A Molecular, Genetic and Biochemical Study o Protein Import in *Arabidopsis*

Amy Joy Baldwin

March 2004

Department of Biology,
University of Leicester

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgements

I would like to thank my supervisor Dr. Paul Jarvis for outstanding supervision, patience, encouragement and resilience throughout this project. I am also grateful to Tony Wardle, Henrik Aronsson, Jocelyn Bédard, Dave Stevenson, Sabina Kovacheva, Sybille Kubis, Ramesh Patel and Penny Dudley for their friendship, advice and assistance. I also wish to show my appreciation to the Biology department as a whole for providing an excellent facility within which to work. I am also grateful to Graham Benskin and June Saddington for the production and maintenance of my *Arabidopsis* plants, at the University of Leicester botanical gardens, and David Twell, Soonki Park, Natalie Allcock and Stephan Hyman for preparation of samples and images, by microscopy. Lastly many thanks must go to my family and friends who have been by my side either mentally or physically the whole time.

A Molecular, Genetic and Biochemical Study of Plastid Protein Import in *Arabidopsis*

By Amy Baldwin

Abstract

Plastids are double membrane-bound organelles that evolved from a cyanobacterial ancestor. The majority of the plastid genes are nuclear-encoded, and plastid proteins are post-translationally imported by translocons at the outer (Toc) and inner (Tic) envelope membranes of the chloroplast. Molecular, genetic and biochemical techniques were used to characterise a family of components of the plastid import apparatus, and a novel screen to identify new loci involved in the import process was conducted.

Toc75 is the protein translocation channel, and there are three Toc75-related sequences in *Arabidopsis*. Fractionation experiments revealed that two of these sequences, *atTOC75-III* and *atTOC75-IV*, encode integral proteins of the outer envelope membrane. The third family member, *atTOC75-I,* is a pseudogene. *atTOC75-III* expression is greatest in young, rapidly expanding tissue, whereas *atTOC75-IV* is expressed constitutively throughout development at a lower level than *atTOC75-III.* The main orthologue of pea Toc75, atToc75-III, is essential for viability. Homozygous *toc75-IV* mutant plants have no visible phenotype. However, etioplasts from both *toc75-IV* seedlings and *atTOC75-IV* overexpressing lines, do display an abnormal phenotype.

Yellow-green *plastid protein import 1* (*ppi1*) plants are null for atToc33, a component of the plastid protein import apparatus. Homozygous *ppi1* mutant seed were treated with the chemical mutagen ethyl methanesulfonate, and screened visually for partially or full restoration of the wild-type (green) phenotype. Five mutant lines were isolated which contain recessive, heritable, single-locus mutations that suppress the *ppi1* phenotype. These five mutations were found to occupy two different loci, namely *sp1* and *sp2*. All five mutations suppress the *ppi1* phenotype in every respect tested. atToc33 is proposed to preferentially control the import of photosynthetic precursor proteins. Characterisation experiments demonstrated that mutant *sp1 ppi1* seedlings contain more photosynthetic proteins than *ppi1,* and that *sp1 ppi1* chloroplasts are able to import the photosynthetic precursor, preOE33, more efficiently than *ppi1* chloroplasts.

Contents

General Introduction

1. *Arabidopsis*, The Model Plant

Arabidopsis is the organism of choice in many laboratories for studying basic biologically processes in plants. It is a small, green angiosperm of the Brassicaceae family that grows naturally in temperate regions. First described by Johannes Thal in 1577, it has been recognised as a model organism only in the last 50 years. *Arabidopsis* has many important advantages for basic research in genetics and molecular biology, which include a fully sequenced small genome (125 Mb total)(*Arabidopsis*-Genome-Initiative, 2000); extensive genetic maps of all five chromosomes; a high seed yield; the ease of cultivation in a small space; rapid development (generation time is around six weeks); ease of cross pollination and full fertility of hybrids; efficient transformation methods using *Agrobacterium tumefaciens*; a large number of mutant lines and genomic resources and a multinational research community of academic, industrial and government laboratories.

There are many resources available in Britain and throughout the world for *Arabidopsis* research. Huge libraries of insertion mutants can be searched online and seed requested from the *Arabidopsis* Biological Research Centre (ABRC) and Nottingham *Arabidopsis* Stock Centre (NASC). In the UK, the Genomic *Arabidopsis* Resource Network (GARNet) provides access to high throughput genomic technologies, such as metabolics, transcriptomics and proteomics. It is part of the Investigating Gene Function (IGF) initiative funded by the Biotechnology and Biological Sciences Research Council (BBSRC), which aims to characterise gene functions of *Arabidopsis* and other model organisms with fully sequenced genomes, such as *Drosophila.*

Extensive resources are also available online. The National Centre for Biotechnology Information (NCBI) based in Bethesda, Maryland, USA, incorporates Entrez, a text-based search and retrieval system for a variety of different literature, nucleotide and protein sequence, genome, taxonomy and strucural databases. For example, the PubMed database includes over 14 million citations for biomedical articles dating back to the 1950's. The citations typically have links to full text articles and other related resources, including articles describing *Arabidopsis*. The BLAST resource can be used to find sequences related to a specific polypeptide or nucleotide (Altschul et al., 1997) and conserved domains can be identified within a protein using their structural search engines. The entire *Arabidopsis* genome sequence can be searched and analysed via this website. The *Arabidopsis* Information Resource (TAIR) is the central point of the online resources for *Arabidopsis*.

TAIR provides news and information alongside search engines for insertional mutants, molecular markers, sequence analysis tools, and many external links.

2. What are Plastids?

Plastids are membrane-bound organelles that are found in plants and algae. Their functions include photosynthesis, starch biosynthesis and storage, tetrapyrrole, cartenoid, lipid and amino acid biosynthesis, and nitrate and sulphate assimilation (Weeden, 1981). With so many different functions, plastids have many structurally different forms, which interconvert depending on the cell type and the environment. Proplastids are undifferentiated plastids that can change into any other type of plastid, such as chloroplasts, which carry out photosynthesis. Amyloplasts are generally found in root cells and are responsible for starch storage. Chromoplasts are pigmented and thus used to draw attention to the fruits and flower petals. Etioplasts which are found in plants growing in the dark. It is also possible for plastid types to interconvert; for example, light stimulates the amyloplasts in potato tubers to convert to chloroplasts, which is visualized when potatoes turn green.

3. The Origin of Plastids

Plastids evolved through a process known as endosymbiosis (Margulis, 1970). Endosymbiosis describes the relationship between two organisms, the endosymbiont living-together, within another organism, the host. Most, if not all, microbial symbioses have a chemical basis; *i.e.* compounds produced by one partner are made use of by the other (Hoffmeister and Martin, 2003). Furthermore, the endosymbiont and host genomes evolve to accommodate the other organism. The relationship between the pea aphid and the γ-proteobacteria, *Buchnera sp.,* can be used to illustrate primary endosymbiosis. *Buchnera* live within specific aphid cells, passing from one generation to the next in the egg. The bacteria are totally dependent on the aphid because they no longer have genes for non-essential amino acid biosynthesis; the host aphid produces these amino acids. The aphid, on the other hand, needs *Buchnera* to synthesize essential amino acids, since sap, the aphids food source, is poor in amino acids. As some of the non-essential amino acids are precursors for the synthesis of essential amino acids, the biosynthesis pathways of the host and symbiont are inextricably intertwined (Shigenobu et al., 2000). They have been co-evolving for 200 million years and now neither the host nor symbiont can survive without the other (Baumann et al., 1995).

The endosymbiotic event in which a eukaryotic cell engulfed a cyanobacterium occurred between 1.2 (Butterfield, 2000) and possibly 1.5 (Jaraux et al., 2001) billion years ago. Three distinct groups of eukaryotes containing the evolved descendents of cyanobacteria, plastids, have evolved from this initial event. These three groups are the green algae (including land plants), the red algae and glaucophytes. Glaucophytes or glaucocystophytes are a small group of algae, the plastids of which still retain the peptidoglycan wall of the endosymbiotic ancestor. Interestingly, organisms exist that have acquired plastids via secondary and even tertiary symbioses (Palmer and Delwiche, 1996). Chlorarachniophyte, for example, is an amoeboid alga that originated from a eukaryote that engulfed another eukaryote, which already contained a plastid. Dinoflagellates are even more complicated with many examples of tertiary endosymbioses, whereby chlorophytes, cryptophytes, heterokont and hapotophytes have been engulfed (Douglas, 1998).

Evolution has united the host and the endosymbiont into a bipartite, co-evolving single organism. Ultimate control of the cell lies in the nucleus, but the plastids feed back information to the nucleus (Jarvis, 2003). The genomes of present day plastids encode ~5% the number of genes as their free-living, cyanobacterial cousins (Abdallah et al., 2000). The plastid genes have either been lost, such as the genes for peptidoglycan cell wall synthesis, or transferred to the nucleus in an on going process (McFadden, 1999). As a result, the *Arabidopsis* genome now contains ~4,500 genes acquired from the cyanobacterial ancestor of plastids, accounting for ~18 % of the nuclear DNA. Interestingly, not all the products of genes transferred from the plastid to the nucleus are targeted back to the plastid, and many localize in other cellular compartments (Abdallah et al., 2000; Martin et al., 2002; Leister, 2003).

One likely explanation for the transfer of plastid genes to the nucleus is Muller's Ratchet; *i.e.*, genomes that only replicate vertically, such as plastid and mitochondrial genomes, and do not share genes horizontally, can accumulation deleterious mutations (Muller, 1932; Moran, 1996). On the other hand, the nuclear genome reproduces sexually and deleterious mutations can be lost through recombination (McFadden, 1999). Plastids may have retained a small portion of their genomes because they require 'on site' structural proteins to maintain the redox balance within bioenergetic membranes. More elaborate regulation involving the nuclear genome may not be fast enough to maintain a redox balance, so these proteins are synthesized within the plastid when and where they are needed (Race et al., 1999).

The transfer of genes to the nucleus has created an interesting problem, because the products of the transferred genes must be localised to the plastid (McFadden, 2001). Plastid proteins are imported from the cytosol by two protein complexes situated in the outer and inner membranes of the plastid double envelope membrane (Keegstra and Cline, 1999; Chen et al., 2000b; Hiltbrunner et al., 2001a; Jarvis and Soll, 2002). The outer envelope membrane protein-conducting channel shares significant sequence identity, structure and gating properties with the *Synechocystis sp.* strain PCC6803 protein, SLR 1227, an export channel (Kaneko et al., 1996; Bolter et al., 1998b; Reumann et al., 1999). *Synechocystis* is a cyanobacterium, and therefore has the same ancestor as plastids. Hypothetically, transfer of the export channel gene to the nucleus would leave the channel protein on the opposite side of the envelope and, therefore, it could potentially operate as an import channel. The proteins that were originally exported by the channel, or their targeting sequences, when fused to the amino (N)-terminus of proteins destined to the plastid could act as plastid localization signals, or transit peptides, that target the plastid proteins to the import channel (McFadden, 2001; Huang et al., 2004). Once the protein is imported into the plastid, the transit peptide sequence is cleaved by a peptidase (Woolhead et al., 2000; Richter and Lamppa, 2003). The peptidase also has a *Synechocystis* homologue, so perhaps the cyanobacterial ancestral peptidase processed the substrate prior to export.

4. Plastid Protein Import

Research into the mechanism by which nuclear encoded plastid proteins are localized to the plastid began in the early 1970s. An 'envelope carrier hypothesis' was proposed by John Ellis (1973) in which he suggested that proteins in the chloroplast envelope mediate import of nuclear-encoded plastid proteins from the cytosol to the stroma. Subsequently, it was discovered that plastid proteins were imported post-translationally, and that precursors contain a cleavable N-terminal targeting sequence that is both necessary and sufficient for targeting to the plastid (Chua and Schmidt, 1978; Highfield and Ellis, 1978; Grossman et al., 1982; Robinson and Ellis, 1985; Flugge and Hinz, 1986; Smeekens et al., 1986; Pain and Blobel, 1987; Keegstra, 1989). Further investigation into plastid protein import has revealed that prior to reaching the plastid, the precursor proteins are maintained in an import-competent state by chaperone proteins (Waegemann et al., 1990; May and Soll, 2000a; Rial et al., 2000), and that precursor proteins are recognised (Sabatini et al., 1982; Cline et al., 1985) and then translocated across the double envelope membrane by two protein complexes: the translocon at the outer envelope membrane of chloroplasts (Toc), and the translocon at the inner

envelope membrane of chloroplasts (Tic) (Schnell et al., 1997). GTP (and low concentrations of ATP) in the cytosol or intermembrane space is required for precursor proteins to bind to the translocon complex, and stromal high concentrations of ATP are required for precursor translocation into the stroma (Olsen et al., 1989; Theg et al., 1989; Young et al., 1999; Schleiff et al., 2003a). Upon translocation into the stroma, the transit peptide is cleaved and the protein is either folded into its final conformation or targeted to one of several internal compartments of the plastid (Keegstra and Cline, 1999; Robinson et al., 2001). Translocon components were originally identified in 1994 and 1995 in several different laboratories (Hirsch et al., 1994; Perry and Keegstra, 1994; Schnell et al., 1994; Wu et al., 1994; Seedorf et al., 1995) and have since been characterised by biochemical, molecular and genetic techniques. The process of plastid protein import is described in greater detail in the following sections. Plastid protein import is discussed in three separate sections as follows: targeting proteins to the plastid (section 5), energy requirements for import (section 6) and the translocon apparatus (section 7).

5. Targeting Preproteins to the Plastid

5.1. Transit Peptides

Since the completion of the *Arabidopsis* genome sequence (The-*Arabidopsis*-Genome-Initiative, 2000), it has become evident that ~3,500 different proteins are directed to the plastid during the typical life of a plant (Abdallah et al., 2000; Leister, 2003). The vast majority of proteins destined for the plastid interior have a cleavable N-terminal transit peptide, which is both necessary and sufficient for import into the plastid (Keegstra, 1989) and functions by interacting with the translocon components and the processing machinery within the plastids. Transit peptides also interact with soluble, cytosolic proteins such as hsp70 and 14-3-3 homologues as described in the next section (5.2.).

No significant sequence identity has been found in the primary structure of transit peptides (Bruce, 2001). However, it may be that sequence identity exists within groups of transit peptides, each different group of transit peptides recognising a different receptor at the plastid surface. If the transit peptide sequences within each group were analysed perhaps a consensus sequence could be identified. Transit peptides do however, typically contain three distinct regions (Bruce, 2001). These are, a short N-terminal domain of ~10 residues beginning with methionine-alanine- and finishing with a glycine or proline, a central domain enriched in

serine or threonine but lacking acidic residues, and a third arginine-rich, C-terminal domain that potentially forms an amphiphilic β-strand (Claros et al., 1997). It is proposed that each domain originated from cyanobacterial 'signal' sequences, *i.e.*, each domain was originally a protein that was exported by the original cyanobacterial export pore.

Since transit peptides are divergent in amino acid sequence, composition and length, it is the secondary structure that is thought to be responsible for the targeting ability of transit peptides (Bruce, 2001). It has been shown that some transit peptides form helical structures in the presence of some lipids. For example, the ferredoxin transit peptide forms helical structures in the presence of negatively charged lipids and mixed micelles, but not in aqueous solution (Horniak et al., 1993; Wienk et al., 2000), and the last ten amino acids of the tobacco small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (SSU) transit peptide have been shown to interact with artificial bilayers containing monogalactosyl diacylglycerol (MGDG), a chloroplast specific lipid (Pinnaduwage and Bruce, 1996). Interestingly, no conformational changes are observed in the presence of phosphatidylcholine, a lipid found more widely in the cell (Horniak et al., 1993). Therefore, prior to reaching the outer envelope translocon complex, the precursor may identify the plastid via the lipid composition of the membrane, or at least the lipids may provide an environment in which the transit peptide forms a structure that is recognisable by stromal or translocon receptors.

In order to allocate roles to different domains within transit peptides, *in vitro* import experiments have been carried out with transit peptides altered in various ways. It has been reported that mutations in specific regions of the pea SSU transit peptide led to less efficient import into isolated chloroplasts *in vitro* (Reiss et al., 1989). However, this is not the case *in vivo* (Rensink et al., 2000). The most widely-used transit peptide in these studies has been the ferredoxin transit peptide. Pilon et al. (1995) conducted sequence comparisons with other transit peptides, and found that features of the ferredoxin transit peptide can be found in other transit peptide sequences, which led them to propose a general structural design for all transit peptides. They proposed that the N-terminus of the ferredoxin transit peptide mediates insertion into mono-galacto-containing lipids whereas the C-terminus mediates the recognition of negatively charged lipids (Pilon et al., 1995). Additionally, Rensink et al. (2000) implicated different areas of the ferredoxin transit peptide in the binding and translocation of the preprotein into the chloroplast. They proposed that the N-terminal end of the transit peptide is required for recognition and binding and that the C-terminal amino acids are necessary for translocation and association with Tic components. A definitive mechanism

for the role transit peptides play in the targeting of precursor proteins to the plastid has yet to be established.

Interestingly, part of the mature protein has been reported to influence the *in vitro* import and intraorganellar targeting of chloroplast protein precursors. It has been demonstrated that if the carboxyl-terminal end of the mature protein is removed the precursor is not localised correctly inside the chloroplast (Ko and Ko, 1992). On another occasion the N-terminal end of the mature region reported to influence import (Rial et al., 2002).

To date, only one protein has been identified that is localized to the interior of the plastid without a cleavable N-terminal transit peptide. This protein is called chloroplast envelope Quinone OxidoReductase Homologue (ceQORH), a peripheral inner envelope membrane protein with no cleavable transit peptide (Miras et al., 2002). A region of the mature protein, but neither at the N- nor the C-terminal end, is responsible for its correct targeting. Other chloroplast proteins lacking transit peptides are those situated in the outer envelope membrane (Schleiff and Klosgen, 2001). These proteins do not pass through the Toc/Tic translocon complexes, but instead insert directly into the outer membrane in an energy-independent manner. Such proteins contain internal targeting sequences and examples include Com70 / SCE70, OEP14 and Toc34 (Wu and Ko, 1993; Tu and Li, 2000; Qbadou et al., 2003).

5.2. Cytosolic Factors

Nuclear-encoded plastid proteins are translated in the cytosol, before being imported into plastids. To prevent them folding into their mature conformation prematurely, cytosolic chaperone proteins bind the precursor proteins upon translation (Jackson-Constan and Keegstra, 2001). Although it has been shown that some proteins do import in an enzymatically active form (America et al., 1994), the small pore size of the Toc complex (14 to 26 Å) suggests that most proteins are required to be only partially folded in order for import to occur (Hinnah et al., 2002). Furthermore, the cytosolic factors may bind the precursor proteins to prevent deleterious enzymatic activity in the wrong cellular compartment.

The first evidence that soluble cytosolic factors were involved in plastid protein import came from experiments conducted on the large, hydrophobic light harvesting chloroplast protein (LHCP) precursor (Waegemann et al., 1990). The LHCP precursor solubilized in 8 M urea is not import competent; however, LHCP dialysised in the presence of soluble leaf extract after

urea treatment partially regained import competence. The factors that made the precursor import-competent were identified as a heat shock protein, hsp70, and another protein that requires ATP (Waegemann et al., 1990).

There are hsp70 proteins in both the cytosol and stroma of plant cells (Marshall et al., 1990). They do not respond to heat stress but maintain precursors in an import competent conformation in an ATP-dependent manner. Cytosolic hsp70 proteins have been shown to interact with a specific site in the preFNR transit peptide (Rial et al., 2000), and the pea stromal hsp70, CSS1, interacts with N-terminus of the SSU transit peptide (May and Soll, 2000a). Algorithms have been developed to search chloroplast transit peptide sequences for hsp70 binding sites, and 75% of the transit peptides searched were found to contain a predicted binding site (Rial et al., 2000).

Another family of proteins that bind precursor transit peptides are the 14-3-3 proteins, which are found in a wide range of biological systems and have a very high degree of sequence conservation. Typically, 14-3-3 proteins are involved in regulation via signal transduction and phosphorylation mechanisms (Aitken et al., 1992). The *Arabidopsis* 14-3-3 gene family encompasses ten members (Wu et al., 1997), and phylogenetic analysis reveals that plant 14-3-3 proteins are descended from two distinct ancestors, which divide the present day homologues into two separate gene classes. Transit peptides are phosphorylated on a serine or threonine residue and 14-3-3 proteins have been shown to bind close to this phosphorylation site (May and Soll, 2000a).

Heterooligomeric complexes containing 14-3-3, probably as a dimer, an hsp70 protein isoform, and the precursor of SSU have an apparent size of 200 kDa (excluding precursor protein) (May and Soll, 2000a). Dissociation of the complex requires ATP and precursors are imported three to four times more efficiently when in a complex with soluble factors in an *in vitro* import experiment, compared to free precursors (May and Soll, 2000a).

5.3. Lipid Interactions

MGDG and digalatosyldiacylglycerol (DGDG) together constitute approximately 80% of total chloroplast lipids (Hartel et al., 1997; Bruce, 1998). Other envelope membrane lipids include sulfolipid and phosphatidylglycerol (Bruce, 1998). The chloroplast envelope is the only cytosolically exposed membrane that contains the galactolipids, MGDG and DGDG

(Bruce, 1998). Therefore, as already mentioned, they are a unique feature of plastids these galactolipids may play a role in plastid protein targeting.

Experiments utilizing a DGDG synthase deficient mutant, *dgd1*, has been used to investigate the requirement for galactolipids (Chen and Li, 1998). DGDG synthase catalyses the addition of a second galactose molecule to MGDG to produce DGDG. Thus, the *dgd1* mutant contains MGDG but less than 10% of wild-type DGDG levels (Dormann et al., 1995). The pale-green colouration and reduced growth of *dgd1* plants suggests that there is a chloroplast defect (Hartel et al., 1997). Chloroplast outer envelope proteins are localised with equal efficiency in *dgd1* mutants and wild-type plants. However, preproteins destined for the interior of the plastid are imported less efficiently than in wild-type preproteins (Chen and Li, 1998). Import was shown to be defective at the binding/docking step (that is supported by 100 μm ATP) (Chen and Li, 1998). Therefore, this evidence suggests that DGDG has a role in the early stages of plastid protein import (also refer to 5.1).

Targeting sequences interact with artificial, protein-free membranes (Bruce, 1998), but when placed in an environment that mimics the membrane environment they undergo a conformational change that results in significant alpha-helical character. Since the primary structure of transit peptides is so divergent, it maybe the secondary structure, induced by the membrane lipids, that enables them to be specifically imported into the plastid (Horniak et al., 1993; Wienk et al., 2000). Since lipids have a regular arrangement, transit peptides may have the capability to disrupt and change the lipid organisation. It might be that the transit peptide acts on the lipids and causes a conformational change and allows the preprotein to access the import apparatus (Bruce, 1998).

6. The Three Stages of Import - Energy Requirements

In vitro, plastid protein import can be divided into three distinct stages based on the energy requirement for each stage. These stages are assumed to correspond to the sequence of events a preprotein undergoes to be imported into a plastid. Plastid protein import, unlike mitochondrial import, only utilizes the energy from nucleotide triphosphate hydrolysis and not the proton motive force (Theg et al., 1989).

Initially, preproteins recognise the translocon apparatus in an energy-independent reversible, manner. During this stage, cross-linking experiments show that precursor proteins associate

more closely with the receptors than the channel (Perry and Keegstra, 1994). Binding of the precursor protein to the translocon complex is irreversible, and involves the N-terminal end of the precursor feeding into the translocon complex to form the early-import intermediate. This process requires energy, which is provided by the translocon receptor from cytosolic GTP hydrolysis (Olsen et al., 1989; Young et al., 1999; Schleiff et al., 2003a). GTP hydrolysis is reported to be enhanced 100-fold in the presence of precursor protein (Olsen et al., 1989; Schleiff et al., 2003a). At this stage of import, the precursor mainly interacts with the channel protein in addition to the receptors and components of the inner membrane translocon (Perry and Keegstra, 1994).

Translocation of the precursor protein into the interior of the plastid requires higher amounts of energy. ATP from inside the plastid is used to translocate the precursor protein, N-terminus first, across the envelope membranes (Olsen et al., 1989; Theg et al., 1989). It has recently been shown that GTP is required to insert the preprotein through the translocon complex (Schleiff et al., 2003a). Schleiff and company suggest that the precursor is fed through the pore in a 'sewing machine-like' mechanism, from outside the plastid by a GTP-dependent component of the outer membrane translocon complex. High amounts of ATP are required in the stromal for translocation to occur probably by the stromal factors so that they pull the precursors into the plastid (Olsen et al., 1989; Theg et al., 1989). Energy is also required by the stromal processing peptidase to cleave the transit peptide from the precursor.

7. The Translocon Complexes

Translocon components were originally identified in 1994 and 1995 in several different laboratories (Hirsch et al., 1994; Kessler et al., 1994; Perry and Keegstra, 1994; Schnell et al., 1994; Wu et al., 1994; Seedorf et al., 1995) and have since been characterised by biochemical, molecular and genetic techniques. The two biochemical methods that were employed to identify the translocon components are i) chemical cross-linking, which involves the labelling of proteins that are in close proximity to precursor proteins with a heterobifunctional cross-linker (Perry and Keegstra, 1994), and ii) co-immunoprecipitation of proteins that directly interact with the SSU transit peptide fused to staphylococcal A protein (Schnell et al., 1994; Wu et al., 1994; Seedorf et al., 1995). The first method identified a 75 kDa protein and an 86 kDa protein. The 86 kDa protein has since been shown to be the proteolytic fragment of a larger 159 kDa protein (Bolter et al., 1998a; Chen et al., 2000a). The second approach also identified a 34 kDa protein and a homologue of a heat shock protein, hsp70. The 75 kDa, 159

kDa and 34 kDa proteins form the core Toc complex (Schleiff et al., 2003b). Translocon components are referred to according to a standardised nomenclature (Schnell et al., 1997), based on their location in the outer or inner plastid envelope membranes, and the molecular mass of the mature protein, e.g., Toc75 (translocon of the outer chloroplast membrane, 75 kDa).

7.1. The Toc Complex

The translocon at the outer envelope membrane of chloroplasts (Toc) consists of regulatory and/or receptor proteins that recognise precursor proteins, and the translocation channel which conducts proteins, across the outer envelope membrane and into the intermembrane space. Precursor proteins are translocated into the stroma by the translocon at the inner envelope membrane (Tic complex). Toc complexes are reported to localize to discrete regions of the chloroplast envelope (Subramanian et al., 2001). Electron microscopy studies have revealed localization of import sites at contact sites (Schnell and Blobel, 1993).

The entire Toc complex has recently been characterized in order to determine its structure, including the relative amounts of each component. It was found that each translocon is ~500 kDa with a ratio of one Toc159 subunit, four or five Toc34 subunits and four Toc75 subunits (Schleiff et al., 2003b). Complex electron micrography reveals that the particles are roughly spherical, which is consistent with the formation of a stable core complex, and that the particles form a solid ring 130 Å in diameter, with a less dense interior structure. The data suggested that each Toc complex has a finger-like central region that separates four curved translocation channels (Schleiff et al., 2003b).

7.1.1. Toc159

Toc159 is, most likely, the first translocon component to interact with cytosolic precursor protein. Under energy-independent conditions, cross-linking experiments showed that Toc159 is the translocon component most closely associated with bound precursor proteins (Perry and Keegstra, 1994; Ma et al., 1996; Kouranov and Schnell, 1997). This interaction is probably through the most N-terminal of the three Toc159 domains, the acidic (A) domain. The high proportion of acidic amino acids in the A-domain suggest that this domain interacts electrostatically with the positively charged transit peptides of precursor proteins. The other two domains of Toc159 are the central GTP-binding (G) domain and the C-terminal membrane anchor (M) domain (Chen et al., 2000a).

Toc159 is both a cytosolic and an integral outer envelope membrane protein (Hirsch et al., 1994; Kessler et al., 1994; Hiltbrunner et al., 2001b). Transfer of soluble Toc159 into the envelope membrane is mediated by the central G-domain of Toc159 and is entirely dependent on GTP hydrolysis (Bauer et al., 2002; Smith et al., 2002). Therefore, if atToc159 is unable to bind GTP, the protein does not insert into the envelope membrane, but remains in the cytosol (Bauer et al., 2002).

The Toc159 G-domain has also been shown to be able to dimerize with Toc34 *in vitro*. The G-domain shares high sequence identity with the GTP-binding domain of Toc34 (Kessler et al., 1994; Sun et al., 2002) and during Toc159 insertion these two homologous domains interact (Smith et al., 2002; Weibel et al., 2003). Integration of atToc159 into the Toc complex requires atToc33 GTPase activity (Wallas et al., 2003). The interaction between the two GTPase domains promotes the association of the atToc159 M-domain with the translocon.

The M-domain is 52 kDa, which is very large and so most likely is not a single membrane spanning domain. Thus, Toc159 may have an additional role, such as forming part of the translocation pore (Jarvis and Soll, 2002). Also, this domain has no extended hydrophobic regions, as would be expected if it were totally immersed within the membrane. The A- and the G-domains protrude into the cytosol (Hirsch et al., 1994; Kessler et al., 1994). In cross-linking experiments, the Toc159 M-domain associated with precursors more during the intermediate stage than during the energy-independent stage (Kouranov and Schnell, 1997). This suggests that the M-domain of Toc159 may be involved in preprotein conductance as well as preprotein recognition. This hypothesis has recently been re-enforced by complementation studies using Toc159 knockout mutant plants (*ppi2*), which are albino and have undifferentiated plastids (Bauer et al., 2000). Expression of the Toc159 M-domain alone was shown to partially complement the *ppi2* mutant phenotype (Lee et al., 2003). Instead of being albino, the complemented *ppi2* plants had half the chlorophyll of wild-type plants and their plastids were nearly fully developed. The M-domain localizes to the envelope membrane but not the cytosol in these transgenic plants (Lee et al., 2003). These data suggest that Toc159 along with Toc75 forms part of the protein conductance channel.

Hence, a simple model for the activity of Toc159 would be firstly interaction with a precursor protein in the cytosol, followed by association of Toc159 to the Toc complex with the help of

Toc34, then import of the precursor protein through the membrane channel partly formed by Toc159 itself, and perhaps recycling of Toc159 back into the cytosol.

7.1.2. Toc34

Toc34 is a chloroplast outer envelope membrane protein that controls preprotein recognition in a GTP-dependent manner (Kessler et al., 1994; Kouranov and Schnell, 1997; Jelic et al., 2002; Sun et al., 2002). As previously mentioned, Toc34 mediates the localization of cytosolic Toc159 to the chloroplast (7.1.1.) (Wallas et al., 2003) and like the other core components of the Toc complex, it was isolated due to its interaction with precursor proteins (Kessler et al., 1994; Schnell et al., 1994; Seedorf et al., 1995). Toc34 has also been shown to associate with Toc75 under non-reducing conditions (Seedorf et al., 1995).

Toc34 is an integral envelope membrane protein (Kessler et al., 1994; Seedorf et al., 1995), that is anchored by a simple C-terminal transmembrane α-helix (Qbadou et al., 2003). The N-terminal end of the protein, which protrudes into the cytosol has typical and conserved motifs of GTP-binding proteins, and includes the G1 to G4 domains of GTP-binding protein family (Bourne et al., 1991; Kessler et al., 1994; Seedorf et al., 1995). Overexpressed and purified Toc34 possesses endogenous GTPase activity (Seedorf et al., 1995). The rate of GTP hydrolysis was reported to be greatly stimulated by precursor protein (Jelic et al., 2002; Schleiff et al., 2003a). However, Toc34 is different from other members of the GTPase family because its G4 domain is not involved in nucleotide specificity or GTPase activity (Aronsson et al., 2003). Unlike other GTP-binding proteins, Toc34 binds the triosphosphate moiety of GTP with a higher affinity than the purine base (Jelic et al., 2002).

X-ray crystallography has shown that pea Toc34 bound to GDP and Mg^{2+} is able to dimerize (Sun et al., 2002). Toc34 also dimerizes in phosphate saline buffer solution at pH 7.2, which is more indicative of biological conditions. Arg128 is an essential residue for dimerization, and dimerization is important for GTPase activity. When arg128 is mutated to ala128, Toc34 (ala128) is exclusively monomeric and GTPase activity is significantly reduced compared to wild-type Toc34 (Sun et al., 2002). These data suggest that each monomer acts as a GTPase activating protein (GAP) on the other monomer (Sun et al., 2002). This is re-enforced by the Toc34 crystal structure, which has features that resemble those found in a small GTPase in complex with a GAP. Toc34 is also proposed to be regulated by phosphorylation (Sveshnikova et al., 2000a; Jelic et al., 2002). Phosphorylation at serine 113, which is close

to the hypothesised triosphosphate binding site, downregulates GTP-binding (Jelic et al., 2002). Conversely, GTP binding influences phosphorylation, but not GDP binding.

One proposed model for Toc34 preprotein recognition involves the precursor recognising unphosphorylated, GTP-bound Toc34, the exchange of GTP for GDP and the release of the precursor towards the translocation channel. Following release of the precursor GDP is also released, and either Toc34 is phosphorylated and becomes inactive or it binds GTP and directs another precursor protein to the translocon channel (Sveshnikova et al., 2000a). It is currently not clear whether Toc34 interacts with preprotein and/or places a role in the localization of cytosolic Toc159 to the plastid envelope membrane.

Like virtually all chloroplast outer envelope membrane proteins, Toc34 inserts directly into the membrane without the need for an N-terminal, cleavable transit peptide (Seedorf et al., 1995). Toc34 is targeted to the plastid envelope membrane by the C-terminal transmembrane α-helical membrane anchor, which is part of the mature protein (Qbadou et al., 2003). Specifically, ten hydrophobic amino acids and the neighbouring C-terminal hydrophilic amino acids are sufficient for proper chloroplast outer membrane insertion (Li and Chen, 1997; May and Soll, 1998). Interestingly, the signal can work in either orientation, at either end of Toc34, and even when the N-terminus of Toc34 is replaced by another hydrophilic passenger protein. The orientation (Ncytosol-Cin) of Toc34 is solely dependent on two positive charges on the N-terminal side of the transmembrane alpha-helix (May and Soll, 1998). Unlike other plastid outer envelope membrane proteins, insertion of atToc34 into the chloroplast outer envelope membrane is assisted by at least two proteinaceous components in the import system (Tsai et al., 1999).

7.1.3. Toc75

Topological studies of pea Toc75 protein structure based on proteolytic digestion, amino acid sequencing, hydrophobicity and computer modelling showed that the protein has a β-barrel structure, comprising either 16 β-strands (Hinnah et al., 1997; Sveshnikova et al., 2000b) or 18 β-strands (Schleiff et al., 2003b). The voltage dependence of the channel open probability resembles those observed for other β-barrel pores re-enforcing the proposed role of Toc75 as the channel of the outer envelope membrane (Hinnah et al., 2002). Electrophysiological studies using heterologously expressed pea Toc75 reconstituted into liposomes have demonstrated that the channel has pore sizes of ~14 Å - 26 Å (Hinnah et al., 2002). Preproteins must therefore have a completely or partially unfolded conformation to pass

through the pore. Chaperone proteins, such as 14-3-3 and hsp70 homologues, are thought to prevent the proteins from folding into their mature forms, therefore, enabling them to pass through the translocon pores (May and Soll, 2000a).

The role of Toc75 is not confined to forming the pore through the envelope membrane. Additionally, pea Toc75 able to specifically recognize a transit peptide without the aid of other Toc components (Hinnah et al., 2002) and it is able to distinguish between different targeting sequences; pea Toc75 can differentiate between a chloroplast transit peptide, a synthetic presequence, and a mitochondrial presequence, based on both electrostatic and conformational interactions (Hinnah et al., 2002).

Toc75 is unique because it is the only chloroplast outer envelope protein known to have a transit peptide (Tranel et al., 1995). Additionally, the transit peptide is unusual in that it is bipartite; the N-terminal end targets the preprotein to the stromal and is removed by the stromal processing peptidase (Tranel and Keegstra, 1996; Inoue et al., 2001) and a poly-glycine tract in the C-terminus of the transit peptide is necessary for proper targeting to the outer envelope (Inoue and Keegstra, 2003). The N-terminal end of the Toc75 transit peptide has been shown to interact with some chloroplast membrane lipids but also a non-chloroplast lipid (Inoue et al., 2001). Toc75 imports into chloroplasts via the same mechanism as internally localized proteins since the Toc75 N-terminal peptide can inhibit import and binding of other precursor proteins (Inoue et al., 2001).

The first experimental evidence that Toc75 homologues are present in plant species other than pea came when the pea Toc75 antibody recognised a 75-kDa protein in isolated chromoplasts from the red bell pepper (Summer and Cline, 1999). Since then, Toc75 homologues have been identified in many plant species, including both monocot and dicot species (Davila-Aponte et al., 2003). Toc75 is also found in multiple plastid types and throughout plant development (Davila-Aponte et al., 2003). Pea Toc75 is expressed in photosynthetic tissue but to a lesser extent in stem and roots (Tranel et al., 1995). Expression is so high in young pea leaves that a partially processed form of pea Toc75 can be observed by western blotting (Tranel et al., 1995). Hence, Toc75 is believed to be part of a general or universal import pathway, common in different plastid-types and all plant species.

7.1.4. Toc64

The most recently identified Toc component, Toc64, copurifies on sucrose density gradients with the isolated Toc complex (Sohrt and Soll, 2000). Under low energy conditions (when the early import intermediate forms) Toc64 can be cross-linked to precursor protein, Toc159 and Toc34, and Tic components, which is consistent with Toc64 being a cytosolically facing, integral outer envelope protein. Like most chloroplast outer envelope membrane proteins Toc64 does not contain a cleavable transit peptide and does not require protease-sensitive components to insert into the membrane.

Analysis of the protein sequence reveals that Toc64 contains an amidase or indole acetamide hydrolase domain and three tetratricopeptide repeats (TPR), which are typically involved in protein-protein interacts. However, one of the three essential active site residues in the amidase domain, a serine, has been changed to a glycine and enzymatic assays detect no amidase activity (Sohrt and Soll, 2000). Additionally, since amidases are normally soluble proteins, it seems unlikely that Toc64 is an amidase. Based on what is known about Toc64, it has been proposed that Toc64 functions early in preprotein translocation, maybe as a docking protein for cytosolic cofactors (Sohrt and Soll, 2000).

7.2. The Translocon Complex at the Inner Envelope Membrane (The Tic Complex)

The Tic complex translocates preproteins from the intermembrane space into the stromal compartment. Precursor translocation by the Tic and Toc complexes probably occurs simultaneously, since Tic components have been shown to associate with Toc components at contact sites (Schnell and Blobel, 1993; Kouranov et al., 1998). Additionally, the Tic complex and associated proteins may be responsible for pulling precursor proteins into the stroma.

The Tic complex is not as well characterised as the Toc complex, in fact, it is unclear how many Tic complexes there are or how the proteins function in protein translocation. Tic components, like the Toc components, were first identified in association with precursor proteins in chemical cross-linking and immunoprecipitation experiments (Kessler and Blobel, 1996; Lubeck et al., 1996; Caliebe et al., 1997; Kouranov et al., 1998; Stahl et al., 1999). Isolated Tic complexes analysed by blue native gel electrophoresis included proteins of 36, 45, 52 (Tic55), 60, 100 and 110 kDa (Caliebe et al., 1997). The main consituents being Tic110 and ClpC (a member of the hsp100 family of chaperone proteins). Conversely, Tic20

and Tic22 were identified by chemical cross-linking to precursor protein, both Tic components interact with Tic110 and Toc components even in the absense of precursor proteins (Kouranov et al., 1998) but other Tic components are not present. The Tic components identified to date are described below.

7.2.1. Tic110

Two roles have been proposed for Tic110; to recruit stromal factors and a constituent of the inner envelope protein import channel. The structure and topology of Tic110 is not clear, circular dichroism indicates that Tic110 forms beta-sheets typical of pore proteins (Heins et al., 2002), alternatively, other results suggest that Tic110 has a high alpha-helical content (Inaba et al., 2003), and protrudes into the stromal compartment. The later hypothesis is reinforced by the fact that Tic110 has a large hydrophilic domain which is resistant to thermolysin, and therefore probably localised in the stroma (Jackson et al., 1998). It has been reported that precursors associate with this soluble region of Tic110 during the late stages of import, perhaps Tic110 acts as an anchor for preproteins where they can associate with processing factors (Inaba et al., 2003).

Evidence to support the role of Tic110 as a constituent of the inner envelope protein import channel stems from one publication. Heins and company (2002) demonstrated that Tic110 reconstituted into planar lipid bilayers forms a voltage-dependent high-conductant channel with a pore size of 1.7 nm. It was reported that transit peptides interact specifically with the channel, which caused the channel opening to flicker as the transit peptides blocked the pore. From this evidence, it is summized that Tic110 has a preprotein binding site and could potentially be the preprotein import channel at the plastid inner membrane (Heins et al., 2002).

7.2.2. Tic40

Tic40, previously identified as Toc36 and Cim/Com44, sits in the stromal side of the plastid inner membrane (Stahl et al., 1999). It is localized, like most other plastid proteins, via the general import pathway with the aid of a cleavable N-terminal transit peptide. A single putative transmembrane domain anchors Tic40 to the envelope membrane, while the remaining hydrophilic domain projects into the stroma (Chou et al., 2003). A TPR (see Toc64 - 7.1.4 - for explanation) motif and a domain with sequence similarity to co-chaperones Sti1p/Hop and Hip are part of this hydrophilic stromal portion. This data, along

with crosslinking experiments that show Tic40 to function during the same stage of import as Tic110 and hsp93 (a stromal chaperone) suggest that Tic40 acts as a co-chaperone in the stromal chaperone complex (Stahl et al., 1999; Chou et al., 2003).

7.2.3. Tic55

As previously mentioned, under native conditions Tic55 is found in complexes with at least five other inner envelope proteins, including predominantly, Tic110 and ClpC (Caliebe et al., 1997). Tic55 is predicted to contain a rieske-type iron-sulphur cluster and a mononuclear iron-binding site. In fact, a reagent that modifies rieske-type proteins inhibits precursor translocation across the inner envelope membrane. Tic55 interacts with precursor protein since it co-purify with a precursor protein in association with Toc86, Toc75, Toc34 and Tic110. Rieske-type proteins are typically involved in electron transfer, however, Caliebe et al. propose that Tic55 is a regulator protein of protein import at the inner envelope membrane, since various other proteins have been identified that use iron-sulphur clusters as prosthetic groups act as biosensors (Hidalgo et al., 1997).

7.2.4. Tic22 and Tic20

A second Tic complex is hypothesised since Tic110 is also found in complexes with Tic22 and Tic20 but not the other Tic components (Kouranov et al., 1998). Tic22 and Tic20 were identified by chemical cross-linking and shown to be closely associated with precursor proteins in the intermediate stage of import (Kouranov and Schnell, 1997). Two different Tic complexes maybe necessary for import of different types of protein, for example, soluble or hydrophobic proteins could behave very differently when passing through the protein-conducting channel or they may need to associate with different stromal factors for further targeting once they have translocated across the double envelope membrane. Since little is known about the Tic complex other import systems might give clues to the mechanism of plastid protein import. Mitochondria have two Tim complexes (translocon at the inner mitochondrial membrane); Tim23, which imports the majority of mitochondrial proteins, and Tim22, which imports integral inner membrane proteins.

Tic22 was the first protein to be identified that is localized in the intermembrane space (Kouranov et al., 1998). It is peripherally associated with the inner membrane and since Tic22 is found along with Tic110 and Tic20 in association with Toc complexes it might either act as a receptor for precursors as they emerge from the Toc complex or mediate contact sites

between the Tic and Toc complexes (Kouranov et al., 1998). Tic22, Tic20 and Tic110 do not form a stable complex but associate with each other only in the presence of the Toc complex whether precursor protein is present or not.

Interestingly, the Tic22 protein may not localize to the periphery of the inner membrane via the general import pathway. Like most precursors destined for the plastid interior, Tic22 precursor does have a transit peptide which is necessary and sufficient for localization, but Tic22 precursor import can not be out competed by import of stromal preprotein suggesting that the protein takes another route (Kouranov et al., 1999). Tic22 import is dependent on protease-sensitive components on the chloroplast surface and ATP (Kouranov et al., 1999).

The role of Tic20 in protein import, is supported by experiments using antisense plants, whereby, the level of Tic20 protein is reduced. Tic20 antisense plants have pale leaves, a reduced accumulation of plastid proteins and other growth defects, which correlate with the reduced levels of Tic20 protein (Chen et al., 2002). Such a severe phenotype suggests that Tic20 has an important role to play in plastid protein import and since Tic20 is a hydrophobic integral membrane protein it would seem logical that it may play a role in the preprotein conducting channel (Kouranov et al., 1998).

7.2.5. Tic62

The most recently characterised Tic component is Tic62, a membrane protein found in isolated Tic complexes with Tic110 and Tic55 (Kuchler et al., 2002). The N-terminal end of Tic62, which faces into the stroma, shares significant sequence identity with eukaryotic NAD(H) dehydrogenases and also the Ycf390-like proteins present in cyanobacteria and non-green algae (Kuchler et al., 2002). Due to these sequence similarities Tic62 is proposed to function as a regulator of protein import by sensing and reacting to the redox state of the organelle (Kuchler et al., 2002). The C-terminus of Tic62 contains a novel, repetitive module that interacts with a ferredoxin-NAD(P)+ oxidoreductase, the last enzyme in the photosynthetic pathway.

7.3. PIRAC

Another protein-conducting channel in the chloroplast envelope membrane has been reported, namely the protein import related anionic channel (PIRAC). Hypothesized to localize to the inner envelope membrane, PIRAC is a 50-picosiemens anionic channel that was identified by

patch-clamp analysis and interacts with the Tic complex (van den Wijngaard and Vredenberg, 1997). The identity of the channel component(s) and the channels function is not yet understood but it is known that precursors only interact with PIRAC in the presence of high amounts of ATP and that GTP alone does not stimulate precursor interaction (van den Wijngaard and Vredenberg, 1997; van den Wijngaard et al., 2000). Unlike the Tic complex, stromal factors are not required for the interaction of precursor with the 50-pS channel (van den Wijngaard et al., 2000).

8. Summary of Envelope Translocation, and Comparison with the Mitochondria Import System

Molecular genetic and biochemical techniques are rapidly increasing our understanding in the field of plastid protein import. The following paragraph summarizes our current knowledge of the subject. The majority of plastid proteins are encoded in the nucleus, translated on cytosolic ribosomes, and imported post-translationally into the plastid (Keegstra and Cline, 1999; Chen et al., 2000b; Hiltbrunner et al., 2001a; Jarvis and Soll, 2002). Most nucleus encoded plastid proteins have a cleavable, amino-terminal transit peptide, which is necessary and sufficient for targeting the precursor to the plastid interior. Prior to reaching the plastid, the precursor proteins are maintained in an import-competent state by chaperone proteins (Waegemann et al., 1990; May and Soll, 2000a; Rial et al., 2000). Precursor proteins are recognised and then translocated across the double envelope membrane by two protein complexes: the translocon at the outer envelope membrane of chloroplasts (Toc), and the translocon at the inner envelope membrane of chloroplasts (Tic). GTP in the cytosol or intermembrane space is required for precursor proteins to bind to the translocon complex, and stromal ATP is required for precursor translocation into the stroma (Olsen et al., 1989; Theg et al., 1989; Young et al., 1999; Schleiff et al., 2003a). Upon translocation into the stroma, the transit peptide is cleaved and the protein is either folded into its final conformation or targeted to one of several internal compartments of the plastid (Keegstra and Cline, 1999).

Greater insight can be gained into plastid protein import by comparing it with other protein import systems, which might employ similar mechanisms to recognise and translocate precursors. Valuable conclusions can be drawn from comparing mitochondrial protein import with plastid protein import since both organelles have evolved from a bacterial progenitor, which originally contained its own genes, the majority of which have now been relocated in the nucleus. Therefore, in both cases a new mechanism evolved to localise proteins within the

organelles. On the other hand, exact comparisons can not be drawn since chloroplast and mitochondrial import differ to a certain degree. The mitochondria require a membrane potential, $\Delta\Psi$, for initial of precursor translocation across the mitochondrial inner membrane (Martin et al., 1991), but the plastids do not (Flugge and Hinz, 1986). Plastids have developed a different mechanism than mitochondria for routing precursor proteins to the inner envelope membrane.

Nevertheless, by relating to the mitochondria system, it is possible to hypothesize a mechanism by which chloroplast precursor proteins are translocated into the chloroplast. Mitochondrial heat shock proteins function as a molecular motor, pulling the precursor across the mitochondrial membranes (Zhang and Glaser, 2002). Heat shock proteins are also thought to be involved in the import of chloroplast precursors, and like mitochondria, a high proportion of chloroplast transit peptides (82.5% of chloroplast transit peptides and 97.0% of plant mitochondrial presequences) contain a Hsp70 binding site (Zhang and Glaser, 2002). Therefore, it is possible to hypothesize, based on experimental data, that chloroplast precursors are also pulled into the organelle via a molecular motor involving Hsp70 proteins. Other insights into chloroplast protein import can also be gained by comparing it to the mitochondrial system, for example, the plastid outer and inner membranes most likely lie in close proximity during translocation (Schnell and Blobel, 1993; Schatz and Dobberstein, 1996).

This thesis endeavours to explore the field of plastid protein import and expand our knowledge into the import machinery and regulatory processes, through molecular genetic and biochemical analyses.

Results I

A Molecular-Genetic Study of the Toc75 Gene Family in *Arabidopsis*

1. Abstract

Toc75 is the protein translocation channel at the outer envelope membrane of chloroplasts. Toc75 was first identified in pea using biochemical approaches. We used a range of molecular, genetic and biochemical techniques to study the *Arabidopsis* Toc75 gene family. Fractionation experiments revealed that two of the three members of this family, *atTOC75-III* and *atTOC75-IV*, encode integral proteins of the outer envelope membrane. The third family member, *atTOC75-I,* is a pseudogene, due to a transposon insertion in the 5' end and numerous frame-shift and mis-sense mutations. The main orthologue of pea Toc75, atToc75-III, is essential for viability since homozygous atToc75-III knockout (*toc75-III*) embryos abort at the two-cell stage. *atTOC75-III* expression is greatest in young, rapidly expanding tissue, whereas *atTOC75-IV* is expressed constitutively throughout development at a lower level than *atTOC75-III.* Homozygous atToc75-IV knockout plants, *toc75-IV,* have no visible phenotype. However, etioplasts from both *toc75-IV* seedlings and *atTOC75-IV* overexpressing lines, do display an abnormal phenotype.

Division of labour. The authors for publication will be as follows: Amy Baldwin, Anthony Wardle, Penny Dudley, Soonki Park, David Twell and Paul Jarvis. Analysis of the *atTOC75-I* gene was carried out by Paul Jarvis. The *toc75-III-1* mutant was isolated from the Koncz laboratory collection by Penny Dudley and the *toc75-IV-1* mutant was isolated from the Sainsbury Laboratory *Arabidopsis* Transposon collection lines by Anthony Wardle. Soonki Park analysed the heterozygous *toc75-III* mutant siliques by Normaski Optics under the supervision of David Twell. Samples for light and electron microscopy were prepared by Natalie Allcock of the University of Leicester Electron Microscope Laboratory. All other experiments were carried out by Amy Baldwin.

2. Introduction

The majority of plastid proteins are translated on cytosolic ribosomes and subsequently imported into plastids (Keegstra and Cline, 1999; Chen et al., 2000b; Hiltbrunner et al., 2001a; Jarvis and Soll, 2001). An N-terminal transit peptide directs the preprotein specifically to the plastid. Upon translocation into the plastid, the transit peptide is cleaved and the protein is either folded into its final conformation or targeted to another compartment of the plastid (Keegstra and Cline, 1999). Preproteins are translocated through the plastid double membrane by two protein complexes: the translocon at the outer envelope membrane of chloroplasts (Toc), and the translocon at the inner envelope membrane of chloroplasts (Tic). Components of the Toc complex were first identified in pea using biochemical approaches (Hirsch et al., 1994; Perry and Keegstra, 1994; Schnell et al., 1994; Wu et al., 1994; Seedorf et al., 1995). To date, these include Toc34, Toc75, Toc64 and Toc159.

Toc34 and Toc159 are related GTPase proteins that share considerable sequence identity within their GTP-binding domains (Kessler et al., 1994). Toc159 exists in two forms: a soluble, cytosolic form, and an integral membrane form (Hiltbrunner et al., 2001b). Localization of Toc159 to the plastid envelope membrane is controlled by its own GTPase domain (Bauer et al., 2002) and is mediated by a direct interaction with the GTPase domain of Toc34 (Smith et al., 2002). Toc159 is thought to function at the chloroplast surface as the initial transit peptide receptor (Hirsch et al., 1994; Perry and Keegstra, 1994; Kouranov and Schnell, 1997). Apart from facilitating the localisation of Toc159 to the outer envelope membrane, Toc34 also interacts directly with the transit peptide of precursor proteins (Sveshnikova et al., 2000a; Schleiff et al., 2002). GTP-binding and phosphorylation of Toc34 may regulate import activity (Sveshnikova et al., 2000a). The most recently identified Toc component, Toc64, contains tetratricopeptide repeat (TPR) protein-protein interactions motifs in a large domain that is exposed to the cytosol. Therefore, this component may act as a docking site for the precursor-complexes either by, interacting with the precursor itself or by binding the accompanying chaperone proteins (May and Soll, 2000b; Sohrt and Soll, 2000). Following binding, the precursor proteins translocate through the channel formed by Toc75 into the chloroplast stroma (via the Tic complex) in a energy-requiring process.

The entire Toc complex has recently been characterized in order to determine its structure and the relative amounts of each component. It was found that each translocon is ~500 kDa and comprises one Toc159 subunit, four or five Toc34 subunits and four Toc75 subunits (Schleiff et al., 2003b). Two dimensional single particle analysis by electron micrography revealed

that the Toc complexes are roughly spherical, which is consistent with the formation of a stable core complex, and that the particles form a solid ring 130 Å in diameter with a less dense interior structure. The data also suggested that each Toc complex has a finger-like central region that separates four curved translocation channels.

Topological studies of pea Toc75 (psToc75) using proteolytic digestion, amino acid sequencing, hydrophobicity and computer modelling showed that Toc75 has a β-barrel structure, comprising either 16 β-strands (Sveshnikova et al., 2000b) or 18 β-strands (Schleiff et al., 2003b). Electrophysiological studies using heterologously expressed psToc75 reconstituted into liposomes have demonstrated that the channel has pore sizes of ~14 - 26 Å (Hinnah et al., 2002). Thus, preproteins must have a completely or partially unfolded conformation to pass through the pore. Chaperone proteins, such as 14-3-3 and hsp70, are thought to prevent the proteins from folding into their mature forms, therefore, enabling them to pass through the import apparatus (Kourtz and Ko, 1997; May and Soll, 2000b; Jackson-Constan and Keegstra, 2001). The voltage dependence of the Toc75 channel open probability resembles those observed for other β-barrel pores, re-enforcing the proposed role of Toc75 as the channel of the outer envelope membrane (Hinnah et al., 2002).

Interestingly, data suggest that the role of Toc75 is not confined to forming the pore through the outer envelope membrane. Electrophysiology measurements of Toc75 reconstituted into planar lipid bilayers showed that Toc75 is able to specifically recognize a transit peptide without the aid of other Toc components (Hinnah et al., 2002). Furthermore, Toc75 is able to distinguish between different targeting sequences: psToc75 can differentiate between a chloroplast transit peptide, a synthetic presequence and a mitochondrial presequence, based on both electrostatic and conformational interactions (Hinnah et al., 2002).

The first experimental evidence that Toc75 homologues are present in plant species other than pea came when a psToc75 antibody was used to identify a 75-kDa protein in isolated chromoplasts from the red bell pepper (Summer and Cline, 1999). Since then, Toc75 homologues have been identified in many plant species, including both monocot and dicot species (Davila-Aponte et al., 2003). Toc75 is also found in multiple plastid types and throughout plant development (Davila-Aponte et al., 2003). Pea Toc75 is expressed throughout photosynthetic tissue but to a lesser extent in stem and roots (Tranel et al., 1995). Hence, Toc75 is believed to be part of a general or universal import pathway, common in different plastid-types and all plant species.

The sequencing of the *Arabidopsis* genome has revealed that this species contains multiple homologues of the majority of translocon components (The-*Arabidopsis*-Genome-Initiative, 2000; Jackson-Constan and Keegstra, 2001); for example, pea Toc34 has two *Arabidopsis* homologues, termed atToc34 and atToc33. Reverse-genetic strategies have been used to examine the specific role of each component, and, in this case, it appears that atToc33 is responsible for the import of photosynthetic precursors, while atToc34 most likely imports non-photosynthetic precursors (Kubis et al., 2003). Interestingly, there are three different Toc75-related sequences in *Arabidopsis*, termed *atTOC75-I*, *atTOC75-III* and *atTOC75-IV* according to their chromosome locations (Jackson-Constan and Keegstra, 2001). In this study, we used molecular, genetic and biochemical techniques to study the functions of these three Toc75-related sequences.

3. Results

3.1. Expression and Structure of the Different Toc75 Family Members

Since the entire *Arabidopsis* genome has now been sequenced (The-*Arabidopsis*-Genome-Initiative, 2000), it is possible to identify all the members of the *Arabidopsis TOC75* gene family by sequence alignment analysis. Previous work revealed three Toc75-related sequences in the *Arabidopsis* genome: *atTOC75-I, atTOC75-III* and *atTOC75-IV* (Jackson-Constan and Keegstra, 2001). Our analysis of the *atTOC75-I* sequence revealed a gypsy-transposable element inserted into the 5' end of the gene (Figure 1a,b). Furthermore, *atTOC75-I* has several frame-shift and mis-sense mutations that suggest it is unlikely to be expressed (Figure 1a,b). To confirm that this is the case, six different primer combinations were used to analyse RNA from wild-type seedlings by RT-PCR. In each case, no *atTOC75-I* gene expression could be detected (data not shown). This evidence, in addition to the absence of any *atTOC75-I* homologous expressed sequence tags (ESTs) in the various online databases and the various mutations already mentioned, leads us to conclude that *atTOC75-I* is a pseudogene.

The *atTOC75-III* gene is known to be expressed since there are twenty-four *atTOC75-III* ESTs currently available in the databases. Conversely, there are no known *atTOC75-IV* ESTs. We therefore used RT-PCR to show that *atTOC75-IV* is expressed in wild-type *Arabidopsis* seedlings (Figure 2b). The correct structures of the two expressed *TOC75* genes was determined by sequencing EST number APZL59c10R from the Kazusa library (for *atTOC75-III*), and by sequencing 5' and 3' RACE products (for *atTOC75-IV*). We have deposited the correct *atTOC75-IV* cDNA sequence in Genbank, .

The structures of the three Toc75-related sequences, shown schematically in Figure 1b, provide further evidence that they are members of the same gene family. With only two exceptions, all three genes have the same exon / intron boundaries; exon 1 is absent from *atTOC75-IV* and there is an intron absent from the *atTOC75-I* pseudogene. The truncation of the 5' end of *atTOC75-IV* means that the atToc75-IV protein contains only eight of the total 16 predicted transmembrane domains (domain numbers 7 and 10 through 16; Figure 1a) found in the other atToc75 proteins (Sveshnikova et al., 2000b). This suggests that the atToc75-IV protein may work as a dimer in order to form a functional, full-sized channel.

Exon 1 of *atTOC75-III* (and presumably also of the pea Toc75 gene, *psTOC75)* encodes the transit peptide that mediates localization to the plastid outer envelope membrane. Since *atTOC75-IV* lacks exon 1, we analysed the atToc75-IV protein with the transit peptide prediction program, TargetP (Emanuelsson et al., 2000). TargetP predicts that atToc75-IV does not have a transit peptide, and so it seems likely that atToc75-IV is not directed to the outer envelope membrane in the same manner as atToc75-III and psToc75. However, it is not uncommon for plastid outer envelope membrane proteins to be inserted into the membrane without the aid of a transit peptide (Schleiff and Klosgen, 2001).

3.2. Phylogenetic Analysis

Our aim was to produce a well-supported phylogeny in order to clarify the evolutionary relationships between the *Arabidopsis* Toc75 homologues. Phylogenetic analysis of the Toc75 amino acid sequences produced a single most parsimonious tree (Figure 1c) of length 1872. The tree contains three main clades. The out-group containing SynToc75 (a *Synechocystis sp.* PCC 6803 protein, with 22% sequence identity to pea Toc75) and the SynToc75-related protein, atToc75-V, contains 362 mutations present in both proteins but not in Toc75 proteins from the two other main clades (Bolter et al., 1998b; Reumann et al., 1999). This was as expected since atToc75-V has previously been shown to be more closely related to its prokaryotic ancestors than to Toc75 from plants (Eckart et al., 2002). With the exception of atToc75-V, the Toc75 proteins from dicot plants group together in the same clade. As expected the Toc75 proteins from the two monocots, *Zea mays* and *Oryza sativa*, formed a single clade with 89 distinct mutations. The *Arabidopsis* Toc75 proteins form a subclade separate from the psToc75 protein, which is supported by a bootstrap value of 55%, meaning that 55% of 1000 possible trees grouped the *Arabidopsis* homologues together.

The phylogeny suggests that the *Arabidopsis TOC75* gene family formed after the pea and *Arabidopsis* species diverged. This means that pea does not necessarily contain genes closely related to either *atTOC75-IV* or *atTOC75-I.* Subsequent to speciation, *atTOC75* duplicated into *atTOC75-III* and another *TOC75* gene. Duplication of this second *TOC75* gene created *atTOC75-I* and *atTOC75-IV.* After *atTOC75-I* and *atTOC75-IV* diverged, various mutations lead to *atTOC75-I* becoming a pseudogene and *atTOC75-IV* becoming truncated.

3.3. Gene Expression Profiles of the Arabidopsis Toc75 Genes

In order to further elucidate their specific roles, we analysed the expression profiles of the *atTOC75-III* and *atTOC75-IV* genes by semi-quantitative RT-PCR. Since other *Arabidopsis* homologues of Toc-related proteins have previously been shown to display some degree of intra-family functional specialization (Kubis et al., 2003; Constan et al., 2004), it seemed possible that similar specialization might exist within the *Arabidopsis* Toc75 family. Expression of the *Arabidopsis TOC75* genes was normalized relative to the expression of the *atEIF4E1* gene (At4g18040), a translation initiation factor that has been shown to be expressed at comparable levels in all other tissue types (Rodriguez et al., 1998).

The expression of *atTOC75-III* was found to be strongest in young, rapidly dividing tissue, and weakest in mature or slow growing tissue, such as 28-day-old rosettes and roots (Figure 2). Unexpectedly, *atTOC75-III* expression is comparatively low in 5-day-old etiolated plants that are rapidly growing. This suggests that atToc75-IV, the only other close homologue of psToc75 that is expressed in *Arabidopsis*, may have a particularly important role in etioplast protein import. The expression of *atTOC75-IV* is more uniform than *atTOC75-III.* Expression of *atTOC75-III* varies 10-fold, while *atTOC75-IV* expression varies just 2-fold.

The data shown on Figure 2 does not provide any information on the relative expression levels of the two genes. However, we believe that the level of *atTOC75-IV* expression is lower than *atTOC75-III* expression, for several reasons. Firstly, we were not able to detect the *atTOC75-IV* transcript by northern blotting (data not shown). Additionally, amplification of *atTOC75-IV* during RT-PCR required five more cycles than amplification of *atTOC75-III* (and *atEIF4E1)* to give a product of a similar intensity (data not shown). Furthermore, there are no ESTs available for *atTOC75-IV* compared to the twenty-four *atTOC75-III* ESTs that are currently available. These reasons lead us to the conclusion that *atTOC75-III* is expressed at a significantly higher level than *atTOC75-IV.*

Like *atTOC75-III, psTOC75* is also expressed most strongly in young, rapidly dividing tissue (Tranel et al., 1995). Additionally, immunoblot data indicated that psToc75 protein accumulates to a lesser degree in roots than green tissue (Tranel et al., 1995). We observed lower *atTOC75-III* expression in roots tissue than young, green tissue (Figure 2). Based the level and pattern of expression, and the high sequence identity with *psTOC75*, we propose that *atTOC75-III* is the main orthologue of *psTOC75.*

3.4. atToc75-III and atToc75-IV are Integral Envelope Membrane Proteins

Pea Toc75 and atToc75-III are chloroplast outer envelope proteins, which are targeted to chloroplasts by N-terminal, bipartite transit peptides (Kessler et al., 1994; Perry and Keegstra, 1994; Tranel et al., 1995; Hiltbrunner et al., 2001b; Inoue and Keegstra, 2003). Like psToc75, atToc75-III is processed via an intermediate form, which can be observed when *in vitro* translated atToc75-III precursor protein is imported into isolated chloroplasts from pea (Figure 3a). It is unlikely that atToc75-IV has an N-terminal transit peptide because the 5' end of the *atTOC75-IV* gene is truncated (Figure 1), and TargetP does not predict that atToc75-IV has a chloroplast transit peptide (Emanuelsson et al., 2000). However, psToc75 and atToc75-III are unique in that no other chloroplast outer envelope membrane proteins identified to date have a transit peptide. Most outer envelope proteins lack a cleavable transit peptide and insert directly into the envelope (Schleiff and Klosgen, 2001). Common features of these proteins are a small to moderate size and in some cases an extended hydrophobic region (Cline and Henry, 1996; Schleiff and Klosgen, 2001). The atToc75-IV protein could be targeted by the same mechanism since it is relatively small, 44 kDa, and appears to have short hydrophobic regions at its N-and C-terminal ends.

To determine if atToc75-IV is a chloroplast protein, we incubated *in vitro* translated atToc75-IV with isolated chloroplasts under import conditions, and then treated the chloroplasts with the protease, thermolysin, which cannot penetrate the envelope membranes (Cline et al., 1984). atToc75-IV protein is substantially more resistant to thermolysin following incubation with chloroplasts yielding protected fragments, than in the absence of chloroplasts (Figure 3b). Therefore, chloroplasts protect atToc75-IV, to a certain extent, from protease digestion, suggesting that atToc75-IV is a chloroplast protein.

The topology of Toc75 has previously been determined by limited proteolytic digestion, N-terminally sequencing of the resulting fragments, and computer modelling (Sveshnikova et al., 2000b). The model predicts that Toc75 transverses the outer chloroplast envelope membrane with 16 amphiphilic beta sheets. We observed partial digestion of the atToc75-IV protein, resulting in fragments of approximately 15, 20, 23, 25 and 27 kDa (Figure 3b). The distinct fragments that increase in size in regular accruements, suggests that atToc75-IV is susceptible to thermolysin at regular intervals along its sequence. This observation could be explained if the protein passes in and out of the chloroplast envelope membrane. Since the proteolytic fragments were not N-terminally sequenced it is not possible to determine whether atToc75-IV has the same topology as psToc75. However, of the three thermolysin-sensitive

sites observed by Sveshnikova et al. (2000b), two are upstream of where the atToc75-IV and psToc75 protein sequences align, but the last one lies 17.3 kDa from the N-terminal end and 26.3 kDa from the C-terminal end. If atToc75-IV does have a similar topology to Toc75, this thermolysin cleavage site could be present in both proteins and explain the thermolysin sensitive ~15 and 27 kDa fragments.

Next, fractionation experiments were conducted to determine if atToc75-IV is an integral envelope membrane protein. The atToc75-III protein was used as a positive control, and was detected by immunoblotting with antibodies raised against either recombinant psToc75 (Tranel et al., 1995) or an atToc75-III specific peptide (Figure 3c). Previously, immunoblotting was used to demonstrate that both antibodies detect atToc75-III protein (Figure 3e). *In vitro* translated atToc75-IV, like atToc75-III, remains associated with isolated chloroplast membranes even after washing with high salt and high pH solutions, suggesting that it is indeed an integral membrane protein (Figure 3c).

Further fractionation experiments were carried out to determine whether atToc75-IV is an envelope or thylakoid membrane protein (Figure 3d). We expected atToc75-IV to be a chloroplast envelope protein for several reasons. Firstly, atToc75-IV is partially sensitive to thermolysin, which does not penetrate the envelope membranes, indicating that atToc75-IV is accessible from outside the chloroplast (Figure 3b). Secondly, to date, only one protein, the chloroplast envelope quinone oxidoreductase homologue (ceQORH), has been shown to be targeted to the chloroplast interior without the aid of a transit peptide (Miras et al., 2002), and it therefore seemed unlikely that atToc75-IV would be thylakoid localized. Thirdly, the chloroplast envelope is where the atToc75-IV homologues, atToc75-III and psToc75 are localized in the outer envelope membrane (Tranel et al., 1995; Hiltbrunner et al., 2001b). To determine the localization of atToc75-IV, we isolated envelope and thylakoid membranes from isolated chloroplasts. atToc75-IV protein was detected in envelope membranes (and total membrane, and total chloroplast preparations), but not in isolated thylakoid membranes (Figure 3d). The control envelope proteins, atToc75-III and atTic110, and the thylakoid control protein, LHCP, were all detected in total chloroplast and total membrane preparations, and the appropriate membrane preparation by immunoblot analysis (Figure 3d). Therefore, we conclude that atToc75-IV is an integral protein of the chloroplast envelope.

3.5. atToc75-IV Inserts Spontaneously into the Chloroplast Envelope

Most outer envelope proteins are synthesized to their mature size without a cleavable transit peptide (Schleiff and Klosgen, 2001). The majority of these proteins do not require energy or any thermolysin-sensitive components for targeting to the chloroplast (Tu and Li, 2000), so it is envisaged that they insert directly into the chloroplast outer envelope rather than utilizing the Toc/Tic complex. To determine if the insertion of atToc75-IV into the envelope membrane of chloroplasts requires energy, we conducted an import assay in the absence of ATP (Figure 4). *In vitro* translated atToc75-IV and preSSU precursors, depleted of ATP, were incubated with ATP-depleted chloroplasts according to previously described procedures (Olsen et al., 1989; Aronsson and Jarvis, 2002). preSSU was found to import normally when ATP was added (Figure 4). However, if ATP was not added, preSSU was not processed, confirming that ATP had been successfully removed from both the chloroplasts and the *in vitro* translated precursors. By contrast, atToc75-IV localized to the membrane fraction whether or not ATP was present (Figure 4), demonstrating that there is no energy requirement for the insertion of atToc75-IV.

3.6. Identification of atToc75-III and atToc75-IV Knock-out Mutants

In order to characterise the *in vivo* roles of the atToc75-III and atToc75-IV proteins, we identified T-DNA or transposon knockout lines for atToc75-III and atToc75-IV. Two independent *toc75-III* mutant alleles were obtained, *toc75-III-1* from the Koncz Laboratory T-DNA collection (Rios et al., 2002) and *toc75-III-2* from Salk Institute T-DNA collection (Alonso et al., 2003). The *toc75-III-2* could be identified by PCR only, since the T-DNA-borne kanamycin-resistance gene was silenced, a feature of ~20% of the Salk Institute lines (Alonso et al., 2003). Similarly, two independent *toc75-IV* mutant alleles were obtained, *toc75-IV-1* from the Sainsbury Laboratory *Arabidopsis* Transposon (SLAT) dSpm collection (Tissier et al., 1999) and *toc75-IV-2* from the Syngenta T-DNA collection (Sessions et al., 2002). In each case, the independent alleles behaved identically, indicating that the phenotypes were due to the insertions and not secondary, unlinked mutations. Unless specifically stated otherwise, the *toc75-III-1* and *toc75-IV-1* alleles were used to generate all the data presented below.

All four knockout lines contained single-site insertions. Seed from heterozygous *toc75-IV* plants segregated approximately three resistant plants for every one sensitive plant when plated on selective medium (Table 2.1). Conversely, seed from heterozygous *toc75-III-1*

plants segregated approximately two resistant plants for every one sensitive plant when plated on selection, and seed from heterozygous *toc75-III-2* plants segregated approximately two heterozygous plants for every one wild-type plant when tested by PCR (Table 2.1). In addition, all plants tested by PCR (60 from *toc75-III-1* and 40 from *toc75-III-2*) were shown to be heterozygous for the appropriate T-DNA insertion. We therefore concluded that homozygous atToc75-III knockout mutations are lethal during an early stage of development.

None of the knockout mutant seedlings (heterozygous *toc75-III* and homozygous *toc75-IV*) had an obvious visible phenotype (Figure 5b). Therefore, double mutants were created in order to assess for functional interactions between the homologues. Double mutants, heterozygous for *toc75-III* and homozygous for *toc75-IV*, do not have an obvious visible phenotype (Figure 5a). This observation is re-enforced by chlorophyll measurements (Figure 5b). Similarly, even in the compromised atToc33 knockout (*ppi1*) background, which displays a yellow-green phenotype (Jarvis et al., 1998), *toc75-IV* homozygous seedlings did not display an obvious visible phenotype (Figure 5).

3.7. The atToc75-III Knockout Mutation is Embryo Lethal

Since homozygous *toc75-III* seedlings could not be identified in the progeny of heterozygous *toc75-III* plants (Figure 5, Table 2.1), we wished to determine the stage at which development terminates in homozygous mutants. Upon looking in the siliques of heterozygous *toc75-III* plants, it was observed that aborted seeds were present at a frequency of ~25% (Table 2.2, Figure 6a). This indicates that the homozygous *toc75-III* mutations are embryo lethal.

To determine the exact stage at which development of homozygous *toc75-III* plants terminates, we used Nomarski optics to observe developing seeds in heterozygous *toc75-III* siliques of different ages (Table 2.3). Approximately 75% of seed were observed to develop normally, passing from the four-cell stage (Figure 6a), through to the eight-cell stage (Figure 6b) and the globular stage (Figure 6c,d), and then finally to the heart stage (Figure 6e) and onwards. However, ~25% of the seed in heterozygous *toc75-III* siliques do not develop beyond the two-cell stage (Figure 6, f-i); *i.e.*, when the normal seed were at the late globular stage (Figure 6d), ~25% of the developing seed had just two nuclei (Figure 6i). When the normal seed reached the heart stage (Figure 6e), ~25% of the seeds aborted (Figure 6j). Additionally, the suspensor of homozygous *toc75-III* seed does not develop normally. In Figure 6i, many smaller nuclei can be observed in the tissue surrounding the developing

embryo. Thus, we conclude that embryo development in *toc75-III* homozygotes arrests at the two-cell stage, and that the mutation ultimately leads to the collapse of the seed.

3.8. Plastid Ultrastructure in the *toc75-IV* Mutant

Since homozygous *toc75-IV* seedlings have no obvious visible phenotype (Figure 5a,b), we sought to identify a mutant phenotype by analysing the ultrastructure of *toc75-IV* plastids. First of all, we analysed the cross-sectional area and overall shape of the plastids, the number of thylakoid membranes per granum, the overall shape, the number of starch grains, plastoglobuli and granal stacks per plastid of 10-day-old chloroplasts, were all statistically analysed, but, no significant differences were observed between *toc75-IV* and wild-type seedlings (data not shown).

The expression of *atTOC75-III* was found to be lowest in dark grown plants (Figure 2), suggesting that *atTOC75-IV* may have a more significant role in etioplasts. We therefore analysed etioplasts from 3-day-old, dark-grown, wild-type and *toc75-IV* seedlings. While the prolammelar bodies and prothylakoids looked normal in *toc75-IV* etioplasts, we observed that some etioplasts contained bodies of cytosol completely enclosed within the plastids. We scored the number of etioplasts with cytosolic inclusions and statistically analysed the data. Although ~20% of the wild-type etioplasts contained these cytosolic inclusions, both *toc75-IV* alleles contained significantly more cytosolic inclusions than wild-type etioplasts (Table 2.4). Given that the data is enumerate (discontinuous) it was analysed using the Chi-squared (χ^2) test, and, because there are only two categories of data (one degree of freedom) we applied the Yates correction. The null hypothesis that there was no significant difference between the wild-type and mutant data was tested (Table 2.4). In both cases, a significant difference was observed between our observed and the expected data (critical value of 3.84, p=0.05). Therefore, we rejected the null hypothesis and concluded that *toc75-IV* etioplasts contain more cytosolic inclusions than wild-type etioplasts.

3.9. Analysis of Plants Overexpressing the *atTOC75-IV* Gene

In order to investigate further the role of atToc75-IV, we transformed wild-type plants with a construct comprising the *atTOC75-IV* cDNA under the control of the cauliflower mosaic virus 35S promotor, using *Agrobacterium tumefaciens*. Thirteen independent lines transformed lines were identified that overexpressed the *atTOC75-IV* gene (Figure 2b). Light-grown plants, overexpressing *atTOC75-IV,* did not display an obvious visible phenotype (data not

shown). However, when we analysed etioplasts from dark-grown seedlings from two of these lines, we observed that etioplasts from seedlings overexpressing *toc75-IV* contained significantly less cytosolic inclusions than wild-type etioplasts (Table 2.4). The chi-squared test was used, to determine whether the mutant data deviated to a significant degree from the wild-type data, in the same manner as previously described. It was observed that the etioplasts of both lines overexpressing *atTOC75-IV* contained significantly less cytosolic inclusions than wild-type etioplasts. Reassuringly, we found that there was no significant difference between the wild-type etioplasts grown on separate occasions (χ^2-value 1.68, p-value >0.05). These data suggest that atToc75-IV has a role in etioplasts. It is not clear whether these structural defects are the result of the failure of *toc75-IV* etioplasts to import certain proteins - for example, cytoskeletal or division apparatus components - or the result of a specific response designed to increase the surface area of metabolically compromised plastids in order to facilitate the exchange of substances with the cytosol.

4. Discussion

Toc75 is the main protein import channel through the plastid outer envelope membrane. It works in concert with receptor proteins, and is capable of recognising, as well as translocating, precursor proteins into the chloroplast (Tranel et al., 1995; Hinnah et al., 2002). Due to the benefits of working with *Arabidopsis*, such as the availability of the complete genome sequence and ease of manipulation and cultivation (The-*Arabidopsis*-Genome-Initiative, 2000; Somerville and Koornneef, 2002), we used a range of molecular, genetic and biochemical techniques to characterise the *Arabidopsis* Toc75 gene family. Three Toc75-related sequences were previously identified in *Arabidopsis* (Jackson-Constan and Keegstra, 2001): here, we show that they are a pseudogene (*atTOC75-I*), the main homologue of pea Toc75, which is essential for viability (*atTOC75-III*), and another, shorter sequence that was not previously known to be expressed but appears to play a role in etioplasts (*atTOC75-IV*) (Jackson-Constan and Keegstra, 2001).

Phylogenetic analysis revealed that the *Arabidopsis* Toc75-related sequences most likely diverged subsequent to the speciation of *Arabidopsis* and pea. Therefore, it is not possible to predict whether or not pea also contains multiple Toc75 homologues. The *Arabidopsis* genome has been reported to have undergone extensive duplication and reshuffling (Blanc et al., 2000). These modifications might have led to the divergence of the atToc75 family, through duplication and subsequent mutations. Curiously, *atTOC75-I* has lost a single intron, subsequent to diverging from the other *TOC75* genes. The loss of a single intron has previously been observed in plants on multiple occasions (Huang et al., 1990; Kumar and Trick, 1993; Hager et al., 1996; Drouin and Moniz de Sa, 1997). One possible mechanism through which *atTOC75-I* might have lost an intron would be through the reverse transcription of processed cellular mRNA to produce a cDNA copy of any one of the *TOC75* expressed genes, which could then partially replace the endogenous copy through homologous recombination (Baltimore, 1985; Derr et al., 1991).

Based on the level and pattern of *atTOC75-III* expression, and the high sequence identity with *psTOC75*, we propose that *atTOC75-III* is the main orthologue of *psTOC75*. The importance of atToc75-III was also demonstrated by the lethality of *toc75-III* embryos. Embryos contain undifferentiated plastids that are normally inherited maternally by the plant zygote. These proplastids do not develop into chloroplasts until the heart stage, when the embryo begins to turn green (Apuya et al., 2001). Since *toc75-III* mutants do not develop beyond the two cell

stage, it seems likely that atToc75-III is vital for the import of proteins into proplastids. Plastid proteins are important for embryo development because plastids synthesize products that are utilized by the rest of the cell as well as the plastid, such products include sugars, fatty acids, terpenoids, and amino acids (Ohlrogge and Browse, 1995; Jarvis et al., 2000). Other mutants have previously been identified that link plastid biogenesis and embryo development (Tsugeki et al., 1996; Uwer et al., 1998; Apuya et al., 2001). The *schlepperless* (*slp*) mutant, for example, contains a T-DNA insertion in a chaperonin-60α gene, which is required for folding and assembly of plastid proteins (Apuya et al., 2001). The *slp* mutant embryos are morphologically and developmentally retarded, eventually dying before reaching maturity.

Our data suggests that the *atTOC75-IV* gene is expressed at a lower level than *atTOC75-III*. The multiple homologues of other Toc components also display differing levels of expression. For example, expression of *atTOC159* is ~10 times higher than its homologues, *atTOC132* and *atTOC120* (Bauer et al., 2000), and maximal expression of *atTOC33* is ~6 times higher than maximal expression of *atTOC34* (Kubis et al., 2003). In each case, the more highly expressed homologue (*atTOC33* or *atTOC159*) has been implicated in the import of photosynthetic precursor proteins. However, our results suggest that this is not the case with the Toc75 family, since embryos containing only proplastids require the atToc75-III protein for development. Functional redundancy also exists between the other Toc component isoforms; *e.g.*, in the absense of atToc33, atToc34 is able to compensate to a certain extent. However, since *toc75-III* embryos abort at just the two cell stage, it appears that atToc75-IV is not able to replace the function of atToc75-III to any significant degree.

It is interesting that the *toc75-III* mutant is embryo lethal, especially given that other Toc mutants isolated to date are not (Jarvis et al., 1998; Bauer et al., 2000; Constan et al., 2004). This observation most likely refects the importance of the role of the atToc75-III protein as the channel the plastid outer envelope membrane, and the inabitility of atToc75-IV to substitute for atToc75-III during embryo biogenesis. The other Toc mutants (Jarvis et al., 1998; Bauer et al., 2000; Constan et al., 2004) are perhaps able to survive due to functional redundancy between the components. Additionally, the fact that the Toc75 protein is able to recognize a transit peptide without the aid of other Toc components (Hinnah et al., 2002) - a function originating from the ancestral Toc75 protein which did not form a complex with Toc159 and Toc34 proteins - perhaps explains why the receptor proteins are not essential for viability, since they can be partly substituted by the channel itself. Infact, since free Toc75 exists in the membranes, which is not associated with Toc complexes, Toc75 may import precursor proteins without the aid of receptor proteins (Schleiff et al., 2003b).

Stoichiometric analysis of the Toc complex reported that each ~500 kDa complex comprises one Toc159 subunit, four or five Toc34 subunits, and four Toc75 subunits (Schleiff et al., 2003b). Since we have demonstrated that *Arabidopsis* encodes two expressed Toc75-related genes, it would be interesting to investigate whether each Toc complex incorporates both or just one type of atToc75. While observing the Toc complex by electron microscopy, Schleiff et al. (2003) sorted the complexes that they observed into eight classes. Class one, which represents 4.5% of the total number of complexes, contained smaller particles that appeared to be significantly different from the other classes (Schleiff et al., 2003b). They concluded that these particles represented either the Toc complex in a different conformation, or a minor experimental artifact, such as partially degraded complex. An alternative explanation is that these particles contain a protein very much like atToc75-IV. Since *atTOC75-IV* is expressed at a lower level than *atTOC75-III,* fewer of the particles incorporating such a putative 'psToc75-IV' would be expected to be observed.

The atToc75-IV protein appears to play a role in etioplasts, since *toc75-IV* etioplasts have significantly more cytosolic inclusions than wild-type etioplasts. Large cytoplasmic inclusions have previously been observed in root plastids of double homozygote *toc132 toc120* seedlings (Kubis et al). atToc132 and atToc120 are atToc159 homologues. The cytosolic inclusions that we observed were not as large, occupying ~20% of the cross sectional surface of the plastid, rather than the *toc132 toc120* root plastids which contained inclusions that occupied up to 50% of the organelles cross-sectional area. It is not clear why these inclusions develop but it maybe a direct result of the plastids inability to import specific structural proteins. On the other hand, the observed increase in surface area to volume ratio may be a method of increasing uptake from the cytosol into the organelle.

We have demonstrated that the *Arabidopsis* genome encodes three Toc75-related sequences; *atTOC75-I, atTOC75-III* and *atTOC75-IV*, with only two of the three sequences being expressed. Fractionation experiments demonstrated that atToc75-IV, like atToc75-III, is an integral chloroplast envelope membrane protein. However, it is not clear whether, atToc75-IV is a channel protein that imports proteins into plastids. Although atToc75-III, appears to have a similar function to that of pea Toc75, the role of atToc75-IV is not yet clear.

Table 2.1 Segregation of the T-DNA-associated antibiotic resistance marker in each Toc75 knockout mutant

Mutant	Antibiotic	Antiobiotic resistant (R)	Antibiotic sensitive (S)	R:S ratio	χ^2
toc75-III-1	Hygromycin	171	91	1.9	0.
toc75-III-2	Kanamycin[a]	35	21	1.7	0.
toc75-IV-1	Phosphinothricin	83	26	3.2	0.
toc75-IV-2	Phosphinothricin	41	14	2.9	0.

[a]The *toc75-III-2* T-DNA insertion was followed by PCR since the associated kanamycin resistance marker wa

[b]Goodness of fit of the observed ratios to 2:1 (*toc75-III*) and 3:1 (*toc75-IV*) was assessed by $\chi 2$ analysis. *P*-v probabilities that the observed ratios differ from 2:1 or 3:1 due to random chance only.

Table 2.2 The frequency of aborted seed in *toc75-III* siliques

Genotype	Normal seed	Infertile ovules	Aborted seed
Wild type	196 (98%)	2 (1%)	3 (1%)
toc75-III-1	434 (76%)	10 (2%)	125 (22%)
toc75-III-2	650 (76%)	11 (1%)	195 (23%)

Maturing siliques from at least four different plants of each genotype were scored for the presence of normal seed, infertile ovules and aborted seed.

Table 2.3 Distribution of embryo phenotypes within single siliques of the *toc75-III-1* mutant

Silique	One-cell	Two-cell[a]	Aborted seed[b]	Four-cell	Eight-cell	16-cell	32-cell	Globula
1	28							
2	12	9		13				
3	7	4		10	4			
4	7	3		14	10			
5	3	6			9	9		
6	2	5	2		2	9	11	5
7		5	3		1	9	16	2
8		4	3			6	12	14

Embryo stages are referred to by the number of cells in the embryo proper, or by the shape of the embryo proper.

[a]Two-cell stage also includes embryos with two free nuclei in the apical cell.

[b]Aborted seed were shrivelled with no visible embryo or endosperm development (see Figure 6).

Table 2.4 The frequency of cytosolic inclusions in the etioplasts of *toc75-IV* and 35S-*atTOC75-IV*

Genotype	Total no. etioplasts	No. with inclusions	Frequency (%)	χ^2
Wild type	54	12	22.2	
toc75-IV-1	62	24	38.7	8.
toc75-IV-2	61	24	39.3	9.
Wild type	100	15	15.0	
35S-atTOC75-IV, line 10	131	4	3.1	1
35S-atTOC75-IV, line 13	123	8	6.5	6.

[a]The data were analysed using a chi-squared test with the Yates' Correction for Continuity. We set c
3.84, *P*-value = 0.05, and in every case the mutant/transgenic data were found to differ significantly
data. A *P*-value of 0.01 indicates that there is 1% probability that the observed differences between
and wild-type data occured by chance alone.

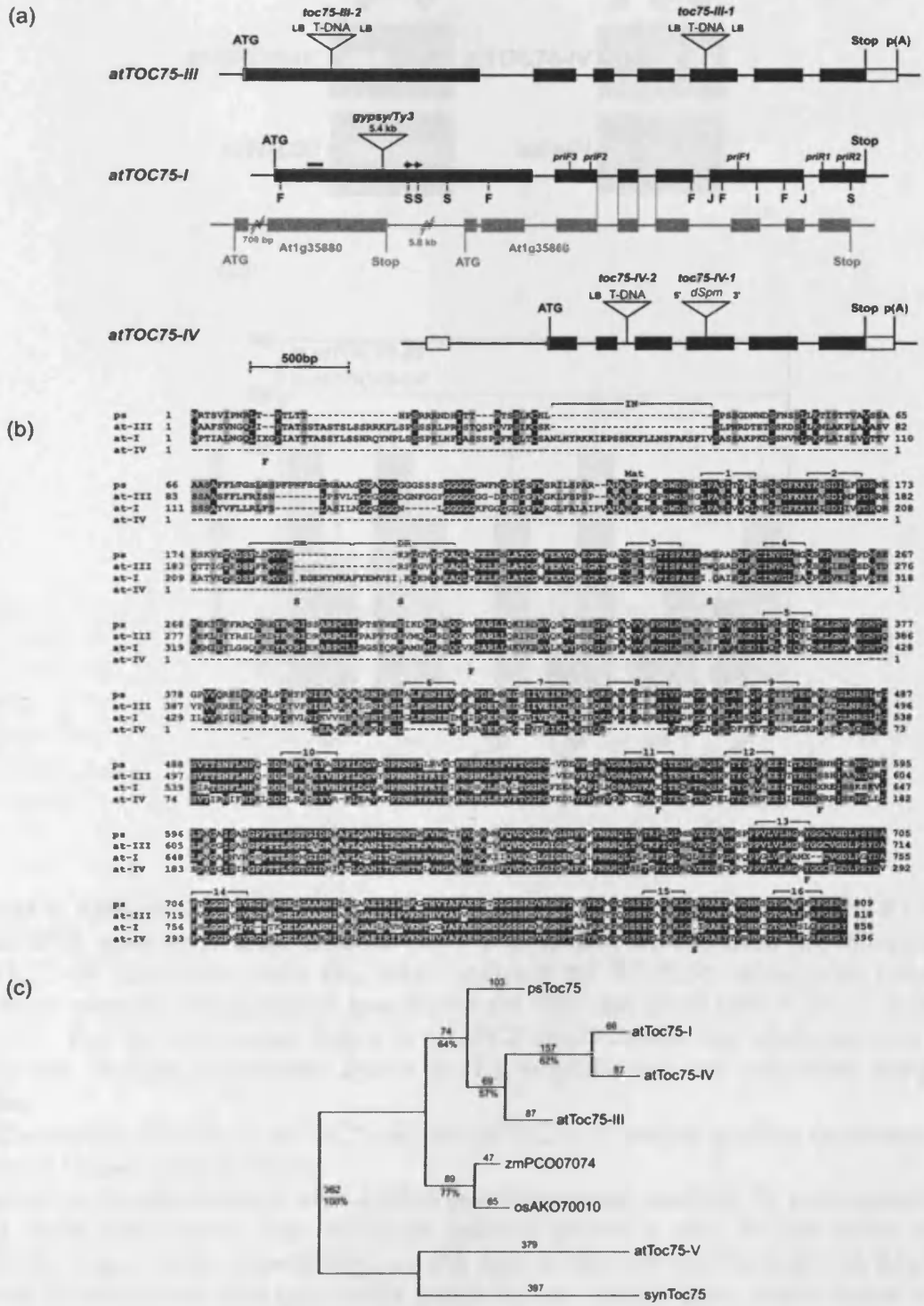
Figure 1 (overleaf). Structural characteristics of the *Arabidopsis* Toc75 gene family.

(a) Schematic diagrams depicting the three *Arabidopsis* Toc75-related sequences. Protein-coding exons are represented by black boxes, and untranslated regions are represented by white boxes; introns are represented by thin lines between the boxes. Mutations in the *atTOC75-I* pseudogene are indicated as follows: F, frame-shift; S, nonsense; J, splice junction; I, missing intron. A 75 bp insertion and 48 bp direct repeat, both in exon 1, are represented by a black bar and two black arrows, respectively. Locations of forward (*priF1*, *priF2* and *priF3*) and reverse (*priR1* and *priR2*) primers used for RT-PCR analysis of *atTOC75-I* are indicated. The grey diagram underneath *atTOC75-I* illustrates how the pseudogene is currently annotated as two separate genes (At1g35880 and At1g35860) in the GenBank database. The locations of T-DNA and transposon insertions are indicated precisely, but the insertion sizes are not to-scale. ATG, translation initiation codon; Stop, translation termination codon; p(A), polyadenylation site; LB, T-DNA left border; RB, T-DNA right border.

(b) Amino acid sequence alignment of pea Toc75 (ps) with the three *Arabidopsis* proteins (at-III, at-I and at-IV). The atToc75-I sequence was derived by conceptual translation of a partially corrected version of the open reading frame illustrated in (a). The locations of frame-shift (F) and nonsense (S) mutations in the atToc75-I sequence, a 25 residue insertion (IN), and a 16 residue direct repeat (DR), are indicated. Residues identical in at least three sequences are highlighted in black, whereas similar residues are highlighted in grey. The location of the first Asp residue of the mature psToc75 protein is shown (Mat), as are the 16 predicted transmembrane domains of psToc75 (Tranel et al., 1995; Sveshnikova et al., 2000b).

(c) Phylogenetic analysis of Toc75-related proteins from *Arabidopsis* and other species. Full length pea, *Arabidopsis* and cyanobacterial amino acid sequences were aligned, together with sequences from maize and rice, and used to produce a phylogenetic tree. Numbers of mutations are given above the clades, with bootstrap values below. Gene and accession numbers for the sequences used are as follows: psToc75 (L36858, S55344); atToc75-III (At3g46740, NP_190258); atToc75-I; atToc75-IV (AY585655, AAT08975); atToc75-V (At5g19620, NP_568378); zmToc75 (AY106148); osToc75 (AK070010); synToc75 (slr1227, NP_440832). Species of origin is indicated as follows: *Pisum sativum*, ps; *Arabidopsis thaliana*, at; *Zea mays*, zm; *Oryza sativa*, os; *Synechocystis* PCC 6803, syn. The cyanobacterial protein, synToc75, was used as the outgroup.

Figure 1, (from previous page)



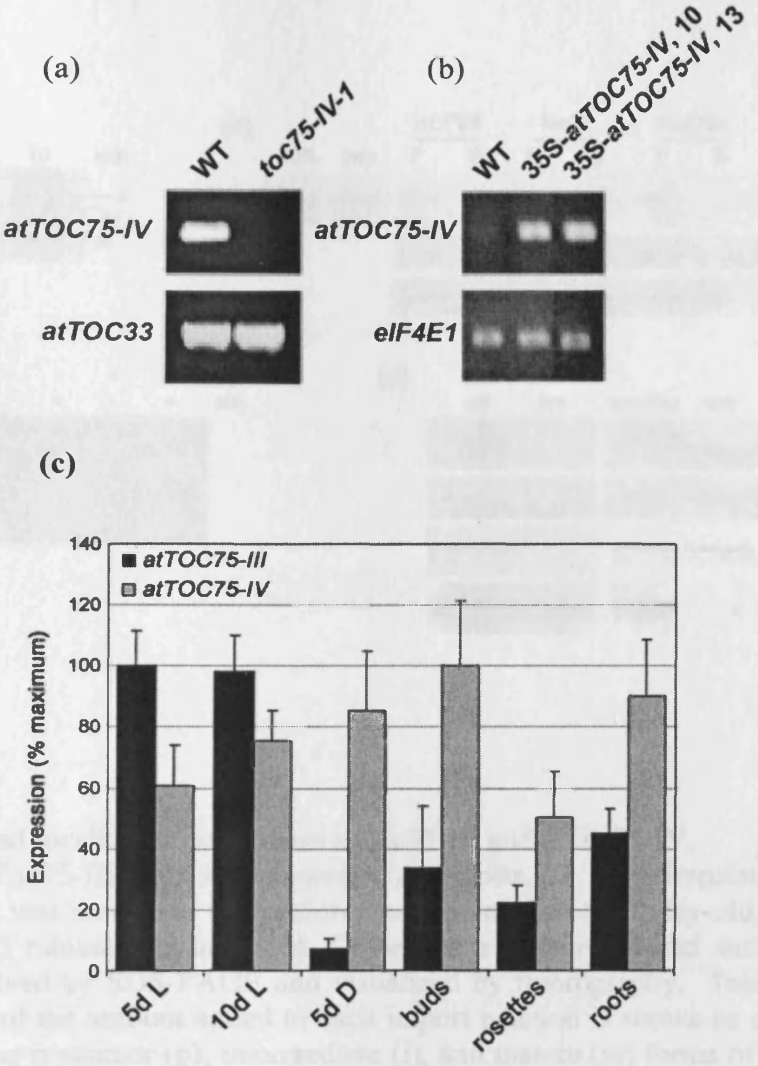


Figure 2. Analysis of *atTOC75-IV* expression in mutant and transgenic lines by RT-PCR.
Total RNA samples isolated from the *toc75-IV-1* mutant (a), and from two independent 35S-*atTOC75-IV* transgenic lines (b), were analysed by RT-PCR, along with corresponding wildtype samples, using primers specific for the indicated genes (*atTOC75-IV*, *atTOC33* and *eIF4E1*). For the experiment shown in (a), PCR amplification was conducted over a total of 40 cycles. For the experiment shown in (b), amplification was conducted using only 25 cycles.
(c) Expression profiles of *atTOC75-III* and *atTOC75-IV* during seedling development and in different tissues of *Arabidopsis*.
Total RNA samples isolated from *Arabidopsis* tissues were analysed by semi-quantitative RT-PCR. RNA was isolated from wild-type seedlings grown *in vitro* for five or ten days in the light (5d L and 10d L, respectively), or five days in the dark (5d D), and from three different tissues of 28-day-old wild-type plants grown on soil (flower buds, rosette leaves and roots). Amplifications were conducted under non-saturating conditions using gene-specific *atTOC75-III*, *atTOC75-IV* and *eIF4E1* primers, and the products were quantified by hybridization with corresponding ^{32}P-labelled cDNA probes. The data for *atTOC75-III* and *atTOC75-IV* data were normalized for *eIF4E1*, and then expressed as a percentage of the maximum level observed for each gene. Values shown are means (±SD) of four independent measurements.

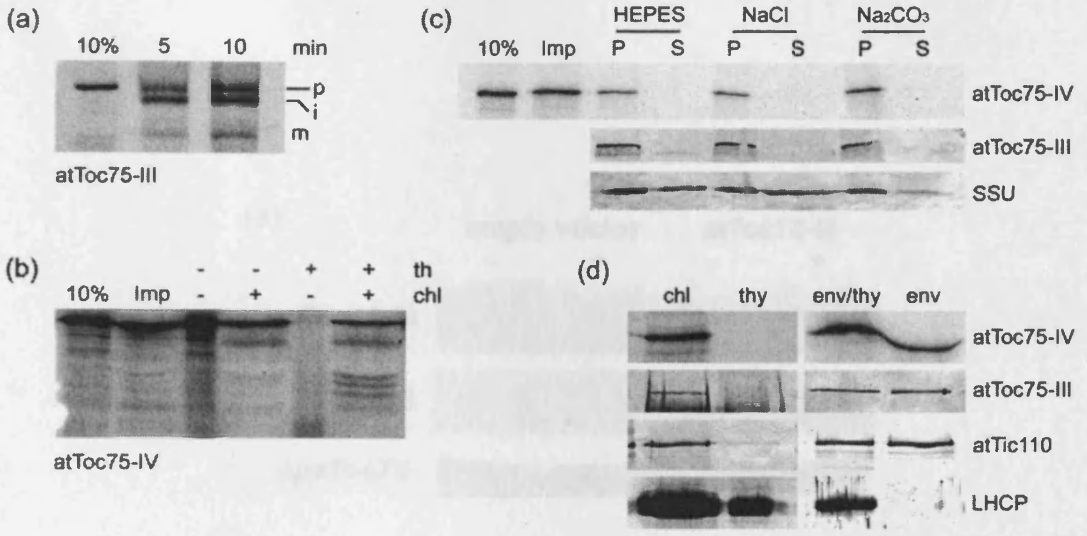


Figure 3. Import and localization studies on atToc75-III and atToc75-IV.
(a) Import of atToc75-III into *Arabidopsis* chloroplasts. *In vitro* translated, ^{35}S-labelled atToc75-III protein was incubated with chloroplasts isolated from 10-day-old, soil-grown pea plants for 5 and 15 minutes, as indicated. Chloroplasts were recovered and the constituent proteins were resolved by SDS-PAGE and visualized by fluorography. Translation product equivalent to 10% of the amount added to each import reaction is shown as a control. Bands corresponding to the precursor (p), intermediate (i), and mature (m) forms of atToc75-III are indicated.
(b) Import of *in vitro* translated, ^{35}S-labelled atToc75-IV into chloroplasts isolated from 10-day-old *Arabidopsis* plants grown *in vitro*. The first lane contains translation product equivalent to 10% of the amount added to each import reaction, and the second lane contains a standard, 20 minute import reaction (Imp). The remaining four lanes show an associated thermolysin treatment experiment: ^{35}S-labelled atToc75-IV was incubated for 20 minutes under import conditions in the presence/absence of chloroplasts, as indicated, before further treatment in the presence/absence of 60 μg/ml thermolysin, as indicated. In each case, chloroplasts were recovered and the constituent proteins were resolved by SDS-PAGE and visualized by fluorography.
(c) High-salt and high-pH treatment of imported atToc75-IV protein. *In vitro* translated,^{35}S-labelled atToc75-IV protein was imported into isolated chloroplasts as described in (b). Recovered chloroplasts were treated for 1 hour with 10 mM HEPES (bursting buffer), 1 M NaCl (high-salt) or 0.1 M Na_2CO_3 (high-pH), before separation into pellet (P) and soluble (S) fractions by centrifugation. Following resolution by SDS-PAGE, atToc75-IV was visualized by fluorography, and atToc75-III and SSU were detected by immunoblotting.
(d) Suborganellar localization studies. *In vitro* translated,^{35}S-labelled atToc75-IV protein was imported into isolated chloroplasts as described in (b). Recovered chloroplasts (chl) were then fractionated by centrifugation to generate thylakoid (thy), total membrane (env/thy) and envelope (env) fractions. Following resolution by SDS-PAGE, atToc75-IV was visualized by fluorography, and atToc75-III, atTic110 and LHCP were detected by immunoblotting. Each lane contains protein equivalent to one import reaction, or 10 million chloroplasts.

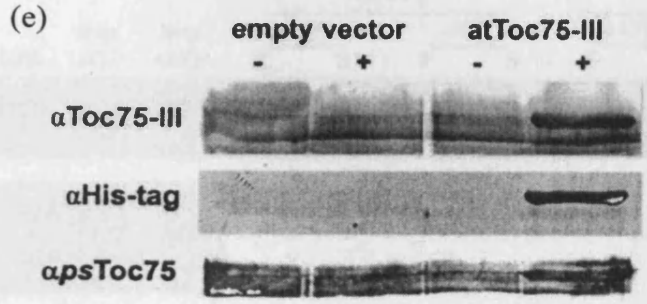


Figure 3. Import and localization studies on atToc75-III and atToc75-IV.
(e) N-terminally His-tagged atToc75-III protein was heterologously expressed using a Lac-inducible promoter. Anti-sera raised against pea Toc75, a short atToc75-III peptide sequence and 6x His, specifically detected induced (+) His-atToc75-III protein (89 kDa) by immunoblotting. In accord, no 89 kDa band was detectible from uninduced (-) cultures or bacteria containing only empty vector, whether induced (+) or not (-).

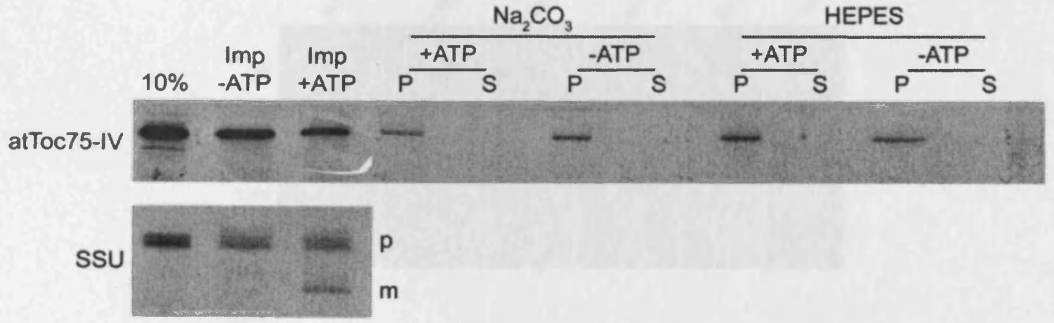


Figure 4. Energetics of atToc75-IV membrane insertion
Chloroplasts isolated from 10-day-old *Arabidopsis* plants grown *in vitro* were incubated in the dark in the presence of 6 μM nigericin for 10 minutes to deplete endogenous ATP. *In vitro* translated, ^{35}S-labelled atToc75-IV and preSSU proteins were passed through Sephadex G-25 to remove small molecules, and then incubated with the chloroplasts under import conditions for 20 minutes in the absence (Imp -ATP) or presence (Imp +ATP) of 5 mM ATP. Chloroplasts recovered from separate import assays were treated for 1 hour with 0.1 M Na_2CO_3 (high-pH) or 10 mM HEPES (bursting buffer), before separation into pellet (P) and soluble (S) fractions by centrifugation. Following resolution by SDS-PAGE, atToc75-IV and preSSU/SSU were visualized by fluorography. Bands corresponding to preSSU (p) and mature SSU (m) are indicated.

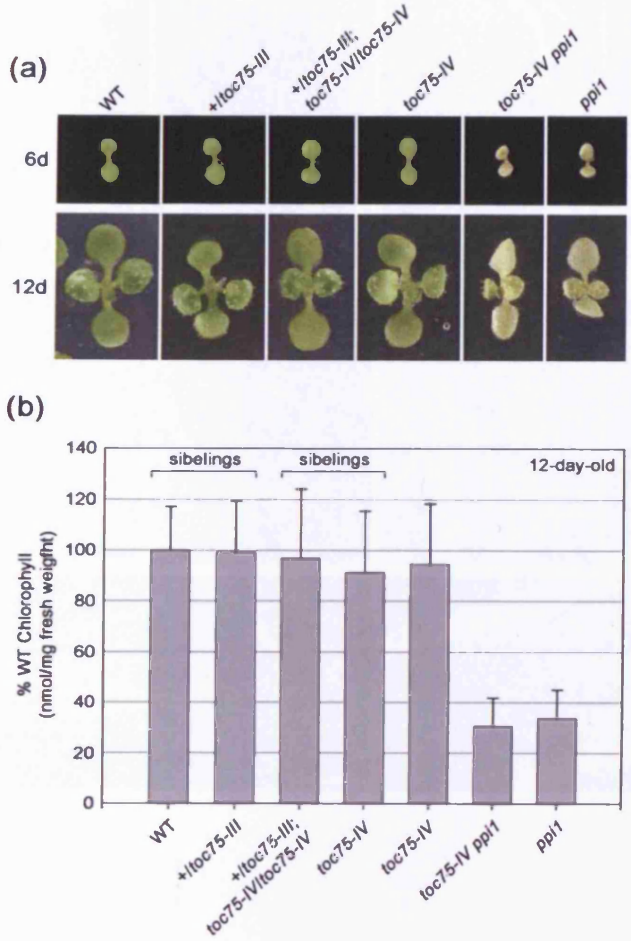


Figure 5. Phenotypic analysis of the *toc75-III* and *toc75-IV* mutants.
(a) Plants grown under long-day conditions *in vitro* for 6 or 12 days post-germination are shown. The presented genotypes are, from left to right: wild type (WT), *toc75-III* heterozygote (+/*toc75-III*), *toc75-III toc75-IV* double mutant (+/*toc75-III*; *toc75-IV*/*toc75-IV*), *toc75-IV* homozygote, *toc75-IV ppi1* double homozygote, and *ppi1* homozygote.
(b) The visible phenotypes of 12-day-old plants of the indicated genotypes were quantified by making chlorophyll measurements. Sibling analysis was conducted on families segregating for the *toc75-III* mutation, which were either wild-type (bars 1 and 2) or homozygous mutant (bars 3 and 4) at the *atTOC75-IV* locus. For bars 1-4, measurements were of samples containing two primary leaves of the same plant; the remainder of each plant was used for genotyping by PCR. For bars 5-7 (*toc75-IV*, *toc75-IV ppi1* and *ppi1*), measurements were of samples containing eight primary leaves from different plants. Values shown are means (±SD) derived from multiple independent measurements: WT (n=11), +/*toc75-III* (n=19), +/*toc75-III*; *toc75-IV*/*toc75-IV* (n=45), *toc75-IV* (n=15), *toc75-IV* (n=20), *toc75-IV ppi1* (n=6), *ppi1* (n=6). The data are expressed as a percentage of the wild-type chlorophyll concentration.

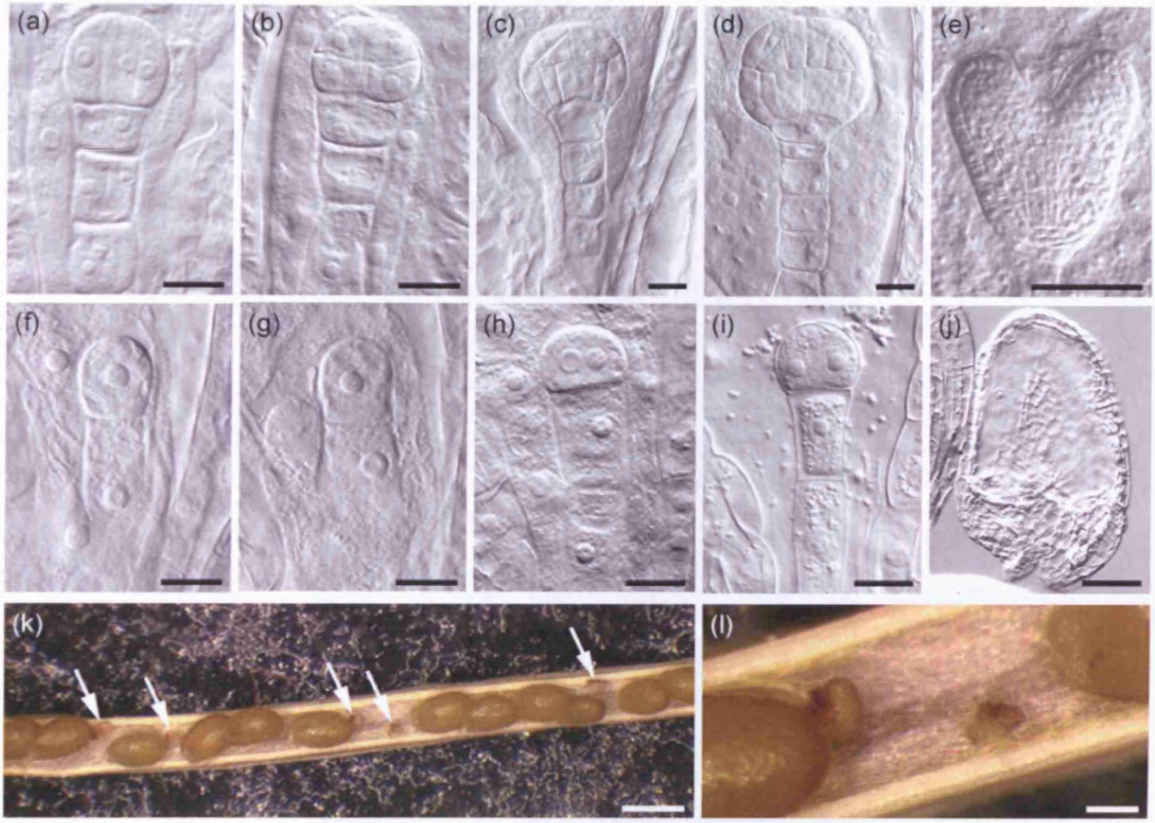


Figure 6. Embryo lethality of the *toc75-III* mutation.
(a-j) Morphology of wild-type (a-e) and mutant (f-j) embryos from the siliques of *toc75-III* heterozygous plants containing normal, heart-stage embryos. The wild-type developmental stage in each pair of images is as follows: (a, f) four-cell embryo stage; (b, g) eight-cell embryo stage; (c, h) 16-cell embryo stage; (d, i) 32-cell embryo stage. Embryo cell stages refer to the number of cells in the embryo proper. A normal, heart-stage embryo (e) and an aborted seed (j) are also shown. Bars = 10μm (a-d, f-i) and 50 μm (e, j).
(k) The appearance of part of a typical silique from a mature *toc75-III* heterozygous plant. Arrows indicate aborted seed. Bar = 0.5 mm.
(l) Higher magnification image of the two central aborted seed shown in (k). Bar = 0.1 mm.

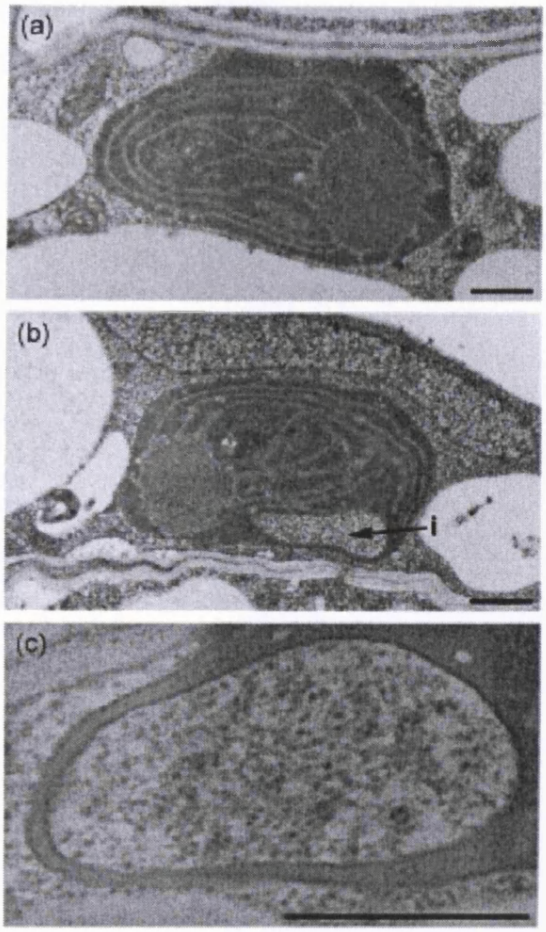


Figure 7. Ultrastructure of etioplasts in the *toc75-IV* mutant.
Typical etioplasts from wild-type (a) and *toc75-IV-2* mutant (b) plants grown *in vitro* in the dark for three days; the *toc75-IV* plastid contains a large inclusion (i) of cytosolic material. In panel (c), an enlargement of a typical cytosolic inclusion within a *toc75-IV-1* mutant etioplasts is shown (c). Bars = 0.5 μm.

Results II

Characterisation of Extragenic Suppressors of a Plastid Translocon Mutant

1. Abstract

Yellow-green *plastid protein import 1* (*ppi1*) plants are null for the atToc33 protein, a component of the plastid protein import apparatus. Homozygous *ppi1* mutant seed were treated with the chemical mutagen, ethyl methsulfonate, and screened visually for partial or full restoration of the wild-type (green) phenotype. Five mutant lines were isolated which contain recessive, heritable, single-locus mutations that suppress the *ppi1* phenotype. These five mutations were found to occupy two different loci, namely *sp1* and *sp2*. All five mutations suppress the *ppi1* phenotype in every respect tested, including having significantly more chlorophyll than *ppi1*, and larger chloroplasts with more thylakoid membranes. Recent data indicate that atToc33 controls the import of photosynthetic precursor proteins preferentially, since the *ppi1* mutant was found to be specifically defective in the expression, chloroplast import, and accumulation of photosynthetic proteins. We show that mutant *sp1 ppi1* seedlings contain more photosynthetic proteins than *ppi1*, and that chloroplasts isolated from *sp1 ppi1* seedlings are able to import the photosynthetic precursor, preOE33, more efficiently than chloroplasts isolated from *ppi1* seedlings.

Division of labour. The authors for publication will be as follows: Amy Baldwin and Paul Jarvis. Samples for light and electron microscopy were prepared by Natalie Allcock of the University of Leicester Electron Microscopy Laboratory. All other experiments were carried out by Amy Baldwin.

2. Introduction

The majority of plastid proteins are encoded in the nucleus, translated on cytosolic ribosomes, and imported post-translationally into the plastid (Keegstra and Cline, 1999; Chen et al., 2000b; Hiltbrunner et al., 2001a; Jarvis and Soll, 2002). Most nucleus encoded plastid proteins have a cleavable, amino-terminal transit peptide, which is necessary and sufficient for targeting the precursor to the plastid interior. Prior to reaching the plastid, the precursor proteins are maintained in an import-competent state by chaperone proteins (Waegemann et al., 1990; May and Soll, 2000a; Rial et al., 2000). Precursor proteins are recognised and then translocated across the double envelope membrane by two protein complexes: the translocon at the outer envelope membrane of chloroplasts (Toc), and the translocon at the inner envelope membrane of chloroplasts (Tic). GTP in the cytosol or intermembrane space is required for precursor proteins to bind to the translocon complex, and stromal ATP is required for precursor translocation into the stroma (Olsen et al., 1989; Theg et al., 1989; Young et al., 1999; Schleiff et al., 2003a). Upon translocation into the stroma, the transit peptide is cleaved and the protein is either folded into its final conformation or targeted to one of several internal compartments of the plastid (Keegstra and Cline, 1999).

Toc and Tic translocon components were originally identified in pea chloroplasts using biochemical techniques (Hirsch et al., 1994; Perry and Keegstra, 1994; Schnell et al., 1994; Wu et al., 1994; Seedorf et al., 1995). Precursor proteins were incubated with isolated chloroplasts under energy-limiting conditions, and proteins in close proximity to the precursor were identified either by chemical cross-linking or by immunoprecipitation. The outer envelope membrane translocon components identified were 75 kDa, 86 kDa and 34 kDa proteins, and were named according to their molecular masses. The 75 kDa protein, Toc75, has since been shown to form the protein-conducting channel through the plastid outer envelope membrane (Hinnah et al., 2002). The 86 kDa protein is a proteolytic fragment of a larger protein, Toc159 (Bolter et al., 1998a; Chen et al., 2000a). Toc159 and the 34 kDa protein, Toc34, are integral outer envelope membrane proteins that control preprotein recognition in a GTP-dependent manner (Hirsch et al., 1994; Kouranov and Schnell, 1997; Jelic et al., 2002; Sun et al., 2002). It has been proposed that Toc159 interacts with precursor proteins in the cytosol and shuttles them to the import apparatus (Hiltbrunner et al., 2001b; Bauer et al., 2002; Smith et al., 2002). It also appears that Toc159 requires Toc34 in order to localize to the plastid outer envelope membrane (Weibel et al., 2003). Toc34 has also been proposed to participate directly in preprotein recognition (Sveshnikova et al., 2000a; Schleiff et al., 2002).

Arabidopsis is now firmly established as the organism of choice in many laboratories for studying basic biological processes in plants. This is not only due to its well defined genetics, ease of cultivation, and efficiency of transformation procedures, but also to the availability of its complete genome sequence (The-Arabidopsis-Genome-Initiative, 2000) and international facilities that provide insertion mutants and various genomic resources (Somerville and Koornneef, 2002). Multiple homologues of many of the Toc and Tic components have been identified in the *Arabidopsis* genome sequence, and are being characterised by both biochemical and molecular-genetic techniques (Jarvis et al., 1998; Bauer et al., 2000; Jackson-Constan and Keegstra, 2001; Constan et al., 2004). The first Toc mutant to be identified, *plastid protein import 1* (*ppi1*), carries a T-DNA insertion which inactivates the *atTOC33* gene, one of the two homologous genes encoding Toc34-related proteins in *Arabidopsis* (Jarvis et al., 1998). Mutant *ppi1* seedlings have a visible, yellow-green phenotype that is associated with the presence of smaller chloroplasts that contain fewer thylakoid membranes and smaller granal stacks (Jarvis et al., 1998). *In vivo* data indicate that atToc33 controls the import of photosynthetic precursor proteins preferentially, since the *ppi1* mutant is specifically defective in the expression, chloroplast, and accumulation of photosynthetic proteins (Kubis et al., 2003).

There are two techniques for linking the sequence and the function of a particular gene: forward genetics and reverse genetics. With the exception of *ppi1*, all of the translocon component mutants described to date have been identified using reverse genetics; *i.e.*, the translocon genes had previously been identified in pea and their functions were subsequently investigated genetically in *Arabidopsis* using specifically identified mutants (Bauer et al., 2000; Gutensohn et al., 2000; Chen et al., 2002; Chou et al., 2003; Constan et al., 2004). Reverse genetics is especially useful for assigning *in vivo* functions to genes or proteins identified using, for example, *in vitro* or biochemical approaches. However, one of the major disadvantages of the reverse genetics approach is that it is not possible to identify novel components. Forward genetics, being the opposite of reverse genetics, enables you to do this, for example, the yellow-green *ppi1* mutant was selected in a mutant screen for plants displaying defective chloroplast biogenesis, and subsequently the gene responsible, *atTOC33*, was identified (Jarvis et al., 1998). Recently, another mutant, *chloroplast import apparatus 2* (*cia2*), was identified in a transgene-based forward genetic screen for mutants with defects in protein import into plastids (Sun et al., 2001). The *CIA2* gene encodes a nuclear-localized protein, which contains a motif conserved within the CCT family of transcription factors, and specifically controls the expression of the translocon genes, *atTOC75* and *atTOC33*. In order

to identify additional novel loci involved in the import of proteins into plastids, we conducted a screen for second-site suppressors of the *ppi1* mutation.

3. Results

3.1. Mutagenesis and Genetics

In order to generate suppressors of the *ppi1* mutation, we treated 7,000 homozygous *ppi1* seed with the chemical mutagen, ethyl methanesulfonate (EMS). The first mutant generation (M_1) seed was grown in 20 separate batches. The resulting second mutant generation (M_2) was visually screened at 7-days and 20-days after germination for seedlings that appeared greener than the yellow-green control *ppi1* seedlings. The M_3 populations contained many plants with mutant phenotypes; for example, there were many albino and dwarf seedlings, indicating that the mutagenesis had been successful. Twenty-one putative suppressor mutants were identified and immediately backcrossed to *ppi1* (Col-0 ecotype) to remove background mutations. Individual seedlings displaying the suppressor phenotypes were selected in the second filial generation (F_2) for further analysis. All 21 suppressor mutations were heritable since the phenotypes persisted in subsequent generations. Five mutants were selected on the basis of their phenotypes, homozygosity for the *ppi1* mutation (as revealed by PCR), and the absence of undesirable secondary mutations. Due to the inherent, slight variability of the *ppi1* phenotype, only the most strongly suppressing lines were selected for further analysis. The mutants that were selected were initially referred to as 2-3-10, 2-4-3, 4-1-16, 5-1 and 14-13 (Figure 1a), and were derived from four different M_2 populations (batches 2, 4, 5 and 14). Two mutants (2-4-3 and 2-3-10) were taken from the same M_2 population (batch 2). Because these two mutants displayed heritably different phenotypes at 7-days-old, it seemed likely that they were different mutations; the 2-4-3 mutant displays a much greener phenotype than the 2-3-10 mutant early in development, but this difference becomes less obvious as the primary leaves emerge (data not shown).

With a view to mapping the suppressor mutations genetically, the *ppi1* mutation was introgressed into the Landsberg *erecta* genetic background (*ppi1* was originally isolated in the Columbia-0 background) (Jarvis et al., 1998). Seven outcrosses into Landsberg *erecta* were done to achieve a theoretical 99% ecotype purity. Homozygous *ppi1* plants in the Landsberg *erecta* background have the same yellow-green phenotype as *ppi1* (Col-0 ecotype) plants, but retain the characteristic rounder leaves and short inflorescences of Landsberg *erecta*. The five selected suppressor mutants were outcrossed to *ppi1* (L*er* ecotype) to assess the inheritance of the *sp* mutations, and to create F_2 populations for genetic mapping. All seedlings in the five resultant F_1 generations were yellow-green, like *ppi1*, and the F_2 generations contained seedlings with both yellow-green and suppressor phenotypes (Table 3.1). The ratio of *ppi1*

like seedlings to suppressor seedlings in all five F_2 populations was not significantly different from 3:1, indicating that each mutant contains a single, recessive mutation.

To determine whether the five *sp* mutants are allelic, pairwise crosses were conducted between the five homozygous *sp ppi1* mutants. The resultant F_1 generation seedlings were scored for yellow-green and suppressor phenotypes. If the *sp* mutations are non-allelic (*i.e.*, the mutations are in different genes), the resulting F_1 generations would be expected to contain only yellow-green plants, since each recessive *sp* mutation would be in the heterozygous state. On the other hand, if the *sp* mutations are allelic, the F_1 generations would be expected to display the suppressor phenotype, since the seedlings would have no wild-type copies of the affected gene. A minimum of 25 seedlings was scored for each of the 20 possible combinations. Since only two of the five lines originated from the same M_1 pool, it was expected that the lines would not all be allelic. Surprisingly, four of the five *sp* mutants (2-3-10, 2-4-3, 5-1 and 14-13) were found to be allelic, since all the F_1 seedlings from crosses between these four lines were found to display the suppressor phenotype. The remaining mutant, 4-1-16, was not allelic with any of the other four mutants since all of the F_1 generations derived from crosses with this mutant exhibited a normal *ppi1* phenotype.

In summary, five mutant lines were isolated which contain recessive, heritable, single-locus mutations that suppress the yellow-green, *ppi1* phenotype. Unexpectedly, four of the mutations were found to be allelic. It is very unlikely that these four mutants all carry the same lesion, however, since they were identified in three different M_2 pools. Hereafter, we refer to the allelic mutants as *spl-1* (2-4-3), *spl-2* (14-13), *spl-3* (5-1) and *spl-4* (2-3-10), and the non-allelic mutant as *sp2* (4-1-16). Heritable differences in phenotype severity between *spl-1* (2-4-3) and *spl-4* (2-3-10) described earlier strongly suggest that these two mutants also carry independent lesions.

To determine the approximate map positions of the five *sp* mutations, seedlings displaying the suppressor phenotype from the appropriate F_2 populations were analysed using standard procedures. Using PCR-based markers placed at approximately 50 cM intervals along the five *Arabidopsis* chromosomes, recombination frequencies were calculated by analysing approximately 60 chromosomes in each case. Each marker had previously been confirmed to be homozygous Landsberg *erecta* ecotype in L*er*-introgressed *ppi1* line. For all four *spl* alleles, recombination frequencies were 50 ± 10% on chromosomes 2, 3, 4 and 5, indicating that the *spl* locus and those particular markers are unlinked. This was also the case at the top of chromosome 1. However, the recombination frequency dropped substantially when a

marker at position 121.7 cM on chromosome 1 (GL2) was analysed. For *spl-1*, the recombination frequency with GL2 was determined to be 0.0%, indicating that the marker (GL2) and the *spl-1* mutation are linked. The F_2 mapping populations for the other three allelic mutants, *spl-2, spl-3* and *spl-4,* also had recombination frequencies of 9.4%, 0.0%, 25.0% with the GL2 marker, respectively, re-enforcing the evidence that they are allelic. The non-allelic *sp2* mutant was also genetically mapped and shown to be on chromosome 2 or 3. Since the *sp2* mutation is not on chromosome 1, this re-enforces the evidence that *sp2* is not allelic with the other four mutations.

3.2. Visible Phenotypes of the Suppressor Mutants

To quantify the extent to which the *sp* mutations suppress the *ppi1* phenotype, the phenotypes of the five suppressor mutants were analysed. Initially, anthocyanin accumulation was measured to rule out the possibility that the change in phenotype was due to increased levels of this purple pigment. However, anthocyanin levels were not found to contribute to the suppressor phenotype in any of the mutants, since all five double homozygous *sp ppi1* mutants have similar levels of anthocyanin to *ppi1* (data not shown). Chlorophyll measurements were made on homozygous 13-day-old mutant seedlings to quantify the degree of recovery with respect to the *ppi1* single mutant (Figure 1b). After 13 days growth, all five mutants contained significantly more chlorophyll per mg tissue than *ppi1* seedlings. However, there was no significant difference in chlorophyll concentration between the four *spl* alleles. The average recovery of the *spl* mutants was 39 ± 6%, if a value of 0% is arbitarily assigned to ppi1 and a value of 100% is assigned to wild type. The *sp2 ppi1* mutant suppressed the *ppi1* phenotype to a lesser extent, since it recovered by only 23 ± 2%. Nevertheless, the concentration of chlorophyll in *sp2 ppi1* was significantly greater than that in *ppi1* single mutants, as revealed by a two-sample t-test carried out using Minitab. The null hypothesis that both means were the same was tested. Statistical analysis calculated that there was only a 0.08% probability that the null hypothesis is true; *i.e.*, there is a 0.08% probability that the means are actually the same and the observed differences occurred by chance. Since 0.08% is less than the significance level of 5%, we rejected the null hypothesis and conclude that *sp2 ppi1* seedlings contain significantly more chlorophyll than *ppi1* seedlings. The fact that *sp2* suppresses *ppi1* to a lesser degree than the non-allelic *spl* mutations re-enforces the fact that they affect different loci, perhaps suppressing *ppi1* via different mechanisms.

3.3. Leaf Anatomy of the sp Mutants

To determine if the *sp* mutations suppress other aspects of the *ppi1* phenotype, in addition to chlorophyll accumulation, the anatomies of cotyledons from 10-day-old seedlings were examined by light microscopy (Figure 2). The larger size of the *sp ppi1* mutant cross-sections compared to *ppi1* reflects the suppression of the *ppi1* phenotype by the *sp* mutations. The cotyledons of wild-type seedlings typically had a single, almost contiguous layer of columnar, palisade mesophyll cells, arranged above several layers of spongy mesophyll cells interspersed with many intercellular air spaces (Figure 2a). In contrast, the palisade mesophyll cells of *ppi1* were barely distinguishable from the spongy mesophyll cells, and the intercellular spaces extended throughout the parenchyma tissue (Figure 2g). The mesophyll cells of the *ppi1* mutant appeared smaller than wild-type mesophyll cells, but the size of the epidermal cells did not appear to be changed. Due to their role in photosynthesis, there are typically three to five times more chloroplasts in palisade mesophyll cells than in other cell types and so defective chloroplast biogenesis may have a more pronounced effect in the palisade mesophyll than in other tissues. The epidermis of the *ppi1* mutant is no longer a smooth layer enclosing the parenchyma tissue; instead, the epidermis undulates, perhaps caused by the irregular underlying cell layers. Similar phenotypes have been observed previously in *Arabidopsis* mutants in which the chloroplasts do not differentiate (Reiter et al., 1994).

Unlike *ppi1*, all five *sp ppi1* mutants contain a recognisable, although not completely contiguous, palisade mesophyll cell layer (Figure 2 b-f). Additionally, the epidermis of all five *sp ppi1* mutants is more uniform and less undulating than that in the *ppi1* mutant, probably due to the more regular arrangement of the underlying cell layer. Furthermore, the mesophyll cells of the four *sp1 ppi1* mutants appear to be larger than those in *ppi1*; mesophyll cells in *ppi1 sp2*, however, were approximately the same size as those in *ppi1* (Figure 2 f). The *sp1-1 ppi1* mutant has fewer intercellular air spaces than the other *sp ppi1* seedlings, which is consistent with the notion that *sp1-1 ppi1* carries a different, more severe lesion than its three alleles.

3.4. Plastid Ultrastructure in the *sp* Mutants

Plastid ultrastructure in mutant *ppi1* seedlings has previously been studied (Jarvis et al., 1998). Chloroplasts in *ppi1* are smaller than those in wild-type seedlings and contain fewer thylakoid membranes with smaller granal stacks, as can be seen in Figure 3 (compare a with

f). The *sp ppi1* chloroplasts appeared to be larger and to contain more thylakoid membranes than *ppi1* chloroplasts (Figure 3). To determine more precisely the extent to which the *sp* mutants suppress the chloroplast biogenesis defect of *ppi1*, we quantified the size of *sp ppi1*, *ppi1* and wild-type chloroplasts. By approximating the cross sectional area of a chloroplast to an ellipse, we calculated that the cross sectional area of *ppi1* chloroplasts is approximately 2.9 times smaller than that of wild-type chloroplasts (Table 3.2). Using the same method, we analysed two *sp1 ppi1* mutants (*sp1-1 ppi1* and *sp1-4 ppi1*) and the *sp2 ppi1* mutant. We found that the cross sectional areas of the *sp ppi1* mutant chloroplasts were between 1.5 and 2.0 times larger than that of *ppi1* chloroplasts (Table 3.2). Since chloroplasts are not always sectioned through their widest girth, measurements varied considerably within a particular mutant sample, which explains the relatively high standard deviations of sample populations. As previously shown, *sp1 ppi1* mutants suppressed the yellow-green *ppi1* phenotype more than the *sp2 ppi1* mutant.

3.5. Biochemical Analyses of the *sp* Mutants

The gene disrupted in *ppi1* encodes a translocon component of the outer plastid envelope, atToc33, which is involved preferentially in the import of photosynthetic precursor proteins (Kubis et al., 2003). In the absence of atToc33, chloroplasts import photosynthetic precursors less efficiently, and so accumulate less photosynthetic proteins. This reduced accumulation of photosynthetic proteins can easily be observed when chloroplast total protein samples are analysed by SDS-PAGE and Coomassie staining (Figure 4a). The three major photosynthetic proteins are all reduced in abundance in *ppi1* chloroplasts, compared with wild-type chloroplasts, even though many other chloroplast proteins appear to accumulate normally (Figure 4a). The *sp1-3 ppi1* mutant contains more of the three major photosynthetic proteins than *ppi1*.

Chloroplasts isolated from *ppi1* seedlings have previously been shown to import the 33 kDa oxygen evolving complex precursor, preOE33, ~80% less efficiently than wild-type chloroplasts (Kubis et al., 2003). To determine whether the yellow-green *ppi1* phenotype is suppressed by an improvement in import, or by another mechanism, we carried out *in vitro* import experiments on isolated chloroplasts (Figure 4b). In three independent import experiments, the *sp1-3 ppi1* mutant chloroplasts imported more *in vitro* translated precursor protein than *ppi1* mutant chloroplasts. If a value of 0% is arbitarily assigned to ppi1 and a value of 100% is assigned to wild type, import into chloroplasts isolated from *sp1-3 ppi1*

recovered by 37.4 ± 11.6%. Therefore, the *spl-3* mutation, and presumably also the three other allelic mutations, suppress the *ppi1* phenotype by improving protein import efficiency.

4. Discussion

A forward genetic screen for second-site suppressors of the atToc33 knockout mutation, *ppi1*, was carried out in order to identify novel loci affecting chloroplast protein import in *Arabidopsis*. The visible, yellow-green phenotype of *ppi1* plants is associated with the presence of smaller chloroplasts containing fewer thylakoid membranes and smaller granal stacks, and reduced efficiency of chloroplast protein import (Jarvis et al., 1998). Two different, recessive mutations, *sp1* and *sp2*, were identified that suppress all tested aspects of the *ppi1* phenotype. Doubly homozygous *sp ppi1* seedlings are greener than *ppi1* seedlings, and contained significantly higher concentrations of chlorophyll than the *ppi1* single mutant (Figure 1b). Chloroplasts of *sp ppi1* seedlings also exhibit the *ppi1* phenotype to a lesser degree than *ppi1* seedlings at the ultrastructural level (Figure 3, Table 3.2). Recent data indicate that atToc33 controls the import of photosynthetic precursor proteins preferentially, since the *ppi1* mutant was found to be specifically defective in the expression, chloroplast import, and accumulation of photosynthetic proteins (Kubis et al., 2003). We show that mutant *sp1 ppi1* seedlings contain more photosynthetic proteins than *ppi1*, and that chloroplasts isolated from *sp1 ppi1* seedlings are able to import the photosynthetic precursor, preOE33, more efficiently than chloroplasts isolated from *ppi1* seedlings.

The *sp1* mutation was mapped genetically to the bottom of chromosome 1, and since no known loci involved in chloroplast protein import are to be found in this region of the genome (Jackson-Constan and Keegstra, 2001; Sun et al., 2001), we conclude that SP1 is a novel component involved, either directly or indirectly, in the import of proteins into plastids. It is likely that we have isolated four independent mutations in the *sp1* locus, since the mutants originated from three different M_2 populations, and the two mutants originating from the same M_2 population had heritably distinct phenotypes (Figure 1). Since, four independent mutations were identified within the same locus, the data suggest that this locus is a particularly susceptible target for our screen. It is not unusual to identify multiple, independent mutations that fall into the same complementation group and in a single screen (Schneider et al., 1997; Levin et al., 1998; Watson et al., 1998). For example, in a screen of 15,000 EMS-mutagenized M_2 seedlings for low levels of arginine decarboxylase (ADC) activity, seven mutants were identified that fall into two complementation groups (Watson et al., 1998).

The primary plastid protein import defect of *ppi1* seedlings leads to more general defects in overall leaf anatomy (Figure 2). Other mutants with defective plastids, such as the *Arabidopsis pale cress* (*pac*) mutant, the tomato *defective chloroplasts and leaves-mutable* (*dcl-m*) and the *Antirrhinum differentiation and greening* (*dag*) mutant, also exhibit defective leaf anatomy (Reiter et al., 1994; Chatterjee et al., 1996; Keddie et al., 1996). In these mutants, chloroplast development from proplastids (and etioplasts, in the case of the *pac* mutant) is blocked. As is the case in *ppi1*, the palisade mesophyll cells in these mutants do not have the characteristic columnar shape, and instead appear more spherical. The cells of *dag* and *dcl* mutants fail to divide and expand to the same extent as wild-type cells, and the leaves of *pac* seedlings contain extensive intercellular air spaces, to a greater extent than *ppi1* (Reiter et al., 1994; Chatterjee et al., 1996; Keddie et al., 1996). Since plastids synthesize lipids, amino acids, and other metabolites, which are utilized in the remainder of the cell, it is not surprising that mutants with plastid protein import defects also have altered leaf anatomy.

Bearing in mind that the *sp* mutations suppresses the *ppi1* phenotype in every respect tested (Figures 1 to 3), and do not create any additional or pleiotropic phenotypes, it is likely that the SP proteins are involved specifically in the regulation or mediation of plastid protein import. The data shown in Figure 4, indicating the partial suppression of the *ppi1* import defect by *sp1*, are particularly informative in this regard. Had we observed suppression of only some aspects of the *ppi1* phenotype, a role for the *SP* loci in processes unrelated to import, such as chlorophyll or lipid biogenesis would have seemed more likely.

We predict that the *ppi1* phenotype can only be suppressed by components that affect protein import prior to, or during the stage which atToc33 normally functions. It is unlikely that proteins that function downstream of this stage could suppress the *ppi1* phenotype. A nuclear-localized factor that upregulates the expression of specific translocon genes has already been identified (Sun et al., 2001). Possible mechanisms for suppression of the *ppi1* phenotype may involve the upregulation of the amount of atToc34 protein, or some other translocon components. While the expression of the *atTOC34* gene is not significantly upregulated in the *ppi1* mutant (Sun et al., 2001; Kubis et al., 2003), but the transgenic overexpression of *atTOC34* compensates for the absence of atToc33, demonstrating a significant functional redundancy between atToc33 and atToc34 (Jarvis et al., 1998).

In addition, there are several potential mechanisms by which the activity of atToc34 protein (or other components) might be upregulated in the *sp* mutants. Since EMS mutagenesis frequently (but not always) creates loss-of-function mutations, possible targets of our screen

include specific, negative regulators of plastid protein import. For example, atToc33 and atToc34 are both GTPases, and GTPase proteins are regulated by other proteins that either accelerate or impede GTPase cycling. The activity of atToc34 could potentially be increased in *sp* mutants if the SP protein normally impedes atToc34 GTPase cycling. As mentioned previously, additional atToc34 compensates for the absence of atToc33 (Jarvis et al., 1998).

Another mechanism by which plastid protein import could be upregulated in *ppi1* would be via a mutation in one of the other translocon components, either preventing it from interacting with some inhibitory factor or by blocking it in an active (GTP-bound) form. This mechanism does not apply in the case of *spl* since it has been mapped to an area of the genome that does not encode any translocon components. Accordingly, we identify two potential targets of this forward genetic screen as transcription factors and inhibitory proteins, which could act directly on the translocon complex, such as the GTPase protein regulators, guanine nucleotide dissociation inhibitors.

Arabidopsis contains multiple homologues of translocon components, some of which have been shown to be developmentally and spatially regulated (Jarvis et al., 1998; Bauer et al., 2000; Gutensohn et al., 2000; Kubis et al., 2003). To date just one transcription factor, CIA2, has been identified that regulates plastid protein import (Sun et al., 2001). It is very likely that there are other, as yet unidentified, components that also regulate plastid protein import. We have identified two loci which, when mutated, suppress the *ppi1* phenotype. One, if not both, of these loci encode a novel component of plastid protein import. Further analyses of these mutants will unravel the process of plastid protein import and how it is regulated.

Table 3.1 - Segregation of *sp* phenotypes in backcrossed F_2 populations, from a cross between homozygous *sp ppi1* and *ppi1* (*Ler* ecotype) plants.

Line number	Phenotype		Ratio	x^2-value[a]	p-value[a]
	ppi1 like	Suppressor			
2-4-3	67	29	2.3	1.125	<0.10
14-13	255	83	3.07	0.016	<0.99
5-1	239	84	2.85	0.125	<0.50
2-3-10	115	36	3.19	0.055	<0.95
4-1-16	226	67	3.37	0.602	<0.10

[a] The hypothesis that the *sp* mutations were segregating in a Mendelian fashion was tested statistically. Since there were more yellow-green seedlings than those with suppressor phenotypes, we hypothesized that the phenotypes were segregating three yellow-green seedlings for each suppressor seedling. Given that the data is enumerate (discontinuous) it was analysed using the Chi squared test, and because there are only two categories of data (one degree of freedom) we applied the Yates correction. In every case, no significant difference was observed between our observed data and the expected data (critical c^2-value 3.84, p = 0.05) and the hypothesis was accepted for all five F_2 populations.

Table 3.2 Statistical analysis of chloroplasts from light micrographs taken from a cross-section through the centre of 10-day-old cotyledons.

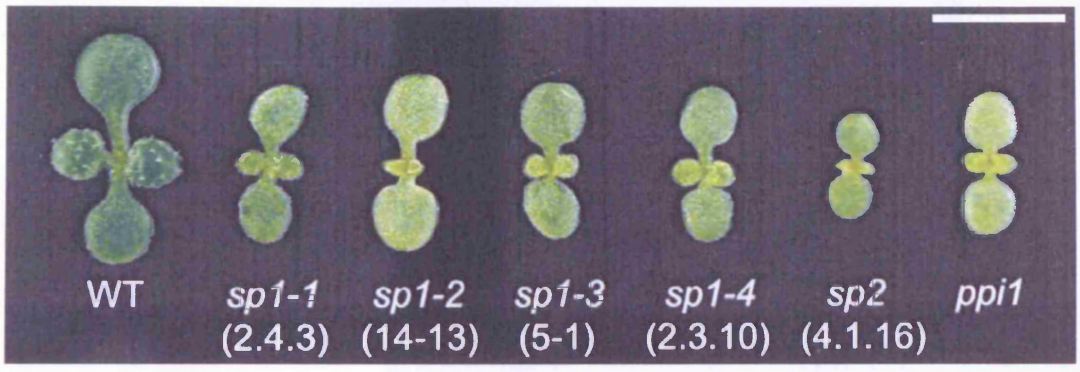
	Length (µm)		Width (µm)		Ellipse (µm^2) [a]		Fold Recovery [b]
	Mean	STDEVP	Mean	STDEVP	Mean	STDEVP	x *ppi1*
WT	5.8	1.2	1.7	0.3	181.9	42.1	2.9
sp1-1	4.9	1.0	1.4	0.5	126.4	53.7	2.0
sp1-4	4.5	1.0	1.3	0.3	107.7	26.4	1.7
sp2	3.9	1.3	1.3	0.5	93.9	49.9	1.5
ppi1	3.6	0.9	1.0	0.3	63.3	22.4	1.0

[a] Formula for the area of an ellipse = Π x length x width

[b] Recovery = mean ellipse area of suppressor mutant / mean ellipse area of *ppi1*

The data were confirmed to indicate significant differences by statistical analysis using ANOVA. The hypothesis that each mean is different from the mean of the *ppi1* chloroplasts was tested. The p-value <0.002 and therefore smaller than the critical value of 0.05 in every case. Therefore, we can conclude that all *sp ppi1* chloroplasts are significantly larger than *ppi1* chloroplasts.

(a)



(b)

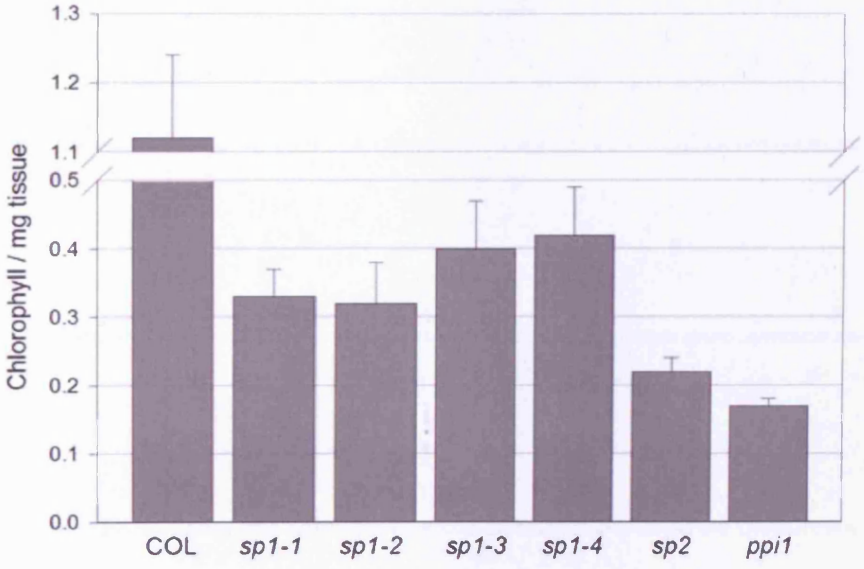


Figure 1. Visible phenotypes of the *sp ppi1* mutants. Photographs (a) and chlorophyll measurements (b) of wild-type (WT), *ppi1 sp* mutants and *ppi1* 10-days-old seedlings. Scale bar = 10 mm.

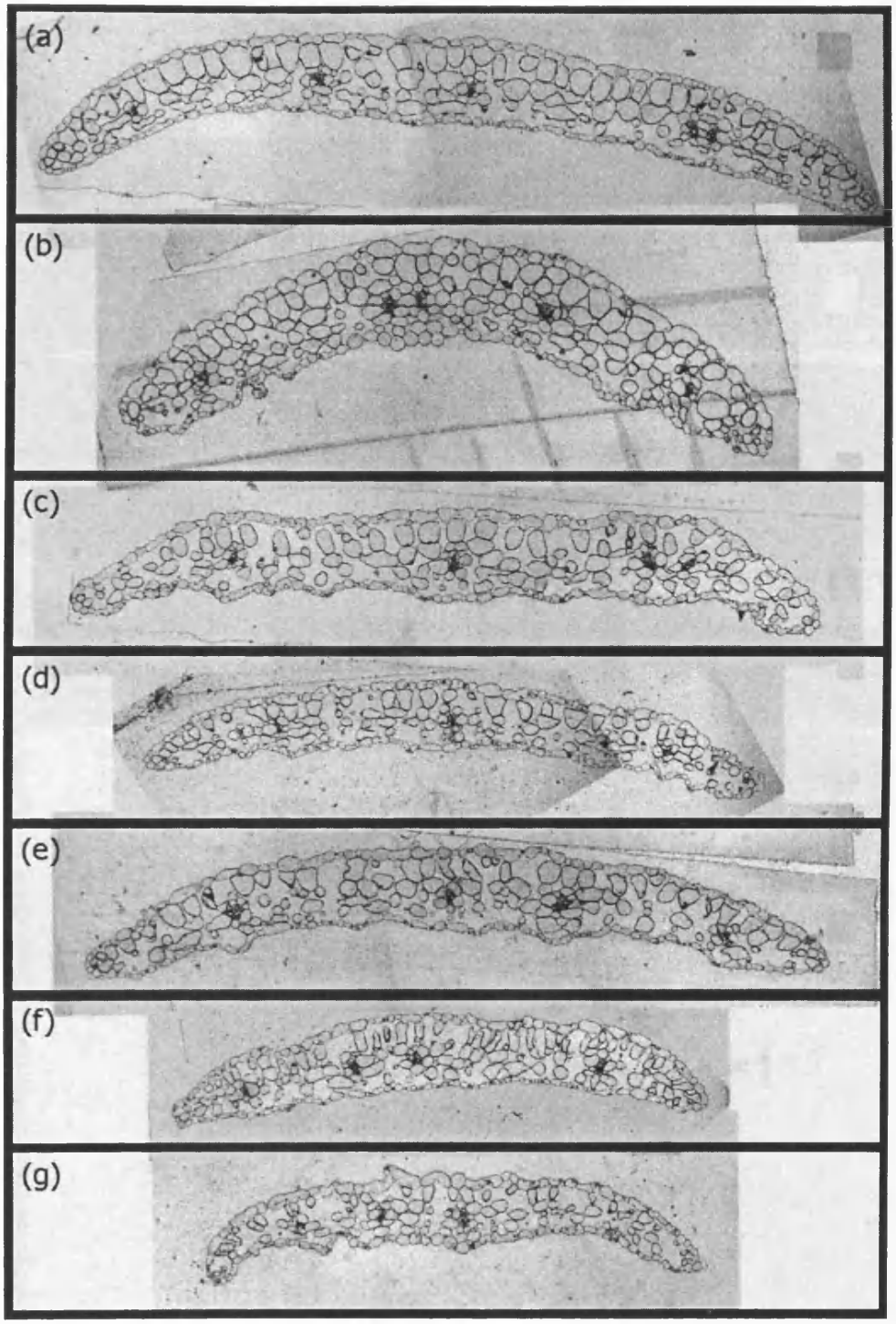


Figure 2. Light microscope images of cross sections through the centre of 10-days-old cotyledons from wild-type (a), *ppi1 sp1-1* (b), *ppi1 sp1-2* (c), *ppi1 sp1-3* (d), *ppi1 sp1-4* (e), *ppi1 sp2* (f) and *ppi1* (g) mutant plants. Scale bar = 1 mm.

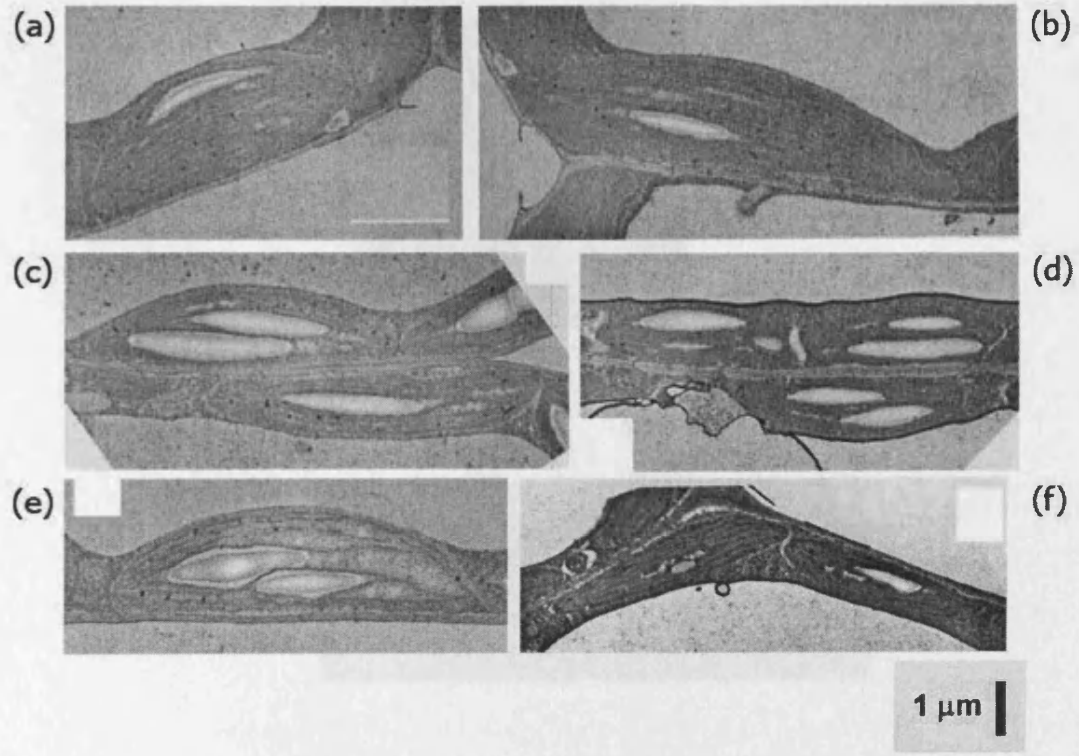


Figure 3. Chloroplast ultrastructure. Representative electron micrograph images of wild-type (a and b), *sp1-1* (c), *sp1-4* (d), *sp2* (e), and *ppi1* (f) chloroplasts from 10-day-old light grown plants. Scale bar = 1 µm.

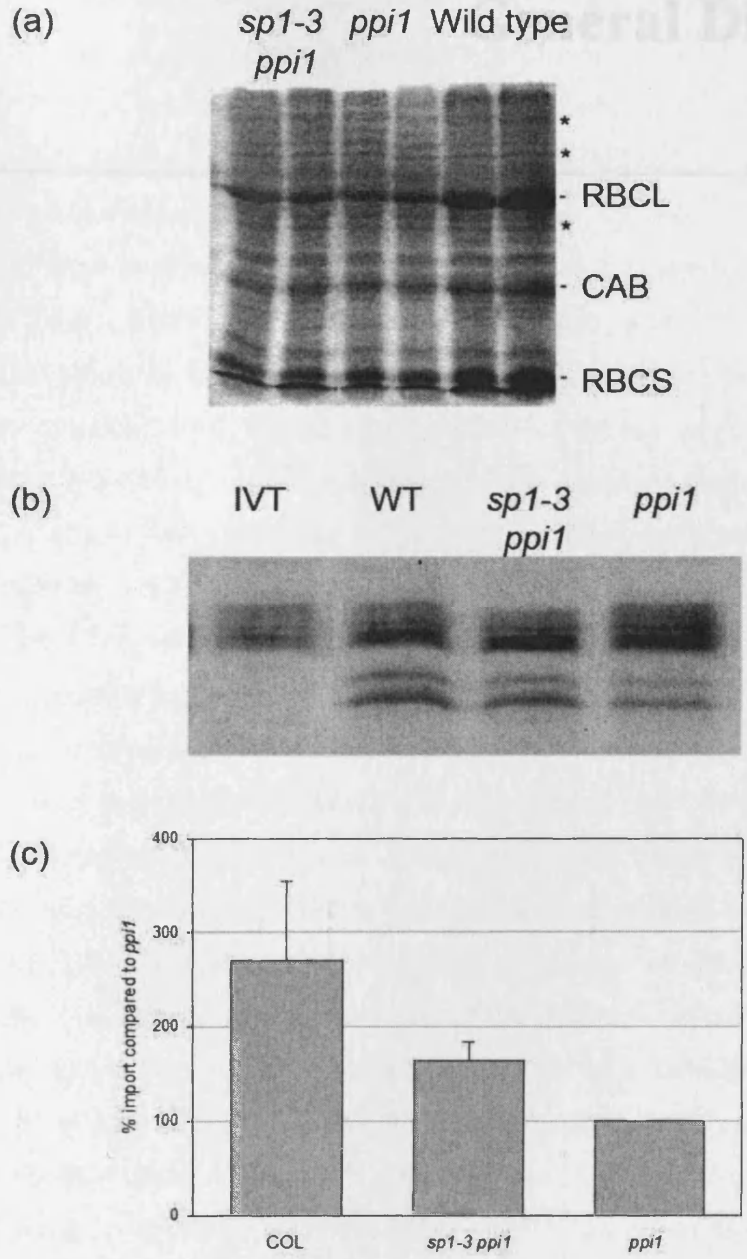


Figure 4. Visualization of total protein content of wild-type, *sp1-3 ppi1* and *ppi1* isolated chloroplasts, and assessment of their import capabilities. (a) Proteins from 10 million chloroplasts (per lane), including the chlorophyll a/b binding protein (CAB), and the large and small subunits of rubisco (RBCL and RBCS, respectively). * indicates protein bands with equal abundance. (b) + (c) *In vitro* translated preOE33 protein was imported into wild-type, *ppi1* and *sp1-3 ppi1* isolated chloroplasts for between 4 and 6 min under standard import conditions (Aronsson and Jarvis, 2002). Precursor (p), intermediate (i) and mature (m) OE33 protein are labelled in (b) and quantification of three independent import experiments is shown in (c); *sp1-3 ppi1* and *ppi1* samples, n = 3, wild-type (WT) sample, n = 2.

General Discussion

In summary, molecular, genetic and biochemical techniques have been used to examine plastid protein import in *Arabidopsis*. Specifically, the *Arabidopsis* Toc75 gene family was characterised in a reverse-genetic study, and a forward-genetic screen was employed to identify novel loci that are involved in plastid protein import, most likely upstream components, such as transcription factors or regulatory proteins.

Three Toc75-related sequences had previously been identified in *Arabidopsis* (The-*Arabidopsis*-Genome-Initiative, 2000; Jackson-Constan and Keegstra, 2001). Characterisation of these sequences revealed that *atTOC75-I* is a pseudogene, that the main homologue of pea Toc75, *atTOC75-III,* is essential for viability, and that *atTOC75-IV,* which was not previously known to be expressed, appears to play a role in etioplasts (Results I). Although the role of atToc75-III, as the main homologue of pea Toc75 seems clear, the precise role of atToc75-IV is still to be explored. It would be interesting to use an inducible promoter to 'knock-down' the expression *atTOC75-III* by RNA interference (RNAi), since this would enable us to see if all parts of the mature plant were affected equally by the absence of the atToc75-III protein, and whether atToc75-IV is able to compensate to any degree during later development. Additionally, since atToc75-IV appears to play a role in etioplasts it would be interesting, but experimentally difficult, to analyse Toc complexes from etioplasts and compare their composition with Toc complexes from other types of plastid, such as chloroplasts and root plastids. It would also be interesting to explore whether translocon components, especially the atToc75-III protein, accumulate to a higher level in etioplasts from *toc75-IV* seedlings than wild-type seedlings by immunoblot analysis. Recently, a protein (atToc64-V / mtOM64) with high sequence identity to the putative chloroplast protein import receptor, Toc64, was shown to be a mitochondrial membrane protein (Chew et al., 2004). The localization studies presented in Results I (Figure 3) do not preclude the possibility that atToc75-IV may also be localized to other subcellular compartments in addition to the chloroplast. Steps are being taken to stably express an atToc75-IV:YFP gene fusion in *Arabidopsis* seedlings.

In the forward-genetic screen, two loci were identified which, when mutated, partially suppress the *ppi1* mutation in every aspect tested (Results II). The *sp1* mutation is a novel mutation since it maps to a region of the genome that is not occupied by any known loci involved in chloroplast protein import (Jackson-Constan and Keegstra, 2001; Sun et al., 2001). In the future, identification of the *sp* genes by map-based cloning, protein localization studies, immunoblot data of other translocon components (especially atToc34), and analysis of the suppression specificity of the *sp* mutations (by crossing the *sp* mutations into other

translocon mutant backgrounds; (Chou et al., 2003; Constan et al., 2004)), will further unravel the process of plastid protein import and how it is regulated.

For example, if a SP protein localizes to the nucleus, like CIA2, it might act as a transcription factor (Sun et al., 2001). Expression of the translocon complex components must be regulated, either transcriptionally or translationally, to enable appropriate regulation of plastid protein import (Dahlin and Cline, 1991). More import complexes are likely to be required in tissue that is rapidly growing, since the plastids will also be rapidly expanding and dividing (Dahlin and Cline, 1991). This was reflected in the pattern of *atTOC75-III* gene expression (Results I, Figure 2), which was highest in young, rapidly dividing tissue. Additionally, the plastid import apparatus must be regulated according to changes in the environment or endogenous, developmental cues (Wan et al., 1996). For example, if the plastids are changing from amyloplasts to chloroplasts (Wan et al., 1996), or the environment changes between dark and light. It would not make sense for plants to produce an import apparatus that preferentially imports photosynthetic precursor proteins (Bauer et al., 2000; Kubis et al., 2003), in non-photosynthetic tissue or in the dark. It is likely that there are many proteins that regulate plastid protein import but have not yet been identified.

To conclude, our knowledge of plastid protein import is steadily increasing, which is due, in part, to public resources that have become available relatively recently, such as the sequence of the *Arabidopsis* genome and various seed, clone and data banks (The-*Arabidopsis*-Genome-Initiative, 2000). Characterization studies, like that of the *Arabidopsis* Toc75 gene family described here, are broadening our knowledge of the currently identified translocon components (Stahl et al., 1999; Constan et al., 2004), and forward-genetic screens, like the *ppi1* suppressor screen reported here, are identifying novel components of the plastid protein import apparatus (Sun et al., 2001).

Materials and Methods

1. Abbreviations

1.1. Units

For the most part, Système International (SI) units were employed. Those units used are summarised below.

k	Kilo (10^3)
m	Milli (10^{-3})
µ	Micro (10^{-6})
n	Nano (10^{-9})
bp	Base pair (of nucleic acid)
Da	Dalton (of protein)
g	Gram
m	Metre
l	Litre
mol	Mole
M	Molar concentration (mol per l)
% (v/v)	Percentage concentration (ml per 100 ml)
% (w/v)	Percentage concentration (g per 100 ml)
°C	Degree Celsius
V	Volt
rpm	Revolution per minute
$O.D._{\lambda}$	Absorbance (subscript = wavelength in nm)
Ci	Curie
h	hour
min	minute
s	second

1.2. Chemicals, Solutions and Media

diH_2O	deionised water
L-broth	Luria-broth
L-agar	Luria-agar
Tris	Tris(hydroxymethyl)aminomethane
EDTA	Ethylenediaminetetra-acetic acid
TE	10 mM Tris, 1 mM EDTA
TER	10 mM Tris, 1 mM EDTA, 10 µg/ml RNase
TBE	0.089 M Tris, 0.089 M boric acid, 2.5 M EDTA, pH 8.0
SSC	Standard saline citrate
SDS	Sodium dodecyl sulfate
BSA	Bovine serum albumin
DTT	Dithiothreitol
dNTP	2'-Deoxyribonucleoside 5'-triphosphate
dATP	2'-Deoxyadenosine 5'-triphosphate
dCTP	2'-Deoxycytosine 5'-triphosphate
dGTP	2'-Deoxyguanosine 5'-triphosphate
dTTP	2'-Deoxythymidine 5'-triphosphate
HEPES	*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid

MES	2-(*N*-morpholino)ethanesulfonic acid
MOPS	3-(*N*-morpholino)propanesulfonic acid
TEMED	*N,N,N',N'*-tetramethylethylenediamine
DMSO	Dimethyl sulfoxide
BAP	6-Benzyl aminopurine
OAc	acetate
Protein Extraction Buffer	PEB
EtBr	Ethidium Bromide
IMS	Industrial methylated spirits
APS	Ammonium persulfate
TBS	Tris buffer saline
TBS-tween	Tris buffer saline, with tween
IPTG	isoproylthiogalatoside
X-Gal	5-bromo-4-chloro-3-indolyl-β-galactoside
MS	Murashig and Skooge

1.3. Biological and Molecular Genetic Terms

DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
A	Adenine
C	Cytosine
G	Guanine
T	Thymine
U	Uracil
cDNA	Complementary DNA
T-DNA	Transfer DNA
PCR	Polymerase chain reaction
RT-PCR	Reverse transcriptase PCR
F1, F2, etc.	Filial generation 1, 2, etc.
M1, M2, etc.	Mutant generation 1, 2, etc.
T1, T2, etc.	Transformant generation 1, 2, etc.
EST	Expressed sequence tag
RACE	Rapid amplification of cDNA ends
PAGE	Polyacrylamide gel electrophoresis
$N_{2(l)}$	Liquid nitrogen
IPA	Propan-2-ol

2. Molecular Biology

DNA was stored at -20°C and RNA and protein samples were stored at -80°C. Unless otherwise stated all non-refrigerated centrifuge steps were carried out in an Eppendorf Centrifuge 5415 D, using a F45-24-11 fixed angle rotor, max speed 16,110 *x g* (13,200 rpm). All refrigerated steps were carried out using a refrigerated Eppendorf Centrifuge 5417, F-45-30-11, fixed angle rotor, max speed 25,000 x g, (16,400 rpm), with the exception for experiments involving chloroplasts and floral dipping, which were carried out using a Sorvall refrigerated RC-5B Super-speed centrifuge.

2.1. Nucleic Acid and Protein Preparation

2.1.a. Plant DNA Extraction

A quick method of preparation was used to obtain DNA from plant tissue for PCR. Approximately 0.15 g plant tissue, previously collected in $N_{2(l)}$ and stored at -80°C, was ground to a powder in a 1.5 ml microcentrifuge tube using a small blue pestle of the appropriate size. The powdered tissue was suspended in 0.5 ml EB (200 mM Tris, pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% SDS) and vortexed briefly to disperse large clumps. Each sample was carried out individually until this point and then stored on ice until all other samples were ready to proceed to the next stage. Samples were centrifuged for 5 min at 13,200 rpm, and subsequently the supernatants were transferred to a fresh 1.5 ml microcentrifuge tube containing 0.5 ml IPA. After 10 min incubation at room temperature, the samples were centrifuged for 7 min at 13,200 rpm. The supernatants were discarded and the pellets washed with 1 ml ice cold 70% (v/v) ethanol. The pellets were dried for approximately 10 min on the bench and resuspended in ~50 µl TER, depending on the size of the pellets. After a further 5 min at room temperature, the samples were centrifuged for 2 min and the supernatant transferred to a fresh tube.

2.1.b. Plant RNA Extraction

RNA was extracted from ~0.15 g plant tissue samples that had previously been collected and stored at -80°C. In the fume hood, a pestle and mortar cooled with liquid nitrogen was used to grind the tissue samples, 150 µl phenol and 500 µl RNA extraction buffer (100 mM NaCl,

10 mM Tris pH 7.5, 1 mM EDTA, 1% (w/v) SDS) to a fine powder. The powder was transferred to a 1.5 ml microcentrifuge tube, and 250 µl chloroform was added prior to a brief vortex and centrifugation at 13,200 rpm for 3 min. The supernatant was transferred to a fresh 1.5 ml microcentrifuge tube and an equal volume of 4 M LiCl was added to the supernatant. The sample was now allowed to incubate on ice for 4 h or overnight in the refrigerator. Each sample was done individually until this point.

After incubation the samples were centrifuged for 10 min at 13,200 rpm, and the pellet was resuspended in 300 µl diH_2O, 3 µl 3 M NaOAc, pH 5.5 and 2.5 volumes of ethanol (EtOH). After incubation for > 1 hr at -20°C the samples were centrifuged for 10 min, and the supernatant was discarded. The pellet was redissolved in 40-100 µl diH_2O, depending on its size.

2.1.c. Rough Plasmid Prep from Bacterial Overnight Cultures

The alkaline lysis protocol employed was an adaptation of that described by Birnboim and Doly (Birnboim and Doly, 1979). Aliquots (5 ml) of L-broth containing the appropriate antibiotic (generally ampicillin for pGEM-T, Promega), in glass 25 ml universal tubes were inoculated with single bacterial colonies (*E. coli* or *A. tumefaciens*) and growth was allowed to proceed overnight at 37°C in a shaking incubator. All subsequent steps were carried out in 1.5 ml microfuge tubes. Cells (between 1.5 and 3.0 ml overnight culture) were pelleted by centrifugation at 13,200 rpm for 1 min in a bench-top microfuge, and then resuspended in 100 µl solution I (50 mM D-glucose, 25 mM Tris-HCl, pH 8.0, 10 mM EDTA). Incubation at room temperature for 5 min was followed by the addition of 200 µl solution II (0.2 M NaOH, 1 % (w/v) SDS; this solution was freshly prepared); samples were mixed by inversion. Incubation on ice was allowed to proceed for 5 min prior to the addition of 150 µl solution III (3 M potassium, 5 M acetate, pH 4.8; solution III was prepared by adding 11.5 ml glacial acetic acid and 28.5 ml distilled water to 60 ml 5 M KOAc). Samples were mixed by inversion, incubated on ice for 5 min, and then centrifuged at 13,200 rpm for 10 min in a bench-top microfuge. Supernatants were transferred to fresh tubes containing 1 ml ice-cold ethanol. Samples were mixed by inversion and allowed to incubate on ice for 5 min. Plasmid DNA was collected by centrifugation at 13,200 rpm for 5 min, washed with 70 % (v/v) ethanol, dried briefly in a vacuum desiccator, and then resuspended in 40 µl TER.

2.1.d. DNA Extraction from Agarose Gels

DNA was extracted from agarose gels using a QIAEX II kit (QIAGEN) to removed free nucleotides and undesirable products of a different size to the required band. The required bands were subsequently cloned. For additional information and the buffer recipes refer to the QIAEX II kit handbook.

DNA fragments were resolved by agarose gel electrophoresis as described later (2.2.a). A clean, sharp razor blade was used to excise the required DNA band from the agarose and excess agarose was removed. The agarose was solubilized by disrupting the hydrogen bonds between sugars in the agarose polymer. This was done with 3 volumes of high salt QX1 buffer to 1 volume of gel. Ten min at 50°C was allowed for solubilisation and adsorption of the DNA to the QIAEXII particles. To keep the silica beads in suspension they were vortexed every 2 min during incubation. The solubilized agarose solution was centrifugation for 30 s at 13,200 rpm and the supernatant was discarded, leaving the DNA immobilised on the silica beads. High salt buffer, QX1 (500 µl), was added to wash away any residual agarose. The solution was centrifuged for 30 s at 13,200 rpm and the supernatant removed. The beads were washed two further times with Buffer PE to remove the residual salt contaminants. To remove any remaining ethanol that could interfere with future enzymatic reactions, the silica beads were allowed to air dry for 10-15 min. The DNA was eluted from the silica particles by resuspending the silica beads in 20 µl of water by vortexing followed by incubation for 5 min at room temperature. The QIAEXII silica particles were collected by centrifugation for 30 s and the purified DNA remained in solution. The supernatant was then transfered to a fresh 1.5 ml microcentrifuge tube.

2.1.e. Total Protein Extraction from *Arabidopsis* Plants

Total protein was prepared for SDS-PAGE from ~0.15 g plant tissue that had been collected on a previous occasion and stored at -80°C. Fresh protein extraction buffer (1 ml) was prepared from a stock (50 mM Tris, pH 6.8, 5% (v/v) glycerol, 0.5% SDS (w/v), 5 mM EDTA) by the addition of 10 mM DTT and 5 µl protease inhibitors (Sigma). Care was taken to keep the samples cold during the primary grinding step to prevent proteolysis. Samples were ground to a powder in a 1.5 ml microcentrifuge tube using a small blue pestle of the appropriate size. Samples was ground further with 100 µl of PEB. Each sample was carried out individually until this point and then stored on ice until all other samples were ready to

proceed to the next stage. Samples were centrifuged for 15 min at 13,200 rpm at 4°C and 50 µl of supernatant was transferred to a fresh tube.

2.2. Electrophoresis

2.2.a. Agarose Gels

Agarose gels were employed to resolve DNA and RNA fragments and to determine the concentration of DNA fragments, using Lambda standards and the intactness of RNA samples. Loading buffer (0.5% Orange G, 50% glycerol, store at 4°C) was added to the samples prior to loading and electrophoresis. Gels generally contained between 0.8 and 1.5 % (w/v) agarose (depending upon the sizes of fragments to be resolved) and 300 µg/l EtBr in 1 x TBE. Agarose was melted and the solution cooled to ~50°C before EtBr was added. For general analyses, small 10 cm x 15 cm gels (Wide Mini-Sub Cell; Bio-Rad) were employed. Larger 20.5 cm x 24 cm gels were used for mapping experiments due the large number of samples. Gels were run in electrophoresis tanks (Bio Rad) containing 0.5 x TBE usually ~10 V/cm. Gels were examined using a short wavelength ultraviolet light transilluminator (UVP) and photographed using a using a BioDoc-it™ system (UVP).

Molecular mass markers (Invitrogen) were used to determine the size of the DNA fragments. 0.05 µg of 1 kb + DNA ladder was added per mm width of well. This DNA produced bands of 12,000, 11,000, 10,000, ... 2000 bp, 1650 bp, 1000 bp, 850 bp, 650 bp, 500 bp, 400 bp, 300 bp, 200 bp and 100 bp.

2.2.b. Denaturing RNA Gels

Denaturing gel electrophoresis was carried out to assess the concentrations and intactness of RNA samples. Denaturing gels consisted of 1.3% (w/v) agarose dissolved in MOPS buffer and 2.55 ml formaldehyde (added just prior to pouring, ~50°C), and were prepared in the same manner as DNA gels. RNA (10-30 µg) samples (total volume not more than 20 µl) were prepared by incubation in 1 x MOPS and 25% formaldehyde/formamide for 20 min at 65°C. EtBr (1 µg) was added to each sample 2-5 min from the start of the 65°C incubation. While the samples were incubating, the gel was pre-run at 60 V in 1 x MOPS with 5% (v/v) formaldehyde. All steps involving formaldehyde were carried out in the fume hood.

Samples were loaded (with 1 x loading buffer) and run in the fume hood at 80 V for more than 2 h. The gel was rinsed with diH_2O to remove traces of formaldehyde before it was viewed using a short wavelength UV transilluminator (UVP) and photographed using a BioDoc-it™ system (UVP).

2.3. Quantification of DNA and RNA

Spectrophotometry was used to calculate the approximate concentrations of double stranded DNA (dsDNA) and RNA of stock solutions. A sample of the stock solution was diluted either 1:100 or 1:200 in diH_2O. The spectrophotometer was calibrated using diH_2O in a thoroughly rinsed cuvette. Measurements were made for each sample at wavelengths 260 nm and 280 nm. The cuvette was rinse thoroughly again and the blank checked with diH_2O. If the blank had drifted an amount significant enough to affect the result values the samples were repeated.

The concentration of the stock dsDNA can be calculated using the equation:

1 O.D. unit 260 nm = 50 ng/µl of dsDNA.

Similarly, 1 O.D. unit 260 nm = 40 ng/µl of RNA

2.3.a. SDS-PAGE

The following stock solutions were prepared for SDS-PAGE:

Solution A: Acrylamide stock solution; 30% (w/v) acrylamide and 0.8% (w/v) bis-acrylamide.

Solution B: 4 x Separating Gel Buffer; 1.5 M Tris-HCl (pH 8.8), 0.4% (w/v) SDS.

Solution C: 4 x Stacking Gel Buffer; 0.5 M Tris-HCl (pH 6.8), 0.4% (w/v) SDS.

5x Sample buffer; 60 mM Tris-HCl (pH 6.8), 25% (v/v) glycerol, 2% (w/v) SDS, 14.4 mM 2-mercaptoethanol, 0.1% (w/v) bromophenol blue.

The Stock solutions mentioned above were stable for several months in the refrigerator.

For 1 l of Coomassie stain: 1.0 g Coomassie Brilliant Blue 250, 450 ml ethanol, 450 ml H_2O, 100 ml glacial acetic acid.

For 1 l Destain: 100 ml methanol, 100 ml glacial acetic acid and 800 ml H_2O.

Gels were cast and run in BioRad Protean III apparatus. Glass plates were cleaned with detergent, distilled water and IMS prior to assembly. Protein gels consist of two phases, the

top stacking phase that concentrates the protein to a fine point and the lower separating gel that separates the proteins according to their size. The separating phase was poured first and allowed to set, before the upper stack phase was poured. For two X% separating gels, X/3 ml Solution A, 2.5 ml Solution B, 7.5-(X/3) ml diH_2O and 50 µl 10% ammonium persulphate (APS) were mixed before the addition of 5 µl TEMED. The gel was immediately poured and overlayed with 50% (v/v) IPA. The gel was allowed to set at room temperature for at least 15 min. The overlying solution was removed before the upper phase was poured. Two stacking phases consisted of 2.3 ml diH_2O, 0.67 ml solution A, 1.0 ml Solution C, 30 µl 10% APS and 5 µl TEMED. The stacking phase was poured directly after the addition of TEMED and the combs were inserted. Again, the gel was allowed to set for at least 15 min at room temperature.

Protein samples prepared in 1 x sample buffer (Laemmli, 1970) and heated to 100°C for 2 min before loading into the wells of the gel. Gels were usually run at 150 V until the bromophenol blue reached the bottom of the gel (~1 h).

Gels were stained for at least 30 min with Coomassie stain, and subsequently destained for several hours or overnight, until the protein bands were clearly visible, and the surrounding gel was virtually clear. Radioactive gels for autoradiography were dried for 2 h at 80°C using a BioRad Gel Dryer, model 583.

2.3.b. Immunoblotting

Protein were transferred from SDS-PAGE gels to Hybond ECL nitrocellulose membrane (AP Biotech) using the Bio-Rad wet blotting system. A sandwich consisting of a sponge, 3 MM filter paper, membrane, the pre-run gel, 3 MM filter paper and another sponge mesh, was constructed and held in place by a frame. The proteins were blotted in cold transfer buffer (1.52 g Tris, 14.44 g glycine, 100 ml methanol made up to 1 l with diH_2O) for exactly 1 h at 100 V. The solutions were kept cold by an ice block placed in the gel tank.

Membranes were prepared for blotting by incubation in TBS-Tween with 5% (w/v) dried milk powder (Marvel), for 1 h. All stages of blotting were incubated on an shaker at room temperature. The membranes were probed with antibodies (1:1,000 and 1:5,000) diluted in TBS-Tween with 5% (w/v) dried milk powder. After 2 h the membranes were rinsed three times, and washed three times for 5 min with TBS-Tween, before the secondary antibody was

applied. The secondary antibody was raised in goat against rabbit and was conjugated to alkaline phosphotase (Sigma). The secondary antibody was applied for 1 h in TBS-Tween with 5% (w/v) dried milk powder, before being washed away by three rinses in TBS-Tween and three, 5 min washes in TBS-Tween. The secondary antibody was visualized using Sigma FAST™ (5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium tablets in suspended in 10 ml diH_2O).

2.4. Sequencing

Sequencing was carried out by the Protein and Nucleic Acid Chemistry Laboratory (PNACL) at the University of Leicester. DNA (1 μg) was sent for sequencing in 8 μl of diH_2O, along with the appropriate primer (10 μl of 1 μM). The sequencing results were viewed using DNA Star, Editseq and Chromas.

2.5. Bacterial Work

2.5.a. Growing Bacteria

E. coli (DH5α) were cultured in L-broth in shaking incubator or on L-agar plates in a bacterial growth cabinet, at 37°C. When applicable, 50 μg/ml ampicillin was applied. *Agrobacterium*, strain GU3101, was grown in YEP-broth (10 g/l yeast extract, 10 g/l peptone, 5 g/l NaCl, pH 7) in a shaking incubator or on YEP-agar in a bacterial growth cabinet, at 30°C. The following concentrations of antibiotic were applied when applicable; kanamycin 30 μg/ml, gentamycin 50 μg/ml, spectomycin 100 μg/ml.

2.5.b. Ligation

Ligations reactions contained 1x ligation buffer (New England BioLabs), 50 ng vector, 3x molar amount insert and T4 DNA ligase (New England BioLabs) in a total volume of 10 μl. Ligations reaction were either carried out for 1-4 h at room temperature or overnight at 4°C, and subsequently transformed into *E.coli*. Alternatively, the pGEM®-T Easy Vector System (Promega) was used and procedures were carried out according to the manufacturers instructions.

2.5.c. *E.coli* (DH5α) Transformation

Aliquots of competent cells (100 µl) in 1.5 ml microfuge tubes were thawed on ice for 10 min. Plasmid DNA (<20 µl) was added to the cells and mixed gently, and incubated on ice for 20 min. Cells were subjected to heat shock at 42°C for 30 s, followed by 2 min incubation on ice. L-broth (800 µl) was added to each sample prior to incubation at 37°C for 1 h. Cells were pelleted by centrifugation at ~10,000 *x g* for 30 s. Most of the supernatant was decanted and the cells were resuspended in the remaining solution (approximately 100 µl) before spreading on L-agar plates containing the appropriate antibiotic. Plates were incubated upside down overnight at 37°C in order to allow the formation of single colonies. If pGEM®-T (Promega) plasmid was transformed the plates supplemented with 0.5 mM isoproylthiogalatoside (IPTG) and 80 µg/ml 5-bromo-4-chloro-3-indolyl-β-galactoside (X-Gal). IPTG induces the activity of the β-galactosidase gene that spans the insertion site in the vector. In plasmids with no insert, β-galactosidase is produced, which hydrolyses X-Gal to produce a deep blue indigo compound. If the DNA fragment has successfully inserted into the plasmid, β-galactosidase cannot be produced and therefore no blue colouration is present. After overnight incubation in the 37°C bacterial growth cabinet, the plates were scored and the white colonies picked individually and cultivated.

2.5.d. *Agrobacterial* Transformation

Aliquots of competent cells (100 µl) in 1.5 ml microfuge tubes were thawed on ice for 10 min. Plasmid DNA (5 µg) was added and mixed gently. The cells were incubated on ice for 5 min prior to being flash frozen in $N_{2(l)}$ and thawed at 37°C. L-broth (1.2 ml) was added to the cells and they were incubated at 30°C in a shaking incubator for 30 min. Cells were pelleted by centrifugation at ~10,000 *x g* for 30 s. Most of the supernatant was decanted and the cells were resuspended in the remaining solution (~100 µl) before spreading on L-agar plates containing the appropriate antibiotic. The cells were grown for up to 3 days in a 30°C bacterial growth cabinet.

2.5.e. Restriction Analysis of Plasmid DNA

For general analyses, between 0.5 μg and 1 μg plasmid DNA (approximately 5 μl small scale plasmid preparation) was used. Enzymes and buffers were obtained from New England BioLabs, and the recommend buffer was used for each enzyme. At least 3 units restriction endonuclease were used per microgram plasmid DNA. The total volume of enzyme used was kept below 10% of the total reaction volume since most restriction enzyme storage buffers contain 50% (v/v) glycerol; reaction buffer glycerol concentrations above 5% (v/v) can result in star activity of the restriction endonuclease (cleavage of sequences similar but not identical to their defined recognition sequences). Digests were generally carried out in a total volume of 20 μl, but volumes were increased when larger amounts of DNA were being digested (for fragment isolation experiments) or restriction endonucleases of low concentration were being used. Digestion was allowed to proceed for at least 60 min (at the temperature recommended by the manufacturer) before analysis using agarose gel electrophoresis.

2.6. Polymerase Chain Reaction

The polymerase chain reaction (PCR) is the *in vitro* enzymatic, amplification of a specific DNA sequence using oligonucleotide primers, Taq DNA polymerase, nucleotides and a DNA template. The two primers flank the region of interest and anneal to opposing strands. It is a cyclic reaction sequence that involves the template denaturing, the primers annealing to the template, and the extension of the primers by Taq DNA polymerase. Each event occurs at a specific temperature.

Conditions and reagent concentrations used for the PCR amplification of DNA varied depending on the primers, the nature of the template and the size of the expected product. Primers were generally between 19 and 22 bases. They were designed by eye to have melting temperatures (T_m) of approximately 60°C. Primers were included in reactions at concentrations of approximately 1 μM each. Each dNTP (AB gene) was included at a concentration of approximately 0.4 mM. *Arabidopsis* genomic DNA templates were added at concentrations of approximately 50 ng per 50 μl reaction; plasmid DNA templates were included at 100 to 1000-fold lower concentrations. Best results were obtained using ExTaq (TaKaRa) polymerase in conjunction with the accompanying PCR buffer at 1 x concentration. ExTaq DNA polymerase was included at a concentration of 2.5 units per 50 μl reaction. Amplifications were typically carried out for 35-40 cycles as follows: i) template denaturation

was carried out at 94°C for 30 s; ii) primer annealing was carried out at temperatures between 55°C and 60°C for 30 s; and iii) elongation was carried out at 72°C for between 1 and 3 dependant upon the length of the expected product (30 s was allowed for each 1,000 bp). Amplifications were generally preceded by a template denaturation step (94°C for 3 min) and followed by 5 min incubation at 72°C. Thermal cyclers used were supplied by Biometra® (T gradient or T3 thermocycler).

2.6.a. Identification of *Arabidopsis* knockout mutants

The mutants used in this study were obtained from the Csaba Koncz Laboratory (Ríos *et al.*, 2002), Salk Institute Genomic Analysis Laboratory (SIGnAL) (Alonso *et al.*, 2003), Sainsbury Laboratory *Arabidopsis* Transposants (SLAT) (Tissier *et al.*, 1999), and Syngenta (Sessions *et al.*, 2002) collections, using published procedures. Mutant identification details are as follows: *toc75-III-1* (Csaba Koncz Laboratory, pool 345, line 3449); *toc75-III-2* (SIGnAL line SALK_15928); *toc75-IV-1* (SLAT, superpool 6, subpool 25); *toc75-IV-2* (Syngenta, line Garlic_192_B10). The *toc75-III-2* mutant was supplied to us via the Nottingham *Arabidopsis* Stock Centre (NASC; accession number N515928). All of the mutants contained T-DNA insertions, except *toc75-IV-1* which contained a stable transposon (*dSpm*) insertion. The kanamycin-resistance marker of the *toc75-III-2* T-DNA insertion was inactive due to silencing effects; this is a problem associated with ~20% of the SIGnAL TDNA lines (Alonso *et al.*, 2003).

Gene-specific primers and insertion-specific primers (for the T-DNA left and right borders, LB and RB, or the 5' and 3' ends of the *dSpm* insertion) were used to follow the mutations by PCR. Plant DNA for PCR analysis was extracted as described previously (Edwards *et al.*, 1991). Primer combinations used for these analyses were as follows: *atTOC75-III* (5' gene-specific, 5'-GAA GCT GAA TTA TAG GAC TCA CAT TTG TAG- 3'; 3' gene-specific, 5'-CTC AGA TGC AGA CAG ACG TGT TAC C-3'; *toc75-III-1* LB, 5'-CTG GGA ATG GCG AAA TCA AGG CAT C-3'; *toc75-III-2* LB, 5'-GCG TGG ACC GCT TGC TGC AAC T-3'); *atTOC75-IV* (5' gene-specific, 5'-CCA ATG TTT GTG GGT CGA GAT T-3' or 5'-CCG AGC TCA TGG ATT TCT TCT TTG TTG TTC AGG-3'; 3' gene-specific, 5'-GCT GAA TGA AAC TTG TTA ATG ATA GT-3'; 5' *dSpm*-specific, 5'- GGT GCA GCA AAA CCC ACA CTT TTA CTT C-3'; 3' *dSpm*-specific, 5'-GTC CAT TTT AGA GTG ACG GCT AAG AGT G-3'; *toc75-IV-2* LB, 5'-TAG CAT CTG AAT TTC ATA ACC AAT CTC GAT ACA C-3'). PCR products corresponding to the insertion junctions were sequenced to verify the

location of each T-DNA or transposon (Figure 1a). Single-locus insertion lines were identified for each mutant (Table 1); for *toc75-IV-1* and *toc75-IV-2*, homozygotes derived from these lines were used for further study.

2.6.b. RT-PCR

First strand cDNA, complementary to the RNA (5 μg) previously extracted, was synthesised in a two part reaction. A 12 μl reaction containing, 5 μg of RNA, 5 μl of CDS-5' primer (5'-$(T)_{17}$(A/G/C)N-3') and diH_2O was placed in a 70°C water bath for 2 min, which allowed the primers to anneal to the RNA strands. The reaction mixture was then incubated on ice for 2 min. This was followed by the addition of the remaining components; 1x PCR buffer, 2 μl of DTT, 1 μl of dNTPs (2.5 mM), and 1 μl of SuperScript™ II RNase H- reverse transcriptase (Invitrogen). This 20 μl reaction mix was heated to 42 °C in a water bath for 60 min, followed by 15 min at 70 °C. During this stage reverse transcriptase elongated the primers to produce the first strand cDNA.

For atTOC75 expression analysis (Results I, Figure 2b) PCR amplifications were performed using the following primer pairs: *atTOC75-I*, forward (*priF1* 5'-TGG GCA TCG ATC GTG TAG CA-3', *priF2* 5'-AGT ATT GTT CCT GAT CCC-3', *priF3* 5'-ATA TCG AGA TCA TGC CAA T-3'), reverse (*priR1* 5'-AAG CTC AGC ACC AAC CTT TA- 3', *priR2* 5'-TGT CCC TCT CTT CTG AAA GCC-3'); *atTOC75-III*, forward (5'-CGT ATC TGG ATG GTG TTT ACA ATC-3'), reverse (5'-GGA ATT CTT AAT ACC TCT CTC CAA ATC GGA AGA AC-3'); *atTOC75-IV*, forward (5'-CCA ATG TTT GTG GGT CGA GAT T-3'), reverse (5'-GGC TGC AGT TAG TAT CTC TCC CCG AAC C-3'); *atTOC33*, forward (5'- GGT CTC TCG TTC GTG AAT GG-3'), reverse (5'-CTG AGC GCC TAT GAT AAG AG-3'); *eIF4E1*, forward (5'-AAA CAA TGG CGG TAG AAG ACA CTC-3'), reverse (5'-AAG ATT TGA GAG GTT TCA AGC GGT GTA AG-3'). Each primer combination spanned an intron. For the expression profiles shown in Figure 2, PCRs used the minimum number of cycles necessary to observe a faint band following staining with ethidium bromide (*atTOC75-III*, 20 cycles; *atTOC75-IV*, 25 cycles; *eIF4E1*, 20 cycles). Products were resolved by agarose gel electrophoresis, blotted onto Hybond NX membrane (Amersham Pharmacia Biotech), and hybridized with gene-specific probes amplified using the primers listed above, according to the manufacturer's instructions (Amersham Pharmacia Biotech). Bands were visualized using a phosphorimager and quantified using ImageQuant software (Molecular Dynamics). *atTOC75-III* and *atTOC75-IV* data were normalized using similar data for the uniformly

expressed translation initiation factor gene, *eIF4E1* (gene number At4g18040; (Rodriguez *et al.*, 1998).

2.6.c. RACE-PCR

The *atTOC75-IV* mRNA was characterized by RACE-PCR using the SMART™ RACE cDNA amplification kit (Clontech) according to the manufacturer's instructions. The genespecific primers used were as follows: 5'RACE (5'-GCT GAA TGA AAC TTG TTA ATG ATA GT-3' and 5'-AAT CTC GAC CCA CAA ACA TTG G-3'); 3'RACE (5'-CCA ATG TTT GTG GGT CGA GAT T-3' and 5'-ACT ATC ATT AAC AAG TTT CAT TCA GC-3'). Sequences of the 5' and 3' RACE-PCR products were determined, and then used to design primers for the amplification of a full-length cDNA clone by RT-PCR: forward, 5'-AAC TGC AGA ACA ATA TAG AGA GAA AGA AG-3'; reverse, 5'- CCA TCG ATG GAA AGA AAG CAG CAC AAA GTC-3'. The PCR product was cloned into pGEM-T Easy (Promega), verified by sequencing, and then subcloned as a *Sac*II/*Spe*I fragment into pBlueScript II KS- (Stratagene). For overexpression in *Arabidopsis* plants, the *atTOC75-IV* cDNA was subcloned as a *Sac*I/*Sal*I fragment into the plant transformation/expression vector, pCHF2, which incorporates a double enhancer version of the cauliflower mosaic virus 35S promoter (Jarvis et al., 1998).

2.7. Rough Mapping

Homozygous *ppi1* (L*er* ecotype) plants were used to cross-pollinate homozygous *sp* mutants, and the resultant F_1 generations were analysed for the segregation of the *sp* phenotype. Seedlings from the F_2 generation displaying the suppressor phenotype were used for genetic mapping of the *sp* mutations. Fourteen genetic markers distributed across the five *Arabidopsis* chromosomes at approximately 60 cM intervals were selected from the *Arabidopsis* Information Resource (TAIR) database. The markers used were: chromosome 1, M235 at 31.90 cM (TAIR accession number, 1945678), GAPB.1 at 59.00 cM (1945602), AF-6 at 114.0 cM (2005285) and GL2 at 121.70 cM (2005270); chromosome 2, CIC9A3R at 16.50 cM (2005299) and F7F1 at 62.00 cM (2005288); chromosome 3, CA1 at 4.24 cM (1945575) and NGA6 at 86.41 cM (1945538); chromosome 4, GA1.1 at 17.20 cM (1945599) and G8300.2 at 81.22 cM (1945676); chromosome 5, PAT1.2 at 5.98 cM (1945631), NGA139 at 50.48 cM (1945524), PHYC.1 at 71.13 cM (1945634) and LFY3 at 116.88 cM (1945613). The recombination frequency was calculated for each locus (RF = number of recombinants / total number of chromosomes tested x 100).

3. Tissue Culture and Whole Plant Manipulation

3.1. *In vitro* Plant Procedures

Arabidopsis seed to be germinated *in vitro* were surface-sterilised by their immersion in: i) 70% (v/v) ethanol and 0.05% (v/v) Triton X-100 for 5 min; ii) 5 % (v/v) sodium hypochlorite; and iii) 100% ethanol for 10 min. Sterilised seed were dried on filter papers in a laminar flow hood. Seed were generally sown on MS (Murashige and Skoog, 1962) (or MS with selection). Plates were sealed with micropore tape (3M) which prevents contamination but allows gaseous exchange. Seed dormancy was broken by stratification at 4°C for 2 to 4 days. Growth was allowed to proceed at 20°C under long day conditions (16 h light, 8 h dark). When plants germinated *in vitro* were required to set seed, they were transferred to soil and allowed to set seed in the greenhouse.

Plants grown on MS with selection were generally scored after 10 days growth. Resistant plants grown on MS containing selection were sufficiently healthy to be treated the same as plants grown on MS without selection; such plants could therefore be transferred to soil and allowed to set seed in the glasshouse. Sensitive plants were never required. The appropriate concentrations for selection were 50 μg/ml kanamycin, 15 μg/ml phosphinothricin, 110 μg/ml gentamycin and 15 μg/ml hygromycin.

3.2. Greenhouse Cultivation

Arabidopsis plants were grown in soil composed of Levingtons F2 Seed and Modular Compost, silver sand and vermiculite (medium 2.0-5.0 mm, Sinclair). Glasshouse temperature was maintained at 20°C throughout the diurnal cycle; glasshouse humidity was not regulated. Between the months of October and March, a photoperiod of 16 h was maintained using artificial lighting. Plants were shaded with retractable green netting between May and August.

Arabidopsis plants were generally grown in 40 cell compartment trays. Each tray holds 2 l compost, 1 l H_2O, 0.2 l vermiculite and 0.6 l sand. *Arabidopsis* plants germinated *in vitro* were transferred to water-saturated soil after ~3 weeks growth. Plants were kept covered with propagator lids for several days following transferral to soil. *Arabidopsis* seed to be germinated in the glasshouse were sown on soil saturated with water. Trays were covered with cling film and placed in a cold room (4°C) for 2 days (stratification to break seed

dormancy). Following stratification, trays were transferred to the greenhouse. The cling film was removed after one to two weeks. All plants were watered daily from the base.

Arabidopsis seed were harvested by gently patting dry siliques to remove the seed after silique maturation and prior to silique dehiscence. All newly-harvested seed were allowed to dry at room temperature for several days prior to sowing.

Greenhouse pests were controlled by both biological and chemical reagents; aphids were controlled by Aphipar (Koppart), Rapid and Polysect; sciarid flies were controlled by yellow sticky traps, and Entonem (Koppert); thrips were controlled by Thripex (Koppert) and blue sticky traps; and powdery mildew was controlled using Fungus Clear (Scott's).

3.3. Transformation of *Arabidopsis* using the Floral Dipping Method

Plant care was found to very important for achieving the best transformation efficiency. Approximately 30 young, healthy *Arabidopsis* plants were grown in 10 cm^2 pots in a greenhouse and when they started flowering their primary inflorscenses were removed. The secondary inflorscenses with a few open flowers were dipped in solution containing *Agrobacterium.*

Cultures of *Agrobacterium* were set up in YEP from glycerol stocks 3 days prior to dipping. The volume (ml) of the culture was calculated by (0.25 x number of pots). The day before dipping, overnight cultures were set up by adding 100 times the current volume of YEP to the *Agrobacterium* cultures in large conical flasks. When the overnight cultures reached an $O.D._{260}$ of 1.8 the following day, the cells were collected by centrifugation at 5,000 rpm for 10 min in a GSA rotor at room temperature. The cells were gently resuspended in 3 volumes of infiltration media (2.2 g/l Murashige-Skoog salts, 1 ml/l Gamborg's vitamins, 10 ng/ml 6-BAP, pH 5.7) so as not to burst the cells. Silwet (200 µl/l) was added immediately prior to submersing the *Arabidopsis* inflorscenses in the bacteria solution.

The *Arabidopsis* plants were supported on bottle lids, upside down, in the bacterial solution for 10 min. This allows the *Agrobacterium* to enter the *Arabidopsis* buds through the locule which is open 5-10 days prior to anthesis. Transformation occurs in the ovules of the developing flowers after the female and male gametophyte cell lineages form, but before the embryo develops beyond a single cell. After 10 min, excess liquid was drained on tissue and

the pots were loosely wrapped in clingfilm to retain humidity, allowing the motile bacteria more time in which to enter the buds.

Two days after dipping the clignfilm was removed and the plants were allowed to set seed. The seed was bulk harvested and the first generation was screened for transformants (T_1).

3.4. EMS Mutagenesis

Approximately 7,000 homozygous *ppi1 Arabidopsis* seed (Jarvis et al., 1998) of the Columbia ecotype were mutagenized by ethyl methanesulfonate (EMS) according to Leyer and Furner (Leyser and Furner, 1992). Two batches of approximately 3,500 seed were soaked overnight in 0.1% (w/v) KCl inside disposable plastic syringes. The ends of the syringes had been removed and the openings were covered with gauze to contain the seed and allow liquid to flow freely inside. The following day the KCl was replaced with 20 ml of 100 mM potassium phosphate solution (pH 5) and 1 ml dimethyl sulphoxide (DMSO), to permeablise the seeds. EMS is highly toxic so all the equipment was sealed in a glove bag (Sigma) inside a fume hood. Once sealed, 225 µl (100 mM) EMS was added to the seeds and solution, and mixed thoroughly. The seed were syringed up and down every hour, for three hours. The EMS solution was chemically inactivated by discarding it into a beaker of sodium thiosulphate crystals. The seeds were washed three times with 100 mM sodium thiosulphate to remove any traces of EMS. Each wash was discarded into the beaker containing the sodium thiosulphate crystals. The seeds were washed with 100 ml diH_2O before being removed from the glove bag and left to dry on 3MM filter paper. The seed were sown a maximum of two days after mutagenesis.

The mutagenized (M_1) seed was grown on soil in 20 separate trays and the progeny (M_2) were screened on MS medium at ~5 seed/cm^2 density. Seven and 20-day-old M_2 seedlings were examined alongside *ppi1* control seedlings for evidence of *ppi1* phenotypic suppression. Putative suppressor mutants were transferred to separate MS plates for ~7 days, before being transferral to soil and growth to maturity.

3.5. Pollination of *Arabidopsis* Plants

Arabidopsis plants to be used as parents in genetic crosses were germinated *in vitro* in order to minimise the possibility of contamination of seed. Plants were transferred to soil as

described above and allowed bolt. Best results were obtained using young, primary inflorescences, although secondary inflorescences (induced by the removal of the primary inflorescence) were also employed. Crosses were carried out using fine point forceps. Between 3 and 4 flower buds were crossed per female parent, and 2 to 3 identical crosses were made (using distinct but genotypically identical plants) for each genetic experiment. Flower buds selected for crossing were the most mature of those whose stigmas were still fully enclosed within the sepals. The apical meristem and all unwanted buds were carefully excised from the inflorescence. Sepals and petals were removed in order to gain access to the stamens. Each bud was then emasculated by the removal of the stamens; anthers were examined at this point and any bud whose anthers had begun to dehisce were discarded. Emasculated flowers were cross-pollinated using at least two buds from the required male parent. Pollination was facilitated by the removal of buds from the male parent and the excision of petals and sepals to expose the stamens. Inflorescences carrying F_1 seed were harvested directly, shortly after the siliques had begun to dehisce. Newly harvested F_1 seed were allowed to dry at room temperature for at least one week prior to sowing. When F_2 seed were required, F_1 seed were germinated *in vitro* in order to minimise the possibility of cross-contamination of seed. F_1 plants were transferred to soil after approximately 3 weeks, and F_2 seed were harvested as described above.

3.6. Chlorophyll Measurements

To extract the chlorophyll from plant tissue, samples were placed in 1 ml of dimethyl formamide (DMF) in 1.5 ml microcentrifuge tubes (Porra et al., 1989) and mixed for 2 h by an orbital shaker, followed by incubation for overnight at 4°C. Due the light sensitive nature of the chlorophyll, subsequent to the submersion of the plants in DMF the microcentrifuge tubes were concealed in a dark environment using aluminium foil and black bin liners. Approximately, 1.5 mg of fresh plant tissue was incubated in 1 ml DMF. Less DMF and more plant tissue was used for seedlings containing less chlorophyll.

After incubation, measurements of the amount of chlorophyll contained within the DMF were made using a spectrophotometer. The absorption was measured at wavelengths 646.8 nm and 663.8 nm. The accuracy of the measurements was maintained by running blanks (DMF) as samples and comparing them to zero values. Measurements were repeated if the blank values deviated from the zero values to a significant number of decimal places as to affect the readings. The amount of chlorophyll was calculated according to the equations;

Chlorophyll *a* = (13.43 x $A_{663.8}$) - (3.47 x $A_{646.8}$)

Chlorophyll *b* = (22.9 x $A_{646.8}$)-(5.38 x $A_{663.8}$)

Total chlorophyll was calculated per mg fresh weight tissue.

4. Biochemical Techniques

4.1. Chloroplast Isolation

Arabidopsis seeds were sterilized, sown on petri plates and grown for 10 days. Each plate contained ~50-200 seeds. Chloroplasts were isolated according to the procedure described by Aronsson and Jarvis (2002), and as follows. For each isolation procedure, 30-35 petri plates of ~10-d-old plants were used. Plants were homogenized for 3-4 s using a polytron (Kinematica PT20). After the first centrifugation, the homogenate (in 0.5-2 ml isolation buffer) was loaded onto a linear Percoll gradient and centrifuged in a swing-out rotor at 7800 *g* for 10 min. Intact chloroplasts were recovered and washed once with HEPES-sorbitol (HS) buffer (50 mM HEPES, 0.3 M sorbitol). The chloroplast yield was calculated using a haemocytometer. Alternatively, tissue was collected from the tips of young pea seedlings, and the chloroplasts were isolated as described above.

4.2. Import Reactions

Import, thermolysin treatment and quantification were performed according to Aronsson and Jarvis (2002). Template DNA for the *in vitro* transcription/translation of preproteins was amplified by PCR from cDNA clones using M13 primers. The preOE33 cDNA clone was obtained from the *Arabidopsis* Biological Resource Centre as EST clone 119E10T7. The *atTOC75-IV* cDNA clone was isolated by RT-PCR and cloned into pBlueScriptII. The *atTOC75-III* PCR product was amplified from an EST, clone name APZL59c10R, also in pBlueScriptII. The *atTOC75-III* cDNA was cloned from pBlueScriptII SK in pBlueScriptII KS in order to utilize to T7 promoter. Transcription/translation was performed using a wheat germ system (Promega) containing [^{35}S]methionine and T7 RNA polymerase according to the manufacturer's instructions.

Import reactions were carried out in HMS buffer containing 20 mM gluconic acid (potassium salt), 10 mM $NaHCO_3$ and 0.2% (w/v) BSA (Seedorf and Soll, 1995). Each 150 µl import assay contained 10^7 chloroplasts (~20 µg chlorophyll when using wild type), 5 mM MgATP, 10 mM methionine, and translation product not exceeding 10% of the total volume. Import reactions were incubated at 26°C for 1-20 min, depending on the experiment. Reactions were stopped with an equal volume of stop buffer (50 mM EDTA, 0.3 M sorbitol, 50 mM HEPES, pH 8.0), and the chloroplasts were pelleted by centrifugation for ~5 s at 16,110 *x g*.

To deplete endogenous ATP, chloroplasts were kept in the dark at room temperature for 20 min and treated with 6 μM nigericin (Sigma). Small molecules, including ATP, were removed from the translation products by Sephadex G-25 filtration (Pharmacia) (Olsen et al., 1989). MgATP was added to import reactions containing dark-adapted chloroplasts and ATP-depleted preprotein at concentrations (50-5000 μM), or was omitted completely. Import was carried out as described above.

4.3. Purification of Chloroplast Membranes by Fractionation

In all fractionation experiments, the chloroplasts were washed twice with HS (50 mM HEPES, 0.3 M sorbitol, pH 7.6) following import, to remove any residual non-associated proteins, and reisolated by centrifugation for 4 min at 3,000 rpm, at 4°C in a Sorvall HB-4 rota.

4.3.a. Total Membrane Preparation and Washes

Following washing and reisolation, the chloroplasts were burst by incubation in BB (10 mM HEPES, 10 mM $MgCl_2$, pH 8.0) for 10 min on ice. Total chloroplast membranes were washed with either BB or 1 M NaCl or 100 mM Na_2CO_3. After 50 min incubation on ice, total membrane fractions were collected by centrifugation at 100,000 *x g* for 1 hr and analysed by SDS-PAGE and immunoblotting.

4.3.b. Envelope Membrane Preparation

To purify envelope membranes burst chloroplasts (section 4.2.) were centrifuged slowly at 4,000 rpm for 8 min at 4°C in a Sorvall HB-4 rota. The supernatant was removed and the pellet resuspended in a further 350 μl BB and centrifuged in the same way, a total a three times. The three 350 μl supernatants were pooled and centrifuged one further time at 4,000 rpm for 8 min at 4°C in a Sorvall HB-4 rota. The remaining membranes (envelope only) were collected from the supernatant by centrifugation at 100,000 g for 1 h.

4.3.c. Thylakoid Membrane Preparation

Thylakoid membranes were prepared according to Rawyler et al, 1992, (Rawyler et al., 1992). Intact chloroplasts were resuspended in 2 ml HS buffer and pelleted through 10 ml 35% (v/v)

Percoll cushion at 2,500 *x g* for 8 min using a refrigerated Sorvall centrifuge. The chloroplasts were resuspended a second time in 2 ml HS and a further 30 ml BB. After incubation for 2 min on ice, 7.2 ml 1.5 M sucrose was added before centrifugation in a swing-out rota at 10,150 rpm for 5 min. The pellet was gently resuspended in 18 ml HS before being added to 8 ml 5% Percoll and further centrifuged at 11,300 rpm in a swing out rota for 15 min. The pellet was resuspended in 30 ml HS and pelleted again at 10,150 rpm for 5 min.The supernatant was poured away and the pellet was resuspended in the remaning solution.

4.4. Thermolysin Treatment

Import reactions were treated with thermolysin to digest any externally localized proteins, since thermolysin can not penetrate the envelope membranes (Cline et al., 1984). Two import reactions (300 µl) were stopped on ice (stop buffer was not added). The chloroplasts were isolated by centrifugation at 4,200 *x g* for 4 min and a refrigerated centrifuge. The chloroplasts were very gently resuspended in 300 µl HMS. At this point the two import reactions were divided equally, one was treated with thermolysin, the other was not (control). Thermolysin (100 µg/ml) and $CaCl_2$ (300 µM) were added to the chloroplasts and they were incubated on ice for 5-40 min in the dark. Digestion was stopped with an equal volume of Stop buffer, the samples were pelleted by centrifugation for ~5 s at 16,110 *x g* and the pellet was resuspended in 2 x sample buffer before analysis by SDS-PAGE (section 2.3.a.).

Appendices

Published Work

Kubis, S., Baldwin, A., Patel, R., Razzaq, A., Dupree, P., Lilley, K., Kurth, J., Leister, D., and Jarvis, P. (2003). The *Arabidopsis ppi1* mutant is specifically defective in the expression, chloroplast import, and accumulation of photosynthetic proteins. Plant Cell **15,** 1859-1871.

Division of Labour

Growth of plants, and isolation and quantification, of chloroplasts for proteomics (Figure 3), and RNA for microarray analysis (Figure 4a), and Northern blot analysis (Figure 4b) was carried out by Amy Baldwin. The remaining work was carried out by other authors.

Relevance to this Thesis

This work provides strong evidence that atToc33 specifically imports photosynthetic proteins (Kubis et al., 2003). However, that is not to say that atToc34 can not replace the role of atToc33, if overexpressed (Jarvis et al., 1998). On the other hand, atToc34 has recently been implicated to have a role in non-photosynthetic tissue (Constan et al., 2004). In the context of Results II, these articles suggest that there are multiple pathways into the chloroplast, and therefore, presumably regulated. Potential SP genes, such as the regulatory or inhibitory proteins described in the discussion of Results II, may be part of this scheme.

The Plant Cell, Vol. 15, 1859–1871, August 2003, www.plantcell.org © 2003 American Society of Plant Biologists

The Arabidopsis *ppi1* Mutant Is Specifically Defective in the Expression, Chloroplast Import, and Accumulation of Photosynthetic Proteins[W]

Sybille Kubis,[a] Amy Baldwin,[a] Ramesh Patel,[a] Azam Razzaq,[b] Paul Dupree,[b] Kathryn Lilley,[b] Joachim Kurth,[c] Dario Leister,[c] and Paul Jarvis[a,1]

[a] Department of Biology, University of Leicester, Leicester LE1 7RH, United Kingdom
[b] Cambridge Centre for Proteomics, University of Cambridge, Cambridge CB2 1QW, United Kingdom
[c] Department for Plant Breeding and Yield Physiology, Max-Planck-Institute for Plant Breeding Research, D-50829 Cologne, Germany

The import of nucleus-encoded proteins into chloroplasts is mediated by translocon complexes in the envelope membranes. A component of the translocon in the outer envelope membrane, Toc34, is encoded in Arabidopsis by two homologous genes, *atTOC33* and *atTOC34*. Whereas *atTOC34* displays relatively uniform expression throughout development, *atTOC33* is strongly upregulated in rapidly growing, photosynthetic tissues. To understand the reason for the existence of these two related genes, we characterized the *atTOC33* knockout mutant *ppi1*. Immunoblotting and proteomics revealed that components of the photosynthetic apparatus are deficient in *ppi1* chloroplasts and that nonphotosynthetic chloroplast proteins are unchanged or enriched slightly. Furthermore, DNA array analysis of 3292 transcripts revealed that photosynthetic genes are moderately, but specifically, downregulated in *ppi1*. Proteome differences in *ppi1* could be correlated with protein import rates: *ppi1* chloroplasts imported the ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit and 33-kD oxygen-evolving complex precursors at significantly reduced rates, but the import of a 50S ribosomal subunit precursor was largely unaffected. The *ppi1* import defect occurred at the level of preprotein binding, which is consistent with a role for atToc33 during preprotein recognition. The data suggest that atToc33 is involved preferentially in the import of photosynthetic proteins and, by extension, that atToc34 is involved in the import of nonphotosynthetic chloroplast proteins.

INTRODUCTION

Like mitochondria, chloroplasts are descended from a free-living prokaryotic organism that entered the eukaryotic lineage through endosymbiosis. During the course of their evolution, chloroplasts relinquished the majority of their genes so that now, >90% of chloroplast proteins are encoded in the nucleus (Leister, 2003). Chloroplast proteins are synthesized in precursor form on cytosolic ribosomes and targeted post-translationally to chloroplasts (Keegstra and Ciine, 1999). Each precursor protein (preprotein) carries an NH_2-terminal targeting signal called the transit peptide, which guides the protein to the chloroplast (Bruce, 2001). Translocon complexes in the outer and inner envelope membranes of chloroplasts (Toc and Tic, respectively) mediate preprotein import in an ATP- and GTP-dependent process (Chen et al., 2000; Hiltbrunner et al., 2001a; Jarvis and Soll, 2002). Proteins are translocated across both membranes simultaneously, in extended conformation, and then processed to their mature size by a stromal processing peptidase. After translocation, newly imported proteins are folded to their final conformation by molecular chaperones or targeted to one of a number of internal compartments via separate protein targeting pathways (Keegstra and Cline, 1999).

Several components of the Toc and Tic complexes have been identified in biochemical studies using isolated pea chloroplasts. The core components of the Toc complex are called Toc159, Toc34, and Toc75, according to their molecular weights (Hirsch et al., 1994; Kessler et al., 1994; Perry and Keegstra, 1994; Schnell et al., 1994; Seedorf et al., 1995; Bölter et al., 1998). Toc159 and Toc34 are related GTPases that have been proposed to act as preprotein receptors. Both proteins are anchored in the outer envelope by COOH-terminal domains and project their GTP binding domains into the cytosol. Although their precise mode of action remains unclear (Sveshnikova et al., 2000; Hiltbrunner et al., 2001b), Toc159 and Toc34 interact with incident preproteins (Ma et al., 1996; Kouranov and Schnell, 1997) and mediate their transfer to the translocation channel, of which Toc75 is a major component. Toc75 has a β-barrel structure comprising 16 amphiphilic β-strands and forms a channel with a pore size of ~14 to 26 Å (Hinnah et al., 2002). The Tic complex is less well characterized, and there is even some disagreement in the literature concerning the identity of its components (Jarvis and Soll, 2002). A 100-kD heat-shock protein (Hsp100) homolog, ClpC, held at the stromal face of the Tic complex has been proposed to drive chloroplast protein import, much like Hsp70 proteins drive protein translocation into mitochondria and the endoplasmic reticulum (Nielsen et al., 1997).

[1] To whom correspondence should be addressed. E-mail rpj3@le.ac.uk; fax 44-116-252-3330.
[W] Online version contains Web-only data.
Article, publication date, and citation information can be found at www.plantcell.org/cgi/doi/10.1105/tpc.012955.

In Arabidopsis, several of these translocon components are encoded by multiple genes. For example, there are two Toc34-related genes in Arabidopsis, *atTOC33* and *atTOC34* (Jarvis et al., 1998; Gutensohn et al., 2000), and four Toc159-related genes, *atTOC159*, *atTOC132*, *atTOC120*, and *atTOC90* (Bauer et al., 2000; Hiltbrunner et al., 2001a). There also are multiple Toc75-related sequences in Arabidopsis, but some other components are encoded by just a single gene (Jackson-Constan and Keegstra, 2001). The reason for this Toc component multiplicity in Arabidopsis is unclear, but it has been proposed that different Toc isoforms have different preprotein recognition specificities (Jarvis et al., 1998). Support for this hypothesis was obtained recently upon characterization of an *atTOC159* knockout mutant called *plastid protein import2* (*ppi2*) (Bauer et al., 2000). Homozygous *ppi2* plants have a seedling-lethal, albino phenotype characterized by pronounced defects in chloroplast biogenesis. Interestingly, whereas the expression and accumulation of proteins involved directly in photosynthesis (photosynthetic proteins) are reduced strongly in *ppi2*, proteins involved in chloroplast functions unrelated to photosynthesis (nonphotosynthetic proteins) are expressed and accumulate normally. These observations led to the suggestion that atToc159 has recognition specificity for photosynthetic proteins and that atToc132 and atToc120 are involved preferentially in the import of nonphotosynthetic proteins (Bauer et al., 2000). The existence of such substrate-specific protein import pathways would prevent the bulk flow of highly abundant, photosynthetic proteins from outcompeting the import of less abundant but equally important nonphotosynthetic proteins.

Previously, we described the identification and preliminary characterization of an *atTOC33* knockout mutant called *ppi1* (Jarvis et al., 1998). Homozygous *ppi1* plants are yellow-green in appearance but, unlike *ppi2* plants, they are able to survive to maturity. This difference in phenotype severity may indicate that any specialization among the Toc34 homologs is less pronounced than what exists among the Toc159 homologs and that there is considerable functional redundancy between atToc33 and atToc34. Indeed, overexpression of atToc34 was shown to complement the atToc33 deficiency of *ppi1* plants (Jarvis et al., 1998). Nevertheless, the possibility remains that under normal, physiological conditions, atToc33 and atToc34 operate preferentially in import pathways with different substrate specificities. Some support for this idea was provided recently by the demonstration that, in vitro, atToc33 but not atToc34 is able to inhibit the binding of ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) small subunit precursor (a photosynthetic preprotein) to chloroplasts in a GTP-dependent manner (Gutensohn et al., 2000). In this study, we took an in vivo approach to further address this question by characterizing the *ppi1* mutant in detail.

RESULTS

atTOC33 Expression Is Induced Strongly in Rapidly Growing Photosynthetic Tissues

We previously determined the expression levels of *atTOC33* and *atTOC34* over a developmental time course at the whole-plant level (Jarvis et al., 1998). We found that *atTOC33* is expressed at levels several times greater than *atTOC34* and that both genes are subject to developmental regulation, with the highest levels of expression occurring in young plants. These data demonstrated that both genes are expressed during periods of intense plastid biogenesis, but they provided no clues regarding the possible functional differences between the proteins, because gene expression levels in individual tissues were not investigated. Although subsequent experiments did provide spatial expression profiles for each gene (Gutensohn et al., 2000), these studies did not address the relative levels of expression of the two genes.

To gain insight into the reasons underlying the existence of two Toc34-related genes in Arabidopsis, we determined the relative expression levels of *atTOC33* and *atTOC34* at different developmental stages and in different tissue types by RNA gel blot analysis (Figure 1). Whereas *atTOC34* was expressed at a relatively uniform, low level at most stages of development, *atTOC33* was induced strongly in photosynthetic tissues undergoing rapid growth (in 5- and 10-day-old light-grown plants and in inflorescence tips). In fully expanded photosynthetic tissues (rosette leaves), the expression of both genes was reduced significantly, although *atTOC33* continued to be expressed at a higher level than *atTOC34*. Most interestingly, in nonphotosynthetic root tissue, *atTOC33* was downregulated specifically such that the two genes were expressed at essentially the same level. Both genes were expressed at low levels in dark-grown plants, which presumably reflects the minimal requirements for plastid biogenesis in these plants. The 10-day-old *ppi1* control indicated that the hybridization probes were specific and confirmed that *atTOC34* expression is not upregulated strongly in the *ppi1* mutant.

Photosynthetic Proteins Are Specifically Deficient in *ppi1* Chloroplasts

The demonstration that *atTOC33* expression was induced strongly in rapidly expanding photosynthetic tissues suggests that atToc33 may be relatively more important for the assembly of the photosynthetic apparatus than atToc34. To investigate this possibility, we used immunoblot analysis to compare the abundance of a range of different photosynthetic and nonphotosynthetic proteins in chloroplasts isolated from 10-day-old wild-type and *ppi1* plants (Figure 2). In this analysis, antibodies against four photosynthetic proteins (Rubisco small subunit [SSU], light-harvesting chlorophyll *a/b* binding protein [LHCII], photosystem I subunit D [PSI-D], and ferredoxin-$NADP^+$ oxidoreductase [FNR]) and four nonphotosynthetic chloroplast proteins (Hsp70, the Hsp100 homolog ClpC, triose phosphate/phosphate translocator [TPT], and coproporphyrinogen oxidase [CPO]) were used. To obtain semiquantitative data, a dilution series of each chloroplast extract was examined with each antibody. By comparing the band intensities obtained, we deduced that each of the photosynthetic proteins was reduced in abundance in *ppi1* chloroplasts by ~50% or more. No evidence for preprotein accumulation in *ppi1* was observed, suggesting that most nonimported preproteins are degraded rapidly or, alternatively, that gene expression is correlated closely

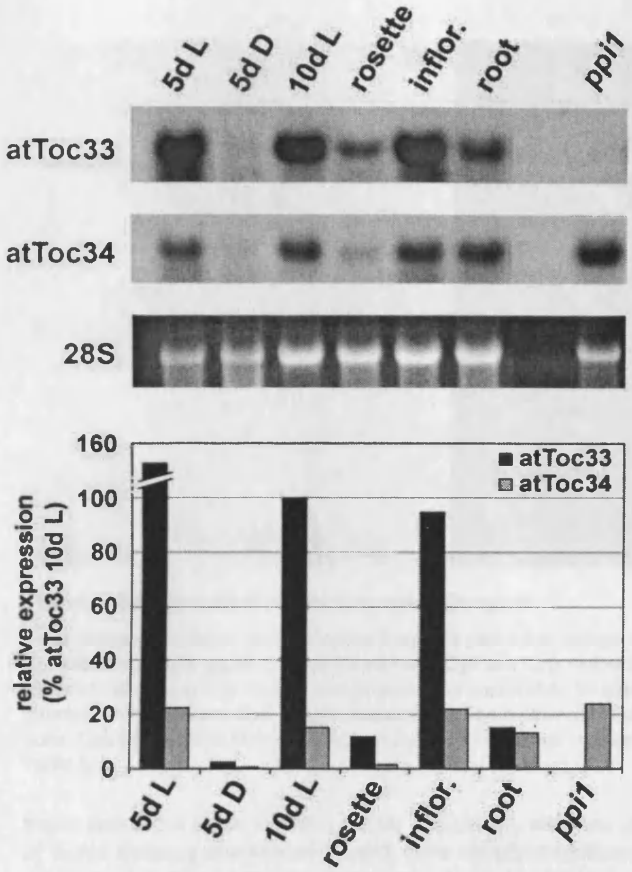


Figure 1. Expression Levels of *atTOC33* and *atTOC34* during Seedling Development and in Different Tissues of Arabidopsis.

Fifteen-microgram samples of total RNA isolated from Arabidopsis tissues were examined by RNA gel blot analysis. RNA was isolated from wild-type seedlings grown in vitro for 5 days in the light (5d L), 5 days in the dark (5d D), and 10 days in the light (10d L), from three different tissues of 28-day-old wild-type plants grown on soil (rosette leaves, young inflorescence tips, and roots), and from *ppi1* seedlings grown in vitro for 10 days in the light (*ppi1*). Filters were probed using ^{32}P-labeled *atTOC33* and *atTOC34* cDNA probes with identical specific activities. rRNA (28S) was used as a loading control. The images shown were obtained after a 3-week exposure, but quantification was performed using a shorter exposure. Relative levels of expression of *atTOC33* and *atTOC34* normalized for 28S rRNA are shown in the graph at bottom; *y* axis values between 100 and 140 have been removed to aid visualization.

with the efficiency of import of the encoded preprotein. By contrast, all four nonphotosynthetic proteins were present at least at wild-type levels in *ppi1* chloroplasts, and some even appeared to be enriched slightly in the mutant.

To corroborate these findings, we next adopted a proteomics approach. The chloroplast proteomes of 10-day-old wild-type and *ppi1* plants were analyzed using difference gel electrophoresis (DIGE). Proteins extracted from isolated chloroplasts were labeled with nonsaturating CyDyeDIGE fluors. In a typical experiment, wild-type proteins were labeled with the Cy5 fluor and mutant proteins were labeled with Cy3. The labeled protein samples were mixed and then resolved in the first dimension by isoelectric focusing and in the second dimension by SDS-PAGE. The Cy5 and Cy3 fluorescent images of a typical gel are shown in Figure 3 (left and right, respectively). Additionally, an overlay of the Cy5 (wild type; false-colored red) and Cy3 (*ppi1*; false-colored green) images has been included to aid the visualization of protein abundance differences (Figure 3, middle): proteins depleted in *ppi1* appear red, whereas those enriched in *ppi1* appear green. The fluorescent signal associated with each protein spot was quantified and used to calculate the degree of enrichment or depletion in *ppi1* chloroplasts.

Based on these data, spots were selected for identification by mass spectrometry, and the results obtained are summarized in Table 1. All of the proteins we identified that were depleted in *ppi1* chloroplasts were components of the photosyn-

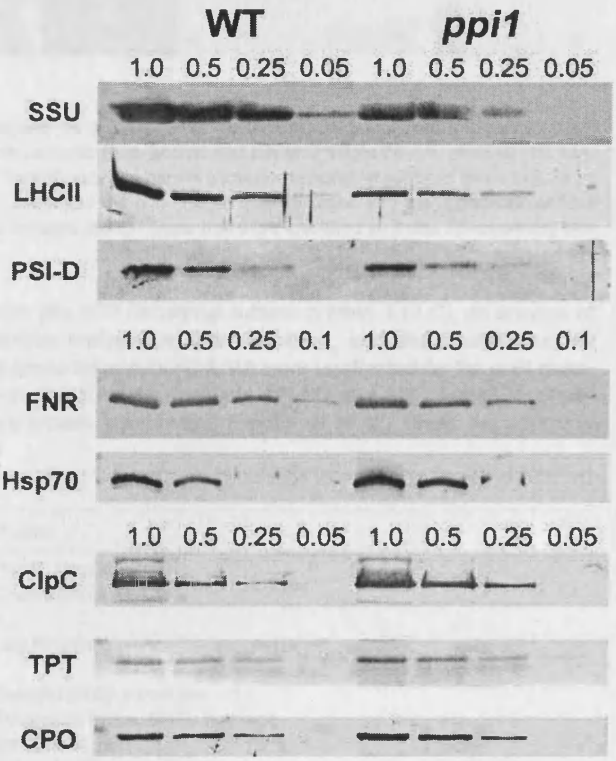


Figure 2. Immunoblot Analysis of Photosynthetic and Nonphotosynthetic Chloroplast Proteins in the *ppi1* Mutant.

Dilution series of chloroplast protein preparations from the wild type (WT) and the *ppi1* mutant (*ppi1*) were separated by SDS-PAGE and transferred to nitrocellulose membranes. Membranes were probed using antibodies against photosynthetic (SSU, LHCII, PSI-D, and FNR) and nonphotosynthetic (Hsp70, ClpC, TPT, and CPO) proteins. In most cases, filters were cut in half and probed for one photosynthetic protein and one nonphotosynthetic protein (the pairings used were SSU/TPT, PSI-D/ClpC, and FNR/Hsp70); the LHCII and CPO data were derived using separate blots. The dilution steps used are given above the gels. Dilution series started with 20 μg for SSU, PSI-D, ClpC, TPT, and CPO, with 10 μg for FNR and Hsp70, and with 2 μg for LHCII. The data shown are representative of three independent experiments.

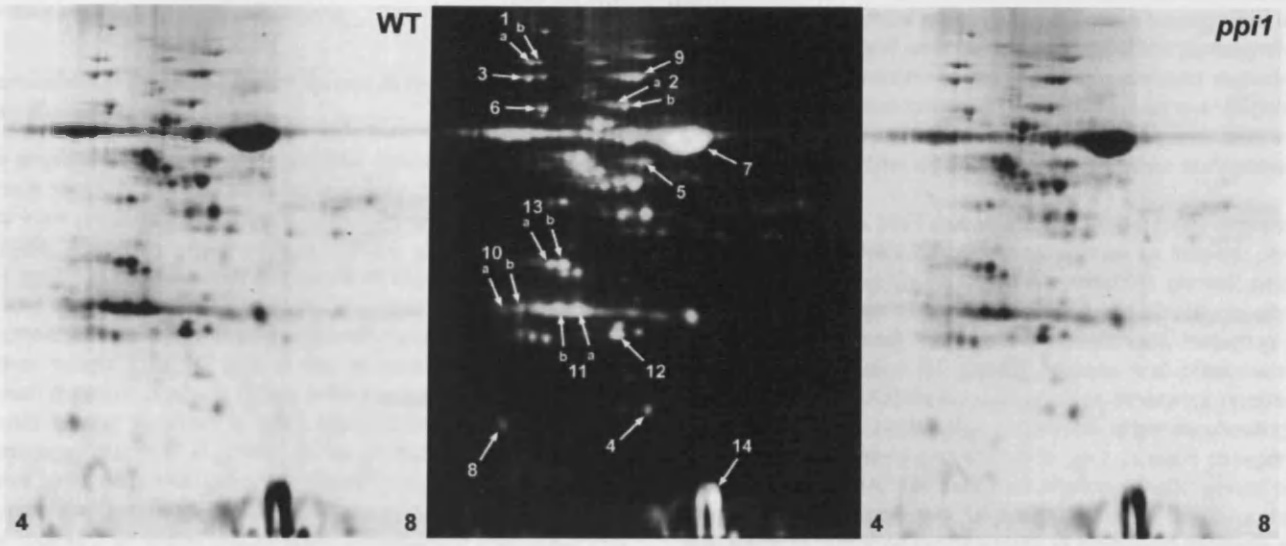


Figure 3. DIGE Analysis of the *ppi1* Chloroplast Proteome.

Total chloroplast protein samples isolated from 10-day-old seedlings were analyzed using CyDyeDIGE technology. Wild-type (WT) and mutant (*ppi1*) protein samples (50 μg each) were labeled with Cy5 and Cy3, respectively. The samples were pooled and run on a single 24-cm, pH-3 to -10, two-dimensional gel and then imaged using parameters appropriate for each fluor. The left and right panels show corresponding sections (pH 4 to 8, as indicated) of the Cy5 and Cy3 images, respectively. The middle panel shows an overlay of the two images after inversion and the application of false color: Cy5 (wild type) is shown in red, and Cy3 (*ppi1*) is shown in green. Labels indicate protein spots that were identified by mass spectrometry (see Table 1).

thetic apparatus (such as SSU, OE33, and LHCII), whereas all of those showing enrichment in *ppi1* were nonphotosynthetic proteins, including molecular chaperones and a component of the plastid genetic system (an EF-Tu–like translation elongation factor). Steady state levels of another genetic system component (the 50S ribosomal subunit protein, L12-C), an enzyme of carbon metabolism (transketolase), and the plastid-encoded Rubisco large subunit (LSU) were unaffected by the *ppi1* mutation. Because the amounts of LSU and SSU protein normally are closely coordinated (Rodermel et al., 1996), we conclude

Table 1. Proteins Identified by DIGE and Mass Spectrometry

Spot No.	Gene No.	Fold Change in *ppi1*[a]	Protein
1	At2g04030	a, 2.41 ± 0.61; b, 1.85 ± 0.09	Hsp90 homolog
2	At1g55490	a, 1.86 ± 0.13; b, 1.64 ± 0.05	Chaperonin 60β
	At3g13470		
3	At4g24280	1.78 ± 0.07	Hsp70 homolog
	At5g49910		
4	At3g62030	1.76 ± 0.29	Peptidyl prolyl isomerase
5	At4g20360	1.72 ± 0.08	Elongation factor, EF-Tu homolog
6	At2g28000	1.64 ± 0.19	Chaperonin 60α
7	RbcL	—	Rubisco large subunit, LSU
8	At3g27850	—	Ribosomal subunit, L12-C
9	At3g60750	—	Transketolase-like protein
10	At3g04790	a, 0.63 ± 0.04; b, 0.56 ± 0.03	Ribulose-5-phosphate isomerase
11a	At4g10340	0.61 ± 0.05	Light-harvesting chlorophyll *a/b* binding protein CP26 (*Lhcb5*)
11b	At1g29910	0.61 ± 0.05	Light-harvesting chlorophyll *a/b* binding protein
	At1g29920		LHCII type 1 (*Lhcb1.1*, *Lhcb1.2*, *Lhcb1.3*, *Lhcb1.5*)
	At1g29930		
	At2g34420		
12	At1g06680	0.58 ± 0.03	23-kD subunit of oxygen-evolving complex, OE23
13	At3g50820	a, 0.56 ± 0.09; b, 0.50 ± 0.06	33-kD subunit of oxygen-evolving complex, OE33
	At5g66570		
14	At1g67090	0.43 ± 0.04	Rubisco small subunit, SSU (*Ats1A*)

[a] Prefixes "a" and "b" denote protein spots that could not be distinguished upon mass spectrometric analysis. Values are means ± SD.

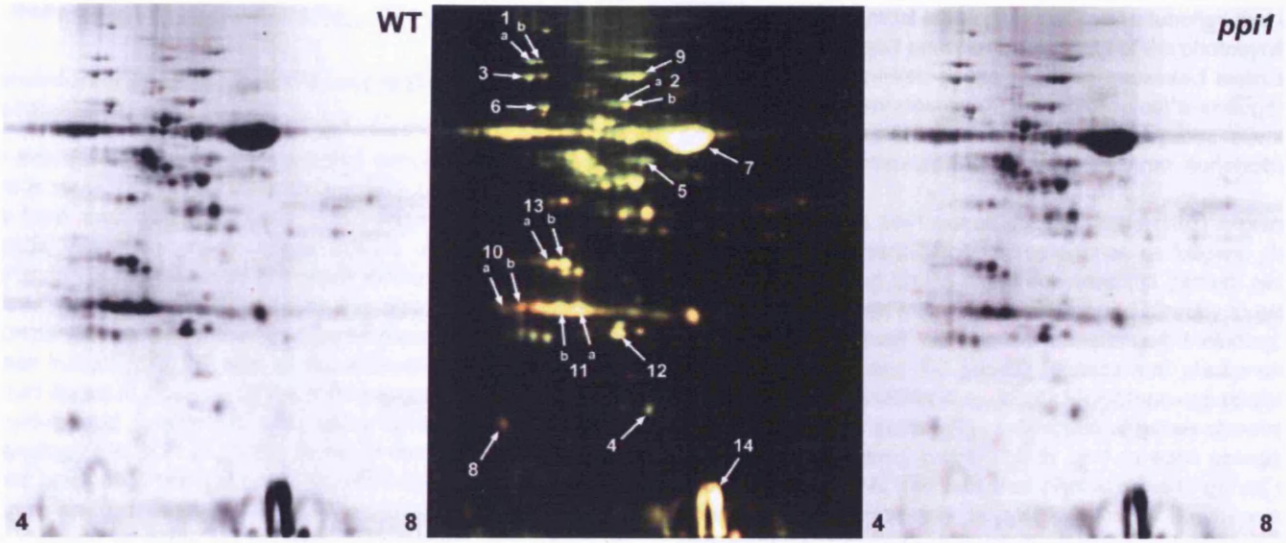


Figure 3. DIGE Analysis of the *ppi1* Chloroplast Proteome.

Total chloroplast protein samples isolated from 10-day-old seedlings were analyzed using CyDyeDIGE technology. Wild-type (WT) and mutant (*ppi1*) protein samples (50 μg each) were labeled with Cy5 and Cy3, respectively. The samples were pooled and run on a single 24-cm, pH-3 to -10, two-dimensional gel and then imaged using parameters appropriate for each fluor. The left and right panels show corresponding sections (pH 4 to 8, as indicated) of the Cy5 and Cy3 images, respectively. The middle panel shows an overlay of the two images after inversion and the application of false color: Cy5 (wild type) is shown in red, and Cy3 (*ppi1*) is shown in green. Labels indicate protein spots that were identified by mass spectrometry (see Table 1).

thetic apparatus (such as SSU, OE33, and LHCII), whereas all of those showing enrichment in *ppi1* were nonphotosynthetic proteins, including molecular chaperones and a component of the plastid genetic system (an EF-Tu–like translation elongation factor). Steady state levels of another genetic system component (the 50S ribosomal subunit protein, L12-C), an enzyme of carbon metabolism (transketolase), and the plastid-encoded Rubisco large subunit (LSU) were unaffected by the *ppi1* mutation. Because the amounts of LSU and SSU protein normally are closely coordinated (Rodermel et al., 1996), we conclude

Table 1. Proteins Identified by DIGE and Mass Spectrometry

Spot No.	Gene No.	Fold Change in *ppi1*[a]	Protein
1	At2g04030	a, 2.41 ± 0.61; b, 1.85 ± 0.09	Hsp90 homolog
2	At1g55490	a, 1.86 ± 0.13; b, 1.64 ± 0.05	Chaperonin 60β
	At3g13470		
3	At4g24280	1.78 ± 0.07	Hsp70 homolog
	At5g49910		
4	At3g62030	1.76 ± 0.29	Peptidyl prolyl isomerase
5	At4g20360	1.72 ± 0.08	Elongation factor, EF-Tu homolog
6	At2g28000	1.64 ± 0.19	Chaperonin 60α
7	RbcL	—	Rubisco large subunit, LSU
8	At3g27850	—	Ribosomal subunit, L12-C
9	At3g60750	—	Transketolase-like protein
10	At3g04790	a, 0.63 ± 0.04; b, 0.56 ± 0.03	Ribulose-5-phosphate isomerase
11a	At4g10340	0.61 ± 0.05	Light-harvesting chlorophyll *a/b* binding protein CP26 (*Lhcb5*)
11b	At1g29910	0.61 ± 0.05	Light-harvesting chlorophyll *a/b* binding protein
	At1g29920		LHCII type 1 (*Lhcb1.1*, *Lhcb1.2*, *Lhcb1.3*, *Lhcb1.5*)
	At1g29930		
	At2g34420		
12	At1g06680	0.58 ± 0.03	23-kD subunit of oxygen-evolving complex, OE23
13	At3g50820	a, 0.56 ± 0.09; b, 0.50 ± 0.06	33-kD subunit of oxygen-evolving complex, OE33
	At5g66570		
14	At1g67090	0.43 ± 0.04	Rubisco small subunit, SSU (*Ats1A*)

[a] Prefixes "a" and "b" denote protein spots that could not be distinguished upon mass spectrometric analysis. Values are means ± SD.

that the moderate SSU deficiency in *ppi1* is insufficient to trigger these coregulatory mechanisms.

Expression of Photosynthetic Genes Is Downregulated Specifically in *ppi1*

The expression of nuclear genes that encode chloroplast proteins is regulated in response to a variety of signals that emanate from plastids (Surpin et al., 2002; Jarvis, 2003). These retrograde signaling mechanisms ensure that nuclear gene expression is coupled with the needs of the cell's chloroplasts, so that the photosynthetic apparatus and other biosynthetic machineries are assembled in stoichiometric fashion. To gain further insight into the role of the atToc33 protein, we conducted a global analysis of the mRNA expression of genes of plastid-related function in *ppi1* using nylon filter DNA array technology (Kurth et al., 2002; Richly et al., 2003). Of the 3292 genes analyzed, 1461 showed significantly different levels of expression from the wild type (see supplemental data online). The majority of these genes were upregulated moderately in *ppi1*, with only 161 genes showing reduced levels of expression. Remarkably, most of the genes of known function that were downregulated in *ppi1* encode components of the photosynthetic apparatus. By contrast, genes showing increased expression levels encode components of the chloroplast's endogenous genetic system, factors involved in metabolic processes unrelated to photosynthesis, and a range of other nonphotosynthetic proteins.

We assigned all 1461 genes showing differential expression in *ppi1* to six different functional categories as follows: photosynthesis, including Calvin cycle enzymes (55 genes); genetic system (152 genes); metabolism not related directly to photosynthesis (170 genes); transport of substances, including metabolites and proteins (60 genes); kinases and phosphatases (55 genes); and putative, hypothetical, or otherwise unclassifiable proteins (969 genes). The proportion of genes showing increased or decreased expression in *ppi1* in each category is shown in Figure 4A. The fact that photosynthetic genes were downregulated specifically in *ppi1* reinforces the hypothesis that atToc33 operates preferentially in an import pathway with

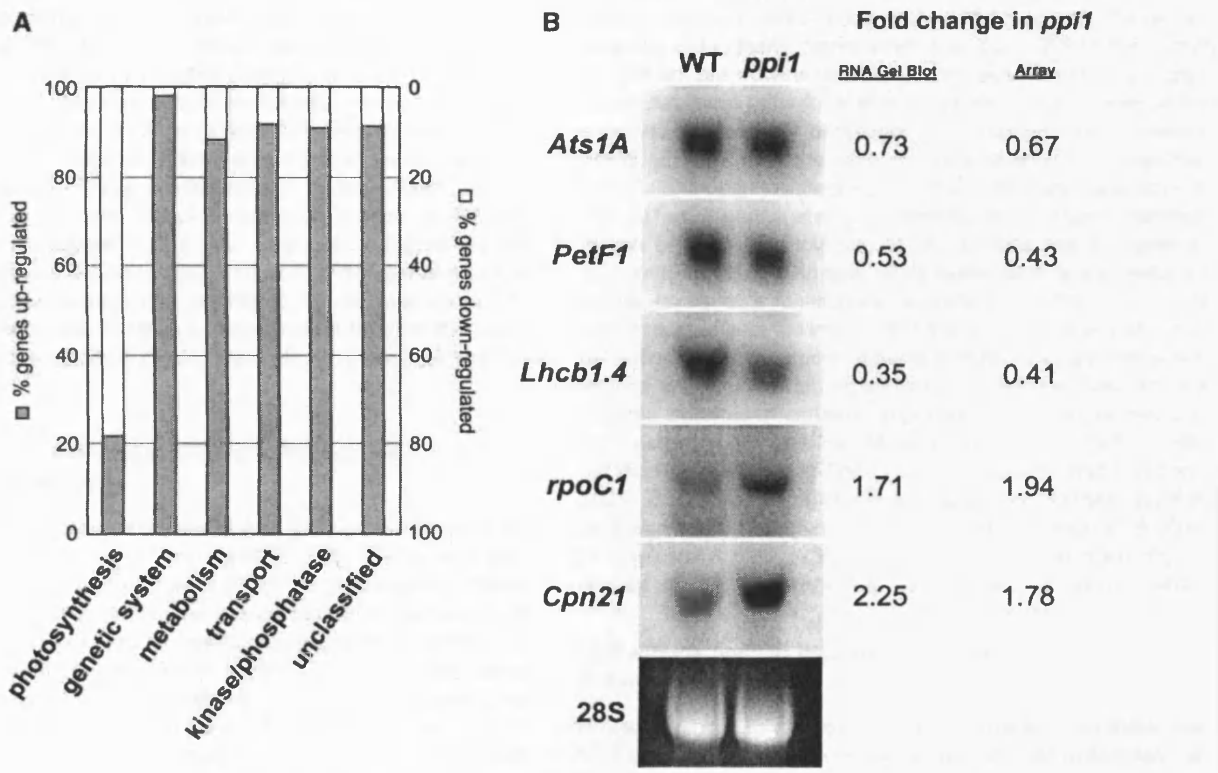


Figure 4. Analysis of the Expression of Nucleus-Encoded Chloroplast Genes in *ppi1*.

(A) Total RNA isolated from 10-day-old wild-type and *ppi1* mutant seedlings was used to probe a nylon filter DNA array carrying 3292 gene-specific tags. All 1461 genes showing differential expression in *ppi1* were placed into six different categories, as described in the text. The proportion of genes within each category showing upregulation (gray bars) and downregulation (white bars) in *ppi1* is shown.

(B) Ten-microgram samples of the RNA described in **(A)** were analyzed further by RNA gel blot analysis. Five of the differentially regulated genes listed in Table 2 were selected for analysis. Filters were hybridized with ^{32}P-labeled probes corresponding to the gene-specific tags described in **(A)**. rRNA (28S) was used as a loading control. Radioactivity associated with each band shown was quantified and used to calculate the fold change in expression in *ppi1* (relative to the wild type [WT]) for each gene; the corresponding data derived from the DNA array experiment (**[A]**; Table 2) are shown for comparison.

specificity for photosynthetic proteins. It is conceivable that the nucleus responds to the perturbation of photosynthetic protein import, communicated via some retrograde signaling pathway, by downregulating the genes that encode the affected proteins.

To further illustrate these gene expression changes, the data for 42 genes representative of four different functional categories (photosynthesis, 12 genes; genetic system, 10 genes; nonphotosynthetic metabolism, 10 genes; and protein targeting, 10 genes) as well as 8 genes belonging to other categories are shown in Table 2; the relative levels of expression of these 50 genes are illustrated further in the supplemental data online. Not only did most of the genes that show downregulated expression in *ppi1* encode components of the photosynthetic apparatus, they also tended to be expressed at a higher level than genes that show upregulated expression (Table 2; see also supplemental data online). The plastome-encoded Rubisco large subunit gene, *rbcL*, is one obvious exception. Also of interest is the observation that all of the protein targeting–related genes were upregulated in *ppi1*, except for the gene that encodes Tic20-I, which is a putative channel component of the inner envelope membrane (Jackson-Constan and Keegstra, 2001). Similarly, genes that encode several different molecular chaperones, some of which have been implicated in chloroplast protein import (Schnell et al., 1994; Nielsen et al., 1997; Jackson-Constan and Keegstra, 2001), also were upregulated in *ppi1* (Table 2; see also supplemental data online).

To corroborate the findings of this DNA array experiment, we examined the expression of five genes listed in Table 2 by RNA gel blot analysis. Three photosynthetic genes (*Ats1A*, *PetF1*, and *Lhcb1.4*) and two nonphotosynthetic genes (*rpoC1* and *Cpn21*) were analyzed in this way. As shown in Figure 4B, the gene expression changes observed using RNA gel blot analysis were quantitatively very similar to those observed using DNA array analysis. Therefore, these results confirm the validity of the gene expression data presented in Figure 4A and Table 2.

ppi1 Affects the Import of Different Preproteins to Different Degrees

Isolated *ppi1* chloroplasts were shown previously to import reduced quantities of in vitro–translated SSU, LHCII, and protochlorophyllide oxidoreductase (POR; a light-sensitive chlorophyll biosynthetic enzyme) over a fixed period of time (Jarvis et al., 1998). However, import rate comparisons for different preproteins were not conducted. To determine if the primary cause of the chloroplast proteome differences observed in *ppi1* (Figures 2 and 3, Table 1) could be a differential effect of the *ppi1* mutation on the import of different preproteins, we compared import rates for three different preproteins using isolated wild-type and mutant chloroplasts. Import rates were determined by measuring the accumulation of imported protein over time. For these experiments, we selected SSU (a component of the photosynthetic dark reactions that occur in the stroma), OE33 (a component of the photosynthetic light reactions that occur in the thylakoids), and L11 (a 50S ribosomal subunit protein and component of the chloroplast's endogenous genetic system) for analysis.

Initial experiments revealed that import rates were linear up to ~10 min for SSU preprotein (preSSU) and preOE33 and up to ~5 min for preL11 (data not shown); therefore, import rate comparisons were conducted within these periods. Resistance to thermolysin treatment confirmed that the processed forms of all three preproteins had been internalized by intact chloroplasts (see supplemental data online). As shown in Figures 5A and 5B, both preSSU and preOE33 were imported into *ppi1* chloroplasts at significantly reduced rates: on average, preSSU import rates were reduced by 42.4 ± 1.9% and preOE33 import rates were reduced by 79.6 ± 7.6%. OE33 is targeted to the thylakoid lumen via a stromal intermediate (Keegstra and Cline, 1999) that could be observed between the precursor and mature OE33 bands of Figure 5B. Because this intermediate had already traversed the envelope, it was regarded as imported protein for quantification purposes. By contrast, the effect of *ppi1* on the import of preL11 was slight, and import rates in the different genotypes were almost identical (Figure 5C).

Although OE33 import was affected more strongly than SSU import in vitro (Figure 5), SSU accumulation was affected more strongly than OE33 accumulation in vivo (Figure 3, Table 1). This finding indicates that factors in addition to import efficiency influence steady state protein levels in *ppi1*. Nevertheless, the data clearly demonstrate that *ppi1* affects the import of different preproteins to different degrees; therefore, we conclude that these differential effects are the primary cause of the selective depletion of photosynthetic proteins in *ppi1* chloroplasts. Exactly how the *ppi1* mutation exerts this preprotein specificity is open to interpretation. The most likely explanation is that atToc33 participates preferentially in an import pathway (which perhaps also involves atToc159) that has a degree of substrate specificity. An alternative explanation is that reduced overall levels of a functionally redundant atToc33 + atToc34 pool in *ppi1* affects the import of those preproteins with the lowest affinity for the import apparatus most strongly. However, such a model does not explain why Arabidopsis has retained multiple isoforms of certain components of the import apparatus, such as Toc34 and Toc159, and only single isoforms of others, such as Tic110 and Tic40 (Jackson-Constan and Keegstra, 2001). Furthermore, this model is inconsistent with the fact that *ppi1* significantly affected the import of preSSU (Figure 5A). One of the reasons that preSSU has been used in so many import studies is that it is imported into chloroplasts with high efficiency.

ppi1 Affects preSSU Import at the Level of Preprotein Binding

Protein import into chloroplasts can be divided into three distinct stages in vitro based on energetic requirements (Olsen et al., 1989; Olsen and Keegstra, 1992; Ma et al., 1996; Kouranov and Schnell, 1997). These stages are assumed to correspond to sequential steps in the import process that occurs in vivo. The first stage of import is energy-independent binding, and this occurs in the absence of ATP. It involves the reversible interaction of preproteins with receptor components of the Toc complex. The second stage of import is early import intermediate formation, which requires low concentrations of ATP (<100 μM) in the intermembrane space. Preproteins at this stage are

Table 2. Selected Genes Showing Differential Expression in *ppi1*

Category/Gene No.	Signal Intensity in the Wild Type	Fold Change in *ppi1*	Gene Name/Product	Function
Photosynthesis				
rbcL	68.84	1.26	Plastid gene *rbcL*, Rubisco large subunit	Calvin cycle
At5g38430	46.48	0.73	*Ats1B*, Rubisco small subunit	Calvin cycle
At4g09650	1.43	0.72	*AtpD*, ATP synthase δ-subunit	ATP synthesis
At1g67740	2.55	0.71	*PsbY*, Y-subunit of PSII	Photosystem II
At5g61410	1.55	0.70	Ribulose-5-phosphate-3-epimerase	Calvin cycle
At1g67090	35.23	0.67	*Ats1A*, Rubisco small subunit *a*	Calvin cycle
At1g76100	0.87	0.65	*PetE1*, plastocyanin	Electron transport
At4g02770	6.58	0.63	*PsaD1*, D-subunit of PSI	Photosystem I
At3g47470	8.17	0.60	*Lhca4*, LHCI polypeptide	Light harvesting
At1g56190	1.16	0.43	Phosphoglycerate kinase	Calvin cycle
At1g60950	2.75	0.43	*PetF1*, ferredoxin	Electron transport
At2g34430	17.01	0.41	*Lhcb1.4*, LHCII polypeptide	Light harvesting
Genetic system				
At1g34380	0.20	2.34	DNA polymerase type I, putative	DNA replication
rpoC1	1.19	1.94	Plastid gene *rpoC1*, RNA polymerase	Transcription
At5g26710	0.91	1.45	Glutamyl-tRNA synthetase	Translation
At3g53460	0.78	1.43	RNA binding protein, CP29	RNA stability
At3g62910	0.47	1.38	Translation releasing factor, RF-1–like	Translation
At4g20980	0.49	1.30	Translation initiation factor, eIF3d-like	Translation
At5g24120	0.13	1.30	σ-like factor	Transcription
At1g32990	0.91	1.30	50S plastid ribosomal protein, L11	Translation
At3g27850	0.85	1.27	50S plastid ribosomal protein, L12-C	Translation
At1g74980	0.23	1.24	30S plastid ribosomal protein, S9	Translation
Metabolism				
At1g69370	0.38	1.55	Chorismate mutase	Amino acid biosynthesis
At3g48560	0.50	1.44	Acetolactate synthase	Amino acid biosynthesis
At4g18240	0.20	1.40	Starch synthase–like protein	Starch biosynthesis
At1g29900	0.78	1.37	Carbamoyl phosphate synthetase, carb	Pyrimidine biosynthesis
At5g57030	0.53	1.34	Lycopene ε-cyclase	Carotenoid biosynthesis
At3g02610	0.42	1.34	Stearoyl–acyl carrier protein desaturase	Fatty acid biosynthesis
At4g25080	0.67	1.31	Mg-protoporphyrin IX methyltransferase	Tetrapyrrole biosynthesis
At2g26670	0.17	1.31	Heme oxygenase, HO1	Tetrapyrrole biosynthesis
At1g09830	1.46	1.28	Phosphoribosylglycineamide synthetase	Purine biosynthesis
At3g11670	0.85	1.26	Digalactosyldiacylglycerol synthase	Galactolipid biosynthesis
Protein targeting				
At5g28750	0.19	1.85	Tha4	Tat pathway, thylakoids
At4g33350	13.64	1.77	Tic22-IV	Protein import
At5g16620	0.23	1.63	Tic40	Protein import
At2g18710	0.33	1.48	SecY	Sec pathway, thylakoids
At5g03940	0.25	1.46	SRP54	SRP pathway, thylakoids
At4g02510	0.68	1.38	Toc159	Protein import
At4g03320	0.43	1.31	Tic20-IV	Protein import
At1g06950	0.30	1.31	Tic110	Protein import
At3g46740	0.78	1.25	Toc75-III	Protein import
At1g04940	2.96	0.67	Tic20-I	Protein import
Other				
At5g20720	0.49	1.78	Cpn21 chaperonin	Molecular chaperone
At4g39960	0.70	1.54	DnaJ homolog	Molecular chaperone
At3g13470	0.59	1.53	Chaperonin cpn60β	Molecular chaperone
At2g28900	0.13	1.48	OEP16 homolog	Outer envelope porin
At4g24280	0.69	1.34	Hsp70 homolog	Molecular chaperone
At5g50920	1.14	1.30	Hsp100 homolog, ClpC-V	Molecular chaperone
At5g24020	0.14	1.28	MinD homolog	Chloroplast division
At5g55280	0.40	1.26	FtsZ homolog	Chloroplast division

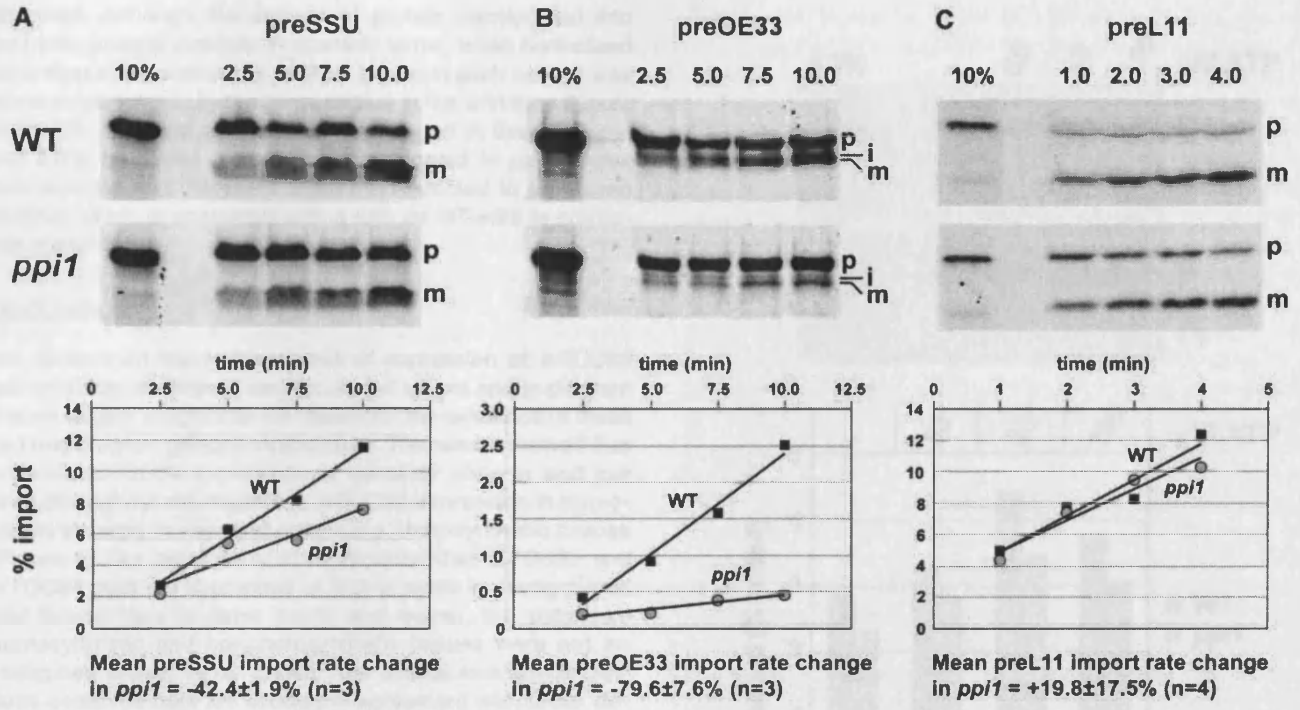


Figure 5. Comparison of Chloroplast Import Rates in the Wild Type and *ppi1* for Different Preproteins.

In vitro–translated, ^{35}S-Met–labeled preSSU **(A)**, preOE33 **(B)**, and preL11 **(C)** were imported into wild-type (WT) and mutant (*ppi1*) chloroplasts for 1.0 to 10.0 min, as indicated. Ten percent of the translation product added to each import reaction was loaded as a control (10%). Precursor (p), intermediate (i; OE33 only), and mature (m) protein forms are indicated at right of the gels. Quantifications for each import experiment are shown in the corresponding graph below each gel. Radioactivity associated with each mature band (intermediate and mature bands together in the case of OE33) was quantified and expressed as a percentage of the total preprotein-associated radioactivity added to each reaction. The data presented in the graphs correspond to the results shown in the gels, which are representative of several independent experiments. Mean import rate changes (±SE) derived from repeated experiments are given at bottom below the graphs.

inserted across the outer envelope membrane and in contact with components of the Tic complex; therefore, progression to this stage of import can be viewed as outer envelope translocation. The final stage of import is complete translocation, which requires high concentrations of ATP (>100 μM) in the stroma. Progression to this stage requires both binding and outer envelope translocation as well as inner envelope translocation and processing by the stromal processing peptidase.

To determine the stage of import affected by *ppi1*, we quantified the amount of preSSU bound to ATP-depleted chloroplasts in import reactions containing different, limiting concentrations of ATP (Figure 6A). All binding reactions were performed in the dark to prevent ATP synthesis, but control import reactions were performed in the light to confirm that the import competence of the chloroplasts was high (data not shown). In the absence of ATP, binding of preSSU to chloroplasts was reduced by ~20% in *ppi1* compared with the wild type (Figure 6A). This finding indicates that *ppi1* affects import at the level of energy-independent binding. Increased preprotein binding as ATP is increased to micromolar concentrations is indicative of early import intermediate formation. Because the reduction in binding observed in *ppi1* remained constant as ATP concentrations were increased up to 100 μM (on average, preSSU binding under ATP-limiting conditions was 22.0 ± 2.3% lower for *ppi1* than for the wild type), it seems unlikely that *ppi1* also affects outer envelope translocation (Figure 6A). The slight decline in the amount of bound preSSU at 100 μM ATP (seen in both genotypes) has been observed previously (Olsen et al., 1989) and presumably reflects the fact that a proportion of the preprotein has been translocated.

The efficiency of inner envelope translocation cannot be measured simply by quantifying the amount of protein imported in reactions containing high ATP concentrations (>100 μM), because this depends not only on the efficiency of inner envelope translocation but also on the efficiency of energy-independent binding and outer envelope translocation. To separate binding and outer envelope translocation from translocation across the inner envelope membrane, we first bound preSSU to wild-type and *ppi1* chloroplasts under ATP-limiting conditions (50 μM). After the formation of import intermediates, we divided each import reaction in two. To half of each reaction, we added 5 mM ATP, whereas to the other half, no additional ATP was added. The amount of translocated protein then was expressed as a percentage of that bound initially, taking into account any differences in binding efficiency (Figure 6B). Once again, an ~20% deficiency in binding/outer envelope translocation was

observed. Although the amount of protein translocated into *ppi1* chloroplasts was less in absolute terms, when normalized according to the amount of preSSU bound in each case, it was found to be essentially the same as that in the wild type (Figure 6B): 9.5% of bound protein was translocated in the wild type, and 9.0% of bound protein was translocated in *ppi1*. These data indicate that the effect of *ppi1* is restricted to preprotein binding, which is consistent with a role for atToc33 in preprotein recognition.

DISCUSSION

We determined the relative levels of expression of *atTOC33* and *atTOC34* at different developmental stages and in different tissues to gain insight into the reason for the existence of these two homologous genes in Arabidopsis. The results showed that whereas *atTOC34* expression is relatively uniform and low level throughout development, *atTOC33* expression is upregulated strongly in new and expanding photosynthetic tissues (Figure 1). Our previous studies revealed that *atTOC33* and *atTOC34* both are expressed at higher levels in young plants and leaves than in older plants and leaves, but individual photosynthetic and nonphotosynthetic tissues were not investigated (Jarvis et al., 1998). The spatial expression patterns observed here are broadly in agreement with those described previously (Gutensohn et al., 2000), but this earlier study failed to address the relative expression levels of the two genes.

Our data correlate nicely with observations made using promoter–β-glucuronidase (GUS) fusions and by in situ hybridization (Gutensohn et al., 2000). The strongest expression of *atTOC34-GUS* was observed in roots and floral organs, whereas *atTOC33-GUS* expression was strongest in leaves and, in flowers, occurred only in the external, photosynthetic tissues of the sepals. In situ hybridization experiments confirmed these floral expression patterns and, furthermore, revealed that in stems, *atTOC33* expression is restricted largely to the photosynthetic mesophyll tissues (Gutensohn et al., 2000). Together, these data suggest that *atTOC33* is particularly important for the biogenesis of photosynthetic proteins. The observation that *ppi1* affects the development of chloroplasts but not root plastids is consistent with this hypothesis (Jarvis et al., 1998; Yu and Li, 2001).

Examination of the *ppi1* chloroplast proteome by immunoblot analysis and DIGE revealed that proteins involved directly in photosynthesis are specifically deficient (Figures 2 and 3, Table 1). Interestingly, we also observed a slight enrichment of several nonphotosynthetic proteins in *ppi1* chloroplasts. One possible explanation for the apparent enrichment of nonphotosynthetic proteins in *ppi1* is a simple concentration effect caused by the selective depletion of photosynthetic proteins. Another possibility is that these proteins are upregulated specifically in *ppi1* in an attempt to compensate for perturbations in chloroplast metabolism. The upregulation of molecular chaperones could be an attempt to maximize the activity and/or half-life of those photosynthetic components that are imported successfully. Whatever the reason, the fact that *ppi1* chloroplasts are able to accumulate increased quantities of these proteins sug-

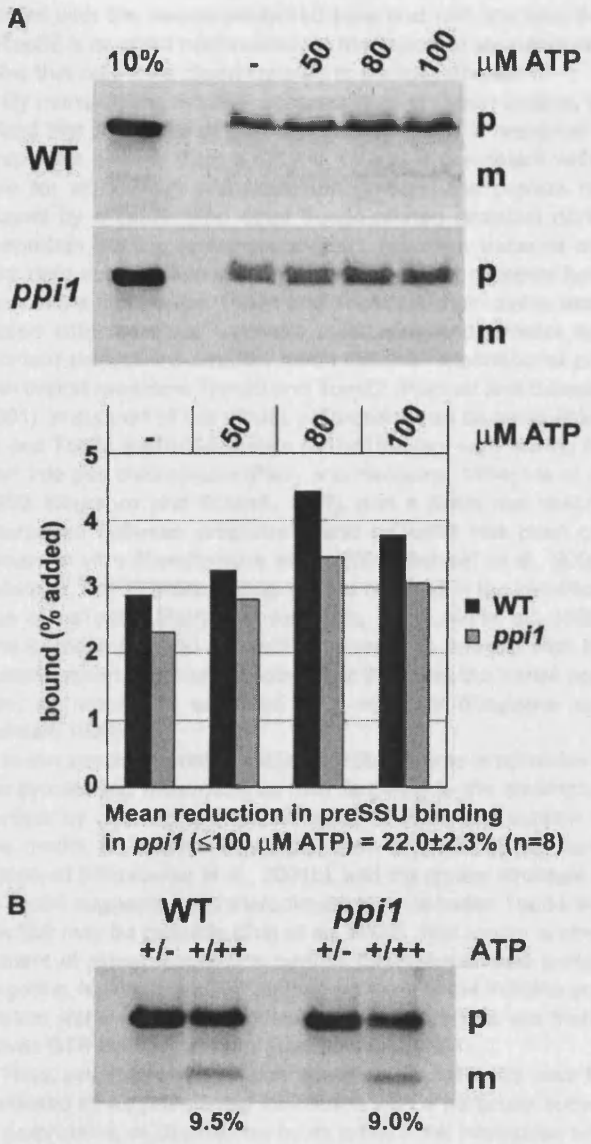


Figure 6. Effect of *ppi1* at Different Stages in the Import of preSSU.

(A) ATP-depleted wild-type (WT) or mutant (*ppi1*) chloroplasts and ^{35}S-Met–labeled preSSU were incubated together in the absence or presence of ATP at the indicated concentrations for 10 min in darkness. Quantification of each chloroplast-bound preSSU band in the assay shown is given in the corresponding graph. The data shown are representative of four independent experiments, from which the mean reduction in preSSU binding (±SE) was derived. m, mature; p, precursor.

(B) Import-chase results. Chloroplasts were incubated with preSSU under ATP-limiting conditions (50 μM). The chloroplasts then were reisolated and divided into two aliquots. Half were incubated in the absence of additional ATP (left lanes; indicated +/−), and the other half were incubated in the presence of 5 mM ATP (right lanes; indicated +/++). Values shown below the fluorograph indicate the amount of imported SSU in the lanes labeled +/++, expressed as a percentage of the amount of bound preSSU in the lanes labeled +/−. The data shown are representative of repeated experiments.

gests that the effect of *ppi1* on protein import is, at least to some degree, preprotein specific.

To gain a more global view of the effects of the *ppi1* mutation, we used DNA array technology to survey the expression of 3292 different genes: 2661 genes with predicted or actual chloroplast transit peptides and 631 others. Remarkably, the *ppi1* mutation resulted in the specific downregulation of highly expressed, photosynthetic genes. By contrast, genes related to the plastid genetic system, nonphotosynthetic metabolism, and various transport processes were upregulated moderately in *ppi1* (Figure 4, Table 2; see also supplemental data online). Comparison of these data with those obtained previously (Kurth et al., 2002; Strand et al., 2003) indicates that the gene expression changes observed in *ppi1* are not a general response to the perturbation of chloroplast biogenesis. An Arabidopsis knockout for the L11 protein of the 50S ribosomal subunit (Pesaresi et al., 2001)—a mutant that has a visible phenotype similar to that of *ppi1*—was found to exhibit upregulated expression of both photosynthetic and plastid genetic system genes (Kurth et al., 2002). On the other hand, more severe defects in chloroplast biogenesis caused by the growth of plants on the herbicide norflurazon resulted in the downregulation of both photosynthetic and plastid genetic system genes (Strand et al., 2003). Therefore, it seems plausible that the specific downregulation of photosynthetic genes observed in *ppi1* is an adaptive response to the reduction in import efficiency of the corresponding proteins.

When we analyzed the import kinetics of three different preproteins in *ppi1*, we found that the effect of the mutation varied according to the preprotein in question (Figure 5). Interestingly, the two preproteins that were affected most strongly (preSSU and preOE33) are components of the photosynthetic apparatus, whereas the preprotein that was affected least (preL11) is a component of the chloroplast's endogenous genetic system. These data indicate that the selective depletion of photosynthetic proteins in *ppi1* most likely is caused by a differential effect of the mutation on the import of different preproteins. Consequent deficiencies in photosynthetic proteins would, in turn, be predicted to cause downregulated expression of the corresponding nuclear genes through retrograde signaling (Surpin et al., 2002; Jarvis, 2003). If we extend this argument to its logical conclusion, one might suppose that all of the genes that are downregulated in *ppi1* encode preproteins that require atToc33 for their import into chloroplasts. However, it also is possible that *ppi1* affects the import of just a few proteins and that it is the consequent deficiencies in these proteins that perturb, indirectly, the expression and/or accumulation of a range of other (functionally related) proteins. Thus, although it seems likely that many of the proteins listed in Tables 1 and 2 use atToc33 preferentially during their import into chloroplasts, it is not necessarily the case that they all do.

Although studies on protein import kinetics in *ppi1* have not been conducted previously, *ppi1* chloroplasts were shown to import reduced quantities of SSU, LHCII, and two different isoforms of POR over a fixed period of time (Jarvis et al., 1998). Although POR is not involved directly in photosynthesis, it is a light-dependent enzyme that catalyzes the committed step in chlorophyll biosynthesis. These import data are therefore consistent with the results presented here and with the idea that atToc33 is involved preferentially in the import of abundant proteins that play roles closely related to photosynthesis.

By manipulating the ATP concentration in import assays, we found that the effect of *ppi1* on preSSU import is restricted to preprotein binding (Figure 6). This finding is consistent with a role for atToc33 in preprotein recognition. The precise role played by atToc33 (and other Toc34-related proteins) during preprotein binding remains uncertain, however, because current data suggest two different models for Toc receptor function. In the first model, Toc34 and Toc159 remain stably associated with the outer envelope membrane and interact with incident preproteins directly, much like the mitochondrial protein import receptors Tom20 and Tom22 (Pfanner and Geissler, 2001). In support of this model, preproteins can be cross-linked to pea Toc34 (psToc34) and/or psToc159 very early during import into pea chloroplasts (Perry and Keegstra, 1994; Ma et al., 1996; Kouranov and Schnell, 1997), and a direct and specific interaction between preproteins and psToc34 has been observed in vitro (Sveshnikova et al., 2000; Schleiff et al., 2002). However, not all cross-linking studies resulted in the identification of psToc34 (Perry and Keegstra, 1994; Ma et al., 1996), and in those that did, psToc34 appeared to interact with the mature region of the preprotein rather than with the transit peptide, as would be expected of a receptor (Kouranov and Schnell, 1997).

In the second model, soluble Toc159 binds to preproteins in the cytosol and then mediates their targeting to the chloroplast surface by docking at membrane-bound Toc34. In support of this model, an abundant cytosolic form of atToc159 has been observed (Hiltbrunner et al., 2001b), and the crystal structure of psToc34 suggests that heterodimerization between Toc34 and Toc159 may be possible (Sun et al., 2002). This model is reminiscent of signal recognition particle (SRP)–dependent protein targeting, in which the GTP binding protein SRP54 initiates preprotein translocation by docking at receptors that are themselves GTP binding proteins (Keenan et al., 2001).

Thus, preprotein recognition specificity by atToc33 may be mediated by its preferential interaction with a particular subset of preproteins, or alternatively by its preferential interaction with one or more of the different Toc159 isoforms. Bearing these possibilities in mind, it is interesting that atToc33 and atToc34 have been shown to exhibit differential competitor properties in an in vitro preSSU chloroplast binding assay: heterologously expressed atToc33 inhibited the binding of preSSU (translated using a rabbit reticulocyte system) to chloroplasts in a GTP-dependent manner, whereas atToc34 did not (Gutensohn et al., 2000). These data support the former possibility and suggest that the specific effect of *ppi1* on photosynthetic protein import is a reflection of differential preprotein binding properties of atToc33 and atToc34. Interestingly, during our analyses of protein import in the *ppi1* mutant, we consistently observed increased binding of unprocessed preL11 to mutant chloroplasts (Figure 5C). Increased abundance of a receptor with specificity for nonphotosynthetic preproteins (not necessarily associated with Toc complexes) might account for this observation. Although *atTOC34* transcript levels were not increased strongly in *ppi1* (Figure 1), it is possible that atToc34 accumulates to a

higher level in *ppi1* as a result of post-transcriptional processes.

Whatever the mode of action of the Toc GTPases, it seems likely that the proposed substrate specificity of the import apparatus is dependent on differences between the transit peptides of different preproteins. Transit peptides share a few general characteristics (e.g., they are rich in hydroxylated residues and deficient in acidic residues), but they differ widely in size and primary sequence and, unlike mitochondrial presequences, they do not share a common secondary structure (Bruce, 2001). This variability, together with the absence of precise information regarding the recognition process, makes it very difficult to predict which preproteins might be imported through a particular Toc complex or import pathway. Nevertheless, examination of the transit peptide sequences of the three preproteins used in this study revealed one property that correlated with import efficiency in the *ppi1* mutant: at pH 7.0, the preSSU, preOE33, and preL11 transit peptides are predicted to have net charges of 2.877, 3.046, and 6.206, respectively. It is conceivable that the higher charge of the preL11 transit peptide favored interaction with atToc34 (or atToc132- and atToc120-related proteins in the wheat germ translation system) and that the less highly charged preSSU and preOE33 transit peptides interact more favorably with atToc33 (or atToc159-related wheat germ proteins).

We have demonstrated that the *ppi1* mutation affects, with remarkable specificity, the expression, chloroplast import, and accumulation of photosynthetic proteins. Furthermore, the *ppi1* import defect appears to be restricted to the preprotein binding step, which is consistent with a role for atToc33 in preprotein recognition. Together, these data suggest that atToc33 is involved preferentially in the recognition of highly abundant, photosynthetic preproteins and that atToc34 may have preference for less abundant, nonphotosynthetic preproteins.

METHODS

Isolation of Arabidopsis Chloroplasts

Chloroplasts were isolated from 10-day-old wild-type and *ppi1* *Arabidopsis thaliana* plants (both Columbia-0 ecotype) grown in vitro as described previously (Aronsson and Jarvis, 2002). Plant material was homogenized for 3 to 4 s (wild type) or 2 to 3 s (*ppi1*) using a Polytron. After the first centrifugation, the homogenate (in 0.5 to 2 mL of isolation buffer) was loaded onto a linear Percoll gradient and centrifuged in a swing-out rotor at 7800*g* for 10 min. Intact chloroplasts were recovered and washed once with HS buffer (50 mM Hepes and 0.3 M sorbitol). The yield and intactness of the chloroplasts were determined as described previously (Aronsson and Jarvis, 2002).

Immunoblot Analysis

Standard methods were used for SDS-PAGE, protein gel blot analysis, and detection. Wild-type and *ppi1* chloroplast preparations were quantified using Bradford reagent (Bio-Rad) and solubilized in 2× SDS-PAGE sample buffer (20% [v/v] glycerol, 10% [v/v] 2-mercaptoethanol, 4% [w/v] SDS, 0.125 M Tris/HCl, pH 6.8, and 0.04% [w/v] bromophenol blue). Amounts of protein loaded per lane ranged from 20 to 0.1 μg. To confirm equal loading of the samples, control gels were stained with Coomassie Brilliant Blue R 250 (Fisher), and five different bands, selected at random, were quantified using ImageQuant software (Molecular Dynamics, Sunnyvale, CA). On average, differences between the samples were ~6%. Where possible, filters were cut across the lanes so that one half could be probed with a "photosynthetic" antibody and the other half could be probed with a "nonphotosynthetic" antibody. Primary antibodies were raised in rabbits and were kindly provided by Kenneth Cline (University of Florida, Gainesville) (Hsp70; raised against pea stromal Hsp70), Ulf-Ingo Flügge (University of Cologne, Germany) (TPT; raised against the spinach protein), Bernhard Grimm (Humboldt University, Germany) (CPO; raised against the tobacco protein), Neil Hoffman (National Science Foundation, Arlington, VA) (LHCII; raised against the pea protein), Kenneth Keegstra (Michigan State University, East Lansing) (ClpC; raised against pea stromal ClpC), Kenton Ko (Queen's University, Kingston, Ontario, Canada) (SSU; raised against the pea protein), and Henrik Scheller (The Royal Veterinary and Agricultural University, Copenhagen, Denmark) (PSI-D and FNR; both raised against barley proteins). The secondary antibody was an anti-rabbit IgG alkaline phosphatase conjugate (Sigma), and the detection reagent was 5-bromo-4-chloro-3-indolyl phosphate/nitroblue tetrazolium alkaline phosphatase substrate (Sigma).

Chloroplast Proteomics

Isolated chloroplasts were precipitated using 0.1 M ammonium acetate in 100% methanol and resuspended in 10 mM Tris/HCl, pH 8.5, 8 M urea, and 2% (w/v) amidosulfobetaine-14 (Calbiochem). Protein concentrations were determined using a detergent-compatible protein quantification kit (Bio-Rad). Protein samples from wild-type and mutant chloroplasts (50 μg each) were labeled with complementary CyDyeDIGE fluors (Cy3 and Cy5; Amersham Biosciences) (Unlu et al., 1997), pooled, and run on either 13- or 24-cm, pH-3 to -10 or pH-4 to -7 immobilized pH gradient strips (Amersham Biosciences). Resolution in the second dimension was performed using Hoeffer SE 600 or Ettan DALT SDS-PAGE systems (Amersham Biosciences). Gels were scanned using a 2920-2D MasterImager (Amersham Biosciences), and images were exported as 16-bit TIFF files. Image analysis was performed using DeCyder software (Amersham Biosciences). Protein spots showing >1.5-fold change in volume ratio were excised manually after colloidal Coomassie blue staining and digested to peptides using trypsin on a MassPrepStation (Micromass, Manchester, UK). The resulting peptides were examined by liquid chromatography/tandem mass spectrometry (Q-Tof, CapLC, Micromass/Waters, www.waters.com). Fragmentation data were used to search the National Centre for Biotechnology Information database using the MASCOT search engine. Multiple chloroplast preparations from each genotype were analyzed. Each combination of samples was run in triplicate for each pH range, including one experiment in which the CyDyeDIGE fluor labeling was reversed. Spots excised from each gel were treated individually so that multiple identifications were performed for each protein.

Protein Import into Arabidopsis Chloroplasts

Template DNA for the in vitro transcription/translation of preproteins was amplified by PCR from cDNA clones using M13 primers. The preSSU and preL11 cDNA clones were described previously (Aronsson and Jarvis, 2002), and the preOE33 cDNA clone was obtained from the ABRC (Columbus, OH) as clone 119E10T7. Transcription/translation was performed using a wheat germ system (Promega) containing ^{35}S-Met and T7 RNA polymerase according to the manufacturer's instructions.

Import reactions were performed in HMS buffer (50 mM HEPES, 3 mM $MgSO_4$, and 0.3 M sorbitol) containing 20 mM gluconic acid (potassium salt), 10 mM $NaHCO_3$, and 0.2% (w/v) BSA (Aronsson and Jarvis, 2002).

Each 150-μL import assay contained 10^7 chloroplasts, 5 mM MgATP, 10 mM Met, and translation product not exceeding 10% of the total volume. Import, thermolysin treatment, and quantification were performed according to Aronsson and Jarvis (2002).

To deplete endogenous ATP, chloroplasts were kept in the dark at room temperature for 20 min. Small molecules, including ATP, were removed from the translation products by Sephadex G 25 filtration (Pharmacia). MgATP was added to import reactions containing dark-adapted chloroplasts and ATP-depleted preprotein at different concentrations (50 to 5000 μM) or was omitted completely. Reactions were incubated in the dark at 25°C for 10 min and then analyzed as usual.

Import-chase experiments used similarly prepared, ATP-depleted chloroplasts and translation products and were performed in the dark. First, 2×10^7 chloroplasts were incubated with translation product in the presence of 50 μM ATP, in a total volume of 300 μL, for 10 min at 25°C. The chloroplasts then were reisolated by spinning through a 35% Percoll cushion in HS buffer at 2000*g* for 6 min and resuspended (with a cut 0.2-mL pipette tip) in 300 μL of buffer containing all of the components of an import reaction except chloroplasts, ATP, and translation product. The reaction then was divided into two equal aliquots of 150 μL. To one aliquot, ATP was added to a final concentration of 5 mM, and to the other aliquot, an equal volume of water was added. The import reactions were allowed to proceed for 10 min at 25°C and then analyzed as described previously.

RNA Isolation

RNA was isolated from plant material grown in vitro (5 and 10 days old) or on soil (28 days old). Plant material (up to 4 g) was ground to a powder in liquid nitrogen, and nucleic acids were extracted with equal volumes (~5 mL each) of 100 mM Tris/HCl, pH 9.0, and phenol in 15-mL tubes. Samples were centrifuged at 3000*g* for 30 min at 4°C and then purified by phenol:chloroform:isoamyl alcohol (25:24:1) extraction. Nucleic acid was precipitated by adding 1.5 mL of 3.0 M NaOAc, pH 5.2, and 2 volumes of 100% ethanol and incubating overnight at −20°C. After centrifugation at 3000*g* for 45 min at 4°C, samples were dried, resuspended in 500 μL of water, and transferred to 1.5-mL tubes. RNA was precipitated by adding 200 μL of 10 M LiCl and incubating on ice overnight. Samples were centrifuged at 20,000*g* for 15 min at 4°C, and the pellets were washed with 70% (v/v) ethanol, dried, and resuspended in sterile water. Heating to 65°C was necessary to dissolve the RNA.

RNA Gel Blot Analysis

Total RNA was fractionated by electrophoresis on 1.3% (w/v) agarose gels containing 1.85% (v/v) formaldehyde and Mops buffer [20 mM 3-(*N*-morpholino)-propanesulfonic acid, 1 mM EDTA, pH 8.0, and 50 mM NaOAc, pH 7.0] and transferred to Hybond NX membranes (Amersham Pharmacia Biotech) in 10× SSC (1× SSC is 0.15 M NaCl and 0.015 M sodium citrate). Membranes were probed with α-^{32}P-dCTP–labeled DNA fragments (Feinberg and Vogelstein, 1983) in 0.3 M sodium phosphate buffer, pH 7.2, containing 7% (w/v) SDS, 1 mM EDTA, pH 8.0, and 2% (w/v) BSA at 65°C overnight. Filters were washed at 65°C twice with 2× SSC and 0.1% (w/v) SDS for 30 min, once with 1× SSC and 0.1% (w/v) SDS for 20 min, and once with 0.5× SSC and 0.1% (w/v) SDS for 10 min. Band quantification was performed using ImageQuant software (Molecular Dynamics).

Identical filters were probed simultaneously with *atTOC33* and *atTOC34* probes. The *atTOC33* and *atTOC34* probes corresponded to full-length cDNAs and were almost identical in length (1132 and 1158 bp, respectively) and GC content (40.99 and 41.28%, respectively). The probes were labeled simultaneously under identical conditions using the same isotope and were shown to have identical specific activities by scintillation counting. Hybridization, washing, and exposure steps were performed simultaneously under identical conditions. Cross-hybridization of the probes between *atTOC33* and *atTOC34* (sequence homology of 50.2%) can be excluded, because the wash stringency used allowed only probe/target combinations with >83% identity to remain hybridized.

DNA Array Analysis

The 1827 gene-specific tags (GSTs) described previously (Kurth et al., 2002) were combined with an additional 1465 GSTs from genes that encode proteins featuring a chloroplast transit peptide predicted by TargetP (Emanuelsson et al., 2000), resulting in a 3292-GST array (Richly et al., 2003). Amplification, quantification, verification, and spotting (in duplicate) of PCR products on nylon filters were performed as described (Varotto et al., 2001; Kurth et al., 2002). At least three experiments with different filters and independent cDNA probes from plant pools were performed for each condition, thus minimizing variation between individual plants, filters, or probes. cDNA probe synthesis was primed by a mixture of oligonucleotides matching the 3292 genes in the antisense orientation and hybridized to the GST array as described (Kurth et al., 2002). Images were read using a Storm PhosphorImager (Molecular Dynamics) imported into the ArrayVision program (version 6.0; Imaging Research, St. Catharines, Ontario, Canada), in which data were normalized with reference to all spots on the array (Kurth et al., 2002). After import of expression data into the ArrayStat program (version 1.0; Imaging Research), a *z* test (nominal α set to 0.05) was performed to identify statistically significant differential expression values. The data obtained for all significantly differentially regulated genes are given in the supplemental data online.

Upon request, materials integral to the findings presented in this publication will be made available in a timely manner to all investigators on similar terms for noncommercial research purposes. To obtain materials, please contact Paul Jarvis, rpj3@le.ac.uk.

ACKNOWLEDGMENTS

We thank Svenja Hester, Julie Howard, Christine Jackson, Adrian Shirlin, and Anthony Wardle for technical assistance and Jocelyn Bédard, Jonathan Combe, Sabina Kovacheva, Enrique López-Juez, and Simon Geir Møller for their insightful comments on the manuscript. We are grateful to Kenneth Cline, Ulf-Ingo Flügge, Bernhard Grimm, Neil Hoffman, Kenneth Keegstra, Kenton Ko, and Henrik Scheller for generously providing antibodies. This work was supported by the Royal Society Rosenheim Research Fellowship (to P.J.), by Biotechnology and Biological Sciences Research Council (BBSRC) Grants 91/C12976 and 91/P12928 (to P.J.), by the Deutsche Forschungsgemeinschaft (to D.L.), and by the Bundesministerium für Bildung und Forschung (to J.K. and D.L.). The GARNet proteomics facility at the University of Cambridge was funded by the BBSRC Investigating Gene Function Initiative.

Received April 15, 2003; accepted June 8, 2003.

REFERENCES

Aronsson, H., and Jarvis, P. (2002). A simple method for isolating import-competent *Arabidopsis* chloroplasts. FEBS Lett. **529,** 215–220.

Bauer, J., Chen, K., Hiltbrunner, A., Wehrli, E., Eugster, M., Schnell, D., and Kessler, F. (2000). The major protein import receptor of plastids is essential for chloroplast biogenesis. Nature **403,** 203–207.

Bölter, B., May, T., and Soll, J. (1998). A protein import receptor in pea chloroplasts, Toc86, is only a proteolytic fragment of a larger polypeptide. FEBS Lett. **441,** 59–62.

Bruce, B.D. (2001). The paradox of plastid transit peptides: Conserva-

tion of function despite divergence in primary structure. Biochim. Biophys. Acta **1541,** 2–21.

Chen, K., Chen, X., and Schnell, D.J. (2000). Mechanism of protein import across the chloroplast envelope. Biochem. Soc. Trans. **28,** 485–491.

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

Feinberg, A.P., and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132,** 6–13.

Gutensohn, M., Schulz, B., Nicolay, P., and Flugge, U.I. (2000). Functional analysis of the two *Arabidopsis* homologues of Toc34, a component of the chloroplast protein import apparatus. Plant J. **23,** 771–783.

Hiltbrunner, A., Bauer, J., Alvarez-Huerta, M., and Kessler, F. (2001a). Protein translocon at the *Arabidopsis* outer chloroplast membrane. Biochem. Cell Biol. **79,** 629–635.

Hiltbrunner, A., Bauer, J., Vidi, P.A., Infanger, S., Weibel, P., Hohwy, M., and Kessler, F. (2001b). Targeting of an abundant cytosolic form of the protein import receptor at Toc159 to the outer chloroplast membrane. J. Cell Biol. **154,** 309–316.

Hinnah, S.C., Wagner, R., Sveshnikova, N., Harrer, R., and Soll, J. (2002). The chloroplast protein import channel Toc75: Pore properties and interaction with transit peptides. Biophys. J. **83,** 899–911.

Hirsch, S., Muckel, E., Heemeyer, F., von Heijne, G., and Soll, J. (1994). A receptor component of the chloroplast protein translocation machinery. Science **266,** 1989–1992.

Jackson-Constan, D., and Keegstra, K. (2001). *Arabidopsis* genes encoding components of the chloroplastic protein import apparatus. Plant Physiol. **125,** 1567–1576.

Jarvis, P. (2003). Intracellular signalling: The language of the chloroplast. Curr. Biol. **13,** R314–R316.

Jarvis, P., Chen, L.J., Li, H., Peto, C.A., Fankhauser, C., and Chory, J. (1998). An *Arabidopsis* mutant defective in the plastid general protein import apparatus. Science **282,** 100–103.

Jarvis, P., and Soll, J. (2002). Toc, Tic, and chloroplast protein import. Biochim. Biophys. Acta **1590,** 177–189.

Keegstra, K., and Cline, K. (1999). Protein import and routing systems of chloroplasts. Plant Cell **11,** 557–570.

Keenan, R.J., Freymann, D.M., Stroud, R.M., and Walter, P. (2001). The signal recognition particle. Annu. Rev. Biochem. **70,** 755–775.

Kessler, F., Blobel, G., Patel, H.A., and Schnell, D.J. (1994). Identification of two GTP-binding proteins in the chloroplast protein import machinery. Science **266,** 1035–1039.

Kouranov, A., and Schnell, D.J. (1997). Analysis of the interactions of preproteins with the import machinery over the course of protein import into chloroplasts. J. Cell Biol. **139,** 1677–1685.

Kurth, J., Varotto, C., Pesaresi, P., Biehl, A., Richly, E., Salamini, F., and Leister, D. (2002). Gene-sequence-tag expression analyses of 1,800 genes related to chloroplast functions. Planta **215,** 101–109.

Leister, D. (2003). Chloroplast research in the genomic age. Trends Genet. **19,** 47–56.

Ma, Y., Kouranov, A., LaSala, S.E., and Schnell, D.J. (1996). Two components of the chloroplast protein import apparatus, IAP86 and IAP75, interact with the transit sequence during the recognition and translocation of precursor proteins at the outer envelope. J. Cell Biol. **134,** 315–327.

Nielsen, E., Akita, M., Davila-Aponte, J., and Keegstra, K. (1997). Stable association of chloroplastic precursors with protein translocation complexes that contain proteins from both envelope membranes and a stromal Hsp100 molecular chaperone. EMBO J. **16,** 935–946.

Olsen, L.J., and Keegstra, K. (1992). The binding of precursor proteins to chloroplasts requires nucleoside triphosphates in the intermembrane space. J. Biol. Chem. **267,** 433–439.

Olsen, L.J., Theg, S.M., Selman, B.R., and Keegstra, K. (1989). ATP is required for the binding of precursor proteins to chloroplasts. J. Biol. Chem. **264,** 6724–6729.

Perry, S.E., and Keegstra, K. (1994). Envelope membrane proteins that interact with chloroplastic precursor proteins. Plant Cell **6,** 93–105.

Pesaresi, P., Varotto, C., Meurer, J., Jahns, P., Salamini, F., and Leister, D. (2001). Knock-out of the plastid ribosomal protein L11 in *Arabidopsis*: Effects on mRNA translation and photosynthesis. Plant J. **27,** 179–189.

Pfanner, N., and Geissler, A. (2001). Versatility of the mitochondrial protein import machinery. Nat. Rev. Mol. Cell Biol. **2,** 339–349.

Richly, E., Dietzmann, A., Biehl, A., Kurth, J., Laloi, C., Apel, K., Salamini, F., and Leister, D. (2003). Co-variations in the nuclear chloroplast transcriptome reveal a regulatory master switch. EMBO Rep. **4,** 491–498.

Rodermel, S., Haley, J., Jiang, C.Z., Tsai, C.H., and Bogorad, L. (1996). A mechanism for intergenomic integration: Abundance of ribulose bisphosphate carboxylase small-subunit protein influences the translation of the large-subunit mRNA. Proc. Natl. Acad. Sci. USA **93,** 3881–3885.

Schleiff, E., Soll, J., Sveshnikova, N., Tien, R., Wright, S., Dabney-Smith, C., Subramanian, C., and Bruce, B.D. (2002). Structural and guanosine triphosphate/diphosphate requirements for transit peptide recognition by the cytosolic domain of the chloroplast outer envelope receptor, Toc34. Biochemistry **41,** 1934–1946.

Schnell, D.J., Kessler, F., and Blobel, G. (1994). Isolation of components of the chloroplast protein import machinery. Science **266,** 1007–1012.

Seedorf, M., Waegemann, K., and Soll, J. (1995). A constituent of the chloroplast import complex represents a new type of GTP-binding protein. Plant J. **7,** 401–411.

Strand, A., Asami, T., Alonso, J., Ecker, J.R., and Chory, J. (2003). Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrinIX. Nature **421,** 79–83.

Sun, Y.J., Forouhar, F., Li Hm, H.M., Tu, S.L., Yeh, Y.H., Kao, S., Shr, H.L., Chou, C.C., Chen, C., and Hsiao, C.D. (2002). Crystal structure of pea Toc34, a novel GTPase of the chloroplast protein translocon. Nat. Struct. Biol. **9,** 95–100.

Surpin, M., Larkin, R.M., and Chory, J. (2002). Signal transduction between the chloroplast and the nucleus. Plant Cell **14** (suppl.), S327–S338.

Sveshnikova, N., Soll, J., and Schleiff, E. (2000). Toc34 is a preprotein receptor regulated by GTP and phosphorylation. Proc. Natl. Acad. Sci. USA **97,** 4973–4978.

Unlu, M., Morgan, M.E., and Minden, J.S. (1997). Difference gel electrophoresis: A single gel method for detecting changes in protein extracts. Electrophoresis **18,** 2071–2077.

Varotto, C., Richly, E., Salamini, F., and Leister, D. (2001). GST-PRIME: A genome-wide primer design software for the generation of gene sequence tags. Nucleic Acids Res. **29,** 4373–4377.

Yu, T.S., and Li, H. (2001). Chloroplast protein translocon components atToc159 and atToc33 are not essential for chloroplast biogenesis in guard cells and root cells. Plant Physiol. **127,** 90–96.

Appendix 2 – *ppi1* (Landsberg *erecta* background)

With a view to mapping the suppressor mutations genetically, the *ppi1* mutation was introgressed into the Landsberg *erecta* genetic background (*ppi1* was originally isolated in the Columbia-0 background). Seven outcrosses into Landsberg *erecta* were done to achieve a theoretical 99% ecotype purity. *ppi1*(L*er* ecotype) have the same yellow-green phenotype as *ppi1*(Col-0 ecotype), but retain the characteristic rounder leaves and short inflorescences of Landsberg *erecta*.

DNA isolated from *ppi1*(L*er* ecotype) was tested by PCR with various markers to determine if the *ppi1* loci had been successfully introgressed into the Landsberg background; i.e., the *ppi1*(L*er* ecotype) was Landsberg *erecta* rather than Columbia ecotype for each marker tested.

Marker	Co-ordinates (cM)	TAIR accession	$ppi1_{Ler}$	Type
Chromosome 1				
EAT	3.87	1945482	Col	SSLP
M235	31.90	1945678	L*er*	CAPS
----------Centromere----------				
GAPB.1	59.00	1945602	L*er*	CAPS
AF-6	114.00	2005285	L*er*	CAPS
GL2	121.70	2005270	L*er*	CAPS
Chromosome 2				
CIC9A3R	16.50	2005299	L*er*	CAPS
----------Centromere----------				
T20K24	35.60	1945652	L*er*	CAPS
F7F1	62.00	2005288	L*er*	CAPS
90J19T7	86.77	1945555	L*er*	CAPS
Chromosome 3				
CA1	4.24	1945575	L*er*	CAPS
----------Centromere----------				
NGA6	86.41	1945538	L*er*	SSLP
Chromosome 4				
GA1.1	17.20	1945599	L*er*	CAPS
----------Centromere----------				
G8300.2	81.22	1945676	L*er*	CAPS
Chromosome 5				
PAT1.2	5.98	1945631	L*er*	indel
NGA139	50.48	1945524	L*er*	SSLP
----------Centromere----------				
PHYC.1	71.13	1945634	L*er*	CAPS
LFY3	116.88	1945613	L*er*	CAPS

Bibliography

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

Aitken, A., Collinge, D.B., van Heusden, B.P., Isobe, T., Roseboom, P.H., Rosenfeld, G., and Soll, J. (1992). 14-3-3 proteins: a highly conserved, widespread family of eukaryotic proteins. Trends Biochem. Sci. **17,** 498-501.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C., and Ecker, J.R. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana.* Science **301,** 653-657.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25,** 3389-3402.

America, T., Hageman, J., Guera, A., Rook, F., Archer, K., Keegstra, K., and Weisbeek, P. (1994). Methotrexate does not block import of a DHFR fusion protein into chloroplasts. Plant Mol. Biol. **24,** 283-294.

Apuya, N.R., Yadegari, R., Fischer, R.L., Harada, J.J., Zimmerman, J.L., and Goldberg, R.B. (2001). The *Arabidopsis* embryo mutant *schlepperless* has a defect in the chaperonin-60alpha gene. Plant Physiol. **126,** 717-730.

Arabidopsis-Genome-Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature **408,** 796-815.

Aronsson, H., and Jarvis, P. (2002). A simple method for isolating import-competent *Arabidopsis* chloroplasts. FEBS Lett. **529,** 215-220.

Aronsson, H., Combe, J., and Jarvis, P. (2003). Unusual nucleotide-binding properties of the chloroplast protein import receptor, atToc33. FEBS Lett **544,** 79-85.

Baltimore, D. (1985). Retroviruses and retrotransposons: the role of reverse transcription in shaping the eukaryotic genome. Cell **40,** 481-482.

Bauer, J., Chen, K., Hiltbunner, A., Wehrli, E., Eugster, M., Schnell, D., and Kessler, F. (2000). The major protein import receptor of plastids is essential for chloroplast biogenesis. Nature **403,** 203-207.

Bauer, J., Hiltbrunner, A., Weibel, P., Vidi, P.A., Alvarez-Huerta, M., Smith, M.D., Schnell, D.J., and Kessler, F. (2002). Essential role of the G-domain in targeting of

the protein import receptor atToc159 to the chloroplast outer membrane. J. Cell Biol. **159,** 845-854.

Baumann, P., Baumann, L., Lai, C.Y., Rouhbakhsh, D., Moran, N.A., and Clark, M.A. (1995). Genetics, physiology, and evolutionary relationships of the genus Buchnera: intracellular symbionts of aphids. Annu. Rev. Microbiol. **49,** 55-94.

Birnboim, H.C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7,** 1513-1523.

Blanc, G., Barakat, A., Guyot, R., Cooke, R., and Delseny, M. (2000). Extensive duplication and reshuffling in the *Arabidopsis* genome. Plant Cell **12,** 1093-1101.

Bolter, B., May, T., and Soll, J. (1998a). A protein import receptor in pea chloroplasts, Toc86, is only a proteolytic fragment of a larger polypeptide. FEBS Lett. **441,** 59-62.

Bolter, B., Soll, J., Schulz, A., Hinnah, S., and Wagner, R. (1998b). Origin of a chloroplast protein importer. Proc. Natl. Acad. Sci. USA **95,** 15831-15836.

Bourne, H.R., Sanders, D.A., and McCormick, F. (1991). The GTPase superfamily: conserved structure and molecular mechanism. Nature **349,** 117-127.

Bruce, B.D. (1998). The role of lipids in plastid protein transport. Plant Mol. Biol. **38,** 223-246.

Bruce, B.D. (2001). The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. Biochim. Biophys. Acta **1541,** 2-21.

Caliebe, A., Grimm, R., Kaiser, G., Lubeck, J., Soll, J., and Heins, L. (1997). The chloroplastic protein import machinery contains a Rieske-type iron- sulfur cluster and a mononuclear iron-binding protein. EMBO J. **16,** 7342-7350.

Chatterjee, M., Sparvoli, S., Edmunds, C., Garosi, P., Findlay, K., and Martin, C. (1996). DAG, a gene required for chloroplast differentiation and palisade development in Antirrhinum majus. EMBO J. **15,** 4194-4207.

Chen, K., Chen, X., and Schnell, D.J. (2000a). Initial binding of preproteins involving the Toc159 receptor can be bypassed during protein import into chloroplasts. Plant Physiol. **122,** 813-822.

Chen, K., Chen, X., and Schnell, D.J. (2000b). Mechanism of protein import across the chloroplast envelope. Biochem. Soc. Trans. **28,** 485-491.

Chen, L.J., and Li, H.M. (1998). A mutant deficient in the plastid lipid DGD is defective in protein import into chloroplasts. Plant J. **16,** 33-39.

Chen, X., Smith, M.D., Fitzpatrick, L., and Schnell, D.J. (2002). *In vivo* analysis of the role of atTic20 in protein import into chloroplasts. Plant Cell **14,** 641-654.

Chew, O., Lister, R., Qbadou, S., Heazlewood, J.L., Soll, J., Schleiff, E., Millar, A.H., and Whelan, J. (2004). A plant outer mitochondrial membrane protein with high

amino acid sequence identity to a chloroplast protein import receptor. FEBS Lett **557,** 109-114.

Chou, M.L., Fitzpatrick, L.M., Tu, S.L., Budziszewski, G., Potter-Lewis, S., Akita, M., Levin, J.Z., Keegstra, K., and Li, H.M. (2003). Tic40, a membrane-anchored co-chaperone homolog in the chloroplast protein translocon. EMBO J. **22,** 2970-2980.

Chua, N.H., and Schmidt, G.W. (1978). In vitro synthesis, transport, and assembly of ribulose 1,5-bisphosphate carboxylase subunits. Basic Life Sci. **11,** 325-347.

Claros, M.G., Brunak, S., and von Heijne, G. (1997). Prediction of N-terminal protein sorting signals. Curr. Opin. Struct. Biol. **7,** 394-398.

Cline, K., and Henry, R. (1996). Import and routing of nucleus-encoded chloroplast proteins. Annu. Rev. Cell. Dev. Biol. **12,** 1-26.

Cline, K., Werner-Washburne, M., Andrews, J., and Keegstra, K. (1984). Thermolysin is a suitable protease for probing the surface of intact pea chloroplasts. Plant Physiol. **75,** 675-678.

Cline, K., Werner-Washburne, M., Lubben, T.H., and Keegstra, K. (1985). Precursors to two nuclear-encoded chloroplast proteins bind to the outer envelope membrane before being imported into chloroplasts. J. Biol. Chem. **260,** 3691-3696.

Constan, D., Patel, R., Keegstra, K., and Jarvis, P. (2004). An outer envelope membrane component of the plastid protein import apparatus plays an essential role in *Arabidopsis*. Plant J. - in press.

Dahlin, C., and Cline, K. (1991). Developmental regulation of the plastid protein import apparatus. Plant Cell **3,** 1131-1140.

Davila-Aponte, J.A., Inoue, K., and Keegstra, K. (2003). Two chloroplastic protein translocation components, Tic110 and Toc75, are conserved in different plastid types from multiple plant species. Plant Mol. Biol. **51,** 175-181.

Derr, L.K., Strathern, J.N., and Garfinkel, D.J. (1991). RNA-mediated recombination in *S. cerevisiae*. Cell **67,** 355-364.

Dormann, P., Hoffmann-Benning, S., Balbo, I., and Benning, C. (1995). Isolation and characterization of an Arabidopsis mutant deficient in the thylakoid lipid digalactosyl diacylglycerol. Plant Cell **7,** 1801-1810.

Douglas, S.E. (1998). Plastid evolution: origins, diversity, trends. Curr. Opin. Genet. Dev. **8,** 655-661.

Drouin, G., and Moniz de Sa, M. (1997). Loss of introns in the pollen-specific actin subfamily members of potato and tomato. J. Mol. Evol. **45,** 509-513.

Eckart, K., Eichacker, L., Sohrt, K., Schleiff, E., Heins, L., and Soll, J. (2002). A Toc75-like protein import channel is abundant in chloroplasts. EMBO Rep. **3,** 557-562.

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

Flugge, U.I., and Hinz, G. (1986). Energy dependence of protein translocation into chloroplasts. Eur. J. Biochem. **160,** 563-570.

Grossman, A.R., Bartlett, S.G., Schmidt, G.W., Mullet, J.E., and Chua, N.H. (1982). Optimal conditions for post-translational uptake of proteins by isolated chloroplasts. In vitro synthesis and transport of plastocyanin, ferredoxin-NADP+ oxidoreductase, and fructose-1,6-bisphosphatase. J. Biol. Chem. **257,** 1558-1563.

Gutensohn, M., Schulz, B., Nicolay, P., and Flugge, U.I. (2000). Functional analysis of the two *Arabidopsis* homologues of Toc34, a component of the chloroplast protein import apparatus. Plant J. **23,** 771-783.

Hager, K.P., Muller, B., Wind, C., Erbach, S., and Fischer, H. (1996). Evolution of legumin genes: loss of an ancestral intron at the beginning of angiosperm diversification. FEBS Lett **387,** 94-98.

Hartel, H., Lokstein, H., Dormann, P., Grimm, B., and Benning, C. (1997). Changes in the composition of the photosynthetic apparatus in the galactolipid-deficient dgd1 mutant of Arabidopsis thaliana. Plant Physiol. **115,** 1175-1184.

Heins, L., Mehrle, A., Hemmler, R., Wagner, R., Kuchler, M., Hormann, F., Sveshnikov, D., and Soll, J. (2002). The preprotein conducting channel at the inner envelope membrane of plastids. EMBO J. **21,** 2616-2625.

Hidalgo, E., Ding, H., and Demple, B. (1997). Redox signal transduction via iron-sulfur clusters in the SoxR transcription activator. Trends Biochem. Sci. **22,** 207-210.

Highfield, P.E., and Ellis, R.J. (1978). Synthesis and transport of the small subunit of chloroplast ribulose bisphosphate carboxylase. Nature **271,** 420-424.

Hiltbrunner, A., Bauer, J., Alvarez-Huerta, M., and Kessler, F. (2001a). Protein translocon at the *Arabidopsis* outer chloroplast membrane. Biochem. Cell Biol. **79,** 629-635.

Hiltbrunner, A., Bauer, J., Vidi, P.A., Infanger, S., Weibel, P., Hohwy, M., and Kessler, F. (2001b). Targeting of an abundant cytosolic form of the protein import receptor at Toc159 to the outer chloroplast membrane. J. Cell Biol. **154,** 309-316.

Hinnah, S.C., Hill, K., Wagner, R., Schlicher, T., and Soll, J. (1997). Reconstitution of a chloroplast protein import channel. EMBO J. **16,** 7351-7360.

Hinnah, S.C., Wagner, R., Sveshnikova, N., Harrer, R., and Soll, J. (2002). The chloroplast protein import channel Toc75: pore properties and interaction with transit peptides. Biophys. J. **83,** 899-911.

Hirsch, S., Muckel, E., Heemeyer, F., von Heijne, G., and Soll, J. (1994). A receptor component of the chloroplast protein translocation machinery. Science **266,** 1989-1992.

Hoffmeister, M., and Martin, W. (2003). Interspecific evolution: microbial symbiosis, endosymbiosis and gene transfer. Environ. Microbiol. **5,** 641-649.

Horniak, L., Pilon, M., van 't Hof, R., and de Kruijff, B. (1993). The secondary structure of the ferredoxin transit sequence is modulated by its interaction with negatively charged lipids. FEBS Lett. **334,** 241-246.

Huang, F., Hedman, E., Funk, C., Kieselbach, T., Schroder, W.P., and Norling, B. (2004). Isolation of outer membrane of *Synechocystis sp.* PCC 6803 and its proteomic characterization. Mol. Cell Proteomics.

Huang, N., Sutliff, T.D., Litts, J.C., and Rodriguez, R.L. (1990). Classification and characterization of the rice alpha-amylase multigene family. Plant Mol Biol **14,** 655-668.

Inaba, T., Li, M., Alvarez-Huerta, M., Kessler, F., and Schnell, D.J. (2003). atTic110 functions as a scaffold for coordinating the stromal events of protein import into chloroplasts. J. Biol. Chem. **278,** 38617-38627.

Inoue, K., and Keegstra, K. (2003). A polyglycine stretch is necessary for proper targeting of the protein translocation channel precursor to the outer envelope membrane of chloroplasts. Plant J. **34,** 661-669.

Inoue, K., Demel, R., de Kruijff, B., and Keegstra, K. (2001). The N-terminal portion of the preToc75 transit peptide interacts with membrane lipids and inhibits binding and import of precursor proteins into isolated chloroplasts. Eur. J. Biochem. **268,** 4036-4043.

Jackson, D.T., Froehlich, J.E., and Keegstra, K. (1998). The hydrophilic domain of Tic110, an inner envelope membrane component of the chloroplastic protein translocation apparatus, faces the stromal compartment. J. Biol. Chem. **273,** 16583-16588.

Jackson-Constan, D., and Keegstra, K. (2001). *Arabidopsis* genes encoding components of the chloroplastic protein import apparatus. Plant Physiol. **125,** 1567-1576.

Jarvis, P. (2003). Intracellular signalling: the language of the chloroplast. Curr. Biol. **13,** R314-R316.

Jarvis, P., and Soll, J. (2002). Toc, Tic, and chloroplast protein import. Biochim. Biophys. Acta **1590,** 177-189.

Jarvis, P., Chen, L.J., Li, H., Peto, C.A., Fankhauser, C., and Chory, J. (1998). An *Arabidopsis* mutant defective in the plastid general protein import apparatus. Science **282,** 100-103.

Jarvis, P., Dormann, P., Peto, C.A., Lutes, J., Benning, C., and Chory, J. (2000). Galactolipid deficiency and abnormal chloroplast development in the *Arabidopsis MGD synthase 1* mutant. Proc. Natl. Acad. Sci. USA **97,** 8175-8179.

Jelic, M., Sveshnikova, N., Motzkus, M., Horth, P., Soll, J., and Schleiff, E. (2002). The chloroplast import receptor Toc34 functions as preprotein-regulated GTPase. Biol. Chem. **383,** 1875-1883.

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

Keddie, J.S., Carroll, B., Jones, J.D., and Gruissem, W. (1996). The DCL gene of tomato is required for chloroplast development and palisade cell morphogenesis in leaves. EMBO J. **15,** 4208-4217.

Keegstra, K. (1989). Transport and routing of proteins into chloroplasts. Cell **56,** 247-253.

Keegstra, K., and Cline, K. (1999). Protein import and routing systems of chloroplasts. Plant Cell **11,** 557-570.

Kessler, F., and Blobel, G. (1996). Interaction of the protein import and folding machineries of the chloroplast. Proc. Natl. Acad. Sci. USA **93,** 7684-7689.

Kessler, F., Blobel, G., Patel, H.A., and Schnell, D.J. (1994). Identification of two GTP-binding proteins in the chloroplast protein import machinery. Science **266,** 1035-1039.

Ko, K., and Ko, Z.W. (1992). Carboxyl-terminal sequences can influence the *in vitro* import and intraorganellar targeting of chloroplast protein precursors. J. Biol. Chem. **267,** 13910-13916.

Kouranov, A., and Schnell, D.J. (1997). Analysis of the interactions of preproteins with the import machinery over the course of protein import into chloroplasts. J. Cell Biol. **139,** 1677-1685.

Kouranov, A., Wang, H., and Schnell, D.J. (1999). Tic22 is targeted to the intermembrane space of chloroplasts by a novel pathway. J. Biol. Chem. **274,** 25181-25186.

Kouranov, A., Chen, X., Fuks, B., and Schnell, D.J. (1998). Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. J. Cell Biol. **143,** 991-1002.

Kourtz, L., and Ko, K. (1997). The early stage of chloroplast protein import involves Com70. J. Biol. Chem. **272,** 2808-2813.

Kubis, S., Baldwin, A., Patel, R., Razzaq, A., Dupree, P., Lilley, K., Kurth, J., Leister, D., and Jarvis, P. (2003). The *Arabidopsis ppi1* mutant is specifically defective in the expression, chloroplast import, and accumulation of photosynthetic proteins. Plant Cell **15,** 1859-1871.

Kuchler, M., Decker, S., Hormann, F., Soll, J., and Heins, L. (2002). Protein import into chloroplasts involves redox-regulated proteins. EMBO J. **21,** 6136-6145.

Kumar, V., and Trick, M. (1993). Sequence complexity of the S receptor kinase gene family in Brassica. Mol. Gen. Genet. **241,** 440-446.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227,** 680-685.

Lee, K.H., Kim, S.J., Lee, Y.J., Jin, J.B., and Hwang, I. (2003). The M domain of atToc159 plays an essential role in the import of proteins into chloroplasts and chloroplast biogenesis. J. Biol. Chem. **278,** 36794-36805.

Leister, D. (2003). Chloroplast research in the genomic age. Trends Genet. **19,** 47-56.

Levin, J.Z., Fletcher, J.C., Chen, X., and Meyerowitz, E.M. (1998). A genetic screen for modifiers of UFO meristem activity identifies three novel FUSED FLORAL ORGANS genes required for early flower development in Arabidopsis. Genetics **149,** 579-595.

Leyser, O.H.M., and Furner, I.J. (1992). Characterisation of three shoot apical meristem mutants of *Arabidopsis thaliana.* Development **116,** 397-403.

Li, H., and Chen, L.J. (1997). A novel chloroplastic outer membrane-targeting signal that functions at both termini of passenger polypeptides. J. Biol. Chem. **272,** 10968-10974.

Lubeck, J., Soll, J., Akita, M., Nielsen, E., and Keegstra, K. (1996). Topology of IEP110, a component of the chloroplastic protein import machinery present in the inner envelope membrane. EMBO J. **15,** 4230-4238.

Ma, Y., Kouranov, A., LaSala, S.E., and Schnell, D.J. (1996). Two components of the chloroplast protein import apparatus, IAP86 and IAP75, interact with the transit sequence during the recognition and translocation of precursor proteins at the outer envelope. J. Cell Biol. **134,** 315-327.

Margulis, L. (1970). Recombination of non-chromosomal genes in Chlamydomonas: assortment of mitochondria and chloroplasts? J Theor Biol **26,** 337-342.

Marshall, J.S., DeRocher, A.E., Keegstra, K., and Vierling, E. (1990). Identification of heat shock protein hsp70 homologues in chloroplasts. Proc. Natl. Acad. Sci. USA **87,** 374-378.

Martin, J., Mahlke, K., and Pfanner, N. (1991). Role of an energized inner membrane in mitochondrial protein import. Delta psi drives the movement of presequences. J Biol Chem **266,** 18051-18057.

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

May, T., and Soll, J. (1998). Positive charges determine the topology and functionality of the transmembrane domain in the chloroplastic outer envelope protein Toc34. J. Cell Biol. **141,** 895-904.

May, T., and Soll, J. (2000). 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. Plant Cell **12,** 53-64.

McFadden, G.I. (1999). Endosymbiosis and evolution of the plant cell. Curr. Opin. Plant Biol. **2,** 513-519.

McFadden, G.I. (2001). Chloroplast origin and integration. Plant Physiol. **125,** 50-53.

Miras, S., Salvi, D., Ferro, M., Grunwald, D., Garin, J., Joyard, J., and Rolland, N. (2002). Non-canonical transit peptide for import into the chloroplast. J. Biol. Chem. **277,** 47770-47778.

Moran, N.A. (1996). Accelerated evolution and Muller's rachet in endosymbiotic bacteria. Proc. Natl. Acad. Sci. USA **93,** 2873-2878.

Muller, H.J. (1932). Some Genetic Aspects of Sex. American Naturalist **66,** 118-138.

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant **15,** 473-497.

Ohlrogge, J., and Browse, J. (1995). Lipid biosynthesis. Plant Cell **7,** 957-970.

Olsen, L.J., Theg, S.M., Selman, B.R., and Keegstra, K. (1989). ATP is required for the binding of precursor proteins to chloroplasts. J. Biol. Chem. **264,** 6724-6729.

Pain, D., and Blobel, G. (1987). Protein import into chloroplasts requires a chloroplast ATPase. Proc. Natl. Acad. Sci. USA **84,** 3288-3292.

Palmer, J.D., and Delwiche, C.F. (1996). Second-hand chloroplasts and the case of the disappearing nucleus. Proc. Natl. Acad. Sci. USA **93,** 7432-7435.

Perry, S.E., and Keegstra, K. (1994). Envelope membrane proteins that interact with chloroplastic precursor proteins. Plant Cell **6,** 93-105.

Pilon, M., Wienk, H., Sips, W., de Swaaf, M., Talboom, I., van 't Hof, R., de Korte-Kool, G., Demel, R., Weisbeek, P., and de Kruijff, B. (1995). Functional domains of the ferredoxin transit sequence involved in chloroplast import. J. Biol. Chem. **270,** 3882-3893.

Pinnaduwage, P., and Bruce, B.D. (1996). In vitro interaction between a chloroplast transit peptide and chloroplast outer envelope lipids is sequence-specific and lipid class-dependent. J. Biol. Chem. **271,** 32907-32915.

Porra, R.J., Thompson, W.A., and Kriedemann, P.E. (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls *a* and *b* extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. Biochim. Biophys. Acta **975,** 384-394.

Qbadou, S., Tien, R., Soll, J., and Schleiff, E. (2003). Membrane insertion of the chloroplast outer envelope protein, Toc34: constrains for insertion and topology. J Cell Sci **116,** 837-846.

Race, H.L., Herrmann, R.G., and Martin, W. (1999). Why have organelles retained genomes? Trends Genet. **15,** 364-370.

Rawyler, A., Meylan, M., and Siegenthaler, P.A. (1992). Galactolipid export from envelope to thylakoid membranes in intact chloroplasts. I. Characterization and involvement in thylakoid lipid asymmetry. Biochim. Biophys. Acta **1104,** 331-341.

Reiss, B., Wasmann, C.C., Schell, J., and Bohnert, H.J. (1989). Effect of mutations on the binding and translocation functions of a chloroplast transit peptide. Proc. Natl. Acad. Sci. USA **86,** 886-890.

Reiter, R.S., Coomber, S.A., Bourett, T.M., Bartley, G.E., and Scolnik, P.A. (1994). Control of leaf and chloroplast development by the Arabidopsis gene pale cress. Plant Cell **6,** 1253-1264.

Rensink, W.A., Schnell, D.J., and Weisbeek, P.J. (2000). The transit sequence of ferredoxin contains different domains for translocation across the outer and inner membrane of the chloroplast envelope. J Biol Chem **275,** 10265-10271.

Reumann, S., Davila-Aponte, J., and Keegstra, K. (1999). The evolutionary origin of the protein-translocating channel of chloroplastic envelope membranes: identification of a cyanobacterial homolog. Proc. Natl. Acad. Sci. USA **96,** 784-789.

Rial, D.V., Arakaki, A.K., and Ceccarelli, E.A. (2000). Interaction of the targeting sequence of chloroplast precursors with Hsp70 molecular chaperones. Eur. J. Biochem. **267,** 6239-6248.

Rial, D.V., Lombardo, V.A., Ceccarelli, E.A., and Ottado, J. (2002). The import of ferredoxin-NADP+ reductase precursor into chloroplasts is modulated by the region

between the transit peptide and the mature core of the protein. Eur J Biochem **269,** 5431-5439.

Richter, S., and Lamppa, G.K. (2003). Structural properties of the chloroplast stromal processing peptidase required for its function in transit peptide removal. J. Biol. Chem. **278,** 39497-39502.

Rios, G., Lossow, A., Hertel, B., Breuer, F., Schaefer, S., Broich, M., Kleinow, T., Jasik, J., Winter, J., Ferrando, A., Farras, R., Panicot, M., Henriques, R., Mariaux, J.B., Oberschall, A., Molnar, G., Berendzen, K., Shukla, V., Lafos, M., Koncz, Z., Redei, G.P., Schell, J., and Koncz, C. (2002). Rapid identification of Arabidopsis insertion mutants by non-radioactive detection of T-DNA tagged genes. Plant J **32,** 243-253.

Robinson, C., and Ellis, R.J. (1985). Transport of proteins into chloroplasts. The effect of incorporation of amino acid analogues on the import and processing of chloroplast polypeptides. Eur. J. Biochem. **152,** 67-73.

Robinson, C., Thompson, S.J., and Woolhead, C. (2001). Multiple pathways used for the targeting of thylakoid proteins in chloroplasts. Traffic **2,** 245-251.

Rodriguez, C.M., Freire, M.A., Camilleri, C., and Robaglia, C. (1998). The *Arabidopsis thaliana* cDNAs coding for eIF4E and eIF(iso)4E are not functionally equivalent for yeast complementation and are differentially expressed during plant development. Plant J. **13,** 465-473.

Sabatini, D.D., Kreibich, G., Morimoto, T., and Adesnik, M. (1982). Mechanisms for the incorporation of proteins in membranes and organelles. J. Cell Biol. **92,** 1-22.

Schatz, G., and Dobberstein, B. (1996). Common principles of protein translocation across membranes. Science **271,** 1519-1526.

Schleiff, E., and Klosgen, R.B. (2001). Without a little help from 'my' friends: direct insertion of proteins into chloroplast membranes? Biochim. Biophys. Acta **1541,** 22-33.

Schleiff, E., Jelic, M., and Soll, J. (2003a). A GTP-driven motor moves proteins across the outer envelope of chloroplasts. Proc. Natl. Acad. Sci. USA **100,** 4604-4609.

Schleiff, E., Soll, J., Kuchler, M., Kuhlbrandt, W., and Harrer, R. (2003b). Characterization of the translocon of the outer envelope of chloroplasts. J. Cell Biol. **160,** 541-551.

Schleiff, E., Soll, J., Sveshnikova, N., Tien, R., Wright, S., Dabney-Smith, C., Subramanian, C., and Bruce, B.D. (2002). Structural and guanosine triphosphate/diphosphate requirements for transit peptide recognition by the cytosolic

domain of the chloroplast outer envelope receptor, Toc34. Biochemistry **41,** 1934-1946.

Schneider, K., Wells, B., Dolan, L., and Roberts, K. (1997). Structural and genetic analysis of epidermal cell differentiation in Arabidopsis primary roots. Development **124,** 1789-1798.

Schnell, D.J., and Blobel, G. (1993). Identification of intermediates in the pathway of protein import into chloroplasts and their localization to envelope contact sites. J. Cell Biol. **120,** 103-115.

Schnell, D.J., Kessler, F., and Blobel, G. (1994). Isolation of components of the chloroplast protein import machinery. Science **266,** 1007-1012.

Schnell, D.J., Blobel, G., Keegstra, K., Kessler, F., Ko, K., and Soll, J. (1997). A consensus nomenclature for the protein-import components of the chloroplast envelope. Trends Cell Biol. **7,** 303-304.

Seedorf, M., and Soll, J. (1995). Copper chloride, an inhibitor of protein import into chloroplasts. FEBS Lett. **367,** 19-22.

Seedorf, M., Waegemann, K., and Soll, J. (1995). A constituent of the chloroplast import complex represents a new type of GTP-binding protein. Plant J. **7,** 401-411.

Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B., Mitzel, T., Katagiri, F., Glazebrook, J., Law, M., and Goff, S.A. (2002). A high-throughput *Arabidopsis* reverse genetics system. Plant Cell **14,** 2985-2994.

Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y., and Ishikawa, H. (2000). Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS. Nature **407,** 81-86.

Smeekens, S., Bauerle, C., Hageman, J., Keegstra, K., and Weisbeek, P. (1986). The role of the transit peptide in the routing of precursors toward different chloroplast compartments. Cell **46,** 365-375.

Smith, M.D., Hiltbrunner, A., Kessler, F., and Schnell, D.J. (2002). The targeting of the atToc159 preprotein receptor to the chloroplast outer membrane is mediated by its GTPase domain and is regulated by GTP. J. Cell Biol. **159,** 833-843.

Sohrt, K., and Soll, J. (2000). Toc64, a new component of the protein translocon of chloroplasts. J. Cell Biol. **148,** 1213-1221.

Somerville, C., and Koornneef, M. (2002). A fortunate choice: the history of Arabidopsis as a model plant. Nat. Rev. Genet. **3,** 883-889.

Stahl, T., Glockmann, C., Soll, J., and Heins, L. (1999). Tic40, a new "old" subunit of the chloroplast protein import translocon. J. Biol. Chem. **274,** 37467-37472.

Subramanian, C., Ivey, R., 3rd, and Bruce, B.D. (2001). Cytometric analysis of an epitope-tagged transit peptide bound to the chloroplast translocation apparatus. Plant J. **25,** 349-363.

Summer, E.J., and Cline, K. (1999). Red bell pepper chromoplasts exhibit in vitro import competency and membrane targeting of passenger proteins from the thylakoidal sec and DeltapH pathways but not the chloroplast signal recognition particle pathway. Plant Physiol. **119,** 575-584.

Sun, C.W., Chen, L.J., Lin, L.C., and Li, H.M. (2001). Leaf-specific upregulation of chloroplast translocon genes by a CCT motif-containing protein, CIA 2. Plant Cell **13,** 2053-2061.

Sun, Y.J., Forouhar, F., Li Hm, H.M., Tu, S.L., Yeh, Y.H., Kao, S., Shr, H.L., Chou, C.C., Chen, C., and Hsiao, C.D. (2002). Crystal structure of pea Toc34, a novel GTPase of the chloroplast protein translocon. Nat. Struct. Biol. **9,** 95-100.

Sveshnikova, N., Soll, J., and Schleiff, E. (2000a). Toc34 is a preprotein receptor regulated by GTP and phosphorylation. Proc. Natl. Acad. Sci. USA **97,** 4973-4978.

Sveshnikova, N., Grimm, R., Soll, J., and Schleiff, E. (2000b). Topology studies of the chloroplast protein import channel Toc75. Biol. Chem. **381,** 687-693.

Theg, S.M., Bauerle, C., Olsen, L.J., Selman, B.R., and Keegstra, K. (1989). Internal ATP is the only energy requirement for the translocation of precursor proteins across chloroplastic membranes. J. Biol. Chem. **264,** 6730-6736.

Tissier, A.F., Marillonnet, S., Klimyuk, V., Patel, K., Torres, M.A., Murphy, G., and Jones, J.D. (1999). Multiple independent defective suppressor-mutator transposon insertions in *Arabidopsis*: a tool for functional genomics. Plant Cell **11,** 1841-1852.

Tranel, P.J., and Keegstra, K. (1996). A novel, bipartite transit peptide targets OEP75 to the outer membrane of the chloroplastic envelope. Plant Cell **8,** 2093-2104.

Tranel, P.J., Froehlich, J., Goyal, A., and Keegstra, K. (1995). A component of the chloroplastic protein import apparatus is targeted to the outer envelope membrane via a novel pathway. EMBO J. **14,** 2436-2446.

Tsai, L.Y., Tu, S.L., and Li, H.M. (1999). Insertion of atToc34 into the chloroplastic outer membrane is assisted by at least two proteinaceous components in the import system. J. Biol. Chem. **274,** 18735-18740.

Tsugeki, R., Kochieva, E.Z., and Fedoroff, N.V. (1996). A transposon insertion in the *Arabidopsis* SSR16 gene causes an embryo-defective lethal mutation. Plant J. **10,** 479-489.

Tu, S.L., and Li, H.M. (2000). Insertion of OEP14 into the outer envelope membrane is mediated by proteinaceous components of chloroplasts. Plant Cell **12,** 1951-1960.

Uwer, U., Willmitzer, L., and Altmann, T. (1998). Inactivation of a glycyl-tRNA synthetase leads to an arrest in plant embryo development. Plant Cell **10,** 1277-1294.

van den Wijngaard, P.W., and Vredenberg, W.J. (1997). A 50-picosiemens anion channel of the chloroplast envelope is involved in chloroplast protein import. J. Biol. Chem. **272,** 29430-29433.

van den Wijngaard, P.W., Demmers, J.A., Thompson, S.J., Wienk, H.L., de Kruijff, B., and Vredenberg, W.J. (2000). Further analysis of the involvement of the envelope anion channel PIRAC in chloroplast protein import. Eur. J. Biochem. **267,** 3812-3817.

Waegemann, K., Paulsen, H., and Soll, J. (1990). Translocation of proteins into isolated chloroplasts requires cytosolic factors to obtain import competence. FEBS **261,** 89-92.

Wallas, T.R., Smith, M.D., Sanchez-Nieto, S., and Schnell, D.J. (2003). The roles of toc34 and toc75 in targeting the toc159 preprotein receptor to chloroplasts. J. Biol. Chem. **278,** 44289-44297.

Wan, J., Blakeley, S.D., Dennis, D.T., and Ko, K. (1996). Transit peptides play a major role in the preferential import of proteins into leucoplasts and chloroplasts. J. Biol. Chem. **271,** 31227-31233.

Watson, M.B., Emory, K.K., Piatak, R.M., and Malmberg, R.L. (1998). Arginine decarboxylase (polyamine synthesis) mutants of Arabidopsis thaliana exhibit altered root growth. Plant J. **13,** 231-239.

Weeden, N.F. (1981). Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. J. Mol. Evol. **17,** 133-139.

Weibel, P., Hiltbrunner, A., Brand, L., and Kessler, F. (2003). Dimerization of Toc-GTPases at the chloroplast protein import machinery. J. Biol. Chem. **278,** 37321-37329.

Wienk, H.L., Wechselberger, R.W., Czisch, M., and de Kruijff, B. (2000). Structure, dynamics, and insertion of a chloroplast targeting peptide in mixed micelles. Biochemistry **39,** 8219-8227.

Woolhead, C., Bolhuis, A., and Robinson, C. (2000). Novel mechanisms for the targeting of proteins into and across chloroplast membranes. Biochem. Soc. Trans. **28,** 491-494.

Wu, C., and Ko, K. (1993). Identification of an uncleavable targeting signal in the 70-kilodalton spinach chloroplast outer envelope membrane protein. J. Biol. Chem. **268,** 19384-19391.

Wu, C., Seibert, F.S., and Ko, K. (1994). Identification of chloroplast envelope proteins in close physical proximity to a partially translocated chimeric precursor protein. J. Biol. Chem. **269,** 32264-32271.

Wu, K., Rooney, M.F., and Ferl, R.J. (1997). The Arabidopsis 14-3-3 multigene family. Plant Physiol. **114,** 1421-1431.

Young, M.E., Keegstra, K., and Froehlich, J.E. (1999). GTP promotes the formation of early-import intermediates but is not required during the translocation step of protein import into chloroplasts. Plant Physiol. **121,** 237-244.

Zhang, X.P., and Glaser, E. (2002). Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. Trends Plant Sci **7,** 14-21.