Chloroplast Biogenesis: Control of Plastid Development, Protein Import, Division and Inheritance

Wataru Sakamoto,a Shin-ya Miyagishima,b and Paul Jarvisc

a Research Institute for Bioresources, Okayama University, Kurashiki, Okayama 710-0046, Japan
b Frontier Research System, RIKEN, Wako, Saitama 351-0198, Japan
c Department of Biology, University of Leicester, Leicester LE1 7RH, United Kingdom
Address correspondence to saka@rib.okayama-u.ac.jp

The Arabidopsis Book © 2008 American Society of Plant Biologists

First published on July 22, 2008: e0110. doi: 10.1199/tab.0110

The chloroplast is a multi-copy cellular organelle that not only performs photosynthesis but also synthesizes amino acids, lipids and phytohormones. The plastid also responds to environmental stimuli such as gravitropism. Biogenesis of chloroplasts is initiated from proplastids in shoot meristems, and involves a series of important events. In the last decade, considerable progress has been made towards understanding various aspects of chloroplast biogenesis at the molecular level, via studies in model systems such as Arabidopsis. This review focuses on two important aspects of chloroplast biogenesis, synthesis/assembly and division/transmission. Chloroplasts originated through endosymbiosis from an ancestor of extant cyanobacteria, and thus contain their own genomes. DNA in chloroplasts is organized into complexes with proteins, and these are called nucleoids. The synthesis of chloroplast proteins is regulated at various steps. However, a majority of proteins are synthesized in the cytosol, and their proper import into chloroplast compartments is a prerequisite for chloroplast development. Fundamental aspects of plastid gene expression/regulation and chloroplast protein transport are described, together with recent proteome analyses of the organelle. Chloroplasts are not de novo synthesized, but instead are propagated from pre-existing plastids. In addition, plastids are transmitted from generation to generation with a unique mode of inheritance. Our current knowledge on the division machinery and the inheritance of plastids is described.

INTRODUCTION

Photosynthesis is undoubtedly the most important process associated with plant life. It converts light energy, captured by pigment-containing light-harvesting antenna, into chemical energy that ultimately sustains plant growth. In plants, photosynthesis occurs exclusively in the chloroplast, the organelle derived through endosymbiosis from a relative of present-day cyanobacteria. It was van Leeuwenhoek in the 1670s who first scientifically described chloroplasts, as green globules in Spirogyra, without completely understanding them (Wise and Hoober, 2006). A chloroplast is defined as a particular type of plastid. These organelles are not synthesized de novo, but are instead propagated from pre-existing plastids via a division process. This division theory was initially hypothesized based on independent cytological analyses performed by Schimper and Meyer in the 1880s. Formation of chloroplasts is initiated from proplastids, an undifferentiated plastid type that is present in the shoot apical meristem (Figure 1A). Responding to light, proplastids develop grana, which are stacks of thylakoid membranes, where the machineries of light harvesting, electron transfer and ATP synthesis are formed. Chloroplasts are not only the site of photosynthesis, but are also responsible for the storage of starch and oil compounds, and for the synthesis of amino acids, lipids and phytohormones. Furthermore, plastids play an active role in environmental sensing, including gravity perception, stomatal opening and closure, and response to pathogen infection.

The first textbook of plastid biology was published in the 1960s by Kirk and Tilney-Bassett (second edition in 1978). It is somewhat surprising that, at that time, very little was stated regarding Arabidopsis (with the exception of the mention of a variegation mutant). Four decades later, however, we have greatly expanded upon the knowledge written in that initial textbook, most notably by utilizing this tiny plant as a model system, since it is highly suitable for modern molecular genetics and systems biology. Here, we focus on several key aspects of plastid biology that are fundamentally important to sustain the organelle’s life cycle, including protein import, division, and inheritance. We place emphasis, in principal, on molecular-genetic studies that are primarily related to work in the Arabidopsis model system. We also discuss gene expression in plastoplasts, its regulatory network, and possible signals that are exchanged between the chloroplast and the nucleus. Due to limitations of space, the biogenesis of the photosynthetic apparatus and metabolic pathways are beyond the scope of this review. Plastid biogenesis has also been described in detail in several textbooks and review articles (Kirk and Tilney-Bassett, 1978; Leister, 2003; Daniell and Chase, 2004; Möller, 2005; Wise and Hoober, 2006; Lopez-Juez, 2007).
GENERAL VIEW ON THE ORIGIN, CONTINUITY AND DIFFERENTIATION OF CHLOROPLASTS

Origin and Continuity of Chloroplasts

Mitochondria and chloroplasts are the descendants of serial endosymbiotic events (Cavalier-Smith, 2004). Mitochondria arose first from an α-proteobacterial ancestor that had been engulfed and enslaved by a primitive eukaryotic host. Chloroplasts arose later (around 1-1.5 billion years ago) from a cyanobacterial ancestor engulfed by a eukaryote in which mitochondria had already been established (Cavalier-Smith, 2004; Reyes-Prieto et al., 2007). Most of the bacterial genes were transferred to the nuclear genome or lost, but both modern organelles nevertheless retain metabolic activities, genetic mechanisms, and protein transport complexes that clearly reflect their prokaryotic origins. A single endosymbiosis of a cyanobacterial ancestor gave rise to the chloroplasts of the Glaucophyta (glaucophyte algae), Rhodophyta (red algae) and Viridiplantae (green algae and land plants). In addition to these groups, many other eukaryotic groups, such as stramenopiles (brown algae), euglenids, dinoflagellates and malarian parasites, also have chloroplasts or non-green plastids. The latter groups acquired plastids via secondary endosymbioses of red or green algae, in which their non-photosynthetic eukaryotic progenitors engulfed and enslaved eukaryotic algae (Cavalier-Smith, 2004; Reyes-Prieto et al., 2007).

In unicellular algae, chloroplasts are usually the only type of plastid present. This is consistent with the fact that vegetative cells of cyanobacteria remain blue-green and photosynthetic throughout their life cycle. Thus, from an evolutionary standpoint, the green chloroplast is the origin of the several different types of plastids that are present in land plants. In contrast with unicellular algae, land plants have evolved systems for plastid differentiation, enabling the formation of plastid types specialized for activities other than photosynthesis (Mullet, 1988; Lopez-Juez and Pyke, 2005; Figure 1). In vascular plants, all plastids including chloroplasts are derived from small, non-green proplastids in meristematic cells. Proplastids normally originate maternally during the formation of plant zygotes, and are transmitted from generation to generation (Mullet, 1988; Lopez-Juez and Pyke, 2005; see the section on Inheritance of Plastids).

Structure of Chloroplasts

Among several types of plastid, the best characterized is the chloroplast in vascular plants. The shape and structure of chloroplasts vary depending on the species, tissue and environmental conditions. It should be noted that while some features are common among species and tissues, some other features have evolved in and are specific to vascular plants. In mature leaf cells, chloroplasts are usually lens-shaped, 5-10 μm in diameter and 2-4 μm in thickness (Figure 1). Each leaf cell usually contains 20 to 100 chloroplasts (Mullet, 1988; Lopez-Juez and Pyke, 2005).

Chloroplasts and all other plastid types are surrounded by two membranes, the outer and the inner envelope membranes. In addition to these membranes, chloroplasts have the thylakoids. Thus, chloroplasts have three membrane systems and three aqueous compartments: the intermembrane space (between two envelopes), the stroma (surrounded by the inner envelope), and the thylakoid lumen (surrounded by thylakoid membrane). The inner envelope membrane and thylakoid membrane were descended from the plasma membrane and thylakoid membrane, respectively, of the engulfed cyanobacterium, whereas the origin of the outer envelope membrane is less clear. The presence of galactolipids and carotenoids and the prokaryotic origin of some outer envelope proteins suggest a link with the cyanobacterial outer membrane (Reumann et al., 2005; Inoue, 2007). In contrast, other lipids in the outer envelope membrane suggest a eukaryotic origin (Douce and Joyard, 1990).

The envelope membranes are the sites for lipid biogenesis (Douce and Joyard, 1990; Joyard et al., 1998), the translocation of nucleus-encoded proteins into plastids from the cytosol (Reumann et al., 2005), and the exchange of molecules across the membranes (Weber et al., 2005). Plastids often have tubular extensions of the two envelope membranes called stromules, which interconnect different plastids (Kwok and Hanson, 2004).

The shape of the thylakoids varies depending upon the lineage and tissue. In cyanobacteria and red algal chloroplasts, phycobilisomes uniformly attach to the outside of long thylakoids, and the thylakoids are arrayed at regular intervals. In green algae and plants, which do not have phycobilisomes, the structure of thylakoids is more complex. Thylakoids extend parallel to the chloroplast main axis; some are short, disc-shaped and organized into stacks called grana, while these grana are interconnected by long, stromal thylakoids (Dekker and Boekema, 2005; Figure 1). The thylakoids appear as discrete units under the transmission electron microscope, but actually form an interconnected compartment, enclosing a single lumen in three dimensions. Photosystems (PSI and PSII) exist on the thylakoid membrane. PSII is limited to granal membranes not in contact with the stroma, while PSI exists exclusively in the thylakoids exposed to the stroma (Dekker and Boekema, 2005). Usually, lipoprotein particles called plastoglobules are associated with the thylakoid membranes (Bréhélin et al., 2007; Figure 1).

The stroma corresponds to the cytosol of the original endosymbiont. It contains all the enzymes needed to carry out the carbon reactions of photosynthesis, and therefore contains starch granules. Nucleoids and ribosomes also exist in the stroma. Each chloroplast contains many nucleoids which are attached to the envelope and thylakoid membranes. The number and location of nucleoids changes depending on the type of plastid and species (Sakai et al., 2004).

Other Types of Plastid and Differentiation

The term plastid originated from the organelle’s plasticity. The plasticity observed in vascular plants has evolved by the acquisition of mechanisms for the activation or inactivation of particular functions of chloroplasts, according to the requirements of specialized tissues.

In vascular plants, proplastids in meristematic tissues differentiate into several different types of plastid depending on the functions which are needed in particular tissues: yellow etioplasts in dark grown leaves, amyloplasts for starch storage, chromoplasts for pigment synthesis, elaioplasts for storing lipids, and leucoplasts for monoterpane synthesis (Mullet, 1988; Lopez-Juez and Pyke,
The differentiation of plastids is reversible and sequential. Therefore, there are spectra of intermediates and even differences within the same subtype. For example, photosynthetic activities are partitioned between mesophyll chloroplasts and bundle sheath chloroplasts in C₄ plants such as maize. Chloroplasts in mesophyll cells have developed granal stacks, while bundle sheath chloroplasts lack grana, are PSII-depleted, and perform most of the reactions of the Calvin cycle (Sheen, 1999).

The best-studied transition is the development of chloroplasts from proplastids (Mullet 1988; Leon et al., 1998). Proplastids in meristematic tissues are colorless, and are usually 0.2-1.0 μm in diameter with very few internal membrane vesicles (Figure 1).
These vesicles are the precursors of the thylakoids, and perhaps are derived from parental proplastids or from invagination of the inner envelope (Benning et al., 2006; Aseeva et al., 2007). Each meristematic cell usually contains 10-20 proplastids. During the development and enlargement of mesophyll cells, the number of plastids per cell increases to 50-100, and plastid size also increases greater than 100-fold. The increase in plastid volume results in part from the development of thylakoid membranes and the accumulation of proteins and lipids required for photosynthesis (Mullet 1988; Leon et al., 1998).

**THE PLASTID GENOME**

In the 1980s, efforts aimed towards the complete sequencing of plastid genomes were made in several plant species, including liverwort, tobacco and rice (Sugiura, 2003). This scientific venture was notably well ahead of the initiation of Arabidopsis nuclear genome sequencing project. To date, the plastid genome sequences of 122 different species have been determined (Organelle Genome Resources, see Table 1). Comparative analyses have been made between the nuclear and chloroplast genomes of Arabidopsis, and the genomes of cyanobacteria and yeast, and in this way the evolution of the genomes via endosymbiotic events has been elucidated (Martin et al., 2002). Here, we place emphasis on the cytological structure and dynamic behavior of plastid genomes during chloroplast biogenesis. Readers are encouraged to refer to additional reviews which focus more on plastid genome structure and evolution (Sugiura, 1992; Maier and Schmitz-Linneweber, 2004). The plastid genome is often called the ‘plastome’, but we do not use that term and instead refer to either ptDNA or the plastid genome.

### Table 1. Databases and websites useful for plastid research

<table>
<thead>
<tr>
<th><strong>Genome</strong></th>
<th><strong>Proteome</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chloroplast Genome DB (Chloroplast Genome Database)</td>
<td>PPDB (Plastid/Plant Proteome Database)</td>
</tr>
<tr>
<td>GObase (The Organelle Genome Database)</td>
<td>SUBA (Arabidopsis Subcellular Proteomic Database)</td>
</tr>
<tr>
<td>Organelle Genome Resources</td>
<td>plpprot (Plastid Proteome Database)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Subcellular localization</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>TargetP (Prediction of organelle targeting sequences)</td>
</tr>
<tr>
<td>PSORT (Prediction of protein sorting signals and localization sites in amino acid sequences)</td>
</tr>
<tr>
<td>PREDOTAR (Prediction of organelle targeting sequences)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Others</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>The Plant Organelles Database</td>
</tr>
<tr>
<td>CyanoBase (The Genome Database for Cyanobacteria)</td>
</tr>
</tbody>
</table>

These vesicles are the precursors of the thylakoids, and perhaps are derived from parental proplastids or from invagination of the inner envelope (Benning et al., 2006; Aseeva et al., 2007). Each meristematic cell usually contains 10-20 proplastids. During the development and enlargement of mesophyll cells, the number of plastids per cell increases to 50-100, and plastid size also increases greater than 100-fold. The increase in plastid volume results in part from the development of thylakoid membranes and the accumulation of proteins and lipids required for photosynthesis (Mullet 1988; Leon et al., 1998).

Plastid differentiation is largely under nuclear control. The nucleus encodes most of the proteins required for metabolic functions in plastids, and many of those needed for plastid gene expression. Nuclear regulation involves multiple aspects, including the selective targeting of nucleus-encoded plastid proteins, the replication of plastid DNA (ptDNA), and the control of transcription and translation within plastids (Mullet 1988; Leon et al., 1998). In addition to such anterograde control, proper plastid differentiation involves retrograde signaling through which the functional and developmental state of the organelle regulates the expression of nuclear genes (Nott et al., 2006; Woodson and Chory, 2008).

### Basic Structure

In Arabidopsis, the complete sequence of the ptDNA was reported in 1999 by a group at the Kazusa Institute (Sato et al., 1999). The Arabidopsis plastid genome consists of a circular DNA of 154 kb in length. It retains a highly conserved genome structure consisting of a pair of inverted repeats (26 kb) that split the circular genome into two parts, termed the small (18 kb) and large (84 kb) single copy regions. The Arabidopsis plastid genome contains 45 RNA-coding genes and 87 protein-coding genes. The functional ptDNA gene products are principally involved in: tran-
scription (RNA polymerase), translation (ribosomal and transfer RNAs, ribosomal proteins), photosynthetic electron transfer (subunits of PSI, PSII, the cytochrome b6/f complex and NAD(P)H dehydrogenase [NDH]), and photosynthetic metabolism (subunits of ATP synthase and Rubisco). Exceptions to these aforementioned examples are two photosynthesis-unrelated housekeeping genes, accD and clpP1, which encode subunits of acetyl-CoA carboxylase and the Clp (Caseinolytic protease) protease, respectively.

Gene arrangements within the ptDNA are also well conserved between the different species (Sugiura, 1992; Maier and Schmitz-Linneweber, 2004). Since plant chloroplasts are descendants from an ancestral cyanobacterium, many genes have retained prokaryotic features and are organized as operons and are co-transcribed. Processing events of polycistronic transcripts is sometimes complex and leads to an accumulation of various RNA molecules. One of the best characterized examples is the psbB operon, in which five genes are co-transcribed and give rise to many poly- and monocistronic RNA molecules (Barkan, 1988; Westhoff and Herrmann, 1988). Genes for rRNAs are also encoded by operons. However, despite the extensive characterization of plastid genomes and their expression, very little is known regarding replication. Several investigations demonstrated that the replication of ptDNA involves an enzyme similar to bacterial DNA polymerase I (Pol-I). In Arabidopsis, two putative Pol-I genes have been identified whose gene products are targeted to chloroplasts (Mori et al., 2003). DNA Pol-I has also been studied in other species, such as rice and tobacco (Kumura et al., 2002; Ono et al., 2007). Collectively, these data suggest that the replication system is shared between chloroplasts and mitochondria, through the dual-targeting of the relevant proteins.

How Plastid DNA Exists in Plastids – Plastid Nucleoids

A single cell contains a variable number of plastids, in which ptDNAs exist in multiple copies in the stroma. For example, a mesophyll cell in Arabidopsis contains approximately 20-200 chloroplasts, and anywhere from 10 to 500 copies of ptDNA per chloroplast (Fuji et al., 1994; Pyke and Leech, 1994). As a result, the copy number of ptDNAs per cell is highly variable. Nonetheless, total DNA isolated from green leaf tissues may contain approximately 20% of ptDNA.

Several questions arise concerning the spatial organization of the multiple plastid genome copies relative to the stroma. Another question is whether or not ptDNAs are capable of forming a complex with proteins similar to bacterial chromosomes. It is unlikely that homologues of a bacterial DNA-binding protein, HU, are encoded in higher plant genomes. However, despite the extensive characterization of plastid genomes and their expression, very little is known regarding replication. Several investigations demonstrated that the replication of ptDNA involves an enzyme similar to bacterial DNA polymerase I (Pol-I). In Arabidopsis, two putative Pol-I genes have been identified whose gene products are targeted to chloroplasts (Mori et al., 2003). DNA Pol-I has also been studied in other species, such as rice and tobacco (Kumura et al., 2002; Ono et al., 2007). Collectively, these data suggest that the replication system is shared between chloroplasts and mitochondria, through the dual-targeting of the relevant proteins.

Is The Amount of Plastid DNA Variable During Chloroplast Differentiation?

Dynamic changes in the morphology of nucleoids implies that DNA contents in each plastid or chloroplast may differ. Quantification of plastid DAPI signals, estimated by using a video-intensified photo-counting microscope, confirmed that each nucleoid contains ~10 copies of ptDNA (Sato et al., 2003; Sakai et al., 2004). These observations also determined that the overall DNA level per plastid increases 5-fold during chloroplast development (Kuroiwa, 1991; Fuji et al., 1994). Furthermore, within mature leaves containing fully-expanded mesophyll cells, there are occasionally few or no detectable DAPI signals (Oldenburg and Ben-dich, 2004; Rowan et al., 2004). Based on these observations, it is suggested that the amount of ptDNA per organelle decreases within mature, old and/or senescing chloroplasts. While the fluctuations in ptDNA levels would be expected to occur to an extent that is detectable via multiple experimental methods, the reduction in old leaves seems to reveal contradictory results. Using fluorescence microscopy, Rowan et al. (2004) showed that ptDNA levels per chloroplast vary significantly, but decrease in older Arabidopsis leaves. Conversely, Li et al. (2006) reported that this reduction was never observed when a DNA hybridization method was employed. It appears that the contradiction is due to technical...
limitations for the quantitative detection of DNA. Although DNA blot analysis may be a reliable method, it is not sensitive to subtle changes. Whereas DAPI stain is a very sensitive, qualitative method for visualizing DNA, it may not result in an accurate estimation of DNA quantity; in particular, a lack of signal should not be taken to indicate that there is no DNA. Hopefully an accurate, novel method that can reliably measure ptDNA content in each chloroplast can be developed in the future. In contrast to leaves, our current studies indicate that ptDNA levels drastically decline during pollen maturation.

PLASTID GENE EXPRESSION AND REGULATION MECHANISMS

Historically, gene expression in chloroplasts has been extensively studied using in vitro systems. These studies have enabled scientists to identify the basic machineries involved in transcription, RNA processing and maturation, and translation (Daniell and Chase, 2004; Cahoon et al., 2006). Such biochemical analyses, which have been conducted since the 1980s, have primarily utilized spinach as a model system. Meanwhile, molecular-genetic analyses in the unicellular alga, Chlamydomonas, and transposon mutagenesis in maize have pioneered methods for deciphering the numerous regulatory factors involved, which mostly act post-transcriptionally and in a gene-specific manner (Barkan and Goldschmidt-Clermont, 2000; Rochaix, 2006). In addition, recent systematic forward- and reverse-genetic approaches in Arabidopsis have enabled us to draw a blueprint of gene expression networks in chloroplasts (Leister, 2003; Leister and Pesaresi, 2005). Here, we briefly overview plastid gene expression, placing particular emphasis on transcriptional and post-transcriptional regulation.
Transcriptional Regulation

RNA Polymerases. Two types of RNA polymerase are known to promote transcription in chloroplasts (Kanamaru and Tanaka, 2004; Shiina et al., 2005; Cahoon et al., 2006) (Figure 3). One is a bacterial type polymerase, as might be expected considering the cyanobacterial origin of chloroplasts, and is indeed encoded by the plastid genome; this polymerase is termed PEP (plastid-encoded polymerase). PEP is a holoenzyme consisting of four subunits (α, β, β′1 and β′2) encoded by the plastidic genes, rpoA, rpoB, rpoC1 and rpoC2, respectively. Eubacterial polymerases such as PEP require an additional factor (‘sigma’) which facilitates promoter recognition (see below). To date, no sigma-like factors have been shown to be encoded in chloroplast genomes. Instead, a nuclear gene coding for a sigma factor was reported (Isono et al., 1997), and was later shown to represent a small gene family in higher plants (Tanaka et al., 1997). In Arabidopsis, six sigma factors (SIG1 to SIG6) are localized in chloroplasts (Shiina et al., 2005; Cahoon et al., 2006).

The second polymerase is a bacteriophage-type polymerase, consisting of a single subunit and bearing similarity to a mitochondrial polymerase (Hedtke et al., 1997). This polymerase is encoded in the nuclear genome, and hence is termed NEP (nucleus-encoded polymerase). The presence of NEP activity in chloroplasts has been suggested for a long time. This hypothesis was substantiated by using transgenic tobacco plants in which one of the PEP subunits was inactivated by chloroplast transformation (Allison et al., 1996); the mutant plants clearly accumulated chloroplast transcripts that were attributable to NEP activity (Hajdukiewicz et al., 1997). In Arabidopsis, three genes encoding NEP have been identified: RpoT3 (or RpoTp) for plastids, RpoT1 (RpoTm) for mitochondria, and RpoT2 (RpoTmp) for both organelles. As a consequence, the transcription of ptDNA is driven by three different polymerases. Presently, any differences between the roles of RpoT2 and RpoT3 remain unclear.

By analogy with eubacterial σ70 promoters, PEP in principle recognizes a conserved promoter motif consisting of -10 and -35 sequence elements. In contrast, NEP recognizes an AT-rich promoter sequence similar to mitochondrial consensus promoters.
of protein synthesis does not always parallel that of transcription (Gruissem et al., 1988). Since chloroplast genes are transcribed as operons, the generation of various mono- and poly-cistronic RNA molecules was presumed to be regulated by complex mechanisms, likely at the level of RNA processing. The importance of post-transcriptional regulation has been reinforced by pioneering genetic works in model organisms such as Chlamydomonas and maize, in which many post-transcriptional factors were identified in a gene-specific manner (Barkan and Goldschmidt-Clermont, 2000; Rochaix, 2006). Those works were accomplished through the characterization of non-photosynthetic mutants, and, for example, maize mutants showing high chlorophyll fluorescence (hcf). In Arabidopsis, the systematic characterization of hcf and other chlorophyll fluorescence mutants in several laboratories has enabled the identification of some novel factors acting on chloroplast gene expression at the post-transcriptional level (Meurer et al., 1996; Shikanai et al., 1999).

The isolation of photosynthetic mutants by chlorophyll fluorescence is based on the fact that the emitted fluorescence reflects the pigment’s photochemical status. If photosynthetic electron flow occurs normally, then the excitation energy of chlorophyll molecules after illumination is used to oxidize water and to drive the passage of electrons through the photosystems and ultimately to NADP⁺. Any mutation that blocks proper electron transfer causes over-reduction of the plastoquinone pool. As a result, the excess excitation energy in chlorophyll cannot be used for photosynthesis and is instead emitted as fluorescence. Thus, tracing the quenching of chlorophyll fluorescence allows a high-throughput screening approach for identifying mutants of interest. A collection of such mutants, showing various fluorescence patterns including hcf, was therefore a good resource for the identification of nucleus-encoded factors that mediate post-transcriptional control in a gene-specific manner (Figure 3). In fact, most of the nuclear factors that were identified as regulators of one or a few chloroplast genes turned out to play roles in RNA processing (including RNA splicing, endo-processing and editing) or translation. Due to limitations of space, we can only refer to the few examples listed below.

The psbB operon is one of the most complex operons conserved in the chloroplast genomes of many higher plants, since it co-transcribes five genes (psbB-psbT-psbH-psbE-psbD) (Barkan, 1988; Westhoff and Herrmann, 1988). Once transcribed, at least three steps – RNA splicing (in the cases of petB and petD), processing (to generate the psbB-psbT, psbH, petB and petD RNA species), and editing (in the case of petB) – are required to generate transcripts that are competent for translation. Forward-genetic analyses have identified three factors involved in the processing of this transcript: HCF107 acts on the cleavage between psbT and psbH (Sané et al., 2005), while HCF152 acts between psbH and petB (Meierhoff et al., 2003) and CRP1 (chloroplast RNA processing1) acts between petB and petD (Fisk et al., 1999). Molecular cloning of the corresponding loci revealed that HCF152 and CRP1 encode proteins that belong to a pentatricopeptide repeat (PPR) protein family (see below and Figure 3). In a separate study, Shikanai’s group characterized a category of mutants that exhibit defects in the NDH complex (Yamazaki et al., 2004). Among these, CRR4 (chlororespiratory reduction) has been shown to encode a PPR protein that plays a role in genespecific RNA editing (Kotera et al., 2005). RNA editing is a mech-

**Post-transcriptional Regulation**

Important roles of post-transcriptional control in chloroplasts have been implicated since the 1980s, when in vivo analyses by run-on transcription and protein pulse labeling revealed that the rate
an organism by which a specific cytidine residue in a primary transcript is edited to uridine. In Arabidopsis chloroplasts, 19 editing sites are known, and CPR4 is involved in one of them, generating an initiation codon (ACG to AUG) in the ndhD gene. In addition, some PPRs affect translation rather than RNA processing. Together, these observations (particularly the discovery of PPR proteins) provide a great body of functional evidence that highlights the importance of post-transcriptional regulation in plastids, as was implicated in earlier studies.

The presence of the PPR family was first implicated via a bioinformatics approach (Lurin et al., 2004; Saha et al., 2007). The PPR domain is characterized by a signature PPR motif that consists of a degenerate 35 residue sequence. Depending on the protein, this motif exists in tandem repeats ranging in number from 2 to 27. Genes encoding PPRs are predominantly found in plant genomes, although a few PPR proteins have been detected in Drosophila and C. elegans. In fact, bioinformatic analysis revealed that 450 and 655 PPR proteins are present in Arabidopsis and rice, respectively. Most of these seem to have N-terminal targeting signals that are predicted to target them to chloroplasts or mitochondria. PPR repeat motifs are predicted to form a structure that serves as a binding site for a single-stranded RNA molecule. In rare cases, PPR domains are suggested to interact with DNA. The PPR family is therefore considered to control organellar gene expression in mitochondria as well as chloroplasts. The abundance of PPRs in plants raises the intriguing question of how this regulatory system has been acquired in the nuclear genome.

Besides PPR proteins and other gene-specific factors, general components in RNA processing and translation were also identified by forward- and reverse-genetic approaches. In chloroplast genomes, ribosomal RNA genes are clustered as an operon and are co-transcribed. After transcription, their respective gene messages are processed. Two types of exonucleases—polynucleotide phosphorylase (PNPase), and RNase R—were shown to participate in this process (Kishine et al., 2004). PNPase seems to simultaneously act on 3′-end maturation of certain mRNAs and tRNA turnover (Walter et al., 2002; Bollenbach et al., 2005). Nuclear factors involved in the splicing of group II introns have been identified (Barkan, 2004; Asakura and Barkan, 2006). Genes encoding ribosomal proteins are found in both the nuclear and plastid genomes. A viable mutant lacking one of the nucleus-encoded plastidic ribosomal proteins has been reported (Pesaresi et al., 2006a). In addition, general factors for translation (sharing similarities with prokaryotic factors) have been identified in Arabidopsis chloroplasts, including: translation initiation factor 2 (cplF2) (Miura et al., 2007), elongation factor G (cplEF-G) (Albrecht et al., 2006), and peptide release factors (cpRF) 1 and 2 (Meurer et al., 2002; Motohashi et al., 2007). Complete loss of such factors results in an embryo-lethal or albino phenotype, but several mutant lines with leaky mutations have been reported. Chloroplast genomes in higher plants seem to contain all the tRNAs necessary for translating chloroplast mRNAs. In contrast, certain tRNAs are missing in mitochondria, and thus must be imported from the cytosol. All of the enzymes required for the aminoacylation of each tRNA are imported from the cytosol. Interestingly, most of the aminoacyl synthases seem to be shared between chloroplasts and mitochondria (Duchêne et al., 2005).

Retrograde Signaling

Because of the partitioning of genetic information, chloroplast development and functions necessarily require input from two different genomes. For example, the multiprotein complexes of photosynthesis are mixtures of nucleus- and chloroplast-encoded subunits; to ensure their proper, stoichiometric assembly, and enable their reorganization in response to developmental or environmental cues, the activities of the nuclear and chloroplast genomes must be coordinated through intracellular signaling.

The pre-eminence of the nucleus in this inter-organellar exchange is beyond doubt. The import of nucleus-encoded proteins itself constitutes a massive flow of information (Jarvis, 2008). Moreover, as discussed earlier, numerous nucleus-encoded regulators mediate stringent, predominantly post-transcriptional control of the expression of chloroplast genes (Rochaix, 2006). Nevertheless, it is also clear that signals emitted by chloroplasts (so-called “retrograde” signals; Figure 3) have profound effects on events in the nucleus (Nott et al., 2006; Pesaresi et al., 2007; Woodson and Chory, 2008). Redox balance within the photosynthetic electron transport (PET) chains, the accumulation of reactive oxygen species (ROS), and the perturbation of plastid gene expression or chlorophyll biosynthesis, all influence nuclear gene expression.

Tetrapyrrole Signaling and gun Mutants. Arabidopsis mutants with defects in retrograde signaling were identified in a reporter-based, forward-genetic screen (Susek et al., 1993). The screening strategy hinged on observations that the transcription of nuclear genes for chloroplast proteins is strongly repressed if chloroplast development is blocked through photooxidative damage (Oelmuller, 1989). The herbicide norflurazon inhibits the formation of photoprotective carotenoids, leading to the photodestruction of the chloroplast interior whilst leaving the rest of the cell intact. Under these circumstances, genes such as those encoding light-harvesting chlorophyll a/b-binding proteins (Lhcb) are strongly repressed. Fusion of an Lhcb promoter to a selectable-marker gene enabled the identification of mutants no longer able to repress Lhcb expression upon norflurazon treatment.

Five independent genomes uncoupled (gun) mutants were identified, four of which (gun2-gun5) interfere with the chlorophyll biosynthetic pathway (Mochizuki et al., 2001; Larkin et al., 2003). Analyses of these mutants culminated in the identification of the tetrapyrrole intermediate, Mg-protoporphyrin-IX, as a key regulator of one particularly important retrograde signaling pathway (Strand et al., 2003). Wild-type plants accumulate Mg-protoporphyrin-IX following norflurazon treatment, triggering a signaling response, whereas the gun2-gun5 mutants are unable to build up sufficient quantities of the intermediate.

The effect of gun1 is somewhat different from that of the other gun mutations (Vinti et al., 2000; Mochizuki et al., 2001). The gun1 protein is a chloroplast-localized PPR domain protein that binds to DNA and localizes at sites of active transcription, but its exact function remains uncertain (Koussevitzky et al., 2007). While the gun2-gun5 mutations block the Mg-protoporphyrin-IX signaling pathway specifically, the effect of gun1 extends to the plastid gene expression-dependent and redox-related pathways as well. This suggests that GUN1 acts downstream in all of these pathways, and that its role is to integrate information from multiple sources.

An ACGT motif was found to be substantially overrepre-
sent in the promoters of retrograde-regulated genes (Kous-
sevitzky et al., 2007). This motif forms the core of the abscisic
acid (ABA) response element, as well as of the light-responsive
G-box, suggesting possible convergence of retrograde and ABA
signaling pathways. When this possibility was investigated, the
ABA-insensitive 4 (abi4) mutant was found to be phenotypically
similar to gun1, with defects in all tested retrograde pathways.
The ABI4 protein is an AP2-type nuclear transcriptional regulator.
A G-box element in the Lhcb promoter, termed CUF1, was pre-
viously shown to mediate responses to plastid signals, as well as
light induction (Strand et al., 2003). Two partially-overlapping bind-
ing motifs exist in CUF1, suggesting a model in which ABI4 and a
light-responsive G-box-binding factor (GBF) compete for access to
the promoter (Koussevitzky et al., 2007). This would explain how
negative retrograde signals from plastids (mediated by ABI4) are
able to override positive light-induced signals (mediated by GBF).
However, not all retrograde-regulated promoters contain both mo-
tifs, implying that this may not be a universal mechanism.

A more complex picture of retrograde signaling emerges upon
consideration of transcriptome responses to genetic lesions or en-
vironmental factors that impinge on chloroplast function. While
some treatments or conditions trigger the en masse up- or down-
regulation of nuclear genes for chloroplast proteins (suggesting
the existence of a “master switch” that perhaps corresponds to
ABI4), many others cause more complex patterns of gene ex-
pression involving the simultaneous up- and down-regulation of
different sets of genes (Biehl et al., 2005). One example is pro-
vided by the ppi f mutant (see section on Protein Transport Sys-
tems), which lacks the chloroplast protein import receptor, atToc33
(Figure 5); this mutation triggers the down-regulation of nuclear
photosynthetic genes specifically, suggesting that retrograde sig-
naling mechanisms exist to prevent the futile expression of pro-
teins not able to reach their final destination (Kubis et al., 2003).

**Plastid Gene Expression and Redox Signaling.** Inhibition of
plastid gene expression, either through the use of genetic muta-
tions or the treatment of plants with inhibitors (e.g. lincomycin),
also triggers the repression of nuclear genes for chloroplast pro-
teins (Gray et al., 2003; Pesaresi et al., 2006b). This response is
light-independent, and is especially (but not exclusively) important
in young seedlings. While plastid gene expression signals are
thought to be distinct from tetrapyrrole signals, the absence of a
normal response in the gun1 mutant implies that this pathway
eventually converges with that triggered by Mg-protoporphyrin-IX.
The identity of signaling intermediates (other than GUN1) in the
plastid gene expression pathway remain elusive.

Redox poise within the photosynthetic machinery and ROS ac-
cumulation both impact on the expression of nuclear genes. The
redox state of the plastoquinone pool (which links PSII with the
cytochrome b6/f complex) is thought to be a key determinant of
PET-mediated retrograde signaling, but other positions along the
PET chain are also responsive (Escoubas et al., 1995; Shao et
al., 2006). During excessive stimulation of the photosynthetic ma-
achinery, reduced forms of PET components predominate, and
ROS formation occurs. These ROS include singlet oxygen, hydro-
gen peroxide and superoxide, which seem to trigger different re-
sponse pathways (Karpinski et al., 1999; Laloi et al., 2007). Retrograde
signaling via the various redox- and ROS-dependent pathways (e.g. under strong light) causes the down-regulation of photosynthetic genes and the up-regulation of anti-oxidative de-
ference genes in the nucleus. Signaling components in these path-
ways have been identified (Lee et al., 2007).

**Plastid Proteases**

Chloroplast proteins are either synthesized in chloroplasts or
transported from the cytosol. Once delivered and/or assembled in
a proper sub-compartment, they must be maintained by ‘protein
quality control’, in which proteases play an essential role. Thus, reg-
ulated proteolysis can be regarded as fine-tuning at the last step of
gene expression. In addition, chloroplastic proteins become sensi-
tive to the inevitable photooxidative damage, which is often caused
by excess light energy. For example, the D1 protein of the PSII re-
action center is considered to be a main target of photodamage
(Aro et al., 1993). As a consequence, D1 is turned over very rapidly
by the repair cycle, in which some proteases (see below) were
shown to play very important roles (Nixon et al., 2005). Also, some
proteases processively degrade partially assembled and mis-folded
proteins (Adam et al., 2006; Sakamoto, 2006). Furthermore, pre-
proteins from the cytosol (containing N-terminal transit peptides)
or precursor proteins synthesized in stroma (e.g. D1 and cy-
tochrome f) undergo maturation by endoproteolytic processing en-
zymes. Proteases involved in these processing and maturation
events have been identified, including stromal processing pepti-
dase (SPP), thylakoid processing peptidase (TPP), transit peptide-
degrading zinc-metalloprotease (termed Zn-MP or PreP), the
cytobtex-terminal peptidase of D1 (CtpA), and the type I signal pep-
tidase (SPaseI for Toc75 maturation).

As exemplified by the ubiquitin-dependent degradation pathway
through the 26S proteasome in the cytosol, proteolysis is generally
accepted to be an important regulatory pathway. While chloroplasts
do not have this pathway, studies in the last decade revealed that
chloroplasts instead contain many prokaryotic-type proteases,
some of which are ATP-dependent and evolutionarily related to the
26S proteasome. The best studied are Clp (Caseinolytic protease)
and FtsH (Flamentious temperature sensitive H) (Adam et al.,
2006; Sakamoto, 2006). Clp is a serine-protease present in
stroma, and is composed of the proteolytic complex Clp/P/R/S/T
and the chaperone complex Clp/C/D. The whole complex com-
prises more than twenty subunits and isomers (Peltier et al., 2004).
For example, four ClpP subunits are present in chloroplasts (ClpP1
is encoded in the chloroplast genome, while ClpP3-6 are nucleus-
encoded). Each isomer seems to be essential for plant viability,
suggesting unique rather than redundant roles for each isomer.
A study with a knockdown line of ClpP6 revealed possible substrates
for Clp: based on these substrates, Clp is suggested to play a role
in the quality control of housekeeping proteins processively, rather
than the proteins for photosynthesis and other metabolic functions
(Sjögren et al., 2006). A recent genetic study in Arabidopsis indi-
cates that Clp may degrade chlorophyllide a oxygenase, an en-
zyme involved in chlorophyll b synthesis, and thus regulate the
chlorophyll degradation pathway (Nakagawara et al., 2007).

FtsH is a zinc-metalloprotease present in the thylakoid mem-
brane, and its protease and chaperone domains exist within a sin-
gle polypeptide. It is embedded in the thylakoid membrane via its
N-terminal transmembrane domains, while the C-terminal pro-
tease domain is present at the stromal side. Nine FtsH isomers
are present in chloroplasts: FtsH2 and FtsH5 are the major iso-
mers forming a hetero-complex (Sakamoto et al., 2003), and the loss of either one results in a leaf variegation phenotype in Arabidopsis (Chen et al., 2000; Sakamoto et al., 2002). As represented by these two isomers, there are two types of FtsH, Type A (FtsH1/5) and Type B (FtsH2/8), that are functionally distinguishable (Zaltsman et al., 2005). Interestingly, the co-existence of two types of FtsH is highly conserved within photosynthetic organisms. Within each type, however, the functions seem interchangeable (Yu et al., 2005). Accumulating studies in Arabidopsis and cyanobacteria demonstrate that FtsH is a major protease involved in the PSII repair cycle (Nixon et al., 2005). Degradation of the photodamaged D1 by FtsH is light-dependent, but the recognition mechanism of the damaged D1 is currently unclear. Recently, pre-proteins of Clp and FtsH subunits (ClpP4 and FtsH1) were shown...
to be substrates for ubiquitination *in vitro* (Shen et al., 2007). Although further studies seem necessary, this observation raises an intriguing possibility that the level of chloroplastic proteases may be controlled in the cytosol through the ubiquitin-proteasome pathway.

Besides Clp and FtsH, a serine-protease, Deg, is known to be present in chloroplasts, and four isomers were so far characterized in Arabidopsis chloroplasts. Deg1, Deg5 and Deg8 are peripherally attached to the lumenal side of the thylakoid membrane, and they were recently demonstrated to play a role in degrading photodamaged D1, and thus in PSII repair (Shen et al., 2007). In contrast, Deg2 is attached to the stromal side of the thylakoid membrane; however, its precise role is unclear (Huesgen et al., 2006). Unlike the situation in chloroplasts, Deg proteins do not appear to play a role in PSII repair in *Synechocystis* (Barker et al., 2006). How lumenal Deg proteases work on photodamaged D1 in the stroma, the transit peptide is removed and the protein either emerges on the stromal side of the envelope, they are cleaved in close proximity.

Besides Clp and FtsH, a serine-protease, Deg, is known to be present in chloroplasts, and four isomers were so far characterized in Arabidopsis chloroplasts. Deg1, Deg5 and Deg8 are peripherally attached to the lumenal side of the thylakoid membrane, and they were recently demonstrated to play a role in degrading photodamaged D1, and thus in PSII repair (Shen et al., 2007). In contrast, Deg2 is attached to the stromal side of the thylakoid membrane; however, its precise role is unclear (Huesgen et al., 2006). Unlike the situation in chloroplasts, Deg proteins do not appear to play a role in PSII repair in *Synechocystis* (Barker et al., 2006). How lumenal Deg proteases work on photodamaged D1 in the stroma, the transit peptide is removed and the protein either emerges on the stromal side of the envelope, they are cleaved in close proximity.

### PROTEIN TRANSPORT SYSTEMS

The plastid genome is greatly reduced, encoding just ~100 different proteins. Thus, >90% of the ~3000 different proteins present in mature plastids are encoded on nuclear DNA and synthesized in the cytosol. Because all plastids within an organism contain the same limited complement of genes, it is the imported proteins that define the developmental fate of the organelle (which may include chloroplast, amyloplast or chromoplast formation).

Nucleus-encoded chloroplast proteins are synthesized in precursor form – each one bearing an amino-terminal targeting signal called a transit peptide – and are imported into the organelle in an active, post-translational targeting process (Soll and Schnell, 2004; Kessler and Schnell, 2006; Jarvis, 2008). This process is mediated by molecular machines in the outer and inner envelope membranes, termed TOC and TIC (Translocon at the outer/inner envelope membrane of chloroplasts), respectively. Upon arrival in the stroma, the transit peptide is removed and the protein either takes on its final conformation or is sorted to one of several internal compartments in a separate targeting process (Figure 4).

### Envelope Translocation

**Transit Peptides.** For most nucleus-encoded proteins of the chloroplast interior, protein import is dependent upon the presence of a transit peptide (Bruce, 2001). Transit peptides engage the translocation machinery directly, and are sufficient to mediate the import of heterologous passenger proteins, such as GFP. As they emerge on the stromal side of the envelope, they are cleaved at a weakly conserved processing site by the stromal processing peptidase (SPP) – a metalloendopeptidase related to the β subunit of the mitochondrial processing peptidase – and then degraded.

Because mistargeting of chloroplast precursor proteins (pre-proteins) would compromise cellular functionality, it is essential that proteins are sorted efficiently and specifically, avoiding other organelles that also accept cytosolically-translated precursor proteins (e.g. mitochondria, peroxisomes and the ER). Thus, one might expect chloroplast transit peptides to share well-defined primary or secondary structural motifs. However, transit peptides are remarkably heterogeneous (Bruce, 2001). They vary in length from 20 to >100 residues, and share no obvious sequence conservation. In fact, their only shared properties appear to be a profusion of hydroxylated residues and a lack of acidic residues, giving them an overall positive charge. In this regard, transit peptides resemble the presequences that mediate import into mitochondria. Interestingly, some preproteins are dual-targeted to both chloroplasts and mitochondria (Duchêne et al., 2005; Millar et al., 2006), clearly indicating functional similarity between the two types of targeting signal.

The lack of conservation amongst transit peptides makes their identification rather difficult. Nevertheless, several programs have been developed that enable their detection with reasonable accuracy (Table 1) (Emauells et al., 2007). Analysis of the Arabidopsis genome sequence using these programs led to estimations of the chloroplast proteome ranging from ~2,000 to >4,000 proteins.

**Cytosolic Factors.** Since chloroplast protein import is a post-translational process (cf. co-translational transport into the ER), it is likely that soluble, cytosolic factors facilitate the routing of precursors from the ribosome to the chloroplast surface. Preproteins are threaded through the envelope membranes in unfolded conformation, and so cytosolic chaperones are thought necessary to prevent their folding or aggregation. Indeed, it is well-documented that Hsp70 chaperones are able to interact with chloroplast transit peptides (Jackson-Constan et al., 2001). It has been suggested that 14–3–3, Hsp70 and Hsp90 proteins facilitate the cytosolic steps of chloroplast targeting, by forming so-called ‘guidance complexes’ (Obadou et al., 2006), but these ideas have not been supported by *in vivo* studies in Arabidopsis and so their significance remains unclear.

**Stages of Envelope Translocation.** Based on energetic requirements determined *in vitro*, chloroplast protein import can be divided into three steps (Olsen and Keegstra, 1992; Young et al., 1999). First, the transit peptide reversibly interacts with receptor components of the TOC complex; this is energy-independent binding. Second, the preprotein becomes deeply inserted into the TOC complex and makes contact with the TIC machinery. Progression to this early import intermediate stage requires low ATP concentrations (~100 μM) in the intermembrane space and GTP, and is irreversible. Finally, the preprotein is completely translocated into the stroma, and the transit peptide is cleaved by SPP. Progression through this step requires high ATP concentrations (~1 mM) in the stroma. Unlike mitochondrial protein import, chloroplast import does not utilize a transmembrane protonotive force. Preproteins likely pass through these different steps seamlessly *in vivo*. Translocation through the two envelope membranes occurs simultaneously, at locations called ‘contact sites’ where they are held in close proximity.

**Recognition and Outer Membrane Translocation.** Preprotein recognition and outer membrane translocation are the two
The mechanism actually employed in vivo incorporates elements of a 'sewing machine' mechanism (Schleiff et al., 2003a). GTP-driven motor to push them through the Toc75 channels using proteins from different Toc34 primary receptors, and act as a centrally located Toc159 might rotate about its axis to accept prepro-region is proposed to comprise a single Toc159 molecule. Central domain (Schleiff et al., 2003b). The four channels are proposed to putative translocation channels surrounding a central finger-like TOC core-complexes revealed a toroid structure comprising four of contact for incident transit peptides (Becker et al., 2004a). Proponents of this model suggest that the soluble Toc159 form observed by others is an experimental artefact, and argue that Toc159 remains stably associated with the membrane throughout the import mechanism. Electron microscopic analysis of purified TOC core-complexes revealed a dimeric configuration, supporting the notion that Toc159 and Toc34 may undergo heterodimerization in vivo (Yeh et al., 2007).

In the second model, membrane-bound Toc34 is the initial point of contact for incident transit peptides (Becker et al., 2004a). Proponents of this model suggest that the soluble Toc159 form observed by others is an experimental artefact, and argue that Toc159 remains stably associated with the membrane throughout the import mechanism. Electron microscopic analysis of purified TOC core-complexes revealed a dimeric configuration, supporting the notion that Toc159 and Toc34 may undergo heterodimerization in vivo (Yeh et al., 2007).

The two models seem to be very different, but it is possible that the mechanism actually employed in vivo incorporates elements of both.

Substrate-Specific Protein Import Pathways. Most components of the import apparatus were identified through biochemical analysis of isolated pea chloroplasts. More recently, Arabidopsis has been widely adopted as an alternative model system (Jarvis et al., 1998; Bauer et al., 2000), due mainly to the availability of its genome sequence. Interestingly, when the genome was scanned for TOC homologues, many components were found to be represented by multiple genes (Jackson-Constan and Keegstra, 2001). For example, Toc34 is encoded by two Arabidopsis genes, termed atTOC33 and atTOC34 (Jarvis et al., 1998), while Toc159 is encoded by four genes, termed atTOC159, atTOC132, atTOC120 and atTOC90 (Bauer et al., 2000). Careful analyses revealed that these gene families encode different receptor isoforms with distinct functions.

Diagram showing the main components implicated in the import of proteins into chloroplasts. Outer envelope membrane components form the TOC complex, while inner envelope membrane components form the TIC complex. Components are identified by their predicted molecular weights (black text), and some key functional domains are indicated (white text). The TOC core-complex is formed by Toc159, Toc34 and Toc75. The former two proteins are receptors that together control preprotein recognition, while Toc75 forms the translocation channel. Different isoforms of the receptors exist in Arabidopsis (red text), and these associate preferentially to form distinct TOC complexes with substrate specificity. This may prevent the bulk flow of abundant precursors from out-competing the import of relatively scarce preproteins during the potentially rate-limiting early stages of import; once this potential bottleneck has been passed, the import pathways may converge at a common TIC machinery. Cytosolic 14-3-3, Hsp70 and Hsp93 proteins may form 'guidance complexes' that direct preproteins to the TOC apparatus. It has been suggested that Toc12, Hsp70 and Tic22 act to facilitate the passage of preproteins across the intermembrane space. The inner membrane translocation channel may be formed by Tic110 and/or Tic20. The former protein is also thought to coordinate late events in import by recruiting stromal chaperones to import sites; Tic110 has been proposed to collaborate with Tic40 and Hsp93 in a putative stromal import motor complex. Upon arrival in the interior, the transit peptide is cleaved by SPP, and other chaperones (Cpn60 or Hsp70) may assist in the folding or onward transport of the mature domain. Finally, the Tic62, Tic55 and Tic32 components may enable the regulation of import in response to redox signals; these components might only be recruited to import sites under certain conditions or for certain preproteins.

Figure 5. The TOC/TIC protein import machinery.
Characterization of an Arabidopsis atToc33 knockout mutant, plastid protein import 1 (ppi1), provided the first in vivo verification of the role of a biochemically-identified translocon component (Jarvis et al., 1998). Later studies on an atToc159 mutant, termed ppi2, yielded the attractive hypothesis that atToc159 is a receptor with specificity for highly-abundant, photosynthetic proteins (Bauer et al., 2000). The ppi2 mutant is albino due to a block in chloroplast differentiation, and, while photosynthetic proteins are deficient in ppi2, non-photosynthetic proteins seem to accumulate normally. During establishment of the photosynthetic apparatus, the import machinery must accommodate massive increases in the expression of key photosynthetic proteins. Existence of a separate receptor system for such proteins would prevent their bulk flow from out-competing the import of less abundant, but equally important non-photosynthetic, housekeeping proteins (Figure 5).

Studies on mutants lacking other receptor isoforms, such as atToc132, atToc120 and atToc34, indicated that these are more important for the biogenesis of non-photosynthetic plastids (Constan et al., 2004; Ivanova et al., 2004; Kubis et al., 2004). The existence of distinct TOC complexes was supported biochemically (Ivanova et al., 2004): the atToc33 isoform of Toc34 predominates in atToc159-containing complexes, whereas the atToc34 isoform predominates in atToc132/atToc120-containing complexes. This accounts nicely for the fact that the ppi1 phenotype is qualitatively similar to that of ppi2 (Kubis et al., 2003). Operation of these substrate-specific import pathways might contribute to the differentiation of different plastid types.

Inner Membrane Translocation. Several putative components of the TIC complex have been identified, but there is considerable disagreement concerning their roles (Figure 5). Preproteins most likely encounter Tic22 first of all, since it resides in the inter-membrane space. It may facilitate the passage of preproteins from TOC to TIC, perhaps functioning in association with an Hsp70 and the inwardly facing J-domain protein, Tic12 (Becker et al., 2004b). The most basic function of the TIC machinery is channel formation, and yet even this function is unclear, since both Tic110 and Tic20 have been proposed to fulfil this role (Chen et al., 2002; Heins et al., 2002). Perhaps both proteins participate in pore formation.

A major component of the TIC machinery is Tic110, which projects a sizeable domain into the stroma (Kessler and Blobel, 1996; Jackson et al., 1998). This stromal domain binds transit peptides, and probably functions to recruit molecular chaperones to the complex (Akita et al., 1997; Inaba et al., 2003). By analogy with the Hsp70-based ‘motors’ that drive transport into mitochondria and the ER, it is thought that stromal chaperones bind to emerging preproteins to ensure unidirectional movement, in a ratchet-type mechanism (Jackson-Constan et al., 2001). In chloroplasts, the relevant chaperone is probably the Hsp100 homologue, Hsp93/CiPc (Nielsen et al., 1997; Kovacheva et al., 2005). The Tic40 protein is an Stl1-domain co-chaperone, and is proposed to control the activity of the Hsp93-based import motor (Chou et al., 2006; Bédard et al., 2007).

Other proteins (Tic62, Tic55 and Tic32) possess redox-related motifs, suggesting roles in the regulation of import in response to redox status (Stengel et al., 2008). It is well documented that chloroplast redox signals influence gene expression, so it would not be surprising if chloroplast import is demonstrated to receive similar control. That the import of some preproteins is influenced by light is consistent with this hypothesis (Hirohashi et al., 2001), since these effects might be mediated by redox signals. However, the precise roles of these redox-related proteins remains unclear.

Non-Canonical Chloroplast Targeting Pathways. For many years, TOC/TIC-mediated import of preproteins with transit peptides was regarded as the unique route for entry into the chloroplast interior. Recent data indicate that alternative targeting signals and pathways exist (Figure 4). For example, proteins lacking cleavable amino-terminal targeting signals are found associated with the inner envelope membrane (Nada and Soll, 2004; Miras et al., 2007). In each case, targeting was shown to proceed without assistance from the TOC machinery. Another exciting development has been the identification of a chloroplast protein targeting pathway involving the endomembrane system (Villarejo et al., 2005). Clients of this pathway possess signal peptides for co-translational transport to the ER, from whence they pass the Golgi prior to final arrival in the chloroplast. These pathways were revealed in part by proteomic analysis, and so are discussed in the next section.

Targeting To The Envelope System

Several mechanisms exist for protein targeting to the chloroplast envelope (Hofmann and Theg, 2005). Most proteins of the outer membrane do not have cleavable targeting signals. Instead, targeting information resides within hydrophobic transmembrane domains. It was originally thought that such proteins insert ‘spontaneously’ into the membrane bilayer, without assistance from an import apparatus. However, a cytosolic sorting factor (AKR2) that mediates transport to the chloroplast surface was recently identified, while other data indicate that insertion employs the Tic75 channel protein, possibly dissociated from other TOC components (Tu et al., 2004; Bae et al., 2008).

One exceptional outer membrane protein is Tic75. This protein possesses a bipartite targeting signal, comprising a standard transit peptide and, immediately downstream of that, an intraorganellar targeting peptide. The latter functions as a ‘stop-transfer’ signal, arresting translocation so that the preprotein can disengage from the translocon and undergo membrane integration (Inoue and Keegstra, 2003).

With the exception of the non-canonical examples mentioned above, proteins of the inner envelope membrane possess a transit peptide and engage the TOC/TIC machinery. Such proteins follow two different targeting routes, referred to as the ‘post-import’ (or ‘conservative sorting’) and ‘stop-transfer’ pathways (Li and Schnell, 2006; Tripp et al., 2007). The former is a two-step process; complete translocation into the stroma is followed by membrane integration of the soluble intermediate in a separate event. This is reminiscent of ‘conservative sorting’ to the mitochondrial inner membrane, so-called because the second-step event is mediated by components of bacterial origin. Mediators of inner membrane insertion in chloroplasts have not been identified. In the ‘stop-transfer’ pathway, hydrophobic transmembrane domains mediate lateral exit from the TIC translocon and membrane integration; this route may be particularly important for polytopic proteins that might otherwise be prone to aggregation.
Targeting To The Thylakoids

The thylakoids contain highly-abundant protein complexes of the photosynthetic light reactions, which comprise both plastid- and nucleus-encoded subunits. The latter are first translocated across the plastid envelope via the TOC/TIC system, and then subsequently engage one of four different pathways for thylakoid targeting (Figure 4) (Jarvis and Robinson, 2004; Guttensohn et al., 2006; Schünemann, 2007).

Thylakoid lumenal proteins are targeted via the so-called Sec and twin-arginine translocase (Tat) pathways. Proteins engaging these translocation systems possess bipartite targeting signals: a standard transit peptide, followed by a lumenal targeting peptide similar to the signal peptides that mediate inner membrane transport in bacteria. The Sec pathway is powered by ATP (consumed by the SecA motor protein) and accepts only unfolded proteins. By contrast, the Tat pathway is proposed to be driven by the thylakoidal proton gradient, in an antiporter mechanism, and is able to accommodate fully-folded substrates; it may serve to deliver proteins that must acquire their final conformation in the stroma, through co-factor binding or oligomerization. Costs estimated for Tat translocation are remarkably high (~ 80,000 protons per protein) (Alder and Theg, 2003), but the energetics of the system have been questioned by its apparent lack of pH-dependence in vivo (Finazzi et al., 2003).

Thylakoid membrane proteins utilize either the signal recognition particle (SRP)-dependent pathway or the so-called ‘spontaneous’ insertion pathway. While the latter proceeds without energy consumption or the involvement of a proteinaceous transport machinery, the former consumes GTP as a consequence of a critical interaction between the SRP (a complex of SRP54 and SRP43, the latter being unique to chloroplasts) and its membrane receptor, FlsY, and is principally concerned with the insertion of polytopic proteins of the light-harvesting complexes. Studies on Arabidopsis mutants lacking components of the SRP pathway have played a key role in its elucidation (Tzvetkova-Chevolleau et al., 2007).

Unlike the TOC/TIC pathway, which shares no clear homology with other translocation systems, at least three of these thylakoid targeting pathways are closely related to protein transport systems of the bacterial inner membrane. They nicely exemplify the ‘conservative sorting’ concept, since the thylakoidal transport events have been retained (or conserved) from the organelle’s prokaryotic origins (hence the name Sec, which is an abbreviation of Secretory). All four pathways are essential for the biogenesis of the photosynthetic machinery, and so have been studied intensively. Most available information relates to nucleus-encoded proteins, but it should be noted that these pathways also target proteins encoded by the plastid genome. For example, a variation on the SRP pathway described above mediates the co-translational insertion of the D1 protein of PSI (Schünemann, 2007).

Finally, the possibility exists that certain other proteins are brought to the thylakoids in transport vesicles that bud off from the inner envelope membrane. While there is no direct evidence for such vesicle-mediated protein traffic, there is considerable evidence that the lipids needed for thylakoid formation are indeed transported in this way (Benning et al., 2006; Aseeva et al., 2007). Bearing in mind that some chloroplast proteins arrive at the organelle in cytosolic transport vesicles (Villarejo et al., 2005), it would not be surprising if similar mechanisms exist for onward transport to the photosynthetic membranes themselves.

PROTEOME ANALYSIS

Completion of genome sequencing projects for Arabidopsis, rice and other species, together with method development for protein identification by mass spectrometry, precipitated the onset of the proteomic era. Because of the extreme complexity of cellular proteomes, and the dynamic-range limitations associated with analyses on such complex mixtures (i.e. the tendency of highly-abundant proteins to mask the presence of others), proteomic studies have mostly focused on isolated subcellular components. In this regard, chloroplasts have received considerable attention (Baginsky and Gruissem, 2004; van Wijk, 2004; Jarvis, 2007). Proteomic analysis can confirm the expression and structure of genes predicted by genome analysis in silico, it can determine subcellular and suborganellar protein localizations, it can provide estimates of protein abundance, and it can even yield information on post-translational modification and multiprotein complex composition. Such information is invaluable, since up to 50% of the ~27,000 protein-coding genes in Arabidopsis are presently of unknown function (Swarbreck et al., 2008).

As discussed earlier, most chloroplast proteins possess an amino-terminal targeting signal, or transit peptide (Jarvis, 2008). Because transit peptides share certain characteristics, it is possible to identify candidate chloroplast proteins in silico by sequence analysis (the TargetP program is a popular choice; Table 1) (Emanuelsson et al., 2007). However, a lack of conservation amongst transit peptides and their similarity to mitochondrial presequences mean that such in silico methods are not totally reliable. Thus, the only dependable method for chloroplast protein identification is direct experimentation. Computational methods suggest that there are ~2,000-4,000 different proteins in chloroplasts, but presently there are just ~1,000 experimentally-verified, Arabidopsis plastid proteins in the PPDB and SUBA databases (Friso et al., 2004; Heazlewood et al., 2007, see Table 1 for the websites). This further emphasizes the need for additional proteome analysis.

Proteome Catalogues

Cataloguing aims to identify all proteins within a particular cellular or organellar compartment. While the chloroplast proteome is substantially smaller than that of an entire cell, it nevertheless comprises several thousand proteins. Thus, many cataloguing experiments have focused on a particular suborganellar compartment (e.g. the thylakoids, envelope, stroma or plastoglobules). Nonetheless, some studies on whole organelles have been conducted, and these have focused on different plastid types (e.g. chloroplasts, amyloplasts and etioplasts).

The Thylakoids. The thylakoid membranes themselves harbor the four multiprotein complexes of the photosynthetic light reactions (PSI, PSII, the cytochrome b6/f complex, and ATP synthase), but also function to form a central aqueous compartment called the lumen. Difficulties associated with the analysis of highly-hydrophobic integral membrane proteins led many thylakoid proteome studies to focus on luminal proteins, or proteins associated peripherally with the membranes (Kieselbach and Schröder, 2003). Yet even this is challenging, since the molar ratio between the most and least abundant luminal proteins may be as high as...
10^6, presenting a major dynamic-range barrier to comprehensive analysis (Peltier et al., 2002).

Lumenal studies revealed an unanticipated level of complexity. Relatively few Arabidopsis lumen proteins were identified experimentally, but information from these enabled in silico estimations for the whole lumen proteome, ranging from ~80 to ~400 proteins (Peltier et al., 2002; Kieselbach and Schröder, 2003; Westerlund et al., 2003); the actual size presumably lies somewhere between these estimates. In addition to the expected photosynthetic proteins, and a significant proportion of unknown function, substantial numbers of candidate lumenal proteins mediate protein folding, processing and proteolysis, anti-oxidative defense, and non-photosynthetic redox reactions; these may serve to repair and maintain normal functionality of the photosynthetic machinery, which experiences substantial redox stress. Interestingly, up to 50% were predicted to be substrates of the Tat pathway, which as discussed earlier is able to transport fully-folded proteins.

The thylakoid membrane itself is dominated by the four photosynthetic complexes, which together comprise ~100 different proteins. Nevertheless, the membrane also contains many proteins associated with the assembly, maintenance and regulation of the complexes. Despite the difficulties that hydrophobic membrane proteins present, proteomic studies achieved near complete coverage of the photosynthetic complexes, and also identified low-abundance components such as those involved in cyclic electron flow around PSI and chlororespiration (Friso et al., 2004; Peltier et al., 2004a).

By combining data from proteomic studies and other reports, the total number of experimentally-verified, thylakoid-associated proteins in Arabidopsis was found to be ~400 (Peltier et al., 2004a; van Wijk, 2004). Of these, ~30% are involved in photosynthesis, ~25% are of unknown function (including proteins with tetratricopeptide repeat [TPR], pentatricopeptide repeat [PPR], DnaJ and rhodanese domains), ~20% mediate protein translocation, folding, processing and proteolysis, and ~10% are involved in oxidative stress defense.

The Envelope System. The envelope is a double-membrane system that forms a semi-permeable barrier between the cytosol and the stroma. It contains the machinery responsible for importing nucleus-encoded proteins (Jarvis, 2008), as well as transporters that exchange ions and metabolites (Weber et al., 2005). It also possesses a unique biochemical machinery (e.g. for the synthesis of plastid membrane components and other lipids), and participates in the communication between plastids and the nucleus (López-Juez and Pyke, 2005; Nott et al., 2006).

Like the thylakoid membrane, the envelope proteome is dominated by hydrophobic integral membrane proteins (Ferro et al., 2003; Froehlich et al., 2003; Rolland et al., 2003). Combined data from proteomic studies and other experiments identified a total of ~400 envelope-associated proteins in Arabidopsis (Peltier et al., 2004a; van Wijk, 2004). Of these, ~30% are of unknown function, 13% mediate protein translocation, folding, processing or degradation, 10% participate in lipid or fatty acid metabolism, and 9% are small molecule transporters. Thus, the proteome clearly reflects the main functions of the envelope, which are quite different from those of the thylakoids. The high proportion of proteins of unknown function indicates that envelope functions are not yet fully understood. Intriguingly, proteins similar to components of the mitochondrial protein import machinery were identified, suggesting that novel protein transport systems operate in the envelope (Ferro et al., 2003).

Protein targeting to the two envelope membranes, and to the intermembrane space, is not well understood (Hofmann and Theg, 2005). Thus, it is difficult to make in silico predictions concerning the composition of the three individual proteomes. Nonetheless, evidence suggests that the outer membrane is characterized by beta-barrel proteins, and that the inner membrane is dominated by polytopic, alpha-helical transporters (Koo and Ohlrogge, 2002; Schleiff et al., 2003c).

By analyzing carefully collected sets of integral proteins from the inner envelope and thylakoid membranes (identified on the basis of published information), the respective proteomes were found to have quite different characteristics (Sun et al., 2004). On average, thylakoid proteins were smaller and more acidic than envelope proteins, and contained fewer cysteine residues. The larger average size of the envelope proteins probably reflects the presence of numerous transporters with multiple membrane spans (Weber et al., 2005), while the pI differences may be linked to pH differences between the compartments (protons are accumulated in the thylakoid lumen during photosynthesis). Cysteines are able to engage directly in redox reactions, and their deficiency in thylakoid proteins might be a measure to reduce oxidative damage.

The Stroma. The carbon reactions of photosynthesis (the Calvin cycle) and other major metabolic pathways are located in the stroma, as are components of the plastid genetic system. One in silico study estimated that the stroma contains up to ~80% of the total chloroplast proteome (Sun et al., 2004). However, it should be noted that many of these proteins may associate permanently or transiently with the thylakoids or the inner envelope membrane, through protein-protein, electrostatic or hydrophobic interactions, or via lipid anchors. Interestingly, the acetyl-coenzyme A carboxylase complex appears to be envelope-associated, placing it near the site of fatty acid use or export (Rolland et al., 2003). Such metabolic channelling has obvious advantages, and there may be many other similar examples.

Over 200 stromal proteins were identified in one study (Peltier et al., 2006). Of these, 26% mediate protein synthesis, folding, proteolysis and sorting, 12% are involved in primary carbon metabolism, including Calvin cycle enzymes, 11% are of unknown function, while 7%, 6%, 4% and 4% mediate the biosynthesis of amino acids, tetrapyrroles, nucleotides and lipids, respectively. Interestingly, proteins of primary carbon metabolism constituted most (~75%) of the total stromal mass; others responsible for protein synthesis, biogenesis and fate represented ~10% of mass, whereas those involved in nitrogen and sulphur assimilation made up ~8%. Other biosynthetic pathways each represented less than 1% of the total mass.

Other studies on the stromal compartment were more focused. For example, all proteins of the plastidic 70S ribosome were identified (Yamaguchi and Subramanian, 2000; Yamaguchi et al., 2000). The ribosome comprises 59 different proteins; 53 are orthologues of bacterial ribosomal proteins, while six are plastid specific. The latter may mediate functions unique to plastid translation and its regulation, such as protein targeting to the thylakoids and control by nuclear factors.

Interestingly, two ribosomal proteins were identified as targets for regulation by the stromal thioredoxin system (Balmer et al., 2003). This system is composed of ferredoxin, ferredoxin-thioredoxin reductase, and thioredoxin, and it links light to the regulation of photosynthetic enzymes and processes such as lipid biosyn-
thesis. Electrons flow from ferredoxin to thioredoxin, which in its reduced state regulates the activity of target proteins by reducing specific disulphides. Proteomic strategies were employed to extend the list of targets of the system (Hisabori et al., 2007). Proteins identified in this way included components of established thioredoxin-regulated pathways (e.g. the Calvin cycle, nitrogen and sulphur metabolism, and protein synthesis) and others not previously recognized as thioredoxin targets (e.g. tetrapyrrole biosynthesis, protein folding, assembly and degradation, starch degradation, DNA replication and transcription, and plastid division). Remarkably, this mainly stromal regulatory network even extends into the thylakoid lumen, where it targets the photosynthetic electron transport chains.

**Plastoglobules.** Plastoglobules are lipid-containing bodies in chloroplasts, chromoplasts and other plastids. They contain various lipidic compounds (e.g. galactolipids, fatty acids, carotenoids, tocopherols and plastequinone), and were previously thought to function as simple lipid stores. However, proteomic analyses identified ~30 different proteins in plastoglobules, revealing a much more complex set of activities (Bréhélin et al., 2007). In addition to the plastoglobulin family of structural proteins, which bind to the surface of the globules and prevent their coalescence, a number of enzymes were identified. Plastoglobules in chloroplasts actively participate in the synthesis of their lipophilic constituents (e.g. alpha-tocopherol [vitamin E], an important anti-oxidant in thylakoids); similarly, those in chromoplasts contain carotenoid biosynthetic enzymes. In chloroplasts, plastoglobules are directly coupled to the thylakoids, suggesting that their contents are in equilibrium and that they contribute to the synthesis of thylakoidal constituents (Austin et al., 2006).

**Whole Organelles.** An extensive study of the whole Arabidopsis chloroplasts identified ~600 different proteins (Kleffmann et al., 2004). Almost complete coverage was achieved for major metabolic pathways (e.g. the Calvin cycle), while coverage for less abundant pathways was only partial. Interestingly, over 30% of the identified proteins were of unknown function. Parallel RNA profiling revealed a correlation between transcript levels and protein abundances in some metabolic pathways, but not others, implying that distinct regulatory mechanisms operate in different pathways.

Proteome studies on amyloplasts identified most enzymes of starch biosynthesis, as expected, but also revealed a surprisingly broad spectrum of biosynthetic capabilities (Balmer et al., 2006a; Stensballe et al., 2008). Like chloroplasts, amyloplasts possess enzymes for nitrogen and sulphur assimilation, and the biosynthesis of amino acids, fatty acids and tetrapyrroles. In comparison with chloroplasts, amyloplasts contain a higher proportion of proteins of carbon, nitrogen and sulphur metabolism, and transport processes. By contrast, chloroplasts contain proportionally more proteins of unknown function, presumably reflecting their more complex activities. While components of the TOC/TIC machinery were identified in amyloplasts, no ribosomal proteins were detected, suggesting that most amyloplast proteins are nucleus-encoded. This makes sense, since the plastid genome is dominated by genes for photosynthetic components. Interestingly, a thioredoxin regulatory network also operates in amyloplasts (Balmer et al., 2006b).

Analyses of etioplasts and chromoplasts have also been conducted (von Zychlinski et al., 2005; Siddique et al., 2006). Proteome comparisons revealed that both have metabolic functions typical of heterotrophic plastids, but that etioplasts nevertheless share significant similarities with chloroplasts—the organelles into which they ultimately develop.

**Protein Targeting Issues**

Surprisingly, when the ~600 proteins identified in the whole chloroplast proteome study mentioned above were analyzed using the TargetP program, only ~60% were predicted to have a transit peptide (Kleffmann et al., 2004). While a more recent study suggests that this may overestimate the number of chloroplast proteins lacking a canonical transit peptide (Zybailov et al., 2008), it is nevertheless interesting to note that considerable numbers were predicted to have putative mitochondrial precursor sequences, signal peptides for ER translocation, or no cleavable targeting signal at all. It is possible that some of these proteins were contaminants from other cellular compartments, or had received incorrect TargetP predictions, but the data nevertheless suggest that protein targeting to chloroplasts is more complex than was previously envisaged (Jarvis, 2008). Until recently, all nucleus-encoded proteins of the chloroplast interior were thought to have a transit peptide for TOC/TIC engagement.

Initial evidence for a more complicated picture of chloroplast protein biogenesis was provided by studies on the Arabidopsis envelope proteome (Ferro et al., 2003). A protein named ceQORH (chloroplast envelope guinone oxidoreductase homologue) was found in the inner membrane, in spite of its lack of a transit peptide. Interestingly, an internal sequence of ~40 residues controls ceQORH localization, in a process that is not mediated by the TOC/TIC machinery (Figure 4) (Miras et al., 2007). Another inner membrane protein, Tic32, was also found to lack a transit peptide, and to be targeted in TOC-independent fashion (Nada and Soll, 2004). Whether Tic32 follows the same import pathway as ceQORH remains to be determined. Nonetheless, it should be noted that Tic32 targeting information resides at its N-terminus, and that the energetic requirements for Tic32 targeting are lower than for ceQORH.

The identification of chloroplast proteins with signal peptides for ER translocation was surprising (Kleffmann et al., 2004). Chloroplast protein traffic through the endomembrane system is well documented in algae and apicomplexan parasites, which have complex plastids (Nassoury and Morse, 2005). While such targeting makes sense in these organisms, due to the complex nature of their plastids and the likely autogenous origin of the outer organelar membrane, it would seem unnecessary in higher plants. Nonetheless, evidence for such a targeting pathway was recently presented (Villarejo et al., 2005). The Arabidopsis carboxylic anhydride 1 (CAH1) protein was found in the stroma, in spite of its predicted signal peptide. Intriguingly, CAH1 could not be imported directly by chloroplasts, but was instead taken up co-translationally by ER microsomes. Moreover, stromal CAH1 is glycosylated; because the relevant glycans are only added in the Golgi, a chloroplast protein transport pathway through the Golgi was inferred (Figure 4). Indeed, application of brefeldin A (an agent that interferes with Golgi-mediated vesicle traffic) obstructed the transport of CAH1, causing its arrest within the endomembrane system. What happens once the vesicles arrive at the chloroplast is surface is less clear, but it is interesting to note that there is considerable evidence for vesicle budding at the inner envelope membrane (Benning et al., 2006; Aseeva et al., 2007). More recently, another
chloroplast protein was shown to follow a similar pathway (Nanjo et al., 2006). The CAH1 pathway may be the vestige of an ancestral targeting mechanism that prevailed during early evolution, and which for some reason has been retained for a few proteins. That some proteins in chloroplasts are predicted to have mitochondrial presequences is much less surprising, since the dual-targeting of certain proteins to both chloroplasts and mitochondria is well documented (Duchêne et al., 2005; Millar et al., 2006). Overall, the emerging picture of chloroplast protein targeting is rather complex, demonstrating that transit peptide prediction in silico cannot provide a complete description of the organelar proteome.

**Comparative Proteomics**

In addition to its role in cataloguing experiments, proteomics has been employed in comparative studies with considerable success. For example, changes in the organellar proteome were studied during de-etiolation or greening (Lonosky et al., 2004; Kleffmann et al., 2007), while the responses of lumenal, stromal and plastoglobular proteomes to low temperature or light stress were characterized (Giacomelli et al., 2006; Goulas et al., 2006; Ytterberg et al., 2006). In another example, chloroplasts from Arabidopsis mutants lacking different TOC receptor isoforms were analyzed (Kubis et al., 2003). Different groups of chloroplast proteins were selectively deficient in the mutants, indicating that the different TOC receptor isoforms possess preprotein recognition specificity (Jarvis, 2008). The data suggested that at least two different import pathways operate in plastids (Figure 5), as discussed in the previous section.

**Oligomerization and Modification**

Many proteins do not function in isolation, but as part of multiprotein complexes. Thus, experimental procedures that maintain oligomeric status are desirable, since they preserve important interaction information. Several complexes of chloroplasts have been purified to homogeneity and analyzed individually; examples include the photosystems, the cytochrome b$_6$f complex (Whitelegge, 2003), the ribosomal subunits (see earlier), and a Clp complex (Peltier et al., 2004b). Additionally, the oligomeric state of the stromal proteome was systematically analysed by Peltier et al. (2006).

Protein modification is another important issue that must be taken into consideration. The molecular mass of an intact protein defines its native covalent state, and so its accurate measurement can reveal modifications mediated either post-transcriptionally, through processes like RNA editing, or post-translationally (Whitelegge, 2003; van Wijk, 2004). Several studies revealed covalent modifications of plastidic proteins, including acetylation, glycosylation, palmitoylation, phosphorylation, and N-terminal methionine excision. Such modifications may influence the activity, interactions or stability of the protein, or anchor it to a membrane.

**PLASTID DIVISION**

Like their free-living ancestors, both chloroplasts and mitochondria divide, thus enabling their continued inheritance by daughter cells after cell division (Boffey and Lloyd, 1988; Kuroiwa et al., 1998; Figure 6). However, most of the ancestral bacterial genes were either lost or transferred to the host nuclear genome during evolution, so that the genomes of both organelles lack sufficient information to mediate their own biogenesis and division. Therefore, unlike in bacteria, division of these organelles is performed and controlled by the eukaryotic nuclear genome (Kuroiwa and Pyke, 1999). Recent structural and molecular-genetic studies have allowed us to understand the mechanisms of plastid division, and in particular the division apparatus including ring structures at the division site (Kuroiwa et al. 1998) and GTPase proteins that are strongly related to components of the bacterial division apparatus (Miyagishima et al., 2003; Osteryoung and Nunnari, 2003; Figure 6). Our current knowledge on plastid division comes from studies in chloroplasts, but it should be noted that some plastid-type-specific mechanisms might exist. As explained below, molecular-genetic works in Arabidopsis have significantly contributed to this field. In addition, plastid division has been studied in ancestral algal chloroplasts. In light of these studies, the division machinery appears to be shared between chloroplasts and other plastid types.
Cytological Aspects

A large number of cytological studies have shown that: (1) plastids multiply by division along with the duplication and separation of their nucleoids; (2) different plastid types, including proplastids, etioplasts, chloroplasts and amyloplasts, are capable of dividing themselves (Boffey and Lloyd, 1988). Electron microscopic observations have also shown that plastids divide by simultaneous constriction of the inner and outer envelopes. Plastids usually multiply by binary fission, but multiple fission has been observed in some species and tissues. Some primitive algae have only one or a few chloroplasts per cell, and the chloroplasts divide synchronously with the cell cycle. In higher plants, however, cells generally contain large numbers of plastids, and the plastids divide non-synchronously, even within the same cell (Figure 6). In addition, plastids continue to divide in developing tissues in which cells expand but do not divide. Many earlier studies demonstrated that the plastid number per cell varies considerably depending on the cell type, developmental stage, and environmental conditions (Boffey and Lloyd, 1988; Pyke, 1999). In spite of these observations, little is understood about how plastid number per cell is controlled at the molecular level, or how it is related to the rate of plastid division.

Importantly, electron microscopic studies have identified electron-dense ring structures encircling the constriction furrow of dividing plastids. This ring structure, called the plastid-dividing (PD) ring, has been observed in several lineages of algae and plants including Arabidopsis (Mita et al., 1986; Hashimoto, 1986; Kuroiwa et al., 1998). In most cases, the PD ring was detected as a double-ring structure, with one ring (the outer PD ring) on the cytoplasmic face of the outer envelope membrane, and the other ring (the inner PD ring) on the stromal face of the inner envelope membrane. In the red alga, Cyanidioschyzon merolae, a middle PD ring was also identified in the intermembrane space. The structure and behavior of the two (or three) rings are different, suggesting that each ring has distinct functions and is composed of distinct sets of proteins (Kuroiwa et al., 1998; Miyagishima et al., 2003). Attempts to identify proteins associated with or comprising the PD rings have been made. In fact, several proteins were shown to exist at the plastid division site (see below), but none of these proteins appears to be an actual component of the PD ring. Thus, the protein components of the PD ring remain unclear.

FtsZ Descended From The Cyanobacterial Endosymbiont

The first protein shown to play a role in chloroplast division was a plant homologue of FtsZ, the key bacterial division protein (Osteryoung and Vierling, 1995; Osteroyung et al., 1998; Strepp et al., 1998). FtsZ is a bacterial GTPase that is structurally similar to tubulin, and which self-assembles into a ring structure beneath the cytoplasmic membrane at the division site. Formation of the FtsZ ring is an initial event at the division site, and initiates the recruitment of other proteins that constitute the bacterial division complex. Of all the proteins involved in bacterial cell division, FtsZ is thought to play an especially important role (Harry et al., 2006). A gene encoding a chloroplast-targeted FtsZ protein was found in the Arabidopsis nuclear genome, and, subsequently, similar FtsZ homologues have been reported for other photosynthetic eukaryotes. Plant FtsZ proteins are most closely related to their cyanobacterial counterparts, which supports an endosymbiotic origin of chloroplasts. Analogous to the bacterial protein, plant FtsZ is localized at the stromal side of the chloroplast division site.

Whereas most bacteria (including cyanobacteria) have only one ftsZ gene, plants have more than two genes that are clustered into two phylogenetic groups: FtsZ1 and FtsZ2. It remains unclear as to why plants contain two types of FtsZ. Nevertheless, depletion of either protein in Arabidopsis disrupts plastid division, suggesting that, instead of being redundant, FtsZ1 and FtsZ2 may have distinct functions. Interestingly, both FtsZ1 and FtsZ2 were shown to co-localize in plastids, even when the localization of FtsZ filaments was altered experimentally (McAndrew et al., 2001; Kuroiwa et al., 2002). A comparison of their primary structures revealed that FtsZ2 contains a short stretch of conserved amino acids at the C-terminus (the C-terminal core domain) very similar to an equivalent region in bacterial FtsZ; by contrast, FtsZ1 does not contain this core domain (Osteryoung et al., 1998; McAndrew et al., 2001). A recent study showed that another plastid division protein, ARC6 (accumulation and replication of chloroplasts 6, see below), specifically interacts with FtsZ2, but not with FtsZ1. ARC6 interacts with FtsZ2 via the core domain, which is absent in FtsZ1, suggesting functional differences between these FtsZ proteins (Maple and Møller, 2007).

In addition to FtsZ1 and FtsZ2, a third FtsZ-like protein (ARC3) was identified in Arabidopsis by map-based cloning of arc3 mutants. In the arc mutants, the number and size of chloroplasts in leaf mesophyll cells differs from that in wild-type plants. Among them, arc2, arc3 and arc5 through to arc12 all have mesophyll cells with reduced numbers of enlarged chloroplasts, suggesting that chloroplast division is defective in these mutants (Pyke, 1999). ARC3 encodes a protein that has an FtsZ-like N-terminal region and a C-terminal domain homologous to a region of phosphatidylinositol-4-phosphate 5-kinase (Shimada et al., 2004). ARC3 is localized on the stromal side of the chloroplast division site (Maple et al., 2007), and forms a complex with at least FtsZ1, FtsZ2 and ARC6 (McAndrew et al., 2008). Functional studies suggest that ARC3 is involved in the placement of the FtsZ ring in chloroplasts (Maple et al., 2007).

Other Division Proteins Descended From The Cyanobacterial Endosymbiont

Given the successful discovery of plastidic FtsZ proteins, reverse-genetic approaches have been taken to identify other plastid division proteins homologous to bacterial components. In addition, conventional forward-genetic approaches were taken using chloroplast division mutants such as the arc mutants. These studies enabled the identification of novel plastid division genes that are likely to be of cyanobacterial origin. Homologues of minD (mini-cell D) (Colletti et al., 2000) and miniE (Itoh et al., 2001), which determine the site forming the FtsZ ring in bacteria, have been found in plant nuclear genomes. Involvement of these genes in the positioning of the chloroplast division site was confirmed by characterizing the corresponding mutants in Arabidopsis (Colletti et al., 2000; Itoh et al., 2001). In a separate work, the arc6 mutation in Arabidopsis was identified by map-based cloning. It was revealed that the
ARC6 locus is orthologous to a cyanobacterial division gene, ftn2. ARC6 contains a DnaJ-like domain and is localized at the chloroplast division site spanning the inner envelope membrane. ARC6 is thus suggested to play a role in stabilizing FtsZ filaments (Vitha et al., 2003). Identification of these bacterial-type division genes, in addition to ftsZ, confirms the hypothesis that much (if not all) of the chloroplast division apparatus is derived from the cyanobacterial ancestor, and that the corresponding genes have been transferred to the nuclear genome during evolution. However, comparative studies between cyanobacteria and plants indicated that the majority of bacterial cell division genes were lost after endosymbiosis (Miyagishima et al., 2005).

**Dynamin and PDV Proteins Originated From The Eukaryotic Host**

In addition to cyanobacteria-derived mechanisms, plastid division requires additional components evolved from the eukaryotic host cell, as represented by a member of dynamin family (Miyagishima et al., 2003; Osteryoung and Nunnari, 2003). Dynamin family proteins contain a GTPase domain and are specifically found in eukaryotic organisms, although a recent study showed that eu bacteria do have proteins distantly related to the eukaryotic dynamin family. The best characterized example in eukaryotes is the dynamin protein that self-assembles into a ring at the neck of clathrin-coated pits. During this process, the dynamin plays a predominant role in pinching off vesicles from the plasma membrane. To date, several proteins have been included in the dynamin family owing to their structural similarity, and each member has been shown to play roles in fission or fusion of distinct eukaryotic membrane systems, for example in mitochondrial division (Praefcke and McMahon, 2004). One member of the dynamin family, named DRP5B (dynamin-related protein 5B), was shown to be involved in plastid division in the red alga, *C. merolae* (Miyagishima et al., 2003), and Arabidopsis (Gao et al., 2003). In Arabidopsis, DRP5B was identified through forward genetics: map-based cloning of the *arc5* locus revealed that the loss of DRP5B results in the arc phenotype. Analogous to conventional dynamins at the plasma membrane, DRP5B localizes at the cytosolic side of the plastid division site. In Arabidopsis *arc5* mutants, chloroplast divisions are arrested at the stage of division-site constriction, supposedly after the FtsZ and PD rings are formed (Pyke, 1999). Recent analyses of Arabidopsis mutants showing a phenotype similar to *arc5* led to identification of PDV1 (plastid division 1), an integral protein of the outer envelope membrane. Database searches found a protein paralogous to PDV1, termed PDV2. Similar to DRP5B, PDV1 is localized to a discontinuous ring at the plastid division site. Based on the mutant phenotypes of these genes, we hypothesize that PDV1 and PDV2 act co-ordinately and recruit DRP5B to the division site (Miyagishima et al., 2006).

**Relationship Between Plastid Division Components**

As aforementioned, many factors of eukaryotic and prokaryotic origin have been identified as components of the chloroplast division machinery in the last decade. Precise roles of these division proteins remain unclear and further characterization is necessary. Nevertheless, combining cytological observations (mainly performed in the red alga, *C. merolae*) with molecular-genetic studies (using available mutants in Arabidopsis) has uncovered (at least part of) the spatio-temporal relationship amongst these components, as follows (Marrison et al., 1999; Miyagishima et al., 2003; Osteryoung and Nunnari, 2003; Maple et al., 2005; Glynn et al., 2007; Maple and Møller, 2007; Figure 6). The FtsZ, PD and dynamin rings form in this order at the division site. Small dynamin patches are discontinuously localized at the division site at the onset of constriction. After this, constriction commences and, at later stages of constriction, dynamin forms a continuous ring structure (Miyagishima et al., 2003). The FtsZ and inner PD rings disappear just before the completion of division. In contrast, remnants of the outer PD ring remain between the daughter chloroplasts, and remnants of the dynamin ring remain clinging to only one daughter chloroplast. Following the completion of division, these remnants of the PD and dynamin rings eventually disappear and dissolve in the cytosol (Miyagishima et al., 2003).

Studies in Arabidopsis suggest that MinD and MinE regulate the positioning of FtsZ ring formation, as in bacteria. ARC3 also appears to play a role in this positioning process. ARC3 is suggested to be involved in the formation or stabilization of the FtsZ ring (Glynn et al., 2007; Maple and Møller, 2007). PDV1 and PDV2 recruit the ARC5/DRP5B dynamin to the division site. A very recent study using the chloroplast division apparatus isolated from *C. merolae* suggested that the rings are connected by unknown factors that span across the two envelope membranes. Interestingly, it was also shown that the isolated complex (the various rings) is capable of constriction after stretching, most likely by the function of the dynamin protein (Yoshida et al., 2006).

**Other Proteins Implicated in Chloroplast Division and Its Regulation**

Disruption of three other genes of cyanobacterial origin, *ALB4 (albino 4)*, *CRL (crumpled leaf)*, *AtSulA/GC1 (giant chloroplast 1)*, and a single gene of eukaryotic origin, *FZL (FZO-like)*, were shown to impair chloroplast division leading to altered chloroplast morphologies in Arabidopsis (reviewed in Glynn et al., 2007; Maple and Møller, 2007). It is currently unclear, however, whether these proteins are directly involved in chloroplast division or how these proteins are involved in the division process. In addition, genes related to the cyanobacterial pathway for peptidoglycan synthesis are present in plants, and these possibly affect chloroplast division. In the moss, *Physcomitrella patens*, disruption of these genes causes defects in chloroplast division, while in Arabidopsis the development of chloroplasts was perturbed in the corresponding mutants. How these genes are involved in chloroplast division awaits further research.

Additional factors, MSL (*MscS*-like protein) and CDT1 (*gdc10*-dependent transcript 1), were identified as regulators of chloroplast division through reverse-genetic approaches. Arabidopsis MSL proteins are homologous to the bacterial mechanosensitive ion channel protein, MscS. In bacteria, MscS proteins are involved in regulating osmotic potential across the cell membrane in response to increased membrane tension induced by osmotic shock. Arabidopsis MSL2 and MSL3 are localized to plastid membranes, and the mesophyll cells in *msl2 msl3* double mutants have a
smaller number of large chloroplasts than those in the wild type. Based on these results, it is suggested that the MSL system plays a role in the density dependent control of chloroplast numbers and volume in mesophyll cells (Haswell and Meyerowitz, 2006).

CDT1 is a member of the pre-replication complex in eukaryotes, and acts as a key regulator of nuclear DNA replication. Arabidopsis CDT1a was shown to be dually targeted to the nucleus and chloroplasts. Interestingly, AtCDT1-RNAi plants show defects in plastid division. This study suggested that AtCDT1a is a component of the nuclear pre-replication complex as in other eukaryotes, and that its additional plastidic form regulates FtsZ ring formation through interaction with ARC6 (Raynaud et al., 2005).

**INHERITANCE OF PLASTIDS**

As implicated in earlier works in the 1900s by Baur and Corens, which were published shortly after the re-discovery of Mendelian inheritance, some genetic traits (in this case exemplified by leaf variegation in *Mirabilis*) were demonstrated to show maternal transmission (Hagemann, 2000). At the present time, a large body of data regarding the mode of inheritance of chloroplasts has been accumulated (Kirk and Tilney-Bassett, 1978; Hagemann and Schröder, 1989; Kuroiwa, 1991; Birky, 1995; Mogensen, 1996; Hagemann, 2004). In plants, chloroplasts and mitochondria are the subject of organelle genetics, since both organelles were derived from endosymbiosis and both contain their own genomes (Birky, 2001). Despite such a long-lasting interest, very little is known about how ptDNA is distributed into daughter plastids upon plastid division. Likewise, we know very little about the molecular mechanisms of plastid inheritance, although many cytological analyses strongly indicate that the behavior of plastids in male gametes is a key issue. In this review, we consider the inheritance mode of plastids, and place particular emphasis on the fate of plastid and ptDNAs during pollen development. We also focus on knowledge related to Arabidopsis, comparing it with other species.

**Types of Plastid Inheritance – Maternal Inheritance is Common But Not Absolute**

To begin this area of investigation, it is logical to ask the question “What happens if two different types of plastid co-exist in the same cell?” Since no event of plastid fusion has been previously demonstrated in higher plants, either one of the plastid types remains, or both types co-exist. We can test these hypotheses by performing protoplast fusions with two different cell types, and then following their plastid types during subsequent plant regeneration (Perl et al., 1991; Wolters et al., 1993). Such experiments suggest that only one type remains, and that the other type is excluded. Based upon these observations, which are sometimes referred to as ‘vegetative segregation’, we assume that two plastid types can co-exist, but it appears that either one or the other is eventually sorted out or lost (leading to ‘homoplasmy’). This ‘sort-out’ mechanism is also applicable during fertilization, if a zygotic cell formed by fertilization contained plastids from both male and female gametes (Birky, 2001). If these two types co-exist, then this condition is called ‘heteroplasmy’.

Plastid inheritance after fertilization in an F1 cross might be: (1) uniparental maternal type; (2) uniparental paternal type; or (3) biparental type. In biparental plants, vegetative segregation leads to the sorting-out of either plastid type, resulting in maternal or paternal homoplasmy in the following generations. This mode of plastidic inheritance has been genetically identified in many angiosperms, and examples of each of the other aforementioned types have also been reported. However, there is an apparent bias in the proportion of occurrence (Kuroiwa, 1991; Mogensen, 1996; Hagemann, 2004). In angiosperms, maternal inheritance dominates over the others: cytological studies (see below) suggest that approximately 80% of all species show maternal inheritance, and that the remaining species show biparental inheritance (Corriveau and Coleman, 1988; Zhang et al., 2003). In contrast, uniparental paternal inheritance of plastids is very rare, and is only reported in the kiwi plant (*Actinidia delicosa*). Thus, one can assume that the behavior of plastids (or ptDNA) during pollen development and fertilization is relevant to plastid inheritance.

**Plastid Behavior During Pollen Development**

Mature angiosperm pollen grains are composed of two or three male reproductive cells (McCormick, 2004). A tetrad formed by meiosis of pollen mother cells gives rise to four microspores. Each microspore undergoes pollen mitosis I (PMI) to divide into a larger vegetative cell and a smaller generative cell. Subsequently, the generative cell undergoes pollen mitosis II (PMII) which results in the formation of two sperm cells. The timing of PMII varies among species, sometimes occurring within the anther, although more commonly it occurs during pollen tube growth. In Arabidopsis, mature pollen is comprised of one vegetative and two sperm cells (PMII completed before pollen maturation, Figure 7). How do plastids behave during this maturation process? At the present time, it appears that there are at least several important steps which determine plastid inheritance by including or excluding plastids. Here we focus on species showing maternal inheritance, and follow Hagemann (2004) in an attempt to categorize these species.

Careful examination of pollen tissues by transmission electron microscopy reveals organelle behavior throughout PMI and PMII. According to the classification described by Hagemann (2004), the exclusion of pollen plastids occurs at different steps. The first step occurs at the stage of PMI, where the asymmetric division results in the formation of a generative cell which is devoid of plastids (Figure 7). This type of exclusion occurs in *Lycopersicon*, and so the relevant plants are referred to as ‘Lycopersicon type’ species. The second step at which exclusion may occur is within the generative cell, in which the incorporated plastids are degraded by unknown mechanisms. This occurs in *Solanum*, and so the relevant plants are termed ‘Solanum type’ species. The last step occurs at the onset of fertilization, and is related to the fact that sperm cells are much smaller in volume than egg cells. As a consequence of this disparity, any plastids present in the sperm will occupy only a limited amount of cytoplasm, and so will not contribute to fertilization and are excluded from the zygote. This type of exclusion occurs in *Triticum*, and so the relevant plants are termed ‘Triticum type’ species. Studies performed by our research group and by other investigators have demonstrated that Arabidopsis is a ‘Lycopersicon type’ species, and that it transmits plastids maternally (Martínez et al., 1997; Nagata et al., 1999) (Figure 7). In some species, while...
plastids are obviously dispensable in the sperm cells, many plastids can nevertheless be observed in the cytoplasm of the vegetative cell of the male gametophyte. In contrast with the maternal inheritance species, plastids containing ptDNAs are detectable in the pollen of biparental species. The latter include Pelagonium, and so these are termed ‘Pelagonium type’ species.

Strictness of Maternal Inheritance

The aforementioned cytological and genetic studies demonstrated that maternal inheritance is a dominant way of plastid transmission in angiosperms. However, this maternal inheritance mode is not absolute, even in the species that were experimentally demonstrated to show maternal inheritance. Specifically, at a relatively low frequency, some paternal plastids are not excluded or degenerated during pollen maturation, and so may be transmitted into the next generation (Azhagiri and Maliga, 2007; Ruf et al., 2007; Svab and Maliga, 2007). Paternal transmission of plastids should be detectable at a certain frequency by following a phenotype (such as yellow or variegated leaves) linked to the paternal plastids in the F1 generation (Hagemann, 2004). Given its occurrence at an extremely low frequency, the detection of paternal inheritance is technically very challenging. Thus, a plastid phenotype which allows for positive and strong selection, such as antibiotic resistance encoded by the plastid genome, can greatly facilitate detailed assessments of the leakage of paternal plastids. In Arabidopsis,

Figure 7. Plastid inheritance in Arabidopsis.

(A) A schematic representation illustrating the transmission of plastids and mitochondria and their nucleoids during pollen development. Note that organelle DNAs are represented as nucleoids detected by DAPI and other fluorescent dyes. Note also that the signals disappear in mature pollen.

(B) Micrographs of a microspore (left) and a mature pollen grain (right) stained by DAPI (ecotype Columbia). Arrowheads indicate the prominent DAPI-stained signals, representing plastid nucleoids.

(C) An example of maternal inheritance revealed by nuclear and plastid DNA polymorphisms. F1 plants were generated by reciprocal crosses between the ecotypes Columbia (Col) and Cape Verde Islands (Cvi). Total DNAs from the parents and the F1 plants were subjected to CAPS (Cleave Amplified Polymorphic Sequence) analyses, to detect both nuclear and plastid DNA polymorphisms (top and bottom panels, respectively). DNA fragments representing Col and Cvi genotypes are indicated by magenta and cyan arrowheads, respectively. The CAPS markers used were G4711 for nuclear DNA, and ndhG for plastid DNA.
plastids are maternally transmitted (‘Lycopersicon type’). Fortunately, a strong selection by spectinomycin resistance is possible due to a mutation which resides in the 16S rRNA gene. Thus, paternally transmitted ptDNA can be detected at a frequency of $3.9 \times 10^{-5}$ (Azhagiri and Maliga, 2007). These results imply that, despite a very low frequency, paternal plastids can overcome all of the aforementioned exclusion steps and still be transmitted into subsequent generations.

**Do Plastid DNAs Decrease During Pollen Development?**

The mode of plastid inheritance (particular maternal inheritance) is in large part the consequence of plastid exclusion or degeneration during pollen maturation and fertilization. However, another interesting question is whether the amount of ptDNA decreases prior to pollen maturation, and whether this reduction is associated with plastid inheritance. A change in the amount of ptDNA has been characterized in cytological studies which detect organelle DNAs through the usage of a fluorescent dye such as DAPI in pollen (Kuroiwa, 1991; Sato et al., 2003; Sakai et al., 2004). Using this method, ptDNAs can be observed as aggregates with variable size, likely representing DNA-protein complexes as plastid nucleoids (described earlier; Figures 2 and 7). Among species showing biparental plastid inheritance, we have frequently observed plastids with detectable DAPI signals (corresponding to nucleoids) in generative or sperm plastids (Nagata et al., 1999). In contrast, such DAPI signals are not detected in maternal inheritance species (Miyamura et al., 1987). This disparity serves as a convenient method for assessing the ‘potential’ for biparental inheritance (Corriveau and Coleman, 1988; Zhang et al., 2003). In fact, it has enabled us to estimate that 80% of angiosperms exhibit maternal inheritance. Although the presence of organellar DAPI signals in sperm cells does not always guarantee biparental inheritance, this observation implies that the amount of ptDNAs that are retained in sperm plastids is controlled, either by preferential replication or by degradation.

We must be cautious when making conclusions that are based upon these observations. DAPI signals (even signals with other fluorescent dyes) do not represent the quantity of DNAs. On the contrary, they tend to show the physiological state in which ptDNA is condensed with associating proteins. Likewise, the absence of DAPI signals does not mean that organelle DNAs are completely missing, as evidenced by several exceptional species showing biparental inheritance. Nevertheless, it should be noted that the ptDNA signals detected by DAPI are missing even in the vegetative cells which are unlike the sperm cells and do not contribute to fertilization (Miyamura et al., 1987; Corriveau, 1991). We recently characterized DAPI-detectable ptDNAs during pollen maturation in Arabidopsis (Figure 7). These DAPI signals were abundant at the initial stage of pollen maturation, were reduced at and around PMII, and then finally disappeared in the vegetative cell of mature pollen (manuscript in preparation). Again, the disappearance of DAPI signals may not represent the actual degradation of DNA. However, at the least it can be concluded that ptDNAs undergo a physiological change during pollen development. We place particular emphasis on this change of ptDNA status, because it may be somewhat related to male-specific degradation of chloroplast DNAs seen in the lower unicellular alga, *Chlamydomonas* (Sager and Lane, 1972). Sears and VanWinkle-Swift (1994) proposed that the diminishment of organelle DNAs detected by DAPI in higher plant pollen may not directly correlate with the mode of plastid inheritance, it may reveal some of the important issues regarding the metabolism of organellar DNA.

**PERSPECTIVES**

Chloroplast biogenesis is one of the most important subjects in plant biology. Complete and comprehensive coverage of this subject area could consume the space of an entire textbook, and so we have elected to focus on several key aspects within this article. In the last two decades, we have seen considerable progress that has significantly increased our understanding of chloroplast biogenesis. Scientists working within the field have benefited significantly from the collective advancements achieved within the Arabidopsis research community. Nevertheless, there are several fundamental questions that remain to be answered. For example, early events associated with thylakoid biogenesis are poorly understood. Thylakoid membranes are considered to be derived from the inner envelope membrane, but the molecular events that underlie their biogenesis are not fully appreciated. It is likely that conventional molecular-genetic studies may be difficult in this case, since fundamental defects in thylakoid formation would likely be detrimental to plant survival. Additionally, the mechanisms by which thylakoids are distributed between daughter cells upon plastid division awaits elucidation in future research. The processes involved in the distribution of chloroplasts (during cell division) or chloroplastic components such as ptDNA (during organelar division) are also poorly understood. It is possible that experiments which aim to shed light on the continuity or maintenance of plastids, rather than their biogenesis or formation, may be the direction of research endeavors in the next decade. Systematic investigations of genome information, gene products, and their tissue-specific patterns of expression and activity will enable us to address unresolved questions within this field.

**ACKNOWLEDGMENTS**

We thank Chieko Saito, Ryo Matsushima, Yusuke Kato, and Yasuyuki Uno for their help in preparing figures. Research in the authors’ laboratories was partly supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (to W.S.), and by the Royal Society Rosenheim Fellowship and the Biotechnology and Biological Sciences Research Council (to P.J.).

**REFERENCES**


Chloroplast Biogenesis: Control of Plastid Development, Protein Import, Division and Inheritance


Inoue, K., and Keegstra, K. (2003). A polyglycine stretch is necessary for proper targeting of the protein translation channel pre-


Chloroplast Biogenesis: Control of Plastid Development, Protein Import, Division and Inheritance

27 of 30


Sakamoto, W., Zaltsman, A., Adam, Z., and Takahashi, Y. (2003). Coordinated regulation and complex formation of yellow variegated1 and yellow variegated2, chloroplastic FtsH metalloproteases in-


