A Molecular-Genetic Study of the Arabidopsis Toc75 Gene Family[1]

Amy Baldwin[3], Anthony Wardle[2], Ramesh Patel[2], Penny Dudley, Soon Ki Park[4], David Twell, Kentaro Inoue, and Paul Jarvis*

Department of Biology, University of Leicester, Leicester LE1 7RH, United Kingdom (A.B., A.W., R.P., P.D., S.K.P., D.T., P.J.); and Department of Plant Sciences, University of California, Davis, California 95616 (K.I.)

Toc75 (translocon at the outer envelope membrane of chloroplasts, 75 kD) is the protein translocation channel at the outer envelope membrane of plastids and was first identified in pea (*Pisum sativum*) using biochemical approaches. The Arabidopsis (*Arabidopsis thaliana*) genome contains three Toc75-related sequences, termed *atTOC75-I*, *atTOC75-III*, and *atTOC75-IV*, which we studied using a range of molecular, genetic, and biochemical techniques. Expression of *atTOC75-III* is strongly regulated and at its highest level in young, rapidly expanding tissues. By contrast, *atTOC75-IV* is expressed uniformly throughout development and at a much lower level than *atTOC75-III*. The third sequence, *atTOC75-I*, is a pseudogene that is not expressed due to a *gypsy/Ty3* transposon insertion in exon 1, and numerous nonsense, frame-shift, and splice-junction mutations. The expressed genes, *atTOC75-III* and *atTOC75-IV*, both encode integral envelope membrane proteins. Unlike atToc75-III, the smaller atToc75-IV protein is not processed upon targeting to the envelope, and its insertion does not require ATP at high concentrations. The *atTOC75-III* gene is essential for viability, since homozygous atToc75-III knockout mutants (termed *toc75-III*) could not be identified, and aborted seeds were observed at a frequency of approximately 25% in the siliques of self-pollinated *toc75-III* heterozygotes. Homozygous *toc75-III* embryos were found to abort at the two-cell stage. Homozygous atToc75-IV knockout plants (termed *toc75-IV*) displayed no obvious visible phenotypes. However, structural abnormalities were observed in the etioplasts of *toc75-IV* seedlings and *atTOC75-IV* overexpressing lines, and *toc75-IV* plants were less efficient at deetiolation than wild type. These results suggest some role for atToc75-IV during growth in the dark.

The majority of plastid proteins are translated on cytosolic ribosomes and subsequently imported into plastids (Keegstra and Cline, 1999; Chen et al., 2000; Hiltbrunner et al., 2001a; Jarvis and Soll, 2001; Jarvis and Robinson, 2004). An amino-terminal transit peptide directs each of these proteins specifically to the plastid. Upon arrival in the stroma, the transit peptide is cleaved and the mature protein is either folded into its final conformation or targeted to another compartment of the plastid (Keegstra and Cline, 1999; Jarvis and Robinson, 2004). Preproteins are translocated through the plastid double membrane envelope by two membrane-bound 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 include Toc34, Toc75, and Toc159, which were first identified in pea (*Pisum sativum*) using biochemical approaches (Hirsch et al., 1994; Perry and Keegstra, 1994; Schnell et al., 1994; Seedorf et al., 1995; Tranel et al., 1995). These three proteins make up the core Toc complex, which was recently characterized with respect to structure and component stoichiometry (Schleiff et al., 2003b). Toc34 and Toc159 are related GTPases that share considerable sequence identity within their GTP-binding domains (Hirsch et al., 1994; Kessler et al., 1994; Seedorf et al., 1995) and which mediate preprotein recognition at the chloroplast surface (Perry and Keegstra, 1994; Kouranov and Schnell, 1997; Sveshnikova et al., 2000b; Becker et al., 2004). Following recognition by Toc159 and Toc34, precursor proteins are transferred to Toc75, the protein that forms the outer envelope translocation channel (Schnell et al., 1994; Tranel et al., 1995; Hinnah et al., 2002). Simultaneous translocation through the Toc and Tic complexes, in an energy-consuming process, subsequently occurs (Keegstra and Cline, 1999; Chen et al., 2000; Hiltbrunner et al., 2001a; Jarvis and Soll, 2001; Jarvis and Robinson, 2004).

Structural and electrophysiological studies on heterologously expressed pea Toc75 (psToc75) reconstituted into liposomes demonstrated that the protein forms

[1] This work was supported by the University of Leicester (Ph.D. studentship to A.B.), by the United States Department of Agriculture Cooperative State Research, Education, and Extension Service (grant no. 2003–02860 to K.I.), by the Royal Society (Rosenheim Research Fellowship to P.J.), and by the Biotechnology and Biological Sciences Research Council (grant nos. 91/C12976, 91/P12928, and 91/C18638 to P.J.).

[2] These authors contributed equally to the paper.

[3] Present address: Department of Plant Sciences, University of California, One Shields Avenue, Davis, CA 95616.

[4] Present address: Division of Plant Bioscience, Kyungpook National University, Daegu 702–701, Korea.

* Corresponding author; e-mail rpj3@le.ac.uk; fax 44–116–252–3330.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.105.063289.

a β-barrel channel with a pore size of approximately 14 to 26 Å (Hinnah et al., 2002). Thus, preproteins must have a completely or partially unfolded conformation to pass through the pore. Topological studies on psToc75 employing proteolytic digestion, amino acid sequencing, hydrophobicity, and computer modeling data suggested that Toc75 possesses either 16 β-strands (Sveshnikova et al., 2000a) or 18 β-strands (Schleiff et al., 2003a). Interestingly, it seems that the role of Toc75 is not confined to preprotein conductance. Electrophysiology measurements on reconstituted Toc75 revealed that the protein is able to specifically recognize a transit peptide without the aid of other Toc components (Hinnah et al., 2002).

Evidence that Toc75-related proteins are present in plant species other than pea was provided by Summer and Cline (1999), who used an anti-psToc75 antibody to identify a 75-kD protein in chromoplasts from red bell pepper. Since then, Toc75 homologs have been identified in many different plants, including both monocotyledonous and dicotyledonous species (Dávila-Aponte et al., 2003). Toc75 can also be found in different plastid types and throughout plant development (Tranel et al., 1995; Summer and Cline, 1999; Dávila-Aponte et al., 2003). In pea, Toc75 is expressed at a high level throughout the photosynthetic tissues and at a lower level in stems and roots (Tranel et al., 1995). Thus, Toc75 is thought to be a constitutive component of the protein import pathway in diverse plastid types and in all higher plant species.

The sequencing of the Arabidopsis (*Arabidopsis thaliana*) genome has revealed that this species contains multiple, homologous isoforms of many of the Toc and Tic components originally identified in pea (Jackson-Constan and Keegstra, 2001). For example, the Arabidopsis genome encodes two Toc34 homologs, termed atToc34 and atToc33 (Jarvis et al., 1998; Gutensohn et al., 2000), and four Toc159 homologs, termed atToc159, atToc132, atToc120, and atToc90 (Bauer et al., 2000; Hiltbrunner et al., 2001a; Ivanova et al., 2004; Kubis et al., 2004). Reverse-genetic strategies have been used to examine the roles of these different isoforms, and it appears that they do exhibit some degree of functional specialization. In the case of the Toc34 family, the atToc33 isoform is preferentially involved in the import of photosynthetic precursors, while the atToc34 isoform is most likely involved in the import of nonphotosynthetic precursors (Gutensohn et al., 2000; Kubis et al., 2003; Constan et al., 2004). Interestingly, there are three different Toc75-related sequences in the Arabidopsis genome, termed *atTOC75-I*, *atTOC75-III*, and *atTOC75-IV* according to their chromosomal locations (Jackson-Constan and Keegstra, 2001). In this study, we used molecular, genetic, and biochemical techniques to investigate the functions of these three sequences. The much more divergent gene, *atTOC75-V/AtOEP80*, was not included in our analysis, since it does not appear to be a true member of the Toc75 family (Eckart et al., 2002; Inoue and Potter, 2004).

RESULTS

Expression and Structure of the Different Arabidopsis Toc75 Gene Family Members

The sequencing of the Arabidopsis genome has enabled the identification, in silico, of putative homologs of proteins originally identified in other species. For example, it was previously reported that there are three psToc75-related sequences in the Arabidopsis genome: *atTOC75-I*, *atTOC75-III*, and *atTOC75-IV* (Jackson-Constan and Keegstra, 2001). The *atTOC75-III* gene is known to be expressed since there are currently 24 different *atTOC75-III* expressed sequence tags (ESTs) present in the databases. Conversely, there are no *atTOC75-I* or *atTOC75-IV* ESTs in the databases (although a partial cDNA sequence was recently released for *atTOC75-IV*; accession no. BX827624). To assess expression of the putative *atTOC75-I* gene (no. At1g35860), RNA isolated from whole, wild-type Arabidopsis seedlings was analyzed extensively by reverse transcription (RT)-PCR, using six different primer combinations (Fig. 1A). In each case, no evidence for *atTOC75-I* gene expression could be detected (data not shown).

To determine the reasons underlying the apparent inactivity of this putative gene, we conducted detailed studies of the surrounding genomic sequence. Our analysis revealed that homology with psToc75 extends outside of the predicted exons of At1g35860 (Fig. 1A) and even into the adjacent predicted gene, At1g35880, 5.8 kb upstream (Fig. 1A). These observations strongly suggested that the gene had been incorrectly annotated and led us to conduct a manual annotation of *atTOC75-I* (deposited in GenBank; accession no. BK005428). The corrected annotation revealed a 5.4-kb *gypsy/Ty3*-related retrotransposon inserted at the 5′ end of the gene (Fig. 1A); this insertion incorporates 395-bp and 392-bp long terminal repeats that share 90% nucleotide sequence identity and is flanked by a perfect 5-bp target site duplication. A related element was previously identified at the *waxy* locus in maize (*Zea mays*; Purugganan and Wessler, 1994). Our analysis also revealed the presence of numerous nonsense, frame-shift, and splice-junction mutations, as well as a 75-bp insertion and a 48-bp direct repeat in exon 1 (Fig. 1A). This evidence, together with the absence of *atTOC75-I* ESTs and our inability to detect *atTOC75-I* expression by RT-PCR, led us to conclude that *atTOC75-I* is a pseudogene.

Next, we investigated the activity and structure of the *atTOC75-IV* gene (no. At4g09080). Expression analysis by RT-PCR provided clear evidence that the gene is active (Fig. 2A), and the nucleotide sequence of the PCR product obtained revealed that *atTOC75-IV* transcripts undergo accurate intron splicing (data not shown). Since no publicly available EST or cDNA clones were available for *atTOC75-IV* when we started these studies, we initiated experiments to identify our own *atTOC75-IV* cDNA clone. Attempts to identify such

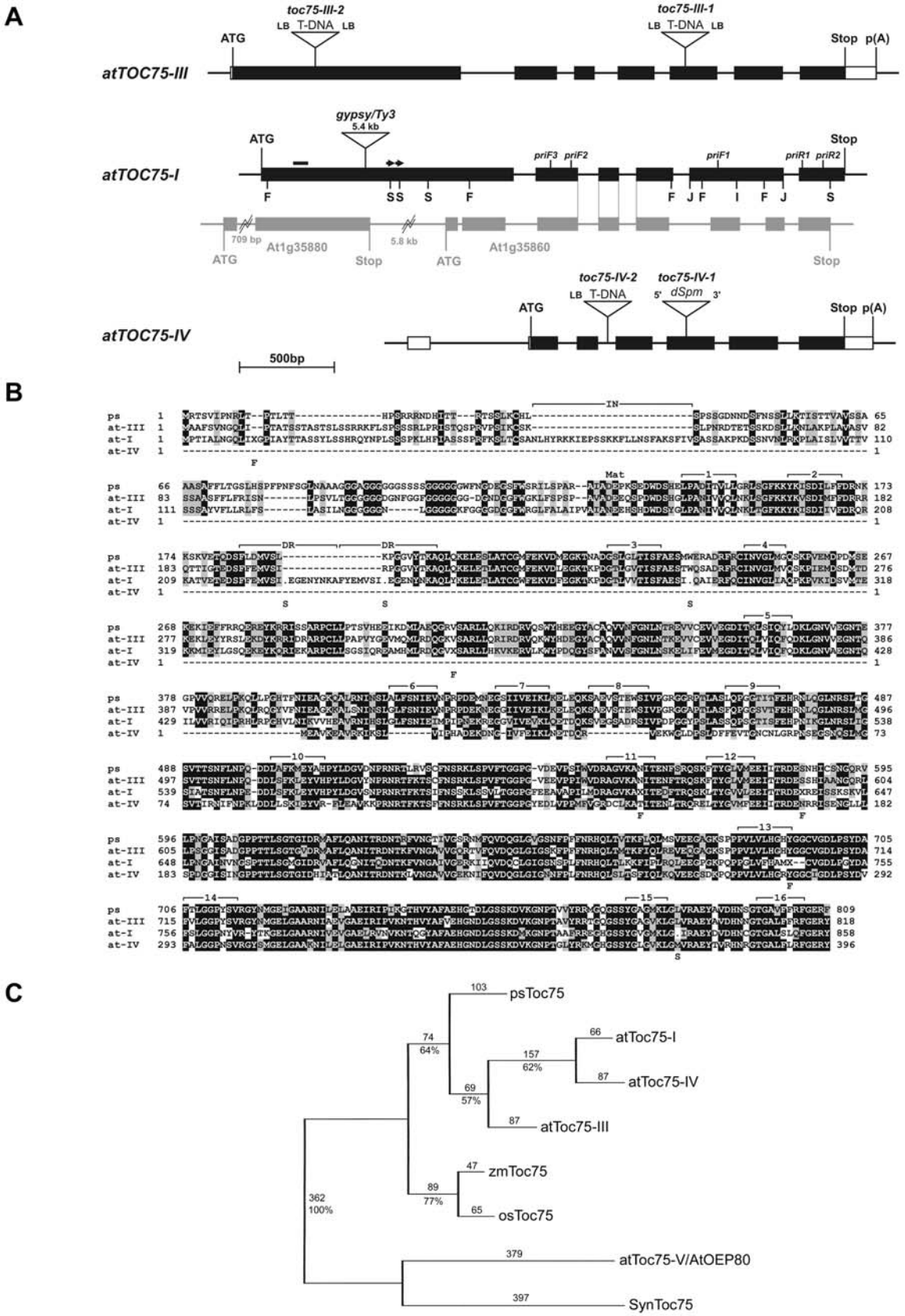

Figure 1. Structural characteristics of the Arabidopsis Toc75 gene family. A, Schematic diagrams depicting the three Arabidopsis Toc75-related sequences. Protein-coding exons are represented by black boxes, and untranslated regions are represented by

a clone by library screening were unsuccessful, presumably due to the low level of expression of the gene, so we instead generated a full-length clone using a combination of RACE-PCR and RT-PCR. The structure of the transcript at its 5′ and 3′ ends was determined by RACE-PCR, and this information was used to design primers for amplification of a full-length cDNA by RT-PCR. The RT-PCR product was cloned and sequenced (information deposited in GenBank; accession nos. AY585655 and AAT08975), and its structure is illustrated in Figure 1A. This experimentally determined structure differs slightly from the predicted structure given as part of the genome sequence annotation (accession no. NM_116977; Inoue and Potter, 2004).

The structure of the at*TOC75-III* gene (no. At3g46740) was determined by sequencing a publicly available EST clone, APZL59c10R, and is shown in Figure 1A. The intron-exon boundaries were found to match exactly those predicted as part of the genome sequence annotation (accession no. AL096859), and the cDNA was essentially the same as a recently released full-length cDNA sequence (accession no. AY127014), except that the 5′ and 3′ untranslated regions were each slightly shorter.

The structures of the three Arabidopsis Toc75-related sequences, shown schematically in Figure 1A, are remarkably similar. All three genes have essentially the same intron-exon boundaries, with only two exceptions: exon 1 is absent from *atTOC75-IV*, and the penultimate intron is absent from the *atTOC75-I* pseudogene. The three Arabidopsis homologs are also very similar to each other, and to psToc75, at the amino acid level (Fig. 1B). Percentage sequence identities shared between the Arabidopsis proteins and the pea protein are as follows: 73% (atToc75-III), 60% (atToc75-IV), and 55% (atToc75-I). The 5′ truncation of the *atTOC75-IV* gene means that the atToc75-IV protein includes only eight of the 16 transmembrane domains (nos. 7 and 10–16; Fig. 1B) predicted to be present in psToc75 (Sveshnikova et al., 2000a).

Phylogenetic Analysis

To determine the evolutionary relationships between the different Arabidopsis Toc75 homologs, we conducted a phylogenetic analysis using amino acid sequences from several different species. In this analysis, we included all three Arabidopsis sequences, the psToc75 sequence, two Toc75 sequences from monocotyledonous species, as well as distantly related sequences from *Synechocystis* PCC 6803 (SynToc75) and Arabidopsis (atToc75-V/AtOEP80); SynToc75 is a cyanobacterial protein that shares only approximately 21% to 22% sequence identity with psToc75 (Bölter et al., 1998; Reumann et al., 1999), and atToc75-V/AtOEP80 shares limited sequence similarity with canonical Toc75 proteins and SynToc75 (Eckart et al., 2002; Inoue and Potter, 2004). Our analysis produced a single most parsimonious tree of length 1,872, containing three main clades (Fig. 1C). The outgroup containing SynToc75 and atToc75-V/AtOEP80 comprises 362 mutations that are present in both proteins but not in Toc75 proteins from the other two clades. With the exception of atToc75-V/AtOEP80, the Toc75 proteins from dicotyledonous species were grouped together in the same clade. Toc75 proteins from the two monocotyledonous species, maize and rice (*Oryza sativa*), formed a separate clade with 89 distinct mutations. The Arabidopsis Toc75 proteins formed a separate subclade, excluding psToc75, which was supported by a bootstrap value of 57%. Applying a cutoff value of 50%, all clades were supported by bootstrap analysis.

The phylogeny suggests that the Arabidopsis Toc75 gene family formed after the pea and Arabidopsis species diverged and implies that pea and other species do not necessarily contain genes closely related

Figure 1. (*Continued.*)
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 shown. The gray diagram underneath *atTOC75-I* illustrates how the pseudogene is currently annotated as two separate genes (At1g35880 and At1g35860) in the GenBank database; vertical, dotted lines indicate four putative splice junctions that are identical in the two gene models. 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. 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 gray. The location of the first Asp residue of mature psToc75 is shown (Mat), as are the 16 predicted transmembrane domains of psToc75 (Tranel et al., 1995; Sveshnikova et al., 2000a). 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, AAM83239); atToc75-I (BK005428); atToc75-IV (AY585655, AAT08975); atToc75-V/AtOEP80 (At5g19620, NP_568378); zmToc75 (AY106148); osToc75 (AK070010); SynToc75 (slr1227, NP_440832). Species of origin is indicated as follows: pea, ps; Arabidopsis, at; maize, zm; rice, os; *Synechocystis* PCC 6803, Syn. The cyanobacterial protein, SynToc75, was used as the outgroup.

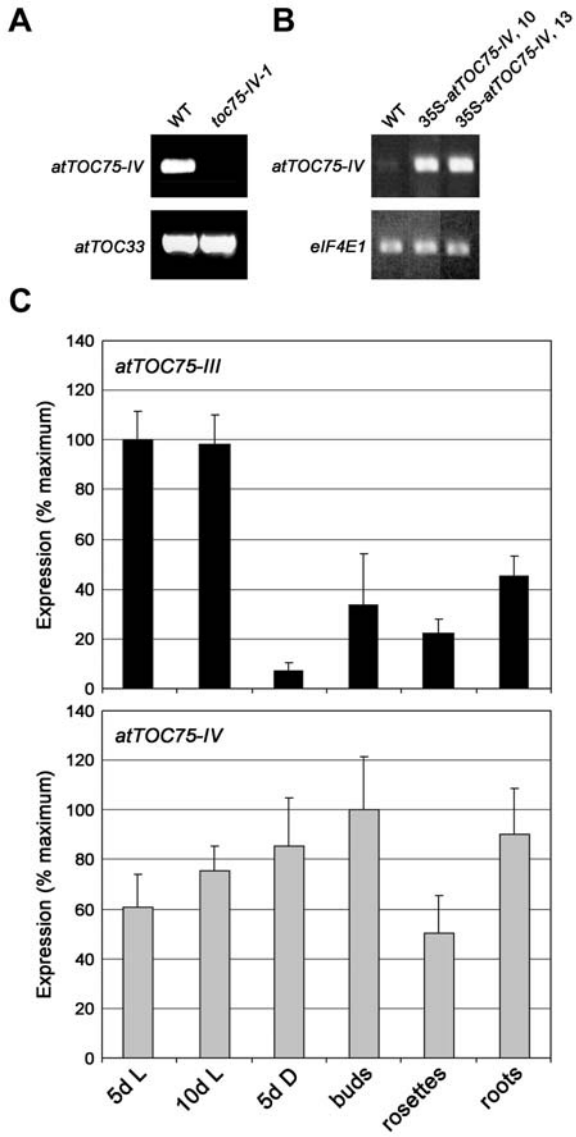

Figure 2. Expression studies on the Arabidopsis Toc75 genes. A and B, Analysis of *atTOC75-IV* expression in mutant and transgenic lines by RT-PCR. Total RNA samples isolated from the *toc75-IV-1* mutant (A) and from two independent 35S-*atTOC75-IV* transgenic lines (B) were analyzed by RT-PCR, along with corresponding wild-type samples, using primers specific for the indicated genes (*atTOC75-IV*, *atTOC33*, and *eIF4E1*). In each case, the rosette leaves of a single mature plant were analyzed; the 35S-*atTOC75-IV* plants used were both homozygous for the transgene. 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. PCR products were resolved by agarose gel electrophoresis and stained with ethidium bromide. 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 analyzed by semiquantitative RT-PCR. RNA was isolated from wild-type seedlings grown in vitro for 5 or 10 d in the light (5 d L and 10 d L, respectively), or 5 d in the dark (5 d D), and from three different tissues of 28-d-old wild-type plants grown on soil (flower buds, rosette leaves, and roots). Amplifications were conducted under nonsaturating conditions using gene-specific *atTOC75-III*, *atTOC75-IV*, and *eIF4E1* primers, and the products were quantified by hybridization with corresponding ^{32}P-labeled cDNA probes. The data for *atTOC75-III* and *atTOC75-IV* 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.

to either *atTOC75-IV* or *atTOC75-I*. After speciation, an Arabidopsis *TOC75* progenitor duplicated to form *atTOC75-III* and another *TOC75* gene. Duplication of the latter then created progenitors of *atTOC75-I* and *atTOC75-IV*. Inactivation of *atTOC75-I* (due to the insertion of the *gypsy/Ty3* retrotransposon or the accumulation of other mutations), and the truncation of *atTOC75-IV*, may have occurred after the final duplication event. Alternatively, retrotransposon insertion may have occurred prior to the duplication event that gave rise to *atTOC75-I* and *atTOC75-IV*.

Expression Profiles of the Arabidopsis Toc75 Genes

Since the members of other Arabidopsis Toc component gene families have been shown to exhibit differential expression patterns and some degree of functional specialization (Bauer et al., 2000; Gutensohn et al., 2000; Kubis et al., 2003; Constan et al., 2004; Ivanova et al., 2004; Kubis et al., 2004), it seemed possible that similar specialization might exist within the Arabidopsis Toc75 family. In an initial attempt to gain insight into the roles of the two expressed Arabidopsis Toc75 genes, *atTOC75-III* and *atTOC75-IV*, we analyzed their expression patterns using semiquantitative RT-PCR. The use of RT-PCR was necessary since the *atTOC75-IV* transcript could not be detected by RNA gel-blot analysis (data not shown). Expression data for the Arabidopsis Toc75 genes were normalized using similar data for a uniformly expressed control gene (translation elongation factor, *eIF4E1*; Rodriguez et al., 1998) and then expressed as a percentage of the maximum value obtained in each case (Fig. 2C).

The data shown in Figure 2C do not provide any information on the relative levels of expression of the two Toc75 genes. However, it is clear that *atTOC75-III* is expressed at much higher levels than *atTOC75-IV* for several reasons. First of all, amplification of comparable amounts of the *atTOC75-III* and *atTOC75-IV* transcripts by RT-PCR required 20 and 25 cycles, respectively (Fig. 2; data not shown). Assuming that each cycle of amplification produces a 2-fold increase in product yield, this difference in cycle number equates to a 32-fold difference in template concentration. These data are broadly consistent with recently reported microarray data (Vojta et al., 2004); although *atTOC75-IV* expression was on the limit of detection in this study, the data indicated a difference in expression between the two genes of approximately 60-fold (Vojta et al., 2004). Finally, while there are 24 different ESTs currently available for the *atTOC75-III* gene, none is available for *atTOC75-IV*.

The expression of *atTOC75-III* was found to be strong in young, rapidly dividing photosynthetic tissues and significantly weaker in mature or slow

growing tissues, such as 28-d-old rosette leaves, and roots (Fig. 2C). Unexpectedly, *atTOC75-III* expression was found to be particularly low in 5-d-old etiolated plants. This suggests that atToc75-IV, the only other close homolog of psToc75 that is expressed in Arabidopsis, may play a relatively more important role in etioplast biogenesis. The expression of *atTOC75-IV* is more uniform than that of *atTOC75-III*; while the expression of *atTOC75-III* varied 10-fold across the samples tested, *atTOC75-IV* expression varied just 2-fold (Fig. 2C). Like *atTOC75-III*, the pea Toc75 gene is also expressed most strongly in young, rapidly dividing tissues (Tranel et al., 1995). Additionally, immunoblot data indicated that psToc75 protein accumulates to a lesser degree in roots than in green tissues (Tranel et al., 1995). These various data on the level and pattern of expression of the Toc75 genes, in combination with the structural characteristics discussed earlier, strongly suggest that *atTOC75-III* encodes the main ortholog of psToc75 in Arabidopsis.

atToc75-IV Is an Integral Protein of the Plastid Envelope

Pea Toc75 is targeted to the chloroplast outer envelope membrane by a bipartite, amino-terminal transit peptide (Tranel et al., 1995; Tranel and Keegstra, 1996). Thus, psToc75 targeting involves two different proteolytic processing steps and a transitional molecule of intermediate size (Tranel et al., 1995; Tranel and Keegstra, 1996). Similarly, atToc75-III is also imported into chloroplasts in a two-step process involving an intermediate (Inoue and Keegstra, 2003; Fig. 3A). Since *atTOC75-IV* lacks the large first exon present in *atTOC75-III* (Fig. 1A), which encodes the bipartite transit peptide, we analyzed the atToc75-IV amino acid sequence using 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 targeted in the same manner as atToc75-III and psToc75. It should be noted that it is not uncommon for outer envelope membrane proteins to be inserted into the membrane without the aid of a transit peptide (Schleiff and Klösgen, 2001). In fact, psToc75 and atToc75-III are unique in that they are the only outer envelope proteins known to utilize a transit peptide for insertion.

Common characteristics of outer envelope membrane proteins that lack a cleavable transit peptide are a small to moderate overall size and, in some cases, a well-defined hydrophobic region (Keegstra and Cline, 1999; Schleiff and Klösgen, 2001). Since atToc75-IV is relatively small (44 kD), we investigated the possibility that it is targeted by a similar mechanism. To this end, in vitro translated atToc75-IV was incubated with isolated pea chloroplasts under conditions favoring protein import. As expected, atToc75-IV efficiently associated with chloroplasts, and, in contrast with atToc75-III, was not processed to a lower molecular mass form during the experiment (Fig. 3A). Unlike the two Toc75 isoforms, a control protein that is normally targeted to peroxisomes (firefly luciferase) was not found to associate significantly with chloroplasts. Imported atToc75-IV was exclusively associated with the membrane fraction of chloroplasts and was resistant to extraction under high pH conditions, strongly suggesting that atToc75-IV is an integral membrane protein. By contrast, a peripheral protein of the inner envelope membrane, psTic22, was susceptible to extraction under alkaline conditions (Fig. 3A), as expected (Kouranov et al., 1998).

To further confirm the proper targeting of imported atToc75-IV, and to provide some clues about its suborganellar localization, we conducted postimport protease treatment experiments (Fig. 3, B–D). Four different precursors were imported into isolated pea chloroplasts, and then the chloroplasts were treated with three different concentrations of trypsin, or in the absence of trypsin; trypsin is a protease that is able to penetrate the outer envelope membrane and so gain access to proteins exposed in the intermembrane space, but not the inner envelope membrane (Cline et al., 1984; Jackson et al., 1998; Kouranov et al., 1999). Bands corresponding to imported proteins were quantified, and the data from two independent experiments are shown in Figure 3B. As expected, atToc75-III was almost completely degraded by trypsin (Tranel and Keegstra, 1996; Inoue and Potter, 2004), whereas tp110-N110 (a truncated derivative of psTic110 that was previously shown to insert into the inner envelope membrane; Jackson et al., 1998) was completely trypsin resistant (Fig. 3B). In contrast with atToc75-III, atToc75-IV was substantially resistant to trypsin treatment, to about the same extent as the intermembrane-space-exposed protein, psTic22 (Kouranov et al., 1999; Fig. 3B). Since the atToc75-IV translation product was not inherently trypsin resistant (Fig. 3C), and since the protease tolerance of the imported protein was dependent upon membrane integrity (Fig. 3D), these data again suggest that atToc75-IV was properly targeted to the chloroplasts in these assays. The fact that atToc75-IV was more trypsin-tolerant than atToc75-III may be interpreted in two different ways: either (1) atToc75-IV is present in the outer envelope membrane, like atToc75-III, but takes on a topology (or forms an assembly with other components) that makes it less accessible to trypsin; or (2) atToc75-IV is located inside of the outer envelope membrane, like psTic22 and psTic110. Given that atToc75-IV is an integral membrane protein (Fig. 3A) and yet is not trypsin-tolerant to the same extent as tp110-110N (Fig. 3B), we favor the former possibility.

To determine if the targeting of atToc75-IV occurs in similar fashion in Arabidopsis, we conducted similar import experiments with isolated Arabidopsis chloroplasts. Like atToc75-IV imported into pea chloroplasts (Fig. 3A), atToc75-IV imported into Arabidopsis chloroplasts (as well as endogenous atToc75-III) was exclusively associated with the membranes and completely resistant to extraction under high salt and high pH conditions (data not shown). In an additional

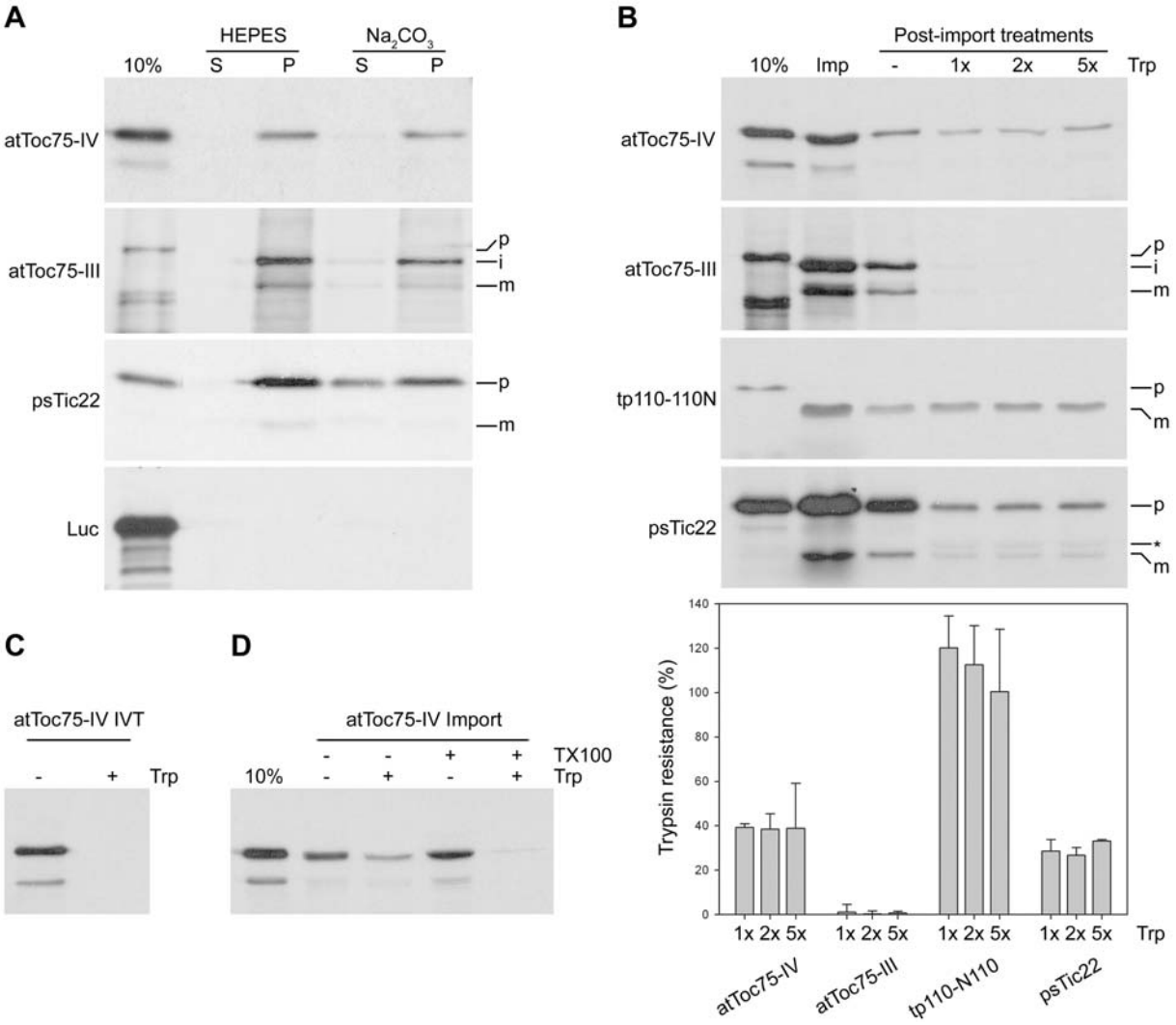

Figure 3. Import and localization studies on atToc75-IV using pea chloroplasts. A, High pH washes of imported proteins. In vitro translated, ^{35}S-labeled precursor proteins (atToc75-IV, atToc75-III, psTic22, and Luciferase, Luc) were incubated with isolated pea chloroplasts for 30 min under import conditions. Intact chloroplasts were reisolated and then treated with either 10 mM HEPES, 10 mM $MgCl_2$, pH 8.0 (bursting buffer), or 0.1 M Na_2CO_3 (high pH), before separation into pellet (P) and soluble (S) fractions by centrifugation. Samples were resolved by SDS-PAGE and visualized by fluorography. In each case, the first lane contains translation product equivalent to 10% of the amount added to each import reaction. The approximately 36-kD band below the main atToc75-IV translation product is presumably the result of translation initiation at Met-72, since this product was repeatedly observed in atToc75-IV translation mixtures. The precursor (p) and putative intermediate (i) and mature (m) forms of atToc75-III are indicated. B, Trypsin treatment of imported proteins. Import of four different precursors (atToc75-IV, atToc75-III, tp110-110N, and psTic22) was conducted as described in A. Reisolated chloroplasts (Imp) were then treated with three different concentrations of trypsin (1×, 2×, and 5× Trp indicates 0.1, 0.2, and 0.5 g trypsin/g chlorophyll, respectively, equivalent to 12.5, 25, and 62.5 ng trypsin/μL), or in the absence of trypsin (−), for 30 min on ice in the dark. Precursor (p), intermediate (i), and mature (m) forms of the various proteins, and a putative trypsin degradation of product of the psTic22 (*), are indicated. Bands corresponding to the imported proteins, in trypsin-treated and control samples, were quantified using ImageJ version 1.32j (http://rsb.info.nih.gov/ij/), and the data were used to produce the graph shown. Plotted values are means (±SD) derived from two independent experiments. The atToc75-III data were derived from the i and m forms, and the psTic22 data were derived from the p and m forms; it was previously demonstrated that cleavage of the psTic22 transit peptide is not essential for import of psTic22 to the intermembrane space (Kouranov et al., 1999). In each case, both bands exhibited essentially the same degree of trypsin resistance. C and D, Control trypsin treatment experiments. C, The in vitro translated atToc75-IV precursor (atToc75-IV IVT) utilized in A and B was treated in the absence (−) and presence (+) of 1× trypsin (Trp), as described in B, in the absence of isolated chloroplasts. D, Import of atToc75-IV was conducted as described in A and B. Samples were then treated with 2× trypsin, as described in B, in the presence or absence of the 1% (v/v) Triton X-100 (TX100), as indicated.

experiment, postimport treatment of the chloroplasts with thermolysin (a protease that cannot penetrate the outer envelope membrane; Cline et al., 1984) was conducted. While atToc75-IV was not proteolytically processed during the import assay (Fig. 4A, compare lanes 1 and 2), the imported protein was nevertheless substantially resistant to thermolysin treatment at the end of the assay (Fig. 4A, compare lanes 5 and 6), suggesting that the protein had become integrated into the chloroplast envelope.

To verify that imported atToc75-IV was associated with the envelope membranes, we conducted fractionation experiments on Arabidopsis chloroplasts containing imported atToc75-IV protein (Fig. 4B). We expected to find that atToc75-IV is envelope-associated for two reasons. First, to date, only two proteins have been shown to gain access to the chloroplast interior without the aid of a transit peptide (Miras et al., 2002; Nada and Soll, 2004). Second, two proteins that are very closely related to atToc75-IV (atToc75-III and psToc75) have been shown to localize to the envelope (Schnell et al., 1994; Tranel et al., 1995; Tranel and Keegstra, 1996; Hiltbrunner et al., 2001b). Like endogenous atToc75-III and atTic110, imported atToc75-IV was detected in the whole chloroplast, total membrane, and envelope membrane preparations, but was completely absent from the thylakoid membrane preparation (Fig. 4B). By contrast, the control protein, light harvesting chlorophyll *a*/*b*-binding protein (LHCP), was detected in the whole chloroplast, total membrane, and thylakoid membrane preparations, but not in the envelope membrane preparation. We therefore conclude that imported atToc75-IV is an integral envelope membrane protein.

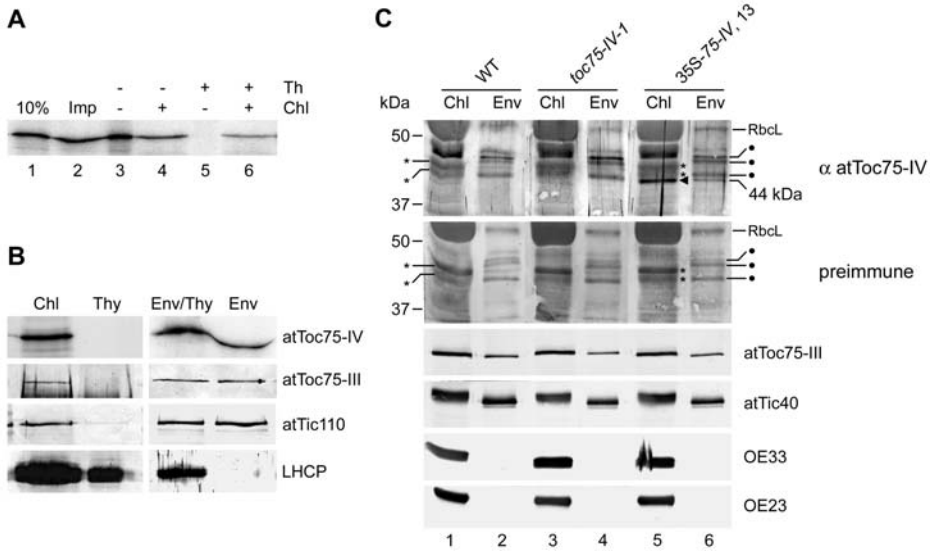

Figure 4. Import and localization studies on atToc75-IV using Arabidopsis chloroplasts. A, Postimport thermolysin treatment experiment. In vitro translated, ^{35}S-labeled atToc75-IV was imported into chloroplasts isolated from 10-d-old Arabidopsis plants grown in vitro. The first lane contains translation product equivalent to 10% of the amount added to each import reaction, and the second lane contains a standard, 20-min import reaction (Imp). The remaining four lanes show an associated thermolysin treatment experiment: ^{35}S-labeled atToc75-IV was incubated for 20 min under import conditions in the presence/absence of chloroplasts (Chl), as indicated, before further treatment in the presence/absence of 60 μg/mL thermolysin (Th), as indicated. In each case, chloroplasts were recovered and the constituent proteins were resolved by SDS-PAGE and visualized by fluorography. B, Suborganellar fractionation experiment. In vitro translated,^{35}S-labeled atToc75-IV protein was imported into isolated chloroplasts as described in A. 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 (i.e. 10 million chloroplasts). C, Immunoblot experiment. Chloroplasts were isolated from 20-d-old wild-type, *toc75-IV-1*, and transgenic 35S-*atTOC75-IV*, line 13 (35S-*75-IV*, 13) plants and then used to prepare envelope membranes as in B. Whole chloroplast (Chl) and envelope membrane (Env) samples (equivalent to 4 μg chlorophyll for the atToc75-IV blots and 2 μg chlorophyll for the control blots) were then analyzed by immunoblotting. Identical blots were probed with antisera derived from two independent rabbits inoculated with an atToc75-IV-specific peptide, and the corresponding preimmune sera (the different rabbits gave similar results and so data are shown for one rabbit only). The antisera specifically detected a protein of 44 kD (the predicted size of atToc75-IV and the apparent molecular mass of in vitro translated atToc75-IV) in Chl and Env samples of the 35S-*atTOC75-IV* line, as indicated by the arrowhead and label at right. Also indicated are bands corresponding to Rubisco large subunit (RbcL), two prominent nonspecific proteins in the Chl samples (indicated with asterisks), and three prominent nonspecific bands in the Env samples (indicated with bullet points); all of these nonspecific bands were also detected by the preimmune sera. Positions of molecular mass standards (kD) are shown at left. Antibodies that recognize atToc75-III, atTic40, and two thylakoid proteins, the 33-kD and 23-kD subunits of the oxygen evolving complex (OE33 and OE23, respectively), were used as controls.

To confirm that atToc75-IV is also localized to chloroplasts in vivo, we probed whole chloroplast protein samples with antibodies raised against an atToc75-IV-specific peptide. Two independent antisera recognized several different protein bands in wild-type and *toc75-IV* mutant chloroplasts, but no differences were observed between the genotypes in either case (Fig. 4C, lanes 1 and 3; data not shown). These data, together with the various lines of evidence indicating low *atTOC75-IV* transcript abundance in wild type (see earlier), suggested that atToc75-IV is a low abundance protein. To facilitate its detection, we also analyzed chloroplast samples from two transgenic lines overexpressing the *atTOC75-IV* cDNA (Fig. 2B). Chloroplasts from both overexpressor lines contained a protein of approximately 44 kD that was specifically detected by both atToc75-IV antisera, but not by the corresponding preimmune sera (Fig. 4C, lane 5; data not shown). The apparent molecular mass of the detected protein (44 kD) matched exactly the apparent molecular mass of the atToc75-IV in vitro translation product (data not shown) and the predicted molecular mass of the atToc75-IV protein (Fig. 1, A and B). The antibodies also detected an approximately 44-kD protein in envelope membranes isolated from the overexpressor line chloroplasts (Fig. 4C, lane 6; data not shown), which implies that the endogenous protein is localized in similar fashion to in vitro translated atToc75-IV (Fig. 4B). These data therefore support our conclusions drawn from the in vitro import experiments, i.e. that atToc75-IV is a chloroplast envelope membrane protein and that it is targeted without the aid of a cleavable transit peptide.

Envelope Insertion of atToc75-IV Does Not Require ATP at High Concentrations

Most outer envelope membrane proteins are synthesized at their mature size without a cleavable transit peptide (Keegstra and Cline, 1999; Schleiff and Klösgen, 2001). Unlike stromal proteins, these outer membrane proteins do not seem to require high ATP concentrations or thermolysin-sensitive envelope components for targeting to the chloroplast, indicating that they do not utilize the complete Toc/Tic complex for import (Keegstra and Cline, 1999; Schleiff and Klösgen, 2001; Tu et al., 2004). In an attempt to shed light on the atToc75-IV targeting mechanism, we conducted import assays under ATP-limiting conditions (Fig. 5). In vitro translated, ATP-depleted atToc75-IV and Rubisco small subunit precursor (preSSU) were each incubated with ATP-depleted chloroplasts under import conditions (Olsen et al., 1989; Aronsson and Jarvis, 2002). The control protein, preSSU, was not imported and processed by the chloroplasts unless 5 mM ATP was included in the import buffer, confirming that ATP had been successfully depleted from both the chloroplasts and the in vitro translated precursors. By contrast, atToc75-IV localized to the membrane fraction whether or not ATP was included in the import buffer (Fig. 5), demonstrating that the membrane insertion of atToc75-IV does not require ATP at the high concentrations necessary for translocation into the stroma. While these data demonstrate that the targeting of atToc75-IV differs significantly from that of preSSU, they do not reveal a difference between the atToc75-IV and atToc75-III/psToc75 targeting mechanisms, since psToc75 targeting was also shown to proceed at low ATP concentrations (approximately 50 μM; Tranel et al., 1995).

Identification of atToc75-III and atToc75-IV Knockout Mutants

To investigate the roles of the atToc75-III and atToc75-IV proteins in vivo, we identified T-DNA or transposon knockout lines for each gene. Two independent alleles of each knockout mutation were identified (Fig. 1A). In each case, the independent alleles behaved identically, indicating that the mutant phenotypes observed were due to the T-DNA or transposon insertions, rather than secondary, unlinked

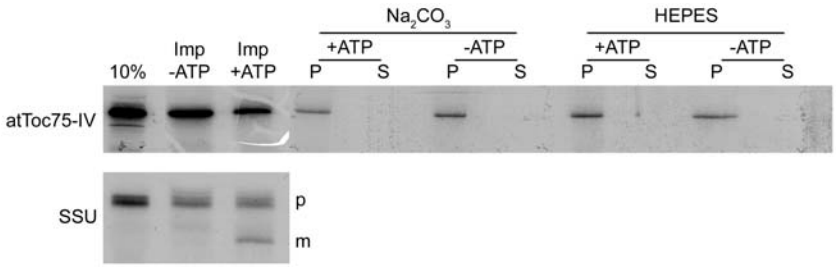

Figure 5. Energetics of atToc75-IV membrane insertion. Chloroplasts isolated from 10-d-old Arabidopsis plants grown in vitro were incubated in the dark in the presence of 6 μM nigericin for 10 min to deplete endogenous ATP. In vitro translated, ^{35}S-labeled atToc75-IV and preSSU proteins were passed through Sephadex G-25 to remove small molecules and then incubated with the chloroplasts under import conditions for 20 min in the absence (Imp −ATP) or presence (Imp +ATP) of 5 mM ATP. Chloroplasts recovered from separate import assays were treated for 1 h with 0.1 M Na_2CO_3 (high pH) or 10 mM HEPES, 10 mM $MgCl_2$, pH 8.0 (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.

mutations. Unless stated otherwise, the *toc75-III-1* and *toc75-IV-1* alleles were used to generate the data presented below.

All four knockout lines contained single-site insertions, based on the segregation of the insertion-associated marker genes (Table I). Families derived from heterozygous *toc75-IV* plants segregated three resistant plants for every one sensitive plant when plated on selective medium (Table I), indicating standard Mendelian inheritance. By contrast, families derived from heterozygous *toc75-III-1* plants contained only two resistant plants for every one sensitive plant, and families derived from heterozygous *toc75-III-2* plants segregated approximately two heterozygous plants for every one wild-type plant when tested by PCR (Table I). Furthermore, all *toc75-III* mutant plants tested by PCR (60 plants for *toc75-III-1* and 35 plants for *toc75-III-2*) were found to be hemizygous for the corresponding T-DNA insertion. These data indicate that homozygous atToc75-III knockout mutations, unlike homozygous atToc75-IV mutations, are lethal during an early stage of development.

None of the single mutant seedlings (*toc75-III* heterozygotes and *toc75-IV* homozygotes) displayed obvious visible or chlorophyll accumulation phenotypes (Fig. 6). The absence of a phenotype in *toc75-III* heterozygotes indicates that the knockout mutation is recessive and that one *atTOC75-III* allele is sufficient to support normal growth under the conditions tested. The absence of a phenotype in *toc75-IV* homozygotes may reflect a degree of redundancy between atToc75-IV and other proteins, such as atToc75-III, or the nonessential nature of atToc75-IV under the conditions tested. In an attempt to reveal a phenotype for atToc75-IV, we generated double mutants by crossing *toc75-IV-1* with *toc75-III-1* and the atToc33 knockout mutant, *ppi1*, which displays protein import defects and a pronounced yellow-green phenotype (Jarvis et al., 1998; Kubis et al., 2003). Double mutants that were heterozygous for *toc75-III* and homozygous for *toc75-IV* (genotype, +/*toc75-III*; *toc75-IV*/*toc75-IV*) did not display any obvious visible phenotypes (Fig. 6). Furthermore, an extensive analysis of siblings among the progeny of plants that were homozygous for *toc75-IV* and heterozygous for *toc75-III* revealed no significant effect of a single *toc75-III* allele in the *toc75-IV* homozygous background (Fig. 6B). These data suggest that, if there is functional redundancy between the two proteins, a single *atTOC75-III* allele is sufficiently potent to compensate for the complete loss of atToc75-IV protein under normal circumstances. Similarly, even in the compromised, import-defective *ppi1* background, the *toc75-IV* mutation did not give rise to any obvious visible phenotypes (Fig. 6).

Like the *toc75-IV* mutants, the previously identified atToc34 knockout mutants, *ppi3-1* and *ppi3-2*, exhibit no obvious visible phenotypes in aerial tissues (Constan et al., 2004). However, the *ppi3* mutants were found to display growth defects in the roots, a result that was interpreted to indicate an important role for atToc34 in the import of nonphotosynthetic proteins. To assess the possibility that the atToc75-IV protein plays a similar role in nonphotosynthetic protein import, we conducted root growth measurements much like those described previously (Constan et al., 2004). However, repeated measurements of plants grown under long-day conditions for 10 d (in total, approximately 100 plants of each genotype were measured) revealed no root length differences between the wild type and *toc75-IV-1* or *toc75-IV-2* (data not shown).

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 (Table I), we determined the stage of development at which the growth of *toc75-III* homozygotes terminates. In siliques of heterozygous *toc75-III* plants, aborted seeds were present at a frequency of approximately 25% (Table II; Fig. 7, K and L). This indicates that homozygous *toc75-III* mutations are embryo lethal, which is consistent with the notion that atToc75-III is the main Arabidopsis ortholog of psToc75 and with the fact that Toc75 plays a key, channel-forming role in the protein import mechanism.

To determine the exact stage at which development of homozygous *toc75-III* embryos terminates, we used Nomarski optics to observe developing seeds in heterozygous *toc75-III* siliques of different ages (Table III). Approximately 75% of embryos were observed to develop normally (Fig. 7, A–E). However, approximately 25% of the embryos in heterozygous *toc75-III* siliques did not develop beyond the two-cell stage (Fig. 7, F–I); i.e. when the normal embryos were at late

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

Mutant	Antibiotic	Antibiotic Resistant (R)	Antibiotic Sensitive (S)	R:S Ratio	χ^2-Value[b]	*P*-Value[b]
toc75-III-1	Hygromycin	171	91	1.9	0.231	0.631
toc75-III-2	Kanamycin[a]	35	21	1.7	0.438	0.508
toc75-IV-1	Phosphinothricin	83	26	3.2	0.076	0.782
toc75-IV-2	Phosphinothricin	41	14	2.9	0.006	0.938

[a]The *toc75-III-2* T-DNA insertion was followed by PCR since the associated kanamycin resistance marker was silenced. [b]Goodness-of-fit of the observed ratios to 2:1 (*toc75-III*) and 3:1 (*toc75-IV*) was assessed by χ^2 analysis. *P*-values are the probabilities that the observed ratios differ from 2:1 or 3:1 due to random chance only.

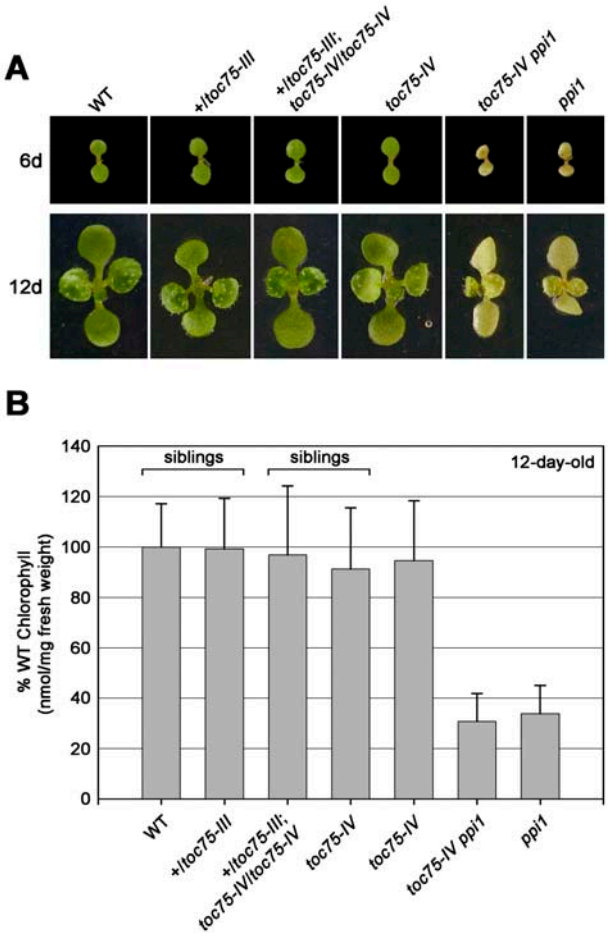

Figure 6. Phenotypic analysis of the *toc75-III* and *toc75-IV* mutants. A, Plants grown under long-day conditions in vitro for 6- or 12-d post-germination are shown. The presented genotypes are, from left to right: wild type (WT), *toc75-III* heterozygote (+/*toc75-III*), *toc75-III toc75-IV* double mutant (+/*toc75-III*; *toc75-IV*/*toc75-IV*), *toc75-IV* homozygote, *toc75-IV ppi1* double homozygote, and *ppi1* homozygote. B, The visible phenotypes of 12-d-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 to 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 to 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: wild type (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.

globular stage (Fig. 7D), approximately 25% of the developing embryos had just two nuclei in the single apical embryo proper cell (Fig. 7I). When the normal embryos reached the heart stage (Fig. 7E), approximately 25% of the seed aborted (Fig. 7J). Additionally, the suspensor of homozygous *toc75-III* seed did not develop normally and was arrested after a single transverse division of the basal cell after the first zygotic division. Moreover, the free nuclear endosperm showed limited divisions and failed to cellularize, leading to nuclear fragmentation (Fig. 7I). Thus, we conclude that embryo development in *toc75-III* homozygotes arrests at the two-cell stage, and that this defect, together with endosperm failure, leads to seed abortion.

Etioplasts in the *toc75-IV* Mutants Have Structural Abnormalities

Since homozygous *toc75-IV* seedlings were found to have no obvious visible phenotypes (Fig. 6), we sought to identify a mutant phenotype by conducting more detailed studies. First, we compared the ability of isolated wild-type and *toc75-IV* mutant chloroplasts to import two different precursor proteins (preSSU and the 33-kD oxygen evolving complex precursor, preOE33), using published procedures (Aronsson and Jarvis, 2002; Kubis et al., 2003). However, no significant differences were observed (data not shown). Next, we used electron microscopy to analyze the ultrastructure of chloroplasts in the primary leaves of 10-d-old, light-grown plants. Electron micrographs of approximately 20 different plastids each from wild type and *toc75-IV-1* were studied and quantified with respect to cross-sectional area and overall shape of the plastid, the number of starch grains, plastoglobuli and granal stacks per plastid, and the number of thylakoid membranes per granum. These data were analyzed statistically, but no significant differences between the wild-type and mutant chloroplasts could be detected (data not shown).

The expression data presented earlier (Fig. 2 and related discussion) suggest that *atTOC75-IV* may play a relatively more important role in etioplasts. We therefore analyzed the ultrastructure of etioplasts from 3-d-old, dark-grown wild-type and *toc75-IV* seedlings. While the prolamellar bodies and prothylakoids of *toc75-IV* etioplasts looked normal, we observed that some etioplasts contained enclosed bodies of cytosol, each apparently surrounded by normal envelope membrane (Fig. 8). We examined a relatively large number of etioplasts from wild-type, *toc75-IV-1*, and *toc75-IV-2* seedlings, scored the number of etioplasts with cytosolic inclusions in each case, and then statistically analyzed the data (Table IV). Although approximately 20% of the wild-type etioplasts contained cytosolic inclusions, both *toc75-IV* alleles

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

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

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%)

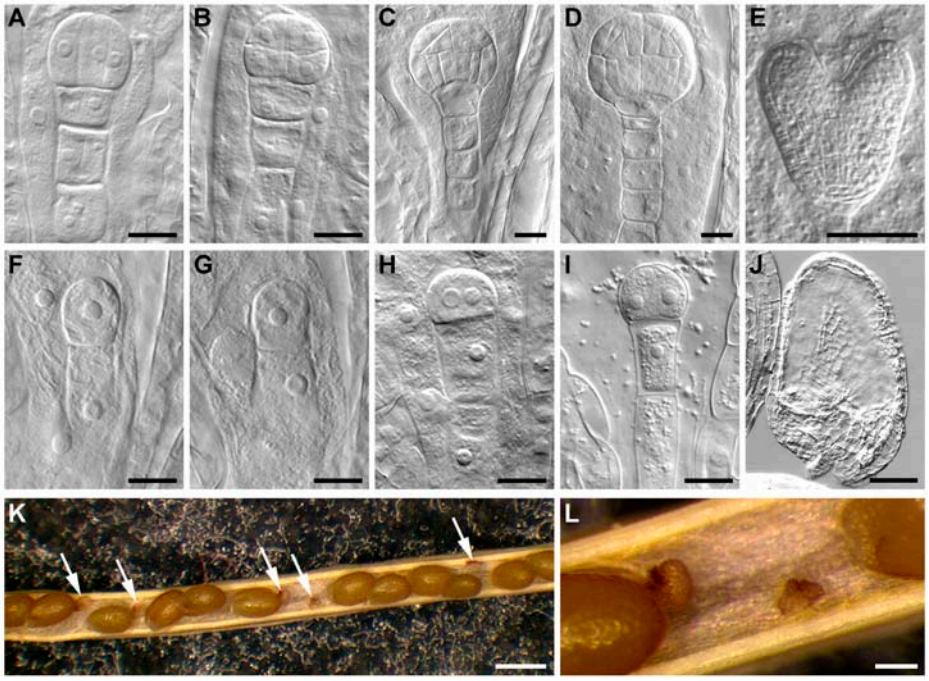

Figure 7. Embryo lethality of the *toc75-III* mutation. A to J, Morphology of normal (A–E) and mutant (F–J) embryos from the siliques of *toc75-III* heterozygous plants. The normal developmental stage in each corresponding pair of images is as follows: A and F, four-cell embryo stage; B and G, eight-cell embryo stage; C and H, 16-cell embryo stage; D and 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 = 10 μm (A–D and F–I) and 50 μm (E and J). K, The appearance of part of a typical silique from a mature *toc75-III* heterozygous plant. Arrows indicate aborted seeds. Bar = 0.5 mm. L, Higher magnification image of the two central aborted seeds shown in K. Bar = 0.1 mm.

contained significantly more cytosolic inclusions than wild-type etioplasts (Table IV). The data were analyzed using the chi-squared test, and, because there are only two categories of data (one degree of freedom), we applied the Yates' correction, which makes the analysis more stringent. The null hypothesis that there was no significant difference between the wild-type and mutant data sets tested was rejected (critical χ^2-value = 3.84, *P*-value = 0.05), and so we conclude that mutant *toc75-IV* etioplasts contain significantly more cytosolic inclusions than wild-type etioplasts.

To investigate further the role of atToc75-IV, we transformed wild-type plants with an *atTOC75-IV* overexpression construct using the floral dip method (Clough and Bent, 1998). Twelve independent transformants were identified and shown to overexpress the *atTOC75-IV* gene by RT-PCR (data not shown). When grown in the light, these plants did not display any obvious phenotypes (data not shown). To assess the effect of *atTOC75-IV* overexpression on the cytosolic inclusions observed in the *toc75-IV* mutants (Table IV; Fig. 8), we examined the etioplasts of dark-grown seedlings in two of these transformed lines by electron microscopy. These two lines were estimated to display at least 30-fold increased levels of *atTOC75-IV* mRNA expression, relative to wild type, by RT-PCR (Fig. 2B). Remarkably, etioplasts from seedlings overexpressing *atTOC75-IV* contained significantly fewer cytosolic inclusions than wild-type etioplasts (Table IV). To assess whether or not the transgenic data deviated to a significant degree from the corresponding wild-type data, we again employed the chi-squared test with the Yates' correction. This analysis confirmed that the etioplasts of both transgenic lines contained significantly fewer cytosolic inclusions than wild-type etioplasts. We also found that there was no significant difference between the two wild-type samples (Table IV; χ^2-value = 1.68, *P*-value = 0.19),

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

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

Silique	One-Cell	Two-Cell[a]	Aborted Seed[b]	Four-Cell	Eight-Cell	16-Cell	32-Cell	Globular	Heart	Total
1	28									28
2	12	9		13						34
3	7	4		10	4					25
4	7	3		14	10					34
5	3	6			9	9				27
6	2	5	2		2	9	11	5		36
7		5	3		1	9	16	2		36
8		4	3			6	12	14	1	40

[a]Two-cell stage also includes embryos with two free nuclei in the apical cell. [b]Aborted seed were shriveled with no visible embryo or endosperm development (see Fig. 7).

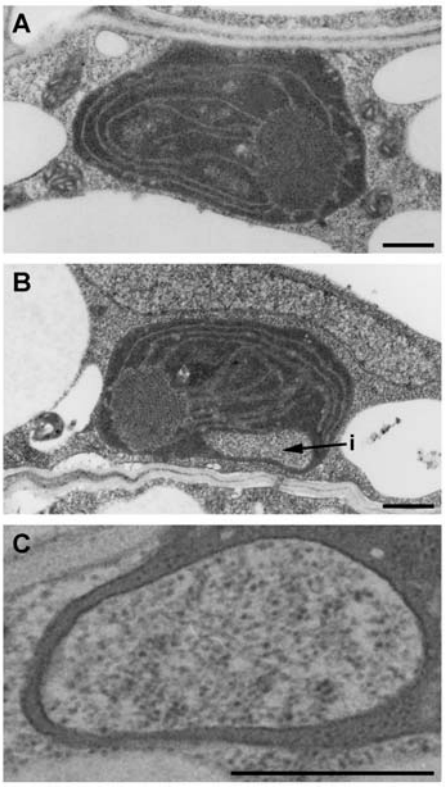

Figure 8. Ultrastructure of etioplasts in the *toc75-IV* mutants. Etioplasts from wild-type (A) and *toc75-IV-2* mutant (B) plants grown in vitro in the dark for 3 d; the *toc75-IV* plastid contains a large inclusion (i) of cytosolic material. In C, an enlargement of a typical cytosolic inclusion within a *toc75-IV-1* mutant etioplast is shown. Bars = 0.5 μm.

which were grown on separate occasions for comparison with the *toc75-IV* mutants and *atTOC75-IV* overexpression lines. Thus, these data strongly suggest that atToc75-IV has a role in etioplasts.

The *toc75-IV* Mutants Exhibit Inefficient Deetiolation

Expression (Fig. 2) and electron microscopy (Table IV; Fig. 8) data pointed toward a role for atToc75-IV in the etioplasts of dark-grown plants. To corroborate these findings, we carefully examined etiolated *toc75-IV* mutant plants for evidence of growth and developmental defects. First of all, we simply measured hypocotyl growth in dark-grown plants. However, repeated measurements of plants grown in the dark for either 5, 6, or 9 d (in total, approximately 800 plants of each genotype were measured) revealed no hypocotyl length differences between the wild type, *toc75-IV-1*, and *toc75-IV-2* (data not shown).

Since the most important function of etioplasts is the initiation of chloroplast development, following exposure to light, we next examined the ability of the *toc75-IV* mutants to undergo deetiolation. Wild-type and mutant plants were grown in the dark for 6 d and then transferred to continuous white light for a further period of 2 d. At the end of the 2-d light period, the plants were scored for visible evidence of deetiolation. Plants with green, expanded cotyledons were scored as having undergone deetiolation, whereas those with yellow, unexpanded cotyledons were scored as having failed to undergo deetiolation (Fig. 9A). This experiment was repeated 11 times, using multiple fresh batches of seed for each genotype (all of which were of the same age and produced by plants grown under identical conditions), including the same seed batches that were used to make the root and hypocotyl length measurements described earlier. Although the experiment was subject to substantial variability, presumably due to the stochastic nature of the deetiolation response and its acute sensitivity to environmental factors, the results revealed a significant difference between both *toc75-IV* mutants and the wild type (Fig. 9, B and C). In several individual experiments, a striking difference was observed between the mutants and wild type (e.g. Fig. 9B). In other experiments, this difference was less pronounced. Overall, however, the combined data of all 11 experiments (in total, >2,500 plants of each genotype were scored) revealed an approximately 50% reduction in deetiolation efficiency in the mutants by comparison with wild type (Fig. 9C). The combined *toc75-IV-1* and *toc75-IV-2* data sets were each statistically significantly different from the wild-type data set, as revealed by a Student's *t* test (*P*-values = 0.000016 and 0.0056, respectively); by contrast, the same test indicated that the two mutant data sets were not significantly different from each other (*P*-value = 0.073). These results demonstrate a physiological role for the atToc75-IV protein, most likely in etioplast biogenesis and/or the etioplast-chloroplast transition. They are also supportive of the ultrastructural data presented in Table IV and Figure 8, as well as of the data indicating that atToc75-IV is a plastidic protein shown in Figures 3 and 4.

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

Genotype	Total No. Etioplasts	No. With Inclusions	Frequency	χ^2-Value[a]	*P*-Value[a]
			%		
Wild type	54	12	22.2		
toc75-IV-1	62	24	38.7	8.77	0.003
toc75-IV-2	61	24	39.3	9.27	0.002
Wild type	100	15	15.0		
35S-atTOC75-IV, line 10	131	4	3.1	13.74	0.0002
35S-atTOC75-IV, line 13	123	8	6.5	6.31	0.01

[a]The data were analyzed using a chi-squared test with the Yates' correction for continuity. We set our critical χ^2-value at 3.84, *P*-value = 0.05, and in every case the mutant/transgenic data were found to differ significantly from the wild-type data. A *P*-value of 0.01 indicates that there is 1% probability that the observed differences between the mutant/transgenic and wild-type data occurred by chance alone.

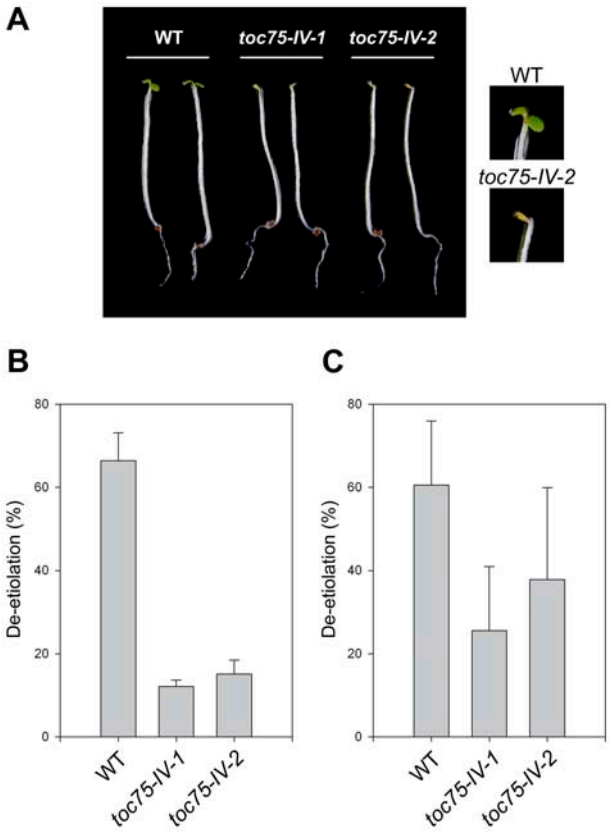

Figure 9. Deetiolation in the *toc75-IV* mutants. Wild-type and mutant Arabidopsis plants were grown in vitro in the dark, from germination, for a period of 6 d and then transferred to continuous white light of standard intensity for a further period of 2 d. At the end of the 2-d light period, plants were scored for evidence of deetiolation. This experiment was repeated 11 times. A, Typical plants from an individual experiment that revealed a striking difference in deetiolation efficiency between the *toc75-IV* mutants and wild type (data from this experiment are shown in B). The insets at right show the cotyledons of representative plants from the main image, at 2-fold higher magnification, and serve to illustrate the presence/absence of a deetiolation response. B, An individual experiment that revealed a striking difference in deetiolation efficiency between the *toc75-IV* mutants and wild type (representative plants from this experiment are shown in A). The data were derived from four separate petri plates per genotype, each one containing approximately 80 plants. Deetiolation frequencies were calculated on a per plate basis and then used to produce the mean values shown (±SD; $n = 4$). C, The combined data of all 11 experiments, each of which comprised three or four separate petri plates per genotype. Deetiolation frequencies were calculated on a per experiment basis and then used to produce the mean values shown (±SD; $n = 11$).

DISCUSSION

The Toc75 protein forms the main protein import channel through the plastid outer envelope membrane (Schnell et al., 1994; Tranel et al., 1995; Hinnah et al., 2002). It works in concert with receptor proteins and is capable of recognizing, as well as translocating, chloroplast precursor proteins (Hinnah et al., 2002; Becker et al., 2004). Three Toc75-related sequences were previously identified in the Arabidopsis genome (Jackson-Constan and Keegstra, 2001). In this study, we used a range of techniques to characterize this small Arabidopsis gene family. Our data indicate that only two of the three family members, *atTOC75-III* and *atTOC75-IV*, are expressed genes. The third gene, *atTOC75-I*, has been inactivated due to a transposon insertion and the accumulation of numerous other mutations (Fig. 1, A and B). While *atTOC75-III* encodes a protein very similar to psToc75, *atTOC75-IV* encodes an amino-terminally truncated derivative (Fig. 1, A and B).

It is interesting to note that Schleiff et al. (2003b) observed a minimal number of smaller Toc particles during their studies of the pea Toc complex, since it is possible that these particles incorporate an atToc75-IV-like protein. However, our phylogenetic analysis revealed that the Arabidopsis Toc75 gene family most likely formed subsequent to the speciation of Arabidopsis and pea (Fig. 1C). Thus, pea and other species do not necessarily contain genes related to *atTOC75-IV* or *atTOC75-I*. The Arabidopsis genome has undergone extensive duplication and reshuffling, and it has therefore been suggested that Arabidopsis is a degenerate tetraploid (Blanc et al., 2000). While it was not possible to map the three Arabidopsis Toc75-related sequences to the major duplications reported by Blanc et al. (2000), it is nevertheless possible that these or similar processes gave rise to the Arabidopsis Toc75 gene family. Interestingly, *atTOC75-I* appears to have lost a single intron subsequent to its divergence from the other genes. The precise loss of one or more introns has previously been observed in plants and other organisms (Häger et al., 1996; Drouin and Moniz de Sa, 1997). One possible mechanism involves the RT of processed cellular mRNA to produce a cDNA copy of the gene in question, or of a closely related gene, which can then partially replace the genomic copy through homologous recombination or gene conversion (Derr et al., 1991; Drouin and Moniz de Sa, 1997).

Based on the level and pattern of *atTOC75-III* expression, and on the high sequence identity shared between atToc75-III and psToc75, we conclude that atToc75-III is the main Arabidopsis ortholog of psToc75. The unique importance of atToc75-III was also demonstrated by the embryo lethality of the homozygous *toc75-III* mutations (Tables I–III; Fig. 7). As has been reported previously, plastids are essential for embryo development, presumably because they synthesize many important products that are utilized by the rest of the cell, including carbohydrates, fatty acids, amino acids, and terpenoids (Uwer et al., 1998; Apuya et al., 2001). The central importance of plastids, and of the proper targeting of plastid proteins, was recently demonstrated by studies on Arabidopsis mutants lacking other components of the protein import apparatus (Hörmann et al., 2004; Kovacheva et al., 2005). Homozygous knockout mutants for the well-characterized Tic complex component, Tic110, resulted in embryo arrest at the globular stage (i.e. at a much later stage than *toc75-III*; Kovacheva et al., 2005),

and similar mutations affecting a new, putative Tic complex component seemed to arrest development even later at the heart stage (Hörmann et al., 2004). It is interesting that, in a large-scale screen for T-DNA-induced mutations affecting seed development, >25% of the reliably identified loci were predicted to encode plastid-localized proteins (McElver et al., 2001), whereas, by contrast, only approximately 12% of all Arabidopsis nuclear genes are predicted to encode proteins of plastid-related function (Leister, 2003).

Various lines of evidence indicate that the *atTOC75-III* gene is expressed at much higher levels than *atTOC75-IV* (see "Results"). In Arabidopsis, the various homologous genes that encode other components of the Toc complex also display differing levels of expression. For example, *atTOC159* is expressed at approximately 5- to 10-fold higher levels than its homologs, *atTOC132* and *atTOC120* (Bauer et al., 2000; Ivanova et al., 2004; Kubis et al., 2004), and maximal expression of *atTOC33* is approximately 6-fold higher than that of *atTOC34* (Jarvis et al., 1998; Kubis et al., 2003). In each case, the more highly expressed homolog (*atTOC159* or *atTOC33*) has been implicated in the preferential import of photosynthetic precursor proteins (Bauer et al., 2000; Kubis et al., 2003, 2004; Ivanova et al., 2004). However, our results suggest that this type of functional specialization does not exist within the Toc75 family. Undifferentiated proplastids are inherited maternally by the plant zygote and do not develop into chloroplasts until the heart stage when the embryo begins to turn green (Apuya et al., 2001). Thus, the fact that homozygous *toc75-III* embryos do not develop beyond the two-cell stage (Fig. 7) strongly suggests that atToc75-III is essential for the import of nonphotosynthetic proteins in developing proplastids. The fact that the mutant embryos abort so early also implies that atToc75-III and atToc75-IV share only limited functional redundancy or that the *atTOC75-IV* gene is not expressed during early embryogenesis.

It is interesting that the *toc75-III* mutation causes embryo lethality, especially since other individual components of the Toc complex (Toc34- and Toc159-related proteins, which are involved in preprotein recognition) are not essential for embryo development (Jarvis et al., 1998; Bauer et al., 2000; Constan et al., 2004; Ivanova et al., 2004; Kubis et al., 2004). In yeast, while mutants lacking the mitochondrial protein import receptors are able to grow on nonfermentable carbon sources, the Tom40 preprotein translocation channel is essential for viability (Pfanner et al., 1997). Since Toc75 is able to recognize transit peptides without the aid of other Toc components (Hinnah et al., 2002), and since free Toc75 exists in the outer envelope membrane (Schleiff et al., 2003b), one might conclude that Toc75 itself constitutes a minimal, functional translocation unit. However, the available data now indicate that the other Toc knockout mutants are able to survive due to functional redundancy between related components. For example, while the atToc33 and atToc34 single knockout mutants, *ppi1* and *ppi3*, are quite healthy and display only mild visible phenotypes, the *ppi1 ppi3* double homozygous genotype is embryo lethal (Constan et al., 2004). Similarly, Arabidopsis double mutants lacking both of the major Toc159 isoforms, atToc159 and atToc132, arrest and die during embryogenesis (Kubis et al., 2004). Thus, the difference in phenotype severity between *toc75-III* and the other Toc knockout mutants most likely reflects the central and uniquely important nature of the role played by atToc75-III in preprotein conductance and the inability of atToc75-IV to substitute for atToc75-III during embryogenesis.

Expression data suggested that the atToc75-IV protein might play a particularly important role during dark growth, or that a phenotype associated with the *toc75-IV* mutations might become more apparent in dark-grown seedlings (Fig. 2). Indeed, when the etioplasts of dark-grown plants were examined by electron microscopy, certain structural defects (an increased frequency of cytosolic inclusions) were observed in the mutants (Table IV; Fig. 8); as one might have expected, a diametrically opposed phenotype was observed in *atTOC75-IV* overexpressing plants (Table IV). Interestingly, elevated frequencies of very similar cytosolic inclusions were observed in the root plastids of atToc132/atToc120 double knockout mutants (Kubis et al., 2004). It is not clear why these inclusions develop or why their frequency is influenced by the expression level of atToc75-IV. One possibility is that they are caused by perturbations in the plastids' ability to import structural proteins required to maintain normal organellar integrity. On the other hand, they may be an adaptive response designed to increase the surface-area-to-volume ratio of the plastids in order to facilitate the uptake of substances, either proteins or small molecules, from the cytosol. Regardless of the reason for their existence, it seems clear that they are associated with defects in plastid functions, since etiolated *toc75-IV* plants were significantly less efficient at deetiolation than the wild type (Fig. 9), a transition during that the differentiation of chloroplasts from etioplasts is of central and integral importance.

Since the envelope insertion of atToc75-IV is apparently not accompanied by proteolytic cleavage (Figs. 3 and 4), it is tempting to speculate that atToc75-IV gains access to the chloroplast using a different mechanism from its larger homolog, atToc75-III. This possibility is particularly interesting if one also assumes that atToc75-IV is a protein import channel with at least some functional similarity to atToc75-III. It may be significant that the targeting of atToc75-III is thought to be dependent upon a preexisting pool of correctly assembled atToc75-III (Tranel et al., 1995; Tranel and Keegstra, 1996; Inoue and Keegstra, 2003); perhaps the role of atToc75-IV is to partially alleviate this "chicken-and-egg" dependency during the early stages of biogenesis. However, the relationship between the atToc75-III and atToc75-IV targeting mechanisms and

the ability of atToc75-IV to act as an import channel are both issues that have not yet been addressed directly, and so it is clear that further work is required before a role can be assigned to the atToc75-IV protein. In this regard, the *toc75-IV* mutants and the *atTOC75-IV* overexpression lines described in this report should prove to be invaluable resources. The *toc75-III* mutants will also be useful tools as we seek to identify functionally important domains of the atToc75-III protein in transgenic complementation studies.

MATERIALS AND METHODS

Sequence Alignment and Phylogenetics

Amino acid sequences were aligned using ClustalW within the BioEdit program (Hall, 1999). Phylogenetic trees were calculated using PAUP* 4.0 b10 (Swofford, 2003). All analyses were performed using branch and bound searches, with the "collapse" option and "furthest" addition sequence selected. No weighting or ordering was imposed on the characters, and any gaps were treated as missing data. Indels were coded separately and appended to the sequence data matrix. Coding of indels was usually binary (deletions 0, insertions 1), but in places where more than one size occurred, the alternatives were coded as 2, 3, etc. Support for clades was estimated by means of non-parametric bootstrap analyses, as implemented in PAUP* 4.0, using 1,000 replicates.

Plant Material and Growth Conditions

All Arabidopsis (*Arabidopsis thaliana*) plants, both wild type and mutant, used in this study were of the Columbia-0 ecotype. For in vitro growth on Murashige and Skoog media, Arabidopsis seeds were surface-sterilized, and plants were grown at 20°C under a long-day cycle (16-h light/8-h dark) of 120 μmol photons m^{-2} s^{-1} white light, as described previously (Aronsson and Jarvis, 2002). When necessary, the following antibiotics were included in the medium at the indicated concentrations: 15 μg/mL hygromycin B (Duchefa, Haarlem, The Netherlands) to select for *toc75-III-1*; 50 μg/mL kanamycin monosulfate (Melford Laboratories, Ipswich, UK) to select for *toc75-III-2*; 10 μg/mL DL-phosphinothricin (Duchefa) to select for *toc75-IV-1* and *toc75-IV-2*; and 110 μg/mL gentamicin sulfate (Duchefa) to select for pCHF2 transformants. For growth on soil, Arabidopsis plants were kept under standard greenhouse conditions, with a long-day cycle (16-h light, 8-h dark) and supplemental lighting when necessary.

Pea (*Pisum sativum*) plants were of the Little Marvel variety (SeedWay, Elizabethtown, PA). Pea seeds were soaked overnight with aeration, sown in vermiculite, and then allowed to grow at 20°C under a 12-h-light/12-h-dark cycle of approximately 30 μmol photons m^{-2} s^{-1} white light for 11 to 14 d.

RNA Isolation, RT-PCR, and RACE-PCR

RNA was isolated according to Kubis et al. (2003). RNA samples (20 μg) were treated with DNAse I according to the manufacturer's instructions (DNA-*free*; Ambion, Austin, TX), quantified spectrophotometrically, and then analyzed on 1.3% (w/v) agarose gels containing 1.85% (v/v) formaldehyde and MOPS buffer (20 mM MOPS, 1 mM EDTA, 50 mM NaOAc, pH 7.0) to confirm concentration and quality.

For RT-PCR analysis, first-strand cDNA was synthesized using 5 μg of total RNA, a primer with the sequence 5′-$(T)_{17}$(A/G/C)N-3′, and SuperScript II RNase H− reverse transcriptase (Invitrogen, Carlsbad, CA), according to the manufacturer's instructions. 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, Uppsala), and hybridized with gene-specific probes amplified using the primers listed above, according to the manufacturer's instructions (Amersham Biosciences, Little Chalfont, UK). Bands were visualized using a phosphorimager and quantified using ImageQuant software (Molecular Dynamics, Sunnyvale, CA). *atTOC75-III* and *atTOC75-IV* data were normalized using similar data for the uniformly expressed translation initiation factor gene, *eIF4E1* (gene no. At4g18040; Rodriguez et al., 1998).

The *atTOC75-IV* mRNA was characterized by RACE-PCR using the SMART RACE cDNA amplification kit (CLONTECH, Palo Alto, CA) according to the manufacturer's instructions. The gene-specific 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, Madison, WI), verified by sequencing, and then subcloned as a *Sac*II/*Spe*I fragment into pBlueScript II KS− (Stratagene, La Jolla, CA). 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).

In Vitro Translation of Precursor Proteins

For the pea import experiments, transcription/translation was performed in a coupled system containing rabbit reticulocyte lysate, [^{35}S]Met, and T7 RNA polymerase, according to the manufacturer's instructions (TNT T7 Coupled Reticulocyte Lysate System; Promega). Plasmid DNA was used as template directly. The plasmids used have been described previously (Jackson et al., 1998; Kouranov et al., 1999; Inoue and Keegstra, 2003), are described in this report (*atTOC75-IV*; see above), or were supplied with the translation system (Luc; Promega).

For the Arabidopsis import experiments, templates were amplified from cDNA clones by PCR using M13 primers. The *atTOC75-III* template was amplified from cDNA clone, APZL59c10R (the insert was subcloned into pBlueScript II KS- to enable use of the T7 promoter), and the *atTOC75-IV* template was amplified from the pBlueScript II KS− clone described above; the preSSU cDNA clone was described previously (Aronsson and Jarvis, 2002). Transcription/translation was performed in a coupled system containing rabbit reticulocyte lysate, [^{35}S]Met, and T7 RNA polymerase, according to the manufacturer's instructions (TNT T7 Quick for PCR DNA; Promega).

Chloroplast Isolation, Import, and Fractionation

Pea chloroplast isolations and in vitro import experiments, including high pH extractions and trypsin treatments, were carried out according to Inoue and Keegstra (2003), except that import reactions were conducted for 30 min instead of 20 min.

Arabidopsis chloroplast isolations and in vitro import experiments were carried out essentially as described by Aronsson and Jarvis (2002), with one exception: the chloroplasts used for the experiments shown in Figure 4C were isolated as described by Fitzpatrick and Keegstra (2001). All import reactions were incubated for 20 min at 25°C unless stated otherwise. Thermolysin reactions contained 60 μg/mL thermolysin, 300 μM $CaCl_2$, and were conducted for 5 min on ice. Control thermolysin reactions lacking chloroplasts contained 10% of the amount of translation mixture added to import reactions, since only approximately 10% of added ^{35}S-labeled atToc75-IV associates with chloroplasts in import assays.

For fractionation studies, chloroplasts were burst by incubation in 10 mM HEPES, 10 mM $MgCl_2$, pH 8.0 (bursting buffer) for 10 min on ice. Total membrane samples were collected by centrifugation at 100,000*g* for 1 h. Thylakoid membranes were prepared according to Rawyler et al. (1992).

Envelope membranes were prepared as follows. Burst chloroplasts were centrifuged at 2,600*g* for 8 min. The supernatant was retained, and then centrifuged again in the same way. This procedure was repeated once more, and then the remaining membranes (envelopes only) were collected from the supernatant by centrifugation at 100,000*g* for 1 h.

The atToc75-III protein was detected using an antibody raised against purified psToc75 (Tranel et al., 1995) or using an antibody raised against a synthetic, atToc75-III-specific peptide (MDSDMTDKEKLEYYRC; Eurogentec S.A., Herstal, Belgium); both antibodies recognized recombinant atToc75-III (data not shown). The atToc75-IV protein was detected using an antibody raised against a synthetic, atToc75-IV-specific peptide (ENRRISENGLLLSPDC; Eurogentec S.A.). The Tic110, Tic40, OE33, OE23 and LHCP antibodies have all been described previously and were gifts from Kenneth Keegstra (Michigan State University; Tic110 and Tic40), Steven Theg (University of California, Davis; OE33 and OE23), and Neil Hoffman (National Science Foundation; LHCP).

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 no. 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 approximately 20% of the SIGnAL T-DNA 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 (Fig. 1A). Single-locus insertion lines were identified for each mutant (Table I); for *toc75-IV-1* and *toc75-IV-2*, homozygotes derived from these lines were used for further study.

Chlorophyll Quantification and Electron Microscopy

Chlorophyll determinations and transmission electron microscopy were carried out as described previously (Porra et al., 1989; Constan et al., 2004). Transmission electron microscopy was carried out at the Electron Microscope Laboratory, Faculty of Medicine and Biological Sciences, University of Leicester.

Phenotypic Analysis of Embryo Development

Siliques of different lengths were dissected on a microscope slide using syringe needle under a dissecting microscope (model Stemi-SV8, Zeiss, Germany). Embryos were cleared with a drop of clearing solution (240 g of chloral hydrate and 30 g of glycerol in 90 mL of water) for 30 min at room temperature. Preparations were examined with a microscope (model BHS, Olympus Optical, Tokyo) equipped for differential interference contrast (model BH2-NIC, Olympus). Images were captured and processed as previously described (Park et al., 1998).

Upon request, all novel materials described in this publication will be made available in a timely manner for noncommercial research purposes, subject to the requisite permission from any third-party owners of all or parts of the material. Obtaining any permissions will be the responsibility of the requestor.

Sequence data from this article have been deposited with the EMBL/GenBank data libraries under accession numbers N515928, At4g18040, AY127014, AL096859, At3g46740, NM_116977, AY585655, AAT08975, At4g09080, BK005428, and BX827624.

ACKNOWLEDGMENTS

We thank Gabino Rios and Csaba Koncz for assisting in the identification of the *toc75-III-1* mutant and Jonathan Jones for the *toc75-IV-1* mutant. We are grateful to Natalie Allcock and Stefan Hyman for transmission electron microscopy, Jocelyn Bédard for assistance during the annotation of *atTOC75-I*, Colin Ferris for help with the phylogenetic analysis, and Henrik Aronsson, Jocelyn Bédard, Sabina Kovacheva, and Sybille Kubis for their comments on the manuscript. We thank Syngenta and SIGnAL for providing additional T-DNA lines. Funding for the SIGnAL indexed insertion mutant collection was provided by the National Science Foundation, and the seed were distributed by the Nottingham Arabidopsis Stock Centre. We are grateful to Neil Hoffman, Kenneth Keegstra, and Steven Theg for kindly providing antibodies and to the Kazusa DNA Research Institute for the *atTOC75-III* cDNA clone, APZL59c10R.

Received March 23, 2005; revised April 26, 2005; accepted April 27, 2005; published May 20, 2005.

LITERATURE CITED

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, et al (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana.* Science **301:** 653–657

Apuya NR, Yadegari R, Fischer RL, Harada JJ, Zimmerman JL, Goldberg RB (2001) The Arabidopsis embryo mutant *schlepperless* has a defect in the *chaperonin-60α* gene. Plant Physiol **126:** 717–730

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

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

Becker T, Jelic M, Vojta A, Radunz A, Soll J, Schleiff E (2004) Preprotein recognition by the Toc complex. EMBO J **23:** 520–530

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

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

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

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

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana.* Plant J **16:** 735–743

Constan D, Patel R, Keegstra K, Jarvis P (2004) An outer envelope membrane component of the plastid protein import apparatus plays an essential role in *Arabidopsis.* Plant J **38:** 93–106

Dávila-Aponte JA, Inoue K, 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 LK, Strathern JN, Garfinkel DJ (1991) RNA-mediated recombination in *S. cerevisiae.* Cell **67:** 355–364

Drouin G, 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, Soll J (2002) A

Toc75-like protein import channel is abundant in chloroplasts. EMBO Rep **3:** 557–562

Edwards K, Johnstone C, Thompson C (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Res **19:** 1349

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

Fitzpatrick LM, Keegstra K (2001) A method for isolating a high yield of *Arabidopsis* chloroplasts capable of efficient import of precursor proteins. Plant J **27:** 59–65

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

Häger KP, Müller B, Wind C, Erbach S, Fischer H (1996) Evolution of legumin genes: loss of an ancestral intron at the beginning of angiosperm diversification. FEBS Lett **387:** 94–98

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser **41:** 95–98

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

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

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

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

Hörmann F, Küchler M, Sveshnikov D, Oppermann U, Li Y, Soll J (2004) Tic32, an essential component in chloroplast biogenesis. J Biol Chem **279:** 34756–34762

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

Inoue K, Potter D (2004) The chloroplastic protein translocation channel Toc75 and its paralog OEP80 represent two distinct protein families and are targeted to the chloroplastic outer envelope by different mechanisms. Plant J **39:** 354–365

Ivanova Y, Smith MD, Chen K, Schnell DJ (2004) Members of the Toc159 import receptor family represent distinct pathways for protein targeting to plastids. Mol Biol Cell **15:** 3379–3392

Jackson DT, Froehlich JE, Keegstra K (1998) The hydrophilic domain of Tic110, an inner envelope membrane component of the chloroplastic protein translocation apparatus, faces the stromal compartment. J Biol Chem **273:** 16583–16588

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

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

Jarvis P, Robinson C (2004) Mechanisms of protein import and routing in chloroplasts. Curr Biol **14:** R1064–R1077

Jarvis P, Soll J (2001) Toc, Tic, and chloroplast protein import. Biochim Biophys Acta **1541:** 64–79

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

Kessler F, Blobel G, Patel HA, Schnell DJ (1994) Identification of two GTP-binding proteins in the chloroplast protein import machinery. Science **266:** 1035–1039

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

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

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

Kovacheva S, Bédard J, Patel R, Dudley P, Twell D, Ríos G, Koncz C, Jarvis P (2005) *In vivo* studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. Plant J **41:** 412–428

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

Kubis S, Patel R, Combe J, Bédard J, Kovacheva S, Lilley K, Biehl A, Leister D, Ríos G, Koncz C, et al (2004) Functional specialization amongst the Arabidopsis Toc159 family of chloroplast protein import receptors. Plant Cell **16:** 2059–2077

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

McElver J, Tzafrir I, Aux G, Rogers R, Ashby C, Smith K, Thomas C, Schetter A, Zhou Q, Cushman MA, et al (2001) Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. Genetics **159:** 1751–1763

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

Nada A, Soll J (2004) Inner envelope protein 32 is imported into chloroplasts by a novel pathway. J Cell Sci **117:** 3975–3982

Olsen LJ, Theg SM, Selman BR, Keegstra K (1989) ATP is required for the binding of precursor proteins to chloroplasts. J Biol Chem **264:** 6724–6729

Park SK, Howden R, Twell D (1998) The Arabidopsis thaliana gametophytic mutation gemini pollen1 disrupts microspore polarity, division asymmetry and pollen cell fate. Development **125:** 3789–3799

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

Pfanner N, Craig EA, Hönlinger A (1997) Mitochondrial preprotein translocase. Annu Rev Cell Dev Biol **13:** 25–51

Porra RJ, Thompson WA, Kriedemann PE (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

Purugganan MD, Wessler SR (1994) Molecular evolution of *magellan*, a maize Ty3/*gypsy*-like retrotransposon. Proc Natl Acad Sci USA **91:** 11674–11678

Rawyler A, Meylan M, Siegenthaler PA (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

Reumann S, Davila-Aponte J, 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

Ríos G, Lossow A, Hertel B, Breuer F, Schaefer S, Broich M, Kleinow T, Jasik J, Winter J, Ferrando A, et al (2002) Rapid identification of *Arabidopsis* insertion mutants by non-radioactive detection of T-DNA tagged genes. Plant J **32:** 243–253

Rodriguez CM, Freire MA, Camilleri C, 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

Schleiff E, Eichacker LA, Eckart K, Becker T, Mirus O, Stahl T, Soll J (2003a) Prediction of the plant beta-barrel proteome: a case study of the chloroplast outer envelope. Protein Sci **12:** 748–759

Schleiff E, Klösgen RB (2001) Without a little help from 'my' friends: direct insertion of proteins into chloroplast membranes? Biochim Biophys Acta **1541:** 22–33

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

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

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

Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B,

Ho P, Bacwaden J, Ko C, et al (2002) A high-throughput Arabidopsis reverse genetics system. Plant Cell **14:** 2985–2994

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

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

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

Swofford DL (2003) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4. Sinauer Associates, Sunderland, MA

Tissier AF, Marillonnet S, Klimyuk V, Patel K, Torres MA, Murphy G, Jones JD (1999) Multiple independent defective *Suppressor-mutator* transposon insertions in Arabidopsis: a tool for functional genomics. Plant Cell **11:** 1841–1852

Tranel PJ, Froehlich J, Goyal A, 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

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

Tu SL, Chen LJ, Smith MD, Su YS, Schnell DJ, Li HM (2004) Import pathways of chloroplast interior proteins and the outer-membrane protein OEP14 converge at Toc75. Plant Cell **16:** 2078–2088

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

Vojta A, Alavi M, Becker T, Hormann F, Küchler M, Soll J, Thomson R, Schleiff E (2004) The protein translocon of the plastid envelopes. J Biol Chem **279:** 21401–21405