cotyledons of soil-grown plants are undifferentiated and lack thylakoid membranes. Major photosynthetic genes such as RbcL, RbcS, and LhcII (CAB) are transcriptionally repressed indirectly due to the chloroplast biogenesis defect. However, it is noteworthy that *ppi2* plastids still import and accumulate the corresponding proteins in small amounts indicating the existence of Toc159 independent import pathways. Moreover, the expression and import of many proteins not involved in photosynthesis such as Toc75 and Tic110 (Bauer et al. 2000) and pE1α (Smith et al. 2004) is not affected in *ppi2*. This led to the conclusion that atToc159 is the major import receptor for photosynthetic protein import and that the residual import of proteins into *ppi2* may be mediated by the remaining Toc159 homologues.

6.1.2 ppi4/toc132/toc120: the knockout mutants of the Toc159 homologues

The phenotypes of knockout mutants of the other Toc159 homologues and complementation experiments with these mutants support the hypothesis of functionally specialized import receptors (Hiltbrunner et al. 2001a, 2004; Ivanova et al. 2004; Kubis et al. 2004). The single *toc120* and *toc90* (*ppi4*) mutants have no visible phenotypes throughout development, *toc132* single mutant plants reveal no or a very slight pale phenotype in young seedlings to clear yellow-green and reticulate phenotype in mature plants depending on the ecotype used (Ivanova et al. 2004; Kubis et al. 2004).

Double knockout plants revealed the functional overlap of the receptors. *Toc120 toc132* double knockout plants were reported either to be embryo or seedling lethal, consistent with their role in the import of essential housekeeping genes and functional redundancy of Toc132 and Toc120 in this process. Correspondingly, overexpression of either of the two genes was sufficient to rescue the *toc120 toc132* mutant phenotype (Kubis et al. 2004). In contrast ectopic expression of atToc159 was not able to complement.

Crosses between *toc90* and other Toc159 homologue mutants did not result in any new visible phenotype expect for the combination of *toc90* (*ppi4*) with *ppi2* (Hiltbrunner et al. 2004). The *toc90* mutation aggravated the *ppi2* albino phenotype. *Ppi2 ppi4* did not accumulate detectable amounts of the photosynthetic protein CAB whilst the import of housekeeping genes was unaffected, indicating that Toc90 – when compared to Toc159 - may have an accessory function in import of photosynthetic proteins (Hiltbrunner et al. 2004).

6.1.3 ppi1/ppi3: the knockout mutants of Toc33 and Toc34

Single mutants of the two Toc34 homologues in *Arabidopsis* display relatively mild phenotypes (Jarvis et al. 1998; Constan et al. 2004a). The *ppi1* mutant carrying a T-DNA insertion in the atToc33 gene reveals a pale green phenotype most pronounced in young leaves (Jarvis et al. 1998). *Ppi1* chloroplasts of young leaves are small and have poorly developed thylakoids. Similar but weaker than in *ppi2* a downregulation of nuclear genes encoding photosynthetic chloroplast proteins was observed in *ppi1* plants (Kubis et al. 2003). Lack of atToc34 in the *ppi3* mutant does not cause any obvious phenotype in green aerial tissues but a slight reduction in root length was observed (Gutensohn et al. 2004; Constan et al. 2004a). The observations that overexpression of atToc34 complements *ppi1* and that the *ppi1 ppi3* double knockout is embryo lethal indicating that the two proteins functionally overlap to support an essential function. Genomics and proteomics data suggest that the two receptors while overlapping have preprotein-recognition specificity, i.e. Toc33 for photosynthetic precursors and Toc34 for housekeeping proteins (Kubis et al. 2003).

6.2 Expression patterns of Toc GTPases

The expression pattern of the different Toc GTPases was analysed in several studies by RT-PCR (Bauer et al. 2000; Yu and Li 2001; Ivanova et al. 2004), RNA-blot (Gutensohn et al. 2000; Kubis et al. 2003, 2004), Affymetrix expression (Vojta et al. 2004) and Western blot analyses (Ivanova et al. 2004). Reasonably consistent results were obtained. Toc159 turned out to be the most abundant and most regulated of the four members of the Toc159 family (Bauer et al. 2000; Kubis et al. 2004; Vojta et al. 2004). It is highly expressed in rapidly growing photosynthetic tissue and downregulated in roots consistent with its proposed function as the major import receptor for photosynthetic precursors. The expression pattern of atToc33 parallels the one of atToc159 pointing to a concerted function of these two Toc GTPases. In contrast atToc90, atToc120, atToc132, and atToc34 show lower and much more uniform expression levels in different tissues and developmental stages than atToc159 or atToc33, indicative of their function in the transport of other and more constitutively expressed precursor proteins.

6.3 Biochemical evidence for functional specialization of chloroplast import receptors

The biochemical studies by (Ivanova et al. 2004) indicated that atToc120 and atToc132 indeed assemble into Toc complexes distinct from those containing Toc159. By sequential immunopurification it was demonstrated that atToc120/atToc132 preferentially assemble with atToc34, whereas atToc159 preferentially assembles with Toc33. Neither Toc120 nor Toc132 was found associated with Toc complexes containing Toc159.

Further evidence for the role of the Toc159 homologues as key determinants of import substrate specificity came from precursor binding studies using transit peptides of some selected photosynthetic and non-photosynthetic constitutively expressed precursor proteins. Toc159 specifically interacted with the transit peptides of two photosynthetic proteins (pSSU, pFd) and the transit peptides of three different non-photosynthetic plastid proteins (pE1α, pL11, pPORA) did not compete for this binding (Smith et al. 2004). The opposite was observed for Toc132 that selectively bound to the transit peptide of the constitutively expressed protein pE1α, but much less to that of a photosynthetic precursor (pSSU) (Ivanova et al. 2004).

6.4 Substrate specificity of Toc-GTPase sub-pathways

The genetic and biochemical studies provided evidence for the existence of structurally and functionally distinct translocons in the outer membrane of plastids as well as the existence of at least two different classes of import substrates (Fig. 3). The two classes of import substrates have been operationally defined as photosynthetic (preferred substrates of Toc159 and 33) and housekeeping proteins (preferred substrates of Toc132, -120, and -34). The molecular basis for the discrimination of these substrates by the Toc-GTPase receptors is not known. It is generally assumed, however, that increased expression and accumulation of a protein in the *ppi1* mutant suggests import via a Toc33-independent import pathway whereas reduced expression and accumulation suggests a Toc33-dependent import pathway. Interestingly, precursors belonging to groups of either upregulated or downregulated showed differential clustering of hydroxylated amino acids in the transit sequences of precursor proteins (Vojta et al. 2004). Thus, the distribution of amino acid residues along in the transit sequences may explain how different receptors discern their favoured substrates.

This may also explain previous observations that transit peptides contain information for the preferential import into a certain plastid type (Wan et al. 1996; Yan et al. 2006). More information on precursor protein subsets that use a specific pathway and the determinants in the transit peptides for one specific pathway is needed to further substantiate this view.

7 Toc/Tic independent “alternative” import pathways into the chloroplast

N-terminal plastid transit peptides are not a general requirement for chloroplast targeting. For example most proteins, which are targeted to the outer envelope membrane of the chloroplast, do not contain a cleavable targeting sequence. Outer envelope proteins were assumed to insert spontaneously (Schleiff and Klosgen 2001; Hofmann and Theg 2005a). But there is evidence that the insertion of several outer membrane proteins like OEP14, *Physcomitrella* OEP64 (Toc64) or DGD1 depends on nucleotides and/or involves proteins at the chloroplast surface (Hofmann and Theg 2005b). The proteins involved and the mechanism of OEP targeting and insertion are largely unknown. But competition studies with a Toc/Tic import substrate hint at an involvement of Toc components in this process. In fact the insertion of OEP14 has been demonstrated to be mediated by Toc75 (Tu and Li 2000; Tu et al. 2004) and it is likely that the insertion of other OEPs depend on the Toc import channel as well.

7.1 Import depending on internal targeting sequences

A number of chloroplast proteins traversing the outer membrane without a cleavable N-terminal transit peptide such as ceQORH (chloroplast envelope quinone oxidoreductase homologue) or Tic32 (Miras et al. 2002; Nada and Soll 2004) have been identified. The targeting information of both proteins is contained in their respective mature sequences, targeting of ceQORH depends on an internal domain of 40 residues (Miras et al. 2002), targeting of Tic32 on the most N-proximal amino acids (Nada and Soll 2004).

7.2 Substrate dependent import

Another plastid protein to be mentioned in the context of alternative import pathways is the protochlorophyllide oxidoreductase A (PORA), an essential enzyme in the light-dependent etioplast to chloroplast transition. Pre-PORA has a cleavable N-terminal transit-peptide, but it is still under dispute whether it is imported via a unique pathway (Reinbothe et al. 2004) or the general Toc/Tic-dependent import pathway (Aronsson et al. 2000, 2003a, 2003b). The import of PORA was reported to depend on its substrate protochlorophyllide (Pchlide) (Reinbothe et al. 1997, 2005; Kim and Apel 2004) and to involve the outer membrane proteins OEP16 and Toc33 (Reinbothe et al. 2004). Recent studies with *Arabidopsis* mutants deficient in OEP16 and Pchlide b indicate that neither OEP16 nor Pchlide b is essential for prePORA import (Philippar et al. 2007).

These examples as well as the detection of other nucleus-encoded proteins without canonical transit peptide sequences in chloroplast proteomes (Friso et al. 2004; Kleffmann et al. 2004) point to the existence of alternative import routes into plastids, independent of the Toc/Tic import pathway (Fig. 4). One such alternative transport route may involve the secretory pathway.

7.3 Protein import via the secretory pathway

Surprisingly at first sight, proteomic studies revealed the existence of many chloroplast proteins that do not have a predicted transit peptide but a predicted signal peptide (SP) promoting ER targeting instead (Friso et al. 2004; Kleffmann et al. 2004). At the same time, a possible import route for such proteins independent of the "general" Toc/Tic pathway was discovered (Chen et al. 2004; Villarejo et al. 2005; Nanjo et al. 2006; Radhamony and Theg 2006). First evidence for a transport of proteins into higher plant plastids via a signal peptide (SP) dependent pathway came from studies by (Chen et al. 2004), who found that the SP of a rice α-amylase (αAmy3) is necessary and sufficient for its dual targeting to the extracellular compartment and to plastids. Later, two other proteins exclusively located in the chloroplast were convincingly demonstrated to traffic from the ER-Golgi system to the chloroplast (Villarejo et al. 2005; Nanjo et al. 2006). Both proteins, *Arabidopsis thaliana* α-carbonic anhydrase (CAH1) and *Oriza sativa* NPP (nucleotide pyrophosphatase/phosphodiesterase) are N-glycosylated plastidal proteins and their chloroplast accumulation is inhibited by brefeldinA – a fungal antibiotic affecting Golgi-mediated vesicular transport. In contrast accumulation of a chloroplast protein taking the Toc/Tic import pathway was not brefeldinA-sensitive (Villarejo et al. 2005). The data from the proteomics studies and the

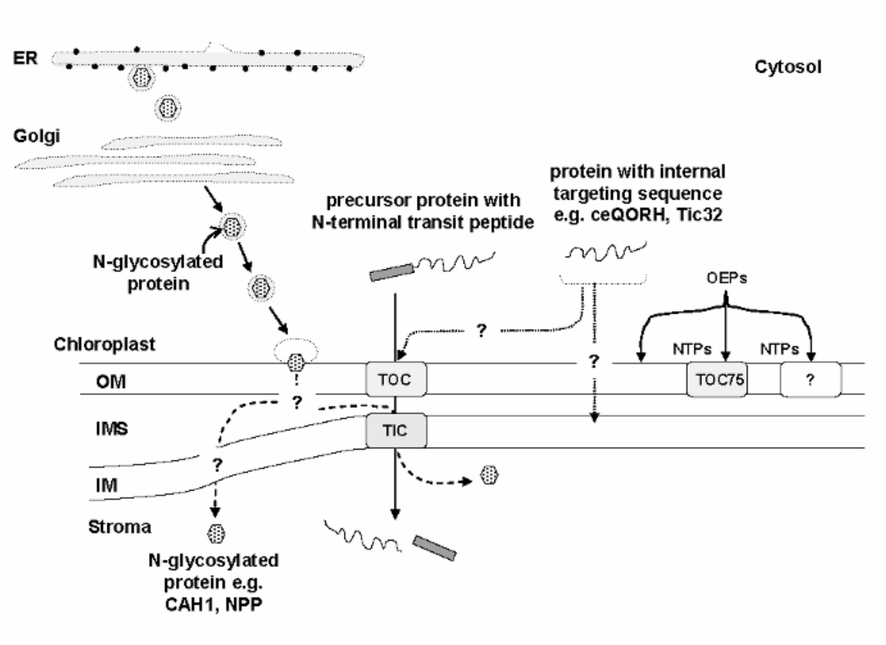


Fig. 4. Import pathways into the chloroplast (from left to right): Some stromal targeted proteins enter the endoplasmic reticulum and are transported via the Golgi apparatus to the chloroplast. During this passage they become glycosylated. They enter the chloroplast by the fusion of a Golgi-derived vesicle with the outer envelope membrane. The import routes of glycosylated proteins inside the chloroplast are not known but might involve Tic components. Precursor proteins with an N-terminal transit-peptide take the general import pathway via the Toc/Tic system. Chloroplast targeted proteins with internal targeting sequence like the inner envelope proteins ceQORH or Tic32 take a so far unknown import pathway most likely independent on the Toc/Tic system. Outer envelope proteins (OEPs) have been demonstrated to insert either spontaneously or dependent on energy (NTPs) and/or proteins in the outer membrane (e.g. Toc75).

immuno-detection of several N-glycosylated proteins in the chloroplast stroma indicates that a larger group of chloroplast proteins may be transported by a trafficking pathway involving the ER-Golgi system.

Although it is well established that the import into the so-called "complex" plastids of many algae and apicomplexan parasites occurs via the secretory pathway (Waller et al. 2000; Nassoury and Morse 2005), the involvement of SP and ER in plastid protein import of higher plants is an exciting new development in the field. Algal chloroplast protein precursors carry bipartite targeting signals consisting of the signal peptide and a stromal targeting domain (Sulli et al. 1999; Kilian and Kroth 2005) suggesting successive action of the secretory and the Toc/Tic pathways. In contrast the newly identified higher plant proteins predicted to use the secretory pathway only bear the signal peptide, suggesting a mechanism diverging from that in algae.

We predict that future research will unravel the components of glycosylated protein trafficking to the chloroplast including those involved in translocation across the envelopes. The mechanisms of the Toc-GTPases, regulation of import and the ATP-driven energetics are still far from being completely resolved and will remain major topics in the field in the next years.

Acknowledgements

We thank the members of the Kessler lab for their support, in particular Sibylle Infanger, Meryll Martin, Gwendoline Rahim and Jana Smutny. The chloroplast import project is supported by SNF grant 3100AO-109667 to FK and in part by the NCCR "Plant Survival".

References

Aronsson H, Combe J, Jarvis P (2003b) Unusual nucleotide-binding properties of the chloroplast protein import receptor, atToc33. FEBS Lett 544:79-85

Aronsson H, Combe J, Patel R, Jarvis P (2006) *In vivo* assessment of the significance of phosphorylation of the *Arabidopsis* chloroplast protein import receptor, atToc33. FEBS Lett 580:649-655

Aronsson H, Sohrt K, Soll J (2000) NADPH:Protochlorophyllide oxidoreductase uses the general import route into chloroplasts. Biol Chem Hoppe Seyler 381:1263-1267

Aronsson H, Sundqvist C, Dahlin C (2003a) POR - import and membrane association of a key element in chloroplast development. Physiol Plant 118:1-9

Asano T, Yoshioka Y, Machida Y (2004) A defect in atToc159 of *Arabidopsis thaliana* causes severe defects in leaf development. Genes Genet Syst 79:207-212

Baldwin A, Wardle A, Patel R, Dudley P, Park SK, Twell D, Inoue K, Jarvis P (2005) A molecular-genetic study of the *Arabidopsis* Toc75 gene family. Plant Physiol 138:715-733

Bauer J, Chen K, Hiltbunner A, Wehrli E, Eugster M, Schnell D, Kessler F (2000) The major protein import receptor of plastids is essential for chloroplast biogenesis. Nature 403:203-207

Bauer J, Hiltbrunner A, Weibel P, Vidi PA, Alvarez-Huerta M, Smith MD, Schnell DJ, Kessler F (2002) Essential role of the G-domain in targeting of the protein import receptor atToc159 to the chloroplast outer membrane. J Cell Biol 159:845-854

Becker T (2005) Preprotein recognition and translocation by the Toc complex - Dissertation der Fakultät für Biologie der Ludwig-Maximilians-Universität München.

Becker T, Hritz J, Vogel M, Caliebe A, Bukau B, Soll J, Schleiff E (2004b) Toc12, a novel subunit of the intermembrane space preprotein translocon of chloroplasts. Mol Biol Cell 15:5130-5144

Becker T, Jelic M, Vojta A, Radunz A, Soll J, Schleiff E (2004a) Preprotein recognition by the Toc complex. EMBO J 23:520-530

Bedard J, Jarvis P (2005) Recognition and envelope translocation of chloroplast preproteins. J Exp Bot 56:2287-2320

Bhushan S, Kuhn C, Berglund AK, Roth C, Glaser E (2006) The role of the N-terminal domain of chloroplast targeting peptides in organellar protein import and miss-sorting. FEBS Lett 580:3966-3972

Bolter B, May T, Soll J (1998) A protein import receptor in pea chloroplasts, Toc86, is only a proteolytic fragment of a larger polypeptide. FEBS Lett 441:59-62

Bourne HR, Sanders DA, McCormick F (1990) The GTPase superfamily: a conserved switch for diverse cell functions. Nature 348:125-132

Bruce BD (2000) Chloroplast transit peptides: structure, function and evolution. Trends Cell Biol 10:440-447

Bruce BD (2001) The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. Biochim Biophys Acta 1541:2-21

Caliebe A, Grimm R, Kaiser G, Lubeck J, Soll J, Heins L (1997) The chloroplastic protein import machinery contains a Rieske-type iron-sulfur cluster and a mononuclear iron-binding protein. EMBO J 16:7342-7350

Chen K, Chen X, Schnell DJ (2000a) Initial binding of preproteins involving the Toc159 receptor can be bypassed during protein import into chloroplasts. Plant Physiol 122:813-822

Chen KY, Li HM (2007) Precursor binding to an 880-kDa Toc complex as an early step during active import of protein into chloroplasts. Plant J 49:149-158

Chen MH, Huang LF, Li HM, Chen YR, Yu SM (2004) Signal peptide-dependent targeting of a rice alpha-amylase and cargo proteins to plastids and extracellular compartments of plant cells. Plant Physiol 135:1367-1377

Chen X, Smith MD, Fitzpatrick L, Schnell DJ (2002) *In vivo* analysis of the role of atTic20 in protein import into chloroplasts. Plant Cell 14:641-654

Cherfils J, Chardin P (1999) GEFs: structural basis for their activation of small GTP-binding proteins. Trends Biochem Sci 24:306-311

Chew O, Lister R, Qbadou S, Heazlewood JL, Soll J, Schleiff E, Millar AH, Whelan J (2004) A plant outer mitochondrial membrane protein with high amino acid sequence identity to a chloroplast protein import receptor. FEBS Lett 557:109-114

Chigri F, Hormann F, Stamp A, Stammers DK, Bolter B, Soll J, Vothknecht UC (2006) Calcium regulation of chloroplast protein translocation is mediated by calmodulin binding to Tic32. Proc Natl Acad Sci U S A 103:16051-16056

Chigri F, Soll J, Vothknecht UC (2005) Calcium regulation of chloroplast protein import. Plant J 42:821-831

Chou ML, Chu CC, Chen LJ, Akita M, Li HM (2006) Stimulation of transit-peptide release and ATP hydrolysis by a cochaperone during protein import into chloroplasts. J Cell Biol 175:893-900

Chou ML, Fitzpatrick LM, Tu SL, Budziszewski G, Potter-Lewis S, Akita M, Levin JZ, Keegstra K, Li HM (2003) Tic40, a membrane-anchored co-chaperone homolog in the chloroplast protein translocon. EMBO J 22:2970-2980

Constan D, Froehlich JE, Rangarajan S, Keegstra K (2004b) A stromal Hsp100 protein is required for normal chloroplast development and function in *Arabidopsis*. Plant Physiol 136:3605-3615

Constan D, Patel R, Keegstra K, Jarvis P (2004a) An outer envelope membrane component of the plastid protein import apparatus plays an essential role in *Arabidopsis*. Plant J 38:93-106

Duy D, Wanner G, Meda AR, von Wiren N, Soll J, Philippar K (2007) PIC1, an ancient permease in *Arabidopsis* chloroplasts, mediates iron transport. Plant Cell 19:986-1006

Friso G, Giacomelli L, Ytterberg AJ, Peltier JB, Rudella A, Sun Q, Wijk KJ (2004) In-depth analysis of the thylakoid membrane proteome of *Arabidopsis thaliana* chloroplasts: new proteins, new functions, and a plastid proteome database. Plant Cell 16:478-499

Fulgosi H, Soll J (2002) The chloroplast protein import receptors Toc34 and Toc159 are phosphorylated by distinct protein kinases. J Biol Chem 277:8934-8940

Gentle IE, Burri L, Lithgow T (2005) Molecular architecture and function of the Omp85 family of proteins. Mol Microbiol 58:1216-1225

Gutensohn M, Fan E, Frielingsdorf S, Hanner P, Hou B, Hust B, Klosgen RB (2006) Toc, Tic, Tat et al.: structure and function of protein transport machineries in chloroplasts. J Plant Physiol 163:333-347

Gutensohn M, Pahnke S, Kolukisaoglu U, Schulz B, Schierhorn A, Voigt A, Hust B, Rollwitz I, Stockel J, Geimer S, Albrecht V, Flugge UI, Klosgen RB (2004) Characterization of a T-DNA insertion mutant for the protein import receptor atToc33 from chloroplasts. Mol Genet Genomics 272:379-396

Gutensohn M, Schulz B, Nicolay P, Flugge UI (2000) Functional analysis of the two *Arabidopsis* homologues of Toc34, a component of the chloroplast protein import apparatus. Plant J 23:771-783

Heins L, Mehrle A, Hemmler R, Wagner R, Kuchler M, Hormann F, Sveshnikov D, Soll J (2002) The preprotein conducting channel at the inner envelope membrane of plastids. EMBO J 21:2616-2625

Hiltbrunner A, Bauer J, Alvarez-Huerta M, Kessler F (2001a) Protein translocon at the *Arabidopsis* outer chloroplast membrane. Biochem Cell Biol 79:629-635

Hiltbrunner A, Bauer J, Vidi PA, Infanger S, Weibel P, Hohwy M, Kessler F (2001b) Targeting of an abundant cytosolic form of the protein import receptor at Toc159 to the outer chloroplast membrane. J Cell Biol 154:309-316

Hiltbrunner A, Grunig K, Alvarez-Huerta M, Infanger S, Bauer J, Kessler F (2004) AtToc90, a new GTP-binding component of the *Arabidopsis* chloroplast protein import machinery. Plant Mol Biol 54:427-440

Hinnah SC, Hill K, Wagner R, Schlicher T, Soll J (1997) Reconstitution of a chloroplast protein import channel. EMBO J 16:7351-7360

Hinnah SC, Wagner R, Sveshnikova N, Harrer R, Soll J (2002) The chloroplast protein import channel Toc75: pore properties and interaction with transit peptides. Biophys J 83:899-911

Hirohashi T, Hase T, Nakai M (2001) Maize non-photosynthetic ferredoxin precursor is mis-sorted to the intermembrane space of chloroplasts in the presence of light. Plant Physiol 125:2154-2163

Hirsch S, Muckel E, Heemeyer F, von Heijne G, Soll J (1994) A receptor component of the chloroplast protein translocation machinery. Science 266:1989-1992

Hofmann NR, Theg SM (2005a) Chloroplast outer membrane protein targeting and insertion. Trends Plant Sci 10:450-457

Hofmann NR, Theg SM (2005b) Protein- and energy-mediated targeting of chloroplast outer envelope membrane proteins. Plant J 44:917-927

Hormann F, Kuchler M, Sveshnikov D, Oppermann U, Li Y, Soll J (2004) Tic32, an essential component in chloroplast biogenesis. J Biol Chem 279:34756-34762

Horniak L, Pilon M, van 't Hof R, de Kruijff B (1993) The secondary structure of the ferredoxin transit sequence is modulated by its interaction with negatively charged lipids. FEBS Lett 334:241-246

Inaba T, Alvarez-Huerta M, Li M, Bauer J, Ewers C, Kessler F, Schnell DJ (2005) *Arabidopsis* tic110 is essential for the assembly and function of the protein import machinery of plastids. Plant Cell 17:1482-1496

Inaba T, Li M, Alvarez-Huerta M, Kessler F, Schnell DJ (2003) atTic110 functions as a scaffold for coordinating the stromal events of protein import into chloroplasts. J Biol Chem 278:38617-38627

Ivanova Y, Smith MD, Chen K, Schnell DJ (2004) Members of the Toc159 import receptor family represent distinct pathways for protein targeting to plastids. Mol Biol Cell 15:3379-3392

Ivey RA, 3rd, Bruce BD (2000) *In vivo* and *in vitro* interaction of DnaK and a chloroplast transit peptide. Cell Stress Chaperones 5:62-71

Jackson-Constan D, Keegstra K (2001a) *Arabidopsis* genes encoding components of the chloroplastic protein import apparatus. Plant Physiol 125:1567-1576

Jackson DT, Froehlich JE, Keegstra K (1998) The hydrophilic domain of Tic110, an inner envelope membrane component of the chloroplastic protein translocation apparatus, faces the stromal compartment. J Biol Chem 273:16583-16588

Jarvis P, Chen LJ, Li H, Peto CA, Fankhauser C, Chory J (1998) An *Arabidopsis* mutant defective in the plastid general protein import apparatus. Science 282:100-103

Jarvis P, Robinson C (2004) Mechanisms of protein import and routing in chloroplasts. Curr Biol 14:R1064-1077

Jelic M, Soll J, Schleiff E (2003) Two Toc34 homologues with different properties. Biochemistry 42:5906-5916

Jelic M, Sveshnikova N, Motzkus M, Horth P, Soll J, Schleiff E (2002) The chloroplast import receptor Toc34 functions as preprotein-regulated GTPase. Biol Chem Hoppe Seyler 383:1875-1883

Kessler F, Blobel G (1996) Interaction of the protein import and folding machineries of the chloroplast. Proc Natl Acad Sci USA 93:7684-7689

Kessler F, Blobel G, Patel HA, Schnell DJ (1994) Identification of two GTP-binding proteins in the chloroplast protein import machinery. Science 266:1035-1039

Kessler F, Schnell DJ (2002) A GTPase gate for protein import into chloroplasts. Nat Struct Biol 9:81-83

Kessler F, Schnell DJ (2004) Chloroplast protein import: solve the GTPase riddle for entry. Trends Cell Biol 14:334-338

Kikuchi S, Hirohashi T, Nakai M (2006) Characterization of the preprotein translocon at the outer envelope membrane of chloroplasts by blue native PAGE. Plant Cell Physiol 47:363-371

Kilian O, Kroth PG (2005) Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids. Plant J 41:175-183

Kim C, Apel K (2004) Substrate-dependent and organ-specific chloroplast protein import in planta. Plant Cell 16:88-98

Kleffmann T, Russenberger D, von Zychlinski A, Christopher W, Sjolander K, Gruissem W, Baginsky S (2004) The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. Curr Biol 14:354-362

Kouranov A, Chen X, Fuks B, Schnell DJ (1998) Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. J Cell Biol 143:991-1002

Kouranov A, Schnell DJ (1997) Analysis of the interactions of preproteins with the import machinery over the course of protein import into chloroplasts. J Cell Biol 139:1677-1685

Kovacheva S, Bedard J, Patel R, Dudley P, Twell D, Rios G, Koncz C, Jarvis P (2005) *In vivo* studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. Plant J 41:412-428

Kovacheva S, Bedard J, Wardle A, Patel R, Jarvis P (2007) Further *in vivo* studies on the role of the molecular chaperone, Hsp93, in plastid protein import. Plant J 50:364-379

Kubis S, Baldwin A, Patel R, Razzaq A, Dupree P, Lilley K, Kurth J, Leister D, Jarvis P (2003) The *Arabidopsis* ppi1 mutant is specifically defective in the expression, chloroplast import, and accumulation of photosynthetic proteins. Plant Cell 15:1859-1871

Kubis S, Patel R, Combe J, Bedard J, Kovacheva S, Lilley K, Biehl A, Leister D, Rios G, Koncz C, Jarvis P (2004) Functional specialization amongst the *Arabidopsis* Toc159 family of chloroplast protein import receptors. Plant Cell 16:2059-2077

Kuchler M, Decker S, Hormann F, Soll J, Heins L (2002) Protein import into chloroplasts involves redox-regulated proteins. EMBO J 21:6136-6145

Kutschera U, Niklas KJ (2005) Endosymbiosis, cell evolution, and speciation. Theory Biosci 124:1-24

Lee KH, Kim SJ, Lee YJ, Jin JB, Hwang I (2003) The M domain of atToc159 plays an essential role in the import of proteins into chloroplasts and chloroplast biogenesis. J Biol Chem 278:36794-36805

Leister D, Schneider A (2003) From genes to photosynthesis in *Arabidopsis thaliana*. Int Rev Cytol 228:31-83

Lopez-Juez E, Pyke KA (2005) Plastids unleashed: their development and their integration in plant development. Int J Dev Biol 49:557-577

Lubeck J, Soll J, Akita M, Nielsen E, Keegstra K (1996) Topology of IEP110, a component of the chloroplastic protein import machinery present in the inner envelope membrane. EMBO J 15:4230-4238

Ma Y, Kouranov A, LaSala SE, Schnell DJ (1996) Two components of the chloroplast protein import apparatus, IAP86 and IAP75, interact with the transit sequence during the recognition and translocation of precursor proteins at the outer envelope. J Cell Biol 134:315-327

Martin T, Sharma R, Sippel C, Waegemann K, Soll J, Vothknecht UC (2006) A protein kinase family in *Arabidopsis* phosphorylates chloroplast precursor proteins. J Biol Chem 281:40216-40223

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci USA 99:12246-12251

May T, Soll J (2000) 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. Plant Cell 12:53-64

McFadden GI (2001) Chloroplast origin and integration. Plant Physiol 125:50-53

Miras S, Salvi D, Ferro M, Grunwald D, Garin J, Joyard J, Rolland N (2002) Non-canonical transit peptide for import into the chloroplast. J Biol Chem 277:47770-47778

Nada A, Soll J (2004) Inner envelope protein 32 is imported into chloroplasts by a novel pathway. J Cell Sci 117:3975-3982

Nakrieko KA, Mould RM, Smith AG (2004) Fidelity of targeting to chloroplasts is not affected by removal of the phosphorylation site from the transit peptide. Eur J Biochem 271:509-516

Nanjo Y, Oka H, Ikarashi N, Kaneko K, Kitajima A, Mitsui T, Munoz FJ, Rodriguez-Lopez M, Baroja-Fernandez E, Pozueta-Romero J (2006) Rice plastidial N-glycosylated nucleotide pyrophosphatase/phosphodiesterase is transported from the ER-golgi to the chloroplast through the secretory pathway. Plant Cell 18:2582-2592

Nassoury N, Morse D (2005) Protein targeting to the chloroplasts of photosynthetic eukaryotes: getting there is half the fun. Biochim Biophys Acta 1743:5-19

Olsen LJ, Keegstra K (1992) The binding of precursor proteins to chloroplasts requires nucleoside triphosphates in the intermembrane space. J Biol Chem 267:433-439

Pain D, Blobel G (1987) Protein import into chloroplasts requires a chloroplast ATPase. Proc Natl Acad Sci USA 84:3288-3292

Paschen SA, Neupert W, Rapaport D (2005) Biogenesis of beta-barrel membrane proteins of mitochondria. Trends Biochem Sci 30:575-582

Perry SE, Keegstra K (1994) Envelope membrane proteins that interact with chloroplastic precursor proteins. Plant Cell 6:93-105

Philippar K, Geis T, Ilkavets I, Oster U, Schwenkert S, Meurer J, Soll J (2007) Chloroplast biogenesis: The use of mutants to study the etioplast-chloroplast transition. Proc Natl Acad Sci USA 104:678-683

Pollmann S, Neu D, Lehmann T, Berkowitz O, Schafer T, Weiler EW (2006) Subcellular localization and tissue specific expression of amidase 1 from *Arabidopsis thaliana*. Planta 224:1241-1253

Pollmann S, Neu D, Weiler EW (2003) Molecular cloning and characterization of an amidase from *Arabidopsis thaliana* capable of converting indole-3-acetamide into the plant growth hormone, indole-3-acetic acid. Phytochemistry 62:293-300

Qbadou S, Becker T, Mirus O, Tews I, Soll J, Schleiff E (2006) The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. EMBO J 25:1836-1847

Radhamony RN, Theg SM (2006) Evidence for an ER to Golgi to chloroplast protein transport pathway. Trends Cell Biol 16:385-387

Reddick LE, Vaughn MD, Wright SJ, Campbell IM, Bruce BD (2007) *In vitro* comparative kinetic analysis of the chloroplast toc GTPases. J Biol Chem 282:11410-11426

Reinbothe C, Lebedev N, Apel K, Reinbothe S (1997) Regulation of chloroplast protein import through a protochlorophyllide-responsive transit peptide. Proc Natl Acad Sci USA 94:8890-8894

Reinbothe S, Pollmann S, Springer A, James RJ, Tichtinsky G, Reinbothe C (2005) A role of Toc33 in the protochlorophyllide-dependent plastid import pathway of NADPH:protochlorophyllide oxidoreductase (POR) A. Plant J 42:1-12

Reinbothe S, Quigley F, Gray J, Schemenewitz A, Reinbothe C (2004) Identification of plastid envelope proteins required for import of protochlorophyllide oxidoreductase A into the chloroplast of barley. Proc Natl Acad Sci USA 101:2197-2202

Reuber TL, Ausubel FM (1996) Isolation of *Arabidopsis* genes that differentiate between resistance responses mediated by the RPS2 and RPM1 disease resistance genes. Plant Cell 8:241-249

Reumann S, Inoue K, Keegstra K (2005) Evolution of the general protein import pathway of plastids (review). Mol Membr Biol 22:73-86

Reumann S, Keegstra K (1999) The endosymbiotic origin of the protein import machinery of chloroplastic envelope membranes. Trends Plant Sci 4:302-307
Rial DV, Arakaki AK, Ceccarelli EA (2000) Interaction of the targeting sequence of chloroplast precursors with Hsp70 molecular chaperones. Eur J Biochem 267:6239-6248
Richter S, Lamppa GK (2003) Structural properties of the chloroplast stromal processing peptidase required for its function in transit peptide removal. J Biol Chem 278:39497-39502
Scheffzek K, Ahmadian MR (2005) GTPase activating proteins: structural and functional insights 18 years after discovery. Cell Mol Life Sci 62:3014-3038
Schleiff E, Jelic M, Soll J (2003b) A GTP-driven motor moves proteins across the outer envelope of chloroplasts. Proc Natl Acad Sci USA 100:4604-4609
Schleiff E, Klosgen RB (2001) Without a little help from 'my' friends: direct insertion of proteins into chloroplast membranes? Biochim Biophys Acta 1541:22-33
Schleiff E, Soll J (2005) Membrane protein insertion: mixing eukaryotic and prokaryotic concepts. EMBO Rep 6:1023-1027
Schleiff E, Soll J, Kuchler M, Kuhlbrandt W, Harrer R (2003a) Characterization of the translocon of the outer envelope of chloroplasts. J Cell Biol 160:541-551
Schnell DJ, Blobel G (1993) Identification of intermediates in the pathway of protein import into chloroplasts and their localization to envelope contact sites. J Cell Biol 120:103-115
Schnell DJ, Kessler F, Blobel G (1994) Isolation of components of the chloroplast protein import machinery. Science 266:1007-1012
Seedorf M, Waegemann K, Soll J (1995) A constituent of the chloroplast import complex represents a new type of GTP-binding protein. Plant J 7:401-411
Siddique MA, Grossmann J, Gruissem W, Baginsky S (2006) Proteome analysis of bell pepper (*Capsicum annuum* L.) Chromoplasts. Plant Cell Physiol 47:1663-1673
Sjogren LL, MacDonald TM, Sutinen S, Clarke AK (2004) Inactivation of the clpC1 gene encoding a chloroplast Hsp100 molecular chaperone causes growth retardation, leaf chlorosis, lower photosynthetic activity, and a specific reduction in photosystem content. Plant Physiol 136:4114-4126
Smith MD, Hiltbrunner A, Kessler F, Schnell DJ (2002) The targeting of the atToc159 preprotein receptor to the chloroplast outer membrane is mediated by its GTPase domain and is regulated by GTP. J Cell Biol 159:833-843
Smith MD, Rounds CM, Wang F, Chen K, Afitlhile M, Schnell DJ (2004) atToc159 is a selective transit peptide receptor for the import of nucleus-encoded chloroplast proteins. J Cell Biol 165:323-334
Sohrt K, Soll J (2000) Toc64, a new component of the protein translocon of chloroplasts. J Cell Biol 148:1213-1221
Soll J (2002) Protein import into chloroplasts. Curr Opin Plant Biol 5:529-535
Stahl T, Glockmann C, Soll J, Heins L (1999) Tic40, a new "old" subunit of the chloroplast protein import translocon. J Biol Chem 274:37467-37472
Sulli C, Fang Z, Muchhal U, Schwartzbach SD (1999) Topology of Euglena chloroplast protein precursors within endoplasmic reticulum to Golgi to chloroplast transport vesicles. J Biol Chem 274:457-463
Sun CW, Chen LJ, Lin LC, Li HM (2001) Leaf-specific upregulation of chloroplast translocon genes by a CCT motif-containing protein, CIA 2. Plant Cell 13:2053-2061

Sun YJ, Forouhar F, Li Hm HM, Tu SL, Yeh YH, Kao S, Shr HL, Chou CC, Chen C, Hsiao CD (2002) Crystal structure of pea Toc34, a novel GTPase of the chloroplast protein translocon. Nat Struct Biol 9:95-100

Sveshnikova N, Soll J, Schleiff E (2000) Toc34 is a preprotein receptor regulated by GTP and phosphorylation. Proc Natl Acad Sci USA 97:4973-4978

Teng YS, Su YS, Chen LJ, Lee YJ, Hwang I, Li HM (2006) Tic21 is an essential translocon component for protein translocation across the chloroplast inner envelope membrane. Plant Cell 18:2247-2257

Theg SM, Bauerle C, Olsen LJ, Selman BR, 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

Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet 5:123-135

Tu SL, Chen LJ, Smith MD, Su YS, Schnell DJ, Li HM (2004) Import pathways of chloroplast interior proteins and the outer-membrane protein OEP14 converge at Toc75. Plant Cell 16:2078-2088

Tu SL, Li HM (2000) Insertion of OEP14 into the outer envelope membrane is mediated by proteinaceous components of chloroplasts. Plant Cell 12:1951-1960

van den Wijngaard PW, Vredenberg WJ (1999) The envelope anion channel involved in chloroplast protein import is associated with Tic110. J Biol Chem 274:25201-25204

van Wijk KJ (2004) Plastid proteomics. Plant Physiol Biochem 42:963-977

Villarejo A, Buren S, Larsson S, Dejardin A, Monne M, Rudhe C, Karlsson J, Jansson S, Lerouge P, Rolland N, von Heijne G, Grebe M, Bako L, Samuelsson G (2005) Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. Nat Cell Biol 7:1224-1231

Vojta A, Alavi M, Becker T, Hormann F, Kuchler M, Soll J, Thomson R, Schleiff E (2004) The protein translocon of the plastid envelopes. J Biol Chem 279:21401-21405

Waegemann K, Soll J (1996) Phosphorylation of the transit sequence of chloroplast precursor proteins. J Biol Chem 271:6545-6554

Wallas TR, Smith MD, Sanchez-Nieto S, Schnell DJ (2003) The roles of toc34 and toc75 in targeting the toc159 preprotein receptor to chloroplasts. J Biol Chem 278:44289-44297

Waller RF, Reed MB, Cowman AF, McFadden GI (2000) Protein trafficking to the plastid of Plasmodium falciparum is via the secretory pathway. EMBO J 19:1794-1802

Wan J, Blakeley SD, Dennis DT, Ko K (1996) Transit peptides play a major role in the preferential import of proteins into leucoplasts and chloroplasts. J Biol Chem 271:31227-31233

Weibel P, Hiltbrunner A, Brand L, Kessler F (2003) Dimerization of Toc-GTPases at the chloroplast protein import machinery. J Biol Chem 278:37321-37329

Wu C, Seibert FS, Ko K (1994) Identification of chloroplast envelope proteins in close physical proximity to a partially translocated chimeric precursor protein. J Biol Chem 269:32264-32271

Yan X, Khan S, Hase T, Emes MJ, Bowsher CG (2006) Differential uptake of photosynthetic and non-photosynthetic proteins by pea root plastids. FEBS Lett 580:6509-6512

Yeh YH, Kesavulu MM, Li HM, Wu SZ, Sun YJ, Konozy EH, Hsiao CD (2007) Dimerization is important for the GTPase activity of chloroplast translocon components atToc33 and psToc159. J Biol Chem 282:13845-13853

Young JC, Hoogenraad NJ, Hartl FU (2003) Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. Cell 112:41-50

Young ME, Keegstra K, Froehlich JE (1999) GTP promotes the formation of early-import intermediates but is not required during the translocation step of protein import into chloroplasts. Plant Physiol 121:237-244

Yu TS, Li H (2001) Chloroplast protein translocon components atToc159 and atToc33 are not essential for chloroplast biogenesis in guard cells and root cells. Plant Physiol 127:90-96

Zhang XP, Glaser E (2002) Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. Trends Plant Sci 7:14-21

Zhong R, Wan J, Jin R, Lamppa G (2003) A pea antisense gene for the chloroplast stromal processing peptidase yields seedling lethals in *Arabidopsis*: survivors show defective GFP import *in vivo*. Plant J 34:802-812

Agne, Birgit
Laboratoire de Physiologie Végétale, Institut de Biologie, Université de Neuchâtel, Rue Emile-Argand 11, 2009 Neuchâtel, Switzerland

Kessler, Felix
Laboratoire de Physiologie Végétale, Institut de Biologie, Université de Neuchâtel, Rue Emile-Argand 11, 2009 Neuchâtel, Switzerland
felix.kessler@unine.ch

Insights into chloroplast proteomics: from basic principles to new horizons

Bianca Naumann and Michael Hippler

Abstract

Many proteomic approaches have been employed to investigate the complex and dynamic proteome of the chloroplast. These range from classical methods like one and two dimensional gel electrophoresis to advanced comparative proteomics strategies such as ICAT or SILAC. Mass spectrometry for protein identification or quantitation plays an important role in most of the methods used and is a fast emerging technology in protein biochemistry. Most proteomic studies of the chloroplast focus on the single compartments of this plant organelle, which greatly reduces the complexity of the sample and thus allows for a more complete and detailed analysis of the complex protein composition. The rapidly developing field of comparative proteomics makes it possible to analyze dynamic protein changes caused, for example, by different developmental stages of a plant, by various stress conditions and distinct genetic backgrounds.

1 The art of proteomics

A great challenge of the post genomic era is to understand how genetic information results in the concerted and dynamic action of gene products to generate function. In contrast to a cell's static genome, the proteome is both complex and dynamic. The proteome is defined as the set of all expressed proteins in a cell, tissue or organism (Wilkins et al. 1999). Proteomics can be defined as the systematic analysis of proteins for their identity, abundance, expression pattern, and function. Proteomics permits a global view on dynamics of biological processes by the systematic analysis of expressed proteins and, in particular, of functional protein complexes. The analysis of a proteome is complicated by the fact that the expressed product of a single gene often represents a protein population that may contain a large amount of micro-heterogeneity. Post-translational modifications (PTM), like phosphorylation, acetylation, glycosylation, protease cleavage, lipidation, or ubiquitination may contribute to the expression profile of a protein. The analysis of such complex protein profiles requires methods that allow high resolution protein separation combined with very sensitive methods for protein identification. Mass spectrometry (MS) has become a powerful tool for peptide and protein identification since it allows sensitive, fast and specific measurement, and

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0224 / Published online: 15 May 2007

thus allows for recognition of peptides and proteins from complex mixtures (Aebersold and Mann 2003; Domon and Aebersold 2006). A typical workflow for proteomic experiments is depicted in Figure 1. Besides protein identification, recognition of post-translational modifications and protein quantitation are important tasks that can be investigated by mass spectrometric experiments. Today whole suites of potent mass spectrometer (MS) are available to fulfill these tasks (Aebersold and Mann 2003; Domon and Aebersold 2006). It is not our aim to discuss the distinct mass spectrometer options available. In the beginning of this review, we would rather like to address issues that are essential for successful mass spectrometric experiments: (i) peptide ionization allowing the entry of peptides into the mass spectrometer, (ii) peptide mass finger printing (PMF) and tandem mass spectrometry (MS/MS), and (iii) algorithms that permit identification of peptides and in turn proteins from mass spectrometric data.

1.1 Prerequisite for biomolecular mass spectrometry: MALDI and ESI Ionization

To enable mass spectrometric analysis of peptide-molecules, they have to be ionized before they can enter the mass spectrometer. The most common ionization methods for biomolecular mass spectrometry are matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI). MALDI is based on the bombardment of sample molecules with a laser light to induce ionization. The sample is pre-mixed with a highly light-absorbing matrix compound. The matrix absorbs the laser energy and transforms it into excitation energy. This leads to the sputtering of matrix molecules, which drag along the analyte ions from the surface of the mixture, and enables the entry of ionized molecules from an intermediate vacuum region into the analyzer of the mass spectrometer, which is under permanent high vacuum. During electrospray ionization, the sample is dissolved in a polar solvent and pumped through a narrow capillary. A high voltage of 3 to 4 kV is usually applied in between the tip of the capillary, which is positioned within the ionization source of the mass spectrometer and the aperture, which represents the entry point to the high vacuum system. In response to this strong electric field, the sample emerging from the tip is dispersed into an aerosol of highly charged droplets. Ultimately, charged sample ions, free from solvent, are released from the droplets, which pass through a sampling cone into an intermediate vacuum region, and from there through a small aperture into the analyzer of the mass spectrometer.

1.2 Peptide mass finger printing and tandem mass spectrometry

MALDI coupled to time-of-flight (TOF) mass spectrometer instruments are commonly used for large-scale protein identification by the peptide mass mapping technique. Peptide masses are determined for specific spots on the analyzer plate by MALDI-MS and these mass maps are then compared to predicted mass maps

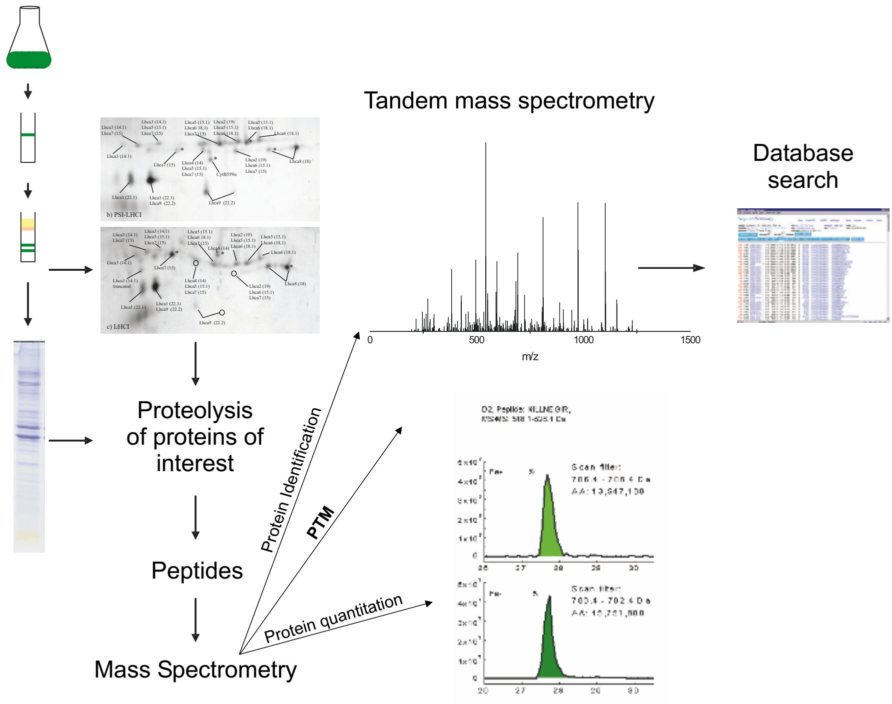


Fig. 1. Typical workflow of a proteomics experiment. Chloroplasts can for example be extracted by centrifugation on a Percoll gradient and than fractionated into their compartments by sucrose gradient centrifugation. Protein samples can than be separated by 1- or 2-dimensional gel electrophoresis and further analyzed with mass spectrometry or tandem mass spectrometry in respect to identification, post-translational modification and quantitation.

in a database to identify the respective protein. MALDI can also be employed to ionize peptides for entry into more complex mass spectrometer, enabling tandem mass spectrometric analyses, such as ion trap, TOF-TOF or quadrupole-TOF mass spectrometer. In contrast to MALDI, ESI allows direct coupling of liquid chromatography (LC) systems to the mass spectrometer. This permits the combination of chromatographic separation of peptides (using on average nano flow rates) and direct elution into the mass spectrometer coupled to mass spectrometric and tandem mass spectrometric analysis of peptides. Besides mass information of the peptide-ion, tandem mass spectrometry produces structural information about specifically selected peptide ions inside the mass spectrometer. MS/MS experiments are performed by colliding a selected ion with inert gas molecules such as argon or helium and subsequent mass measurement of the fragment ions yielding fragmentation mass spectra. Importantly, fragmentation occurs preferably at the peptide bonds resulting in y- and b-type ions that represent fragment ions harboring either C- or N-terminus, respectively. This information can then be assembled to generate structural information regarding the intact molecule and enable direct amino acid sequencing of peptides. To identify peptides from mass spectrometric and tandem mass spectrometric data, algorithms are available that take advantage of protein and DNA database information to correlate peptide sequences with mass spectrometric data information. In addition *de novo* amino acid sequencing from MS/MS data is feasible. Since today's mass spectrometer become more and more sensitive and faster in data recording, the bottleneck of the mass spectrometric experiment seems to be the evaluation of these data.

1.3 Database searching

A set of distinct peptide masses obtained from proteolytic cleavage (mostly tryptic cleavage) of a protein and subsequent mass spectrometric mass measurement can be used to identify proteins. In this approach such a set of peptide masses is mapped against an *in silico* digest of a protein sequence database. The approach, called peptide mass fingerprinting (Mann et al. 1993; Giddings et al. 2003), is however vulnerable to the complexity of the mixture. With the increase of peptide species in the mixture the possible combinations increase exponentially, thus, making a correct protein assignment difficult. In addition peptide masses may not be unique in sequence databases. Therefore, identification of a protein via PMF depends critically on the mass accuracy and the mass resolution of the mass spectrometer.

The introduction of MS/MS spectra for peptide and protein identification included the fragmentation pattern of a peptide as a supplementary constraint in addition to the peptide mass, thus, rendering the peptide identification and in turn the protein identification more reliable. Sequence tags (Mann and Wilm 1994; Shevchenko et al. 1996) amend the mass of the peptide, as in PMF, with a short partial amino acid sequence, which is determined from the spectrum and its position within the peptide. Thus, four parameters define a sequence tag: a) its mass, b) its partial amino acid sequence, c) the mass before the start of the partial amino acid sequence in the peptide, and d) the remaining mass after the end of the partial amino acid sequence within the peptide. Partial sequences can be searched for in sequence databases, usually presented as plain text-files in fasta-format (Pearson and Lipman 1988). Furthermore, the masses of the resulting fragments of the *in silico* digest of these files are used to filter the results. Another filter is presented through the positioning of the partial amino acid sequence. These three filters are very restrictive and more discriminating than searching with mere masses alone. Therefore, it is in widespread use today, with new developments reported regularly (Bafna and Edwards 2001; Sunyaev et al. 2003; Tabb et al. 2003; Savitski et al. 2005).

Another approach developed around the same time as sequence tagging makes use of the complete MS/MS spectrum (Eng et al. 1994). It uses cross correlation to compare the acquired mass spectra to theoretically derived spectra from sequences in a database. This algorithm, named Sequest, along with Mascot (Perkins et al. 1999), which employs sequence search, ion search, PMF, and introduces a probabilistic based scoring scheme for the first time, are the so called industry standards for software in this area today. Besides Sequest and Mascot, numerous other tools that match mass spectrometric data to sequence databases are available (for reviews see Kapp et al. 2005; Shadforth et al. 2005). Although database search is able to identify peptides from complex mixtures, it obviously fails if there is no database available, or if other obstacles hinder the identification. In these cases, *de novo* amino acid sequencing may be of use.

1.4 *De Novo* sequencing

De novo sequencing algorithms seek to determine the underlying peptide sequence from the mass spectrometric information alone. The rational behind this approach is that peptides dissociate into predictable fragments. Looking at y-ions alone clearly shows that the difference in-between two consecutive y-ions in a spectrum represents the mass of one or multiple amino acid. Other ion-types may provide additional and supporting information in this scenario. The best case occurs when a complete fragment ion ladder of at least one ion-type is present. The inherent problem in *de novo* sequencing is, however, that it is not known which peak represents which ion-type in a given MS/MS spectrum *a priori*.

A number of *de novo* amino acid sequencing programs have been described and are in use today (Dancik et al. 1999; Fernandez-de-Cossio et al. 2000; Chen et al. 2001; Taylor and Johnson 2001; Bafna and Edwards 2003; Ma et al. 2003). These programs face other limitations. They are usually computational intensive and dependent on high quality spectra (Spengler 2004; Yan et al. 2005). For these reasons, they are quite limited in practice. *De novo* amino acid sequencing information together with mass information could be used for error-tolerant searching of DNA and in particular genomic DNA databases. Therefore, there is a need to connect *de novo* sequencing approaches with database searching algorithms.

1.5 Linking database searching and *de novo* sequencing

The GenomicPeptideFinder (GPF) connects *de novo* sequencing with database search (Fig. 2) (Allmer et al. 2006). The aim of GPF is to employ mass spectrometric data for genomic data mining. It enables detection of intron-split and/or alternatively spliced peptides from MS/MS data when deduced from genomic DNA (Allmer et al. 2004). As depicted in Figure 2, prior to GPF search, mass spectra are submitted to *de novo* amino acid sequencing by PEAKS (Ma et al. 2003). The predictions are converted to queries for GPF and searched against the six-frame translation of a genomic DNA database. For this error tolerant search small sub sequences of the *de novo* prediction are mapped to the six-frame translation of a genomic database. The proximity of a match, usually 2100 base pairs upstream and downstream, is investigated in more detail. This time shorter sequence fragments are searched in the extended region. All matches are tried out whether they, when they are joined, define a tryptic peptide that would explain the precursor

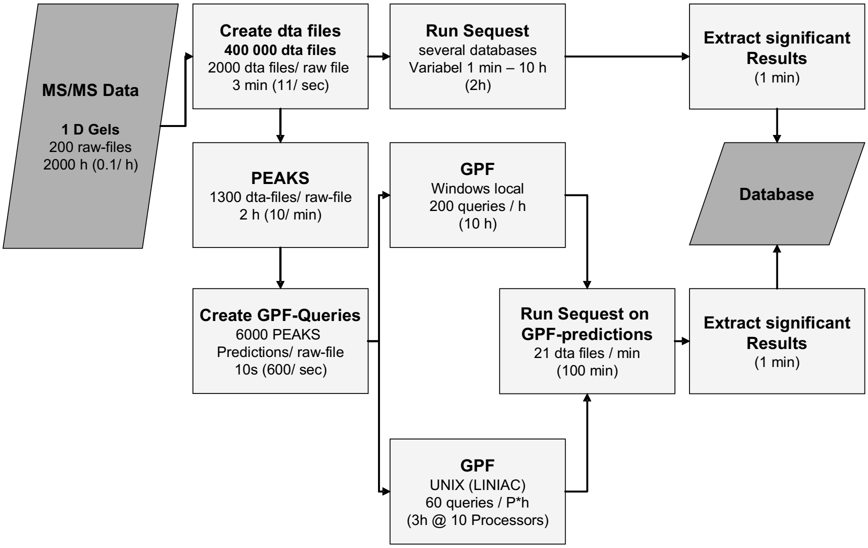


Fig. 2. Computational processing of the MS/MS spectra acquired by mass spectrometry. Certain data associated with each process such as processing time is presented in the boxes above, each box representing a distinct process. Processing times were calculated for one PC if not indicated otherwise. Most triply charged dta-files were not submitted to *de novo* prediction analysis. The GPF core is the same in both PC and UNIX distribution. Most GPF processing was done on the LINIAC Cluster (University of Pennsylvania, Philadelphia, USA). Figure taken from Allmer et al. (2006).

mass within the error of the mass spectrometer used. Joining the matched sequences allows splicing out of intervening sequences. In order to allow for sequencing errors of the *de novo* algorithm, the intervening sequence is checked along the reading frames of the bordering matches whether they can explain the precursor mass, if completely removing it, renders the mass too low. All *de novo* sequence predictions are used in this fashion. All resulting peptides are stored in a fasta-file and are submitted to database search. The new sequences function as the database and they are correlated using Sequest against the original mass spectra which gave rise to the *de novo* predictions.

This approach was used to study the thylakoid proteome of *Chlamydomonas reinhardtii* (Allmer et al. 2006). The concerted action of Sequest and GPF allowed identification of 2622 distinct peptides. In total 448 peptides were identified by GPF analysis alone including 98 intron-split peptides, resulting in the identification of novel proteins, improved annotation of gene models, and evidence of alternative splicing. It is predictable that the combination of *de novo* sequencing from MS/MS spectra in conjunction with error-tolerant GPF performance will be of help to explore nuclear gene structures and identify alternative splicing in eukaryotic organism with complex genomes.

1.6 Strategies for the analysis of proteome dynamics

Traditionally, a standard technique for proteome analysis combines protein separation by high-resolution (isoelectric focusing (IEF)/SDS-PAGE) two-dimensional gel electrophoresis (2-DE) with mass spectrometry or tandem MS identification of selected protein spots. The 2-DE technique has been used for the separation, detection and quantification of individual proteins present in a complex sample in combination with mass spectrometry and database searching for the identification of the separated proteins (as reviewed in Aebersold and Mann 2003; Gorg et al. 2004; Wittmann-Liebold et al. 2006).

In 2-DE, proteins are separated in first and second dimension according to their isoelectric point and molecular mass, respectively. In the first dimension proteins are fractionated by isoelectric focusing. Hereby separation of proteins is achieved through electrophoresis in a pH gradient gel system (using a gel strip with embedded pH gradient). Proteins migrate in the gel according to their charge at the respective pH and will accumulate at their isolelectric point (IP) where the positive and negative charges of the peptide are balanced so that they do not display a charge to the outside and do not migrate in an electric field anymore. In the second dimension proteins from the gel strip are run into a SDS-PAGE and separated according to their molecular mass. Before separation, proteins in the gel strip are treated with sodium dodecyl sulfate (SDS) along with other reagents to ensure that they are denatured and carry an appropriate negative charge. After separation, the gel is stained (i.e. Coomassie brilliant blue, silver, fluorescence dyes) and further analyzed. Protein spots of interest can be excised and digested with a site-specific protease (often trypsin). The resulting peptides are further investigated by mass spectrometry. This combination of methods is employed as a tool to detect and dissect dynamic changes in the proteome of a cell or tissue in response to changes in the physiological environment, the developmental state or internal perturbations, such as mutations. Fluorescence 2-DE Difference Gel Electrophoresis (DIGE) (Unlu et al. 1997) represents a new development in 2-DE. The use of multiple distinct fluorescent dyes to label protein samples prior to 2-DE PAGE allows multiple samples to be co-separated and visualized on one 2-DE gel.

Although, classical two-dimensional gel electrophoresis is a powerful tool, it faces a number of limitations especially when it comes to the separation of highly hydrophobic membrane proteins and proteins that possess basic isoelectric points. Hydrophobic proteins tend to precipitate at their isoelectric point in the non-detergent isoelectric focusing. In addition, they are often very heterogeneous in their physico-chemical properties what makes it difficult to achieve comprehensive, reproducible and comparable protein maps of membrane fractions (Ephritikhine et al. 2004).

New experimental approaches that are independent of 2-DE have been developed recently to overcome these limitations and allow comparative analysis of a protein between experimental and control samples in "solution", enabling a quantitative overview of the dynamically altered proteome.

Differential isotopic labeling strategies can also be employed to distinguish proteins from control and experimental conditions. Besides crosslinking of proteins isolated from cells grown under different conditions with isotopically labeled and unlabeled chemical probes (ICAT, isotope-coded affinity tag) (Gygi et al. 1999), proteins can be metabolically labeled with stable isotopes by growing cells in isotopically enriched media (SILAC, stable isotope-labeling of amino acids in cell culture). Experimental and control cell pools are then mixed, digested with enzymes and analyzed by LC-MS/MS for protein quantification (Oda et al. 1999; Ong et al. 2002). Moreover, mass spectrometry can be used to achieve absolute quantitation of proteins. For this purpose, proteotypic peptides that distinctively recognize a protein can be chemically synthesized holding stable isotopes (e.g. ^{13}C, ^{15}N etc.) at a single amino acid so that their masses will differ from the mass of the analyte, thereby permitting differentiation by MS and MS/MS methods. The absolute concentration of a protein can be calculated from the signal intensities derived from the analyte and from those of the internal standard (Zhu and Desiderio 1996). Tryptic peptides derived from the analyzed proteins and synthetic isotopically labeled internal standards were employed in absolute quantification of proteins in solution (Barr et al. 1996; Barnidge et al. 2003) and, recently, in-gel (Gerber et al. 2003).

In recent years alternative approaches have been developed that make use of the coupling between liquid chromatography and tandem MS (LC-MS/MS) (as reviewed in (Peng and Gygi 2001). The power of such a strategy can for example be illustrated by the identification of more than 70 proteins from the yeast ribosome in a single analysis. This approach was performed by analyzing tryptic peptides, which derived from digestion of the whole complex, by multi-dimensional liquid chromatography (mudPIT) coupled to MS/MS (Link et al. 1999).

Our current understanding of the organization of a proteome-wide interaction network points to its enormous complexity. It has become apparent that on average, every fourth protein in a proteome might be shared between protein complexes of different function. Two rather impressive proteome approaches were described for systematic analyses of components of multi-protein complexes from baker yeast *Saccharomyces cerevisiae* (Gavin et al. 2002; Ho et al. 2002). In one approach about 10% of predicted yeast proteins were used as baits to discover protein-protein interactions. Fascinatingly, 3,617 associated proteins were identified by mass spectrometry, covering about 25% of the yeast proteome (Ho et al. 2002). In the other approach distinct genes were tagged with an expression cassette encoding for protein A and the calmodulin binding protein. Protein complexes that contain a tagged protein could be isolated by tandem-affinity purification and the individual components be analyzed by MS/MS (Gavin et al. 2002).

The importance and power of proteomics for the exploration of plant proteomes and in particular plastid proteomics will be discussed in depth in the following section.

2 Proteomics of the chloroplast and its compartments

The chloroplast is a highly dynamic and complex cell organelle and a major part of its metabolism is involved in photosynthesis and related energy producing processes. However, it also produces amino acids and lipids as well as secondary metabolites like isoprenoids. A chloroplast can be divided into several compartments: the double layered envelope membrane, the soluble stroma and the thylakoid membrane enclosing the lumen. Despite of its endosymbiotic origin, most of the chloroplast proteins are nuclear encoded (about 90%). Plastidic genetic material is organized in so-called nucleoids and expressed via the chloroplast's transcription and translation machinery. Most of the chloroplast proteins synthesized in the cytosol are imported via the Toc and Tic (translocon at the outer/inner envelope membrane) translocation machinery. They contain an N-terminal chloroplast transit peptide (cTP) that is necessary for recognition at the outer membrane and is cleaved off after the passage into the stroma. Proteins targeted to the lumen of the thylakoids have to overcome a second barrier: the thylakoid membrane. Therefore, they carry a bipartite transit peptide, the more N-terminal region determining the chloroplast targeting. Once the initial transit peptide is cleaved, the formerly masked second transit peptide can be assessed and determines the next target, for example the thylakoid lumen (lTP).

To cope with the large amount of functions in cell metabolism, chloroplasts require a large amount of enzymes and other multi protein complexes. Therefore, a number of about 3000-4000 predicted chloroplast proteins in *Arabidopsis* is easily imaginable (Peltier et al. 2002; Schubert et al. 2002; Kieselbach and Schroder 2003; Baginsky and Gruissem 2004; Kleffmann et al. 2004; Richly and Leister 2004; Sun et al. 2004; van Wijk 2004). Only one proteomic study so far took on the difficult task to analyze the complete chloroplast proteome. In a study on *Arabidopsis thaliana* Kleffmann et al. (2004) were able to improve the dynamic range by using multi dimensional chromatography and additional enrichment of envelope membrane proteins to identify a set of 636 proteins, mostly associated with energy production or metabolic processes. But even with advanced proteomic strategies, the analysis of the whole chloroplast proteome is a difficult operation and therefore the chloroplast proteome is often divided in different subproteomes composed of the proteins from its subcompartments.

2.1 Envelope membranes

The chloroplast is a closed compartment located in the cytosol of the cell. It is completely surrounded by a double membrane, the so-called envelope. Since a wide range of different metabolic processes take place in the chloroplast, it is necessary for it to be functionally fully integrated into the plant cell, requiring an unproblematic exchange of metabolites and signals with the other compartments. Therefore, the envelope membrane is the site of many transport systems that facilitate the transport of not only carbohydrates but also phosphates, amino acids, protons and different metal-ions (Joyard et al. 1998; Rolland et al. 2003).

About 90% of the chloroplast proteins are synthesized in the cytosol and must be transported through the envelope membranes. This is achieved with the Toc/Tic protein import machinery. The Toc proteins, located in the outer membrane, are able to recognize chloroplast transit peptides and guide the immature proteins through the outer envelope. Here the Tic complex takes over and translocates the preprotein into the stroma where the TP is cleaved and chaperones fold the proteins into their functional conformation (Bedard and Jarvis 2005). The envelope membranes are also involved in the production and metabolism of different lipids such as structural membrane constituents, carotenoids and prenylquinones. Lipid compounds can also be further metabolized to signal molecules that are active in, for example, growth regulation or plant defense (Joyard et al. 1998; Rolland et al. 2003).

Predictions of envelope membrane components are difficult to make since a lot of these proteins do not contain cTP. Ferro et al. (2002) were able to determine typical properties for internal envelope membrane proteins, like a strong hydrophobicity based on several transmembrane domains (TMD), a pI larger than 8.8 and a Res/TM value (amino acid residues/TMD) of less than 100. Their prediction with ChloroP (Emanuelsson et al. 1999) and a manual check on these criteria resulted in 136 potential envelope proteins in *Arabidopsis*. Koo and Ohlrogge (2002) used a combination of TargetP (Emanuelsson et al. 2000), the TMD predictor TMHMM and a manual rejection of known thylakoid proteins to predict a number of 541 inner envelope candidate proteins. A new prediction approach was used by Schleiff et al. (2003) based on the idea that most of the outer envelope proteins are embedded in the membrane by a β-barrel structure. Their prediction was based on the combination of a computational β-barrel analysis, the determination of the isoelectric point, a TargetP analysis and a manual selection. This resulted in a pool of 891 putative outer envelope proteins. The candidate proteins derived from these predictions can be seen as starting points to design experimental approaches to characterize the envelope proteome. But experimental proteomic analysis of integral membrane proteins has always proven to be difficult because of the hydrophobic nature of the proteins and their highly dynamic expression. Therefore, a variety of extraction strategies such as the solubilization of proteins with organic solvents such as chloroform/methanol mixtures or treatments of membranes with alkaline substances or salts but also different fractionation strategies like SCX columns have been used (Ephritikhine et al. 2004; Rolland et al. 2006). Subsequent protein identification was mostly done by nano/LC coupled mass spectrometry (Ferro et al. 2000, 2002, 2003; Froehlich et al. 2003). A wide variety of proteins have been identified with these different approaches. A head count done by Peltier et al. (2004b) resulted in a number of 429 identified proteins located in the chloroplast envelope of *Arabidopsis*. Many of these proteins have no known function yet another big part of them works in protein translocation or metabolism. Froehlich et al. (2003) were able to identify many components of the *Arabidopsis* Toc/Tic protein import complex in a large scale proteomic analysis also including proteases and chaperones associated with these complexes. A blue native PAGE (BN-PAGE) analysis in pea revealed the molecular organization of the Toc core complex consistent of Toc159, Toc75 and Toc34 (Kikuchi et al.

2006) with an estimated size of about 800-1000 kDa and a stoichiometry of 1:3:3 (Toc159:Toc75:Toc34). It also became clear that the A-domain of Toc159 is involved in stabilizing the association of Toc34 with the complex. As for the Tic complex, the analysis by BN-PAGE uncovered a new subunit, Tic62, and demonstrated that this subunit together with Tic110 and Tic55 forms a core protein complex in the Tic translocon (Kuchler et al. 2002). As mentioned above, not only proteins have to be transported through the envelope membranes. Therefore, it is not surprising that all proteome studies of the envelope so far yielded in a large number of identified transporters for metabolites, ions or other organic components. Very abundant in most studies are, just to mention a few, oxoglutarate/malate-, phosphate/triosephosphate-, sugar- or ABC-type transporters (Seigneurin-Berny et al. 1999; Ferro et al. 2003; Froehlich et al. 2003). Another large group of identified proteins is involved in lipid metabolism. Among these are synthases, desaturases and acyltransferases that metabolize fatty acids, glycerolipids, pigments or prenylquinones. Noteworthy are also enzymes like the allene oxide synthase that is involved in the metabolism of oxylipins, which are signal components in plant growth and defense reactions. Other identified proteins were, for example, components involved in the response of the plant to oxidative stress like superoxide dismutase or ascorbate peroxidase (Seigneurin-Berny et al. 1999; Ferro et al. 2003; Froehlich et al. 2003). In conclusion, these findings demonstrate that the envelope membranes are not only an important transport machinery, but additionally represent a specialized and essential part in chloroplast metabolism.

2.2 Stroma and chloroplast ribosome

The chloroplast stroma is enclosed by the envelope membranes. Most importantly, this compartment is the site of the light independent photosynthetic reaction, the Calvin cycle, but also the oxidative pentose phosphate pathway and glycolysis are located here. The stroma contains the chloroplast DNA as well as its complex translational machinery in the form of the 70S chloroplast ribosomes. The stromal proteome is considered to be quite intricate and a prediction made by Sun et al. (2004) resulted in the number of 3387 putative stromal proteins. Due to it's complexity, most studies of the chloroplast stroma focused on single protein components and so far only one large scale proteomics study aimed to investigate the complete *Arabidopsis* stromal proteome (Peltier et al. 2006). A two dimensional approach was used combining colorless-native PAGE (CN-PAGE) in the first and SDS-PAGE in the second dimension. Gels were subjected to analysis with MALDI-TOF or LC-ESI-MS/MS as well as to a semi quantitative approach using staining with CyproRuby followed by image analysis. A number of 241 nonredundant proteins was identified and sorted into functional categories. Interestingly, a significant part of the proteins (26%) was involved in protein metabolism such as synthesis, folding, sorting, and proteolysis. As expected numerous proteins associated with carbon metabolism were also detected (12%). These enzymes from the Calvin cycle, the oxidative pentose phosphate pathway and glycolysis made up about three quarters of the stromal protein mass. Another 21% of the proteins had functions in nucleotide synthesis and degradation, amino acid metabolism or tetrapyrrole synthesis. It was not possible to assign a function to 11% of the identified proteins. Additionally, the authors extensively searched existing literature to determine the oligomeric state of the identified proteins in order to demonstrate the importance of functional paralogues in complex metabolic pathways.

2.2.1 The chloroplast ribosomes

The proteome of chloroplast ribosomes has an important role in the expression of chloroplast encoded genes and was therefore investigated in great detail. 2-DE-PAGE coupled with protein sequencing, HPLC, LC/MS and mass spectrometric analysis showed that the 70S ribosomes of spinach chloroplasts contain no less than 59 proteins; 33 in the 50S and 25 in the 30S subunit as well as a 70S complex associated ribosome recycling factor (Yamaguchi and Subramanian 2000; Yamaguchi et al. 2000). The 30S subunit that is responsible for mRNA binding and initiation of translation contains 21 *E. coli* orthologues as well as four plastid specific ribosomal proteins. In the 50S subunit, where the peptide synthesis takes place, 31 *E. coli* orthologues as well as two plastid specific ribosomal proteins were identified. These plastid specific ribosomal proteins termed PSRP1-6 are proposed to perform functions unique to plastid translation. The subsequent analysis of the chloroplast ribosomes of the unicellular green algae *Chlamydomonas rheinhardtii* using SDS-PAGE, LC MS/MS and mudPIT analysis resulted in the identification of 21 30S and 28 50S proteins as well as two proteins (RAP 38 and RAP 41) solely associated with the 70S complex (Yamaguchi et al. 2002, 2003). As for the 30S subunit in addition to several *E. coli* orthologues, the authors identified a PSRP3 homologue from spinach as well as a novel S1- domain containing protein named PSRP7. They were also able to determine that due to N-terminal extensions or other inserted sequences, three of the *Chlamydomonas* 30S ribosomal proteins are unusually large. In contrast to that, the composition of the 50S subunit is more conserved (27 *E. coli* orthologues, one spinach homologue, similar sizes). The authors, therefore, concluded that the differences in the composition of the 30S subunit might be related to unique features of *Chlamydomonas* chloroplast translational regulation, especially concerning mRNA discrimination at the 30S complex. In contrast to that, the enzymatic function of forming peptidyl bonds between amino acids in the 50S subunit is more conserved between organisms.

2.2.2 Plastoglobuli

Plastoglobuli (PG) are an additional component of the chloroplast stroma representing lipid containing structures that are mostly attached to the thylakoid membranes. They are known to accumulate α-tocopherol, plastoquinone, and triacylglycerols. Recently, the plastoglobuli proteome of *Arabidopsis* chloroplasts from WT plants grown under different light conditions and from the *clpr2-1* mutant that over accumulates PG were analyzed. NanoLC-ESI-MS/MS and a stable isotope labeling strategy were used to reveal the protein composition and acquire a functional model of these structures as well as to characterize their function in chloroplast metabolism (Ytterberg et al. 2006). It became clear that plastoglobuli contain a specific proteome mostly consisting of proteins from the fibrillin family forming the coating of the particles. In addition to these, several proteins involved in lipid metabolism, quinone synthesis and regulation as well as a number of aldolases involved in Calvin cycle and/or glycolysis could be identified. The authors, therefore, concluded that PG are not only storage facilities, but have a defined function in several metabolic pathways. They state that PG represent a connection in-between the thylakoid and the inner envelope membrane in the context of the metabolism of small molecules essential in thylakoid function and protection such as tocopherols and quinones. The important role of PG in tocopherol metabolism was also demonstrated in a parallel study in *Arabidopsis* (Vidi et al. 2006). The authors could show the localization of the tocopherol cyclase VTE1 within the PG using MS/MS, immunogold, and fluorescence labeling. In addition, they were able to identify a number of unclassified proteins and proteins from the plastid lipid associated protein/fibrillin-like family. Proteins involved in chloroplast processes like sugar and abscisic acid metabolism as well as jasmonic acid biosynthesis were also found.

2.2.3 The ferredoxin/thioredoxin system

The ferredoxin/thioredoxin system of the chloroplast is a key component of the regulation of photosynthetic enzymes in response to light facilitated through a number of redox processes. In an approach to find potential thioredoxin targets in the chloroplast, Balmer et al. (2003) bound thioredoxin f and m where one of the active Lys residues was replaced by Ser to a column in order to trap interacting stromal proteins from spinach. Specific proteins eluted from the column by addition of reducing equivalent (DTT) were then analyzed by 2-DE gels and MS. This led to the identification of several known along with a large number of so far unknown thioredoxin targets. It could be established that thioredoxins are also involved in regulation of so far unrecognized processes like isoprenoid, tetrapyrrole, or vitamin biosynthesis, protein assembly, folding, and degradation as well as processes involved in carbohydrate metabolism and DNA replication and transcription. In a subsequent study, the same group used a similar strategy including an affinity chromatography with WT thioredoxin f to trap proteins forming protein complexes based on protein/protein interactions (Balmer et al. 2004). The data revealed 27 so far unrecognized partners for protein/protein interaction with thioredoxin and indicated that not all of the thioredoxin targets identified to date that are able to form covalent interactions are able to interact electrostatically. The interaction partners cover a wide variety of chloroplast functions like Calvin cycle, translation, protein assembly and folding or other biosynthetic processes. The authors concluded that the formation of electrostatic complexes may help in the efficient transfer of electrons from photosystem I over the ferredoxin/thioredoxin system to the target proteins that can then be differentially regulated.

2.2.4 The Clp protease complex

An important component of the chloroplast stroma is the Clp protease complex, which, for example, degrades misfolded or unassembled proteins in an ATP dependent manner. Peltier et al. (2001; 2004a) identified and characterized a Clp protease complex of about 350 kDa in the stroma of *Arabidopsis* by employing BN-PAGE, CN-PAGE and native IEF/SDS-PAGE in combination with MALDI and ESI-MS. It could be shown that the complex is partially associated with the thylakoid membrane and consists of eleven different Clp proteins, five of which are serin-type proteases (ClpP1, 3-6) present in 1-3 copies per complex, four are non-proteolytic (ClpR1-4) and two have chaperone functions (ClpS1, 2) (Peltier et al. 2001, 2004a). This data, in addition to a detailed analysis of a ClpR2 deficient mutant of *Arabidopsis* with comparative quantitation using iTRAQ (isobaric tags for relative and absolute quantitation), and other approaches indicate a central role of the Clp protease complex proteins in plastid homeostasis as well as in chloroplast biogenesis and plant development (Peltier et al. 2004a; Rudella et al. 2006).

2.3 Thylakoid membrane

Oxygenic photosynthesis is the predominant function of the thylakoid membrane within the chloroplast. To perform this function and transform light energy into chemical energy in the form of ATP and NADPH, the tylakoid membrane contains four large multisubunit complexes involved in photosynthetic electron transfer and ATP synthesis. These are photosystem II (PSII), the cytochrome b_6f ($cytb_6f$) complex, photosystem I (PSI), and the ATP synthase. These complexes are distributed between two distinct membrane types. The stacked grana lamellae contain most of the PSII as well as $cytb_6f$ complexes, whereas the stroma lamellae contain most of the PSI as well as $cytb_6f$ and ATP synthase complexes (Timperio et al. 2004). In addition, the thylakoid membrane harbors proteins involved in assembly and maintenance of the lipid bilayer and the proteins therein. These include proteins for folding, incorporation, modification and degradation of the photosynthetic complexes as well as components of a complex transport machinery (Friso et al. 2004). The thylakoid proteome is a highly dynamic system since it requires the ability to adapt to changing environmental conditions such as increasing light intensities or changing temperature in order to especially protect the photosynthetic machinery from damage (Aro et al. 2005). Analysis of the thylakoid proteome proved to be difficult due to the strong prominence of the photosynthetic proteins that enhance the problem of dynamic resolution (van Wijk 2004) and due to the strong hydrophobicity of the integral membrane proteins (Whitelegge et al. 2006). Therefore, a wide range of approaches was employed to identify and analyze the components of the membrane and its associated proteins. The most comprehensive investigations of the thylakoid proteome involved combinations of several fractionation strategies to avoid the problems mentioned above. Ultracentrifugation, aqueous polymer two phase partitioning and two dimensional SDS-PAGE combined with MS were applied in a study of the *Synechocystis sp. PCC*

6803 thylakoid proteome (Srivastava et al. 2005). This yielded a pure integral membrane fraction where 76 proteins could be identified. Only 14 of these had transmembrane domains whereas the other proteins were peripherally located, most likely on the cytosolic side of the membrane. With the applied strategies the authors were able to not only resolve the abundant photosynthetic proteins, but also a multitude of proteins having functions in protein sorting mechanisms, pigment biosynthesis, protein folding or hypothetical proteins with yet unknown function. In a literally "in depth analysis" of the thylakoid proteome of *Arabidopsis* salt, detergent, and organic solvent extraction in combination with different multidimensional protein separation techniques, enzymatic, and non-enzymatic protein cleavages, MALDI and ESI-MS as well as bioinformatic approaches were used to generate an overview of peripheral and integral thylakoid proteins (Friso et al. 2004). Three distinct subproteomes could be separated: a peripheral, a peripheral but tightly associated and an integral membrane proteome. The peripheral subproteome was combined with proteins previously identified in a study of the lumenal proteome from the same group (Peltier et al. 2002) resulting in a set of 99 proteins that were dominated by the proteins of the oxygen evolving complex from PSII but also contained a number of unique new proteins. The hydrophobic fractions included 134 proteins with 76 of them having one or more known or predicted transmembrane domains. In the complete set of 198 non-redundant proteins, many (42%) photosynthetic proteins could be detected representing 85% of the known proteins in the four multisubunit complexes. In addition, a large number (15%) of new proteins with unknown function was identified such as two rubredoxins, a metallochaperone and a new DnaJ-domain protein. Other proteins found were involved in translation, metabolism or protein fate as well as in the protection from oxidative stress. The combination of extraction strategies used here allowed the detection of low abundant as well as small and very hydrophobic proteins that can hardly be resolved with standard 2-DE gel approaches. In a subsequent study, the same group developed a faster fractionation strategy employing three phase partitioning (TPP) of salt stripped thylakoids combined with RP-nano-LC-ESI-MS/MS. The authors again combined the data from the TPP approach with the analysis mentioned before (Friso et al. 2004) as well as their results of the lumenal and peripheral proteome of the thylakoids (Peltier et al. 2002) to achieve a combined dataset of more than 300 proteins. Whereas all other fractionation strategies mostly identified photosynthetic or other abundant thylakoid proteins, the TPP revealed a whole new level of lower abundant proteins, of which 50% were unknown. Others were involved in chlorophyll/prenyl lipid biosynthesis, protein sorting or degradation, stress defense or signaling. These studies show the significant improvement of dynamic resolution by the improvement of fractionation strategies and, concomitant with that, an increasing understanding of the thylakoid membrane proteome of higher plants. The combined datasets including functional assignments as well as other data from chloroplast proteomic studies from the van Wijk group can be accessed via the Plastid Proteome Database at Cornell (http://ppdb.tc.cornell.edu/). Still there are other approaches employed to characterize the thylakoid proteome. In a recent publication, Allmer et al. (2006) were able to create a comprehensive overview on the thylakoid proteome of

Chlamydomonas reinhardtii by combining SDS-page and ESI MS/MS of purified thylakoid membrane fractions obtained by ultracentrifugation with a high throughput genomic data mining strategy. This resulted in the detection of numerous low abundant proteins consisting of, for example, a novel light-harvesting protein, the STT7 kinase and a DegP-like protease along with numerous proteins with unknown function.

2.3.1 Analysis of the photosynthetic machinery

Numerous studies of thylakoid proteins focused on the proteins of the photosynthetic machinery to elucidate the composition of the different supercomplexes and their response towards changing environmental conditions. BN-PAGE is a method that is often employed to investigate the native state and thereby the interaction and the composition of the photosynthetic complexes. Combined with a second dimension SDS-PAGE and MS, this approach additionally enables the identification of subunits of the complexes. Studies in several higher plants gave insights in the supramolecular organization of the photosynthetic structures and revealed the existence of supercomplexes of PSI and PSII as well as of different forms of the photosystemI-light harvesting complexI (PSI-LHCI) and the PSII core complex. The employment of native gels also made it possible to resolve dimeric LHCI as well as monomeric and trimeric light harvesting complexII (LHCII), thus, supporting the results from diverse X-ray crystallographic analyses as well as the subunit organization and composition of the cytochrome b_6f complex and the ATP synthase (Heinemeyer et al. 2004; Ciambella et al. 2005; Granvogl et al. 2006). Still BN-PAGE analysis struggle to detect very small and hydrophobic proteins most likely due to the fact that these proteins have less tryptic cleavage sites and therefore often escape mass spectrometric detection (Granvogl et al. 2006). Also the dynamic resolution is not sufficient enough to detect very low abundant proteins making the discovery of unknown proteins very difficult. However, BN-PAGE represents a good alternative to the resolution of sucrose gradient for the separation of protein complexes. Additionally, this method can be used to show alterations in the protein complex structure in response to environmental factors or the metabolic state of the cell as was shown for *Chlamydomonas* cell grown in either photoautotrophic or photoheterotrophic conditions (Rexroth et al. 2003).

PSI-LHCI Complex. High performance liquid chromatography, used after prefractionation of the photosynthetic complexes through solubilisation and either differential or ultra centrifugation, can also be employed to analyze the thylakoid protein composition on an intact molecule level (Huber et al. 2004; Timperio et al. 2004). In this context, Zolla and Timperio (2000) were able to resolve most of the components of the spinach PSI-LHCI complex and to determine their molecular masses. In addition, it was shown on five dicotyledonous and four monocotyledonous plants that all species investigated possessed isoforms of the Lhca1 protein (Zolla et al. 2002). With classical denaturating one and two dimensional gel electrophoresis and MS/MS Storf et al. (2004) were for the first time able to identify the very low expressed Lhca5 gene product on the protein level. The authors showed the existence of several isoforms of all Lhca proteins (Lhca1-4) in tomato, demonstrating the presence of different populations of PSI in higher plants. A two dimensional SDS-PAGE approach was used to separate even highly hydrophobic proteins from the green alga *Chlamydomonas* when sample preparation was adapted towards the specific requirements of membrane proteins and resulted in a 2-DE map resolving LHCI and LHCII components in addition to other thylakoid proteins (Hippler et al. 2001). Using this improved 2-DE technique and mass spectrometry in a subsequent publication, Stauber et al. (2003) were able to establish a detailed 2-DE map of the *Chlamydomonas* light harvesting proteins demonstrating the expression of nine different Lhca as well as eight Lhcb proteins. In addition, a differential modification of Lhcbm3 and Lhcbm6 could be shown. According to these studies, it appeared that the LHCI of *Chlamydomonas* is significantly larger than that of higher plants. In an analysis of extracted PSI-LHCI and LHCI complexes by western blotting and 2-DE, Takahashi et al. (2004) additionally demonstrated that the core of the LHCI complex in *Chlamydomonas* is different from the dimeric structures of that of higher plants and forms a stable oligomeric complex that is able to assemble in the absence of the PSI core. It could further be demonstrated that three of the Lhca proteins (Lhca2, 3, 9) are only able to associate with the LHCI when they are stabilized by the presence of the PSI core what might be due to a functional role of these subunits in excitation energy transfer.

PSII-LHCII Complex. The PSII-LHCII complex was also object of detailed analysis of its structure and function using a number of different proteomic approaches. One study used SDS-PAGE in combination with protein sequencing and MS to analyze a highly purified PSII-LHCII complex containing a HIS-tag on the *psbB* gene product from *Synechocystis* (Kashino et al. 2002). This resulted in the detection of all known PSII subunits and some novel proteins representing potential candidates for the functional regulation of PSII. Using RP-HPLC-ESI-MS analysis on several different dicotyledonous and monocotyledonous plants, it was possible to characterize unique chromatographic patterns for each species. It became clear, that in monocots the LHCII complex appeared as mono- and trimers, whereas in dicots the trimeric form seems predominant. In addition, several isoforms of Lhcb1, 3, and 6 were detected that were predicted from the sequences of the multigene families coding for these proteins, but could not be resolved well with traditional gel based approaches (Huber et al. 2001; Zolla et al. 2003). These isoforms may have a role in the adaptation of PSII to different light conditions.

2.3.2 Post-translational modifications

Post-translational modifications, especially phosphorylation of photosynthetic membrane proteins, also play an essential role in the redistribution of excess light energy from the LHCII in PSII to the PSI in a process known as state transition that includes a migration of LHCII proteins from the grana to the stroma thylakoid regions (Vener et al. 2001; Timperio and Zolla 2005). Employing mass spectrometry on thylakoid membranes from *Arabidopsis* that were enriched for phosphoproteins using immobilized metal affinity chromatography (IMAC), Vener et al. (2001) were able to map phosphorylation sites of the central photosynthetic proteins including LHCII, D1, D2, CP43 and PsbH. In a subsequent analysis using this approach, they were able to identify in addition to the previously known phosphoproteins three more phosphorylation sites in a peptide from CP29, an expressed membrane protein, and for the first time in a PSI protein, namely PsaD (Hansson and Vener 2003). It became clear that the phosphorylation sites of all these *Arabidopsis* proteins are located at threonine residues near the N-terminus of the protein (Vener et al. 2001; Hansson and Vener 2003). In an analysis of the trypsin shaved thylakoid membrane of *Chlamydomonas*, enriched for phosphoproteins Turkina et al. (2004) detected a very unusual LHCII protein. It was demonstrated that the CP29 in its mature form still contained its transit peptide and solely showed a N-terminal methionine excision as well as a phosphorylation and acetylation site. This might represent an evolutionary compromise to keep the TP and in turn its functionally important phosphorylation site (Turkina et al. 2004). It is of note that the phosphorylation sites of *Chlamydomonas* Lhcbm proteins are also closer to the N-terminus as expected from predicted cleavage sites of the transit peptides (Stauber et al. 2003; Turkina et al. 2006b). Interestingly, N-terminal processed forms of Lhcbm proteins in *C. reinhardtii* exist, which lack these phosophorylation site, suggesting a novel type of regulation for *Chlamydomonas* Lhcbm proteins (Stauber et al. 2003). RP-HPLC can also be used to detect PTMs as shown on the PSII-LHCII complexes of pea and spinach where phosphorylation of for example D1, D2, CP43, two Lhcbs and PsbH could be resolved (Gomez et al. 2002). This approach was also capable of actually showing the migration of LHCII proteins like Lhcb2 and several isoforms of Lhcb1 from the grana to the stroma regions (Timperio and Zolla 2005). But it became clear that the phosphorylation state of the proteins was not the determining factor for the movement of the proteins, since some migrated in their unmodified form, but rather structural changes in the thylakoid organization due to PTMs as well as the pigment composition might be the cause for the migration of LHCII proteins to the PSI. Another important feature of phosphorylation is the maintenance of PSII by controlling the turnover of its reaction center proteins (Vener et al. 2001). This has, for example, been demonstrated by employing BN-gels to visualize the photoinhibition repair cycle on the basis of different PSII complexes represented in the native gel (Aro et al. 2005). With the employment of a wide variety of proteomic approaches including diverse fractionation strategies it became possible to rapidly increase the understanding of the complex organization, function, and regulation of the chloroplast thylakoid membrane and the complexes of oxygenic photosynthesis located therein. The rapidly evolving field of proteomics makes it possible to slowly overcome the limitations created by the physico-chemical character of a protein, its post-translational modification as well as dynamic resolution problems caused by the high abundance of the photosynthetic proteins.

2.4 Thylakoid lumen

The thylakoid lumen is the space enclosed by the thylakoid membranes. It is known that it contains proteins involved in oxygenic photosynthesis like the extrinsic subunits of PSII PsbO, PsbP, and PsbQ that function in the stabilization of the water oxidizing complex as well as the electron carrier protein plastocyanin. With the increased use of 2-DE separation techniques combined with mass spectrometry it became clear that the lumen harbors a lot more proteins than the ones involved directly in primary aspects of photosynthesis (Kieselbach et al. 1998). After the genome of *Arabidopsis thaliana* was sequenced the efforts to identify these proteins increased and yielded in the identification of so far about 40 to 50 proteins from *Arabidopsis*, spinach, and pea (Kieselbach and Schroder 2003; Sun et al. 2004). A part of these proteins was related to the already known subunits of PSII. Isoforms with unknown functions of PsbO and plastocyanin were found as well as many proteins with a PsbP domain and the PSII assembly factor Hcf136 (Kieselbach et al. 2000; Peltier et al. 2000, 2002; Schubert et al. 2002). The xanthophyll cycle enzyme violaxanthin deepoxidase as well as different peroxidases have an important function in the protection of PSII from oxidative stress and were identified in several studies (Peltier et al. 2002; Schubert et al. 2002). The lumenal proteome was also shown to be rich in putative immunophilins. They belong to the cyclophilin-type peptidy-prolyl *cis-trans* isomerases (PPIases) or to the FKBP-type PPIases that are involved in protein folding and might as well work in chaperoning or have regulatory functions (Peltier et al. 2002; Schubert et al. 2002; Kieselbach and Schroder 2003). Numerous proteases of different types were identified in the lumen as well. Their roles include, among others, the processing of the D1 protein (Schubert et al. 2002; Kieselbach and Schroder 2003). A surprisingly large group of proteins with a pentapeptide repeat were also found under the lumenal proteins, but their sequences did not show any other known functional domains (Peltier et al. 2000, 2002; Schubert et al. 2002; Kieselbach and Schroder 2003). Proteins located on the thylakoid periphery of the lumen were additionally investigated by Peltier et al. (2000; 2002) and contained a large number of fibrillins, which might function in carotenoid storage and a ClpS1 protease from the thylakoid associated Clp protease complex.

To help in the identification of possible lumenal proteins, efforts were also put in the prediction of the complete lumenal proteome. Predictions of proteins with plastidal and lumenal transit peptides made with TargetP and SignalP (Nielsen et al. 1997) resulted in only 80 putative proteins for *Arabidopsis* (Schubert et al. 2002; Kieselbach and Schroder 2003) and at least 200 candidate proteins for *Arabidopsis* and pea (Peltier et al. 2000, 2002). The difficulties with the prediction programs used led to the development of LumenP (Westerlund et al. 2003), a neural network predictor especially designed to identify lumenal target sequences. A combined prediction using TargetP and LumenP on the *Arabidopsis* open reading frames resulted in 417 proteins to be potentially located in the thylakoid lumen. An independent comparison of predictions from SignalP with predictions from LumenP yielded in a number of 285 and 291 lumenal proteins, respectively. Still, only 150 proteins were predicted by both algorithms and from the 100 proteins that were different between the programs, only 26 overlapped with the 53 experimentally shown lumenal proteins, which the authors collected from literature (Sun et al. 2004). In the course of the combined efforts to investigate the proteome of the lumen, it became clear that in order to transport proteins from the chloroplast stroma through the thylakoid membrane into the lumen, the delta pH dependent

TAT pathway (requiring a twin Arginine motive in the transit peptide) plays a much more pronounced role than the ATP driven Sec pathway, and additionally that the twin arginine motive makes a prediction of a lumenal protein more reliable than any other amino acid sequence in the lTP (Westerlund et al. 2003; Sun et al. 2004). It is noteworthy that all types of predictions have the tendency to be erroneous and have to be verified with experimental data.

3 Predictions and collections of the chloroplast proteome

The knowledge of the subcellular localization of so far uncharacterized proteins can provide valuable information to elucidate their function in specific metabolic processes and to get more insights into cellular functions as a whole (Heazlewood et al. 2005). For this reason, efforts were made to develop software tools like SignalP (Nielsen et al. 1997), ChloroP (Emanuelsson et al. 1999), TargetP (Emanuelsson et al. 2000), or Predotar (Small et al. 2004) to predict N-terminal transit peptides which determine the subcellular localization of the proteins as well as their putative cleavage sites. The major drawback of this strategy is that even TargetP, the currently most successful prediction program, has a prediction efficiency of only 70-85% for vascular plant gene products, whereas its prediction accuracy is even less when it comes to green alga (Kleffmann et al. 2006). This high number of erroneous predictions has several reasons. One is that transit peptides are usually not well conserved and quite diverse in length and amino acid composition. Some proteins of the outer envelope membrane are also known to have no target peptide at all (Jarvis 2004; Richly and Leister 2004; Lunn 2006). Another problem is the limited number of proteins that are usually used to train the predictors making it not surprising that results can be inconsistent when using different prediction tools (Heazlewood et al. 2005). Even though most of the used training sets contain a subset of algae proteins, predictions for these are often especially erroneous. The problem here might be based on the fact that algal transit peptides are on average 32 amino acids shorter than those of vascular plants and therefore more difficult to predict with the current mixed training sets used (Gomez et al. 2003).

Nevertheless, in addition to the limitations of computer algorithms, the targeting of some proteins will never be recognizable with such tools. The existence of proteins without target sequences and the employment of unconventional import strategies such as via the secretory pathway and dual targeting, demonstrates the need for alternative approaches to elucidate organelle targeting (for reviews see: Millar et al. 2006; Radhamony and Theg 2006). Such limitations will result in an underestimation of organellar proteomes as, for example, the proteome of the chloroplast. Still computer based predictions of the possible localization of a protein can be very useful since they can give indications on the presence of low abundant proteins that would not be detectable in an experimental proteomic approach due to dynamic range limitations (Newton et al. 2004). Nevertheless, experimental proof is essential to verify the predicted location for these proteins.

With the sequencing of the *Arabidopsis thaliana* genome, this plant became the primary subject to predict chloroplast proteins with estimated numbers ranging from about 3000 to 4000 proteins for the complete organelle (Peltier et al. 2002; Schubert et al. 2002; Kieselbach and Schroder 2003; Baginsky and Gruissem 2004; Kleffmann et al. 2004; Richly and Leister 2004; Sun et al. 2004; van Wijk 2004). These collected sets of protein predictions were then used to predict the localization of proteins within the different compartments of the chloroplast (see chapters in the article). Sun et al. (2004) made a large effort to improve these strategies by collecting a set of about 250 proteins with different compartment locations from published data. The analysis of biochemical properties such as the cysteine content, the protein-size or the number of transmembrane domains revealed that, for example, the cysteine content of proteins in the thylakoid membranes or the lumen is much lower than in proteins of the envelope membranes. The authors were also able to show differences in the size, pI and number of transmembrane domains within proteins from different sub compartments. These characterizations can be helpful in determining the location of a protein and to further optimize prediction programs.

Recently, a number of databases were created containing huge collections of experimental as well as in silico data to provide information on the proteomes of cellular as well as organellar components, especially from *Arabidopsis*.

Heazlewood et al. (2006) developed SUBA, a subcellular database that combines information from mass spectrometric and fluorescent protein experiments, as well as several other databases, to a total of more than 6700 nonredundant proteins that can be assigned to ten distinct locations within the cell. The SUBA database gives the user the possibility to build protein sets out of published or newly imported data including the results of different prediction programs and compare these with the database entries (http://www.plantenergy.uwa.edu.au/applications/suba/index.php). plprot, developed by Kleffmann et al. (2006) is a database that contains collected experimental data on different plastid types like chloroplasts, etioplasts, and undifferentiated plastids. It contains 2043 partially redundant protein entries from *Arabidopsis*, rice, and BY2 cells including a variety of information on the single proteins. It also features a plastid type comparison that makes it possible to compare proteins within the different datasets (http://www.plprot.ethz.ch/). A comprehensive chloroplast specific database, PPDB, is available from the lab of KJ van Wijk (Friso et al. 2004) (http://ppdb.tc.cornell.edu/). It combines predicted and experimentally identified proteins from all suborganellar locations, including also information on fluorescent protein experiments from the SUBA database. Protein entries include annotations, predicted and experimentally determined molecular and biophysical properties, as well as information on protein-protein interactions, results from comparative proteomics studies, functional classifications, and schematics on biochemical pathways.

4 Comparative proteomics

With the increasing amount of information about the various subproteomes of cell organelles, it became clear that in order to understand complex biological processes it is not only important to understand gene expression but also to analyze complex protein expression patterns that are strongly influenced by post-transcriptional and post-translational modification processes (Steen and Pandey 2002). Therefore, much effort was put into the comparison of protein abundance in different sets of samples that derived from, for instance, different developmental stages of the organelle or from cells under diverse biotic or abiotic stress conditions. Several methods used rely on gel systems like the classical two-dimensional SDS-PAGE with isoelectric focusing as a first dimension or the difference gel electrophoresis. Since these methods cope with the known problems of gel based protein separation, like the difficulties in the resolution of strong hydrophobic membrane proteins, non-gel-based methods were developed that employ stable isotopes like ICAT, SILAC, or iTRAQ (isobaric tags for relative and absolute quantitation). Nevertheless, in chloroplast proteomics, most comparative studies so far rely on 2-DE and concomitant image analysis and mass spectrometry.

4.1 Plant and chloroplast development

Several studies used proteomics to investigate developmental stages of whole leaves as done in rice (Zhao et al. 2005) or on senescent leaves of white clover (Wilson et al. 2002). This included the analysis of a chloroplast fraction and demonstrated the organized breakdown of organelles and macromolecules and increased levels of proteins involved in remobilization of nutrients that are relocated to developing plant parts. 2-DE and a detailed data evaluation were used to characterize chloroplast biogenesis in maize that led to the identification of 26 unique spots on gels from 5 different time points. These proteins were mostly from the light reaction and the carbon assimilation cycle of photosynthesis but also chaperones and other metabolic enzymes (Lonosky et al. 2004). In a plant cell, a fully differentiated chloroplast can, dependent on the tissue it is located in, serve a variety of functions as for example known from mesophyll and bundle sheath cells in C4 plants. This aspect was analyzed in detail using maize chloroplast stroma for 2-DE and image analysis as well as differential labeling with cleavable ICAT and a comparison of unlabeled stroma proteins by LC-ESI-MS (Majeran et al. 2005) (Figure 3). The three approaches proved to be complementary and resulted in the identification of 400 proteins from a wide variety of pathways and a detailed overview of differential protein accumulation in chloroplasts from mesophyll or bundle sheath cells that are mainly due to the metabolic differentiation of the two cell types. The authors additionally provided evidence for a differential regulation of plastid gene expression, protein biogenesis and protein fate and presented a number of so far unknown proteins that are specifically expressed in one tissue and probably have central functions in the C4 plant metabolism.

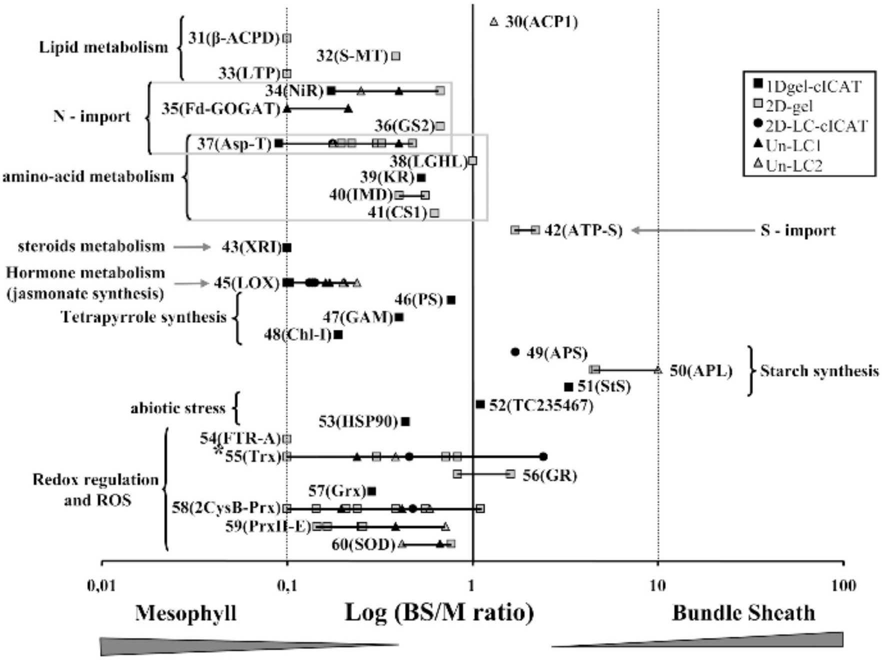


Fig. 3. Overview of bundle sheath: mesophyll ratios of selected proteins. Values were determined by three complementary methods, 2DE and image analysis, differential labeling with cleavable ICAT, and a comparison of unlabeled stroma proteins by LC-ESI-MS. Taken with permission from Majeran et al. (2005).

4.2 Biotic stress

Comparative proteomics was also used to characterize the answers of plants to biotic or abiotic stress factors. The response of *Arabidopsis* to a pathogen attack of *Pseudomonas syringae* pv. *tomato*, for example, was shown to induce phosphorylation of leaf proteins whose abundances could be compared using iTRAQ (Jones et al. 2006a). The attack also led to the induction of characteristic protein changes in the total soluble leaf proteome. In addition, chloroplast and mitochondrial fractions that mainly contained defense related antioxidants and metabolic enzymes were analyzed with 2-DE (Jones et al. 2006b). In a study on *Nicotiana benthamiana*, thylakoid membranes 2-DE enabled the identification of the PsbO and PsbP proteins as well as at least four different isoforms respectively. When plants were inoculated with the Spanish strain of pepper mild mottle virus PsbP proteins selectively decreased dramatically in abundance as compared to PsbO proteins, which can be seen as a counteraction to the virus infection by regulating photosynthetic activity probably as a basic defense mechanism (Perez-Bueno et al. 2004).

4.3 Abiotic stress

4.3.1 Light and temperature

Abiotic stress can be caused by several environmental conditions. For plants, light is the most important requirement for life but excess light can severely damage the cells especially with the accumulation of reactive oxygen species (ROS). As was shown for *Arabidopsis* chloroplasts by 2-DE and image analysis the largest damage can be observed in the protein complexes involved in photosynthesis or in protein metabolism whereas proteins with increased abundance under high light conditions were mainly defense related like scavengers for ROS or chaperones (Phee et al. 2004). In a later study on the effects of high light, Giacomelli et al. (2006) analyzed the *Arabidopsis* thylakoid lumenal and peripheral proteomes as well as the thylakoid associated plastoglobuli from wild type and the vtc2-2 mutant, containing only 20-30% of WT level of ascorbate, using proteomics and physiological experiments. Seven proteins were found upregulated in both genotypes under high light conditions: YCF37, four members of the fibrillin protein family, Fru-biphosphate aldolase-1 and a flavin reductase-related protein, of which the latter three types of proteins were located in the PG. The authors, therefore, concluded that the PG are probably involved in the synthesis and accumulation of α-tocopherol and quinones which are major antioxidants and that breakdown of carotenoids and turnover of lipids/fatty acids might also be a plastoglobuli associated process. High light response of the soluble proteome led to the upregulation of only the Ser type IV thylakoid protease SPPA whereas many other proteases showed no significant response. Since ascorbate is an important antioxidant, the analysis of the ascorbate-mutant was expected to reveal an additional level of specific stress responses to high light, but finally, ascorbate deficiency showed only small effects on the stress reaction since the mutant, compared to the WT, only differentially accumulated proteins that belonged to known stress response functions like the superoxide dismutases or some chaperones. In a study on whole *Chlamydomonas* cells, including two very high light resistant mutant strains, Forster et al. (2006) were able to identify new candidate proteins possibly involved in high light resistance like a DEAD box RNA helicase-like protein, NAB1 and RB38 where the latter two might be especially involved in the stability and function of the LHC and PSII under high light conditions. The effects of different light and temperature conditions on thylakoid pigment binding proteins, especially of Lhcb family members, was investigated in maize (Caffarri et al. 2005). Physiological analysis showed that plants grown in low temperature had an increased photoinhibitory damage, and in combination with high light, also an increased non-photochemical quenching. Low temperature and high light additionally increased the LHCII content whereas PSII core and PSI-LHCI proteins decreased in abundance. 2-DE gels also demonstrated a differential accumulation of several individual Lhcb1-3 proteins under different growth conditions indicating that diverse LHCII isoforms corresponding to multiple Lhcb1-3 genes might be needed for the acclimation to changing light and temperature conditions. Also low temperature alone can lead to significant changes in the chloroplast proteome.

Goulas et al. (2006) characterized lumenal and stromal proteins of *Arabidopsis* plants that were grown at low temperatures for different periods of time using a DIGE approach (Fig. 4). Most changes in the protein pattern could be observed after ten days in the cold. Changes in the stromal proteome included the upregulation of Rubisco whereas the other Calvin cycle enzymes decreased. An increase was also observed for some proteins from PSI and several enzymes related to oxidative stress. In the lumen only a few proteins like PsbP1 and PsbO2 increased abundance in the cold and some immunophilins showed variable responses.

4.3.2 CO_2 and iron

Another important aspect in plant life is the availability of carbon dioxide in order to sustain oxygenic photosynthesis. Under low CO_2 conditions, especially algae induce CO_2 concentrating mechanisms (CCM). However, even before these are completely developed, carbon dioxide limitation leads to a specific redox-dependent phosphorylation of two proteins, Lci5 and UEP (unknown expressed protein) that could be characterized in extrinsic thylakoid protein preparations using IMAC and ESI/TOF in *Chlamydomonas* chloroplasts (Turkina et al. 2006a). The data also indicated that thylakoids might contain a redox-dependent protein kinase specifically activated in the early stages of CCM.

Since plants are carrying out photosynthesis, they are highly dependent on metal components like iron or copper that function as cofactors in many enzyme complexes. Changes occurring in the thylakoid membrane at the onset of iron deficiency were analyzed in detail in several studies. Moseley et al. (2002) employed fluorescence emission analysis, immunoblots and 2-DE gel electrophoresis on thylakoid membranes of the green alga *Chlamydomonas*. They could demonstrate that iron deficiency leads to a functional uncoupling of the LHCI antenna from the PSI core leading to an impaired efficiency in the excitation energy transfer. Furthermore, adaptation to iron deficiency leads to a distinct and highly coordinated remodeling of the photosynthetic antenna complexes of LHCI and to a pronounced decrease in the abundance of PSI. Importantly, the remodeling of the photosynthetic apparatus became evident before a chlorotic phenotype was visible.

In a following publication Naumann et al. (2005) used a SILAC strategy to further characterize the remodeling process of the LHCI in *Chlamydomonas* and showed that, whereas Lhca5, 1, 7, and 8 are reduced in abundance, Lhca4 and 9 are induced. They also demonstrated that the N-terminal processing of Lhca3 occurs at a functionally assembled PSI-LHCI complex and could therefore be regarded as a key regulatory step in the remodeling process (see Fig. 4). Employing the SILAC approach, it was shown that the onset of iron deficiency additionally leads to alterations in the abundance of a variety of thylakoid proteins not directly involved in primary photosynthetic processes (Naumann et al., unpublished data). Iron deficiency was also analyzed on thylakoid membranes of sugar beet with IEF-SDS-PAGE as well as BN-SDS-PAGE in combination with mass spectrometry (Andaluz et al. 2006) (Fig. 4). In this study, it was demonstrated that iron deficiency leads to a pronounced decrease in proteins involved in photosynthetic

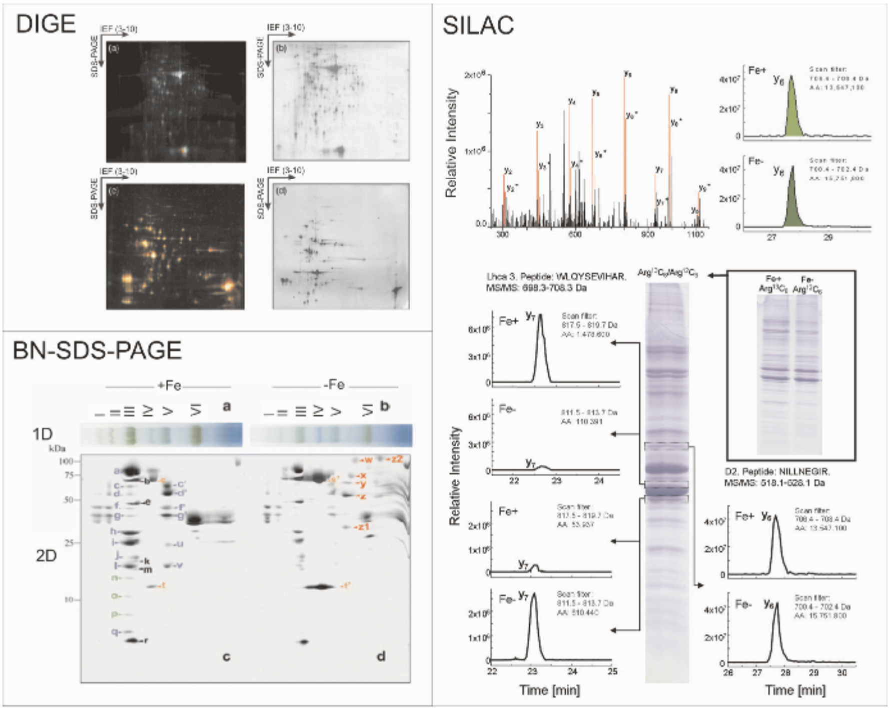


Fig. 4. Examples of comparative proteomics approaches, DIGE analysis of stromal and lumenal proteins of *Arabidopsis* plants, grown at low temperatures for different periods of time (Goulas et al. 2006), BN-SDS-PAGE of thylakoid membranes from Fe-sufficient and Fe-deficient samples from sugar beet (Andaluz et al. 2006), and fragmentation spectrum and elution profile for abundance calculation of a SILAC labeling experiment with labeled arginine on thylakoids from *Chlamydomonas* cells grown under Fe-limiting or standard conditions (Naumann et al. 2005). All figures taken with permission.

electron transport, whereas Calvin cycle enzymes partially increase in abundance. Also increased were proteins from biosynthetic pathways as well as stress related proteins. Comparing the two methods employed, it became clear that they are complementary since BN-SDS-PAGE is better suited for the resolution of hydrophobic membrane proteins, whereas IEF-SDS-PAGE has a higher resolving power. The use of proteomics to investigate plastid protein dynamics, adaptation, and acclimation to diverse stress conditions demonstrate that it is an emerging key technique to tackle these kinds of questions in plastid biology. A short summary of methods used in comparative proteomic approaches as well as their field of application is presented in Table 1; selected examples are additionally illustrated in Figure 4.

Table 1. Overview of approaches used in comparative proteomics experiments

Proteomic approach	Protein identification	Topic	Reference
2DE and image analysis	MALDI-TOF	Chloroplast biogenesis (Maize)	Lonosky et al. (2004)
	MALDI-TOF, LC-MS/MS, IB*	Leaf development (Rice)	Zhao et al. (2005)
	MALDI-TOF	Leaf senenscence (White clover)	Wilson et al. (2002)
	N-terminal sequencing, IB	OEC after pathogen attack (*Nicotiana benthamiana*)	Perez-Bueno et al. (2004)
	LC-MS/MS	Defence proteome after pathogen attack (*Arabidopsis*)	Jones et al. (2006b)
	LC-MS/MS, IB	Effects of iron deficiency on thylakoid proteome (*Chlamydomonas*)	Moseley et al. (2002)
	MALDI-TOF (/TOF)	Effects of iron deficiency on thylakoid proteome (*Beta vulgaris*)	Andaluz et al. (2006)
	LC-MS/MS, IB	Effects of high light on thylakoid proteome (*Arabidopsis*)	Giacomelli et al. (2006)
	MALDI-TOF	Effects of high light on chloroplast proteome (*Arabidopsis*)	Phee et al. (2004)
	MALDI-TOF	Effects of high light on chloroplast proteome (*Chlamydomonas*)	Forster et al. (2006)
	IB	Effects of diverse light and temperature on Lhcb gene products (Maize)	Caffarri et al. (2005)
	MALDI-TOF, LC-MS/MS	Proteomes of bundle sheath and mesophyll cells (Maize)	Majeran et al. (2005)
DIGE	MALDI-TOF (/TOF)	Effects of low temperature on chloroplast lumen and stromal proteome (*Arabidopsis*)	Goulas et al. (2006)
IMAC	ESI-MS/MS, IB	Phosphorylation of extrinsic thylakoid proteins under CO_2 limitation (*Chlamydomonas*)	Turkina et al. (2006a)
iTRAQ	LC-MS/MS	Defense phosphoproteome of leaves after pathogen attack (*Arabidopsis*)	Jones et al. (2006a)
BN-SDS-PAGE	MALDI-TOF (/TOF)	Effects of iron deficiency on thylakoid proteome (*Beta vulgaris*)	Andaluz et al. (2006)
SILAC	LC-MS/MS, IB	Remodeling of PSI in iron deficiency (*Chlamydomonas*)	Naumann et al. (2005)
cICAT and comparative Ion chromatography with LC-MS	LC-MS/MS	Proteomes of bundle sheath and mesophyll cells (Maize)	Majeran et al. (2005)

*IB: immunoblot

5 Conclusion

In the recent years, many proteomic studies of the chloroplast revealed a large number of proteins with yet unknown function. This calls for further analysis to elucidate the function and role in physiological processes of these gene products by reverse genetics experiments. It also became clear that comparative proteomics is an important tool to analyze dynamic changes in the chloroplast proteome. It is foreseeable that absolute quantitation, especially of individual subunits stemming from multiprotein complexes will open new perspectives in analyzing structures and functions in response to changing cellular environment. In this context, it will also be important to focus on the pronounced dynamic and variability of a proteome specifically due to post-translational modifications. Proteomics will also play a role in increasing the understanding of protein/protein interactions, and in combination with detailed gene expression studies help to improve our knowledge of the biology in complex cell organelles like the chloroplast in a whole systems perspective.

Acknowledgement

We are grateful to Mia Terashima and Dr. Jens Allmer for fruitful discussions and critical reading of the manuscript.

References

Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. Nature 422:198-207

Allmer J, Markert C, Stauber EJ, Hippler M (2004) A new approach that allows identification of intron-split peptides from mass spectrometric data in genomic databases. FEBS Lett 562:202-206

Allmer J, Naumann B, Markert C, Zhang M, Hippler M (2006) Mass spectrometric genomic data mining: Novel insights into bioenergetic pathways in *Chlamydomonas reinhardtii*. Proteomics 6:6207-6220

Andaluz S, Lopez-Millan AF, De Las Rivas J, Aro EM, Abadia J, Abadia A (2006) Proteomic profiles of thylakoid membranes and changes in response to iron deficiency. Photosynth Res 89:141-155

Aro EM, Suorsa M, Rokka A, Allahverdiyeva Y, Paakkarinen V, Saleem A, Battchikova N, Rintamaki E (2005) Dynamics of photosystem II: a proteomic approach to thylakoid protein complexes. J Exp Bot 56:347-356

Bafna V, Edwards N (2001) SCOPE: a probabilistic model for scoring tandem mass spectra against a peptide database. Bioinformatics 17:3-21

Bafna V, Edwards N (2003) On *de novo* Interpretation of Tandem Mass Spectra for Peptide Identification. RECOMB 2003:9-18

Baginsky S, Gruissem W (2004) Chloroplast proteomics: potentials and challenges. J Exp Bot 55:1213-1220

Balmer Y, Koller A, del Val G, Manieri W, Schurmann P, Buchanan BB (2003) Proteomics gives insight into the regulatory function of chloroplast thioredoxins. Proc Natl Acad Sci USA 100:370-375

Balmer Y, Koller A, Val GD, Schurmann P, Buchanan BB (2004) Proteomics uncovers proteins interacting electrostatically with thioredoxin in chloroplasts. Photosynth Res 79:275-280

Barnidge DR, Dratz EA, Martin T, Bonilla LE, Moran LB, Lindall A (2003) Absolute quantification of the G protein-coupled receptor rhodopsin by LC/MS/MS using proteolysis product peptides and synthetic peptide standards. Anal Chem 75:445-451

Barr JR, Maggio VL, Patterson DG Jr, Cooper GR, Henderson LO, Turner WE, Smith SJ, Hannon WH, Needham LL, Sampson EJ (1996) Isotope dilution--mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I. Clin Chem 42:1676-1682

Bedard J, Jarvis P (2005) Recognition and envelope translocation of chloroplast preproteins. J Exp Bot 56:2287-2320

Caffarri S, Frigerio S, Olivieri E, Righetti PG, Bassi R (2005) Differential accumulation of Lhcb gene products in thylakoid membranes of *Zea mays* plants grown under contrasting light and temperature conditions. Proteomics 5:758-768

Chen T, Kao MY, Tepel M, Rush J, Church GM (2001) A dynamic programming approach to de novo peptide sequencing via tandem mass spectrometry. J Comput Biol 8:325-337

Ciambella C, Roepstorff P, Aro EM, Zolla L (2005) A proteomic approach for investigation of photosynthetic apparatus in plants. Proteomics 5:746-757

Dancik V, Addona TA, Clauser KR, Vath JE, Pevzner PA (1999) *De novo* peptide sequencing via tandem mass spectrometry. J Comput Biol 6:327-342

Domon B, Aebersold R (2006) Mass spectrometry and protein analysis. Science 312:212-217

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol 300:1005-1016

Emanuelsson O, Nielsen H, von Heijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci 8:978-984

Eng JK, Mccormack AL, Yates JR (1994) An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. J Am Soc Mass Spectrom 5:976-989

Ephritikhine G, Ferro M, Rolland N (2004) Plant membrane proteomics. Plant Physiol Biochem 42:943-962

Fernandez-de-Cossio J, Gonzalez J, Satomi Y, Shima T, Okumura N, Besada V, Betancourt L, Padron G, Shimonishi Y, Takao T (2000) Automated interpretation of low-energy collision-induced dissociation spectra by SeqMS, a software aid for *de novo* sequencing by tandem mass spectrometry. Electrophoresis 21:1694-1699

Ferro M, Salvi D, Brugiere S, Miras S, Kowalski S, Louwagie M, Garin J, Joyard J, Rolland N (2003) Proteomics of the chloroplast envelope membranes from *Arabidopsis thaliana*. Mol Cell Proteomics 2:325-345

Ferro M, Salvi D, Riviere-Rolland H, Vermat T, Seigneurin-Berny D, Grunwald D, Garin J, Joyard J, Rolland N (2002) Integral membrane proteins of the chloroplast envelope:

identification and subcellular localization of new transporters. Proc Natl Acad Sci USA 99:11487-11492

Ferro M, Seigneurin-Berny D, Rolland N, Chapel A, Salvi D, Garin J, Joyard J (2000) Organic solvent extraction as a versatile procedure to identify hydrophobic chloroplast membrane proteins. Electrophoresis 21:3517-3526

Forster B, Mathesius U, Pogson BJ (2006) Comparative proteomics of high light stress in the model alga *Chlamydomonas reinhardtii*. Proteomics 6:4309-4320

Friso G, Giacomelli L, Ytterberg AJ, Peltier JB, Rudella A, Sun Q, Wijk KJ (2004) In-depth analysis of the thylakoid membrane proteome of *Arabidopsis thaliana* chloroplasts: new proteins, new functions, and a plastid proteome database. Plant Cell 16:478-499

Froehlich JE, Wilkerson CG, Ray WK, McAndrew RS, Osteryoung KW, Gage DA, Phinney BS (2003) Proteomic study of the *Arabidopsis thaliana* chloroplastic envelope membrane utilizing alternatives to traditional two-dimensional electrophoresis. J Proteome Res 2:413-425

Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Hofert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier MA, Copley RR, Edelmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, Superti-Furga G (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415:141-147

Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci USA 100:6940-6945

Giacomelli L, Rudella A, van Wijk KJ (2006) High light response of the thylakoid proteome in *Arabidopsis* wild type and the ascorbate-deficient mutant vtc2-2. A comparative proteomics study. Plant Physiol 141:685-701

Giddings MC, Shah AA, Gesteland R, Moore B (2003) Genome-based peptide fingerprint scanning. Proc Natl Acad Sci USA 100:20-25

Gomez SM, Bil KY, Aguilera R, Nishio JN, Faull KF, Whitelegge JP (2003) Transit peptide cleavage sites of integral thylakoid membrane proteins. Mol Cell Proteomics 2:1068-1085

Gomez SM, Nishio JN, Faull KF, Whitelegge JP (2002) The chloroplast grana proteome defined by intact mass measurements from liquid chromatography mass spectrometry. Mol Cell Proteomics 1:46-59

Gorg A, Weiss W, Dunn MJ (2004) Current two-dimensional electrophoresis technology for proteomics. Proteomics 4:3665-3685

Goulas E, Schubert M, Kieselbach T, Kleczkowski LA, Gardestrom P, Schroder W, Hurry V (2006) The chloroplast lumen and stromal proteomes of *Arabidopsis thaliana* show differential sensitivity to short- and long-term exposure to low temperature. Plant J 47:720-734

Granvogl B, Reisinger V, Eichacker LA (2006) Mapping the proteome of thylakoid membranes by *de novo* sequencing of intermembrane peptide domains. Proteomics 6:3681-3695

Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17:994-999

Hansson M, Vener AV (2003) Identification of three previously unknown in vivo protein phosphorylation sites in thylakoid membranes of *Arabidopsis thaliana*. Mol Cell Proteomics 2:550-559

Heazlewood JL, Tonti-Filippini J, Verboom RE, Millar AH (2005) Combining experimental and predicted datasets for determination of the subcellular location of proteins in *Arabidopsis*. Plant Physiol 139:598-609

Heazlewood JL, Verboom RE, Tonti-Filippini J, Small I, Millar AH (2006) SUBA: the *Arabidopsis* Subcellular Database. Nucleic Acids Res 35:D213-D218

Heinemeyer J, Eubel H, Wehmhoner D, Jansch L, Braun HP (2004) Proteomic approach to characterize the supramolecular organization of photosystems in higher plants. Phytochemistry 65:1683-1692

Hippler M, Klein J, Fink A, Allinger T, Hoerth P (2001) Towards functional proteomics of membrane protein complexes: analysis of thylakoid membranes from *Chlamydomonas reinhardtii*. Plant J 28:595-606

Ho Y, Gruhler A, Heilbut A, Bader GD, Moore L, Adams SL, Millar A, Taylor P, Bennett K, Boutilier K, Yang L, Wolting C, Donaldson I, Schandorff S, Shewnarane J, Vo M, Taggart J, Goudreault M, Muskat B, Alfarano C, Dewar D, Lin Z, Michalickova K, Willems AR, Sassi H, Nielsen PA, Rasmussen KJ, Andersen JR, Johansen LE, Hansen LH, Jespersen H, Podtelejnikov A, Nielsen E, Crawford J, Poulsen V, Sorensen BD, Matthiesen J, Hendrickson RC, Gleeson F, Pawson T, Moran MF, Durocher D, Mann M, Hogue CW, Figeys D, Tyers M (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. Nature 415:180-183

Huber CG, Timperio AM, Zolla L (2001) Isoforms of photosystem II antenna proteins in different plant species revealed by liquid chromatography-electrospray ionization mass spectrometry. J Biol Chem 276:45755-45761

Huber CG, Walcher W, Timperio AM, Troiani S, Porceddu A, Zolla L (2004) Multidimensional proteomic analysis of photosynthetic membrane proteins by liquid extraction-ultracentrifugation-liquid chromatography-mass spectrometry. Proteomics 4:3909-3920

Jarvis P (2004) Organellar proteomics: chloroplasts in the spotlight. Curr Biol 14:R317-R319

Jones AM, Bennett MH, Mansfield JW, Grant M (2006a) Analysis of the defence phosphoproteome of *Arabidopsis thaliana* using differential mass tagging. Proteomics 6:4155-4165

Jones AM, Thomas V, Bennett MH, Mansfield J, Grant M (2006b) Modifications to the *Arabidopsis* defense proteome occur prior to significant transcriptional change in response to inoculation with pseudomonas syringae. Plant Physiol 142:1603-1620

Joyard J, Teyssier E, Miege C, Berny-Seigneurin D, Marechal E, Block MA, Dorne AJ, Rolland N, Ajlani G, Douce R (1998) The biochemical machinery of plastid envelope membranes. Plant Physiol 118:715-723

Kapp EA, Schutz F, Connolly LM, Chakel JA, Meza JE, Miller CA, Fenyo D, Eng JK, Adkins JN, Omenn GS, Simpson RJ (2005) An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: sensitivity and specificity analysis. Proteomics 5:3475-3490

Kashino Y, Lauber WM, Carroll JA, Wang Q, Whitmarsh J, Satoh K, Pakrasi HB (2002) Proteomic analysis of a highly active photosystem II preparation from the cyanobacterium *Synechocystis* sp. PCC 6803 reveals the presence of novel polypeptides. Biochemistry 41:8004-8012

Kieselbach T, Bystedt M, Hynds P, Robinson C, Schroder WP (2000) A peroxidase homologue and novel plastocyanin located by proteomics to the *Arabidopsis* chloroplast thylakoid lumen. FEBS Lett 480:271-276

Kieselbach T, Hagman, Andersson B, Schroder WP (1998) The thylakoid lumen of chloroplasts. Isolation and characterization. J Biol Chem 273:6710-6716

Kieselbach T, Schroder WP (2003) The proteome of the chloroplast lumen of higher plants. Photosynth Res 78:249-264

Kikuchi S, Hirohashi T, Nakai M (2006) Characterization of the preprotein translocon at the outer envelope membrane of chloroplasts by blue native PAGE. Plant Cell Physiol 47:363-371

Kleffmann T, Hirsch-Hoffmann M, Gruissem W, Baginsky S (2006) plprot: a comprehensive proteome database for different plastid types. Plant Cell Physiol 47:432-436

Kleffmann T, Russenberger D, von Zychlinski A, Christopher W, Sjolander K, Gruissem W, Baginsky S (2004) The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. Curr Biol 14:354-362

Koo AJ, Ohlrogge JB (2002) The predicted candidates of *Arabidopsis* plastid inner envelope membrane proteins and their expression profiles. Plant Physiol 130:823-836

Kuchler M, Decker S, Hormann F, Soll J, Heins L (2002) Protein import into chloroplasts involves redox-regulated proteins. Embo J 21:6136-6145

Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates JR 3rd (1999) Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol 17:676-682

Lonosky PM, Zhang X, Honavar VG, Dobbs DL, Fu A, Rodermel SR (2004) A proteomic analysis of maize chloroplast biogenesis. Plant Physiol 134:560-574

Lunn JE (2006) Compartmentation in plant metabolism. J Exp Bot 58:35-47

Ma B, Zhang K, Hendrie C, Liang C, Li M, Doherty-Kirby A, Lajoie G (2003) PEAKS: powerful software for peptide *de novo* sequencing by tandem mass spectrometry. Rapid Commun Mass Spectrom 17:2337-2342

Majeran W, Cai Y, Sun Q, van Wijk KJ (2005) Functional differentiation of bundle sheath and mesophyll maize chloroplasts determined by comparative proteomics. Plant Cell 17:3111-3140

Mann M, Hojrup P, Roepstorff P (1993) Use of mass spectrometric molecular weight information to identify proteins in sequence databases. Biol Mass Spectrom 22:338-345

Mann M, Wilm M (1994) Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Anal Chem 66:4390-4399

Millar AH, Whelan J, Small I (2006) Recent surprises in protein targeting to mitochondria and plastids. Curr Opin Plant Biol 9:610-615

Moseley JL, Allinger T, Herzog S, Hoerth P, Wehinger E, Merchant S, Hippler M (2002) Adaptation to Fe-deficiency requires remodeling of the photosynthetic apparatus. Embo J 21:6709-6720

Naumann B, Stauber EJ, Busch A, Sommer F, Hippler M (2005) N-terminal processing of Lhca3 Is a key step in remodeling of the photosystem I-light-harvesting complex under iron deficiency in *Chlamydomonas reinhardtii*. J Biol Chem 280:20431-20441

Newton RP, Brenton AG, Smith CJ, Dudley E (2004) Plant proteome analysis by mass spectrometry: principles, problems, pitfalls and recent developments. Phytochemistry 65:1449-1485

Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10:1-6

Oda Y, Huang K, Cross FR, Cowburn D, Chait BT (1999) Accurate quantitation of protein expression and site-specific phosphorylation. Proc Natl Acad Sci USA 96:6591-6596

Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1:376-386

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85:2444-2448

Peltier JB, Cai Y, Sun Q, Zabrouskov V, Giacomelli L, Rudella A, Ytterberg AJ, Rutschow H, van Wijk KJ (2006) The oligomeric stromal proteome of *Arabidopsis thaliana* chloroplasts. Mol Cell Proteomics 5:114-133

Peltier JB, Emanuelsson O, Kalume DE, Ytterberg J, Friso G, Rudella A, Liberles DA, Soderberg L, Roepstorff P, von Heijne G, van Wijk KJ (2002) Central functions of the lumenal and peripheral thylakoid proteome of *Arabidopsis* determined by experimentation and genome-wide prediction. Plant Cell 14:211-236

Peltier JB, Friso G, Kalume DE, Roepstorff P, Nilsson F, Adamska I, van Wijk KJ (2000) Proteomics of the chloroplast: systematic identification and targeting analysis of lumenal and peripheral thylakoid proteins. Plant Cell 12:319-341

Peltier JB, Ripoll DR, Friso G, Rudella A, Cai Y, Ytterberg J, Giacomelli L, Pillardy J, van Wijk KJ (2004a) Clp protease complexes from photosynthetic and non-photosynthetic plastids and mitochondria of plants, their predicted three-dimensional structures, and functional implications. J Biol Chem 279:4768-4781

Peltier JB, Ytterberg AJ, Sun Q, van Wijk KJ (2004b) New functions of the thylakoid membrane proteome of *Arabidopsis thaliana* revealed by a simple, fast, and versatile fractionation strategy. J Biol Chem 279:49367-49383

Peltier JB, Ytterberg J, Liberles DA, Roepstorff P, van Wijk KJ (2001) Identification of a 350-kDa ClpP protease complex with 10 different Clp isoforms in chloroplasts of *Arabidopsis thaliana*. J Biol Chem 276:16318-16327

Peng J, Gygi SP (2001) Proteomics: the move to mixtures. J Mass Spectrom 36:1083-1091

Perez-Bueno ML, Rahoutei J, Sajnani C, Garcia-Luque I, Baron M (2004) Proteomic analysis of the oxygen-evolving complex of photosystem II under biotec stress: Studies on *Nicotiana benthamiana* infected with tobamoviruses. Proteomics 4:418-425

Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551-3567

Phee BK, Cho JH, Park S, Jung JH, Lee YH, Jeon JS, Bhoo SH, Hahn TR (2004) Proteomic analysis of the response of *Arabidopsis* chloroplast proteins to high light stress. Proteomics 4:3560-3568

Radhamony RN, Theg SM (2006) Evidence for an ER to Golgi to chloroplast protein transport pathway. Trends Cell Biol 16:385-387

Rexroth S, Meyer zu Tittingdorf JM, Krause F, Dencher NA, Seelert H (2003) Thylakoid membrane at altered metabolic state: challenging the forgotten realms of the proteome. Electrophoresis 24:2814-2823

Richly E, Leister D (2004) An improved prediction of chloroplast proteins reveals diversities and commonalities in the chloroplast proteomes of *Arabidopsis* and rice. Gene 329:11-16

Rolland N, Ferro M, Ephritikhine G, Marmagne A, Ramus C, Brugiere S, Salvi D, Seigneurin-Berny D, Bourguignon J, Barbier-Brygoo H, Joyard J, Garin J (2006) A versatile method for deciphering plant membrane proteomes. J Exp Bot 57:1579-1589

Rolland N, Ferro M, Seigneurin-Berny D, Garin J, Douce R, Joyard J (2003) Proteomics of chloroplast envelope membranes. Photosynth Res 78:205-230

Rudella A, Friso G, Alonso JM, Ecker JR, van Wijk KJ (2006) Downregulation of ClpR2 leads to reduced accumulation of the ClpPRS protease complex and defects in chloroplast biogenesis in *Arabidopsis*. Plant Cell 18:1704-1721

Savitski MM, Nielsen ML, Zubarev RA (2005) New data base-independent, sequence tag-based scoring of peptide MS/MS data validates Mowse scores, recovers below threshold data, singles out modified peptides, and assesses the quality of MS/MS techniques. Mol Cell Proteomics 4:1180-1188

Schleiff E, Eichacker LA, Eckart K, Becker T, Mirus O, Stahl T, Soll J (2003) Prediction of the plant beta-barrel proteome: a case study of the chloroplast outer envelope. Protein Sci 12:748-759

Schubert M, Petersson UA, Haas BJ, Funk C, Schroder WP, Kieselbach T (2002) Proteome map of the chloroplast lumen of *Arabidopsis thaliana*. J Biol Chem 277:8354-8365

Seigneurin-Berny D, Rolland N, Garin J, Joyard J (1999) Technical advance: differential extraction of hydrophobic proteins from chloroplast envelope membranes: a subcellular-specific proteomic approach to identify rare intrinsic membrane proteins. Plant J 19:217-228

Shadforth I, Crowther D, Bessant C (2005) Protein and peptide identification algorithms using MS for use in high-throughput, automated pipelines. Proteomics 5:4082-4095

Shevchenko A, Jensen ON, Podtelejnikov AV, Sagliocco F, Wilm M, Vorm O, Mortensen P, Boucherie H, Mann M (1996) Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc Natl Acad Sci USA 93:14440-14445

Small I, Peeters N, Legeai F, Lurin C (2004) Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics 4:1581-1590

Spengler B (2004) *De novo* sequencing, peptide composition analysis, and composition-based sequencing: a new strategy employing accurate mass determination by fourier transform ion cyclotron resonance mass spectrometry. J Am Soc Mass Spectrom 15:703-714

Srivastava R, Pisareva T, Norling B (2005) Proteomic studies of the thylakoid membrane of *Synechocystis* sp. PCC 6803. Proteomics 5:4905-4916

Stauber EJ, Fink A, Markert C, Kruse O, Johanningmeier U, Hippler M (2003) Proteomics of *Chlamydomonas reinhardtii* light-harvesting proteins. Eukaryot Cell 2:978-994

Steen H, Pandey A (2002) Proteomics goes quantitative: measuring protein abundance. Trends Biotechnol 20:361-364

Storf S, Stauber EJ, Hippler M, Schmid VH (2004) Proteomic analysis of the photosystem I light-harvesting antenna in tomato (*Lycopersicon esculentum*). Biochemistry 43:9214-9224

Sun Q, Emanuelsson O, van Wijk KJ (2004) Analysis of curated and predicted plastid subproteomes of *Arabidopsis*. Subcellular compartmentalization leads to distinctive proteome properties. Plant Physiol 135:723-734

Sunyaev S, Liska AJ, Golod A, Shevchenko A (2003) MultiTag: multiple error-tolerant sequence tag search for the sequence-similarity identification of proteins by mass spectrometry. Anal Chem 75:1307-1315

Tabb DL, Saraf A, Yates JR 3rd (2003) GutenTag: high-throughput sequence tagging via an empirically derived fragmentation model. Anal Chem 75:6415-6421

Takahashi Y, Yasui TA, Stauber EJ, Hippler M (2004) Comparison of the subunit compositions of the PSI-LHCI supercomplex and the LHCI in the green alga *Chlamydomonas reinhardtii*. Biochemistry 43:7816-7823

Taylor JA, Johnson RS (2001) Implementation and uses of automated *de novo* peptide sequencing by tandem mass spectrometry. Anal Chem 73:2594-2604

Timperio AM, Huber CG, Zolla L (2004) Separation and identification of the light harvesting proteins contained in grana and stroma thylakoid membrane fractions. J Chromatogr A 1040:73-81

Timperio AM, Zolla L (2005) Investigation of the lateral light-induced migration of photosystem II light-harvesting proteins by nano-high performance liquid chromatography electrospray ionization mass spectrometry. J Biol Chem 280:28858-28866

Turkina MV, Blanco-Rivero A, Vainonen JP, Vener AV, Villarejo A (2006a) CO2 limitation induces specific redox-dependent protein phosphorylation in *Chlamydomonas reinhardtii*. Proteomics 6:2693-2704

Turkina MV, Kargul J, Blanco-Rivero A, Villarejo A, Barber J, Vener AV (2006b) Environmentally modulated phosphoproteome of photosynthetic membranes in the green alga *Chlamydomonas reinhardtii*. Mol Cell Proteomics 5:1412-1425

Turkina MV, Villarejo A, Vener AV (2004) The transit peptide of CP29 thylakoid protein in *Chlamydomonas reinhardtii* is not removed but undergoes acetylation and phosphorylation. FEBS Lett 564:104-108

Unlu M, Morgan ME, Minden JS (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. Electrophoresis 18:2071-2077

van Wijk KJ (2004) Plastid proteomics. Plant Physiol Biochem 42:963-977

Vener AV, Harms A, Sussman MR, Vierstra RD (2001) Mass spectrometric resolution of reversible protein phosphorylation in photosynthetic membranes of *Arabidopsis thaliana*. J Biol Chem 276:6959-6966

Vidi PA, Kanwischer M, Baginsky S, Austin JR, Csucs G, Dormann P, Kessler F, Brehelin C (2006) Tocopherol cyclase (VTE1) localization and vitamin E accumulation in chloroplast plastoglobule lipoprotein particles. J Biol Chem 281:11225-11234

Westerlund I, Von Heijne G, Emanuelsson O (2003) LumenP--a neural network predictor for protein localization in the thylakoid lumen. Protein Sci 12:2360-2366

Whitelegge JP, Laganowsky A, Nishio J, Souda P, Zhang H, Cramer WA (2006) Sequencing covalent modifications of membrane proteins. J Exp Bot 57:1515-1522

Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, Hochstrasser DF (1999) Protein identification and analysis tools in the ExPASy server. Methods Mol Biol 112:531-552

Wilson KA, McManus MT, Gordon ME, Jordan TW (2002) The proteomics of senescence in leaves of white clover, *Trifolium repens* (L.). Proteomics 2:1114-1122

Wittmann-Liebold B, Graack HR, Pohl T (2006) Two-dimensional gel electrophoresis as tool for proteomics studies in combination with protein identification by mass spectrometry. Proteomics 6:4688-4703

Yamaguchi K, Beligni MV, Prieto S, Haynes PA, McDonald WH, Yates JR 3rd, Mayfield SP (2003) Proteomic characterization of the *Chlamydomonas reinhardtii* chloroplast ribosome. Identification of proteins unique to the 70 S ribosome. J Biol Chem 278:33774-33785

Yamaguchi K, Prieto S, Beligni MV, Haynes PA, McDonald WH, Yates JR 3rd, Mayfield SP (2002) Proteomic characterization of the small subunit of *Chlamydomonas* *reinhardtii* chloroplast ribosome: identification of a novel S1 domain-containing protein and unusually large orthologs of bacterial S2, S3, and S5. Plant Cell 14:2957-2974

Yamaguchi K, Subramanian AR (2000) The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast). J Biol Chem 275:28466-28482

Yamaguchi K, von Knoblauch K, Subramanian AR (2000) The plastid ribosomal proteins. Identification of all the proteins in the 30 S subunit of an organelle ribosome (chloroplast). J Biol Chem 275:28455-28465

Yan B, Pan C, Olman VN, Hettich RL, Xu Y (2005) A graph-theoretic approach for the separation of b and y ions in tandem mass spectra. Bioinformatics 21:563-574

Ytterberg AJ, Peltier JB, van Wijk KJ (2006) Protein profiling of plastoglobules in chloroplasts and chromoplasts. A surprising site for differential accumulation of metabolic enzymes. Plant Physiol 140:984-997

Zhao C, Wang J, Cao M, Zhao K, Shao J, Lei T, Yin J, Hill GG, Xu N, Liu S (2005) Proteomic changes in rice leaves during development of field-grown rice plants. Proteomics 5:961-972

Zolla L, Rinalducci S, Timperio AM, Huber CG (2002) Proteomics of light-harvesting proteins in different plant species. Analysis and comparison by liquid chromatography-electrospray ionization mass spectrometry. Photosystem I. Plant Physiol 130:1938-1950

Zolla L, Timperio AM (2000) High performance liquid chromatography-electrospray mass spectrometry for the simultaneous resolution and identification of intrinsic thylakoid membrane proteins. Proteins 41:398-406

Zolla L, Timperio AM, Walcher W, Huber CG (2003) Proteomics of light-harvesting proteins in different plant species. Analysis and comparison by liquid chromatography-electrospray ionization mass spectrometry. Photosystem II. Plant Physiol 131:198-214

Hippler, Michael
Institute of Plant Biochemistry and Biotechnology, University of Muenster, Hindenburgplatz 55, 48143 Muenster, Germany
mhippler@uni-muenster.de

Naumann, Bianca
Institute of Plant Biochemistry and Biotechnology, University of Muenster, Hindenburgplatz 55, 48143 Muenster, Germany

List of abbreviations

2-DE: two dimensional gel electrophoresis
BN-PAGE: blue-native polyacrylamide gel electrophoresis
CCM: CO_2 concentrating mechanism
CN-PAGE: colorless-native polyacrylamide gel electrophoresis
cTP: chloroplast transit peptide

cytb$_6$f: cytochrome b$_6$f complex
DIGE: fluorescence two dimensional Difference Gel Electrophoresis
ESI: electrospray ionization
GPF: genomic peptide finder
HPLC: high performance liquid chromatography
IB: immunoblot
ICAT: isotope-coded affinity tag
IEF: isoelectric focusing
IMAC: immobilized metal affinity chromatography
IB: immunoblot
IP: isoelectric point
LC: liquid chromatography
LHCI: light harvesting complex I
LHCII: light harvesting complex II
lTP: lumenal transit peptide
MALDI: matrix assisted laser desorption ionisation
MS/MS: tandem mass spectrometry
MS: mass spectrometry, mass spectrometer
mudPIT: multidimensional protein identification technology
PAGE: polyacrylamide gel electrophoresis
PG: plastoglobuli
PMF: peptide mass finger printing
PSI: photosystem I
PSII: photosystem II
PTM: post-translational modification
ROS: reactive oxygen species
RP: reversed phase
SCX: strong cation exchange
SDS: sodium dodecyl sulfate
SILAC: stable isotope-labeling of amino acids in cell culture
TMD: transmembrane domain
TOF: time of flight
TPP: three phase partitioning
WT: wild type

Plastid-nucleus communication: anterograde and retrograde signalling in the development and function of plastids

Katharina Bräutigam, Lars Dietzel, and Thomas Pfannschmidt

Abstract

Plastids are organelles that are a unique feature of plant cells. They represent an important metabolic and genetic compartment that is essential for almost all aspects in the life of a plant. Its endosymbiotic origin requires the establishment of novel signalling pathways between the organelle and the nucleus of the host cell. During evolution, therefore, a complex regulatory network evolved that couples development and function of the organelles to that of the cell. Nowadays, the nucleus controls most aspects of plastids by providing proteins essential for plastid processes. This 'anterograde' signalling, however, is complemented by a backward flow of information from the plastid to the nucleus. This 'retrograde' signalling represents a feedback control that reports the functional state of the organelle to the nucleus. This means that extensive communication between the two compartments is established. This helps the plant to perceive and respond properly to varying environmental influences and to developmental signals at the cellular level. The interaction and mutual dependency of anterograde and retrograde signals are discussed with respect to recent observations. Models are presented that provide a unifying view of the different known pathways.

1 Introduction

Plastids are organelles typical for plant cells. They originated from an endosymbiotic event in which a heterotrophic eukaryotic cell engulfed a photosynthetic cyanobacterium. During establishment of this endosymbiosis, the cyanobacterium lost the majority of its genes to the nucleus of the host cell, but remained a site for photosynthesis and other important biochemical pathways. This gene transfer gave the host cell control over development and function of the endosymbiont. The evolutionary result (around two billion years ago) was an autotrophic, eukaryotic cell able to perform photosynthesis with the help of a new organelle. Three lines of organisms evolved from this cell: the glaucocystophyta, the rhodophyta (the "red line"), and the chlorophyta (the "green line"). From the latter, the plants evolved 450–500 million years ago (Delwiche 1999; Stoebe and Maier 2002; Timmis et al. 2004). Typical higher plant plastids of today are surrounded by two

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0243 / Published online: 6 July 2007

membranes, the inner membrane originated from the cyanobacterial ancestor, the outer membrane from the engulfing host cell. They exhibit an astonishingly high morphological plasticity depending on the tissue where they are located, for example, green chloroplasts are found within photosynthetic tissues while coloured chromoplasts are present within flowers or fruits (Buchanan et al. 2002). Plastids still contain a genome of their own, the so-called plastome that is always the same independent from the morphological form. It is generally believed to be of plasmid-like structure with a size of around 120–200 kb, however, ultra-structural investigations suggest much more variable structures and sizes (Bendich 2004). The encoded set of genes is relatively conserved among higher plants and covers ca. 100–130 different genes. These mainly encode components for the photosynthetic apparatus and the plastid-own gene expression machinery that is responsible for the expression of the plastome (Sugiura 1992). At present 92 different plastomes have been sequenced (see http://www.ncbi.nlm.nih.gov/genomes/ORGANELLES/), providing much information for our understanding of the phylogeny of plastids (Martin et al. 2002; Howe et al. 2003). The great majority of the plastid proteome, however, is encoded in the nucleus; therefore, plastids are generally regarded to be genetically semi-autonomous. Depending on plastid type and developmental or functional state of the plant estimates range from 2500–4500 different proteins being located in plastids (Abdallah et al. 2000; Kleffmann et al. 2004; van Wijk 2004). Beside their major role as site for photosynthesis plastids are involved in almost all biosynthetic pathways of the cell making them an integral functional and biochemical cell compartment that is essential for the life of a plant.

Many aspects touched by this article are reviewed in much more detail by other contributions to this book and the reader interested in such details is referred to them. This review aims to discuss the problems that higher plant cells have to overcome in the communication between the nucleus and the plastids as well as in the coordination of the different gene expression machineries. Furthermore, it summarizes the current knowledge on the signalling routes acting either from the nucleus toward the plastid, the forward or anterograde signalling, or from the plastid to the nucleus, the backward or retrograde signalling.

2 Major problems of coordination and communication between plastids and nucleus

The evolutionary integration of an additional genetic and biochemical compartment generated a number of problems for the eukaryotic host cell that were solved by a massive gene transfer from the endosymbiont to the nucleus, a re-organisation of metabolic pathways in the cell, and the establishment of novel signalling routes controlling development and function of the new organelle (Martin and Schnarrenberger 1997). With the development of multicellular organisms possessing various plastid forms these signalling routes have been supplemented with additional pathways controlling tissue-specific formation of plastid forms. Plant cells of today exhibit a complex signalling network that helps plastids to function in a controlled fashion depending on the demands of the respective cell or tissue. The task of this signalling network is focussed on three major problems in the co-ordination and communication between plastids and the nucleus of a cell: tissue-specific development, highly different gene copy numbers and integration of varying signals (Fig. 1).

2.1 Tissue specificity of plastid development

As already mentioned plastids can develop into various different forms (Herrmann et al. 1992; Waters and Pyke 2004; Lopez-Juez and Pyke 2005). Starting from an inherited undifferentiated proplastid, chloroplasts (for photosynthesis), amyloplasts (for starch storage), elaioplasts (for oil storage), or chromoplasts (in coloured tissues) develop. The decisive determinant for the development of a specific plastid type is the tissue environment of the host cell. Thus, the nucleus of a cell has to receive signals from neighbouring cells reporting the surrounding cellular context. In addition, it also has to receive environmental signals such as light that initiate, for example, phytochrome-mediated photomorphogenesis including etioplast-chloroplast transition. These endogenous and exogenous signals (Fig. 1) finally control the fate of a cell and of all organelles inside. Typically in a cell (regardless which kind of cell) only one type of plastid can be found indicating that plastid and cell development are tightly coupled. This requires a highly sophisticated integration of developmental and environmental signals that influences the respective plastid development. Such a complex integration suggests exclusive anterograde control of plastid development, however, several mutants with plastid defects also exhibit clear changes in leaf tissue morphology (Aluru et al. 2006). These observations suggest that retrograde signalling from plastids affects, in turn, tissue development pointing to a mutual control (see below).

2.2 The gene copy number problem

Typically, multi-subunit protein complexes in plastids consist of a patchwork of plastid and nuclear encoded subunits. A good example for this is the subunit composition of the photosystems. The core proteins are uniformly encoded in the plastome while the peripheral subunits are encoded in the nucleus (Race et al. 1999). A correct assembly of such complexes, therefore, requires a precise coordination in the expression of photosynthesis genes between the two compartments to result in stoichiometric amounts of protein subunits. Difficulties in this event, however, are not only generated by the spatial separation of the coding compartments, but also by the highly different gene copy number. A nuclear encoded photosynthetic subunit normally is encoded by a single gene or by a pair of duplicated genes. The latter is true for around 70% of all nuclear genes in *Arabidopsis* (The *Arabidopsis* Genome Initiative 2000). In contrast, a plastid encoded subunit might be encoded by up to 100 copies of the same gene since the plastome is highly polyploidic

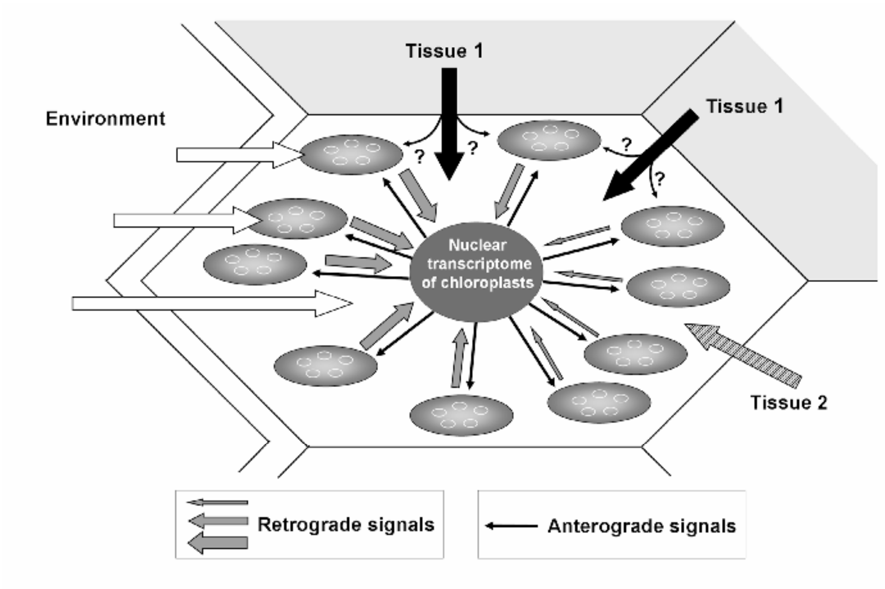


Fig. 1. Problems of coordination and communication between plastids and nucleus. The scheme depicts a higher plant cell with a nucleus (central circle) surrounded by several plastids (ovals). On the left the cell adjoins the environment (separated by a double line), on top and on the right it adjoins cells of other tissues. External signals from these areas enter the cell (large arrows representing signals from the environment (white), neighbouring tissue 1 (black), and a different neighbouring tissue 2 (hatched)). These external signals are perceived by appropriate receptors in cell membrane, cytosol, or by the plastids itself. The interplay between perception of the cell and the plastids is largely unknown, for example, it is unknown whether signals from neighbouring tissue are detected only by the cell or also by plastids in parallel (indicated by question marks). The environmental factor light is perceived at the same time by cytosolic photoreceptors and photosynthesis within plastids, however, the integration of resulting responses is largely unknown. Anterograde signals from the nucleus and retrograde signals from the plastids (indicated in the legend) contribute to an extensive communication between the two compartments leading to an integrated response on cell-external signals. Gradients in such external signals, for example, light quality gradients might initiate retrograde signals of different strengths depending on the cellular position of the perceiving plastid.

(Sugiura 1992). If one considers that a single cell may contain up to 100 plastids this theoretically generates an up to 10,000-fold excess of a single plastid gene over a single nuclear gene (indicated by several white circles in the plastids in Fig. 1). For instance, in *Arabidopsis* the assembly of photosystem II requires a coordinated expression of, for example, the reaction centre protein D1 (encoded by the single plastid gene *psbA*) and the 33 kD protein of the water splitting apparatus (encoded by the two nuclear genes *PsbO1* and *PsbO2*). Despite the discrepancy in gene copy number the plant cell is able to express these PSII components in such a way that functional photosystem II particles are assembled without extensive accumulation of un-needed subunits. Thus, coordination processes have to integrate the activities of two evolutionary divergent gene expression machineries (eukaryotic in the nucleus and (mainly) prokaryotic in the plastid) in two different compartments with extreme different gene copy numbers.

2.3 Integration of plastid responses within the cell

As mentioned above, higher plant cells contain up to 100 plastids that are distributed over the whole cytoplasm. Because of gradients in intensity or quality of endogenous or exogenous signals (e.g. hormones or light) within tissues or cells these plastids are not all in an identical functional state. This, therefore, might cause differences in the retrograde signals of the individual plastids even in the same cell (Fig. 1, indicated by different sizes of retrograde arrows). When these signals of different strength or nature reach the nucleus, each single signal is either processed separately or the different signals are integrated into one coarse signal of more general character. The first possibility would require the ability of the nucleus to localise the origin of the different signals and to direct the respective response in an appropriate manner. The other possibility would result in an averaged response and the different demands of the single plastids have to be considered by another mechanism such as by the respective plastid itself.

3 Anterograde signalling

3.1 The nuclear control principle

As outlined above, the nucleus encodes almost all plastid proteins. This provides a very simple way of control. Only those processes in plastids can be active for which all necessary protein components are expressed properly in nucleus and cytosol and transported into the organelle. Thus, control of nuclear gene expression decides about plastid type and its metabolic activities (Gruissem 1989; Emes and Tobin 1993; Leon et al. 1998). Despite the simplicity of this regulation principle the control mechanisms itself are very complex. Only a few hundred enzymes are really necessary to perform the biochemical steps in plastids (Neuhaus and Emes 2000) but a few thousand proteins appear to be required for the regulation. Genomic and proteomic approaches in the last decade have uncovered a much higher multiplicity of proteins within plastids than anticipated so far (Abdallah et al. 2000; Kleffmann et al. 2004; Peltier et al. 2004; van Wijk 2004). At present, no distinct function has been determined for many of the identified components. The main reasons are incomplete data base information, the low abundance of many regulatory proteins makes biochemical analyses difficult, and the fact that many proteins and their functions are still unknown. Nevertheless, all the proteins that have so far been identified to be involved in the regulation of plastid development and function can be assigned to a few major functional categories. These categories include: i) establishment of plastid proteins or protein complexes including import, sorting, and assembly, ii) expression and regulation of the plastid genome, iii) modification of plastid enzyme or pathway activities by, for example, post-translational modifications, and iv) plastid division. Furthermore, plastid development is affected by tissue-specific and environmental influences, which may involve proteins acting in the categories mentioned above. All these topics are reviewed in other chapters in this issue and, therefore, are not provided here in molecular detail. The reader with deeper insights in such research fields is referred to these contributions.

Nevertheless, some of such anterograde signalling pathways or mechanisms are of great importance to our understanding of the mutual communication between plastids and nucleus. Since this communication is decisive for plastid development and function in general, relevant anterograde signalling pathways are discussed more deeply. In this way, we aim to analyse the present knowledge on both signalling directions in order to provide a unifying view.

3.1.1 Anterograde control of plastid protein import

All nuclear encoded plastid proteins have to be imported into the organelle. This requires that the proteins pass the surrounding double membrane of the plastids. During evolution most proteins acquired N-terminal transit peptides that direct the protein through the Tic-Toc (*translocon of inner chloroplast membrane, translocon of outer chloroplast membrane*) complex to the respective inner-plastidial location, i.e., in chloroplasts to stroma, thylakoid membrane, or lumen (Soll 2002). The protein content of a plastid, however, is highly dependent on the general developmental stage and function of the respective tissue, for example, photosynthetic leaf or non-photosynthetic root. Accordingly, anterograde control has to determine which proteins are imported and, therefore, which plastid type is developed in a given tissue. Developmental and spatial specificity of this process is achieved by varying subunit composition of the Tic-Toc complexes resulting in different substrate specificities (Soll and Schleiff 2004). Since Tic-Toc complexes consist of nuclear encoded components the protein import machinery is under exclusive nuclear control. It has been demonstrated that distinct but homologous Tic-Toc import pathways exist that are responsible for the import of specific classes of substrates (Kessler and Schnell 2006). However, the import into chloroplasts is suggested to be sensitive to the redox state of the organelle, which could couple the import of proteins to the functional state of the plastid (Küchler et al. 2002). This also indicates a possible environmental control of plastid protein import.

3.1.2 Nuclear control of plastid transcription

The number of the plastid encoded proteins (87 in *Arabidopsis thaliana*) is relatively small when compared to that of nuclear encoded ones. However, they are of exceptional importance for plastid function in general and, therefore, their expression is strongly regulated to be expressed properly (Barkan and Goldschmidt-Clermont 2000). Plastidic genes can be roughly grouped into photosynthesis genes and genes for components of the plastid gene expression machinery such as RNA polymerases or ribosomes. In addition, tRNAs and rRNAs are also encoded (Sugiura 1995). The gene composition of the plastome is relatively conserved among higher plants and always includes the core subunits of the photosynthesis protein complexes that are the pacemakers for the assembly of the large multi-subunit complexes of the thylakoid membrane (Race et al. 1999). Expression of these genes, therefore, has an important influence on early chloroplast development when the photosynthetic machinery is being built up.

Plastid transcription is performed by two different RNA polymerase activities, a nuclear encoded phage-like single subunit polymerase called NEP (for nuclear encoded polymerase) and a plastid encoded bacteria-like RNA polymerase called PEP (for plastid encoded polymerase) (Hess and Börner 1999; Liere and Maliga 2001). The latter enzyme is a multi-subunit complex that is formed by proteins from the plastid genes *rpoA*, *rpoB*, *rpoC1*, and *rpoC2*. This complex represents the so-called core-enzyme that is able to catalyse transcription but that depends on the activity of nuclear encoded sigma factors for specific promoter recognition (Link 1996; Allison 2000). Thus, the plastid-encoded enzyme is still under nuclear control. In *Arabidopsis*, we know six different sigma factors at present (Shiina et al. 2005). In addition, biochemical studies suggest further structural improvement of the PEP core enzyme during light-induced etioplast-chloroplast transition by addition of other subunits with further properties, for example, a CKII kinase activity or a superoxide dismutase (Pfannschmidt et al. 2000; Ogrzewalla et al. 2002; Suzuki et al. 2004). Furthermore, a number of eukaryotic-like transcription factors do likely exist in plastids, which may interact with the PEP enzyme (Sato 2001; Wagner and Pfannschmidt 2006; Schwacke et al. 2007). The function of such additional factors is largely unknown, but it is proposed that they may mediate environmental influences to PEP (Sato 2001; Wagner and Pfannschmidt 2006; Schwacke et al. 2007).

The NEP enzyme is encoded by three different nuclear genes (*RpoTs*). One protein is targeted to mitochondria (*RpoTm*), one to plastids (*RpoTp*), and one to both organelles (*RpoTmp*). Therefore, two NEP enzymes exist in plastids (Liere and Börner 2006). NEP and PEP act in close interrelationship during plastid development since the plastid *rpo* genes are exclusively transcribed by the NEP enzyme. When PEP has been established it starts to transcribe genes for photosynthesis. Current models suggest a cascade of transcription events in the early plastid development with a high NEP activity in young and undifferentiated cells, which then declines in parallel to the establishment of the PEP enzyme. The importance of the NEP activity can be observed in the *SCABRA3* mutant of *Arabidopsis* (Hricova et al. 2006). The *SCABRA3* gene encodes the RpoTp RNA polymerase. Its lack causes severely impaired plant growth and reduced pigmentation of all green tissues. Recent work on a *Sig2* knockout mutant uncovered a further interesting relationship. Sig2 was found to be responsible for the exclusive recognition of a number of tRNA promoters by PEP (Kanamaru et al. 2001; Kanamaru and Tanaka 2004). Among these the *trnE* gene was identified that encodes the glutamyl-tRNA. This molecule is not only involved in plastid translation but also represents the

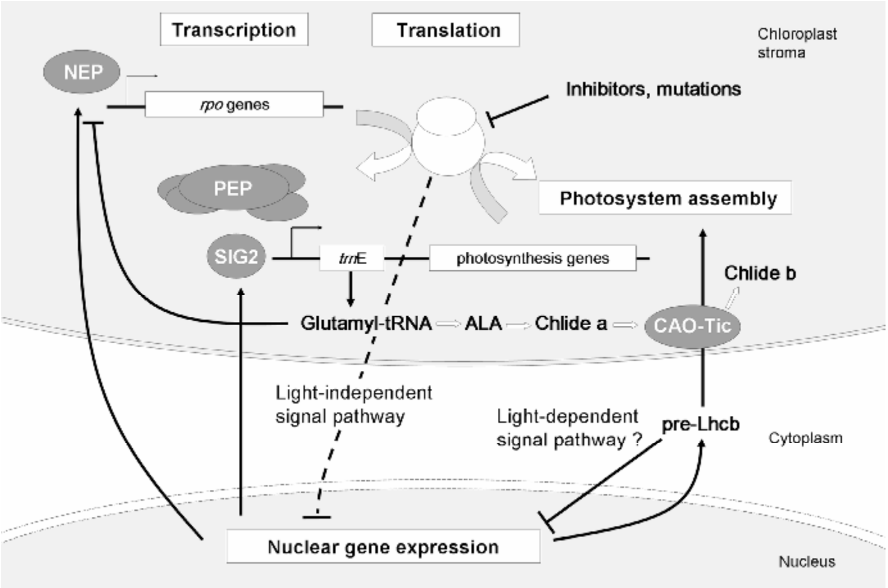


Fig. 2. Retrograde signals from plastid gene expression. The cellular compartments nucleus, cytosol, and chloroplast as well as a plastid ribosome (double white oval) are depicted schematically. Genes are given as white bars on black lines with a thin arrow indicating the transcription start. Thick black arrows represent signalling pathways, broken lines indicate putative or still unclear branches. White arrows indicate synthesis, white grey arrows represent translation, affected processes are in white boxes, identified or potential candidates for intermediate protein signalling components are shown as grey ovals with white letters. Three nuclear encoded components (NEP, SIG2, pre-Lhcb) indicate anterograde influences or signals on plastid gene expression. SIG2 is necessary for PEP-mediated transcription of *trnE* that functions both as 5-aminolevulinic (ALA) precursor and as inhibitor of NEP activity upon maturation of the plastid (compare Section 3.1.1). Inhibition of ALA synthesis may negatively affect the import of Lhcb proteins, which could repress *Lhcb* transcription. Inhibition of translation by either inhibitors or mutations indirectly causes the same effects by inhibition of PEP subunit expression. For further functional details see text.

precursor molecule for amino levulinic acid that is required for chlorophyll synthesis (compare Fig. 2). Consistently, downregulation of *trnE* expression resulted in a pale phenotype of the mutants. Surprisingly, these mutants exhibited high NEP activity also in older plastids. Further work demonstrated that glutamyl-tRNA molecules bind to NEP and cause its functional downregulation in chloroplasts (Hanaoka et al. 2005). Thus, with the NEP-dependent onset of PEP a negative feedback mechanism on NEP activity by PEP-dependent transcription of the *trnE* gene is initiated (Fig. 2). However, the NEP enzyme is not completely inactive in green tissues since it is required for the exclusive expression of some specific genes (*clpP, adh*) (Liere and Börner 2006). This complex mutual regulation of NEP and PEP activity is controlled by the action of nuclear encoded components demonstrating that the nucleus exerts a very strong influence on plastid transcription.

3.1.3 Nuclear encoded proteins in posttranscriptional regulation events

Posttranscriptional processes represent a further major level of regulation in plastid gene expression (Deng and Gruissem 1987; Mullet and Klein 1987). Studies from the last twenty years indicate that plastid transcripts undergo an extensive and complex maturation process that includes the splicing of introns, the maturation of poly-cistronic primary transcripts and the stabilisation or degradation of the mature transcript by binding of regulatory proteins to hair-pin structures at both 5'- and 3'-ends (Nickelsen 2003; Bollenbach et al. 2004). In addition, several transcript editing sites have been discovered in plastids (Freyer et al. 1997). All these processing events involve specific proteins that, as far as identified, are encoded exclusively in the nucleus. In the case of the editing process, it was found that each editing site has its specific protein responsible for the respective post-transcriptional step (Bock and Koop 1997). Proteins of the pentatricopeptide repeat (PPR) family were identified to be involved in the mediation of this specificity (Shikanai 2006). Furthermore, PPR proteins were also found to facilitate the trans-splicing of *rps12* pre-mRNA in maize (Schmitz-Linneweber et al. 2006). Many basic components of translation such as rRNA as well as protein components of ribosomes and all tRNAs are encoded in the plastids. Nevertheless, a great number of ribosomal proteins are imported from the cytosol including translation initiation factors. Much work in this research field has been done in *Chlamydomonas*. Abundant evidence demonstrates that translation initiation is responsive to environmental influences. Furthermore, translation of many proteins (especially of the photosynthetic apparatus) might occur in parallel of complex assembly and/or protein folding and, therefore, must be seen in context with such processes (Cohen and Mayfield 1997; Danon 1997; Stern et al. 1997; Zhang et al. 2000). The control of posttranscriptional processes provides a wide field for anterograde control and we are still far away from complete understanding of the multiple interactions of these mechanisms (see related chapters in this book).

3.1.4 Posttranslational modifications controlling plastid metabolism

The activity of enzymes or protein complexes in plastids is largely regulated by posttranslational modifications such as phosphorylation of aminoacids or reduction/oxidation of sulfhydryl groups. These modifications have been found for many proteins in plastids and especially in chloroplasts. They require specific modifying proteins, for instance, kinases, phosphatases, disulfide isomerases, or dehydrogenases. Well-studied examples are the numerous regulation processes that optimise and protect the photosynthetic machinery of chloroplasts (Aro and Andersson 2001; Blankenship 2002). This also includes the targeted degradation of plastid proteins. Several chloroplast proteases have been characterised in the last decade, including Clp and FtsH proteases. Studies on knockout mutants revealed strong phenotypes (see Section 3.2.1) indicating that controlled degradation is of great importance for the biogenesis of chloroplasts (Adam et al. 2006).

Many of such modifying proteins are a prerequisite for sensing of environmental signals (see below). They perform their regulatory functions when all plastid machineries are built up and when the work-flow of these machineries has to be modified or regulated to optimise the respective function. Thus functional establishment of these proteins represents the final stage of the maturation of a plastid and marks the end of the individual plastid development. A fully matured plastid, now, is completely functional and the anterograde control of the nucleus is reduced to the supply of proteins that are required due to turnover, various repair mechanisms, or to specific demands of the plastid. This requires a permanent communication between nucleus and plastid, which is discussed in more detail below.

3.2 Developmental signals

The processes described in Section 3.1 provide a broad spectrum of targets for regulatory proteins affecting plastid development and function. Anterograde control can act in several of these processes in parallel, thus creating a complex network. Endosymbiosis occurred very early in evolution far before multi-cellular plants or algae evolved. The anterograde signalling network, therefore, was established in parallel with the signalling routes controlling the proper development of the multi-cellular body of a plant (Dyall et al. 2004). This most probably is the reason why the development of a distinct plastid type is so tightly coupled to the function of its respective host cell. Therefore, it is impossible to understand the molecular nature of anterograde signalling without including defined developmental or physiological models. Furthermore, one must distinguish between developmental signals that control the expression of plastid-localized proteins and developmental signals that affect the function of such proteins within the plastids. In many cases studies on mutants with defects in plastid development were very useful to understand the respective underlying molecular mechanisms. Some of them are discussed below.

3.2.1 Control of early chloroplast development

The expression of nuclear encoded plastid proteins occurs in a temporally and spatially coordinated manner that correlates with the tissue context of the host cell. Therefore, one of the most interesting developmental stages is the establishment of functionally differentiated plastids from their non-functional ancestors. Most studies reported so far are focussed on the development of chloroplasts.

Leaves of monocot plants such as barley grow from a basal meristem resulting in a gradient of cells of increasing age with the oldest cells at the top and the youngest at the bottom of a leaf. This provides a very useful physiological system to study the properties of chloroplasts of different age and allows for dissecting molecular processes that occur during the maturation from very young plastids to fully active chloroplasts. Although proteins are imported the whole lifetime of a plastid the most dramatic changes occur in the early stages of plastid development when the functional protein complement is built up and the plastid type and function is determined. By comparing base, middle, and top sections of four and six day-old barley leaves, it could be determined that in or near the basal meristematic region plastid DNA replication and plastid transcription is highest followed by cells building up the photosynthetic complexes (Baumgartner et al. 1989). These observations fit into the current model about the functional dependency of the PEP enzyme on the action of the NEP enzyme in early plastid development (compare Section 3.1.2, Liere and Maliga 2001).

In contrast to monocot plants, leaves of dicot plants do not have such a defined meristematic zone and grow in a more complex spreading way. Developmental studies can be performed here only by examining leaves in several consecutive steps during seedling development. In pea such studies uncovered a comparable sequence of events with stages of increased plastid DNA replication, followed by stages of high transcriptional rates (Dubell and Mullet 1995a, 1995b). The spreading leaf development of dicot plants is characterised by an early phase of accelerated cell division followed by a phase of cell expansion. During the cell division stage plastids divide concomitantly with the cells. Because of this, plastids in dicot leaves are inherited arbitrarily to the daughter cells of a dividing cell. The process leads to cell sectors that represent clones of cells. This effect becomes obvious during the so-called sorting-out of green versus white plastids in the green and white tissues of variegation mutants (Tilney-Basset 1978). Typically, plastids in white sectors display poorly developed membrane structures and lack chlorophyll and carotenoid pigments while plastids in green sectors appear normal. The lack of pigments is often only a side effect of the genetic lesion and made it difficult to understand the true nature of the defect in plastid development. However, considerable progress has been made in recent years in understanding molecular causes for variegation in a number of mutants (Sakamoto 2003, see also below). Since the green leaf sections allow the plants to grow and to survive, variegation mutants are typically non-lethal in contrast to plants with defects leading to an albino phenotype. Therefore such mutants provide good models for investigating nuclear effects on plastid development in dicot plants (Aluru et al. 2006). A good example is the long known *immutans* mutant from *Arabidopsis* that was recently found to be defective in a plastid-localised alternative oxidase. This enzyme functions as a terminal oxidase in carotenogenesis that transfers electrons from the plastoquinol pool to oxygen. In this way, it serves as a kind of safety valve during early chloroplast development when electrons from the desaturation reactions of carotenoid biogenesis are transferred to the PQ pool. Lack of this oxidase leads to over-reduction of the PQ pool and subsequently to accumulation of phytoene making the plastid more susceptible to light-induced photodamage that then impairs chloroplast development (Aluru et al. 2001; Aluru and Rodermel 2004). However, the major problem in understanding the variegated phenotype is the fact that genetically identical cells develop white and green sectors. It has been, therefore, hypothesised that the individual plastids in *immutans* mutant cells may have different threshold values for photodamage allowing some of them to survive while others are destroyed or damaged (Aluru et al. 2006). Consistent with this model is the observation that formation of white sectors is enhanced under high-light intensities when a higher demand for excitation energy quenching by carotenoids exists. Another interesting *Arabidopsis* pigmentation mutant is *var2 (variegated 2)*. The *var2* gene encodes a plastid-localised metalloprotease of the FtsH family, which has twelve members in *Arabidopsis*. Nine of them are located in plastids. The Var2 protein forms complexes with the FtsH proteins 1, 5, and 8 that are localised in the thylakoid-membrane. It has been suggested that these complexes are involved in the PSII repair cycle that occurs during photoinhibition. As in *immutans* a varying threshold level of individual plastids has been proposed that defines the number of FtsH complexes necessary for the formation of green sectors (Chen et al. 2000; Takechi et al. 2000; Nixon et al. 2005; Aluru et al. 2006). Although the precise molecular mechanism leading to the variegation is not yet understood, these mutants demonstrate that mutations in nuclear genes for plastid proteins can disrupt the process of normal chloroplast development and that this disruption is not necessarily associated with essential primary plastid functions (compare Section 3.1.4).

3.2.2 Development of differential plastid types

Despite the observed convertibility of plastids into different types our knowledge of the molecular mechanisms underlying the development of plastid types other than chloroplasts is poor. It was observed that kinetin affects proplastids in potato tubers in such a way that they were converted into amyloplasts (Mingo Castel et al. 1991). Amyloplasts in turn were converted into chloroplasts under the influence of light (Ljubicic et al. 1998). In tobacco bright yellow 2 (BY2) cell cultures amyloplast formation from proplastids or undifferentiated leucoplasts could be initiated by depletion of auxin and addition of cytokinin. Interestingly, this differentiation required the transcription of nuclear encoded genes for components of plastid starch biosynthesis (Miyazawa et al. 1999) suggesting that hormonal signals trigger expression of nuclear genes for plastid proteins and subsequently induce the respective plastid type and function. A similar signalling cascade was found in chloroplast-chromoplast conversion during tomato fruit ripening. This process could be correlated with a dramatic increase in the expression of genes for carotenoid biogenesis enzymes again suggesting that the plastid type was defined by proteins imported from the cytoplasm after a special set of genes was activated (Giovannoni 2004). A recent study on cauliflower identified the *Orange* (*Or*) gene that confers the accumulation of high levels of β-carotene. The gene encodes a plastid-associated protein containing a DnaJ Cys-rich domain. The function of the Or protein is associated with processes that trigger the differentiation of proplastids or other colourless plastids into chromoplasts (Lu et al. 2006).

A very special case is the formation of differential chloroplasts in C4 plants since this requires a spatially different expression of photosynthesis genes in a closely related area. The mesophyll cells contain chloroplasts with normal morphology that generate ATP, reduction equivalents and oxygen, however, CO_2 is fixed into a C4 carbonic acid. The acid is then used in specialised chloroplasts of the bundle sheath cells. Such chloroplasts exhibit only poorly developed thylakoids and do not produce oxygen. However, they can decarboxylate the C4 carbonic acid and can perform the Calvin cycle. It was found that the nuclear *RbcS* gene encoding the small regulatory subunit of RubisCO, the CO_2 fixating enzyme in the Calvin cycle, is specifically expressed only in the bundle sheath cells (Ewing et al. 1998). Mutants with defective bundle sheath cells such as *bundle sheath defective 2* (*Bsd2*) were found to be impaired in a protein that is necessary for posttranslational control of RbcL accumulation, the large subunit of the RubisCO, that is plastid encoded (Brutnell et al. 1999). Thus, the targeted control of RubisCO formation by either regulating directly the nuclear gene or by indirect stabilisation of the plastid encoded subunit appears to be a crucial factor for the generation of bundle sheath chloroplasts. The question, how the nucleus of the bundle sheath cell obtains information about the cell position within the tissue, however, remains to be resolved.

Both, the conversion of plastid types and the generation of chloroplasts with differential functions appear to be dependent on the temporally or spatially controlled expression of specific nuclear genes and represent a further facet of the nuclear control principle.

3.2.3 Control of plastid division

Plastids can only originate from other pre-existing plastids. Therefore, all early developmental effects on plastid form, function, and number are inherently coupled to the ability of plastids to divide and proliferate. Microscopically, plastid division resembles the division of bacterial cells (Pyke 1999). Recent genetic and genomic approaches led to the identification of components that are essential for plastid division and that turned out to be indeed of prokaryotic origin. For instance, the *arc* (*accumulation and replication of chloroplasts*)-mutants exhibit altered or reduced plastid numbers per cell (Pyke and Leech 1994) and helped to identify several proteins necessary for plastid division at various steps of the process. Furthermore, a homologue of the bacterial *FtsZ* gene (encoding the polymerising protein of the contraction ring) was found to be targeted into chloroplasts (Osteryoung and Vierling 1995). All identified proteins are nuclear encoded and thus help to couple plastid division to the developmental stage of the host cell. It remains, however, still a mystery how the nucleus can perceive and trigger the number of the existing plastids in the cell. As known so far the size of a cell and the number of plastids inside of it appear to correlate suggesting a tight control of ‘chloroplast compartment size’ (Cookson et al. 2003; Lopez-Juez and Pyke 2005). The nuclear encoding of the division apparatus enables the cell to synchronise cell and plastid division. This is of eminent importance in early stages of plant development, for example, during leaf development (compare above). It should be noted that the *arc* mutants are impaired in plastid division but not in plastid development. This indicates that development and division of plastids are controlled by different or separate programmes.

3.3 Environmental control of plastid development

Plastid development is not only triggered by internal signals that control the expression of specific plastid proteins. It is also very much influenced by signals from the environment in which the plant is growing. One of the most important factors is light, which becomes obvious by the developmental peculiarities of leaf plastids when seedlings germinate and grow in complete darkness. Under these conditions, seedlings perform skotomorphogenesis and plastids develop into etioplasts. This plastid type can be seen as a poorly developed yellowish precursor form of chloroplasts that is arrested at a point of development when light is absolutely indispensable. The etioplast contains no thylakoid membrane structures and is characterised by the prolamellar body, a paracristallinic body generated by regular clusters of NADPH, the chlorophyll precursor protochlorophyllide and the NADPH protochlorophyllide oxidoreductase A (POR A) that accumulate to large amounts in darkness (Von Wettstein et al. 1995). Upon illumination POR A rapidly converts protochlorophyllide into chlorophyllide, the prolamellar body disassembles and thylakoid membranes develop. The POR A enzyme itself is also light labile and, under illumination, is replaced by a stable form, POR B (Armstrong et al. 1995). In consecutive steps the normal build-up of the chloroplast is performed which includes the expression of a great number of nuclear genes. They are activated by cytosolic photoreceptors such as phytochromes and cryptochromes (Link 1991). Import of such generated gene products then leads for instance to a restructuring of the PEP enzyme from the etioplastidic B to the chloroplastic A form (Pfannschmidt and Link 1994; Link 1996) and to the establishment of the photosynthetic apparatus and its related metabolic pathways. While light is essential for the redox reaction performed by the POR enzyme within the plastid, it moreover functions in the cytosol as a signal that triggers morphological programmes finally leading to the photomorphogenesis of the whole seedling. This became clear with the isolation of *Arabidopsis* mutants that exhibit photomorphogenesis in the dark. It turned out that such *det* (*de-etiolated*) and *cop* (*constitutively photomorphogenic*) mutants were defective in repressors of photomorphogenic programmes that globally mediate light responses. These are not only important for plastid proteins but also for many others (Schäfer and Bowler 2002; Wang and Deng 2003; Lorrain et al. 2006). Therefore, it is difficult to separate light-dependent plastid development from the general light-dependency in plant development. Nevertheless, recent studies on *cue* (*cab underexpressed*) mutants demonstrate that it is possible to inhibit light induction of nuclear photosynthesis genes without affecting other light responses suggesting that the light control of plastid development can be uncoupled from such other programmes (Vinti et al. 2005). Thus exogenous environmental signals and endogenous developmental signal work hand in hand in the control and regulation of plastid development, maturation, and adaptation.

4 Retrograde signalling

Plastids, however, are not only recipients of signals since they can also function as a signal generator especially when they have reached their full functional capacity. This ability has been named "retrograde signalling" to define and separate it from the above described anterograde signalling. This topic has attracted much interest in the last decade and a lot of excellent reviews have been written about it (Goldschmidt-Clermont 1998; Beck 2001; Brown et al. 2001; Jarvis 2001; Papenbrock and Grimm 2001; Rodermel 2001; Surpin et al. 2002; Gray et al. 2003; Strand 2004; Leister 2005; Nott et al. 2006). As a first definition, we can describe "retrograde signalling" as a functional feedback signal from the plastid that informs the nucleus about changes in specific processes occurring in the plastid compartment and induces appropriate changes in the expression of nuclear genes for plastid and possibly other proteins. It became quite clear in the last years that not a single but multiple signals of various natures exist that are active under different conditions. Nevertheless, the molecular identity of what really acts as the signal(s) that leave(s) the plastid is still enigmatic. In this review, we describe the role of retrograde signalling within the cell, integrate recent novel findings in the existing models, and propose possible signalling mechanisms with a special emphasis on the interactions between these different models.

4.1 Signals depending on plastid gene expression

Studies on the *albostrians* mutant of *Hordeum vulgare* led to the first proposal of a signal originating from the plastid and affecting gene expression events within the nucleus. This recessive nuclear mutant exhibits white tissue stripes in which the plastids lack ribosomes. Because all plastids in these cells are inherited from undifferentiated meristem cells a basal, still unknown mechanism must prevent ribosome formation and consequently plastid translation. Since decreased phosphoribulokinase and NADPH-glyceraldehyde-3-phosphate dehydrogenase activities were detected in these tissues it was suggested that the plastid might provide a signal that influences the nuclear expression of these enzymes (Bradbeer et al. 1979).

Further studies demonstrated that a whole set of photosynthesis-related genes was downregulated in the white tissues of the mutant (Hess et al. 1991). Treatment of various plant species with inhibitors of prokaryotic translation such as chloramphenicol, erythromycin, lincomycin, and streptomycin, which inhibit plastid but not cytosolic translation, prevented the light-dependent induction of nuclear gene expression. This indicated that the observed decrease in nuclear gene expression in *albostrians* was no side-effect of the mutation but a true effect from plastid translation. Interestingly, application of such translation inhibitors affected nuclear transcription only when performed within the first 2–3 days of seedling development. This suggested that either plastid translation generated a signalling system only in early plastid development or the plastid never reached the stage to be able to send a signal (Gray et al. 2003). Other studies revealed that light induction of nuclear *Lhcb* and *RbcS* gene expression was also prevented when plastid transcription was inhibited in early seedling development by application of tagetitoxin or rifampicin, both inhibitors of the PEP enzyme at this stage (Matthews and Durbin 1990; Pfannschmidt and Link 1997). In addition, a preventing effect on nuclear gene expression was reported for nalidixic acid, a prokaryotic DNA gyrase inhibitor, when applied in the first days of seedling development (Gray et al. 1995). Furthermore, it could be shown that lincomycin and erythromycin treatments negatively affect nuclear transcription also in the dark when applied to the mutants *lip1* (*light-independent photomorphogenesis 1*) from pea or *cop1* (*constitutively photomorphogenic 1*) from *Arabidopsis*. Both mutants exhibit photomorphogenesis in the dark (Sullivan and Gray 1999). This indicates that light itself is not an essential factor for this signalling pathway, but does not exclude that light-responsive developmental programmes are involved. The correlating time frames in which translation, transcription, and DNA replication inhibitors were effective strongly suggest that in early plastid development the gene expression within the plastid generates a signal or a signal generating process that is essential for plastid development itself as well as for the coordinated expression of nuclear encoded plastid proteins. The molecular nature of this signal or the way how it is transduced are still unknown, however, recent findings have shed more light on this topic.

The present knowledge of the protein import apparatus (see above) suggests that it works only in one direction, thus it seems unlikely that the gene expression machinery generates a protein that is transported directly out of the plastid. However, the basic subunits of the PEP enzyme are encoded in the plastome and, therefore, their expression is susceptible to all the inhibitors mentioned. As described above (Section 3.1.1), the PEP enzyme requires sigma-factor 2 for specific transcription initiation of several plastid encoded tRNAs including the *trnE* gene (Kanamaru et al. 2001; Hanaoka et al. 2005). The *trnE* gene encodes the glutamyl-tRNA that functions as the precursor molecule of 5-aminolevulinic acid (ALA), a basic component of chlorophyll biosynthesis. Therefore, any disruption of the PEP-related expression of the *trnE* gene can be expected to affect chlorophyll biosynthesis, i.e., downregulation of *trnE* expression should reduce the amount of synthesised protochlorophyllide (Pchlide a) and subsequently of chlorophyllide a (Chlide a). Chlide a is oxygenated to Chlide b by the enzyme chlorophyllide a oxygenase (CAO). CAO-deficient mutants such as *Chlorina* from *Arabidopsis* fail to accumulate Chl *b* and, interestingly, also the Chl *b* binding proteins from LHCII. A recent study suggests an important role for CAO in the regulation of Lhcb1 and Lhcb4 protein import (Reinbothe et al. 2006). Isolated chloroplasts from the *Chlorina* mutant displayed drastic reduction in the import of *in vitro* ^{35}S-labelled Lhcb1 and Lhcb4 proteins while the control protein, ^{35}S-labelled plastocyanin precursor, was imported in the same way as by wild type chloroplasts. Since CAO was found to be a component of the inner-envelope membrane generating a novel Tic sub-complex with Tic 40, Tic 22, and Tic 20 it was hypothesized that Chlide a binding to CAO and its conversion into Chlide b may prevent the Lhcb precursor from slipping back into the cytosol and supporting its import (Reinbothe et al. 2006). One can speculate that this mechanism might be also involved in the transduction of a plastid signal generated by the gene expression machinery. If the *trnE* gene is not transcribed by the PEP enzyme, the Lhcb precursors will stick within a CAO associated translocon complex due to the lack of Chlide a and b. This could then repress the expression of the Lhcb genes in the nucleus by a simple negative feedback mechanism originating from the accumulating non-imported Lhcb precursor proteins (Fig. 2). Such a model would be consistent with all observations reporting a prevention of light-induced nuclear transcription by the action of a signal from the plastid expression machinery.

However, there are also observations that do not fit into this model such as the inhibition of plastid translation in dark-grown *lip1* mutants (see above). Apparently, plastids in the very early stage of seedling development (within the first two days) produce a signal that is independent from light and hence independent from the production of Chlide (Sullivan and Gray 1999). Furthermore, a recent study on *Arabidopsis* mutants with defects in organellar ribosomal L11 proteins demonstrate that plastid and mitochondrial translation synergistically affect nuclear gene expression (Pesaresi et al. 2006). Only the double mutant but not the single mutants exhibited strong downregulation of nuclear photosynthesis genes indicating that signals from both organelles cooperate. The same was observed in mutants with downregulation of the nuclear encoded prolyl-tRNA synthetase 1 (PRORS1), an enzyme that exhibit dual targeting to plastids and mitochondria. PRORS1 is essential for translation in the organelles and null alleles of this gene are embryo-lethal. Leaky mutants, however, survived but exhibited strong downregulation of nuclear photosynthesis gene transcription. It appeared that this downregulation is independent from light and oxidative stress.

The present experimental data described above suggest that two different retrograde signals originate from plastid gene expression. One signal is light-dependent and might be mediated by accumulating Lhcb precursors as hypothesised above, the other one is light-independent and is mediated by a presently unknown mechanism (Fig. 2).

4.2 Retrograde signals depending on pigment synthesis

Tetrapyrroles and carotenoids are the two major groups of pigments involved in light-sensing, light-harvesting and energy quenching. To assure optimal photosynthesis the expression of pigment binding proteins like the light-harvesting complexes (LHC) is coupled to the biosynthesis of chlorophyll and carotenoids. The plastidic localization of all enzymes involved in pigment biosynthesis and light-harvesting and the nuclear localization of the corresponding genes requires bidirectional communication pathways between chloroplast and nucleus (Rüdiger and Grimm 2006).

The observation that greening is accompanied by the expression of chlorophyll binding proteins led to the assumption that chlorophyll precursors could play a role in signalling events. The first indication for tetrapyrroles acting as a plastid signal came from Johanningmeier and Howell (Johanningmeier and Howell 1984). They discovered a lowered level of the nuclear *Lhc* transcript upon feeding C*hlamydomonas reinhardtii* with the iron chelator dipyridyl. It blocks a late step in chlorophyll biosynthesis within the plastid and leads to accumulation of Mg-protoporphyrin IX (Mg-proto IX). The same effect could be shown later in higher plants (Kittsteiner et al. 1991). Furthermore, nuclear genes *Hsp70a/b/c* coding for heat-shock proteins were induced by direct feeding of *Chlamydomonas reinhardtii* cells with Mg-proto IX (Kropat et al. 1997). Therefore, this molecule was proposed to be a signal regulating nuclear gene expression.

The involvement of chlorophylls or chlorophyll precursors in retrograde signalling was further supported by investigations on plants with defects in carotenoid synthesis. Mutants with carotenoid deficiency accumulated less *Lhc*-mRNA whereas other nuclear gene products acting in the cytosol were not affected (Mayfield and Taylor 1984). The same result could be obtained using the herbicide norflurazon (NF). NF inhibits the plastid enzyme phytoene desaturase (PDS) that catalyses an early step in carotenoid biosynthesis (Chamovitz et al. 1991). Lack of carotenoids not only results in a reduced capacity for light-harvesting but also increases photo-oxidative damage of chlorophylls, which, in turn, leads to destruction of the thylakoid membrane. Further investigations indicated that intact plastids are essential for a proper expression of the nuclear genes *RbcS* and *Lhcb* (Oelmüller and Mohr 1986). Since the NF-promoted photo-oxidative effects were found to be restricted to chloroplasts (Mayfield et al. 1986; Puente et al. 1996), it was claimed that NF-treated plastids generate a “plastid signal” (Taylor 1989; Oelmüller 1989).

4.2.1 Signals derived from tetrapyrrole biosynthesis

The regulation pattern of tetrapyrrole synthesis is rather complex and involves regulatory steps within the plastid as well as tight communication with mitochondria (Papenbrock and Grimm 2001; Grimm 2003). Due to this complexity, this review concentrates only on steps of interest for retrograde signalling.

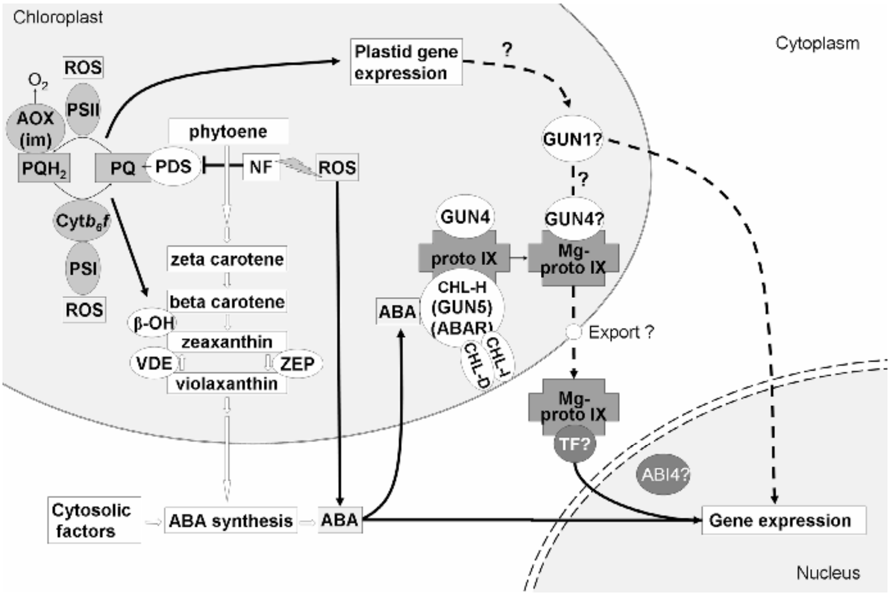
One potential plastid signal generated early in the tetrapyrrole synthesis pathway was implied by the *Arabidopsis laf6* (*long after far-red*) mutant, which was found to be defective in an envelope-located ABC-transporter (atABC1). The mutant accumulates protoporphyrin IX (proto IX) and the resulting phenotype displays an impairment of phytochrome A-mediated responses, i.e., phytochrome-regulated nuclear genes for chalcone synthase (*Chs*), ferredoxin:NADP:oxidoreductase (*Fnr*), and *Lhcb* are affected. An overexpresser of atABC1 in turn accumulated less amounts of proto IX (Moller et al. 2001). Therefore, it was suggested that either proto IX is transported into the cytosol acting as a plastid signal or that atABC1, which possesses ATPase activity, may provide energy for transport of proto IX and other intermediates of tetrapyrrole biosynthesis (Moller et al. 2001). In another model, it was then hypothesised that the observed effects could be explained by an involvement of atABC1 in the regulation of ferrochelatase activity in plastid haem synthesis (Cornah et al. 2003). Recently, atABC1 was demonstrated to be related to the *E. coli* SufB protein and was renamed into AtNAP1. It appears to represent an atypical plastidic SufB-like protein that plays an important role in Fe-S cluster assembly and regulation of iron homeostasis (Xu et al. 2005). How this function relates to plastid signalling has to be determined in the future.

For identification of nuclear-encoded components involved in retrograde signalling Susek and colleagues chose an ingenious genetic approach (Susek et al. 1993). They fused the *Lhcb1.2 (CAB3)* promoter that was known to be responsive to plastid signals (Mayfield and Taylor 1984; Oelmüller et al. 1986) to two reporter genes conferring hygromycin resistance and β-glucuronidase activity and transformed *Arabidopsis* wild type with this construct. The resulting transgenic line was mutagenised with ethyl methane sulfonate (EMS). The EMS mutant collection was then grown in presence of NF and screened for individuals displaying *Lhcb* gene expression despite the NF-mediated repressive plastid signal. By this means six mutants were isolated that were termed *gun* (*genomes uncoupled*) mutants (Susek et al. 1993).

The appearance of the *gun* mutants is very heterogeneous ranging from pale yellowish to phenotypes not distinguishable from wild type. Most of the *gun* mutants (*gun2-gun5*) were mapped to tetrapyrrole biosynthesis genes (Surpin et al. 2002). This underlines the importance of concerted action of pigment synthesis and LHC protein expression. The major finding related to plastid signalling was that all disturbances of the chlorophyll synthesis pathway lead to accumulation of Mg-proto IX and provoke a downregulation of *Lhc* genes in wild type (Strand et al. 2003). *gun2* and *gun3* are allelic with *hy1* and *hy2* (Mochizuki et al. 2001) encoding haem oxygenase and phytochromobiline synthase, respectively. Phytochromobiline is the photoreactive molecule in phytochromes, thus, its lack causes the *hy* phenotype. Both mutants accumulate haem that in turn inhibits *trnE*-reductase, the first committed step of chlorophyll biosynthesis preventing the accumulation of Mg-proto IX. Since Mg-proto IX is thought to be involved in retrograde signalling a very important role can be assigned to *gun4* and *gun5*. Both enzymes participate in a key regulatory step of chlorophyll synthesis - the introduction of magnesium into proto IX. *gun5* was identified as CHL-H (a subunit of the Mg-chelatase) and *gun4* was found to be an activator of the Mg-chelatase. An interaction of GUN4 and CHL-H was shown by co-immunoprecipitation and *in vitro* experiments (Larkin et al. 2003). The action of both enzymes seems to play a crucial role in tetrapyrrole mediated retrograde signalling (Mochizuki et al. 2001; Larkin et al. 2003; Strand 2004). The property of GUN4 to bind both the substrate (proto IX) and the product of the chelatase reaction (Mg proto IX) might enable this protein to act as a global controller of tetrapyrrole fluxes through the pathway. It is a soluble 22 kDa protein that can be found in envelope, stroma, and thylakoid fractions. The envelope-bound form of GUN4 might be an additional route for a chlorophyll-related retrograde signal (Larkin et al. 2003; Davison et al. 2005; Verdecia et al. 2005). GUN4 can mask Mg-proto IX in order to avoid accumulation of phototoxic chlorophyll precursors since it binds Mg-proto IX with higher affinity than proto IX. The resulting Mg-proto IX-GUN4 complex is no longer able to activate the Mg-chelatase resulting in a lower Mg-proto IX level. Together with the finding that a complete AtGUN4-knockout results in a pale yellowish light-sensitive phenotype this underlines the important role of GUN4. Thus, indications accumulate that the Mg-proto IX concentration

Fig. 3. (overleaf) Model for pigment-dependent retrograde signalling. The cellular compartments nucleus, cytosol and chloroplast as well as components of the photosynthetic electron transport chain within the chloroplast are depicted schematically. Thick black arrows represent signalling pathways. Broken lines indicate putative or unclear branches. Thin black arrows indicate electron transfers, white arrows indicate synthesis pathways, affected processes are given in white boxes, identified or potential candidates for intermediate protein signalling components are shown as grey ovals with white letters. Tetrapyrroles are indicated by dodecagonal symbols. Three major origins of pigment dependent signals are shown. Norflurazon (NF) inhibits the phytoene desaturase (PDS) and subsequent carotenoid synthesis. This lowers the capability of plastids to scavenge reactive oxygen species (ROS) concomitantly produced by photosynthesis especially under high-light stress (indicated by a flash) and to synthesise abscisic acid (ABA) precursors. PDS needs the alternative oxidase (AOX/*immutans*) for final electron transfer toward oxygen to replenish PQ as electron sink for the desaturation reaction. The redox states of PET and PQ are able to regulate the expression of beta-hydroxlyase (ß-OH), zeaxanthin epoxidase (ZEP) and violaxanthin deepoxidase (VDE), important enzymes of the carotenoid synthesis pathway. In this way, PET affects carotenoid and ABA synthesis. ABA levels, therefore, may act as a plastid signal. PQ redox state also affects plastid gene expression, therefore, connecting pigment- and photosynthesis-related signals. Two further signals originate from chlorophyll biosynthesis. Mg-proto IX accumulates during NF treatment as well as under high-light stress. The envelope-located Mg-chelatase complex controls the tetrapyrrole pathway and consists of the subunits CHL-H (GUN5, the chelating enzyme), its activator GUN4 (a potential sensor for Mg-proto IX levels) and the subunits CHL-D and CHL-I. From here a signal depending on chlorophyll synthesis leaves the plastid. Whether Mg-proto IX itself acts as a signal molecule or a GUN4-related signalling pathway is involved is currently unclear. In addition, CHL-H was reported to be an ABA receptor (ABAR) that provides a link between the NF-induced carotenoid and ABA deficiency and the "*gun*"-signalling pathway. It is possible that the two pathways interact. Mutants lacking the transcription factor ABI4 exhibit *gun*-like expression profiles. Promoters of nuclear photosynthesis genes responding to plastid signals, for example, *Lhcb* and *RbcS* carry ABA responsive elements pointing to ABA as the responsible plastid signal. It might be possible that ABI4 mediates the *gun*-related signal to gene expression level. GUN1 is different from other GUN components since it reacts both on inhibition of plastid gene expression and on inhibition of carotenoid synthesis by NF. Therefore GUN1 may unify the two signalling pathways. For further functional details see text.

alone cannot account for the signal. One hypothesis suggests that GUN4 senses Mg-proto IX levels and transmits the signal to the cytosol *via* an unknown mechanism (Larkin et al. 2003). The other model favours Mg-proto IX to be transported out of plastids into the cytosol where it binds to a signalling protein or directly to a transcription factor (Strand 2004). Direct feeding of Mg-proto IX to *A. thaliana* protoplasts resulted in lowered *Lhcb* expression (Strand et al. 2003). However, clear evidence for Mg-proto IX transport over the plastid envelope is still missing. This phototoxic compound might represent an emergency state of the plastid that causes a downregulation of respective genes for photosynthesis and upregulation of light stress related genes in the nucleus. Indications for this hypothesis resulted from array data exhibiting respective expression profiles (Strand et al. 2003).



The *gun5* mutant carries a single amino acid mutation in the CHL-H subunit of Mg-chelatase. This enzyme unifies three functions: Mg chelation, sensing and signalling the status of chlorophyll biosynthesis and perception of abscisic acid (ABA; see below)). CHL-H recruits proto IX possibly activated by GUN4 and transfers it to the light activated [CHL-I/CHL-D] complex where the chelatase reaction takes place (Willows and Hansson 2003). *gun5* mutants are able to produce reasonable amounts of Mg-proto IX and chlorophyll but they reveal a clear *gun* phenotype, which implies that Mg-proto IX concentration alone cannot trigger the signal. This fact is substantiated by the finding that mutants defective in CHL-I subunit accumulate even lower levels of Mg-proto IX than *gun5* but do not exhibit a *gun* phenotype (Mochizuki et al. 2001). This indicates that GUN4 and CHL-H represent the two most reasonable candidates involved in the Mg-proto IX-related retrograde signalling pathway.

Two other *gun* mutants were isolated in the screen described above. *gun6* is not identified yet but first investigations suggest that it is presumably not involved in retrograde signalling (Susek et al. 1993; Mochizuki et al. 1996). In contrast, *gun1* seems to play a very interesting role in this process. *gun1* is also not mapped yet, however, it shows a unique phenotype that is different from all other *gun* mutants. Beside the *gun* phenotype upon NF treatment *gun1* plants respond to treatment with inhibitors of plastid translation such as chloramphenicol (Susek et al. 1993). This suggests that *gun1* could act as a putative signal transducer that possibly integrates signals from chlorophyll biosynthesis and plastid gene expression (see also Section 4.1). Array data indicate that *gun1* mutants exhibit expression profiles different from *gun2* and *gun5* upon application of NF (Strand et al. 2003). This supports the view that the *gun1* gene product may act on a different signalling route then the other *gun* gene products (Fig. 3).

Further complication in understanding the chlorophyll-related retrograde signal came from a recent report that points to a connection between *gun* and ABA signalling (Shen et al. 2006). The *Arabidopsis cch* (*constitutive chlorina*) mutant, which is a stronger allele of *gun5*, was found to be defective in ABA-related responses. The *cch* mutation causes not only a *gun* phenotype and an expectable lowered chlorophyll level, but also lacks ABA-related stomatal closure (Shen et al. 2006). Furthermore, Shen and colleagues revealed that CHL-H can bind ABA and thus represents a potential ABA receptor (ABAR) (Fig. 3). ABA exogenously applied to wild type plants led to increased Mg-proto IX but lowered chlorophyll levels. This suggests a crosstalk between ABA and *gun* signalling pathways at the level of CHL-H (Fig. 3). Interestingly, only the *cch* mutant but not the *gun5* mutant exhibited typical ABA-insensitive responses such as varying stomatal aperture, root length, and germination. This implies that the *gun*- and ABA signalling pathway in part may interact, but that they mediate also very distinct reactions. How CHL-H/ABAR functions in non-green tissues remains an interesting field of study.

Since nuclear encoded photosynthesis genes are targets of plastid derived signals (Strand et al. 2003; Baier and Dietz 2005; Fey et al. 2005a, b) it is urgently needed to study them on promoter level. One of the best studied promoters is the *Lhcb1.2* (CAB3) promoter that was used for the *gun* mutant screening. It was shown by (Strand et al. 2003) that a mutated CUF1 (*cab upstream factor1*) element (Terzaghi and Cashmore 1995) cloned into the *gun5* (CHL-H) mutant and wild type displayed no response to the plastid signal suggesting that the CUF1-element is needed for the transcriptional regulation of the *Lhcb* promoter. Furthermore, the 5'-region of CUF1 within the *Lhcb1* promoter carries a putative S (*sugar responsive*)-box that was previously shown in the *RbcS*-promoter to be responsive to ABA and sugar (Acevedo-Hernandez et al. 2005). The 5'-truncated *Lhcb1* promoter in the *gun5* background shows only a fractional amount of *Lhcb* transcription (Strand et al. 2003). Since *gun5* is the mutated allele of CHL-H of the Mg-chelatase and CHL-H is thought to be an ABA receptor this points to a combined ABA and Mg-proto IX regulation at the *Lhcb* promoter site. Interestingly, the *abi4* (*ABA insensitive 4*) mutant of *Arabidopsis* exhibits a weak *gun* phenotype (Nott et al. 2006). ABI4 is a transcription factor that binds to the S-Box within the *RbcS* promoter (Acevedo-Hernandez et al. 2005). It, therefore, may mediate the retrograde regulation of nuclear encoded genes (Fig. 3). These observations suggest that the *gun*-phenotype is possibly caused by the combined action of two signals, a tetrapyrrole-related one and an ABA-related one.

4.2.2 Signals from carotenoid and ABA biosynthesis

In order to avoid phototoxic effects and to guarantee optimal light-harvesting, carotenoid synthesis is closely coupled to chlorophyll synthesis and expression of *Lhc* genes (Herrin et al. 1992; Anderson et al. 1995). Carotenoid synthesis requires a multi-step pathway beginning from isoprenoid precursors and ending up in carotenoids like lutein, neoxanthin, xanthophylls, and finally ABA (Bartley et al. 1994). The first interesting step in terms of regulation is the condensation of geranylgeranyl diphosphate to phytoene catalysed by the phytoene synthase (PSY). The expression of this enzyme is induced by phytochrome, but in mature chloroplasts chlorophyll seems to regulate PSY activity (Fraser et al. 2000). The next step in carotenoid synthesis is the desaturation reaction from phytoene to zeta-carotene that is performed by PDS (see above). This enzyme requires plastoquinone as electron sink involving an alternative terminal oxidase (AOX) that finally transfers the electrons to oxygen. Interestingly, AOX was found to be mutated in the *Arabidopsis* variegation mutant *immutans* (compare 3.2.1) resulting in accumulation of phytoene due to the over-reduction of the PQ pool. From the present data the phytoene desaturation by PDS seems to be a key step in the co-regulation of carotenoid biosynthesis and nuclear gene expression. If zeta-carotene is not synthesised any more due to either inhibition by norflurazon or by a genetic mutation, the chloroplast will become highly susceptible to photo-oxidation. Furthermore, ABA synthesis is blocked which represses the ABA-regulated defence responses that normally occur when ROS accumulate.

The next major regulatory step takes place at the level of xanthophyll cycle pigments namely zeaxanthin, antheraxanthin, and violaxanthin (Woitsch and Romer 2003). It was found that beta-hydroxylase (ß-OH) (converting beta-carotene into zeaxanthin) and zeaxanthin epoxidase (ZEP) (catalysing the reaction from zeaxanthin over antheraxanthin to violaxanthin) are co-expressed in the same light-dependent manner like *Lhcb* on mRNA level (Woitsch and Romer 2003). Using the electron transport inhibitors DCMU and DBMIB (compare 4.3.1) they revealed further that at least ZEP and ß-OH expression depends on the redox state of the PQ pool. The expression of violaxanthin de-epoxidase (VDE) was downregulated after treatments with both inhibitors suggesting a dependency on a redox-active component following the PQ pool. These findings imply that the redox state of PET may partially regulate carotenoid synthesis and, as a consequence, also ABA biosynthesis (Fig. 3).

ABA synthesis might be also influenced by other plastidic events, for example, lumenal ascorbate availability which was shown in the *Arabidopsis vtc1* mutant (Pastori et al. 2003; Baier and Dietz 2005). These findings suggest a close connection between PET, plastidic redox state, carotenoid, and chlorophyll biosynthesis. As a consequence the ABA level is affected. This might be a signal that transduces information about the light status of the chloroplast toward the nucleus. Since the last steps of ABA synthesis are cytosolic (Seo and Koshiba 2002) it seems possible that cytosolic events additionally contribute to the plastid signals that affect nuclear gene expression.

4.3 Redox signals from chloroplasts

Photosynthetic light energy fixation responds sensitively in its efficiency to environmental changes like alterations in light intensity, light quality, or temperature (Anderson et al. 1995; Niyogi 2000; Haldrup et al. 2001; Allen 2003; Pfannschmidt 2005). To maintain high photosynthetic efficiency under such a variety of environmental conditions, several compensation mechanisms have been evolved

(Walters 2005). They include, for example, regulation of enzyme activities (Scheibe 1991; Buchanan et al. 1994), adaptation of plastid gene expression (Link 2003; Pfannschmidt and Liere 2005), and changes in nuclear gene expression (Baier and Dietz 2005; Fey et al. 2005a). As regulating parameter in these mechanisms, the reduction-oxidation (redox) states of various components of the photosynthetic machinery have been identified. Thus, redox-dependent signals participate in the intracellular communication between chloroplasts and nucleus.

'Redox signals' can originate from or can be sensed by components or pools of compounds that exist either in a reduced or in an oxidized form. Gain of electrons (or hydrogen atoms) results in a reduced state, whereas loss of electrons (or hydrogen atoms) generates the oxidized state. Therefore, redox reactions are characterized by the transfer of electrons or hydrogen atoms between molecules. Photosynthesis is one of the most prominent redox processes. Its light-driven chemistry consists of a series of redox reactions involving structural components of the photosynthetic apparatus and functionally coupled pools of redox-active compounds. Redox signals from the photosynthetic electron transport (PET) chain have been shown to report the functional state of PET to chloroplast and nuclear gene expression machineries (Pfannschmidt 2003). In this way, photosynthesis exerts control over the expression of its own genes across two different cellular compartments. In addition, there is evidence for an extended influence of redox signals from PET on the expression of nuclear genes involved in processes other than photosynthesis (see below).

Reactive oxygen species (ROS) like singlet oxygen ($^1O_2$), superoxide anion (O_2) and H_2O_2, are continuously produced as photosynthetic by-products and are detoxified by scavenging mechanisms to prevent oxidative stress. Under high-light and other stress conditions, however, production of reactive oxygen intermediates exceeds detoxification capacity and ROS can accumulate. Recent studies revealed that singlet oxygen and hydrogen peroxide activate two different or distinct pathways of chloroplast-to-nucleus signalling (Kovtun et al. 2000; Meskauskiene et al. 2001; op den Camp et al. 2003; Apel and Hirt 2004). In addition, signals from the ROS scavenging processes *via* glutathione biosynthesis may provide a further signalling route to control nuclear defence genes (Mullineaux and Rausch 2005).

4.3.1 Signals from photosynthetic electron transport

The influence of PET on nuclear gene expression could be demonstrated in several algae and higher plants (Rodermel 2001; Pfannschmidt et al. 2003; Baier and Dietz 2005; Beck 2005; Fey et al. 2005a; Leister 2005). First clear evidence for such an influence came from experiments using the unicellular green algae *Dunaliella tertiolecta* and *Dunaliella salina* (Escoubas et al. 1995; Maxwell et al. 1995). Transcription of *Lhcb* genes was increased in high-light acclimated cells of *Dunaliella teriolecta* when shifted to low light conditions. The application of site-specific electron transport inhibitors DCMU and DBMIB (Trebst 1980) demonstrated that this increased *Lhcb* transcription is indeed coupled to PET. Data furthermore point to the redox state of the plastoquinone pool (PQ) as controlling parameter. *Lhcb* transcription was stimulated in low-light and DCMU treatment (blocking PET at the Q_B site leading to oxidized PQ) whereas reduced transcription was observed in high-light or DBMIB treatment (blocking PET at Q_O leading to reduced PQ) (Escoubas et al. 1995). In a similar approach, *Lhcb* transcription and LHCII protein content was analyzed in response to changes in light intensity and growth temperature in the related algae *Dunaliella salina* (Maxwell et al. 1995). This study takes into account that a given light intensity represents high-light condition in a low temperature environment but low-light condition under high temperature. Shifting cultures under constant light from low to higher temperature relaxed PSII excitation pressure in a similar way as a shift from high-light to low-light condition under constant ambient temperature. An increase of *Lhcb* transcripts was observed upon relaxation of PSII excitation pressure. Again, the data suggest regulation of *Lhcb* transcription by the redox state of the PQ pool.

Studying acclimation to altered light intensities in *Lemna perpusilla* demonstrated that the redox state of the PQ pool can also regulate *Lhcb* transcription and LHCII protein content in a higher plant (Yang et al. 2001). A cytb_6f deficient mutant was found to be locked in the state of high-light acclimation and did not show induction of LHCII under low-light conditions as was observed in wild type plants. Application of DCMU could abolish this effect, thus pointing to the PQ pool as the light intensity sensor. Further experiments with *Arabidopsis* cell cultures and transgenic *Arabidopsis* plants harbouring a reporter gene fused to the *Lhcb2* promoter investigated the effect of plastid redox state by application of DCMU in comparison to the effect of sugar on *Lhcb* gene expression (Oswald et al. 2001). Interestingly, it was found that *Lhcb* expression responded to both sugar content and redox signals. Upon sugar depletion, an increase in *Lhcb* transcript level was observed which could be blocked by the application of DCMU thus implying a connection between PET and sugar signalling. In a study with winter rye, *Lhcb* gene expression was found to respond to changes in PET that were generated by varying light and temperature conditions (Pursiheimo et al. 2001). The authors concluded that the regulation of nuclear gene expression was mediated by the redox state of electron acceptors of PSI, which contrasts the results described above. Investigation of four-week-old tobacco plants carrying a chimeric *PetE* gene construct from pea revealed that DCMU treatment decreased accumulation of pea *PetE* and endogenous *Lhcb1* transcripts. However, transcriptional rates as determined by nuclear run-on assays revealed upregulation of the pea *PetE* construct while tobacco *Lhcb1* genes were decreased. This suggests multiple effects of PET acting at different levels of gene expression, i.e., at transcriptional and posttranscriptional levels (Sullivan and Gray 2002). PET regulation of posttranscriptional mechanisms rather than transcription itself was uncovered also in other studies. Light-induced accumulation of pea-ferredoxin-1 (Fed-1) transcript in transgenic tobacco was observed even when the transgene was fused to a constitutive promoter. Responsiveness to DCMU suggested an influence of the PET on transcript amounts (Petracek et al. 1997, 1998). Upregulation of the nuclear *Apx2* gene in response to chloroplast signals was observed in *Arabidopsis* after high-light treatment (Karpinski et al. 1997, 1999). *Apx2* codes for a cytosolic ascorbate peroxidase that catalyses the reduction of hydrogen peroxide to water and that can be induced by H_2O_2. Ascorbate peroxidases play a crucial role under stress conditions and will be discussed further below. However, experimental data also suggested an involvement of the PQ pool as an early signal in the regulation of *Apx2* expression (Karpinski et al. 1997, 1999). This conclusion is in contrast to recent findings on *Arabidopsis* (Fryer et al. 2003), but gained further support from a different study on transgenic tobacco (Yabuta et al. 2004). Here, further experiments are needed to understand the molecular details of this regulation.

In a different experimental approach, light quality rather then light intensity was changed. This allows for experiments under low-light conditions avoiding stress-mediated side effects. Emission spectra of artificial light sources can be adjusted in such a way that they preferentially excite either photosystem I or II (Melis 1991; Walters and Horton 1994; Allen and Pfannschmidt 2000). PSI- and PSII-favouring light conditions lead to rather oxidized or reduced redox systems in the electron transport chain, respectively. Such light quality induced imbalances in excitation energy are counterbalanced on a timescale of days by photosystem stoichiometry adjustment as has been shown for several cyanobacteria, algae, and higher plants (Melis 1991; Allen 1995; Pfannschmidt 2003). The described experimental system was used to study the influence of signals from PET on nuclear gene expression. Transgenic tobacco lines containing nuclear PSI gene promoters fused to a reporter gene showed that *PsaD, PsaF*, and *PetE* promoters are activated by the reduction of PET components (Pfannschmidt et al. 2001). *PsaD, PsaF*, and *PetE* code for the PSI subunits D and F, and for plastocyanin (PC), all of which associated with electron transport around PSI. Application of electron transport inhibitors DCMU and DBMIB in this system further revealed that the redox state of the PQ pool regulates PC promoter activity whereas *PsaD* and *PsaF* promoters respond to PET-derived signals downstream of PQ. In the same experimental setup, the promoter of the nuclear non-photosynthetic gene *nia2* (coding for nitrate reductase) responded also to signals from PET when analyzed in *Lemna, Arabidopsis*, and tobacco (Sherameti et al. 2002) demonstrating that this regulation is not restricted to photosynthesis genes. The data discussed above suggest that redox signals from the PQ pool are involved in both high-light and low-light responses. Whether these signals are transduced *via* two independent pathways (Fig. 4) is not clarified yet.

Additional indication for an influence of PET on nuclear gene expression came from the *Arabidopsis* mutant *cue-1* (*chlorophyll a/b binding protein underexpressing*) (Streatfield et al. 1999). The mutant, that lacks the phosphoenolpyruvate/phosphate translocator PP1, exhibits light intensity dependent underexpression of *Lhc* genes. Reduced flux through the shikimate pathway and measurements of rapid induction kinetics of chlorophyll a fluorescence furthermore suggest a reduced PQ pool size. However, reduced chlorophyll and carotenoid contents could also account for the observed *Lhc* underexpression.

Early experiments studying the influence of PET-derived signals focused on single or a limited number of genes. Technical advances in recent years now allow more extended analyses. The influence of light quality induced redox signals was assessed on a larger scale in *Arabidopsis* by expression analysis using a macroarray for ca. 3300 nuclear genes (Fey et al. 2005b). Most of the genes covered by

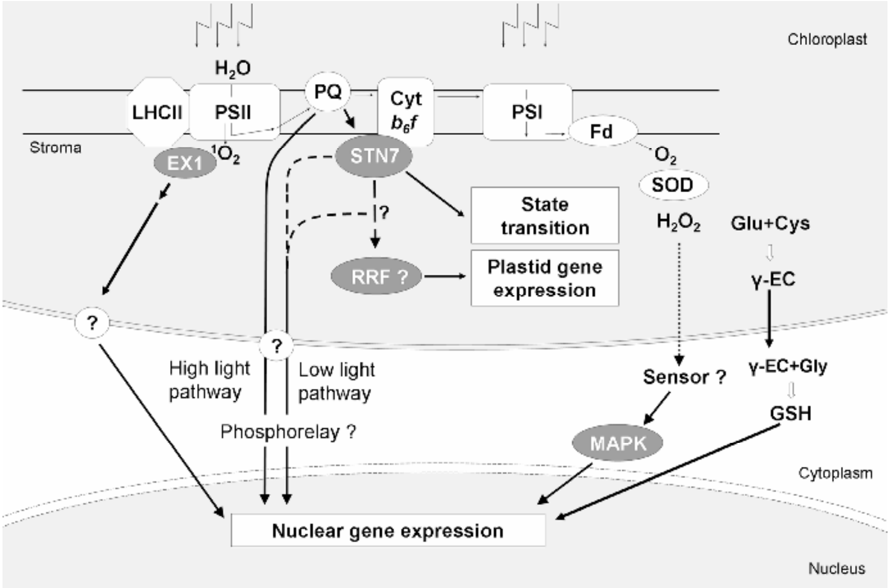


Fig. 4. Chloroplast redox signals and putative signal transduction mechanisms to the nucleus. The cellular compartments nucleus, cytosol, and chloroplast as well as components of the photosynthetic electron transport chain within the chloroplast are depicted schematically. Thick black arrows represent signalling pathways, broken lines indicate putative or still unclear branches. Thin black arrows indicate electron transfers, dotted lines represent diffusion, and white arrows indicate syntheses, affected processes are in white boxes, identified or potential candidates for intermediate protein signalling components are shown as grey ovals with white letters. Four major origins of chloroplast redox signals are shown: singlet oxygen ($^1O_2$) from PSII, signals directly originating from PET, hydrogen peroxide (H_2O_2) generated by PSI and γ-glutamylcysteine (γ-EC) representing the precursor molecule for cytosolic glutathione. Singlet oxygen-related signals are mediated by the protein *EXECUTER1* (EX1). How the signal leaves the plastid is not understood yet. The redox state of the PQ pool appears to be involved in high-light and low-light responses. The sensor for its redox state appears to be the thylakoid-associated kinase STN7 that may initiate a phosphorylation cascade (*via* a postulated redox responsive factor (RRF)) that mediates effects on plastid gene expression. A branch of this pathway leaves the plastid to coordinate the expression of nuclear photosynthesis genes. Hydrogen peroxide generated by superoxide dismutase (SOD) might act as a third signal leaving the plastid by activating cytosolic MAP kinase cascades. For further functional details see text.

this macroarray encode nuclear encoded chloroplast proteins (Kurth et al. 2002; Richly et al. 2003). Light quality changes in combination with DCMU treatments revealed that 286 genes are regulated by photosynthetic redox signals. Most of the identified genes code for putative proteins or proteins of unknown function. Genes with an assigned function cover, for example, gene expression, metabolism, or signal transduction and are not restricted to photosynthesis underlining the great influence of photosynthetic redox signals on nuclear gene expression. In a subsequent study, expression of ca. 8000 randomly selected *Arabidopsis* genes was monitored in response to different light qualities and light intensities (Piippo et al. 2006). Analyzing expression profiles, the authors concluded that under the studied conditions, nuclear gene expression responds to signals from stromal redox components on the reducing site of PSI as well as to other metabolic cues.

4.3.2 Transduction of PET signals

While studies during the last decade provide ample evidence that redox signals from PET influence expression of many nuclear genes, little is known about the signal transduction mechanisms. The redox signal has to be sensed, transduced across the chloroplast envelope into the cytosol, and finally transmitted to the nucleus in order to exert transcriptional control. So far, none of these steps within the signalling pathway has been understood at the molecular level.

To date, the PQ pool represents the best characterized origin of a redox signal from PET (see above). Early studies in *Dunaliella tertiolecta* observed that inactivation of protein phosphatase activities by inhibitors reduced acclimation responses on high-light to low-light shifts that were shown to depend on the PQ redox state (Escoubas et al. 1995). The authors, therefore, hypothesised that the redox state of the PQ pool is transduced to the nucleus *via* a phosphorylation cascade: a redox sensing kinase is thought to phosphorylate an unknown plastidic protein. After signal transfer across the chloroplast envelope, a cytosolic kinase is assumed to phosphorylate a *Lhcb* gene repressor (Escoubas et al. 1995; Durnford and Falkowski 1997). The identities and functional connections between the involved signalling partners in this model, however, remain to be clarified. A later study using a similar experimental setup and organism observed several protein complexes at the *Lhcb* promoter indicating a rather complex regulation of *Lhcb* gene expression (Chen et al. 2004). In this study, *Lhcb* was further found to respond to the trans-thylakoid pH gradient beside the redox state of the PQ pool, suggesting that more than one redox signal could be involved. In transgenic tobacco, the *PsaF* promoter was shown to respond to signals from the PET (Pfannschmidt et al. 2001). An independent study demonstrated that this promoter can also be induced by a cytosolic kinase even when functional plastids are absent (Chandok et al. 2001) supporting the idea of a cytosolic phosphorylation cascade in the transduction of redox signals (Fig. 4).

In another line of argumentation, it has been discussed that the redox state of the PQ pool regulates both short- and long-term acclimation processes at low-light conditions (Allen and Pfannschmidt 2000). Recent data indeed suggest a hierarchical and/or coupled action of the two responses (Allen and Pfannschmidt 2000; Pursiheimo et al. 2001; Bonardi et al. 2005; Pfannschmidt 2005). Thus, it can be hypothesised that the PQ signalling pathway that controls chloroplast gene expression might be also part of the pathway toward the nucleus. Under PQ-reducing conditions, activation of an LHCII kinase was proposed that phosphorylates the mobile part of the LHCII antenna in response to the redox state of the PQ pool (Allen 2003). Studies on green algae and higher plants indicate that the redox activation of the kinase requires a functional cyt b_6f complex involving plastoquinol binding to the PQ oxidising site of the complex (Vener et al. 1997; Zito et al. 1999; Hou et al. 2003). Phosphorylated LHCII then migrates from PSII to PSI, extends PSI antenna size and thus compensates excitation imbalances to a certain degree. If reducing conditions prevail, transcription of plastid genes (e.g. *psaAB*) will be activated (Pfannschmidt et al. 1999, 2003) and will trigger photosystem stoichiometry adjustment. The latter process involves both, plastid and nuclear gene expression (Fig. 4).

The existence of an LHCII kinase has long been proposed (Allen and Race 2002). Recently, a good candidate has been discovered in *Chlamydomonas reinhardtii* (Depège et al. 2003). The *stt7* mutant was deficient in state transition, and reversible phosphorylation of LHCII under altered light conditions could not be observed. The affected gene, interestingly, is a nuclear gene coding for a plastid localized serin/threonine kinase associated to the thylakoid membrane. The *Arabidopsis* orthologue *stn7* is also required for short-term acclimation and LHCII phosphorylations (Bellafiore et al. 2005; Bonardi et al. 2005). STN7 might therefore represent an ideal candidate for sensing the redox state of the PQ pool and for transducing the signal *via* phosphorylation to further signalling components (Fig. 4). Whether signalling of the two processes branches already at the point of signal perception or whether they act in series is still unknown. In our model, a downstream redox-responsive factor (RRF) is proposed that controls plastid gene expression. Studies reporting that plastid transcription can be regulated by phosphorylation are in line with this argument (Tiller and Link 1993). It is still not known how the signal from the thylakoid membrane is transduced to the plastid transcription machinery. TSP9, a small 9 kDa protein associated to PSII represents a possible candidate for signal transduction within the chloroplast. Under reducing conditions, the protein can be partially released from PSII and its C-terminus contains a basic domain that potentially functions in DNA-binding (Carlberg et al. 2003; Fey et al. 2005a). A functional relationship to long-term acclimation responses, however, still remains to be established.

It was additionally shown that lack of STN7 in *Arabidopsis* also impairs long-term light quality acclimation suggesting that STN7 coordinates both, state transitions, and long-term acclimation responses under changing light conditions (Bonardi et al. 2005). STN7 could provide a link between photosynthetic efficiency and gene expression in the chloroplast and in the nucleus. First experimental data, however, indicate a minor role of STN7 in regulating nuclear gene expression under moderate light condition as differential gene expression between *stn7* mutant and wild type was only reported for stress-responsive genes under stress conditions (Tikkanen et al. 2006). The present picture became even more complex by a recent study on *Chlamydomonas rheinhardtii* investigating seven mutants with different defects in the cyt b_6f complex. In all mutants the light induction of tetrapyrrole biosynthetic genes was either abolished or strongly reduced. This was not observed in other photosynthesis mutants, in wild type treated with DBMIB or in the state transition mutant *stt7* indicating that not the PQ redox state signals the light induction of chlorophyll biosynthesis genes. The mutant analyses point to the integrity of the Q_o site in the cyt b_6f complex as a decisive determinant in the regulation of these nuclear genes (Shao et al. 2006). Although one has to be careful with general conclusions this study demonstrates the multiplicity of redox signals originating from PET.

At present potential candidates for sensing the redox state of PET and for transducing redox signals within the chloroplast have been identified while the precise functional connections still remain to be shown. Moreover, little is known about how the signal is transferred across the chloroplast envelope. Nevertheless, several studies suggest a phosphorelay for signal transduction within the cytosol in order to regulate gene expression in the nucleus. Much more experimental data will be required to identify further partners and their functional relationships in the signalling pathway.

4.3.3 Reactive oxygen species: hydrogen peroxide (H_2O_2)

Hydrogen peroxide, which represents a further chloroplast redox signal, is the principal ROS in plants. In chloroplasts, it is mainly generated at PSI when excitation energy is available to excess. This can be the case, for example, under high light stress, or light in combination with chilling, nutrient starvation, or drought. Over-reduction of the electron transport chain leads to transfer of electrons from reduced ferredoxin to oxygen (Mehler reaction) (Mehler 1951). The resultant superoxide is detoxified by superoxide dismutase (SOD) to hydrogen peroxide and can accumulate in this form (Mullineaux and Karpinski 2002; Apel and Hirt 2004). H_2O_2 can then be reduced to water by ascorbate peroxidases (APX) using ascorbate as electron donor and requiring glutathione to restore the electron donor.

In plants, multiple genes for APX exist encoding proteins that are specifically targeted to different cell compartments. Whereas chloroplast ascorbate peroxidases are constitutively expressed, cytosolic ascorbate peroxidases can be induced under oxidative burst conditions (Shigeoka et al. 2002; Nott et al. 2006). Studies in *Arabidopsis* could show that the high-light induced expression of nuclear genes *apx1* and *apx2* correlated with the generation of H_2O_2. Furthermore, externally applied H_2O_2 could induce cytosolic ascorbate peroxidases even in darkness, in this way mimicking photo-oxidative stress (Karpinski et al. 1997; Foyer and Noctor 1999; Karpinski et al. 1999). Infiltration of high-light-exposed leaves with catalase, an enzyme that detoxifies H_2O_2, could diminish induction of *apx2* gene expression thus pointing towards H_2O_2 as a retrogade signal (Karpinski et al. 1999). As already discussed (4.3.1), *apx2* is also controlled by the redox state of the PQ pool (Karpinski et al. 1997). The same was found for *Arabidopsis elip2* (*early light-inducible protein*) that belongs to the *Lhc*-superfamily (Kimura et al. 2003). Therefore, a combined action of the two signals was suggested. In a recent study, initial induction of tobacco *apx2* was assigned to the redox state of the PQ pool while the later response was attributed to the observed levels of H_2O_2 (Yabuta et al. 2004).

The example of the *apx* genes demonstrates that H_2O_2 generated in the chloroplast results in changes of nuclear genes. Its impact on the nuclear transcriptome of *Arabidopsis* cell cultures was assessed on a broader scale by external application of hydrogen peroxide. Microarray analysis revealed H_2O_2-sensitivity for 1-2% of the analyzed genes including stress-related and defence genes (Desikan et al.

2001). In this context, further studies are available that support these findings (Vandenabeele et al. 2003; Davletova et al. 2005).

Up to date, little is known about the exact nature and site of sensing of plastid generated hydrogen peroxide. Nevertheless, one of the best models for the transduction of a redox signal exists for H_2O_2. Hydrogen peroxide is thought to diffuse freely across the chloroplast envelope and to activate a mitogen-activated protein kinase (MAPK) cascade in the cytosol. This then affects gene expression events in the nucleus (Fig. 4) (Kovtun et al. 2000; Apel and Hirt 2004). Such a signal transduction *via* phosphorylation minimally involves a MAPKKK-MAPKK-MAPK module linked to an upstream receptor and downstream targets. In *Arabidopsis* H_2O_2 specifically activates MAPKK kinases ANP1 or MEKK1 that eventually activate the MAPKs MPK3, MPK4 and MPK6. MAPKs are known to phosphorylate a variety of substrates including transcription factors thus providing a possible link to the level of nuclear gene expression.

Although H_2O_2 is produced in response to various different stresses the induced responses, for example, to pathogen attack or high-light clearly differ. This raises the question of how the signal is specified. Differences in the local distribution of H_2O_2 or its targets and dosage of H_2O_2 as well as the existence of additional signals have been suggested to confer specificity of the response (Kovtun et al. 2000; Beck 2005). In addition, the involvement of H_2O_2 in auxin responses (Kovtun et al. 2000) and ABA signalling (Desikan et al. 2001) further suggests that H_2O_2 participates in complex signalling networks.

4.3.4 Reactive oxygen species: singlet oxygen ($^1O_2$)

Singlet oxygen represents a distinct non-radical reactive oxygen species that influences nuclear gene expression and activates several stress response pathways (op den Camp et al. 2003). While continuously generated at PSII, its production rapidly increases under excess light conditions. This mainly results from over-reduction of electron transfer compounds around PSII which facilitates formation of triplet state P680 and enhanced production of singlet oxygen by energy transfer. Because of its short half-life of about ~ 200 ns, singlet oxygen induces oxidative damage at its site of generation (op den Camp et al. 2003). Under various stresses, several ROS are formed at the same time making it difficult to discern the action of chemically distinct reactive oxygen species (Apel and Hirt 2004).

Studies with the *Arabidopsis flu* (*fluorescent*) mutant overcame these difficulties as singlet oxygen production can specifically be induced upon a dark-to-light shift without parallel induction of hydrogen peroxide. The mutant was discovered in a screen for novel factors involved in tetrapyrrole biosynthesis and accumulates free protochlorophyllide (Pchlide) when put into darkness. Upon re-illumination the mutant exhibits enhanced singlet oxygen production that leads to growth inhibition and cell death while mutant plants grown under continuous light develop like wild type (Meskauskiene et al. 2001). It could be demonstrated that the induced responses are not caused by the toxicity of singlet oxygen itself but that they result from an activation of a genetically determined stress programme, which requires a protein called *EXECUTER1* (*EX1*) (Wagner et al. 2004). This could be demonstrated by a *flu/ex1* double mutant that accumulated Pchlide in the dark and generated similar amounts of singlet oxygen after illumination as observed in the *flu* single mutant. However, the singlet oxygen-mediated stress responses observed in *flu* plants did not occur in the double mutant, which grew like wild type. Thus, mutation of a single gene was sufficient to abolish the stress responses induced by singlet oxygen (Wagner et al. 2004). *EX1* therefore represents a good candidate for sensing of singlet oxygen or for the transduction of singlet oxygen-derived signals (Fig. 3). A detailed study of the singlet oxygen-induced cell death response in the *Arabidopsis flu* mutant further revealed that this response is promoted by an ethylene-, salicylic acid-, and jasmonic acid-dependent signalling pathway while it is blocked by a jasmonic acid precursor (Danon et al. 2005).

The impact of singlet oxygen on nuclear genes expression was analyzed in the *flu* mutant by whole genome transcript profiling (op den Camp et al. 2003). In mature *flu* plants, expression of about 5% of the genome changed in response to reillumination. Among these genes, 70 were identified as specifically activated by singlet oxygen. This was done by using stringent data selection criteria and by comparison to global expression profiles of the *flu* mutant obtained under H_2O_2-generating treatments with paraquat. Expression studies in seedlings of the *flu* mutant confirmed the previous findings (Danon et al. 2005). Interestingly, several of the singlet oxygen responsive genes were involved in biosynthesis or signalling of ethylene, jasmonic acid, or salicylic acid as well as further yet unspecified signalling. Although, *EX1* represents a good candidate also for signal transduction from singlet oxygen within the chloroplast to the level of nuclear gene expression, a functional link especially for the mediation of the signal over the envelope membrane remains to be established.

4.3.5 Glutathione as mediator of stress responsive gene expression

As already mentioned reduced glutathione (GSH) is involved in the scavenging of reactive oxygen species since it is needed for re-reduction of the primary scavenger ascorbate (see Section 4.3.3). Recent observations suggest that GSH may also act as a plastid signal that controls the expression of stress defence genes and a respective model has been proposed (Mullineaux and Rausch 2005). GSH is an important component of plant cell chemistry that is involved in many enzymatic and non-enzymatic reactions. Therefore, it is difficult to establish a direct link between glutathione redox state, glutathione level and the control of gene expression. However, there are several reports demonstrating that changes in expression of stress defence genes, for example, *Apx2* or *Pr1* (*pathogenesis related 1*) correlate with changes in cellular glutathione content (Mullineaux and Rausch 2005). A recent report about the *Arabidopsis* mutant *rax1-1* (*regulator of Apx2*) describes a first direct link (Ball et al. 2004). The mutant is impaired in the glutathione synthetase 1 (GSH1), an essential enzyme for GSH synthesis, resulting in lowered cellular GSH content. However, the mutant exhibits a constitutive expression of the *Apx2* gene that is normally expressed only under stress such as high-light illumination (Karpinski et al. 1997). Thus, a low cellular GSH level resulting from oxidative stress in chloroplasts may activate defence genes in the nucleus. A possible way in which such a lower content can be signalled came from investigations of the GSH synthesis pathway. GSH is a tri-peptide composed of cysteine (the final product of sulphur assimilation), glutamate, and glycine. Its synthesis depends on the assimilation of sulphur, nitrogen, and carbon and is under low-light intensities (< 100 μmol m^{-2} s^{-1} photons) also dependent on the PET rate (Noctor et al. 2002; Kopriva and Rennenberg 2004; Ogawa et al. 2004). In a first step, γ-glutamylcysteine (γ-EC) is synthesised from glutamate and cysteine by GSH1. It is further processed to GSH *via* the addition of glycine by the enzyme glutathione synthetase 2 (GSH2). A recent study on the GSH1 and GSH2 transcript structures in *Arabidopsis* indicates that the first step of GSH synthesis is confined to the plastid compartment while the second one predominantly occurs in the cytosol (Wachter et al. 2005). Thus, plastid γ-EC represents the precursor molecule for cytosolic GSH and must be transported out of the organelle. Since the synthesis is dependent on photosynthesis (see above) sudden environmental impacts on photosynthetic rates may affect γ-EC synthesis, which affects subsequently the GSH synthesis in the cytosol. Thus, the export of the precursor molecule might be a plastid signal (Mullineaux and Rausch 2005). So far, this trans-plastidial route for GSH biosynthesis has been demonstrated only in *Arabidopsis*. Whether this is true also for other plants and whether the γ-EC export represents a plastid signal of general importance in all plants has to be verified in future.

4.4 Plastid signals controlling tissue development

The above described mechanisms focus on ROS signals from chloroplasts that influence nuclear gene expression. Some of them induce appropriate defence mechanisms that are essential to avoid or protect from photo-oxidative damage. The protection from, for example, light stress is of vital interest for the cell and the plant as a whole. However, if stresses induce ROS production that exceeds the detoxification capabilities for a long time, then the defence and repair mechanisms will not be able to compensate the damages and the cell might be targeted to an induced cell death. The analysis of the variegation mutants *immutans* and *var2* led to the proposal that these detoxification capacities of the individual chloroplasts in a cell may vary (Aluru et al. 2006). Together with the segregation of the plastids during early tissue development this may lead to the variegated phenotype of the mutants (see above). Interestingly, these studies uncovered a further unexpected influence of plastids on the host cell. All retrograde signalling pathways described so far regulate processes that help to optimise or adapt pathways related to the plastid function. The observations on various variegation mutants suggest that there exist also a function for plastid signals in leaf development. Beside *immutans* and *var2* also, for example, *cue1* (*cab underexpressed 1*) from *Arabidopsis* (Streatfield et al. 1999), *DAG* (*differentiation and greening*) from *Antirrhinum* (Chatterjee et al. 1996), *DAL* (DAG-like) from *Arabidopsis* (Bellaoui and Gruissem 2004), DCL (*defective chloroplasts and leaves*) as well as *ghost* both from tomato (Keddie 1996; Josse 2000) have been analysed. In leaf cross-sections of the white tissues of all these plants a reduction or a change in palisade cell expansion can be found. It appears that the defect in chloroplast formation has also a distinct effect on the cell layer where these plastids are mainly active. Comparable effects have been reported also for the *SCABRA3* mutant indicating that the function of the plastid-localised RpoTp RNA polymerase is required for mesophyll cell proliferation (Hricova et al. 2006). This suggests a tight coupling of plastid and host cell development. How the information about the chloroplast defect is sensed and how this affects the tissue development is not clear yet, but it is reasonable to assume that such white plastids perform a reduction or a change of the signal transmitted *via* the above described signalling pathways due to their functional defect. This might block or reduce further steps in palisade tissue development. Thus, the retrograde signals represent an important source of information also for developmental programmes of whole tissues (compare 3.2.1).

5 Conclusions and perspectives

This review demonstrates that development and function of plastids in plant cells require the combined action of the nucleus and the plastids themselves. Without the temporally and spatially controlled expression of nuclear encoded plastid proteins the establishment of functionally active plastids is impossible, thus, the nucleus exerts a tight control over these organelles. However, the nucleus is blind for the specific demands of individual plastids in the cell that may vary either to environmental or developmental changes. This implies that plastids beside their metabolic functions also act as sensor for changes in various external and internal signals, which they report to the nucleus. These retrograde signals originate from different plastid processes such as gene expression, photosynthesis, pigment synthesis, and stress responses. The present data indicate that these signalling pathways interact or function in a combined manner forming a complex network. This in turn is connected to multiple anterograde signals. The studies discussed in this review represent examples that indicate how deeply plastids are embedded into the cellular regulation networks. Because of the multiplicity of mutual interaction, it is impossible to cover all interactions in one single review. Thus the role of sugars and metabolic signals could be only slightly indicated and the interested reader is referred to several excellent reviews that cover these topics extensively (Geigenberger et al. 2005; Gupta and Kaur 2005; Couee et al. 2006; Wingler et al. 2006). We are certainly just beginning to understand the complex relationships that exist between the various signalling pathways into and out of the compartments of a plant cell. The rapid technical progress, however, will provide us with tools to understand cellular responses to external and internal signal at a network level. Thus, the next few years promise considerable progress in our understanding of the connection of cellular processes and the underlying regulatory networks during plant development and environmental acclimation.

Acknowledgements

Our work was supported by grants from the DFG, the "NWP" and "Excellence in Science" programmes of Thuringia to TP and to the DFG research group FOR 387.

References

Abdallah F, Salamini F, Leister D (2000) A prediction of the size and evolutionary origin of the proteome of chloroplasts of *Arabidopsis*. Trends Plant Sci 5:141-142

Acevedo-Hernandez G, Leon P, Herrera-Estrella L (2005) Sugar and ABA responsiveness of a minimal RBCS light-responsive unit is mediated by direct binding of ABI4. Plant J 43:506-519

Adam Z, Rudella A, van Wijk KJ (2006) Recent advances in the study of Clp, FtsH and other proteases located in chloroplasts. Curr Opin Plant Biol 9:234-240

Allen JF (1995) Thylakoid protein phosphorylation, state1-state 2 transitions, and photosystem stoichiometry adjustment: redox control at multiple levels of gene expression. Physiol Plant 93:196-205

Allen JF (2003) State transitions--a question of balance. Science 299:1530-1532

Allen JF, Pfannschmidt T (2000) Balancing the two photosystems: photosynthetic electron transfer governs transcription of reaction centre genes in chloroplasts. Philos Trans R Soc Lond B Biol Sci 355:1351-1359

Allen JF, Race HL (2002) Will the real LHC II kinase please step forward? Sci STKE 2002:PE43

Allison LA (2000) The role of sigma factors in plastid transcription. Biochimie 82:537-548

Aluru MR, Bae H, Wu D, Rodermel SR (2001) The *Arabidopsis* immutans mutation affects plastid differentiation and the morphogenesis of white and green sectors in variegated plants. Plant Physiol 127:67-77

Aluru MR, Rodermel SR (2004) Control of chloroplast redox by the IMMUTANS terminal oxidase. Physiol Plant 120:4-11

Aluru MR, Yu F, Fu A, Rodermel S (2006) *Arabidopsis* variegation mutants: new insights into chloroplast biogenesis. J Exp Bot 57:1871-1881

Anderson JM, Chow WS, Park YI (1995) The grand design of photosynthesis: Acclimation of the photosynthetic apparatus to environmental cues. Photosynth Res 46:129-139

Aro EM, Andersson B (2001) Regulation of photosynthesis. Dordrecht: Kluwer Academic Publishers

Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu Rev Plant Biol 55:373-399

Armstrong GA, Runge S, Frick G, Sperling U, Apel K (1995) Identification of NADPH:protochlorophyllide oxidoreductases A and B: A branched pathway for light-dependent chlorophyll biosynthesis in *Arabidopsis thaliana*. Plant Physiol 108:1505-1517

Baier M, Dietz KJ (2005) Chloroplasts as source and target of cellular redox regulation: a discussion on chloroplast redox signals in the context of plant physiology. J Exp Bot 56:1449-1462

Ball L, Accotto GP, Bechtold U, Creissen G, Funck D, Jimenez A, Kular B, Leyland N, Mejia-Carranza J, Reynolds H, Karpinski S, Mullineaux PM (2004) Evidence for a direct link between glutathione biosynthesis and stress defense gene expression in *Arabidopsis*. Plant Cell 16:2448-2462

Barkan A, Goldschmidt-Clermont M (2000) Participation of nuclear genes in chloroplast gene expression. Biochimie 82:559-572

Bartley GE, Scolnik PA, Giuliano G (1994) Molecular biology of carotenoid biosynthesis in plants. Annu Rev Plant Physiol Plant Mol Biol 45:287-301

Baumgartner BJ, Rapp JC, Mullet JE (1989) Plastid transcription and DNA copy number increase early in barley chloroplast development. Plant Physiol 89:1011-1018

Beck CF (2001) Signaling pathways in chloroplast-to-nucleus communication. Protist 152:175-182

Beck CF (2005) Signaling pathways from the chloroplast to the nucleus. Planta 222:743-756

Bellafiore S, Bameche F, Peltier G, Rochaix JD (2005) State transitions and light adaptation require chloroplast thylakoid protein kinase STN7. Nature 433:892-895

Bellaoui M, Gruissem W (2004) Altered expression of the *Arabidopsis* ortholog of *DCL* affects normal plant development. Planta 219:819-825

Bendich AJ (2004) Circular chloroplast chromosomes: the grand illusion. Plant Cell 16:1661-1666

Blankenship RE (2002) Molecular mechanisms of photosynthesis. Oxford: Blackwell Science Ltd

Bock R, Koop HU (1997) Extraplastidic site-specific factors mediate RNA editing in chloroplasts. EMBO J 16: 3282-3288

Bollenbach TJ, Schuster G, Stern DB (2004) Cooperation of endo- and exoribonucleases in chloroplast mRNA turnover. Prog Nucleic Acid Res Mol Biol 78:305-337

Bonardi V, Pesaresi P, Becker T, Schleiff E, Wagner R, Pfannschmidt T, Jahns P, Leister D (2005) Photosystem II core phosphorylation and photosynthetic acclimation require two different protein kinases. Nature 437:1179-1182

Bradbeer JW, Atkinson YE, Börner T, Hagemann R (1979) Cytoplasmic synthesis of plastid polypeptides may be controlled by plastid synthesized RNA. Nature 279:816-817

Brown EC, Somanchi A, Mayfield SP (2001) Interorganellar crosstalk: new perspectives on signaling from the chloroplast to the nucleus. Genome Biol 2:REVIEWS1021

Brutnell TP, Sawers RJH, Mant A, Langdale JA (1999) Bundle sheath defective2, a novel protein required for post-translational regulation of the rbcL gene of maize. Plant Cell 11:849-864

Buchanan BB, Gruissem W, Jones RL (2002) Biochemistry and molecular biology of plants. Somerset: John Wiley and Sons Inc.

Buchanan BB, Schurmann P, Jacquot JP (1994) Thioredoxin and metabolic regulation. Sem Cell Biol 5:285-293

Carlberg I, Hansson M, Kieselbach T, Schroder WP, Andersson B, Vener AV (2003) A novel plant protein undergoing light-induced phosphorylation and release from the photosynthetic thylakoid membranes. Proc Natl Acad Sci USA 100:757-762

Chamovitz D, Pecker I, Hirschberg J (1991) The molecular basis of resistance to the herbicide norflurazon. Plant Mol Biol 16:967-974

Chandok MR, Sopory SK, Oelmüller R (2001) Cytoplasmic kinase and phosphatase activities can induce PsaF gene expression in the absence of functional plastids: evidence that phosphorylation/dephosphorylation events are involved in interorganellar crosstalk. Mol Gen Genet 264:819-826

Chatterjee M, Sparvoli S, Edmunds C, Garosi P, Findlay K, Martin C (1996) *DAG*, a gene required for chloroplast differentiation and palisade development in *Antirrhinum majus*. EMBO J 15:4194-4207

Chen M, Choi Y, Voytas DF, Rodermel SR (2000) Mutations in the *Arabidopsis* Var2 locus cause leaf variegation due to the loss of a chloroplast FtsH protease. Plant J 22:303-313

Chen YB, Durnford DG, Koblizek M, Falkowski PG (2004) Plastid regulation of Lhcb1 transcription in the chlorophyte alga *Dunaliella tertiolecta*. Plant Physiol 136:3737-3750

Cohen A, Mayfield SP (1997) Translational regulation of gene expression in plants. Curr Opin Biotech 8:189-194

Cookson PJ, Kiano JW, Shipton CA, Fraser PD, Romer S, Schuch W, Bramley PM, Pyke KA (2003) Increases in cell elongation, plastid compartment size and phytoene synthase activity underlie the phenotype of the high pigment-1 mutant of tomato. Planta 217:896-903

Cornah JE, Terry MJ, Smith AG (2003) Green or red: what stops the traffic in the tetrapyrrole pathway? Trends Plant Sci 8:224-230

Couee I, Sulmon C, Gouesbet G, El Amrani A (2006) Involvement of soluble sugars in reactive oxygen species balance and responses to oxidative stress in plants. J Exp Bot 57:449-459

Danon A (1997) Translational regulation in the chloroplast. Plant Physiol 115:1293-1298

Danon A, Miersch O, Felix G, Camp RG, Apel K (2005) Concurrent activation of cell death-regulating signaling pathways by singlet oxygen in *Arabidopsis thaliana*. Plant J 41:68-80

Davison PA, Schubert HL, Reid JD, Iorg CD, Heroux A, Hill CP, Hunter CN (2005) Structural and biochemical characterization of Gun4 suggests a mechanism for its role in chlorophyll biosynthesis. Biochemistry 44:7603-7612

Davletova S, Rizhsky L, Liang H, Shengqiang Z, Oliver DJ, Coutu J, Shulaev V, Schlauch K, Mittler R (2005) Cytosolic ascorbate peroxidase 1 is a central component of the reactive oxygen gene network of *Arabidopsis*. Plant Cell 17:268-281

Delwiche C (1999) Tracing the thread of plastid diversity through the tapestry of life. Am Nat 154:S164-S177

Deng XW, Gruissem W (1987) Control of plastid gene expression during development: the limited role of transcriptional regulation. Cell 49:379-387

Depège N, Bellafiore S, Rochaix JD (2003) Role of chloroplast protein kinase Stt7 in LHCII phosphorylation and state transition in *Chlamydomonas*. Science 299:1572-1575

Desikan R, S AH-M, Hancock JT, Neill SJ (2001) Regulation of the *Arabidopsis* transcriptome by oxidative stress. Plant Physiol 127:159-172

Dubell AN, Mullet JE (1995a) Continuous far-red light activates plastid DNA-synthesis in pea leaves but not full cell enlargement or an increase in plastid number per cell. Plant Physiol 109:95-103

Dubell AN, Mullet JE (1995b) Differential transcription of pea chloroplast genes during light-induced leaf development transcription - Continuous far-red light activates chloroplast transcription. Plant Physiol 109:104-112

Durnford DG, Falkowski PG (1997) Chloroplast redox regulation of nuclear gene transcription during photoacclimation. Photosynth Res 53:229-241

Dyall SD, Brown MT, Johnson PJ (2004) Ancient invasions: From endosymbionts to organelles. Science 340:253-257

Emes MJ, Tobin AK (1993) Control of metabolism and development in higher plant plastids. Int Rev Cyt 145:149-216

Escoubas JM, Lomas M, LaRoche J, Falkowski PG (1995) Light intensity regulation of cab gene transcription is signaled by the redox state of the plastoquinone pool. Proc Natl Acad Sci USA 92:10237-10241

Ewing RM, Jenkins GI, Langdale JA (1998) Transcripts of maize RbcS genes accumulate differentially in C-3 and C-4 tissues. Plant Mol Biol 36:593-599

Fey V, Wagner R, Bräutigam K, Pfannschmidt T (2005a) Photosynthetic redox control of nuclear gene expression. J Exp Bot 56:1491-1498

Fey V, Wagner R, Bräutigam K, Wirtz M, Hell R, Dietzmann A, Leister D, Oelmüller R, Pfannschmidt T (2005b) Retrograde plastid redox signals in the expression of nuclear genes for chloroplast proteins of *Arabidopsis thaliana*. J Biol Chem 280:5318-5328

Foyer CH, Noctor G (1999) Plant biology - Leaves in the dark see the light. Science 284:599-601

Fraser PD, Schuch W, Bramley PM (2000) Phytoene synthase from tomato (*Lycopersicon esculentum*) chloroplasts - partial purification and biochemical properties. Planta 211:361-369

Freyer R, Kiefer-Meyer MC, Kossel H (1997) Occurrence of plastid RNA editing in all major lineages of land plants. Proc Natl Acad Sci USA 94:6285-6290

Fryer MJ, Ball L, Oxborough K, Karpinski S, Mullineaux PM, Baker NR (2003) Control of ascorbate peroxidase 2 expression by hydrogen peroxide and leaf water status during excess light stress reveals a functional organisation of *Arabidopsis* leaves. Plant J 33:691-705

Geigenberger P, Kolbe A, Tiessen A (2005) Redox regulation of carbon storage and partitioning in response to light and sugars. J Exp Bot 56:1469-1479

Giovannoni JJ (2004) Genetic regulation of fruit development and ripening. Plant Cell 16:S160-S170

Goldschmidt-Clermont M (1998) Coordination of nuclear and chloroplast gene expression in plant cells. Int Rev Cytol 177:115-180

Gray JC, Sornarajah R, Zabron AA, Duckett CM, Khan MS (1995) Chloroplast control of nuclear gene expression. In: Mathijs P (ed) Photosynthesis, from light to biosphere. Dordrecht: Kluwer, pp 543-550

Gray JC, Sullivan JA, Wang JH, Jerome CA, MacLean D (2003) Coordination of plastid and nuclear gene expression. Philos Trans R Soc Lond B Biol Sci 358:135-144

Grimm B (2003) Regulatory mechanisms of eukaryotic tetrapyrrole biosynthesis. In: Kadish KM, Smith KM, Guilard R (eds) The Porphyrin Handbook, Vol 12. San Diego: Academic Press, pp 1–32

Gruissem W (1989) Chloroplast gene expression: how plants turn their plastids on. Cell 56: 161-170

Gupta AK, Kaur N (2005) Sugar signalling and gene expression in relation to carbohydrate metabolism under abiotic stresses in plants. J Biosci 30:761-776

Haldrup A, Jensen PE, Lunde C, Scheller HV (2001) Balance of power: a view of the mechanism of photosynthetic state transitions. Trends Plant Sci 6:301-305

Hanaoka M, Kanamura M, Fujiwara M, Takahashi H, Tanaka K (2005) Glutamyl-tRNA mediates a switch in RNA polymerase use during chloroplast biogenesis. EMBO Rep 6:545-550

Herrin D, Battey J, Greer K, Schmidt G (1992) Regulation of chlorophyll apoprotein expression and accumulation. Requirements for carotenoids and chlorophyll. J Biol Chem 267:8260-8269

Herrmann RG, Westhoff P, Link G (1992) Biogenesis of plastids in higher plants. Wien: Springer Verlag

Hess WR, Börner T (1999) Organellar RNA polymerases of higher plants. Int Rev Cytol 190:1-59

Hess WR, Schendel R, Börner T, Rüdiger W (1991) Reduction of mRNA levels for two nuclear encoded light regulated genes in the barley mutant albostrians is not correlated with phytochrome content and activity. J Plant Physiol 138:292-298

Hou CX, Rintamäki E, Aro E-M (2003) Ascorbate-mediated LHCII protein phosphorylation - LHCII kinase regulation in light and in darkness. Biochemistry 42:5828-5836

Howe CJ, Barbrook AC, Koumandou VL, Nisbet RER, Symington HA, Wightman TF (2003) Evolution of the chloroplast genome. Phil Trans Roy Soc Lond B Biol Sci 358:99-106

Hricova A, Quesada V, Micol JL (2006) The *SCABRA3* nuclear gene encodes the plastid RpoTp RNA polymerase, which is required for chloroplast biogenesis and mesophyll cell proliferation in *Arabidopsis*. Plant Physiol 141:942-956

Jarvis P (2001) Intracellular signalling: the chloroplast talks! Curr Biol 11:R307-310

Johanningmeier U, Howell S (1984) Regulation of light-harvesting chlorophyll-binding protein mRNA accumulation in *Chlamydomonas reinhardtii*. Possible involvement of chlorophyll synthesis precursors. J Biol Chem 259:13541-13549

Josse EM (2000) A plastid terminal oxidase associated with carotenoid desaturation during chromoplast differentiation. Plant Physiol 123:1427-1436

Kanamaru K, Nagashima A, Fujiwara M, Shimada H, Shirano Y, Nakabayashi K, Shibata D, Tanaka K, Takahashi H (2001) An *Arabidopsis* sigma factor (SIG2)-dependent expression of plastid encoded tRNAs in chloroplasts. Plant Cell Physiol 42:1034-1043

Kanamaru K, Tanaka K (2004) Roles of chloroplast RNA polymerase sigma factors in chloroplast development and stress response in higher plants. Biosci Biotechnol Biochem 68:2215-2223

Karpinski S, Escobar C, Karpinska B, Creissen G, Mullineaux PM (1997) Photosynthetic electron transport regulates the expression of cytosolic ascorbate peroxidase genes in *Arabidopsis* during excess light stress. Plant Cell 9:627-640

Karpinski S, Reynolds H, Karpinska B, Wingsle G, Creissen G, Mullineaux P (1999) Systemic signaling and acclimation in response to excess excitation energy in *Arabidopsis*. Science 284:654-657

Keddie JS (1996) The *DCL* gene from tomato is required for chloroplast development and palisade cell morphogenesis in leaves. EMBO J 15:4208-4217

Kessler F, Schnell DJ (2006) The function and diversity of plastid protein import pathways: a multilane GTPase highway into plastids. Traffic 7:248-257

Kimura M, Manabe K, Abe T, Yoshida S, Matsui M, Yamamoto YY (2003) Analysis of hydrogen peroxide-independent expression of the high-light-inducible ELIP2 gene with the aid of the ELIP2 promoter-luciferase fusions. Photochem Photobiol 77:668-674

Kittsteiner U, Brunner H, Rüdiger W (1991) The greening process in cress seedlings. II. Complexing agents and 5-aminolevulinate inhibit accumulation of cab-mRNA coding for the light-harvesting chlorophyll a/b protein. Physiol Plant 81:190-196

Kleffmann T, Russenberger D, Von Zychlinski A, Christopher W, Sjölander K, Gruissem W, Baginsky S (2004) The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. Curr Biol 14:354-362

Kopriva S, Rennenberg H (2004) Control of sulphate assimilation and glutathione synthesis: interaction with N and C metabolism. J Exp Bot 55:1831-1842

Kovtun Y, Chiu WL, Tena G, Sheen L (2000) Functional analysis of oxidative stress-activated mitogen-activated protein kinase cascade in plants. Proc Natl Acad Sci USA 97:2940-2945

Kropat J, Oster U, Rüdiger W, Beck C (1997) Chlorophyll precursors are signals of chloroplast origin involved in light induction of nuclear heat-shock genes. Proc Natl Acad Sci USA, 94:14168-14172

Küchler M, Decker S, Hormann F, Soll J, Heins L (2002) Protein import into chloroplasts involves redox-regulated proteins. EMBO J 21:6136-6145

Kurth J, Varotto C, Pesaresi P, Biehl A, Richly E, Salamini F, Leister D (2002) Gene-sequence-tag expression analyses of 1,800 genes related to chloroplast functions. Planta 215:101-109

Larkin RM, Alonso JM, Ecker JR, Chory J (2003) GUN4, a regulator of chlorophyll synthesis and intracellular signaling. Science 299:902-906

Leister D (2005) Genomics-based dissection of the cross-talk of chloroplasts with the nucleus and mitochondria in *Arabidopsis*. Gene 354:110-116

Leon P, Arroyo A, Mackenzie S (1998) Nuclear control of plastid and mitochondrial development in higher plants. Ann Rev Plant Physiol Plant Mol Biol 49:453-480

Liere K, Börner T (2006) Transcription of plastid genes. In: Grasser KD (ed) Regulation of transcription in plants. Oxford: Blackwell Publishing Ltd., pp 184-224

Liere K, Maliga P (2001) Plastid RNA Polymerases in higher plants. In: Aro EM, Andersson B (eds) Regulation of photosynthesis. Dordrecht: Kluwer Academic Publishers, pp 29 - 49

Link G (1991) Photoregulated development of chloroplasts. In: Bogorad L, Vasil IK (eds) The photosynthetic apparatus: molecular biology and operation. Cell culture and somatic cell genetics of plants. San Diego: Academic Press, pp 365-394

Link G (1996) Green life: Control of chloroplast gene transcription. BioEssays 18:465-471

Link G (2003) Redox regulation of chloroplast transcription. Antioxid Redox Signal 5:79-87

Ljubicic JM, Wrischer M, Ljubicic N (1998) Formation of the photosynthetic apparatus in plastids during greening of potato microtubers. Plant Physiol Biochem 36:747-752

Lopez-Juez E, Pyke KA (2005) Plastids unleashed: their development and their integration in plant development. Int J Dev Biol 49:557-577

Lorrain S, Genoud T, Fankhauser C (2006) Let there be light in the nucleus! Curr Opin Plant Biol 9:509-514

Lu S, Van Eck J, Zhou X, Lopez AB, O'Halloran DM, Cosman KM, Conlin BJ, Paolillo DJ, Garvin DF, Vrebalov J, Kochian LV, Küpper H, Earle ED, Cao J, Li L (2006) The cauliflower *Or* gene encodes a DnaJ cysteine-rich domain-containing protein that mediates high levels of β-carotene accumulation. Plant Cell 18:3594-3605

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci USA 99:12246-12251

Martin W, Schnarrenberger C (1997) The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. Curr Genet 332:1-18

Matthews DE, Durbin RD (1990) Tagetitoxin inhibits RNA synthesis directed by RNA polymerases from chloroplasts and *Escherichia coli*. J Biol Chem 265:493-498

Maxwell DP, Laudenbach DE, Huner N (1995) Redox regulation of light-harvesting complex II and cab mRNA abundance in *Dunaliella salina*. Plant Physiol 109:787-795

Mayfield S, Nelson T, Taylor W, Malkin R (1986) Carotenoid synthesis and pleiotropic effects in carotenoid-deficient seedlings of maize. Planta 169:23-32

Mayfield S, Taylor W (1984) Carotenoid-deficient maize seedlings fail to accumulate light-harvesting chlorophyll a/b binding protein (LHCP) mRNA. Eur J Biochem 144:79-84

Mehler AH (1951) Studies on reactions of illuminated chloroplasts. I. Mechanism of the reduction of oxygen and other Hill reagents. Arch Biochem 33:65-77

Melis A (1991) Dynamics of photosynthetic membrane-composition and function. Biochim Biophys Acta 1058:87-106

Meskauskiene R, Nater M, Goslings D, Kessler F, op den Camp R, Apel K (2001) FLU: a negative regulator of chlorophyll biosynthesis in *Arabidopsis thaliana*. Proc Natl Acad Sci USA 98:12826-12831

Mingo Castel AM, Pelacho AM, De Felipe MC (1991) Amyloplast division in kinetin induced potato tubers. Plant Sci 73:211-217

Miyazawa Y, Sakai A, Miyagishima S-Y, Takano H, Kawano S, Kuroiwa T (1999) Auxin and cytokinin have opposite effects on amyloplast development and the expression of starch synthesis genes in cultured Bright Yellow 2-cells. Plant Physiol 121:461-469

Mochizuki N, Brusslan J, Larkin R, Nagatani A, Chory J (2001) *Arabidopsis* genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. Proc Natl Acad Sci USA 98:2053-2058

Mochizuki N, Susek R, Chory J (1996) An intracellular signal transduction pathway between the chloroplast and nucleus is involved in de-etiolation. Plant Physiol 112:1465-1469

Moller S, Kunkel T, Chua N (2001) A plastidic ABC protein involved in intercompartmental communication of light signaling. Genes Dev 15:90-103

Mullet JE, Klein RR (1987) Transcription and RNA stability are important determinants of higher plant chloroplast RNA levels. EMBO J 6: 1571-1579

Mullineaux P, Karpinski S (2002) Signal transduction in response to excess light: getting out of the chloroplast. Curr Opin Plant Biol 5:43-48

Mullineaux PM, Rausch T (2005) Glutathione, photosynthesis and the redox regulation of stress-responsive gene expression. Photosynth Res 86:459-474

Neuhaus E, Emes MJ (2000) Nonphotosynthetic metabolism in plastids. Ann Rev Plant Physiol Plant Mol Biol 51:111-140

Nickelsen J (2003) Chloroplast RNA-binding proteins. Curr Genet 43:392-399

Nixon PJ, Barker M, Boehm M, de Vries R, Komenda J (2005) FtsH-mediated repair of the photosystem II complex in response to light stress. J Exp Bot 56:357-363

Niyogi KK (2000) Safety valves for photosynthesis. Curr Opin Plant Biol 3:455-460

Noctor G, Gomez L, Vanacker H, Foyer C (2002) Interactions between biosynthesis, compartmentation and transport in the control of glutathione homeostasis and signalling. J Exp Bot 53:1283-1304

Nott A, Jung HS, Koussevitzky S, Chory J (2006) Plastid-to-nucleus retrograde signaling. Annu Rev Plant Biol 57:739-759

Oelmüller R (1989) Photooxidative destruction of chloroplasts and its effect on nuclear gene-expression and extraplastidic enzyme levels. Photochem Photobiol 49:229-239

Oelmüller R, Levitan I, Bergfeld R, Rajasekhar V, Mohr H (1986) Expression of nuclear genes as affected by treatments acting on the plastids. Planta 168:482-492

Oelmüller R, Mohr H (1986) Photooxidative destruction of chloroplasts and its consequences for expression of nuclear genes. Planta 167:106-113

Ogawa K, Hatano-Iwasaki A, Yanagida M, Iwabuchi M (2004) Level of glutathione is regulated by ATP-dependent ligation of glutamate and cysteine through photosynthesis in *Arabidopsis thaliana*: mechanism of strong interaction of light intensity with flowering. Plant Cell Physiol 45:1-8

Ogrzewalla K, Piotrowski M, Reinbothe S, Link G (2002) The plastid transcription kinase from mustard (*Sinapis alba* L.) - A nuclear-encoded CK2-type chloroplast enzyme with redox- sensitive function. Eur J Biochem 269:3329-3337

op den Camp RG, Przybyla D, Ochsenbein C, Laloi C, Kim C, Danon A, Wagner D, Hideg E, Gobel C, Feussner I, Nater M, Apel K (2003) Rapid induction of distinct stress responses after the release of singlet oxygen in *Arabidopsis*. Plant Cell 15:2320-2332

Osteryoung KW, Vierling E (1995) Conserved cell and organelle division. Nature 376:473-474

Oswald O, Martin T, Dominy PJ, Graham IA (2001) Plastid redox state and sugars: Interactive regulators of nuclear-encoded photosynthetic gene expression. Proc Natl Acad Sci USA 98:2047-2052

Papenbrock J, Grimm B (2001) Regulatory network of tetrapyrrole biosynthesis-studies of intracellular signalling involved in metabolic and developmental control of plastids. Planta 213:667-681

Pastori G, Kiddle G, Antoniw J, Bernard S, Veljovic-Jovanovic S, Verrier P, Noctor G, Christine H (2003) Leaf vitamin C contents modulate plant defense transcripts and regulate genes that control development through hormone signaling. Plant Cell 15:939-951

Peltier JB, Ytterberg AJ, Sun Q, van Wijk KJ (2004) New functions of the thylakoid membrane proteome of *Arabidopsis thaliana* revealed by a simple, fast and versatile fractionation strategy. J Biol Chem 279:49367-49383

Pesaresi P, Masiero S, Eubel H, Braun HP, Bhushan S, Glaser E, Salamini F, Leister D (2006) Nuclear photosynthetic gene expression is synergistically modulated by rates of protein synthesis in chloroplasts and mitochondria. Plant Cell 18:970-991

Petracek ME, Dickey LF, Huber SC, Thompson WF (1997) Light-regulated changes in abundance and polyribosome association of ferredoxin mRNA are dependent on photosynthesis. Plant Cell 9:2291-2300

Petracek ME, Dickey LF, Nguyen TT, Gatz C, Sowinski DA, Allen GC, Thompson WF (1998) Ferredoxin-1 mRNA is destabilized by changes in photosynthetic electron transport. Proc Natl Acad Sci USA 95:9009-9013

Pfannschmidt T (2003) Chloroplast redox signals: how photosynthesis controls its own genes. Trends Plant Sci 8:33-41

Pfannschmidt T (2005) Acclimation to varying light qualities: Toward the functional relationship of state transitions and adjustment of photosystem stoichiometry. J Phycol 41:723-725

Pfannschmidt T, Liere K (2005) Redox regulation and modification of proteins controlling chloroplast gene expression. Antioxid Redox Signal 7:607-618

Pfannschmidt T, Link G (1994) Separation of two classes of plastid DNA-dependent RNA polymerases that are differentially expressed in mustard (Sinapis alba L.) seedlings. Plant Mol Biol 25:69-81

Pfannschmidt T, Link G (1997) The A and B forms of plastid DNA-dependent RNA polymerase from mustard (*Sinapis alba* L.) transcribe the same genes in a different developmental context. Mol Gen Genet 257:35-44

Pfannschmidt T, Nilsson A, Allen JF (1999) Photosynthetic control of chloroplast gene expression. Nature 397:625-628

Pfannschmidt T, Ogrzewalla K, Baginsky S, Sickmann A, Meyer HE, Link G (2000) The multisubunit chloroplast RNA polymerase A from mustard (*Sinapis alba* L.) - Integration of a prokaryotic core into a larger complex with organelle-specific functions. Eur J Biochem 267:253-261

Pfannschmidt T, Schütze K, Brost M, Oelmüller R (2001) A novel mechanism of nuclear photosynthesis gene regulation by redox signals from the chloroplast during photosystem stoichiometry adjustment. J Biol Chem 276:36125-36130

Pfannschmidt T, Schütze K, Fey V, Sherameti I, Oelmüller R (2003) Chloroplast redox control of nuclear gene expression - A new class of plastid signals in interorganellar communication. Antioxid Redox Signal 5:95-101

Piippo M, Allahverdiyeva Y, Paakkarinen V, Suoranta UM, Battchikova N, Aro EM (2006) Chloroplastd-mediated regulation of nuclear genes in *Arabidopsis thaliana* in the absence of light stress. Physiol Genom 13:142-152

Puente P, Wei N, Deng X (1996) Combinatorial interplay of promoter elements constitutes the minimal determinants for light and developmental control of gene expression in *Arabidopsis*. EMBO J 15:3732-3743

Pursiheimo S, Mulo P, Rintamäki E, Aro EM (2001) Coregulation of light-harvesting complex II phosphorylation and lhcb mRNA accumulation in winter rye. Plant J 26:317-327

Pyke KA (1999) Plastid division and development. Plant Cell 11:549-556

Pyke KA, Leech RM (1994) A genetic analysis of chloroplast division and expansion in *Arabidopsis thaliana*. Plant Physiol 104:201-207

Race HL, Herrmann RG, Martin W (1999) Why have organelles retained genomes? Trends Genet 15:364-370

Reinbothe C, Bartsch S, Eggink LL, Hoober K, Brusslan J, Andrade-Paz R, Monnet J, Reinbothe S (2006) A role for chlorophyllide a oxygenase in the regulated import and stabilization of light-harvesting chlorophyll a/b proteins. Proc Natl Acad Sci USA 103:4777-4782

Richly E, Dietzmann A, Biehl A, Kurth J, Laloi C, Apel K, Salamini F, Leister D (2003) Covariations in the nuclear chloroplast transcriptome reveal a regulatory masterswitch. EMBO Rep 4:491-498

Robinson C, Thompson SJ, Woodhead C (2001) Multiple pathways used for the targeting of thylakoid proteins in chloroplasts. Traffic 2:245-251

Rodermel S (2001) Pathways of plastid-to-nucleus signaling. Trends Plant Sci 6:471-478

Rüdiger W, Grimm B (2006) Chlorophyll metabolism, an overview. In: Grimm B, Porra RJ, Rüdiger W, Scheer H (eds) Advances in Photosynthesis and Respiration. Dordrecht, The Netherlands: Springer, pp 133-146

Sakamoto W (2003) Leaf-variegated mutations and their responsible genes in *Arabidopsis thaliana*. Genes Genet Syst 78:1-9
Sato N (2001) Was the evolution of plastid genetic machinery discontinuous? Trends Plant Sci 6:151-155
Schäfer E, Bowler C (2002) Phytochrome-mediated photoperception and signal transduction in higher plants. EMBO Rep 3:1042-1048
Scheibe R (1991) Redox-modulation of chloroplast enzymes. A common principle for individual control. Plant Physiol 96:1-3
Soll J, Schleiff E (2004) Protein import into chloroplasts. Nat Rev Mol Cell Biol 5:198-208
Schmitz-Linneweber C, Williams-Carrier RE, Williams-Voelker PM, Kroeger TS, Vichas A, Barkan A (2006) A pentatricopeptide repeat protein facilitates the trans-splicing of the maize chloroplast rps12 pre-mRNA. Plant Cell 8:2650-2663
Schwacke R, Fischer K, Ketelsen B, Krupinska K, Krause K (2007) Comparative survey of plastid and mitochondrial targeting properties of transcription factors in *Arabidopsis* and rice. Mol Genet Genomics (doi:10.1007/s00438-007-0214-4)
Seo M, Koshiba T (2002) Complex regulation of ABA biosynthesis in plants. Trends Plant Sci 7:41-48
Shao N, Vallon O, Dent R, Niyogi KK, Beck CF (2006) Defects in the cytochrome b_6/f complex prevent light-induced expression of nuclear genes involved in chlorophyll biosynthesis. Plant Physiol 141:1128-1137
Shen Y, Wang X, Wu F, Du S, Cao Z, Shang Y, Wang X, Peng C, Yu X, Zhu S (2006) The Mg-chelatase H subunit is an abscisic acid receptor. Nature 443:823-826
Sherameti I, Sopory S K, Trebicka A, Pfannschmidt T, Oelmüller R (2002) Photosynthetic electron transport determines nitrate reductase gene expression and activity in higher plants. J Biol Chem 277:46594-46600
Shigeoka S, Ishikawa T, Tamoi M, Miyagawa Y, Takeda T, Yabuta Y, Yoshimura K (2002) Regulation and function of ascorbate peroxidase isoenzymes. J Exp Bot 53:1305-1319
Shikanai T (2006) RNA editing in plant organelles: machinery, physiological function and evolution. Cell Mol Life Sci 63:698-708
Shiina T, Tsunoyama Y, Nakahira Y, Khan MS (2005) Plastid RNA polymerases, promoters, and transcription regulators in higher plants. Int Rev Cytol 244:1-68
Soll J (2002) Protein import into chloroplasts. Curr Opin Plant Biol 5:529-535
Stern D, Higgs D, Yang J (1997) Transcription and translation in chloroplasts. Trends Plant Sci 2:308-315
Stoebe B, Maier UG (2002) One, two, three: nature's tool box for building plastids. Protoplasma 219:123-130
Strand A (2004) Plastid-to-nucleus signalling. Curr Opin Plant Biol 7:621-625
Strand A, Asami T, Alonso J, Ecker JR, Chory J (2003) Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrinIX. Nature 421:79-83
Streatfield SJ, Weber A, Kinsman EA, Häusler RE, Li J, Post-Beittenmiller D, Kaiser WM, Pyke KA, Flügge UI, Chory J (1999) The phosphoenolpyruvate/phosphate translocator is required for phenolic metabolism, palisade cell development, and plastid-dependent nuclear gene expression. Plant Cell 11:1609-1622
Sugiura M (1992) The chloroplast genome. Plant Mol Biol 19:149-168
Sugiura M (1995) The chloroplast genome. Essays Biochem 30:49-57

Sullivan JA, Gray JC (1999) Plastid translation is required for the expression of nuclear photosynthesis genes in the dark and in roots of the pea lip1 mutant. Plant Cell 11:901-910

Sullivan JA, Gray JC (2002) Multiple plastid genes regulate the expression of the pea plastocyanin gene in pea and transgenic tobacco plants. Plant J 32:763-774

Surpin M, Larkin R, Chory J (2002) Signal transduction between the chloroplast and the nucleus. Plant Cell 14:327-338

Susek R, Ausubel F, Chory J (1993) Signal transduction mutants of *Arabidopsis* uncouple nuclear CAB and RBCS gene expression from chloroplast development. Cell 74:787-799

Suzuki JY, Ytterberg AJ, Beardslee TA, Allison L, van Wijk KJ, Maliga P (2004) Affinity purification of the tobacco plastid RNA polymerase and in vitro reconstitution of the holoenzyme. Plant J 40:164-172

Takechi K, Sodmergen P, Murata M, Motoyoshi F, Sakamoto W (2000) The YELLOW VARIEGATED (VAR2) locus encodes a homologue of FtsH, an ATP-dependent protease in *Arabidopsis*. Plant Cell Physiol 41:1334-1346

Taylor WC (1989) Regulatory interactions between nuclear and plastid genomes. Ann Rev Plant Physiol Plant Mol Biol 40:211-233

The *Arabidopsis* Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796-815

Terzaghi W, Cashmore A (1995) Light-regulated transcription. Annu Rev Plant Physiol Plant Mol Biol 445:74

Tikkanen M, Piippo M, Suorsa M, Sirpio S, Mulo P, Vainonen J, Vener AV, Allahverdiyeva Y, Aro EM (2006) State transitions revisited-a buffering system for dynamic low light acclimation of *Arabidopsis*. Plant Mol Biol 62:779-793

Tiller K, Link G (1993) Phosphorylation and dephosphorylation affect functional characteristics of chloroplast and etioplast transcription systems from mustard (Sinapis alba L.). EMBO J 12:1745-1753

Tilney-Basset RAE (1978) The inheritance and behaviour of plastids. Amsterdam: Elsevier

Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nature Rev Genet 5:123-136

Trebst A (1980) Inhibitors in electron flow: tools for the functional and structural localization of carriers and energy conservation sites. Methods Enzymol 69:675-715

van Wijk KJ (2004) Plastid proteomics. Plant Physiol Biochem 42:963-977

Vandenabeele S, Van Der Kelen K, Dat J, Gadjev I, Boonefaes T, Morsa S, Rottiers P, Slooten L, Van Montagu M, Zabeau M, Inze D, Van Breusegem F (2003) A comprehensive analysis of hydrogen peroxide-induced gene expression in tobacco. Proc Natl Acad Sci USA 100:16113-16118

Vinti G, Fourrier N, Bowyer JR, Lopez-Juez E (2005) *Arabidopsis* cue mutants with defective plastids are impaired primarily in the photocontrol of the expression of photosynthesis-associated nuclear genes. Plant Mol Biol 57:343-357

Vinti G, Hills A, Campbell S, Bowyer J, Mochizuki N, Chory J, Lopez-Juez E (2000) Interactions between hy 1 and gun mutants of *Arabidopsis*, and their implications for plastid/nuclear signalling. Plant J 24:883-894

Vener A, van Kan PJ, Rich PR, Ohad II, Andersson B (1997) Plastoquinol at the quinol oxidation site of reduced cytochrome bf mediates signal transduction between light and protein phosphorylation: Thylakoid protein kinase deactivation by a single-turnover flash. Proc Natl Acad Sci USA 94:1585-1590

Verdecia MA, Larkin RM, Ferrer JL, Riek R, Chory J, Noel JP (2005) Structure of the Mg-chelatase cofactor GUN4 reveals a novel hand-shaped fold for porphyrin binding. PLoS Biol 3: e151

Von Wettstein D, Gough S, Kannangara CG (1995) Chlorophyll biosynthesis. Plant Cell 7:1039-1057

Wachter A, Wolf S, Steininger H, Bogs J, Rausch T (2005) Differential targeting of GSH1 and GSH2 is achieved by multiple transcription initiation: implications for the compartmentation of glutathione biosynthesis in the Brassicaceae. Plant J 41:15-30

Wagner D, Przybyla D, Op den Camp R, Kim C, Landgraf F, Lee KP, Wursch M, Laloi C, Nater M, Hideg E, Apel K (2004) The genetic basis of singlet oxygen-induced stress responses of *Arabidopsis thaliana*. Science 306:1183-1185

Wagner R, Pfannschmidt T (2006) Eukaryotic transcription factors in plastids - Bioinformatic assessment and implications for the evolution of gene expression machineries in plants. Gene 381:62-70

Walters RG (2005) Towards an understanding of photosynthetic acclimation. J Exp Bot 56:435-447

Walters RG, Horton P (1994) Acclimation of *Arabidopsis thaliana* to the light environment: Changes in composition of the photosynthetic apparatus. Planta 195:248-256

Wang HY, Deng XW (2003) Dissecting the phytochrome A-dependent signalling network in higher plants. Trends Plant Sci 8:172-178

Waters M, Pyke KA (2004) Plastid development and differentiation. In: Moller SG (ed) Plastids. Oxford: Blackwell, pp 30-59

Willows R, Hansson M (2003) Mechanism, structure, and regulation of magnesium chelatase. In: Kadish KM, Smith K, GuilardR (eds) The Porphyrin Handbook II, Vol 13. San Diego: Academic Press, pp 1-47

Wingler A, Purdy S, MacLean JA, Pourtau N (2006) The role of sugars in integrating environmental signals during the regulation of leaf senescence. J Exp Bot 57:391-399

Woitsch S, Römer S (2003) Expression of xanthophyll biosynthetic genes during light-dependent chloroplast differentiation. Plant Physiol 132:1508-1517

Xu XM, Adams S, Chua N-H, Moller SG (2005) AtNAP1 represents an atypical SufB protein in *Arabidopsis* plastids. J Biol Chem 280:6648-6654

Yabuta Y, Maruta T, Yoshimura K, Ishikawa T, Shigeoka S (2004) Two distinct redox signaling pathways for cytosolic APX induction under photooxidative stress. Plant Cell Physiol 45:1586-1594

Yang DH, Andersson B, Aro EM, Ohad I (2001) The redox state of the plastoquinone pool controls the level of the light-harvesting chlorophyll a/b binding protein complex II (LHC II) during photoacclimation. Photosynth Res 68:163-174

Zhang LX, Paakkarinen V, van Wijk KJ, Aro EM (2000) Biogenesis of the chloroplast-encoded D1 protein: Regulation of translation elongation, insertion, and assembly into photosystem II. Plant Cell 12:1769-1781

Zito F, Finazzi G, Delosme R, Nitschke W, Picot D, Wollman FA (1999) The Q_O site of cytochrome b6f complexes controls the activation of the LHCII kinase. EMBO J 18:2961-2969

Bräutigam, Katharina
Institute for General Botany and Plant Physiology, Junior Research Group, Friedrich-Schiller-University Jena, Dornburger Str. 159, 07743 Jena, Germany

Dietzel, Lars
Institute for General Botany and Plant Physiology, Junior Research Group, Friedrich-Schiller-University Jena, Dornburger Str. 159, 07743 Jena, Germany

Pfannschmidt, Thomas
Institute for General Botany and Plant Physiology, Junior Research Group, Friedrich-Schiller-University Jena, Dornburger Str. 159, 07743 Jena, Germany
Thomas.Pfannschmidt@uni-jena.de

List of abbreviations

ALA: 5-aminolevulinic acid
Chl: chlorophyll
Chlide: chlorophyllide
GSH: reduced glutathione
Mg-proto IX: Mg-protoporphyrin IX
NF: norflurazon
PDS: phytoene desaturase
PET: photosynthetic electron transport
PSI: photosystem I
PSII: photosystem II
PQ: plastoquinone
ROS: reactive oxygen species

The genetic transformation of plastids

Hans-Ulrich Koop, Stefan Herz, Timothy J Golds, and Jörg Nickelsen

Abstract

Biolistic delivery of DNA initiated plastid transformation research and still is the most widely used approach to generate transplastomic lines in both algae and higher plants. The principal design of transformation vectors is similar in both phylogenetic groups. Although important additions to the list of species transformed in their plastomes have been made in algae and in higher plants, the key organisms in the area are still the two species, in which stable plastid transformation was initially successful, i.e., *Chlamydomonas reinhardtii* and tobacco. Basic research into organelle biology has substantially benefited from the homologous recombination-based capability to precisely insert at predetermined loci, delete, disrupt, or exchange plastid genome sequences. Successful expression of recombinant proteins, including pharmaceutical proteins, has been demonstrated in *Chlamydomonas* as well as in higher plants, where some interesting agronomic traits were also engineered through plastid transformation.

1 Introduction

Plants are defined as the organisms containing plastids. Plant cells are operating and functioning through the integrated expression networks of nuclear, mitochondrial, and plastid genes. The capability of using genetic transformation for changing components of the integrated networks allows – in basic research – to study the interplay between the different genomes. In applied research, genetic transformation can optimize plants for their performance in natural or artificial environments and can introduce new functions such as the production of recombinant proteins or novel metabolites. Stable genetic transformation of plastids was first introduced for *Chlamydomonas* almost 20 years ago (Boynton et al. 1988; Blowers et al. 1989), and was successfully applied to the higher plant *Nicotiana tabacum* L. (tobacco) soon afterwards (Svab et al. 1990). In both species transformation involves a single or very few plastid DNA molecules initially, which leads to cells or organisms containing genetically different plastomes. These are termed "heteroplasmic" (Fig. 1). Distribution of plastid DNA molecules (and, in higher plants, plastids) among the daughter cells originating from mitosis is a statistical process. As a consequence, segregation of different plastid DNA molecules occurs. Under appropriate selection this process leads to cells (organisms) containing only transformed plastomes, which are called "homoplasmic" (Fig. 1). Several

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0225 / Published online: 15 May 2007

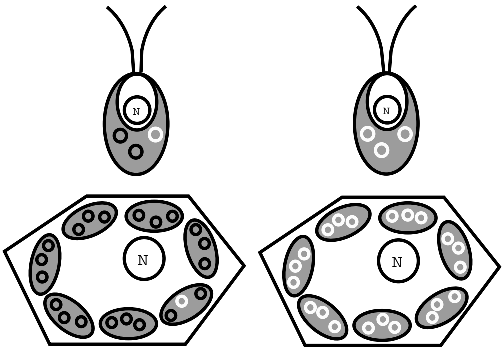


Fig. 1. Schematic representation of heteroplasmic (left) and homoplasmic (right) cells of *Chlamydomonas* (top) and a higher plant (bottom). N: nucleus; plastids are in grey; plastid DNA molecules are depicted as black (wild type) or white (transformed) circles. Note that drawings are not to scale and numbers of DNA molecules and (in higher plants) plastids are far too low.

hundred original scientific reports and numerous reviews (Table 1) on research in the area have been published in the meantime. This review attempts at serving as a reference article and at providing an actual update on this research.

1.1 Plastid biology in *Chlamydomonas* and tobacco

The basic technology of genetic transformation of plastids was developed for *Chlamydomonas reinhardtii*, and most of the technical features, like transfer of DNA via particle bombardment and some of the selection markers used, were directly applicable for higher plants. One could argue therefore that *Chlamydomonas* can serve as a model for plastid transformation in higher plants. This view is attractive since, due to the presence of only one plastid per cell and the much shorter generation time, in *Chlamydomonas* it takes only four to six weeks to reach homoplasmy, where in tobacco four to six months are necessary. As pointed out by Maliga (1993), there are however some important differences in plastid biology between single-celled algae and higher plants, which need to be kept in mind when trying to directly transfer techniques between the different groups of organisms. These differences include, e.g., the size of the genome and the genes present, the number of genome copies per cell, and the number of nucleoids per plastid. The morphology and intracellular location of plastids, and the occurrence of tissue specific plastid forms in higher plants are further differences. Furthermore, the fate of plastids before and after fertilization, the option to switch between photoautotrophic, mixotrophic and heterotrophic growth conditions in *Chlamydomonas,* whereas the target tissues for plastid transformation in higher plants always are heterotrophic in vitro cultures, and, finally, the shorter time to reach homoplasmy (Fig. 1) in the single-celled alga as compared to higher plants (Table 2) are also factors different in the two model systems.

Table 1. Reviews on genetic transformation of plastids.

Author(s)	Year	Title
Howe CJ	1988	Organelle transformation.
Butow RA, Fox TD	1990	Organelle transformation: shoot first, ask questions later.
Maliga P	1993	Towards plastid transformation in flowering plants.
Maliga P et al.	1993	Plastid engineering in land plants: a conservative genome is open to change.
Dix PJ, Kavanagh TA	1995	Transforming the plastome: genetic markers and DNA delivery systems.
Rochaix JD	1995	*Chlamydomonas reinhardtii* as the photosynthetic yeast
Rochaix JD	1997	Chloroplast reverse genetics: new insights into the function of plastid genes.
Bock R	1998	Analysis of RNA editing in plastids.
Kofer W et al.	1998a	PEG-mediated plastid transformation in higher plants.
Bock R	2000	Sense from nonsense: how the genetic information of chloroplasts is altered by RNA editing.
Bogorad L	2000	Engineering chloroplasts: an alternative site for foreign genes, proteins, reactions and products.
Daniell H	2000	Genetically modified food crops: current concerns and solutions for next generation crops.
Hager M, Bock R	2000	Enslaved bacteria as new hope for plant biotechnologists.
Heifetz PB	2000	Genetic engineering of the chloroplast.
Nickelsen J, Kück U	2000	The unicellular green alga *Chlamydomonas reinhardtii* as an experimental system to study chloroplast RNA metabolism.
Bock R	2001	Transgenic plastids in basic research and plant biotechnology.
Daniell H et al.	2001a	Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants.
Heifetz PB, Tuttle AM	2001	Protein expression in plastids.
van Bel AJ et al.	2001	Novel approach in plastid transformation.
Daniell H	2002	Molecular strategies for gene containment in transgenic crops.
Daniell H, Dhingra A	2002	Multigene engineering: dawn of an exciting new era in biotechnology.
Daniell H et al.	2002	Milestones in chloroplast genetic engineering: an environmentally friendly era in biotechnology.
Maliga P	2002	Engineering the plastid genome of higher plants.
Staub JM	2002	Expression of recombinant proteins via the plastid genome.
Maliga P	2003	Progress towards commercialization of plastid transformation technology.

Author(s)	Year	Title
Walmsley AM, Arntzen CJ	2003	Plant cell factories and mucosal vaccines.
Bock R	2004	Studying RNA editing in transgenic chloroplasts of higher plants.
Bock R, Khan MS	2004	Taming plastids for a green future.
Franklin SE, Mayfield SP	2004	Prospects for molecular farming in the green alga *Chlamydomonas reinhardtii*.
Lorence A, Verpoorte R	2004	Gene transfer and expression in plants.
Maliga P	2004	Plastid transformation in higher plants.
Ramesh VM, Bingham SE, Webber AN	2004	A simple method for chloroplast transformation in *Chlamydomonas reinhardtii*.
Tregoning J et al.	2004	New advances in the production of edible plant vaccines: chloroplast expression of a tetanus vaccine antigen, TetC.
Xiong L, Sayre RT	2004	Engineering the chloroplast encoded proteins of *Chlamydomonas*.
Daniell H et al.	2005a	Chloroplast-derived vaccine antigens and other therapeutic proteins.
Daniell H et al.	2005b	Breakthrough in chloroplast genetic engineering of agronomically important crops.
Daniell H et al.	2005c	Chloroplast genetic engineering to improve agronomic traits.
Khan MS et al.	2005	Phage phiC31 integrase: a new tool in plastid genome engineering.
Ma JK et al.	2005	Molecular farming for new drugs and vaccines. Current perspectives on the production of pharmaceuticals in transgenic plants.
Maliga P	2005	New vectors and marker excision systems mark progress in engineering the plastid genome of higher plants.
Mayfield SP, Franklin SE	2005	Expression of human antibodies in eukaryotic micro-algae.
Nugent JM, Joyce SM	2005	Producing human therapeutic proteins in plastids.
Chase CD	2006	Genetically engineered cytoplasmic male sterility.
Daniell H	2006	Production of biopharmaceuticals and vaccines in plants via the chloroplast genome.
Dhingra A, Daniell H	2006	Chloroplast genetic engineering via organogenesis or somatic embryogenesis.
Lu XM et al.	2006	Chloroplast transformation.
Lutz KA et al.	2006a	Construction of marker-free transplastomic tobacco using the Cre-loxP site-specific recombination system.
Bock R	2006	Plastid biotechnology: prospects for herbicide and insect resistance, metabolic engineering and molecular farming.

Table 2. Features of *Chlamydomonas reinhardtii* and *Nicotiana tabacum* as model species for plastid transformation[a].

Feature		*Chlamydomonas reinhardtii*	*Nicotiana tabacum*
features of the organism			
organization		single-celled	multicellular, highly differentiated
time to reach homoplasmy[b]		three to four weeks	three to four months
ploidy level	haploid	+	(-)
	diploid	+	+
culture conditions	autotrophic	+	-
	mixotrophic	+	-
	heterotrophic	+	+
transformable	nucleus	+	+
	mitochondrion	+	-
	plastid	+	+
genome sequenced	nucleus	+	-
	mitochondrion	+	-
	plastid	+	+
features of the organelle			
plastid morphology		cup-shaped	lentiform
plastid size (µm diameter)		eight to ten	five to ten
plastid fusion after fertilization		+	-
plastids per cell[c]		1	100
plastid type	proplastid	-	+
	etioplast	-	+
	chloroplast	+	+
	leucoplast	-	+
	chromoplast	-	+
	gerontoplast	-	+
eyespot		+	-
pyrenoid		+	-
features of the plastome			
nucleoids per plastid		10	10-50
plastome copies per cell[c]		80	500-10000
plastome size		203.395 bp	155.943 bp
size of inverted repeat		21,2 kbp	26,4 kbp
protein genes		69	101
RNA genes		40	45
GC content		37%	34%
coding sequences		38%	49%
short dispersed repeats		20%	-

[a] Compiled from GenBank entries BK000554 (*Chlamydomonas*) and Z00044 (tobacco), respectively, and from Grossman et al. (2003), Maliga (1993), Maul et al. (2002), Rochaix (1995), and Yukawa et al. (2005).
[b]See Fig. 1
[c] In the case of tobacco the term "cell" refers to fully developed mesophyll cells

2 General procedures

Differences in the structure of the organism between (single-celled) algae and multicellular and highly-differentiated higher plants primarily have an influence on the selection process in plastid transformation, while the basic processes of introduction of DNA into the organelle, of integration of sequences into the plastid DNA and of gene expression control are similar. Therefore, we will first describe procedures, which have been used irrespective of the species in question and will then address algae and higher plants separately.

2.1 Gene transfer methods

Stably transformed lines have primarily been generated by using two methods to deliver transforming DNA into plastids, the particle gun-mediated biolistic process and treatment of isolated protoplasts with polyethylene glycol (PEG) [for a detailed description of the particle bombardment process see e.g. Boynton et al. 1988; Lutz et al. 2006a; for PEG treatment, see Kofer et al. 1998a)]. The mechanism of entry of the transforming DNA is assumed to be by mechanical impact: microprojectiles supposedly, after passing the cell wall, penetrate the organelle's envelope, thus, carrying the DNA inside. It is not known whether or how a chloroplast envelope would reseal after penetration. The mechanism of DNA entry after PEG-treatment is even less clear. The assumption is that PEG produces transient 'holes', in the plasma membrane through which DNA can enter into the cell (Paszkowski et al. 1984). This would lead to deposition of plasmids into the cytosol, although it remains completely unknown how the DNA could subsequently reach the inside of the plastids. If, however, there is transfer of DNA from the cytosol into the plastids, then it is conceivable that also with particle bombardment plasmids are primarily delivered into the cytosol and enter the organelle afterwards. Particle bombardment (Boynton et al. 1988) is the method primarily used for the genetic transformation of plastids in algae as well as in higher plants. PEG-treatment of protoplasts (Golds et al. 1993) was successfully used in a number of higher plant species (see Table 9). In tobacco, plastid transformation is highly efficient irrespective of the methods used for DNA delivery. Another, but less efficient, technique is vortexing of cell-wall deficient algal cells with glass beads (Kindle et al. 1991). A femtosyringe-based microinjection procedure was used to deliver a GFP gene into plastids (Knoblauch et al. 1999; van Bel et al. 2001), and transient expression was clearly achieved, stable transformants were, however, not described. Earlier reports on *Agrobacterium*-mediated plastid transformation were never subsequently confirmed (de Block et al. 1985; Venkateswarlu and Nazar 1991).

2.2 Transformation vectors

Naturally, the vector design depends on the purpose of a specific experiment. In contrast to stable transformation, for transient expression a plasmid carrying a functional expression cassette would suffice, and no sequences necessary for stable integration are required.

2.2.1 Transient expression

Relatively few reports have been published on transient expression in plastids in vivo. Note, that gene products detected in experiments analysing transient expression might at least in part be due to transcription from sequences integrated via co-integrate formation (Klaus et al. 2004), if vectors containing extended plastome sequences were used. Daniell et al. (1990) reported expression of chloramphenicol acetyl transferase in bombarded tobacco suspension cells, but no expression was found after electroporation of suspension cell-derived protoplasts. The expression was assigned to plastids (supposedly leucoplasts and not chloroplasts). Transient expression of GUS following particle bombardment in tobacco (Ye et al. 1990; Daniell et al. 1991; Seki et al. 1995) or PEG-treatment of leaf protoplasts in *Nicotiana plumbaginifolia* (Spörlein et al. 1991) was also reported. In the protoplasts, the GUS protein, which was co-purified with plastids, was proteinase stable, in contrast to protein derived from a nucleo/cytosolic reporter construct. GFP served as a reporter for transient expression after particle bombardment (Hibberd et al. 1998) and after femtosyringe-mediated microinjection (Knoblauch et al. 1999). Expression is quite cumbersome to detect and difficult to quantify after particle bombardment or microinjection. Thus, a versatile, reliable and easy to use system for quantitative transient expression studies is still missing.

2.2.2 Stable transformation

The organelle's recombination system requires sequences on the transformation vector with sufficient homology to the target plastome to allow for homologous recombination. Such 'homologous flanks' are generally about 1 kbp in length. Shorter flanks would presumably reduce recombination efficiency, while significantly longer flanks cause technical problems with vector construction. Expression cassettes in plastid transformation vectors require regulatory elements, such as promoters, 5' UTRs, ribosome binding sites and 3' UTRs, which are compatible with the plastid gene expression machinery. A heterologous transcription system can also be used, consisting of a foreign RNA polymerase and an expression cassette, which is equipped with a suitable promoter (McBride et al. 1994). Further modifications consist of a "downstream box" for enhanced translation efficiency (Kuroda and Maliga 2001a; Herz et al. 2005), fusion and/or purification tags for enhanced protein stability and facilitation of protein extraction (Leelavathi and Reddy 2003), and protease cleavage sites, if authentic starting amino acids are required for a desired protein end product (Staub et al. 2000). Artificial operons may

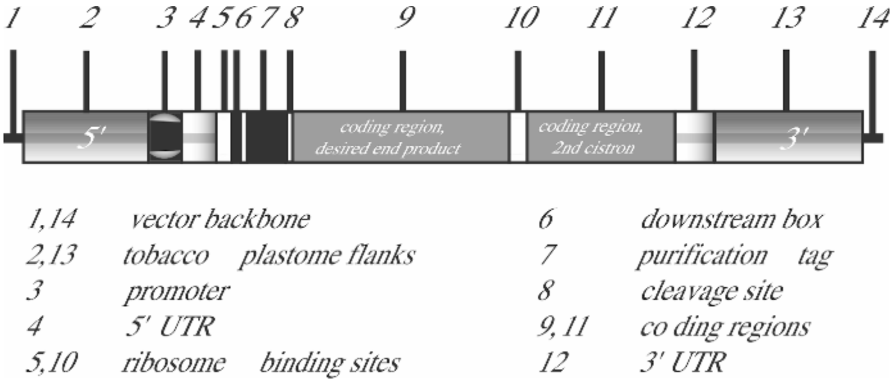


Fig. 2. Elements of a dicistronic plastid transformation vector.

contain one or more cistrons (Staub and Maliga 1995a; Lössl et al. 2003; Arai et al. 2004; Quesada-Vargas et al. 2005; Herz et al. 2005). A plastid transformation vector for the insertion of a dicistronic operon is depicted in Fig. 2. A selection marker gene is certainly also required for the transformation process. If the second cistron is used for this purpose, it is safe to assume that the protein encoded by the first cistron is also transcribed in the selected lines, due to read-through transcription. Alternatively, a selection marker cassette could be positioned elsewhere on the same transformation vector or on a different transformation vector and used in a co-transformation approach, which works efficiently in plastid transformation (Carrer and Maliga 1995; Herz et al. 2005). Up to four genes combined in a single operon were successfully introduced into the tobacco plastome (Nakashita et al. 2001; Lössl et al. 2003; Arai et al. 2004; Quesada-Vargas et al. 2005), and it might be possible to co-introduce and co-express even higher numbers of cistrons.

Not all elements given in Fig. 2 are absolutely required. Separate promoters are not necessary, if transcription is mediated by endogenous transcription start signals (Staub and Maliga 1995a). Such "operon extension vectors" are described in detail by Herz et al. (2005). Interestingly, the level of expression may even be higher if transcription is controlled by an endogenous rather than a separate promoter. It is also possible to use incomplete expression cassettes on co-transforming, separate transformation vectors, since complete and functional expression cassettes can be assembled from such "split vectors" by homologous recombination inside the plant after transformation (Herz et al. 2005).

Lutz et al. (2004) demonstrated that the integrase of phage phiC31 can be used for integrating foreign DNA into the tobacco plastome, if in a preceding transformation step suitable recognition sequences (attB) had been introduced into the target plastome. The real advantage of this approach remains to be demonstrated. Integration of attB elements relies on the endogenous recombination system – which was speculated to be rate-limiting – and the integrase has to be supplied either via another stable transformation or transiently.

2.2.3 Episomal maintenance of foreign sequences

Considerable effort was invested in studying the possibility of introducing constructs that are not integrating into the plastid chromosome and are maintained as episomes. In *Chlamydomonas*, Kindle et al. (1994) found highly amplified plasmid copies that were capable of correcting a photosynthetic growth defect. Suzuki et al. (1997) analyzed the transformed lines further and found characteristic rearrangements in both copies of the inverted repeat. Attempts to generate plastid transformants using vectors without any homology to the recipient plastome failed. Re-transformation of the lines that contained amplified plasmid copies with standard-type vectors surprisingly led to loss of the amplified plasmids. Thus, the mechanisms leading to establishment of plasmids apparently capable of autonomously replicating in plastids are presently not understood. Interestingly, plastid transformation in *Euglena gracilis* (Doetsch et al. 2001) also involved episomal elements. In tobacco, a potentially autonomously replicating element, NICE 1, was described (Staub and Maliga 1994, 1995b). However, as correctly stated by the authors, replication of this element could have also occurred while integrated into the plastid chromosome (see also: Klaus et al. 2004). Interestingly, Mühlbauer et al. (2002) did not find any influence on plastid replication activity after inactivation or deletion of the NICE 1 sequences from the tobacco plastome.

2.3 Marker gene removal

Marker removal approaches are useful, when the number of available selection markers is limited and multiple consecutive transformation steps are required for generating a desired end product. Furthermore, expression of marker genes constitutes an unnecessary metabolic burden on transplastomic plants. In addition, public concern requests removal of antibiotic resistance marker genes from transgenic plants intended for human consumption or animal feed. Three different strategies are available for transplastomic lines: direct repeat-mediated loop-out recombination, segregation of different plastomes after use of separate transformation vectors for selection marker and gene of interest, and marker excision using site-specific recombinases.

2.3.1 Direct repeat-mediated loop-out recombination

The highly active recombination system of plastids leads to loop-out recombination of introduced sequences, if transformation generates direct repeats in the plastome. This needs to be considered when designing transformation vectors, since undesired loss of sequences might follow otherwise (Maliga et al. 1993; Zou et al. 2003). On the other hand, direct repeat-mediated loop-out recombination can also be used for marker removal as initially shown for *Chlamydomonas* by Fischer et al. (1996) and later for higher plants by Iamtham and Day (2000) and Durfourmantel et al. (2006). The marker is maintained as long as a selective pressure is present. After removal of selection, marker gene sequences are excised; however, transformants cannot be distinguished phenotypically from wild type lines. Therefore, this system benefits from the availability of a secondary selection system, e.g., herbicide resistance (Iamtham and Day 2000). Interestingly, the same approach can be applied for targeted gene inactivation in chloroplasts (Kode et al. 2005, 2006). In a different approach, Klaus et al. (2004) used transformation vectors with a different architecture. They positioned the selection marker cassette outside the homologous flanks, such that the marker can never become stably integrated. This approach was stimulated by their observation that recombination via a single flank occurs routinely, leading to the formation of vector co-integrates, which are later resolved through secondary recombination events. Secondary recombination within co-integrate structures automatically results either in plastomes identical to those of the acceptor lines or marker-free plastomes containing the gene of interest. Klaus et al. (2004) applied phenotypical selection using pigmentation. Alternatively, use of a secondary selection marker, PCR screening or other visible markers could also be conceived as a means to assist in detecting the regenerates containing the gene of interest. A difference in using vectors with the marker gene in the vector backbone lies in the fact that loop-out recombination can only occur after integration of vector sequences into the plastome, whereas in the process described by Fischer et al. (1996) and Iamtham and Day (2000) the marker gene might be lost even prior to the transformation event itself. Whether this might reduce transformation efficiency, is not known.

2.3.2 Co-transformation and segregation

Co-transformation was first shown to be possible in tobacco plastids by Carrer and Maliga (1995). In *Chlamydomonas*, Fischer et al. (1996) inserted the resistance marker into an essential gene. Thus insertion of the marker, i.e., disruption of the essential gene, could not be driven to homoplasmy and segregation allowed for the recovery of lines containing the gene of interest but not the marker gene. Ye et al. (2003) used two different vectors in tobacco and a scheme, which was initially based on spectinomycin as the selective inhibitor and subsequently on an herbicide. The rational behind this scheme is that initial selection with herbicides is not possible in plastid transformation, whereas after enrichment for transplastomes, the level of herbicide tolerance might be sufficient to, in a heteroplasmic situation, allow for segregation of lines, which carry the herbicide but not the antibiotic resistance genes. Indeed, 20% of the recovered lines fulfilled this criterion.

2.3.3 Use of site-specific recombinases

CRE recombinase-mediated marker removal from transplastomic tobacco was independently reported from two different groups (Hajdukiewicz et al. 2001; Corneille et al. 2001). CRE recombinase, derived from the P1 bacteriophage, mediates insertion or excision of sequences, provided that recognition elements, *loxP* sites, are present on the recombination substrate molecules. Marker gene removal, thus, requires directly repeated *loxP* elements flanking the marker gene in the plastome. CRE recombinase can be expressed from a nuclear expression cassette, translated in the cytosol and then introduced into the plastid through the organelle's import machinery. When executing this approach, additional plastome rearrangements were found that were not necessarily only due to 'cryptic' *lox* sites in the plastome (Corneille et al. 2003) but were either based on short direct repeats or on recombination 'hot spots'. CRE recombinase seems to generally increase recombination activity in the plastome. Introduction of the recombinase into the transplastomic lines can either be established by *Agrobacterium*-mediated stable or transient (Lutz et al. 2006a, 2006b) nuclear transformation or by crossing a transplastomic line with a suitable nuclear transformant. Marker removal through *Agroinfiltration*-based transient expression (Lutz et al. 2006a, 2006b) is efficient and clearly preferable since removal of stably integrated expression cassettes from the nuclear genome is not necessary. In addition, plastome rearrangements no longer occur once the CRE recombinase is absent.

3 Plastid transformation in algae

3.1 Expression control elements

In *Chlamydomonas reinhardtii*, extensive efforts have been made to identify the crucial cis-acting determinants that regulate chloroplast gene expression by systematic site-directed mutagenesis of plastid 5′ and 3′ regions after co-integration or co-transformation of selectable marker genes (see 3.2). As a consequence, several elements that affect RNA stability and translational activities have been mapped especially in the 5′ UTRs of various chloroplast mRNAs (see Herrin and Nickelsen 2004; and chapters by Stern and Danon in this issue).

Expression of foreign genes in algae was performed with only a limited set of 5′ and 3′ regions as listed in Table 4. However, the analysis of reporter gene expression after systematic testing of various combinations of these 5′ and 3′ regulatory elements revealed that the *atpA* and *psbD* 5′ regions including the respective promoters and 5′ UTRs confer the highest expression rates for both the *uidA* and *gfp* reporter genes (Ishikura et al. 1999; Kasai et al. 2003; Barnes et al. 2005). In contrast, the *rbcL* and *psbA* 5′ regions produce less mRNA and protein while the nature of the 3′ UTR had only a small impact on reporter gene expression. Overall, a direct correlation of mRNA and protein levels was observed with some notable exceptions (Barnes et al. 2005; Kato et al. 2006).

To date, it remains to be clarified whether similar to the situation in vascular plants (see 4.1), viral, or artificial cis-regulatory elements work also in *C. reinhardtii*. However, it was demonstrated that neither the *psbA* 5′ region from wheat nor the spinach *psbB* 5′ region consisting of the promoter and 5′ UTR each were capable of producing stable *aadA* reporter gene mRNA in *C. reinhardtii* chloroplasts despite the fact that the genes were efficiently transcribed (Nickelsen 1999). This suggests that the post-transcriptional principles of chloroplast gene expression in algae and plants differ to some extent.

3.2 Resistance marker genes

Three principle strategies have been used for selecting chloroplast transformants of *Chlamydomonas reinhardtii*. The initial selection scheme was based on the wide range of chloroplast mutants with photosynthetic defects, which had been isolated during several decades of classical genetic work. These mutants were complemented by the respective intact wild type genes resulting in restored photo-autotrophy. For instance, Boynton et al. (1988) used in their pioneer work a *C. reinhardtii atpB* deletion mutant, which they transformed with an *atpB* gene fragment. Another example is represented by the *tscA* gene that enabled the restoration of photosystem I activity in the chloroplast mutant *H13* (Goldschmidt-Clermont et al. 1991).

A second strategy involved the use of mutations within rRNA genes that confer resistances to antibiotics like spectinomycin, streptomycin, or erythromycin (for an overview see Goldschmidt-Clermont 1998). Moreover, mutations in the *psbA* gene conferring resistance to herbicides like metribuzin or DCMU were used for selection of transformants (Przibilla et al. 1991; Newman et al. 1992) and in the red alga *Porphyridium sp*., a mutant form of the chloroplast-encoded acetohydroxyacid synthase (AHAS) gene allowed the selection of chloroplast transformants using the herbicide sulfometuron methyl (Lapidot et al. 2002).

Finally, a third - and nowadays commonly applied - strategy is based on the expression of bacterial genes whose gene products inactivate antibiotics. The *aadA* gene from *Escherichia coli* conferring resistance to spectinomycin and streptomycin is widely used in *C. reinhardtii* (Goldschmidt-Clermont 1991) and, more recently, the *aphA-6* gene from *Acinetobacter baumannii* has also been shown to be suitable for selecting chloroplast transformants on kanamycin- or amikacin-containing media (Bateman and Purton 2000).

3.3 Targeted inactivation

Although the long-standing isolation of chloroplast mutants of *Chlamydomonas reinhardtii* had already enabled one to assign distinct functions to several chloroplast genes, the establishment of the chloroplast transformation system by Boynton et al. (1988) immediately opened the door for the systematic inactivation of chloroplast genes of unknown function. The first targeted gene disruption affected PsaC, a subunit of PS I, which was shown to be essential for PS I activity (Takahashi et al. 1991). At the same time and as mentioned above (3.2), the chloroplast *tscA* locus was mapped by biolistic complementation of the mutant strain *H13* and shown to encode a small RNA which is required for the trans-splicing process generating mature *psaA* mRNA and, thus, active PS I (Goldschmidt-Clermont et al. 1991). In the meantime, 36 genes of the *C. reinhardtii* genome have been inactivated, which are listed in Table 3, representing an updated version of the one published by Grossman et al. (2003). Only six genes turned out to be essential, i.e., could not be brought to homoplasmy. These include three genes for subunits of the chloroplast-encoded RNA polymerase, a ribosomal protein gene, the *clpP*

Table 3. Inactivated chloroplast genes in *Chlamydomonas reinhardtii*.

Gene	Inactivation status	Reference
RNA-polymerase		
rpoB1	heteroplasmic	Fischer et al. 1996
rpoB2	heteroplasmic	Fischer et al. 1996
rpoC2	heteroplasmic	Fischer et al. 1996
photosystems		
psaA	homoplasmic	Redding et al. 1999
psaB	homoplasmic	Redding et al. 1999
psaC	homoplasmic	Takahashi et al. 1991
psaJ	homoplasmic	Fischer et al. 1999
tscA	homoplasmic	Goldschmidt-Clermont et al. 1991
ycf3	homoplasmic	Boudreau et al. 1997a
ycf4	homoplasmic	Boudreau et al. 1997a
psbA	homoplasmic	Bennoun et al. 1986
psbC	homoplasmic	Rochaix et al. 1989
psbD	homoplasmic	Erickson et al. 1986
psbE	homoplasmic	Morais et al. 1998
psbH	homoplasmic	Summer et al. 1997; O´Connor et al. 1998
psbI	homoplasmic	Kunstner et al. 1995
psbK	homoplasmic	Takahashi et al. 1994
psbT	homoplasmic	Ohnishi and Takahashi 2001
psbZ	homoplasmic	Swiatek et al. 2001
petA	homoplasmic	Kuras and Wollman 1994
petB	homoplasmic	Kuras and Wollman 1994
petD	homoplasmic	Kuras and Wollman 1994
petG	homoplasmic	Berthold et al. 1995
petL	homoplasmic	Takahashi et al. 1996
atpA	homoplasmic	Drapier et al. 1998
atpB	homoplasmic	Shepherd et al. 1979
atpE	homoplasmic	Robertson et al. 1990
RUBISCO		
rbcL	homoplasmic	Spreitzer et al. 1985
ribosomal proteins		
rps3	heteroplasmic	Liu et al. 1993
protease		
clpP	heteroplasmic	Huang et al. 1994; Majeran et al. 2000
chlorophyll synthesis		
chlB	homoplasmic	Li et al. 1993
chlL	homoplasmic	Suzuki and Bauer 1992
chlN	homoplasmic	Choquet et al. 1992
others		
cemA	homoplasmic	Rolland et al. 1997
ccsA	homoplasmic	Xie and Merchant 1996
ORF1995	heteroplasmic	Boudreau et al. 1997b

gene and ORF1995. Recently a procedure was described which allows the analysis of the function of such essential genes by reducing the gene product levels. This strategy, named translational attenuation, is based on the finding that reduced

Table 4. Chloroplast expression of foreign genes in *Chlamydomonas reinhardtii*.

Protein	Expression	Insertion site	Expression construct	Reference
reporter proteins				
ß-glucuronidase (*Escherichia coli*)	0.08%	*rbcL-psaB*	*PatpA 5'atpA uidA 3'atpA*	Ishikura et al. 1999
ß-glucuronidase (*Escherichia coli*)	0.009%	*rbcL-psaB*	*PrbcL 5'rbcL uidA 3'rbcL*	Ishikura et al. 1999
ß-glucuronidase (*Escherichia coli*)	34.4 nmol/h mg	*atpB-IR*	*PpetD 5'petD uidA 3'rbcL*	Sakamoto et al. 1993
luciferase (*Renilla reniformis*)	n.a	*tscA-chlN*	*PatpA 5'atpA rluc 3'atpA*	Minko et al. 1999
luciferase (*Vibrio harveyi*, codon adapted)	450 U/µg	*psbA-5SrRNA*	*PpsbA 5'psbA luxCt 3'rbcL*	Mayfield and Schultz 2004
luciferase (*Photinus pyralis*, codon adapted)	variable	*psbN-psbT*	*PpsbD 5'psbD lucCP 3'atpB*	Matsuo et al. 2006
luciferase (*Photinus pyralis*, codon adapted)	variable	*ORF2971-psbD*	*PtufA 5'tufA lucCP 3'atpB*	Matsuo et al. 2006
GFP (*Aequorea aequorea*)	0.006%	*psbA-5SrRNA*	*PrbcL 5'rbcL GFPncb 3'rbcL*	Franklin et al. 2002
GFP (*Aequorea aequorea*, codon adapted)	0.5%	*psbA-5SrRNA*	*PrbcL 5'rbcL GFPct 3'rbcL*	Franklin et al. 2002
other proteins				
RecA (*Escherichia coli*)	n.a	*atpB-IR*	*PatpA 5'atpA recA 3'rbcL*	Cerrutti et al. 1995
fusion of VP1 and cholera toxin B (FMDV and *Vibrio cholerae*)	3%	*chlL*	*PatpA 5'atpA CTBVP1 3'rbcL*	Sun et al. 2003
large single-chain antibody (*Homo sapiens*)	n.a	*psbA-5SrRNA*	*PrbcL 5'rbcL HSV8-lsc 3'rbcL*	Mayfield et al. 2003
large single-chain antibody (*Homo sapiens*)	n.a	*psbA-5SrRNA*	*PatpA 5'rbcL HSV8-lsc 3'rbcL*	Mayfield et al. 2003
allophycocyanin (*Spirulina maxima*)	2%	*chlL*	*PatpA 5'rbcL apcAapcB 3'rbcL*	Su et al. 2005

protein synthesis rates which are obtained after alteration of the AUG start codon can already cause severe phenotypes (Chen et al. 1993). Correspondingly, after mutation of the *clpP* initiation codon to AUU the degradation of the cytochrome b_6f complex was affected suggesting that ClpP is involved in quality control of this photosynthetic complex (Majeran et al. 2000). Most inactivated genes encode photosynthetic functions and, thus, are not essential for cell viability on acetate-containing medium (Table 3).

Several site-directed mutants for distinct amino acids in diverse photosynthetic subunits were generated which provides a very detailed view on the structure/function relationships in photosynthesis (for a review see: Xiong and Sayre 2004; Marin-Navarro and Moreno 2006).

3.4 Introduced genes, expressed proteins

Despite the extraordinary significance of the chloroplast transformation system in *Chlamydomonas reinhardtii* for elucidating scientific aspects, biotechnological applications were considered only relatively recently. Nevertheless, as compiled in Table 4, several foreign genes have now successfully been expressed in the algal chloroplast. Besides reporter genes like β-glucuronidase, luciferase, and green fluorescent protein (GFP), high-yield expression (3% of total soluble protein) of a fusion protein consisting of VP1 protein from the foot-and-mouth disease virus and cholera toxin B subunit has been achieved. Antigenicity was demonstrated suggesting that transplastomic *C. reinhardtii* cells might be a source for mucosal vaccines (Sun et al. 2003). In addition, a fully active human antibody directed against glycoprotein D of the herpes simplex virus was expressed in the alga (Mayfield et al. 2003) verifying that pharmaceutical proteins can be synthesized in *C. reinhardtii* chloroplasts. An enhancement of gene expression was observed after adaptation of codon-usage of foreign genes to the plastid codon usage. This appears to reflect an important aspect for future algal biotechnological applications (Franklin and Mayfield 2004).

3.5 Transformed species

Although recent years have seen substantial improvements in genetic engineering of the nuclear genomes of a variety of algae including several multicellular seaweeds like *Porphyra*, *Gracilaria*, *Ulva*, and *Laminaria* (Qin et al. 2005), to date, only three chloroplast genomes from algae have successfully been transformed. Besides *C. reinhardtii*, the chloroplasts of *Euglena gracilis* were transformed with an *aadA* cassette which contained *E. gracilis* expression control elements and shown to be resistant to spectinomycin (Doetsch et al. 2001). However, despite the presence of suitably-sized homologous flanking chloroplast DNA sequences, the transforming DNA was not stably integrated into the chloroplast genome but, instead, was inherited as an episomal element during continuous selection on antibiotics (Doetsch et al. 2001). Further work is required to elucidate the potential of this transformation system, which represents the first one for an alga containing complex chloroplasts, a feature that developed during secondary endosymbiosis (Delwiche 1999). Moreover, this system might pave the way for the genetic engineering of complex plastids from other algae of higher oecological and/or economical importance like diatoms or brown algae.

In contrast to *E. gracilis*, the unicellular red alga *Porphyridium spec*. containing primary chloroplasts can be stably transformed after integration of the transforming DNA into the chloroplast genome (Lapidot et al. 2002). Single crossover events have been observed after homologous recombination-mediated integration of a mutant AHAS gene conferring resistance to the herbicide SMM (see 3.2) into the chloroplast genome. However, homoplasmy was not reached under the applied experimental conditions leaving the question open whether transformants can be maintained under non-selective conditions. Interestingly, transformation rates were shown to significantly increase after synchronization of cell cultures in light/dark regimes and particle bombardment immediately after the dark phase (Lapidot et al. 2002). This procedure might be valuable also for other algal species, which have so far not been accessible to chloroplast transformation.

4 Plastid transformation in higher plants

4.1 Expression control elements

Quite a number of different regulatory elements have been tested for heterologous gene expression in plastids of higher plants (Table 5). Only very few of the elements are routinely used in plastid expression vectors (see also Table 8): the strong constitutive plastid 16S rRNA promoter in combination with the viral T7G10-5'-UTR (Staub et al. 2000; Kuroda and Maliga 2001b) or alternatively with a synthetic ribosomal binding site (rbs) consisting of the terminal 18 bp of the *rbcL*-5'-UTR (Svab and Maliga 1993). The light-regulated *psbA* control elements (promoter, 5'-UTR and 3'-UTR) are also frequently used (Staub and Maliga 1993; Fernandez-San Millan et al. 2003). These control elements have been shown to generally generate superior expression levels. Very high expression levels could also be obtained with the T7-system (promoter and 5'-UTR) relying on nuclear expressed and plastid imported T7-polymerase (McBride et al. 1994) or with operon extension vectors under the control of strong endogenous promoters (Staub and Maliga 1995a; Herz et al. 2005). Sometimes a T7-terminator was introduced in addition to a plastid 3'-UTR to ensure termination, when T7-polymerase was used to transcribe transplastomic genes (Magee et al. 2004b; Lössl et al. 2005).

As expression in plastids is predominantly controlled at the post-transcriptional level (Stern et al. 1997), the 5'-UTR is an important determinant of the expression level (Eibl et al. 1999). Another important feature is the N-terminal sequence of the gene of interest, which can be modified by fusion tags (Kuroda and Maliga 2001a; Herz et al. 2005).

A potential problem using control elements homologous to endogenous control elements is the risk of undesired recombination events (Svab and Maliga 1993). One such example was recently described for the *psbA*-3'-UTR (Rogalski et al. 2006). To avoid this potential problem some groups used plastid control elements from different species (Reddy et al. 2002; Zhou et al. 2006). However, homologous elements have frequently been used without reported recombination problems.

In some cases the 5'-UTR of the gene of interest was used as a ribosomal binding site and no extra 5'-UTR was included, especially when polycistronic operons have been introduced into the plastome (e.g. De Cosa et al. 2001; Madoka et al. 2002; Lössl et al. 2003).

Most 3'-UTRs do not terminate transcription, rather they merely act as processing and stabilising elements (Stern and Gruissem 1987). No substantial differences in the suitability of different 3'-UTRs for expression vectors have been reported (Eibl et al. 1999), so the 3'-UTR seems to be only of minor importance compared to promoter and 5'-UTR.

Table 5. Regulatory elements used in higher plant plastid transformants[a].

Regulatory element	Reference
promoters	
16S rRNA	Svab and Maliga 1993
psbA	Staub and Maliga 1993
T7G10[b,c]	McBride et al. 1994
clpP	Sriraman et al. 1998
trc[b]	Newell et al. 2003
rbcL	Herz et al. 2005
PHS[b.d]	Buhot et al. 2006
atpI	Wurbs et al. 2007
5'-untranslated regions	
rbcL (rbs)	Svab and Maliga 1993
psbA	Staub and Maliga 1993
T7G10[b]	Staub et al. 2000
atpB	Kuroda and Maliga 2002
clpP	Kuroda and Maliga 2002
rpl22	Herz et al. 2005
psbC	Herz et al. 2005
psaB	Herz et al. 2005
IREScp148[b]	Herz et al. 2005
atpI	Wurbs et al. 2007
3'-untranslated regions	
psbA	Staub and Maliga 1993
rps16	Zoubenko et al. 1994
rbcL	Eibl et al. 1999
rpl32	Eibl et al. 1999
rrnB	Newell et al. 2003
Ta[b]	Buhot et al. 2006

[a] Expression control elements were used in various combinations.
[b] Regulatory elements not of plastid origin: *trc* (*E. coli),* PHS (*E. coli groE* heat shock promoter), T7G10 (phage T7 gene 10 promoter), IREScp148 (internal ribosome entry site of the coat protein of a crucifer-infecting tobamovirus), *Ta* (*E. coli* threonine attenuator).
[c] T7-RNA polymerase needed.
[d] Chimeric transcription factor needed.

4.2 Inducible gene expression

A number of reasons make inducible gene expression in plastids highly desirable. If an economically feasible pre- or post-harvest induction were available, metabolic drain during growth and development could be avoided. Furthermore, negative effects of gene product(s) or metabolic changes caused by novel gene products might be a problem, if expression were constitutive (Lössl et al. 2003; Herz et al. 2005; Chakrabarti et al. 2006). Finally, it would be very valuable for basic research, if plastid gene expression could be switched on and off at will and at desired time-points.

Expression of plastid genes is not primarily controlled at the transcriptional level through regulated promoters that supply differential gene expression in response to physiological, developmental, or tissue specificity parameters. Therefore, inducible expression in plastids cannot be achieved using endogenous plastid control elements. External control was first described using a plastid transgene under control of the phage T7 promoter in combination with T7 polymerase encoded by a nuclear transgene and imported into the organelle (McBride et al. 1994). Controlled expression is achieved to a certain extent (Magee et al. 2004a), and negative effects observed during constitutive expression of genes of interest (Lössl et al. 2003) were avoided, when the same genes were transcribed by an ethanol induced T7 polymerase (Lössl et al. 2005). The system is, however, not optimal. The T7 promoter is recognized in in vitro experiments by the nucleus encoded plastid RNA polymerase (Lerbs-Mache 1993). This would, if true also in vivo, lead to background expression in the non-induced state. Furthermore, expression of some plastid genes is altered in the presence of T7 polymerase even if the genes do not contain a T7 promoter (Magee and Kavanagh 2002), and the low level of expression typical for most nuclear inducible promoters in the absence of an inducer may be sufficient to cause an undesirable phenotype (Magee et al. 2004a, 2007). Buhot et al. (2006) reported using the eubacterial *E. coli* groE heat shock promoter, which is not recognized by the plastid transcription machineries. Controlled expression was achieved through transient expression from a nuclear expression cassette of a chimeric sigma factor that mediates the interaction of the plastid encoded plastid RNA polymerase (PEP) and the eubacterial promoter. It remains to be seen how the system performs if combined with an inducible nuclear promoter.

Yet another approach towards inducible gene expression in plastids is based on CRE recombinase-mediated excision of the selection marker gene leaving its AUG translation start codon behind (Tungsuchat et al. 2006). Thus, a gene of interest lacking an own start codon is brought into contact with the non-excised start codon of the excised marker gene. The advantage of the system lies in the fact that it is not sensitive to read-through transcription. Control is executed by generating a translatable open reading frame and GFP was used as the reporter protein. Prior to excision there is no detectable GFP, while accumulation of GFP is found to constitute up to 0.3% of the total cellular protein after excision. Again, a transgene expressed from the nucleus is required to trigger plastid expression: primary transplastomic lines harbouring an inactive gene of interest were transformed in a second step in their nuclear genome using *Agrobacterium*-mediated gene transfer. Once the activation has occurred it cannot be reversed, and it remains to be seen, how the approach can be adapted for practical purposes.

A direct induction system, which is independent of nuclear gene expression, is based on constitutive repression of a plastid transgene by the lac repressor and induction with isopropyl-ß-D-galactopyranoside (IPTG) (Mühlbauer and Koop 2005). Increase of the level of reporter protein (GFP) was about 20-fold. This system is also not optimal, since there is low-level expression in the non-induced state. It is, however, attractive, since post-harvest induction is possible (Mühlbauer and Koop 2005), avoiding spraying of IPTG in the open field, which might be ecologically undesirable.

All the approaches towards inducible plastid gene expression developed so far are useful for basic research and for lab-scale expression studies. Inducible expression for production-scale application remains a prominent challenge in plastid transformation technology.

4.3 Resistance marker genes and selection schemes

In comparison to nuclear transformation protocols the number of selection genes successfully used for plastid transformation is relatively small (Table 6). With one exception all the direct selection markers provide resistance to the aminoglycoside antibiotics spectinomycin, streptomycin, and kanamycin. These compounds inhibit protein synthesis by specifically binding to the organelle's prokaryotic 70S ribosomes. Pioneering work with tobacco transformation was achieved using plastid marker genes isolated from plants that were resistant to streptomycin and spectinomycin (Svab et al. 1990; Staub and Maliga 1992). Two specific point mutations in the *rrn16* gene (Spc^+ and Str^+) and one mutation in the *rps12* gene (Str^+) alter ribosome structure and prevent antibiotic binding. Similar gene sequences, cloned from the *Solanum nigrum* plastome, have been successfully used for transformation of tobacco (Kavanagh et al. 1999) and more recently tomato (Nugent et al. 2005). However, much higher efficiencies of transformation have been reported using dominant chimeric antibiotic resistance genes. The most universally used marker is the *aadA* gene, which detoxifies spectinomycin and streptomycin (Goldschmidt-Clermont 1991; Svab and Maliga 1993). Translational fusions between *aadA* and *gfp* (FLARE-S) have also been used to generate bifunctional proteins that can be used for visual tracking of the transformation process (Khan and Maliga 1999). Marker genes giving resistance to kanamycin, *nptII* (Carrer et al. 1993) and *aphA-6* (Huang et al. 2002) have also been described. A novel approach for cotton plastid transformation involved the simultaneous use of *nptII* and *aphA-6* to detoxify kanamycin. The double gene/single selection strategy was shown to be more efficient than using the *aphA-6* gene alone (Kumar et al. 2004b).

To date only one non-antibiotic resistance marker has been described for direct selection of plastid transformants, the *badh* gene from spinach (Daniell et al. 2001b). In tobacco, extraordinarily high transformation efficiencies were claimed using this gene in combination with the selection agent betaine aldehyde, which is

Table 6. Selection genes for higher plant plastid transformation.

Selection agent	Gene	Mutation, Enzyme	Reference[a]
direct selection			
spectinomycin	*rrn16*	point mutation in 16S rRNA	Svab et al. 1990
streptomycin	*rrn16*	point mutation in 16S rRNA	Svab et al. 1990
streptomycin	*rps12*	point mutation in rps12	Staub and Maliga 1992
spectinomycin and streptomycin	*aadA*	aminoglycoside 3' adenyltransferase	Svab and Maliga 1993
spectinomycin and streptomycin	*gfp* + *aadA*	green fluorescent protein fused with aminoglycoside 3' adenyltransferase (FLARE-S)	Khan and Maliga 1999
kanamycin	*nptII*	neomycin phosphotransferase II	Carrer et al. 1993
kanamycin	*aphA-6*	aminoglycoside phosphotransferase	Huang et al. 2002
betaine aldehyde	*badh*	betaine aldehyde dehydrogenase	Daniell et al. 2001b
secondary selection			
phosphinothricin (glyphosinate ammonium)	*bar*	phosphinothricin acetyltransferase	Iamtham and Day 2000
glyphosate	*epsps*	resistant form of 5-enolpyruvylshikimate-3-phosphate synthase	Ye et al. 2003
isoxaflutole	*hppd*	4-hydroxyphenylpyruvate dioxygenase	Dufourmantel et al. 2007
negative selection			
5-fluorocytosine	*codA*	cytosine deaminase	Serino and Maliga 1997

[a] Only the first publication on each marker is cited.

inactivated to non-toxic glycine betaine. It should be noted, however, that no further reports verifying the system have been published.

Secondary selection genes, while not suitable for direct selection, can be used to confer a selective advantage where a dominant population of transformed plastid chromosomes has first been established using antibiotic selection. Such markers are particularly useful for counter-selection strategies, which result in the removal of antibiotic resistance markers from transformed plants. Genes conferring resistance to the herbicides phosphinothricin/glyphosinate ammonium (Iamtham and Day 2000; Ye et al. 2003), glyphosate (Ye et al. 2003) or isoxaflutole (Dufourmantel et al. 2007) have all been used successfully in this way.

Bacterial cytosine deaminase (*codA*) has been shown to be a suitable negative selection marker for tobacco plastid transformation. Cytosine deaminase converts the selection agent 5-fluorocytosine to a toxic metabolite 5-fluorouracil and leads to cell death (Serino and Maliga 1997). Cells that do not express the enzyme grow normally when plated on 5-fluorocytosine. Corneille et al. (2001) later demonstrated the functionality of the negative selection system for monitoring the excision of *codA* using the CRE-*lox* recombination system.

Higher plant plastid transformation necessitates the development of selection systems to meet highly demanding criteria. Selective advantage must be generated on two levels, that of the plastid and that of the individual cell. A typical tobacco mesophyll cell contains as many as 100 plastids each with up to 100 plastome copies (see Table 2 for an overview of plastid biology). Although the precise mechanism of plastid transformation is unknown, it can be speculated that it is a rare event, perhaps initially only occurring as one transformed molecule within a single plastid. Appropriate selection conditions must be chosen to amplify the transformed molecules such that they become the dominant plastome type. The removal of all wild type plastomes can prove difficult and sometimes very time consuming. Conventionally this has been performed by making cycles of repeated regeneration from leaf explants on selection medium, such that cell division and organelle segregation ultimately lead to stable homoplasmic tissues (Svab and Maliga 1993). Dix and Kavanagh (1995) have described the possible benefit of using plastid genes carrying point mutations to speed up the process of selecting for homoplasmic transformants. Recessive-type markers as opposed to dominant selectable markers such as *aadA* do not cause localized detoxification of the selection agent, which could conceivably maintain heteroplastomy. However, much lower transformation frequencies are generally obtained using genes carrying point mutations compared to the dominant selection markers. A novel selection system was described by Klaus et al. (2003) to improve selection of transformants and also accelerate segregation towards homoplasmy. Firstly, homoplastomic pigment-deficient mutants were produced following site-specific deletion of photosynthesis-related genes using the *aadA* gene and spectinomycin selection (see section 4.4). These acceptor lines were propagated in vitro and used as an alternative to wild type plants for re-transformation using reconstitution vectors carrying *aphA-6* together with foreign sequences of interest. Transformants recovered after kanamycin selection had a wild type appearance due to complementation of the previously deleted plastome sequences and these regenerants could clearly be distinguished from untransfomed tissues. Surprisingly, PCR showed that the primary regenerants were already homoplasmic, suggesting that green tissues have a strong selective advantage over pigment deficient ones.

4.4 Targeted inactivation

Reverse genetic analysis is quite straightforward in tobacco due to the precise recombination system active within plastids. To date, 38 genes of the tobacco plastome have been inactivated to analyse or confirm their function (Table 7). Inactivation or deletion of plastid genes has been achieved by site-specific integration of a dominant marker (e.g. Burrows et al. 1998), replacement with a frame-shifted mutant (Horvath et al. 2000), or CRE/*lox* mediated excision (Kuroda and Maliga 2003). Recently, a deletion method based on the insertion of a direct repeat,

Table 7. Inactivated chloroplast genes in *Nicotiana tabacum*.

Gene	Inactivation status	Reference
RNA-polymerase		
rpoA	homoplasmic	Serino and Maliga 1998; De Santis-Maciossek et al. 1999; Klaus et al. 2003
rpoB	homoplasmic	Allison et al. 1996; De Santis-Maciossek et al. 1999
rpoC1	homoplasmic	Serino and Maliga 1998; De Santis-Maciossek et al. 1999
rpoC2	homoplasmic	Serino and Maliga 1998
tRNA		
$trnV_{GAC}$	homoplasmic	Corneille et al. 2001; Hajdukiewicz et al. 2001
photosystems		
psaJ	homoplasmic	Schöttler et al. 2007a
psbA	homoplasmic	Baena-Gonzales et al. 2003
psbE	homoplasmic	Swiatek et al. 2003a
psbF	homoplasmic	Swiatek et al. 2003a
psbI	homoplasmic	Schwenkert et al. 2006
psbJ	homoplasmic	Hager et al. 2002; Swiatek et al. 2003a
psbL	homoplasmic	Swiatek et al. 2003a
petA	homoplasmic	Monde et al. 2000; Klaus et al. 2003
petB	homoplasmic	Monde et al. 2000
petD	heteroplasmic	Monde et al. 2000
petL	homoplasmic	Fiebig et al. 2004; Schöttler et al. 2007b
ycf3	homoplasmic	Ruf et al. 1997; Klaus et al. 2003
ycf6 (petN)	homoplasmic	Hager et al. 1999
ycf9 (lhbA, psbZ)	heteroplasmic, homoplasmic	Mäenpää et al. 2000; Ruf et al. 2000; Baena-Gonzales et al. 2001; Swiatek et al. 2001
RUBISCO		
rbcL	homoplasmic	Kanevski and Maliga 1994; Kode et al. 2006
acetyl-CoA-carboxylase		
accD	heteroplasmic	Kode et al. 2005
NDH complex		
ndhA	heteroplasmic	Kofer et al. 1998b
ndhB	homoplasmic	Shikanai et al. 1998; Horvath et al. 2000
ndhC	homoplasmic, heteroplasmic	Burrows et al. 1998; Kofer et al. 1998b
ndhH	heteroplasmic	Kofer et al. 1998b
ndhF	homoplasmic	Martin et al. 2004
ndhI	heteroplasmic	Kofer et al. 1998b
ndhJ	homoplasmic	Burrows et al. 1998
ndhK	homoplasmic, heteroplasmic	Burrows et al. 1998; Kofer et al. 1998b
DNA replication		
oriA	homoplasmic	Mühlbauer et al. 2002
oriB	heteroplasmic	Mühlbauer et al. 2002

Gene	Inactivation status	Reference
RNA binding		
sprA	homoplasmic	Sugita et al. 1997
ribosomal proteins		
rps14	heteroplasmic	Ahlert et al. 2003
rps18	heteroplasmic	Rogalski et al. 2006
protease		
clpP1	heteroplasmic	Shikanai et al. 2001; Kuroda and Maliga 2003
hypothetical chloroplast open reading frames		
ycf1	heteroplasmic	Drescher et al. 2000
ycf2	heteroplasmic	Drescher et al. 2000
ycf10 (cemA)	homoplasmic	Swiatek et al. 2003b

Note: alternative gene names are given in brackets.

flanking the gene to be deleted and the selection marker was described (Kode et al. 2006). A subsequent loop-out recombination then eliminates the desired gene together with the selection marker.

Homoplasmic plant lines could be obtained, for most inactivated genes, allowing clear assignment of an observed phenotype. Although many of these mutants were defective or impaired in photosynthesis, the lines could be grown readily on sugar-containing media. However, in a few cases only heteroplasmic inactivation could be obtained suggesting an essential role of the gene even under heterotrophic conditions. These genes comprise *ycf1* and *ycf2* whose function is not yet clear (Drescher et al. 2000), the protease subunit gene *clpP1*, which is essential for shoot development (Kuroda and Maliga 2003) and the ß-carboxyl transferase subunit encoded by *accD*, which is required for fatty acid synthesis (Kode et al. 2005). Plastid ribosomal proteins (e.g. S14 and S18) seem to be essential for cell survival in tobacco, but not necessarily in all higher plants (Rogalski et al. 2006; Ahlert et al. 2003). The genes coding for plastidic NAD(P)H dehydrogenase seem to be dispensable under optimal growth conditions (Burrows et al. 1998; Kofer et al. 1998b; Horvath et al. 2000).

4.5 Introduced genes, expressed proteins

To date, a large number of heterologous genes have been expressed in plastids of higher plants including reporter proteins to monitor efficiency of regulatory elements, modified endogenous proteins, agronomic traits like herbicide resistance, insect resistance, pathogen resistance, output traits such as pharmaceutical proteins, vaccines or bioplastics, and a diverse group of heterologous enzymes (Table 8). The absence of a glycosylation system and the prokaryotic nature of the plastid expression system make the plastid compartment an unsuitable system for some proteins, whereas many others have been successfully expressed. The reported expression levels range from 0.001 to over 40% of the total soluble protein (TSP).

Very high expression levels (> 10% TSP) seem in some cases to delay plant development or result in a chlorotic phenotype (Tregoning et al. 2003; Chakrabarti et al. 2006). Given the differences in methods of quantification, the reported levels of expression need to be interpreted with some care. Most of the reported expression levels are maximum values, which were obtained under optimal conditions. Stable proteins such as GUS accumulate *in planta* such that the highest levels are found in mature plants (Herz et al. 2005), whereas proteins more susceptible to degradation like interferon (Leelavathi and Reddy 2003) or VP6 (Birch-Machin et al. 2004) occur at higher levels in young leaves. Depending on the regulatory elements, light conditions also influence the expression level (Fernandez-San Millan et al. 2003; Watson et al. 2004; Herz et al. 2005; Wirth et al. 2006). In general, the expression level in plastids is higher than with conventional nuclear expression in plants, but lower than the levels obtained with recent transient expression technology (Gleba et al. 2005). However, it should be clear that no expression system is universally suitable for every protein. The characteristics of the protein of interest have to fit with the chosen expression system. Unfortunately, this cannot be predicted in advance, and needs to be tested experimentally. As such there are also examples for proteins, which could not be expressed in plastids like haemoglobin (Magee et al. 2004b), ß-zein (Bellucci et al. 2005), or haemagglutinin (Lelivelt et al. 2005).

Almost all proteins were expressed in tobacco plastids except GUS (tobacco and petunia), neomycin phosphotransferase (tobacco, cotton), GFP (tobacco, potato, lettuce, poplar and rice), AAD-GFP (tobacco, rice and *Lesquerella*), HPPD (tobacco and soybean), Bt-toxin (tobacco, oilseed rape and soybean), BADH (tobacco and carrot), lycopene-ß-cyclase (tobacco and tomato), and haemagglutinin (lettuce). See Table 9 for additional information.

Whereas most expression studies in plastids rely on the endogenous PEP/NEP polymerases, there is also the possibility to use an orthologous polymerase such as the T7-polymerase to achieve transcription in plastids. Expression of a plastid-localised *uidA* gene by the aid of a nuclear expressed and plastid-targeted T7-polymerase resulted in very high transcript and protein levels (McBride et al. 1994). High transcript levels do, however, not necessarily result in high levels of translated protein (Magee et al. 2004a, 2004b). There is growing evidence that correct folding and proteolytic stability of the target protein are more important determinants of the expression level than transcription and translation efficiency (Birch-Machin et al. 2004). When GUS was fused to the N-terminus of interferon-γ the expression level increased from 0.1 to 6% and the half-life of the fusion protein increased from 6 to 48 hours compared to the unmodified interferon-γ although both versions were under the control of identical regulatory elements (Leelavathi and Reddy 2003). Similar results were obtained with recombinant epidermal growth factor (Wirth et al. 2006).

Unlike in many other expression systems, codon usage plays only a minor role in the plastid expression system of *N. tabacum*, probably because of the relatively balanced codon frequency (Maliga 2003). Nevertheless, heterologous gene expression was modestly increased (up to 2.5-fold), if the codon usage was adjusted to the relatively AT-rich plastid genome of tobacco (Ye et al. 2001; Tregoning et al. 2003). On the other hand, at least in vitro translation efficiencies do not always correlate with codon usage (Nakamura and Sugiura 2007). Although mRNA editing occurs in resident plastome genes, no editing of heterologous genes has ever been observed.

Staub and co-workers (2000) established an elegant expression system for mature somatotropin in plastids by fusing the mature somatotropin domain to an ubiquitin domain, which is only processed to mature protein by endogenous cytosolic ubiquitin-protease during the extraction procedure but not in the intact plastid. However, one additional amino acid was removed from the N-terminus in most of the processed somatotropin. This could arise from incorrect processing by cytosolic ubiquitin-protease or from a secondary protease activity. In fact, most endogenous proteins expressed in plastids are processed post-translationally by methionine-aminopeptidase and/or peptide-deformylase (Giglione and Meinnel 2001). In the case of the RUBISCO large subunit even two N-terminal amino acids are removed post-translationally (Houtz et al. 1989). Currently, little is known about post-translational modifications of recombinant proteins in plastids. Analysis of recombinant hydroxyphenyl-pyruvate dioxygenase (HPPD) in plastids showed that the starting methionine was cleaved off, but no further modifications were detected (Dufourmantel et al. 2007). However, when tetanus toxin (TetC) was expressed in tobacco plastids the initiator methionine was not removed post-translationally, but around half of the TetC was expressed as a slightly larger, modified protein (Tregoning et al. 2003). Comparative analysis of mature amino-terminal sequences of twelve recombinant proteins expressed in chloroplasts suggests that recombinant proteins comply with the N-terminal processing rules proposed for endogenous plastid proteins (Fernandez-San Millan et al. 2007).

Recently lipidation and functional activity of a recombinant bacterial lipoprotein expressed in tobacco chloroplasts was reported (Glenz et al. 2006). The protein was only lipidated when the appropriate signal sequence was present. This is also a prerequisite for lipidation in bacteria and cyanobacteria. The main fraction of the protein was lipidated but unlipidated protein and lipoprotein variants were also present. Another important aspect is the correct formation of disulfide bonds, which can be achieved in the cytosol of prokaryotic hosts like *E. coli* only in specially modified strains (Bessette et al. 1999). It was shown that all disulfide bonds of somatotropin where formed correctly inside plastids (Staub et al. 2000), making it a suitable host for disulfide-containing proteins.

To date most recombinant proteins have been extracted from green leaves, but in some plant species other organs like seeds, fruits, or tubers present attractive sources for protein extraction, because of advantages in transportation and storage. However, expression in chloroplasts seems to be much higher compared to other plastid types, such as amyloplasts or chromoplasts. Expression of an AAD-GFP fusion protein (FLARE-S) was detected in non-green tissues including petals and roots of transplastomic tobacco (Khan and Maliga 1999). However, the expression level of GFP in potato tubers was only 0.05% TSP compared to 5% TSP in green tissues (Sidorov et al. 1999). Kumar et al. (2004a), on the other hand, report only a minor decrease of BADH-expression in carrot roots compared to carrot leaves. In transplastomic tomato fruits the expression level of the *aadA* selection marker under control of the constitutive 16S-promoter was half as high as in the green leaves (Ruf et al. 2001). High expression of recombinant HPPD under control of the light-regulated *psbA* promoter and 5'-UTR was reported in transplastomic tobacco leaves, but also at a lower level in seeds and petals, whereas no expression was detectable in roots (Dufourmantel et al. 2007). In soybean expression of *Bt*-toxin was detected in leaves, stems and seeds but not in root tissue (Dufourmantel et al. 2005).

Recombinant HPPD (4-hydroxyphenylpyruvate dioxygenase) in transplastomic tobacco and soybean provided improved tolerance to the herbicide isoxaflutole compared to nuclear transgenic plants (Dufourmantel et al. 2007). But in the case of EPSPS (5-enolpyrovylshikimate-3-phosphate synthase) expression, transplastomic lines showed no higher resistance to the herbicide glyphosate than nuclear transformants, despite much lower expression levels of EPSPS in the nuclear transformants (Ye et al. 2001). The reason for the different resistance levels might be the alternative mode of action of glyphosate (inhibitor of aromatic amino acid biosynthesis) and isoxaflutole (inhibitor of tocopherol- and plastoquinone-biosynthesis). Glyphosate is toxic for all cell types whereas isoxaflutole is only toxic to photosynthetic cells. Thus, plastid expression of HPPD is particularly well suited since only expression in the chloroplast is needed whereas plastid expression of recombinant proteins in non-green tissues is generally much lower, limiting the efficiency of EPSPS in these cells (Dufourmantel et al. 2007). Plastidic expression of PAT (phosphinothricin acetyltransferase) resulted in high tolerance to the herbicide phosphinothricin, an inhibitor of glutamine biosynthesis (Lutz et al. 2001; Kang et al. 2003b).

Besides herbicide resistance, another promising area for transplastomic plants is metabolic engineering. The expression of chorismate pyruvate lyase in plastids yields p-hydroxybenzoic acid, which is a precursor for liquid crystal polymers (Viitanen et al. 2004). Recently, the ß-carotene level in transplastomic tomato fruits was shown to be increased by expression of bacterial lycopene-ß-cyclase, which converts lycopene into ß-carotene (Wurbs et al. 2007). Lycopene-ß-cyclase from the fungus *Phycomyces blakesleeanus* could not be expressed successfully due to mRNA instability (Wurbs et al. 2007).

Plastid-localised expression of the *phb*-operon from *Ralstonia eutropha* has also been described (Lössl et al. 2003; Arai et al. 2004; Lössl et al. 2005). The *phb*-operon encodes ß-ketothiolase, acetyl-CoA reductase and PHB synthase. These enzymes catalyse the synthesis of polyhydroxybutyrate, which is a biodegradable plastic, from the plastidic precursor acetyl-coenzyme A. The expression of functional polycistronic operons is a major advantage of plastid transformation over other transformation methods in plants. However, change of metabolic flux or product toxicity may enforce regulation of the genes or pathways that are introduced (Lössl et al. 2005).

The expression of the bacterial *cry*-operon comprising ORF1, ORF2, and *cry2Aa2* is another example for the expression of a large polycistronic operon in plastids (De Cosa et al. 2001). ORF2 supports crystallisation of the *Bt*-toxin leading to the formation of Bt-crystals within the plastids. However, the quoted expression level of 46% total soluble protein is somewhat misleading as extracts

Table 8. Proteins expressed in plastids of higher plants.

Protein	Expression	Insertion site	Expression construct	Reference
reporter proteins				
ß-glucuronidase	2.5%	*trnV-16S*	*PpsbA 5'psbA uidA 3'psbA*	Staub and Maliga 1993
ß-glucuronidase	0.5%	*trnN-trnR*	*Prrn 5'T7G10 uidA aadA 3' rpl32*	Herz et al. 2005
ß-glucuronidase	3,7%	*trnN-trnR*	*Prrn 5'T7G10 5AAsyn-uidA aadA 3' rpl32*	Herz et al. 2005
ß-glucuronidase	1.5%	*trnS-orf74*	*Prrn 5'T7G10 5AA-uidA aadA 3' rpl32*	Herz et al. 2005
ß-glucuronidase	3.8%	*rps12-orf131*	*Prrn 5'T7G10 5AAsyn-uidA aadA 3'rpl32*	Herz et al. 2005
ß-glucuronidase	10.8%	*psbA-trnH*	*OpsbA 5'T7G10 5AAsyn-uidA aadA 3'rpl32*	Herz et al. 2005
ß-glucuronidase	20-30%	*rps12-trnV*	*PT7G10 5'T7G10 uidA 3'psbA*	McBride et al. 1994
neomycin phos. transf.	1.0%	*rbcL-accD*	*Prrn 5'rbcL 5AArbcl-neo 3'psbA*	Carrer et al. 1993
neomycin phos. transf.	0.3%	*rps12-trnV*	*Prrn 5'clpP neo 3'rbcL*	Kuroda and Maliga 2002
neomycin phos. transf.	0.8%	*rps12-trnV*	*Prrn 5'atpB neo 3'rbcL*	Kuroda and Maliga 2002
neomycin phos. transf.	7%	*rps12-trnV*	*Prrn 5'atpB 14AAatpB-neo 3'rbcL*	Kuroda and Maliga 2001a
neomycin phos. transf.	10.8%	*rps12-trnV*	*Prrn 5'rbcL 14AArbcl-neo 3'rbcL*	Kuroda and Maliga 2001a
neomycin phos. transf.	0.16%	*rps12-trnV*	*Prrn 5'T7G10 10AApts-neo 3'rbcL*	Kuroda and Maliga 2001b
neomycin phos. transf.	16.4%	*rps12-trnV*	*Prrn 5'T7G10 10AAT7G10-neo 3'rbcL*	Kuroda and Maliga 2001b
neomycin phos. transf.	23%	*rps12-trnV*	*Prrn 5'T7G10 3AAsyn-neo 3'rbcL*	Kuroda and Maliga 2001b
GFP	5%	*rps12-trnV* potato	*Prrn 5'rbs gfp 3'rps16*	Sidorov et al. 1999
GFP	5.5%	*rbcL-accD*	*Prrn 5'rbs gfp 3'rrnB*	Newell et al. 2003
GFP	36%	*rbcL-accD* lettuce	*PpsbA 5'psbA gfp 3'rps16*	Kanamoto et al. 2006
GFP	n.a.[c]	*rbcL-accD* poplar	*PpsbA 5'psbA gfp 3'rps16*	Okumura et al. 2006

Protein	Expression	Insertion site	Expression construct	Reference
AAD-GFP fusion protein (FLARE-S)	8%	*rps12-trnV*	*Prrn 5'atpB 14AAatpB-aadA-gfp 3'psbA*	Khan and Maliga 1999
AAD-GFP fusion protein (FLARE-S)	18%	*rps12-trnV*	*Prrn 5'rbcL 14AArbcl-aadA-gfp 3'psbA*	Khan and Maliga 1999
CTB-GFP fusion protein	21%	*trnI-trnA*	*Prrn 5'rbs aadA 5'psbA ctb-gfp 3'psbA*	Limaye et al. 2006
eYFP	n.a.[c]	*rps12-trnV*	*Pphs 5'rbs eyfp 3'ta*	Buhot et al. 2006
plastid proteins				
acetyl-CoA carboxylase	17-63 pmol / min mg	*accD*	*Prrn 5'accD accD 3'accD*	Madoka et al. 2002
RUBISCO (large subunit)	wild type level	*rbcL-replacement*	*PrbcL 5'rbcL rbcL-histag 3'rbcL*	Rumeau et al. 2004
RUBISCO (small subunit)	wild type level	*trnI-trnA*	*PpsbA 5'psbA rbcS 3'psbA*	Dhingra et al. 2004
RUBISCO (bacterial [a])	1/3 wild type level	*rbcL-replacement*	*PrbcL 5'rbcL rbcM aadA 3'rps16*	Whitney and Andrews 2001
herbicide resistance				
EPSPS	n.a.[c]	*rbcL-accD*	*Prrn 5'rbs aadA epsps 3'psbA*	Daniell et al. 1998
EPSPS	0.001%	*rps12-trnV*	*Prrn 5'rbcL CP4bact 3'rps16*	Ye et al. 2001
EPSPS	0.002%	*rps12-trnV*	*Prrn 5'rbcL CP4syn 3'rps16*	Ye et al. 2001
EPSPS	0.2%	*rps12-trnV*	*Prrn 5'T7G10 CP4bact 3'rps16*	Ye et al. 2001
EPSPS	0.3%	*rps12-trnV*	*Prrn 5'T7G10 CP4syn 3'rps16*	Ye et al. 2001
EPSPS	10%	*rps12-trnV*	*Prrn 5'T7G10 14AAgfp-CP4syn 3'rps16*	Ye et al. 2001
PAT	7%	*rps12-trnV*	*Prrn 5'atpB 14AAatpB-bar 3'rbcL*	Lutz et al. 2001
PAT	n.a.[c]	*trnI-trnA*	*Prrn 5'rbs aadA bar 3'psbA*	Kang et al. 2003b
HPPD	n.a.[c]	*rps12-orf131*	*Prrn 5'rbs hpd 3'rbcL*	Falk et al. 2005
HPPD	5%	*rbcL-accD*	*PpsbA 5'psbA hppd 3'rbcL*	Dufourmantel et al. 2007
HPPD	5%	*rps12-trnV* soybean	*Prrn 5'T7G10 hppd 3'rbcL*	Dufourmantel et al. 2007

Protein	Expression	Insertion site	Expression construct	Reference
insect resistance				
Bt toxin	3%	*rbcL-accD*	*PpsbA 5'psbA(rice) cry1Ia5 3'psbA(rice)*	Reddy et al. 2002
Bt toxin	n.a.[c]	*rps12-trnV soybean*	*Prrn 5'T7G10 cry1Ab 3'rbcL*	Dufourmantel et al. 2005
Bt toxin	n.a.[c]	*rps7-ndhB* oilseed rape	*Prrn 5'rbs cry1Aa10 3'psbA(rice)*	Hou et al. 2003
Bt toxin	3-5%	*rps12-trnV*	*Prrn 5'rbcL cry1Ac 3'rps16*	McBride et al. 1995
Bt toxin	2-3%	*rbcL-accD*	*Prrn 5'rbs aadA cry2Aa2 3'psbA*	Kota et al. 1999
Bt toxin	46.1%[b]	*trnI-trnA*	*Prrn 5'rbs aadA ORF1 ORF2 cry2Aa2 3'psbA*	De Cosa et al. 2001
Bt toxin	10%	*trnI-trnA*	*Orrn 5'cry cry9Aa2 3'rbcL*	Chakrabarti et al. 2006
pathogen resistance				
MSI-99	n.a.[c]	*rps12-trnV*	*Prrn 5'rbs msi99 aadA 3'psbA*	DeGray et al. 2001
pharmaceutical proteins				
somatotropin	0.2%	*rps12-trnV*	*PpsbA 5'psbA hgh 3'rps16*	Staub et al. 2000
somatotropin	1%	*rps12-trnV*	*PpsbA 5'psbA ubq-hgh 3'rps16*	Staub et al. 2000
somatotropin	7%	*rps12-trnV*	*Prrn 5'T7G10 ubq-hgh 3'rps16*	Staub et al. 2000
HSA	0.02%[b]	*trnI-trnA*	*Prrn 5'rbs aadA hsa 3'psbA*	Fernandez-San Millan et al. 2003
HSA	11.1%[b]	*trnI-trnA*	*PpsbA 5'psbA hsa 3'psbA*	Fernandez-San Millan et al. 2003
insulin like growth factor	33%[d]	*trnI-trnA*	*PpsbA 5'psbA igf 3'psbA*	Daniell et al. 2005a
interferon α5	n.a.	*trnI-trnA*	*PpsbA 5'psbA ifnA5 3'psbA*	Daniell et al. 2005a
interferon α2b	18%[d]	*trnI-trnA*	*PpsbA 5'psbA ifnA2b 3'psbA*	Daniell et al. 2005a
interferon-γ	0.1%	*rbcL-accD*	*PpsbA 5'psbA ifnG 3'psbA*	Leelavathi and Reddy 2003
interferon-γ	6%	*rbcL-accD*	*PpsbA 5'psbA histag-uidA-ifnG 3'psbA*	Leelavathi and Reddy 2003
haemoglobin	n.d.	*rbcL-accD*	*PT7G10 5'T7G10 hba hbb 3'rps16-T7G10*	Magee et al. 2004b

Protein	Expression	Insertion site	Expression construct	Reference
Guy's 13 antibody	n.a.	*trnI-trnA*	*Prrn 5'rbs igA-G 3'psbA*	Daniell et al. 2005a
single-chain camel antibody fragment	low level	*rps12-trnV*	*PT7G10 5'T7G10 abl 3'rps16-T7G10*	Magee et al. 2004a
epidermal growth factor	n.d.	*16S-trnI*	*PpsbA 5'psbA hegf 3'rps16*	Wirth et al. 2006
epidermal growth factor	low level	*16S-trnI*	*PpsbA 5'psbA 186AAuidA-hegf 3'rps16*	Wirth et al. 2006
vaccines				
TetC (tetanus)	10%	*rps12-trnV*	*Prrn 5'atpB tetC(bact) 3'rbcL*	Tregoning et al. 2003
TetC (tetanus)	25%	*rps12-trnV*	*Prrn 5'T7G10 tetC(bact) 3'rbcL*	Tregoning et al. 2003
TetC (tetanus)	10%	*rps12-trnV*	*Prrn 5'T7G10 tetC(syn) 3'rbcL*	Tregoning et al. 2003
LT-B (enterotoxigenic *E. coli*)	2.5%	*trnI-trnA*	*Prrn 5'rbs aadA ltb 3'psbA*	Kang et al. 2003a
LTK63 (enterotoxigenic *E. coli*)	3.7%	*trnI-trnA*	*Prrn 5'rbs aadA ltk63 3'psbA*	Kang et al. 2004
CT-B (cholera)	4.1%	*trnI-trnA*	*Prrn 5'rbs aadA ctb 3'psbA*	Daniell et al. 2001c
VP6 (rotavirus)	3%	*rbcL-accD*	*Prrn 5'rbs vp6 3'rrnB*	Birch-Machin et al. 2004
VP6 (rotavirus)	0.6%	*rbcL-accD*	*PpsbA 5'psbA vp6 3'rrnB*	Birch-Machin et al. 2004
2L21 peptide (virulent canine parvovirus)	31%	*trnI-trnA*	*PpsbA 5'psbA ctb-2l21 3'psbA*	Molina et al. 2004
2L21 peptide (virulent canine parvovirus)	23%	*trnI-trnA*	*PpsbA 5'psbA gfp-2l21 3'psbA*	Molina et al. 2004
PA (anthrax)	18%	*trnI-trnA*	*PpsbA 5'psbA pag 3'psbA*	Watson et al. 2004
F1-V (plague)	14.8%	*trnI-trnA*	*PpsbA 5'psbA caF1-lcrV 3'psbA*	Daniell et al. 2005a
Haemagglutinin (influenza)	n.d.	*trnI-trnA lettuce*	*Prrn(lettuce) 5'rbs aadA ha 3'psbA(lettuce)*	Lelivelt et al. 2005
VP1 (foot and mouth disease)	2-3%	*trnK-psbA*	*Prrn 5'rbs vp1 3'psbA*	Li et al. 2006a
lipoprotein A (lyme disease)	1%	*rbcL-accD*	*PpsbA 5'psbA ospA-histag 3'psbA*	Glenz et al. 2006
lipoprotein A (lyme disease)	10%	*rbcL-accD*	*PpsbA 5'psbA ospA-histag (without signal sequence) 3'psbA*	Glenz et al. 2006

Protein	Expression	Insertion site	Expression construct	Reference
NS3 (hepatitis C)	2%[d]	*n.a.*	*PpsbA 5'psbA ns3 3'psbA*	Daniell 2006
ORF2 fragment (hepatitis E)	0.1%	*trnM-trnG*	*PpsbA(rice) 5'psbA (rice) e2 3'psbA(rice)*	Zhou et al. 2006
VCA (Epstein-Barr virus)	0.004%	*rbcL-accD*	*PpsbA(rice) 5'psbA (rice) vca 3'psbA (rice)*	Lee et al. 2006a
spike protein subunit (SARS)	0.2%	*rbcL-accD*	*PpsbA 5'psbA histag-s1 3'psbA*	Li et al. 2006b
LecA surface antigen (amebiasis)	7%[d]	*n.a.*	*PpsbA 5'psbA lecA 3'psbA*	Daniell 2006
enzymes				
mercuric ion reductase; organomercurial lyase	n.a.[c]	*trnI-trnA*	*Prrn 5'rbs aadA merA merB 3'psbA*	Ruiz et al. 2003
xylanase	6%	*rbcL-accD*	*PpsbA 5'psbA(rice) xynA 3'psbA(rice)*	Leelavathi et al. 2003
chorismate pyruvate lyase (CPL)	35%	*trnI-trnA*	*PpsbA 5'psbA ubiC 3'psbA*	Viitanen et al. 2004
betaine aldehyde dehydrogenase	9 nmol/min mg	*trnI-trnA*	*Prrn 5'rbs aadA badh 3'psbA*	Daniell et al. 2001b
betaine aldehyde dehydrogenase	10-13 nmol/min mg	*trnI-trnA carrot*	*Prrn 5'rbs aadA 5'T7G10 badh 3'rps16*	Kumar et al 2004a
ß-ketothiolase	14.7 units/mg	*trnI-trnA*	*PpsbA 5'psbA phaA 3'psbA*	Ruiz and Daniell 2005
trehalose-6-phosphate synthase	5 µmol / min mg	*trnI-trnA*	*Prrn 5'rbs aadA tps1 3'psbA*	Lee et al. 2003
anthranilate synthase (α-subunit)	n.a.[c]	*rpl32-trnL*	*Prrn 5'rbs asa2 3'rpl32*	Zhang et al. 2001a
lycopene-ß-cyclase	n.a.[c]	*trnfM-trnG* (tomato and tobacco)	*PatpI 5'atpI crtY 3'rps16*	Wurbs et al. 2007
lycopene-ß-cyclase	n.d.	*trnfM-trnG* (tomato and tobacco)	*PatpI 5'atpI carRA 3'rps16*	Wurbs et al. 2007
bio-plastics				
PBP (GVGVP)	n.a.	*trnI-trnA rbcL-accD*	*Prrn 5'rbs aadA eg121 3'psbA*	Guda et al. 2000
PHB operon	10-160 ppm PHB	*rbcL-accD*	*Prrn 5'rbs aadA phbC phbA phbB 3'psbA*	Arai et al. 2004
PHB operon	1383 ppm PHB	*trnN-trnR*	*PT7G10 5'T7G10 phbC phbA phbB 3'phbB-rbcL-T7*	Lössl et al. 2005

Protein	Expression	Insertion site	Expression construct	Reference
storage protein				
ß-zein	n.d.	*trnM-trnG*	*Prrn 5'rbs g2 3'rbcL*	Bellucci et al. 2005

Note: Not all expression construct variants could be included. The expression data refers to percentage of total soluble protein (TSP). When this data was not available, enzyme activity or amount of end product (in ppm) is shown, n.a. data not available, n.d. no expression detectable.
The insertion-site shows the endogenous genes between which the insertion of the expression-cassette takes place. If not stated otherwise, all genes were inserted into the tobacco plastome. The expression cassette designates the promoter (P) or operon extension (O); 5'-UTR (5') (rbs, synthetic ribosomal binding site derived from the *rbcL* 5'-UTR); gene(s) present in the cistron, components of fusion-proteins are connected by hyphens, if truncated versions were used the number of amino acids (AA) is indicated (syn, synthetic sequence; pts, synthetic plastid downstream sequence; or name of the ORF); and 3'-UTR (3') at the end of the cistron. If not stated otherwise, all control elements are from tobacco.
[a] Homodimeric rbcM from *R. rubrum*, for an overview of additional RUBISCO-variants expressed in tobacco plastids see Andrews and Whitney (2003); [b] Quantification is based on protein solubilised in 50 mM NaOH; [c] Only biological activity determined; [d] No details given.

solubilised in 50 mM NaOH were used for quantification. The precise recombination mechanism in plastids allows exact modification of endogenous proteins. Replacement of endogenous RUBISCO (large subunit) by a RUBISCO-protein containing a C-terminal HisTag did not alter RUBISCO expression levels, which is the most abundant plant protein with 30-65% TSP. But the transplastomic plants accumulated high amounts of zinc, due to the presence of the HisTag (Rumeau et al. 2004). RUBISCO was also the target of more intensive modifications, e.g., replacement of plant *rbcL* by bacterial *rbcM* (Whitney and Andrews 2001). Description of the various modifications would exceed the scope of this article and they are excellently reviewed in Andrews and Whitney (2003).

Many of the enzymes listed in Table 8 are associated with beneficial agronomic traits: trehalose-6-phosphate synthase conferring drought tolerance (Lee et al. 2003), ß-ketothiolase conferring male sterility (Ruiz and Daniell 2005), betaine-aldehyde dehydrogenase (BADH) conferring salt tolerance (Kumar et al. 2004a), mercuric ion reductase resp. organomercurial lyase enabling phytoremediation (Ruiz et al. 2003).

Vaccines are the most prevalent class of pharmaceutical proteins expressed in plastids of higher plants. To date 14 different vaccines have been expressed in tobacco plastids and all extracted and analysed proteins have shown immune response in animals. It has often been proposed to use plant-made vaccines directly as edible vaccines, taking advantage of cheap production cost and easy application (Tregoning et al. 2004; Daniell et al. 2005a; Daniell 2006). However, edible vaccines would have to face the high standards of pharmaceutical production and potential risk of amalgamation with food plants (Fox 2006). It is, therefore, more likely that for human vaccines, the proteins would be extracted, purified and formulated as with existing production procedures. Nevertheless with an increasing world population the need for cheap vaccine production also increases, making plastid expression systems an attractive alternative.

4.6 Transformed species

Plastid transformation technology for tobacco was first described over 15 years ago (Svab et al. 1990). However, despite numerous additional publications describing improvements in the efficiency of tobacco transformation the transfer of the technology to other plants has proven relatively difficult. Table 9 summarizes the current status of higher plant plastid transformation. It should be stressed that in addition to tobacco, fertile homoplasmic plants have only been described for *N. plumbaginifolia*, tomato, soybean, *Lesquerella*, cotton, petunia, and lettuce. Furthermore, with four exceptions (potato, tomato, soybean, and lettuce) the remaining species are all documented as single publications only. As such there is no great depth of knowledge in the field regarding reproducibility and potential for improvement. The favoured method for transformation has been particle bombardment using explants as target tissue (e.g. leaves, callus, or suspension cells). Various efficiencies have been observed using this approach, as many as 40 leaf bombardments were needed to obtain a single plastid transformant from *Arabidopsis* (Sikdar et al. 1998), whereas in soybean (Dufourmantel et al. 2004) and, recently, in tomato (Wurbs et al. 2007) one transformant per shot or better have been described.

Less widely used is PEG-mediated plastid transformation of protoplasts. Some success has been reported in *N. plumbaginifolia* (O'Neill et al. 1993), tomato (Nugent et al. 2005), lettuce (Lelivelt et al. 2005), cauliflower (Nugent et al. 2006), and the moss *Physcomitrella* (Sugiura and Sugita 2004). The difficulties in isolating and culturing protoplasts and obtaining good plating efficiencies after treatment with PEG, are most probably a major restricting factor using this approach.

Direct comparison of transformation efficiencies between species is inappropriate, as vector constructs are rarely identical and different selection systems were utilized. However, it is readily apparent that even the best efficiencies reported are generally much lower than those typically obtained in tobacco, where bombardment can yield one to fourteen transformants per shot with leaves (Svab and Maliga 1993; Daniell et al. 2001c), four or more events per plate of bombarded cell suspension cells (Langbecker et al. 2004), and three to 47 transformants can be obtained for every million protoplasts treated with PEG (Koop et al. 1996). There are several reasons given in the literature for the lower transformation efficiencies observed in non-tobacco species, including, reduced activity of plastid homologous recombination (Sikdar et al. 1998), a focus on green tissues containing fully developed chloroplasts (Bogorad 2000), and use of heterologous elements for vector construction (Skarjinskaia et al. 2003). While the influence of these factors cannot be excluded there is no collective evidence that any are limiting progress in the field. The critical components for success are more likely rapid transformation protocols allowing for the efficient treatment of large numbers of cells or explants,

Table 9. Development of plastid transformation systems for higher plants.

Species[a]	System[b]	Genes[c]	Efficiency[d]	Status	Reference
Nicotiana tabacum (tobacco)	PG, leaves	*Nicotiana tabacum* rrn16 (Spc^+/Str^+)	3 lines from 148 shots	homoplasmic T_0 plants and T_1 progeny	Svab et al. 1990
	PG, leaves	aadA	84 lines from 79 shots	homoplasmic T_0 plants and T_1 progeny	Svab and Maliga 1993
	PEG, ppts	*Nicotiana tabacum* rrn16 (Spc^+/Str^+)	5 lines from 1.0 x 10^6 treated ppts	homoplasmic T_0 plants, T_1 not described	Golds et al. 1993
	PEG, ppts	aadA	118 lines from 6.0 x 10^6 treated ppts (best 47 from 1.0 x 10^6)	homoplasmic T_0 plants and T_1 progeny	Koop et al. 1996
	PG, cell suspension	aadA, gfp	best > 4 per shot[e]	homoplasmic T_0	Langbecker et al. 2004
Nicotiana plumbaginifolia (tex mex tobacco)	PEG, ppts	*Nicotiana tabacum* 16S rRNA (Spc^+/Str^+)	2 lines from 10^6 treated ppts	homoplasmic T_0 plant and T_1 progeny	O'Neill et al. 1993
Arabidopsis thaliana (mouse ear cress)	PG, leaves	aadA	2 lines from 201 shots (best 1 from 40)	homoplasmic T_0 plants but not fertile	Sikdar et al. 1998
Solanum tuberosum (potato)	PG, leaves	aadA, gfp	6 lines from 150 shots (best 2 from 12)	homoplasmic T_0 plants. no seed, tubers	Sidorov et al. 1999
	PG, leaves	aadA, gfp	14 lines from 282 shots	homoplasmic T_0 plants. no seed, tubers	Nguyen et al. 2005
Oryza sativa (rice)	PG, cell suspension	FLARE-S (aadA + gfp)	12 lines from 25 shots	heteroplastomic T_0 plants, no T_1 progeny	Khan and Maliga 1999
	PG, callus	aadA, gfp	2 lines from 120 shots	heteroplastomic T_0 and T_1 plants	Lee et al. 2006b
Glycine max (soybean)	PG, cell suspension	aadA	1 line from 984 shots	heteroplastomic callus, no plants regenerated	Zhang et al. 2001b
	PG, callus	aadA	18 lines	homoplasmic	Dufourman-

Species[a]	System[b]	Genes[c]	Efficiency[d]	Status	Reference
			from 8 shots	T_0 plants and T_1 progeny	tel et al. 2004
	PG, callus	aadA, cry1Ab	1 line from 11 shots	homoplasmic T_0 plant and T_1 progeny resistant to larval damage.	Dufourmantel et al. 2005
	PG, callus	aadA, hppd	1 line from 14 shots	homoplasmic T_0 plant and T_1 progeny, resistant to herbicide	Dufourmantel et al. 2007
Lycopersicon esculentum (tomato)	PG, leaves	aadA	6 lines from 60 shots (best 3 from 20)	homoplasmic T_0 plants and T_1 progeny	Ruf et al. 2001
	PEG, ppts	*Nicotiana tabacum* rrn16 (Spc^+,Str^+) *S. nigrum*, rrn16 (Spc^+) and rps12 (Str^+)	1 line from every 1.5 x 10^6 ppts treated	homoplasmic T_0 plants and T_1 progeny	Nugent et al. 2005
	PG, leaves	aadA, crtY, carRA	1-2 lines per shot	homoplasmic T_0 and T_1 progeny	Wurbs et al. 2007
Lesquerella fendleri (bladder pod)	PG, leaves	FLARE-S (aadA + gfp)	2 lines from 51 shots	segregating T_1 progeny from a grafted shoot	Skarjinskaia et al. 2003
Brassica napus (oilseed rape)	PG, cotyledon petioles	aadA, cry1Aa10	4 lines from 1000 explants (number of shots not described)	T_0 plants heteroplasto-mic, resistant to larval damage, T_1 progeny not described	Hou et al. 2003
Physcomitrella patens (spreading earth-moss)[f]	PEG, ppts	aadA	14 lines from 1.4 x 10^6 treated ppts	heteroplasmic and homoplasmic lines	Sugiura and Sugita 2004
Gossypium hirsutum (cotton)	PG, callus	aphA-6, nptII	30 lines from 199 shots (best 13 from 31 shots)	homoplasmic T_0 plants and T_1 progeny	Kumar et al. 2004b
Daucus carota	PG, callus	aadA, badh	9 lines	homoplasmic	Kumar et al.

Species[a]	System[b]	Genes[c]	Efficiency[d]	Status	Reference
(carrot)			from 284 shots (best 4 from 30 shots)	T_0 plants with increased salt tolerance, T_1 progeny not described	2004a
Petunia hybrida (petunia)	PG, leaves	aadA, gus	3 lines from 31 shots	homoplasmic T_0 plants and T_1 progeny	Zubko et al. 2004
Solanum rickii (wild nightshade)	PG, internode sections	aadA	2 lines from 1 shot	T_0 plants no T_1 progeny described	Matveena et al. 2005
Lactuca sativa (lettuce)	PEG, ppts	aadA, gfp, HA	9 lines from 5.6 x 10^6 treated ppts	homoplasmic T_0 plants and T_1 progeny, no expression of HA	Lelivelt et al. 2005
	PG, leaves	aadA, gfp	6 lines from 10 shots	homoplasmic T_0 plants and T_1 progeny	Kanamoto et al. 2006
Brassica oleracea (cauliflower)	PEG, ppts	aadA	1 line from 3.0 x 10^6 treated ppts	homoplasmic T_0 plant, no progeny	Nugent et al. 2006
Populus alba (poplar)	PG, leaves	aadA, gfp	10 lines from 30 shots	homoplasmic T_0 plants (5-10 years required for sexual maturity)	Okumura et al. 2006
Marchantia polymorpha (liverwort)[f]	cell suspension	aadA	30 lines from 10 shots (best 24 from 5 shots)	homoplasmic callus lines	Chiyoda et al. 2007

[a] For tobacco (*N. tabacum*) only representative papers are given. Other species are listed in the order in which they were first published together with subsequent additional reports.
[b] PG (particle gun), PEG (polyethylene glycol), ppts (protoplasts).
[c] *aadA* (aminoglycoside 3'-adenyltransferase), *nptII* (neomycin phosphotransferase), *aphA-6* (aminoglycoside phosphotransferase), *gfp* (green fluorescent protein), *gus* (ß-glucuronidase), *cry1Ab/cry1Aa10 (Bt* crystal toxin proteins), *hppd* (4-hydroxyphenylpyruvate dioxygenase), *crtY* (lycopene ß-cyclase from *Erwinia herbicola*), *carRA* (lycopene ß-cyclase from *Phycomyces blakesleeanus*), *HA* (haemagglutinin).
[d] Average efficiency for published work, direct comparison of results is difficult since different transformation and selection regimes were employed and in some cases putative transformants were not all analyzed in detail. Where appropriate optimal transformation results are given.
[e] A range of bombardment parameters tested.
[f] Moss species are listed together with higher plants.

construction of species specific transformation vectors, a suitable selection marker and use of tissues with a high regeneration capacity such that fertile plants can be recovered. Of particular merit is the recent report describing the extension of plastid transformation technology from herbaceous plants to the woody tree species poplar (Okumura et al. 2006). In contrast, limited success has been achieved with monocotyledonous plants. To date, there are only two reports on rice, both of which describe integration of foreign sequences into the plastome but no homoplasmic plants were recovered (Khan and Maliga 1999; Lee et al. 2006b).

When species, closely related to tobacco, prove difficult to transform in their plastome, an interesting approach can be used exploiting the fact that plastids in tobacco can be transformed. Kuchuk et al. (2006) transformed the plastomes of five different recalcitrant solanaceous species after transferring their plastids into tobacco; thus, generating cytoplasmic hybrids with tobacco supplying the nuclear genome and the other species donating the cytoplasmic genomes.

5 Perspectives

Plastid transformation offers a basic tool for the study of plastid gene function and regulation but has also opened up the possibility to use the technology for commercial applications. The very high expression levels observed for recombinant proteins make the system ideal for applications involving plant-made-pharmaceuticals. Tobacco has received the most attention, since it is easily transformed and is a non-food crop. To date, over 50 different recombinant proteins have been expressed in tobacco. A major class of these proteins includes vaccine-related antigens. Considerable progress has also been made in the last few years for plastid-based expression in edible crop species. While it is highly unlikely that edible vaccines will meet with regulatory approval for humans such approaches could conceivably be useful for animal vaccination or serve as an alternative to tobacco as a production platform. The ongoing challenge will be to demonstrate that a plant-based production system offers an effective alternative to conventional fermenter production.

Plastid transformants offer an additional advantage compared to nuclear transformants for genetic safety, since transgenes are maternally inherited in most crops. Improved safety coupled with high expression and the ease of selectable marker elimination may lead to a new generation of transgenic crops expressing useful agricultural traits.

The most striking limitation of plastid transformation is the lack of substantial progress with monocotyledonous species, which include the agriculturally important cereal crops. However, the outlook is encouraging; a combination of improved transformation technologies coupled with an increase in the number of groups working in the field should deliver reproducible systems for these crops in the coming years.

In algae, challenges for genetic engineering of chloroplasts include the further optimization of foreign gene expression. This will initially involve the model system *Chlamydomonas reinhardtii*. The development of transformation protocols for complex plastids of ecologically or economically relevant groups like diatoms and brown algae will provide important tools for basic as well as applied studies.

References

Ahlert D, Ruf S, Bock R (2003) Plastid protein synthesis is required for plant development in tobacco. Proc Natl Acad Sci USA 100:15730-15735

Allison LA, Simon LD, Maliga P (1996) Deletion of *rpoB* reveals a second distinct transcription system in plastids of higher plants. EMBO J 15:2802-2809

Andrews TJ, Whitney SM (2003) Manipulating ribulose bisphosphate carboxylase/oxygenase in the chloroplasts of higher plants. Arch Biochem Biophys 414:159-169

Arai Y, Shikanai T, Doi Y, Yoshida S, Yamaguchi I, Nakashita H (2004) Production of polyhydroxybutyrate by polycistronic expression of bacterial genes in tobacco plastid. Plant Cell Physiol 45:1176-1184

Baena-Gonzalez E, Gray JC, Tyystjarvi E, Aro EM, Maenpaa P (2001) Abnormal regulation of photosynthetic electron transport in a chloroplast *ycf9* inactivation mutant. J Biol Chem 276:20795-20802

Baena-Gonzalez E, Allahverdiyeva Y, Svab Z, Maliga P, Josse EM, Kuntz M, Maenpaa P, Aro EM (2003) Deletion of the tobacco plastid *psbA* gene triggers an upregulation of the thylakoid-associated NAD(P)H dehydrogenase complex and the plastid terminal oxidase (PTOX). Plant J 35:704-716

Barnes D, Franklin S, Schultz J, Henry R, Brown E, Coragliotti A, Mayfield SP (2005) Contribution of 5´- and 3´-untranslated regions of plastid mRNAs to the expression of *Chlamydomonas reinhardtii* chloroplast genes. Mol Gen Genomics 274:625-636

Bateman JM, Purton S (2000) Tools for chloroplast transformation in *Chlamydomonas*: expression vectors and new dominant selectable marker. Mol Gen Genet 263:404-410

Bellucci M, De Marchis F, Mannucci R, Bock R, Arcioni S (2005) Cytoplasm and chloroplasts are not suitable subcellular locations for beta-zein accumulation in transgenic plants. J Exp Bot 56:1205-1212

Bennoun P, Spierer-Herz M, Erickson J, Girard-Bascou J, Pierre Y, Delosme M, Rochaix JD (1986) Characterization of photosystem II mutants of *Chlamydomonas reinhardtii* lacking *psbA* gene. Plant Mol Biol 6:151-160

Berthold DA, Schmidt CL, Malkin R (1995) The deletion of *petG* in *Chlamydomonas reinhardtii* disrupts the cytochrome b6f complex. J Biol Chem 270:29293-29298

Bessette PH, Aslund F, Beckwith J, Georgiou G (1999) Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. Proc Natl Acad Sci USA 96:13703-13708

Birch-Machin I, Newell CA, Hibberd JM, Gray JC (2004) Accumulation of rotavirus VP6 protein in chloroplasts of transplastomic tobacco is limited by protein stability. Plant Biotechnol J 2:261-271

Blowers AD, Bogorad L, Shark KB, Sanford JC (1989) Studies on *Chlamydomonas* chloroplast transformation: foreign DNA can be stably maintained in the chromosome. Plant Cell 1:123-132

Bock R (1998) Analysis of RNA editing in plastids. Methods 15:75-83

Bock R (2000) Sense from nonsense: how the genetic information of chloroplasts is altered by RNA editing. Biochimie 82:549-557

Bock R (2001) Transgenic plastids in basic research and plant biotechnology. J Mol Biol 312:425-438

Bock R (2004) Studying RNA editing in transgenic chloroplasts of higher plants. Methods Mol Biol 265:345-356

Bock R (2006) Plastid biotechnology: prospects for herbicide and insect resistance, metabolic engineering and molecular farming. Curr Opin Biotechnol 17:1-7

Bock R, Khan MS (2004) Taming plastids for a green future. Trends Biotechnol 22:311-318

Bogorad L (2000) Engineering chloroplasts: an alternative site for foreign genes, proteins, reactions and products Trends Biotechnol 18:257-263

Boudreau E, Takahashi Y, Lemieux C, Turmel M, Rochaix JD (1997a) The chloroplast *ycf3* and *ycf4* open reading frames of *Chlamydomonas reinhardtii* are required for the accumulation of the photosystem I complex. EMBO J 15:6095-6104

Boudreau E, Turmel M, Goldschmidt-Clermont M, Rochaix JD, Sivan S, Michaels A, Leu S (1997b) A large open reading frame (ORF1995) in the chloroplast DNA of *Chlamydomonas reinhardtii* encodes an essential protein. Mol Gen Genet 253:649-653

Boynton JE, Gillham NW, Harris EH, Hosler JP, Johnson AM, Jones AR, Randolph-Andersen BL, Robertson D, Klein TM, Shark KB, Sanford JC (1988) Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. Science 240:1534-1538

Buhot L, Horvath E, Medgyesy P, Lerbs-Mache S (2006) Hybrid transcription system for controlled plastid transgene expression. Plant J 46:700-707

Burrows PA, Sazanov LA, Svab Z, Maliga P, Nixon PJ (1998) Identification of a functional respiratory complex in chloroplasts through analysis of tobacco mutants containing disrupted plastid *ndh* genes. EMBO J 17:868-876

Butow RA, Fox TD (1990) Organelle transformation: shoot first, ask questions later. Trends Biochem Sci 15:465-468

Carrer H, Hockenberry TN, Svab Z, Maliga P (1993) Kanamycin resistance as a selectable marker for plastid transformation in tobacco. Mol Gen Genet 241:49-56

Carrer H, Maliga P (1995) Targeted insertion of foreign genes into the tobacco plastid genome without physical linkage to the selectable marker gene. Biotechnology 13:791-794

Cerutti H, Johnson AM, Boynton JE, Gillham NW (1995) Inhibition of chloroplast DNA recombination and repair by dominant negative mutants of *Escherichia coli recA*. Mol Cell Biol 15:3003-3011

Chakrabarti SK, Lutz KA, Lertwiriyawong B, Svab Z, Maliga P (2006) Expression of the *cry9Aa2 B.t.* gene in tobacco chloroplasts confers resistance to potato tuber moth. Transgenic Res 15:481-488

Chase CD (2006) Genetically engineered cytoplasmic male sterility. Trends Plant Sci 11:7-9

Chen X, Kindle KL, Stern DB (1993) Initiation codon mutations in the *Chlamydomonas* chloroplast *petD* gene result in temperature-sensitive photosynthetic growth. EMBO J 12:3627-3635

Chiyoda S, Linley PJ, Yamato KT, Fukuzawa H, Yokota A, Kohchi T (2007) Simple and efficient plastid transformation system for the liverwort *Marchantia polymorpha* L. suspension-culture cells. Transgenic Res 16:41-49

Choquet Y, Rahire M, Girard-Bascou J, Erickson J, Rochaix JD (1992) A chloroplast gene is required for the light-independent accumulation of chlorophyll in *Chlamydomonas reinhardtii*. EMBO J 11:1697-1704

Corneille S, Lutz K, Svab Z, Maliga P (2001) Efficient elimination of selectable marker genes from the plastid genome by the CRE-*lox* site-specific recombination system. Plant J 27:171-178

Corneille S, Lutz KA, Azhagiri AK, Maliga P (2003) Identification of functional *lox* sites in the plastid genome. Plant J 35:753-762

Daniell H (2000) Genetically modified food crops: current concerns and solutions for next generation crops. Biotechnol Genet Eng Rev 17:327-352

Daniell H (2002) Molecular strategies for gene containment in transgenic crops. Nat Biotechnol 20:581-586

Daniell H (2006) Production of biopharmaceuticals and vaccines in plants via the chloroplast genome. Biotechnol J 1:1071-1079

Daniell H, Dhingra A (2002) Multigene engineering: dawn of an exciting new era in biotechnology. Curr Opin Biotechnol 13:136-141

Daniell H, Vivekananda J, Nielsen BL, Ye GN, Tewari KK (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

Daniell H, Krishnan M, McFadden BE (1991) Transient expression of ß-glucuronidase in different cellular compartments following biolistic delivery of foreign DNA into wheat leaves and calli. Plant Cell Rep 9:615-619

Daniell H, Datta R, Varma S, Gray S, Lee SB (1998) Containment of herbicide resistance through genetic engineering of the chloroplast genome. Nat Biotechnol 16:345-348

Daniell H, Streatfield SJ, Wycoff K (2001a) Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. Trends Plant Sci 6:219-226

Daniell H, Muthukumar B, Lee SB (2001b) Marker free transgenic plants: engineering the chloroplast genome without the use of antibiotic selection. Curr Genet 39:109-116

Daniell H, Lee SB, Panchal T, Wiebe PO (2001c) Expression of the native cholera toxin B subunit gene and assembly as functional oligomers in transgenic tobacco chloroplasts. J Mol Biol 311:1001-1009

Daniell H, Khan MS, Allison L (2002) Milestones in chloroplast genetic engineering: an environmentally friendly era in biotechnology. Trends Plant Sci 7:84-91

Daniell H, Chebolu S, Kumar S, Singleton M, Falconer R (2005a) Chloroplast-derived vaccine antigens and other therapeutic proteins. Vaccine 23:1779-1783

Daniell H, Kumar S, Dufourmantel N (2005b) Breakthrough in chloroplast genetic engineering of agronomically important crops. Trends Biotechnol 23:238-245

Daniell H, Ruiz ON, Dhingra A (2005c) Chloroplast genetic engineering to improve agronomic traits. Methods Mol Biol 286:111-138

De Block M, Schell J, Van Montagu M (1985) Chloroplast transformation by *Agrobacterium tumefaciens*. EMBO J 4:1367-13732

De Cosa B, Moar W, Lee SB, Miller M, Daniell H (2001) Overexpression of the *Bt cry2Aa2* operon in chloroplasts leads to formation of insecticidal crystals. Nat Biotechnol 19:71-74

De Santis-Maciossek G, Kofer W, Bock A, Schoch S, Maier RM, Wanner G, Rudiger W, Koop HU, Herrmann RG (1999) Targeted disruption of the plastid RNA polymerase genes *rpoA, B* and *C1*: molecular biology, biochemistry and ultrastructure. Plant J 18:477-489

DeGray G, Rajasekaran K, Smith F, Sanford J, Daniell H (2001) Expression of an antimicrobial peptide via the chloroplast genome to control phytopathogenic bacteria and fungi. Plant Physiol 127:852-862

Delwiche CF (1999) Tracing the thread of plastid diversity through the tapestry of life. Am Nat 154:164-177

Dhingra A, Daniell H (2006) Chloroplast genetic engineering via organogenesis or somatic embryogenesis. Methods Mol Biol 323:245-262

Dhingra A, Portis AR Jr, Daniell H (2004) Enhanced translation of a chloroplast-expressed RbcS gene restores small subunit levels and photosynthesis in nuclear RbcS antisense plants. Proc Natl Acad Sci USA 101:6315-6320

Dix PJ, Kavanagh TA (1995) Transforming the plastome: genetic markers and DNA delivery systems. Euphytica 85:29-34

Doetsch N, Favreau M, Kuscuoglu N, Thompson M, Hallick RB (2001) Chloroplast transformation in *Euglena gracilis*: splicing of a groupIII twintron transcribed from a transgenic *psbK* operon. Curr Genet 39:49-60

Drapier D, Suzuki H, Levy H, Rimbault B, Kindle KL Stern DB, Wollman FA (1998) The chloroplast *atpA* gene cluster in *Chlamydomonas reinhardtii*. Functional analysis of a polycistronic transcription unit. Plant Physiol 117:629-641

Drescher A, Ruf S, Calsa T Jr, Carrer H, Bock R (2000) The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. Plant J 22:97-104

Dufourmantel N, Pelissier B, Garcon F, Peltier G, Ferullo JM, Tissot G (2004) Generation of fertile transplastomic soybean. Plant Mol Biol 55:479-489

Dufourmantel N, Tissot G, Goutorbe F, Garcon F, Muhr C, Jansens S, Pelissier B, Peltier G, Dubald M (2005) Generation and analysis of soybean plastid transformants expressing *Bacillus thuringiensis* Cry1Ab protoxin. Plant Mol Biol 58:659-668

Dufourmantel N, Dubald M, Matringe M, Canard H, Garcon F, Job C, Kay E, Wisniewski JP, Ferullo JM, Pelissier, B, Sailland A, Tissot G (2007) Generation and characterization of soybean and marker-free tobacco plastid transformants overexpressing a bacterial 4-hydroxyphenylpyruvate dioxygenase which provides strong herbicide. Plant Biotechnol J 5:118-133

Eibl C, Zou Z, Beck A, Kim M, Mullet J, Koop HU (1999) In vivo analysis of plastid *psbA*, *rbcL* and *rpl32* UTR elements by chloroplast transformation: tobacco plastid gene expression is controlled by modulation of transcript levels and translation efficiency. Plant J 19:333-345

Erickson J, Rahire M, Malnoe P, Girard-Bascou J, Pierre Y, Bennoun P, Rochaix JD (1986) Lack of D2 protein in a *Chlamydomonas reinhardtii psbD* mutant affects photosystem II stability and D1 expression. EMBO J 5:1745-1754

Falk J, Brosch M, Schafer A, Braun S, Krupinska K (2005) Characterization of transplastomic tobacco plants with a plastid localized barley 4-hydroxyphenylpyruvate dioxygenase. J Plant Physiol 162:738-742

Fernández-San Millán A, Mingo-Castel A, Miller M, Daniell H (2003) A chloroplast transgenic approach to hyper-express and purify human serum albumin, a protein highly susceptible to proteolytic degradation. Plant Biotechnol J 1:71-79

Fernández-San Millán A, Farran I, Molina A, Mingo-Castel AM, Veramendi J (2007) Expression of recombinant proteins lacking methionine as N-terminal amino acid in plastids: Human serum albumin as a case study. J Biotechnol 127:593-604

Fiebig A, Stegemann S, Bock R (2004) Rapid evolution of RNA editing sites in a small non-essential plastid gene. Nucleic Acids Res 32:3615-3622

Fischer N, Stampacchia O, Redding K, Rochaix JD (1996) Selectable marker recycling in the chloroplast. Mol Gen Genet 251:373-380

Fischer N, Boudreau E, Hippler M, Drepper F, Haehnel W, Rochaix JD (1999) A large fraction of PsaF is nonfunctional in photosystem I complexes lacking the PsaJ subunit. Biochemistry 36:93-102

Fox JL (2006) Turning plants into protein factories. Nat Biotechnol 24:1191-1193

Franklin SE, Mayfield SP (2004) Prospects for molecular farming in the green alga *Chlamydomonas reinhardtii.* Curr Opin Plant Biol 7:159-165

Franklin S, Ngo B, Efuet E, Mayfield SP (2002) Development of a GFP reporter gene for *Chlamydomonas reinhardtii* chloroplast. Plant J 30:733-744

Giglione C, Meinnel T (2001) Organellar peptide deformylases: universality of the N-terminal methionine cleavage mechanism. Trends Plant Sci 6:566-572

Gleba Y, Klimyuk V, Marillonnet S (2005) Magnifection - a new platform for expressing recombinant vaccines in plants. Vaccine 23:2042-2048

Glenz K, Bouchon B, Stehle T, Wallich R, Simon MM, Warzecha H (2006) Production of a recombinant bacterial lipoprotein in higher plant chloroplasts. Nat Biotechnol 24:76-77

Golds TJ, Maliga P, Koop HU (1993) Stable plastid transformation in PEG-treated protoplasts of *Nicotiana tabacum*. Biotechnology 11:95-97

Goldschmidt-Clermont M (1991) Transgenic expression of aminoglycoside adenine transferase in the chloroplast: a selectable marker for site-directed transformation of *Chlamydomonas*. Nucleic Acids Res 19:4083-4089

Goldschmidt-Clermont M (1998) Chloroplast Transformation. In: Rochaix JD, Goldschmidt-Clermont M, Merchant S (eds) The molecular biology of chloroplasts and mitochondria in *Chlamydomonas*. Kluwer Acad Pub: Netherlands pp 139-149

Goldschmidt-Clermont M, Choquet Y, Girard-Bascou J, Michel F, Schirmer-Rahire M, Rochaix JD (1991) A small chloroplast RNA may be required for trans-splicing in *Chlamydomonas reinhardtii*. Cell 65:135-143

Grossman AR, Harris EE, Hauser C, Lefebvre PA, Martinez D, Rokhsar D, Shrager J, Silflow CD, Stern D, Vallon O, Zhang Z (2003) *Chlamydomonas reinhardtii* at the crossroads of genetics. Eukar Cell 2:1137-1150

Guda C, Lee SB, Daniell H (2000) Stable expression of a biogradable protein-based polymer in tobacco chloroplasts. Plant Cell Rep 19:257-262

Hager M, Bock R (2000) Enslaved bacteria as new hope for plant biotechnologists. Appl Microbiol Biotechnol 54:302-310

Hager M, Biehler K, Illerhaus J, Ruf S, Bock R (1999) Targeted inactivation of the smallest plastid genome-encoded open reading frame reveals a novel and essential subunit of the cytochrome b(6)f complex. EMBO J 18:5834-5842

Hager M, Hermann M, Biehler K, Krieger-Liszkay A, Bock R (2002) Lack of the small plastid-encoded PsbJ polypeptide results in a defective water-splitting apparatus of photosystem II, reduced photosystem I levels, and hypersensitivity to light. J Biol Chem 277:14031-14039

Hajdukiewicz PT, Gilbertson L, Staub JM (2001) Multiple pathways for Cre/lox-mediated recombination in plastids. Plant J 27:161-170

Heifetz PB (2000) Genetic engineering of the chloroplast. Biochimie 82:655-666

Heifetz PB, Tuttle AM (2001) Protein expression in plastids. Curr Opin Plant Biol 4:157-161

Herrin D, Nickelsen J (2004) Chloroplast RNA processing and stability. Photosyn Res 82:301-314

Herz S, Füssl M, Steiger S, Koop HU (2005) Development of novel types of plastid transformation vectors and evaluation of factors controlling expression. Transgenic Res 14:969-982

Hibberd JM, Linley PJ, Khan MS, Gray JC (1998) Transient expression of green fluorescent protein in various plastid types following microprojectile bombardment. Plant J 8:627-632

Horvath EM, Peter SO, Joet T, Rumeau D, Cournac L, Horvath GV, Kavanagh TA, Schafer C, Peltier G, Medgyesy P (2000) Targeted inactivation of the plastid *ndhB* gene in tobacco results in an enhanced sensitivity of photosynthesis to moderate stomatal closure. Plant Physiol 123:1337-1350

Hou BK, Zhou YH, Wan LH, Zhang ZL, Shen GF, Chen ZH, Hu ZM (2003) Chloroplast transformation in oilseed rape. Transgenic Res 12:111-114

Houtz RL, Stults JT, Mulligan RM, Tolbert NE (1989) Post-translational modifications in the large subunit of ribulose bisphosphate carboxylase/oxygenase. Proc Natl Acad Sci USA 86:1855-1859

Howe CJ (1988) Organelle transformation. Trends Genet 4:150-152

Huang C, Wang S, Chen L, Lemieux C, Otis C, Turmel M, Liu XQ (1994) The *Chlamydomonas* chloroplast *clpP* gene contains translated large insertion sequences and is essential for cell growth. Mol Gen Genet 244:151-159

Huang FC, Klaus SM, Herz S, Zou Z, Koop HU, Golds TJ (2002) Efficient plastid transformation in tobacco using the *aphA*-6 gene and kanamycin selection. Mol Genet Genomics 268:19-27

Iamtham S, Day A (2000) Removal of antibiotic resistance genes from transgenic tobacco plastids. Nat Biotechnol 18:1172-1176

Ishikura K, Takaoka Y, Kato K, Sekine M, Yoshida K, Shinmyo A (1999) Expression of a foreign gene in *Chlamydomonas reinhardtii* chloroplast. J Biosci Bioengin 87:307-314

Kanamoto H, Yamashita A, Asao H, Okumura S, Takase H, Hattori M, Yokota A, Tomizawa K (2006) Efficient and stable transformation of *Lactuca sativa* L. cv. Cisco (lettuce) plastids. Transgenic Res 15:205-217

Kanevski I, Maliga P (1994) Relocation of the plastid *rbcL* gene to the nucleus yields functional ribulose-1,5-bisphosphate carboxylase in tobacco chloroplasts. Proc Natl Acad Sci USA 91:1969-1973

Kang TJ, Loc NH, Jang MO, Jang YS, Kim YS, Seo JE, Yang MS (2003a) Expression of the B subunit of *E. coli* heat-labile enterotoxin in the chloroplasts of plants and its characterization. Transgenic Res 12:683-691

Kang TJ, Seo JE, Loc NH, Yang MS (2003b) Herbicide resistance of tobacco chloroplasts expressing the *bar* gene. Mol Cells 16:60-66

Kang TJ, Han SC, Kim MY, Kim YS, Yang MS (2004) Expression of non-toxic mutant of *Escherichia coli* heat-labile enterotoxin in tobacco chloroplasts. Protein Expr Purif 38:123-128

Kasai S, Yoshimura S, Ishikura K, Takaoka Y, Kobayashi K, Kato K, Shinmyo A (2003) Effect of coding regions on chloroplast gene expression in *Chlamydomonas reinhardtii*. J Biosci Bioengin 95:276-282

Kato K, Ishikura K, Kasai S, Shinmyo A (2006) Efficient translation destabilizes transcripts in chloroplasts of *Chlamydomonas reinhardtii*. J Biosci Bioengin 101:471-477

Kavanagh TA, Thanh ND, Lao NT, McGrath N, Peter SO, Horvath EM, Dix PJ, Medgyesy P (1999) Homeologous plastid DNA transformation in tobacco is mediated by multiple recombination events. Genetics 152:1111-1122

Khan MS, Maliga P (1999) Fluorescent antibiotic resistance marker for tracking plastid transformation in higher plants. Nat Biotechnol 17:910-915

Khan MS, Khalid AM, Malik KA (2005) Phage phiC31 integrase: a new tool in plastid genome engineering. Trends Plant Sci 10:1-3

Kindle KL, Richards KL, Stern DB (1991) Engineering the chloroplast genome: techniques and capabilities for chloroplast transformation in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA 88:1721-1725

Kindle KL, Suzuki H, Stern DB (1994) Gene amplification can correct a photosynthetic growth defect caused by mRNA instability in *Chlamydomonas* chloroplasts. Plant Cell 6:187-200

Klaus SM, Huang FC, Eibl C, Koop HU, Golds TJ (2003) Rapid and proven production of transplastomic tobacco plants by restoration of pigmentation and photosynthesis. Plant J 35:811-821

Klaus SM, Huang FC, Golds TJ, Koop HU (2004) Generation of marker-free plastid transformants using a transiently cointegrated selection gene. Nat Biotechnol 22:225-229

Knoblauch M, Hibberd JM, Gray JC, van Bel AJ (1999) A galinstan expansion femtosyringe for microinjection of eukaryotic organelles and prokaryotes. Nat Biotechnol 17:906-909

Kode V, Mudd EA, Iamtham S, Day A (2005) The tobacco plastid *accD* gene is essential and is required for leaf development. Plant J 44:237-244

Kode V, Mudd EA, Iamtham S, Day A (2006) Isolation of precise plastid deletion mutants by homology-based excision: a resource for site-directed mutagenesis, multi-gene changes and high-throughput plastid transformation. Plant J 46:901-909

Kofer W, Eibl C, Steinmüller K, Koop HU (1998a) PEG-mediated plastid transformation in higher plants. In vitro Cell Dev Biol Plant 34:303-309

Kofer W, Koop HU, Wanner G, Steinmuller K (1998b) Mutagenesis of the genes encoding subunits A, C, H, I, J and K of the plastid NAD(P)H-plastoquinone-oxidoreductase in tobacco by polyethylene glycol-mediated plastome transformation. Mol Gen Genet 258:166-173

Koop HU, Steinmuller K, Wagner H, Rossler C, Eibl C, Sacher L (1996) Integration of foreign sequences into the tobacco plastome via polyethylene glycol-mediated protoplast transformation. Planta 199:193-201

Kota M, Daniell H, Varma S, Garczynski SF, Gould F, Moar WJ (1999) Overexpression of the *Bacillus thuringiensis (Bt)* Cry2Aa2 protein in chloroplasts confers resistance to plants against susceptible and *Bt*-resistant insects. Proc Natl Acad Sci USA 96:1840-1845

Kuchuk N, Sytnyk K, Vasylenko M, Shakhovsky A, Komarnytsky I, Kushnir S, Gleba Y (2006) Genetic transformation of plastids of different *Solanaceae* species using tobacco cells as organelle hosts. Theor Appl Genet 113:519-527

Kumar S, Dhingra A, Daniell H (2004a) Plastid-expressed betaine aldehyde dehydrogenase gene in carrot cultured cells, roots, and leaves confers enhanced salt tolerance. Plant Physiol 136:2843-2854

Kumar S, Dhingra A, Daniell H (2004b) Stable transformation of the cotton plastid genome and maternal inheritance of transgenes. Plant Mol Biol 56:203-216

Kunstner P, Guardiola A, Takahashi Y, Rochaix JD (1995) A mutant strain of *Chlamydomonas reinhardtii* lacking the chloroplast photosystem II *psbI* gene grows photoautotrophically. J Biol Chem 270:9651-9654

Kuras R, Wollman FA (1994) The assembly of cytochrome b6f complexes: an approach using genetic transformation of the green alga *Chlamydomonas reinhardtii*. EMBO J 13:1019-1027

Kuroda H, Maliga P (2001a) Sequences downstream of the translation initiation codon are important determinants of translation efficiency in chloroplasts. Plant Physiol 125:430-436

Kuroda H, Maliga P (2001b) Complementarity of the 16S rRNA penultimate stem with sequences downstream of the AUG destabilizes the plastid mRNAs. Nucleic Acids Res 29:970-975

Kuroda H, Maliga P (2002) Overexpression of the *clpP* 5'-untranslated region in a chimeric context causes a mutant phenotype, suggesting competition for a *clpP*-specific RNA maturation factor in tobacco chloroplasts. Plant Physiol 129:1600-1606

Kuroda H, Maliga P (2003) The plastid *clpP*1 protease gene is essential for plant development. Nature 425:86-89

Langbecker CL, Ye G-N, Broyles DL, Duggan LL, Xu CW, Hajdukiewicz PTJ, Armstrong CL, Staub JM (2004) High-frequency transformation of undeveloped plastids in tobacco suspension cells. Plant Physiol 135:39-46

Lapidot M, Raveh D, Sivan A, Shoshana A, Shapira M (2002) Stable chloroplast transfromation of the unicellular red alga *Porphyridium spec*. Plant Physiol 129:7-12

Lee SB, Kwon HB, Kwon SJ, Park SC, Jeong MJ, Han SE, Byun MO, Daniell H (2003) Accumulation of trehalose within transgenic chloroplasts confers drought tolerance. Mol Breed 11:1–13

Lee MY, Zhou Y, Lung RW, Chye ML, Yip WK, Zee SY, Lam E (2006a) Expression of viral capsid protein antigen against Epstein-Barr virus in plastids of *Nicotiana tabacum* cv. SR1. Biotechnol Bioeng 94:1129-1137

Lee SM, Kang KH, Chung H, Yoo SH, Xu XM, Lee SB, Cheong JJ, Daniell H, KimM (2006b) Plastid transformation in the monocotyledonous cereal crop, rice (*Oryza sativa*) and transmission of transgenes to their progeny. Mol Cells 21:401-410

Leelavathi S, Reddy VS (2003) Chloroplast expression of His-tagged GUS-fusions: a general strategy to overproduce and purify foreign proteins using transplastomic plants as bioreactors. Mol Breed 11:49-58

Leelavathi S, Gupta N, Maiti S, Ghosh A, Reddy VS (2003) Overproduction of an alkali- and thermo-stable xylanase in tobacco chloroplasts and efficient recovery of the enzyme. Mol Breed 11:59-67

Lelivelt CL, McCabe MS, Newell CA, Desnoo CB, van Dun KM, Birch-Machin I, Gray JC, Mills KH, Nugent, JM (2005) Stable plastid transformation in lettuce (*Lactuca sativa* L.). Plant Mol Biol 58:763-774

Lerbs-Mache S (1993) The 110-kDa polypeptide of spinach plastid DNA-dependent RNA polymerase: Single-subunit enzyme or catalytic core of multimeric enzyme complexes? Proc Natl Acad Sci USA 90:5509-5513

Li J, Goldschmidt-Clermont M, Timko MP (1993) Chloroplast-encoded *chlB* is required for light-independent protochlorophyllide reductase activity in *Chlamydomonas reinhardtii*. Plant Cell 5:1817-29

Li Y, Sun M, Liu J, Yang Z, Zhang Z, Shen G (2006a) High expression of foot-and-mouth disease virus structural protein VP1 in tobacco chloroplasts. Plant Cell Rep 25:329-333

Li HY, Ramalingam S, Chye ML (2006b) Accumulation of recombinant SARS-CoA spike protein in plant cytosol and chloroplasts indicate potential for development of plant-derived oral vaccines. Exp Biol Med 231:1346-1352

Limaye A, Koya V, Samsam M, Daniell H (2006) Receptor-mediated oral delivery of a bioencapsulated green fluorescent protein expressed in transgenic chloroplasts into the mouse circulatory system. FASEB J 20:959-961

Liu XQ, Huang C, Xu H (1993) The unusual *rps3*-like ORF172 is functionally essential and structurally conserved in *Chlamydomonas*. FEBS Lett 336:225-230

Lorence A, Verpoorte R (2004) Gene transfer and expression in plants. Methods Mol Biol 267:329-350

Lössl A, Eibl C, Harloff HJ, Jung C, Koop HU (2003) Polyester synthesis in transplastomic tobacco (*Nicotiana tabacum* L.): Significant contents of polyhydroxybutyrate are associated with growth reduction. Plant Cell Rep 21:891-899

Lössl, A, Bohmert, K, Harloff, H, Eibl, C, Mühlbauer, S, Koop, HU (2005) Inducible trans-activation of plastid transgenes: expression of the *R. eutropha phb* operon in transplastomic tobacco. Plant Cell Physiol 46:1462-1471

Lu XM, Yin WB, Hu ZM (2006) Chloroplast transformation. Methods Mol Biol 318:285-303

Lutz KA, Knapp JE, Maliga P (2001) Expression of *bar* in the plastid genome confers herbicide resistance. Plant Physiol 125:1585-1590

Lutz KA, Corneille S, Azhagiri AK, Svab Z, Maliga P (2004) A novel approach to plastid transformation utilizes the phiC31 phage integrase. Plant J 37:906-913

Lutz KA, Svab Z, Maliga P (2006a) Construction of marker-free transplastomic tobacco using the Cre-loxP site-specific recombination system. Nature Protoc 1:900-910

Lutz KA, Bosacchi MH, Maliga P (2006b) Plastid marker-gene excision by transiently expressed CRE recombinase. Plant J 45:447-456

Ma JK, Barros E, Bock R, Christou P, Dale PJ, Dix PJ, Fischer R, Irwin J, Mahoney R, Pezzotti M, Schillberg S, Sparrow P, Stoger E, Twyman RM; European Union Framework 6 Pharma-Planta Consortium (2005) Molecular farming for new drugs and vaccines. Current perspectives on the production of pharmaceuticals in transgenic plants. EMBO Rep 6:593-599

Madoka Y, Tomizawa KI, Mizoi J, Nishida I, Nagano Y, Sasaki Y (2002) Chloroplast transformation with modified *accD* operon increases acetyl-CoA carboxylase and causes extension of leaf longevity and increase in seed yield in tobacco. Plant Cell Physiol 43:1518-1525

Mänpää P, Gonzalez EB, Chen L, Khan MS, Gray JC, Aro EM (2000) The *ycf 9* (orf 62) gene in the plant chloroplast genome encodes a hydrophobic protein of stromal thylakoid membranes. J Exp Bot 51:375-382

Magee AM, Kavanagh TA (2002) Plastid genes transcribed by the nucleus-encoded plastid RNA polymerase show increased transcript accumulation in transgenic plants expressing a chloroplast-localized phage T7 RNA polymerase. J Exp Bot 53:2341-2349

Magee AM, Coyne S, Murphy D, Horvath EM, Medgyesy P, Kavanagh TA (2004a) T7 RNA polymerase-directed expression of an antibody fragment transgene in plastids causes a semi-lethal pale-green seedling phenotype. Transgenic Res 13:325-337

Magee AM, Horvath EM, Kavanagh TA (2004b) Pre-screening plastid transgene expression cassettes in *Escherichia coli* may be unreliable as a predictor of expression levels in chloroplast-transformed plants. Plant Sci 166:1605-1611

Magee AM, MacLean D, Gray JC, Kavanagh TA (2007) Disruption of essential plastid gene expression caused by T7 RNA polymerase-mediated transcription of plastid transgenes during early seedling development. Transgenic Res: in press (doi 10.1007/s11248-006-9045-z)

Majeran W, Wollman F-A, Vallon O (2000) Evidence for a role of ClpP in the degradation of the chloroplast cytochrome b_6f complex. Plant Cell 12:137-149

Maliga P (1993) Towards plastid transformation in flowering plants. TIBTECH 11:101-107

Maliga P (2002) Engineering the plastid genome of higher plants. Curr Opin Plant Biol 5:164-172

Maliga P (2003) Progress towards commercialization of plastid transformation technology. Trends Biotechnol 21:20-28

Maliga P (2004) Plastid transformation in higher plants. Annu Rev Plant Biol 55:289-313

Maliga P (2005) New vectors and marker excision systems mark progress in engineering the plastid genome of higher plants. Photochem Photobiol Sci 4:971-976

Maliga P, Carrer H, Kanevski I, Staub J, Svab Z (1993) Plastid engineering in land plants: a conservative genome is open to change. Philos Trans R Soc Lond B Biol Sci 342:203-208

Marin-Navarro J, Moreno J (2006) Cysteines 449 and 459 modulate the reduction-oxidation conformational changes of ribulose 1.5-bisphosphate carboxylase/oxygenase and the translocation of the enzyme to membranes during stress. Plant Cell Environ 29:898-908

Martin M, Casano LM, Zapata JM, Guera A, del Campo EM, Schmitz-Linneweber C, Maier RM, Sabater B (2004) Role of thylakoid Ndh complex and peroxidase in the protection against photo-oxidative stress: fluorescence and enzyme activities in wild-type and ndhF-deficient tobacco. Physiol Plant 1221:443-452

Matsuo T, Onai K, Okamoto K, Minagawa J, Ishiura M (2006) Real-time monitoring of chloroplast gene expression by a luciferase reporter: evidence for nuclear regulation of chloroplast circadian period. Molec Cell Biol 26:863-870

Matveeva NA, Shakhovskii AM, Kuchuk NV (2005) Stable transformation of *Solanum rickii* chloroplast DNA. Tsitol Genet 39:3-8

Maul JE, Lilly JW, Cui L, dePamphilis CW, Miller W, Harris EH, Stern DB (2002) The *Chlamydomonas reinhardtii* plastid chromosome: islands of genes in a sea of repeats. Plant Cell 14:2659-2679

Mayfield SP, Schultz J (2004) Development of a luciferase reporter gene, *luxCt*, for *Chlamydomonas reinhardtii* chloroplast. Plant J 37:449-458

Mayfield SP, Franklin SE (2005) Expression of human antibodies in eukaryotic microalgae. Vaccine 23:1828-1832

Mayfield SP, Franklin SE, Lerner RA (2003) Expression and assembly of a fully active antibody in algae. Proc Natl Acad Sci USA 100:438-442

McBride KE, Schaaf DJ, Daley M, Stalker DM (1994) Controlled expression of plastid transgenes in plants based on a nuclear DNA-encoded and plastid-targeted T7 RNA polymerase. Proc Natl Acad Sci USA 91:7301-7305

McBride KE, Svab Z, Schaaf DJ, Hogan PS, Stalker DM, Maliga P (1995) Amplification of a chimeric *Bacillus* gene in chloroplasts leads to an extraordinary level of an insecticidal protein in tobacco. Biotechnology (NY) 13:362-365

Minko I, Holloway SP, Nikaido S, Carter M, Odom OW, Johnson C, Herrin DL (1999) *Renilla* luciferase as a vital reporter for chloroplast gene expression in *Chlamydomonas*. Mol Gen Genet 262:421-425

Molina A, Hervás-Stubbs S, Daniell H, Mingo-Castel A, Veramendi J (2004) High-yield expression of a viral peptide animal vaccine in transgenic tobacco chloroplasts. Plant Biotechnol J 2:141-153

Monde RA, Zito F, Olive J, Wollman FA, Stern DB (2000) Post-transcriptional defects in tobacco chloroplast mutants lacking the cytochrome b6/f complex. Plant J 21:61-72

Morais F, Barber J, Nixon PJ (1998) The chloroplast-encoded alpha subunit of cytochrome b-559 is required for assembly of the photosystem II complex in both the light and the dark in *Chlamydomonas reinhardtii*. J Biol Chem 273:29315-29320

Mühlbauer SK, Koop HU (2005) External control of transgene expression in tobacco plastids using the bacterial lac repressor. Plant J 43:941-946

Mühlbauer SK, Lössl A, Tzekova L, Zou Z, Koop HU (2002) Functional analysis of plastid DNA replication origins in tobacco by targeted inactivation. Plant J 32:175-184

Nakamura M, Sugiura M (2007) Translation efficiencies of synonymous codons are not always correlated with codon usage in tobacco chloroplasts. Plant J 49:128-134

Nakashita H, Arai Y, Shikanai T, Doi Y, Yamaguchi I (2001) Introduction of bacterial metabolism into higher plants by polycistronic transgene expression. Biosci Biotechnol Biochem 65:1688-1691

Newell CA, Birch-Machin I, Hibberd JM, Gray JC (2003) Expression of green fluorescent protein from bacterial and plastid promoters in tobacco chloroplasts. Transgenic Res 12:631-634

Newman S, Harris EH, Johnson AM, Boynton JE, Gillham NW (1992) Nonrandom distribution of chloroplast recombination events in *Chlamydomonas reinhardtii*: Evidence for a hotspot and an adjacent cold region. Genetics 132:413-429

Nguyen, TT, Nugent, G, Cardi, T, Dix PJ (2005) Generation of homoplasmic plastid transformants of a commercial cultivar of potato (*Solanum tuberosum* L.) Plant Sci 168:1495-1500

Nickelsen J (1999) Transcripts containing the 5´ untranslated regions of the plastid genes *psbA* and *psbB* from higher plants are unstable in *Chlamydomonas reinhardtii* chloroplasts. Mol Gen Genet 262:768-771

Nickelsen J, Kück U (2000) The unicellular green alga *Chlamydomonas reinhardtii* as an experimental system to study chloroplast RNA metabolism. Naturwiss 87:97-107

Nugent JM, Joyce SM (2005) Producing human therapeutic proteins in plastids. Curr Pharm Des 11:2459-2470

Nugent GD, Ten Have M, van der Gulik A, Dix PJ, Uijtewaal BA, Mordhorst AP (2005) Plastid transformants of tomato selected using mutations affecting ribosome structure. Plant Cell Rep 24:341-349

Nugent GD, Coyne S, Nguyen TT, Kavanagh T, Dix PJ (2006) Nuclear and plastid transformation of *Brassica oleracea* var. botrytis (cauliflower) using PEG-mediated uptake of DNA into protoplasts. Plant Sci 170:135-142

O´Connor HE, Ruffle SV, Cain AJ, Deak Z, Vass I, Nugent JH, Purton S (1998) The 9 kDa phosphoprotein of photosystem II. Generation and characterization of *Chlamydomonas* mutants lacking PSII-H and a site-directed mutant lacking the phosporylation site. Biochem Biophys Acta 1364:63-72

Ohnishi N, Takahashi Y (2001) PsbT polypeptide is required for efficient repair of photodamaged photosystem II reaction center. J Biol Chem 276:33798-33804

Okumura S, Sawada M, Park YW, Hayashi T, Shimamura M, Takase H, Tomizawa K (2006) Transformation of poplar (*Populus alba*) plastids and expression of foreign proteins in tree chloroplasts. Transgenic Res 15:637-646

O'Neill C, Horvath GV, Horvath E, Dix PJ, Medgyesy P (1993) Chloroplast transformation in plants: polyethylene glycol (PEG) treatment of protoplasts is an alternative to biolistic delivery systems. Plant J 3:729-738

Paszkowski J, Shillito RD, Saul M, Mandak V, Hohn T, Hohn B, Potrykus I (1984) Direct gene transfer to plants. EMBO J 3:2717-2722

Przibilla E, Heiss S, Johanningmeier U, Trebst A (1991) Site-specific mutageneisis of the D1 subunit of photosystem II in wild-type *Chlamydomonas*. Plant Cell 3:169-174

Qin S, Jiang P, Tseng C (2005) Transforming kelp into a marine bioreactor. Trends Biotech 23:265-268

Quesada-Vargas T, Ruiz ON, Daniell H (2005) Characterization of heterologous multigene operons in transgenic chloroplasts: transcription, processing, and translation. Plant Physiol 138:1746-1762

Ramesh VM, Bingham SE, Webber AN (2004) A simple method for chloroplast transformation in *Chlamydomonas reinhardtii*. Methods Mol Biol 274:301-307

Redding K, Cournac L, Vassiliev IR, Golbeck JH, Peltier G, Rochaix JD (1999) Photosystem I is indispensable for photoautotrophic growth, CO_2 fixation, and H_2 photoproduction in *Chlamydomonas reinhardtii*. J Biol Chem 274:10466-10473

Reddy VS, Leelavathi S, Selvapandiyan A, Raman R, Giovanni F, Shukla V, Bhatnagar RK (2002) Analysis of chloroplast transformed tobacco plants with *cry1Ia5* under rice *psbA* transcriptional elements reveal high level expression of *Bt* toxin without imposing yield penalty and stable inheritance of transplastome. Mol Breed 9:259–269

Robertson D, Boynton JE, Gillham NW (1990) Cotranscription of the wild-type chloroplast *atpE* gene encoding the CF1/CF0 epsilon subunit with 3′ half of the *rps7* gene in *Chlamydomonas reinhardtii* and characterization of frameshift mutations in *atpE*. Mol Gen Genet 221:155-163

Rochaix JD (1995) *Chlamydomonas reinhardtii* as the photosynthetic yeast. Ann Rev Genet 29:209-230

Rochaix JD (1997) Chloroplast reverse genetics: New insights into the function of plastid genes. Trends Plant Sci 2:419-425

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

Rogalski M, Ruf S, Bock R (2006) Tobacco plastid ribosomal protein S18 is essential for cell survival. Nucleic Acids Res 34:4537-45645

Rolland N, Dome AJ, Amoroso G, Sultemeyer DF, Joyard J, Rochaix JD (1997) Disruption of the plastid *ycf10* open reading frame aafects uptake of inorganic carbon in the chloroplast of *Chlamydomonas*. EMBO J 16:6713-6726

Ruf S, Kossel H, Bock R (1997) Targeted inactivation of a tobacco intron-containing open reading frame reveals a novel chloroplast-encoded photosystem I-related gene. J Cell Biol 139:95-102

Ruf S, Biehler K, Bock R (2000) A small chloroplast-encoded protein as a novel architectural component of the light-harvesting antenna. J Cell Biol 149:369-378

Ruf S, Hermann M, Berger IJ, Carrer H, Bock R (2001) Stable genetic transformation of tomato plastids and expression of a foreign protein in fruit. Nat Biotechnol 19:870-875

Ruiz ON, Daniell H (2005) Engineering cytoplasmic male sterility via the chloroplast genome by expression of {beta}-ketothiolase. Plant Physiol 138:1232-1246

Ruiz ON, Hussein HS, Terry N, Daniell H (2003) Phytoremediation of organomercurial compounds via chloroplast genetic engineering. Plant Physiol 132:1344-1352

Rumeau D, Bécuwe-Linka N, Beyly A, Carrier P, Cuiné S, Genty B, Medgyesy P, Horvath EM, Peltier G (2004) Increased zinc content in transplastomic tobacco plants expressing a polyhistidine-tagged Rubisco large subunit. Plant Biotechnol J 2:389-399

Sakamoto W, Kindle K, Stern DB (1993) In vivo analysis of *Chlamydomonas* chloroplast *petD* gene expression using stable transformtion of beta-glucuronidase translational fusions. Proc Natl Acad Sci USA 90:497-501

Schöttler MA, Flügel C, Thiele W, Stegemann S, Bock R (2007a) The plastome-encoded PsaJ subunit is required for efficient photosystem I excitation, but not for plastocyanin oxidation in tobacco. Biochem J: in press (doi:10.1042/BJ20061573)

Schöttler MA, Flügel C, Thiele W, Bock R (2007b) Knock-out of the plastid-encoded PetL subunit results in reduced stability and accelerated leaf age-dependent loss of the cytochrome b6f complex. J Biol Chem 282:976-985

Schwenkert S, Umate P, Dal Bosco C, Volz S, Mlcochová L, Zoryan M, Eichacker LA, Ohad I, Herrmann RG, Meurer J (2006) PSBI affects the stability, function, and phosphorylation patterns of photosystem II assemblies in tobacco. J Biol Chem 281:34227-34238

Seki M, Shigemoto N, Sugita M, Sugiura M, Koop HU, Irifune K, Morikawa H (1995) Transient expression of ß-glucuronidase in plastids of various plant cells and tissues delivered by a pneumatic particle gun. J Plant Res 108:235-240

Serino G, Maliga P (1997) A negative selection scheme based on the expression of cytosine deaminase in plastids. Plant J 12:697-701

Serino G, Maliga P (1998) RNA polymerase subunits encoded by the plastid *rpo* genes are not shared with the nucleus-encoded plastid enzyme. Plant Physiol 117:1165-1170

Shepherd HS, Boynton JE, Gillham NW (1979) Mutations in nine chloroplast loci of *Chlamydomonas* affecting different photosynthtic functions. Proc Natl Acad Sci USA 76:1353-1357

Shikanai T, Endo T, Hashimoto T, Yamada Y, Asada K, Yokota A (1998) Directed disruption of the tobacco *ndhB* gene impairs cyclic electron flow around photosystem I. Proc Natl Acad Sci USA 95:9705-9709

Shikanai T, Shimizu K, Ueda K, Nishimura Y, Kuroiwa T, Hashimoto T (2001) The chloroplast *clpP* gene, encoding a proteolytic subunit of ATP-dependent protease, is indispensable for chloroplast development in tobacco. Plant Cell Physiol 42:264-273

Sidorov VA, Kasten D, Pang SZ, Hajdukiewicz PT, Staub JM, Nehra NS (1999) Technical advance: Stable chloroplast transformation in potato: use of green fluorescent protein as a plastid marker. Plant J 19:209-216

Sikdar SR, Serino G, Chaudhuri S, Maliga P (1998) Plastid transformation in *Arabidopsis thaliana*. Plant Cell Rep 18:20-24

Skarjinskaia M, Svab Z, Maliga P (2003) Plastid transformation in *Lesquerella fendleri*, an oilseed *Brassicacea*. Transgenic Res 12:115-122

Spörlein B, Streubel M, Dahlfeld G, Westhoff P, Koop HU (1991) PEG-mediated plastid transformation: a new system for transient gene expression assays in chloroplasts. Theor Appl Genet 82:717-722

Spreitzer RJ, Goldschmidt-Clermont M, Rahire M, Rochaix JD (1985) Nonsense mutations in the *Chlamydomonas* chloroplast gene that codes for the large subunit of ribulose bisphosphate carboxylase/oxygenase. Proc Natl Acad Sci USA 82:5460-5464

Sriraman P, Silhavy D, Maliga P (1998) The phage-type *PclpP-53* plastid promoter comprises sequences downstream of the transcription initiation site. Nucleic Acids Res 26:4874-4879

Staub JM (2002) Expression of Recombinant Proteins via the Plastid Genome. In: Vinci VA, Parekh SR (eds) Handbook of industrial cell culture: Mammalian, microbial and plant cells. Humana Press: Totowa, NJ

Staub JM, Maliga P (1992) Long regions of homologous DNA are incorporated into the tobacco plastid genome by transformation. Plant Cell 4:39-45

Staub JM, Maliga P (1993) Accumulation of D1 polypeptide in tobacco plastids is regulated via the untranslated region of the *psbA* mRNA. EMBO J 12:601-606

Staub JM, Maliga P (1994) Extrachromosomal elements in tobacco plastids. Proc Natl Acad Sci USA 91:7468-7472

Staub JM, Maliga P (1995a) Expression of a chimeric *uidA* gene indicates that polycistronic mRNAs are efficiently translated in tobacco plastids. Plant J 7:845-848

Staub JM, Maliga P (1995b) Marker rescue from the *Nicotiana tabacum* plastid genome using a plastid/*Escherichia coli* shuttle vector. Mol Gen Genet 249:37-42

Staub JM, Garcia B, Graves J, Hajdukiewicz PT, Hunter P, Nehra N, Paradkar V, Schlittler M, Carroll JA, Spatola L, Ward D, Ye G, Russell DA (2000) High-yield production of a human therapeutic protein in tobacco chloroplasts. Nat Biotechnol 18:333-338

Stern DB, Gruissem W (1987) Control of plastid gene expression: 3' inverted repeats act as mRNA processing and stabilizing elements, but do not terminate transcription. Cell 24:1145-1157

Stern DB, Higgs DC, Yang J (1997) Transcription and translation in chloroplasts. Trends Plant Sci 2:308-315

Stern DB, Bassi R, Herrmann RG, Wollman FA (2001) The chloroplast gene *ycf9* encodes a photosystem II (PSII) core subunit, PsbZ, that participates in PSII supramolecular architecture. Plant Cell 13:1347-1367

Su ZL, Qian KX, Tan CP, Meng CX, Qin S (2005) Recombination and heterologous expression of allophycyanin gene in the chloroplast of *Chlamydomonas reinhardtii*. Acta Biochim Biophys Sinica 37:709-712

Sugita M, Svab Z, Maliga P, Sugiura M (1997) Targeted deletion of *sprA* from the tobacco plastid genome indicates that the encoded small RNA is not essential for pre-16S rRNA maturation in plastids. Mol Gen Genet 257:23-27

Sugiura C, Sugita M (2004) Plastid transformation reveals that moss tRNA-CCG is not essential for plastid function. Plant J 40:314-321

Summer EJ, Schmid VH, Bruns BU, Schmidt GW (1997) Requirement for the H phosphoprotein in photosystem II of *Chlamydomonas reinhardtii*. Plant Physiol 113:1359-1368

Sun M, Qian K, Su N, Chang H, Liu J, Chen G (2003) Foot-and-mouth disease virus VP1 protein fused with cholera toxin B subunit expressed in *Chlamydomonas reinhardtii* chloroplast. Biotechnol Lett 25:1087-1092

Suzuki JY, Bauer CE (1992) Light-independent chlorophyll biosynthesis: involvement of the chloroplast gene *chlL* (*frxC*). Plant Cell 4:929-940

Suzuki H, Ingersoll J, Stern DB, Kindle KL (1997) Generation and maintenance of tandemly repeated extrachromosomal plasmid DNA in *Chlamydomonas* chloroplasts. Plant J 11:635-648

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

Svab Z, Maliga P (1993) High-frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene. Proc Natl Acad Sci USA 90:913-917

Swiatek M, Kuras R, Sokolenko A, Higgs D, Olive J, Cinque G, Muller B, Eichacker LA, Stern DB, Bassi R, Herrmann RG, Wollman FA (2001) The chloroplast gene *ycf9* encodes a photosystem II (PSII) core subunit, PsbZ, that participates in PSII supramolecular architecture. Plant Cell 13:1347-1367

Swiatek M, Regel RE, Meurer J, Wanner G, Pakrasi HB, Ohad I, Herrmann RG (2003a) Effects of selective inactivation of individual genes for low-molecular-mass subunits on the assembly of photosystem II, as revealed by chloroplast transformation: the *psbEFLJ* operon in *Nicotiana tabacum*. Mol Genet Genomics 268:699-710

Swiatek M, Greiner S, Kemp S, Drescher A, Koop HU, Herrmann RG, Maier RM (2003b) PCR analysis of pulsed-field gel electrophoresis-purified plastid DNA, a sensitive tool to judge the hetero-/homoplasmic status of plastid transformants. Curr Genet 43:45-53

Takahashi Y, Goldschmidt-Clermont M, Soen S-Y, Franzen LG, Rochaix JD (1991) Directed chloroplast transformation in *Chlamydomonas reinhardtii*: Insertional inactivation of the *psaC* gene encoding the iron-sulfur protein destabilizes photosystem I. EMBO J 10:2033-2040

Takahashi Y, Matsumoto H, Goldschmidt-Clermont M, Rochaix JD (1994) Directed disruption of the *Chlamydomonas* chloroplast *psbK* gene destabilizes the photosystem II reaction center complex. Plant Mol Biol 24:779-788

Takahashi Y, Rahire M, Breyton C, Popot JL, Joliot P, Rochaix JD (1996) The chloroplast *ycf7* (*petL*) open reading frame of *Chlamydomonas reinhardtii* encodes a small functionally important subunit of the cytochrome b6f complex. EMBO J 15:3498-3506

Tregoning JS, Nixon P, Kuroda H, Svab Z, Clare S, Bowe F, Fairweather N, Ytterberg J, van Wijk KJ, Dougan G, Maliga P (2003) Expression of tetanus toxin Fragment C in tobacco chloroplasts. Nucleic Acids Res 31:1174-1179

Tregoning J, Maliga P, Dougan G, Nixon PJ (2004) New advances in the production of edible plant vaccines: chloroplast expression of a tetanus vaccine antigen, TetC. Phytochem 65:989-994

Tungsuchat T, Kuroda H, Narangajavana J, Maliga P (2006) Gene activation in plastids by the CRE site-specific recombinase. Plant Mol Biol 61:711-718

van Bel AJ, Hibberd J, Prufer D, Knoblauch M (2001) Novel approach in plastid transformation. Curr Opin Biotechnol 12:144-149

Venkateswarlu K, Nazar RN (1991) Evidence for T-DNA mediated gene targeting to tobacco chloroplasts. Biotechnology 9:1103-1105

Viitanen PV, Devine AL, Khan MS, Deuel DL, Van Dyk DE, Daniell H (2004) Metabolic engineering of the chloroplast genome using the *Escherichia coli ubiC* gene reveals that chorismate is a readily abundant plant precursor for p-hydroxybenzoic acid biosynthesis. Plant Physiol 136:4048-4060

Walmsley AM, Arntzen CJ (2003) Plant cell factories and mucosal vaccines. Curr Opin Biotechnol 14:145-150

Watson J, Koya V, Leppla SH, Daniell H (2004) Expression of *Bacillus anthracis* protective antigen in transgenic chloroplasts of tobacco, a non-food/feed crop. Vaccine 22:4374-4384

Whitney SM, Andrews TJ (2001) Plastome-encoded bacterial ribulose-1,5-bisphosphate carboxylase/oxygenase (RUBISCO) supports photosynthesis and growth in tobacco. Proc Natl Acad Sci USA 98:14738-14743

Wirth S, Segretin ME, Mentaberry A, Bravo-Almonacid F (2006) Accumulation of hEGF and hEGF-fusion proteins in chloroplast-transformed tobacco plants is higher in the dark than in the light. J Biotechnol 125:159-172

Wurbs D, Ruf S, Bock R (2007) Contained metabolic engineering in tomatoes by expression of carotenoid biosynthesis genes from the plastid genome. Plant J 49:276–288

Xie Z, Merchant S (1996) The plastid-encoded *ccsA* gene is required for heme attachment to chloroplast c-type cytochromes. J Biol Chem 271:4632-4639
Xiong L, Sayre RT (2004) Engineering the chloroplast encoded proteins of *Chlamydomonas*. Photosynth Res 80:411-419
Ye GN, Daniell H, Sanford JC (1990) Optimization of delivery of foreign DNA into higher-plant chloroplasts. Plant Mol Biol 15:809-819
Ye GN, Hajdukiewicz PT, Broyles D, Rodriguez D, Xu CW, Nehra N, Staub JM (2001) Plastid-expressed 5-enolpyruvylshikimate-3-phosphate synthase genes provide high level glyphosate tolerance in tobacco. Plant J 25:261-270
Ye GN, Colburn SM, Xu CW, Hajdukiewicz PT, Staub JM (2003) Persistence of unselected transgenic DNA during a plastid transformation and segregation approach to herbicide resistance. Plant Physiol 133:402-410
Yukawa M, Tsudzuki T, Sugiura M (2005) The 2005 version of the chloroplast DNA sequence from tobacco (*Nicotiana tabacum*). Plant Mol Bio Rep 23:359-365
Zhang XH, Brotherton JE, Widholm JM, Portis AR Jr (2001a) Targeting a nuclear anthranilate synthase alpha-subunit gene to the tobacco plastid genome results in enhanced tryptophan biosynthesis. Return of a gene to its pre-endosymbiotic origin. Plant Physiol 127:131-141
Zhang XH, Portis AR, Wildholm JM (2001b) Plastid transformation of soybean suspension cultures. J Plant Biotechnol 3:39-44
Zhou YX, Lee MY, Ng JM, Chye ML, Yip WK, Zee SY, Lam E (2006) A truncated hepatitis E virus ORF2 protein expressed in tobacco plastids is immunogenic in mice. World J Gastroenterol 12:306-312
Zou Z, Eibl C, Koop HU (2003) The stem-loop region of the tobacco *psbA* 5'UTR is an important determinant of mRNA stability and translation efficiency. Mol Genet Genomics 269:340-349
Zoubenko OV, Allison LA, Svab Z, Maliga P (1994) Efficient targeting of foreign genes into the tobacco plastid genome. Nucleic Acids Res 25:3819-3824
Zubko MK, Zubko EI, van Zuilen K, Meyer P, Day A (2004) Stable transformation of petunia plastids. Transgenic Res 13:523-530

Golds, Timothy J
Research Centre Freising, Icon Genetics AG, Lise-Meitner-Straße 30, D 85354 Freising, Germany

Herz, Stefan
Research Centre Freising, Icon Genetics AG, Lise-Meitner-Straße 30, D 85354 Freising, Germany

Koop, Hans-Ulrich
Faculty of Biology, Department I, Botany, Ludwig-Maximilians-Universität München, Menzinger Straße 67, D 80638 München, Germany
koop@lmu.de

Nickelsen, Jörg
Faculty of Biology, Department I, Botany, Ludwig-Maximilians-Universität München, Menzinger Straße 67, D 80638 München, Germany

Index