The Maize *tha4* Gene Functions in Sec-Independent Protein Transport in Chloroplasts and Is Related to *hcf106*, *tatA*, and *tatB*

Macie B. Walker, Laura M. Roy, Eric Coleman, Rodger Voelker, and Alice Barkan

Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403

Abstract. Proteins are translocated across the chloroplast thylakoid membrane by a variety of mechanisms. Some proteins engage a translocation machinery that is derived from the bacterial Sec export system and require an interaction with a chloroplast-localized SecA homologue. Other proteins engage a machinery that is SecA-independent, but requires a transmembrane pH gradient. Recently, a counterpart to this Δ pH mechanism was discovered in bacteria. Genetic studies revealed that one maize protein involved in this mechanism, HCF106, is related in both structure and function to the bacterial *tatA* and *tatB* gene products. We describe here the mutant phenotype and molecular cloning of a second maize gene that functions in the Δ pH mechanism. This gene, *thylakoid assembly 4 (tha4*), is required specifically for the translocation of proteins that engage the Δ pH pathway. The sequence of the *tha4* gene product resembles those of the maize *hcf106* gene and the bacterial *tatA* and *tatB* genes. Sequence comparisons suggest that *tha4* more closely resembles *tatA*, and *hcf106* more closely resembles *tatB*. These findings support the notion that this sec-independent translocation mechanism has been highly conserved during the evolution of eucaryotic organelles from bacterial endosymbionts.

Key words: protein export • thylakoid • *Mutator* • maize • Sec-independent

THE targeting of proteins to the chloroplast thylakoid membrane and the bacterial cytoplasmic membrane involves several conserved mechanisms (for reviews see Settles and Martienssen, 1998; Dalbey and Robinson, 1999; Keegstra and Cline, 1999). Chloroplasts contain homologues of the bacterial proteins SecA and SRP54 (cpSecA and cpSRP54), both of which function in the targeting of specific substrates to the thylakoid membrane. Other thylakoid proteins engage a targeting machinery that was, until recently, thought to be unique to the chloroplast. This mechanism, referred to as the Δ pH pathway because of its dependence upon a transmembrane pH gradient (Cline et al., 1992), functions in the absence of nucleoside triphosphates and soluble factors in vitro (for review see Schnell, 1998). NH_2-terminal signal sequences target proteins to either the chloroplast Δ pH or cpSecA systems. These resemble bacterial signal sequences and are cleaved by a signal peptidase in the thylakoid lumen (for review see Schnell, 1998). Much attention has been focused on deciphering the features within the signal sequence that distinguish Δ pH from cpSecA substrates. One clear discriminant is the twin arginine motif: two arginine residues immediately precede the hydrophobic domain of the signal sequence in Δ pH pathway substrates, and both residues are essential to engage the Δ pH machinery (Chaddock et al., 1995; Henry et al., 1997). Previously, just one protein that functions in the thylakoid Δ pH system had been identified, HCF106 in maize (Voelker and Barkan, 1995; Settles et al., 1997).

Recently, a Sec-independent export mechanism was discovered in bacteria that is related to the thylakoid Δ pH system. This mechanism functions in the export of proteins that bind complex redox cofactors (for reviews see Settles and Martienssen, 1998; Dalbey and Robinson, 1999). That export of such proteins might involve a mechanism related to the thylakoid Δ pH system was first suggested by the fact that many secreted redox proteins have signal sequences with a twin arginine motif, as do Δ pH pathway substrates (Berks, 1996). This relationship was recently confirmed in several ways. First, the export of several *Escherichia coli* proteins whose signal sequences have the twin arginine motif requires the function of the *tatA* and *tatB* genes, which have sequence similarity to maize *hcf106* (Sargent et al., 1998; Weiner et al., 1998). Second, bacterial

Address correspondence to Alice Barkan, University of Oregon, Institute of Molecular Biology, 1370 Franklin Boulevard, Klamath 297, Eugene, OR 97403. Tel.: (541) 346-5145. Fax: (541) 346-5891. E-mail: abarkan@molbio.uoregon.edu

L.M. Roy's current address is Department of Biology, University of California, San Diego, CA 92093.

E. Coleman's current address is 345 Cinnamon Drive, Satellite Beach, FL 32937.

R. Voelker's current address is Oregon Department of Agriculture, Salem, OR 97310.

twin arginine signal sequences direct the transport of proteins via the thylakoid Δ pH system in chloroplasts (Mori and Cline, 1998; Wexler et al., 1998). The bacterial version of this system has been named the tat[1] system because it is involved in twin arginine translocation (Sargent et al., 1998).

The Δ pH mechanism has attracted considerable attention because it differs in fundamental ways from the intensively studied Sec mechanism. For example, translocation via the Δ pH system in vitro requires neither nucleoside triphosphates nor soluble factors (for review see Schnell, 1998), and there is evidence that both the bacterial and thylakoid systems can accommodate proteins with tertiary structure (Creighton et al., 1995; Clark and Theg, 1997; Hynds et al., 1998; Santini et al., 1998; Dalbey and Robinson, 1999). Nonetheless, the thylakoid membrane remains highly impermeable to ions during the translocation of proteins via the Δ pH pathway (Teter and Theg, 1998). The structure and mechanism of the Δ pH/tat translocation machinery are unknown. To identify additional proteins that are involved in this process, we sought new mutations in maize that specifically disrupt the thylakoid Δ pH system. We describe here the phenotype of one such mutation, which defines a new gene, *thylakoid assembly 4* (*tha4*). The *tha4* gene was cloned by transposon tagging, revealing a gene product that resembles the plant protein HCF106 and bacterial proteins implicated in the tat export mechanism.

Materials and Methods

Plant Material

tha4-ml was recovered in a screen of F2 families derived from a maize line with active *Mutator* (*Mu*) transposons. Numerous *tha4-m1*/+ plants were propagated in parallel for several generations by crossing to inbred lines. Heterozygous plants were self-pollinated to recover homozygous mutant seedlings. *tha4* mutant seedlings used for DNA and RNA extraction were identified initially by their subtle chlorophyll deficiency, and then confirmed by immunoblot analysis of leaf proteins. Plants used in these experiments were grown for 10–12 d at 26°C in 16 h of light (400 μE/m^2/s) and 8 h of dark. Basal leaf tissue used for RNA extraction was obtained from the basal 0.5 cm of the second leaf of 10-d-old seedlings (inbred line B73) grown according to this regime. Etiolated leaf tissue used for RNA extraction was obtained from B73 seedlings that were grown in the complete absence of light for 9 d. *tha4* was mapped to chromosome 1L by crossing *tha4-m1*/+ plants with stocks harboring various B-A translocations (Beckett, 1978).

Extraction and Analysis of DNA and RNA

Total DNA was extracted from maize seedlings and analyzed by Southern hybridization with digoxigenin-labeled *Mu* probes, as described previously (Voelker et al., 1997). Total leaf RNA was extracted with TRIzol Reagent (GIBCO BRL) according to the manufacturer's instructions. RNA gel blot hybridizations were performed as described previously (Barkan, 1998). Because the amount of *tha9*-specific cDNA sequence was insufficient to generate a sensitive gene-specific probe for RNA gel blots (~40 nucleotides), RNase protection assays were used to distinguish the *tha4* and *tha9* mRNAs in different maize tissues. RNase protection assays were performed as described in Voelker et al. (1997). The radioactive antisense probes were generated from near full-length *tha4* and *tha9* cDNA sequences. The *tha4* probe was derived by in vitro transcription of the *tha4* cDNA clone encoding the complete *tha4* open reading frame in pGEM-3Z (see Isolation and Analysis of cDNA). The *tha9* probe was generated by in vitro transcription of a derivative of the *tha9* cDNA clone described below, which had been truncated to delete sequences corresponding to the polyA tail.

Cloning of Genomic DNA

To clone the 1.9-kb XhoI fragment containing the *Mu1* insertion linked to *tha4-m1*, DNA from a homozygous *tha4-ml* mutant was first digested with XhoI and fractionated in an agarose gel. DNA was extracted from a gel slice containing DNA fragments of 1.8–2 kb, by using QIAEX beads according to the manufacturer's instructions (Qiagen). The DNA was ligated into pBluescript SK+ (Stratagene) that had been treated with XhoI and calf intestine phosphatase and electroporated into XL1-Blue MRF′ cells (Stratagene). Colony lifts were probed with a radiolabeled *Mu1* probe, leading to the identification of clone A (see Fig. 4 A).

A genomic library derived from the maize inbred line B73 (a gift of Doug Rice, Pioneer Hi-Bred, Johnston, Iowa) was screened to obtain sequence information both upstream and downstream of clone A. A radiolabeled 190-bp fragment derived from gene-specific sequences in the 3′ untranslated region (UTR) of the *tha4* cDNA (see next section for cDNA isolation) identified two overlapping genomic clones, which contained clone A sequences within 2.5- and 11-kb XbaI fragments, respectively. The 11-kb XbaI fragment was digested with SacI to yield a 2-kb fragment containing clone A sequences. The 2.5-kb XbaI and 2-kb SacI fragments were subcloned into a modified pBluescript SK+ vector and used as templates for sequencing. DNA sequences were analyzed by Yanling Wang in the Institute of Molecular Biology DNA Sequencing Facility (University of Oregon, Eugene, OR).

Primers used for PCR analysis of the *tha4* locus in the revertant sector (see Fig. 4) were as follows: primer M, 5′ CGAAATGGCACCGTGTTACAC 3′; primer N, 5′ GGGAACCACCACGGGTATC 3′; and *Mu* primer, 5′ AGAGAAGCCAACGCCAWCGCCTCYATTT 3′.

Isolation and Analysis of cDNA

To isolate a *tha4* cDNA, a maize leaf cDNA library (Fisk et al., 1999) was screened by PCR using primers designed to amplify the 3′ ends of *tha4* cDNAs. For this purpose, *tha4* gene primers were chosen that mapped to sequences in clone A encoding amino acids just downstream of the predicted transmembrane domain. The initial PCR used a *tha4* gene primer (5′ CCAAGCAGCTCCCCGAGATC 3′) and a vector primer (5′ AGGGTTTTCCCAGTCACGAC 3′) according to the following profile: 94°C/4 min, followed by 30 cycles of 94°C/1 min, 60°C/45 s, 72°C/2 min, and a final extension at 72°C/5 min. A second round of PCR was performed with a nested *tha4* gene primer (5′ ATCGGCAAGACCGTCAAGAGC 3′) and an EcoRI-oligo dT primer (5′CGGAATTC$(T)_{17}$) according to the profile: 94°C/4 min, followed by an initial two cycles of 94°C/1 min, 37°C/45 s, 72°C/2 min, followed by 30 cycles of 94°C/1 min, 60°C/45 s, 72°C/2 min, and a final extension at 72°C/5 min. Amplifications were performed in 50-μl reactions containing 50 mM KCl, 10 mM Tris-HCl, pH 9, 1% Triton X-100, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 5 μl cDNA library stock or 1 μl of initial PCR, and 250 ng of each primer plus *Taq* DNA polymerase. The PCR product was cloned into pBluescript SK+ and its DNA sequence was determined. The cDNA library was screened using this PCR clone as a probe, yielding two types of cDNA encoding highly similar proteins but with distinct 3′ UTR sequences. This revealed the presence of a closely related gene in the maize genome, which we named *tha9*. Authentic *tha4* cDNAs were distinguished from tha9 cDNAs by four criteria: (1) their 3′ UTR detected an mRNA that accumulated to reduced levels in *tha4* mutant seedlings (see Fig. 5); (2) their 3′ UTR detected only authentic *tha4* genomic clones (i.e., that matched clone A) when used to screen a genomic library of B73 DNA; (3) a nearly full-length, spliced *tha4* cDNA obtained by reverse transcriptase-PCR (see below) was identical throughout a 382 bp region of overlap to genomic clone A; and (4) partial unspliced *tha4* cDNAs recovered from the cDNA library contained intron sequence that matched intron sequence in the *tha4* clone obtained from the B73 genomic library.

Because library screens failed to yield full-length *tha4* cDNAs, a cDNA containing the entire coding region of the *tha4* mRNA was obtained by reverse transcription–PCR amplification of poly $(A)^+$ seedling leaf RNA from the inbred maize line B73 (Pioneer Hi-Bred). cDNA synthesis was catalyzed by M-MuLV reverse transcriptase (Promega) and primed with a *tha4* 3′ UTR gene-specific primer (5′ CTTCAATACGTAGAAGCTC

1. *Abbreviations used in this paper: Mu, Mutator*; OE16, 16-kD subunit of the oxygen-evolving complex associated with photosystem II; OE23, 23-kD subunit of the oxygen-evolving complex associated with photosystem II; PC, plastocyanin; tat, twin arginine translocation; UTR, untranslated region.

3′). The cDNA was amplified in a PCR reaction containing a gene primer spanning the start codon (5′ AGCAGGCATGGGGATAC 3′) and a nested 3′UTR primer (5′ GGATATGAACTGCTAACTCG3′), by using the profile: 94°C/4 min, followed by 30 cycles of 94°C/1 min, 60°C/45 s, 72°C/1 min, and a final extension at 72°C/5 min. Amplification buffer included 10% glycerol. The PCR product was cloned into pGEM-3Z (Promega). The *tha4* cDNA sequence has been entered in GenBank/EMBL/DDBJ under accession number AF145755. The *tha9* cDNA sequence has been entered in GenBank/EMBL/DDBJ under accession number AF145756.

Sequence Alignments

Sequences of THA4 (accession number AF145755), THA9 (accession number AF145756), HCF106 (accession number AAC01571), TatA (accession number CAA06724), and TatB (accession number CAA06725) were aligned using ClustalW 1.7 (Thompson et al., 1994) and BoxShade (Bioinformatics Group, ISREC).

Extraction and Analysis of Protein and Fractionation of Chloroplasts

Methods for the extraction of leaf protein, SDS-PAGE, and immunoblot analysis are described in Barkan (1998). Chloroplasts were isolated as described in Voelker and Barkan (1995) and fractionated as follows. Unless otherwise noted, all buffers included protease inhibitors at the following concentrations: 2 μg/ml aprotinin, 5 mM EDTA, 2 μg/ml leupeptin, 1 μg/ml pepstatin, and 1 mM PMSF. Chloroplasts were lysed by resuspension in ice-cold 10 mM Hepes-KOH, pH 8, at a chlorophyll concentration of 0.3 mg/ml for 10 min. Thylakoid membranes were separated from the stroma by centrifugation for 5 min at 12,000 *g*, and resuspended to a chlorophyll concentration of 0.3 mg/ml in thylakoid resuspension buffer (10 mM Hepes-KOH, pH 8, 100 mM sucrose). For the salt-washed membranes, the membrane fraction was resuspended to a chlorophyll concentration of 0.2 mg/ml chlorophyll in either 150 mM sodium carbonate or 2 M sodium bromide and incubated on ice for 30 min, followed by centrifugation at 12,000 *g* for 5 min. The pellets were rinsed in HS (50 mM Hepes-KOH, pH 8, 330 mM sorbitol), centrifuged at 12,000 *g* for 5 min, and resuspended in HS. Proteins from the salt-washed supernatants were precipitated by the addition of BSA to 0.5 mg/ml and TCA to 10% (wt/vol), incubation on ice for 30 min, and centrifugation at 12,000 *g* for 20 min. The pellet was resuspended in HS. For protease treatments, aliquots of unwashed thylakoid membranes at 0.2 mg/ml chlorophyll in thylakoid resuspension buffer (without protease inhibitors) were treated with either 0.2 mg/ml thermolysin or 0.03 mg/ml proteinase K on ice for 30 min. Thermolysin was inactivated by the addition of 1/20 vol of 0.5 M EDTA and 3 vol of thermolysin stop buffer (HS plus 20 mM EDTA). Proteinase K was inactivated by the addition of 1/20 vol of 40 mM PMSF, 1/10 vol of 0.1 M EGTA, and 3 vol of proteinase K stop buffer (HS plus 10 mM EGTA). Samples were centrifuged at 12,000 *g* for 5 min, and the membrane pellets were resuspended in thermolysin stop buffer or proteinase K stop buffer, as appropriate. In control experiments, to test the protease accessibility of lumenal proteins in the absence of intact thylakoid membranes, Triton X-100 (0.1%) was added before protease addition.

Antisera

The near identity of THA4 and THA9 prevented the development of a THA4-specific antiserum (see Fig. 6 B). An antigen was generated by ligating a fragment of the *tha9* cDNA encoding the COOH-terminal 63 amino acids (amino acids 106–169) into the pET28-C(+) vector (Novagen), to yield an in-frame fusion with sequences encoding an NH_2-terminal 6× histidine tag. The fusion protein was purified on a nickel column and injected into rabbits for the production of polyclonal antiserum. The antiserum obtained does not recognize HCF106 (data not shown).

A recombinant HCF106 fragment was generated by subcloning a 0.8-kb SacI-HindIII fragment from an *hcf106* cDNA clone into the pET28-C(+) vector (Novagen), to yield an in-frame fusion with sequences encoding an NH_2-terminal 6× histidine tag. This fusion protein included amino acids 75–243 of the predicted HCF106 gene product. The fusion protein was purified on a nickel column and injected into rabbits for the production of polyclonal antiserum. The anti-HCF106 antiserum does not detect THA4/THA9 on immunoblots (data not shown).

Antisera to the 33-, 23-, and 16-kD subunits of the oxygen-evolving complex (OE33, OE23, and OE16, respectively) and to plastocyanin (PC) were described previously (Voelker and Barkan, 1995).

Results

A Mutation in the tha4 Gene Disrupts the Δ pH–Dependent Thylakoid Targeting Pathway

The reference allele of *tha4, tha4-m1*::*Mu1* (hereafter referred to as *tha4-m1*) was detected in a seedling screen of the F2 progeny of maize plants with active *Mu* transposons. Homozygous mutant seedlings were subtly chlorophyll-deficient and died after the development of three to four leaves. In these ways, *tha4* mutants resembled many previously described maize mutants that lack subsets of thylakoid membrane proteins (for review see Barkan, 1998). Immunoblot data presented previously (Roy and Barkan, 1998) showed that *tha4-m1* mutants accumulate only 20% of the normal levels of the core subunits of photosystems I and II and the cytochrome b_6f complex. However, the thylakoid ATP synthase and the major light harvesting chlorophyll a/b binding protein accumulate normally. The pigmentation and thylakoid protein deficiencies of *tha4-m1* mutants were similar to those of *hcf106* and *tha1* mutants, which have defects in the translocation of proteins across the thylakoid membrane (Voelker and Barkan, 1995). To explore the possibility that *tha4-m1* likewise disrupts thylakoid protein targeting, the abundance and processing of proteins found in the thylakoid lumen were assessed with immunoblots. Two cpSecA substrates, PC and OE33, accumulated to normal levels in *tha4-m1* mutants (Fig. 1). In contrast, OE23 and OE16, which are translocated via the Δ pH pathway, accumulated to only 10–20% of wild-type levels. Furthermore, antibodies to OE23 and OE16 detected higher molecular weight proteins in *tha4-m1* mutants with the sizes predicted for precursors that have retained their lumenal targeting sequences. Thus, the *tha4-m1* phenotype strongly resembles that of *hcf106* mutants, in which the Δ pH pathway is disrupted (Voelker and Barkan, 1995).

Increased accumulation of incompletely processed OE23 and OE16 could result from either a defect in the lumenal processing protease or in the movement of the proteins across the thylakoid membrane. To distinguish between these possibilities, the intrachloroplast location of the precursors was determined by fractionating *tha4-m1* mutant chloroplasts to separate the stroma from the thylakoid membrane vesicles. Precursors of both OE23 and OE16 were enriched in the stromal fraction, whereas the mature

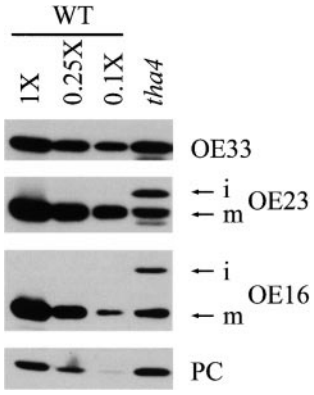

Figure 1. Immunoblot analysis of lumenal proteins in *tha4-m1* mutants. 5 μg of total leaf protein or the indicated dilutions of the wild-type (WT) sample were analyzed. Proteins were detected with monospecific antisera for OE33, OE23, OE16, or PC. Arrows indicate intermediate (i) or mature (m) forms of OE23 and OE16.

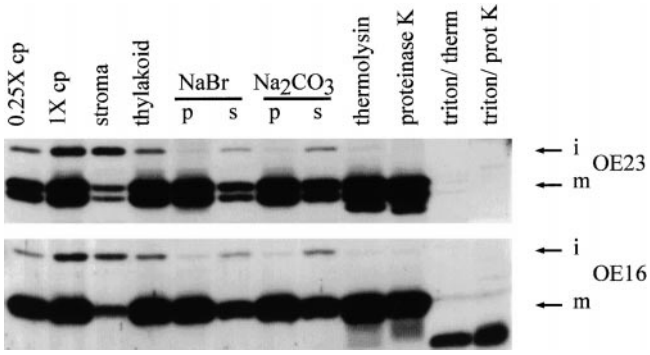

Figure 2. Localization of precursors to OE16 and OE23 in fractionated *tha4-m1* chloroplasts. Chloroplasts (cp) from mutant seedlings were hypotonically lysed and centrifuged to separate thylakoid membrane vesicles (thylakoid) from stroma. To remove peripheral proteins on the stromal face of the membrane, the thylakoid fraction was washed with either Na_2CO_3 or NaBr. The resulting pellet (p) and supernatant (s) are indicated. Alternatively, the thylakoid fraction was treated with proteinase K or thermolysin with and without Triton X-100. Proteins were detected on immunoblots with monospecific antisera for OE23 or OE16. With the exception of lane 1 (0.25× cp), each lane contains protein isolated from the same number of chloroplasts. The stromal intermediate (i) and mature (m) forms of OE23 and OE16 are indicated. Triton was included in control protease reactions to demonstrate that mature OE23 and OE16 are susceptible to protease in the absence of intact membranes. Maize OE23 is often detected as a doublet, perhaps because it is encoded by a gene family (our unpublished results).

forms were found predominantly in the thylakoid membrane fraction (Fig. 2). A small proportion of each precursor remained bound to the thylakoid membrane. To determine whether these bound precursors had been translocated across the membrane, the stromal face of the thylakoid vesicles was treated with carbonate, NaBr, or proteases. The carbonate and NaBr treatments caused some disruption of the vesicles, as revealed by the recovery of a small fraction of mature OE23 and OE16 in the supernatant; wild-type thylakoid membranes treated in the same fashion behaved similarly (see Fig. 7 B). Nonetheless, it is clear that the membrane-bound precursors were extrinsic proteins on the stromal face of the membrane: they were removed by treatment with carbonate or NaBr, and were selectively degraded by the proteases thermolysin or proteinase K (Fig. 2). These results provide strong evidence that the accumulation of incompletely processed OE23 and OE16 in *tha4-m1* mutants results from a defect in their translocation to the thylakoid lumen.

Taken together, these results strongly suggest that *tha4*, like *hcf106*, functions in the Δ pH–dependent system for translocating proteins across the thylakoid membrane. The *tha4* gene was mapped to chromosome 1L by crossing with a series of B-A translocation stocks (Beckett, 1978). In contrast, *hcf106* maps to chromosome 2, indicating that *tha4-m1* defines a new gene.

To address the possibility that *tha4-m1* disrupts the Δ pH pathway by interfering with the accumulation of membrane bound HCF106, the abundance and location of HCF106 were monitored in *tha4-m1* mutants. HCF106 accumulates to normal levels in *tha4-m1* mutants (Fig. 3 A). The same *tha4-m1* chloroplast fractions used in Fig. 2 were probed with anti–HCF106 antibody (Fig. 3 B), revealing that HCF106 is tightly associated with the thylakoid membrane in *tha4-m1* mutants. Furthermore, HCF106 is highly susceptible to proteases in *tha4-m1* thylakoid membrane vesicles (Fig. 3 B). This fractionation behavior is the same as that described for HCF106 in wild-type chloroplasts (Settles et al., 1997; see Fig. 7). Therefore, *tha4-m1* does not alter the accumulation of HCF106 or its insertion into the thylakoid membrane.

Molecular Cloning of the tha4 Gene

tha4-m1 arose in a maize line with active *Mu* transposons and was somatically unstable, suggesting that it was caused by the insertion of a *Mu* transposon. To identify *Mu* insertions that were genetically linked to *tha4*, DNA from mutants derived from diverse branches of the *tha4-m1* pedigree were analyzed by Southern hybridization, using probes corresponding to each member of the *Mu* family (Bennetzen, 1996). A *Mu1* probe detected a 1.9-kb XhoI fragment in all mutants that was absent in closely related homozygous wild-type plants (data not shown). Probes corresponding to the other members of the *Mu* family failed to detect a genetically linked insertion.

The 1.9-kb XhoI fragment was cloned from a size-enriched genomic library of *tha4-m1* mutant DNA (Fig. 4 A, clone A). Southern blots of wild-type and *tha4-m1* mutant DNAs were probed with the genomic sequence flanking the cloned *Mu1* insertion, revealing that all mutant and no wild-type plants were homozygous for the cloned insertion (data not shown). Longer clones corresponding to this region were isolated from a genomic library of wild-type DNA. To test whether clone A contained the insertion that is the cause of the *tha4-m1* mutant phenotype, the structure of the corresponding genomic region was monitored in a revertant sector that appeared on a *tha4-m1* mutant leaf. DNA extracted from a dark green revertant sector and from the flanking, slightly paler mutant tissue was analyzed with PCR using primer pairs designed to selec-

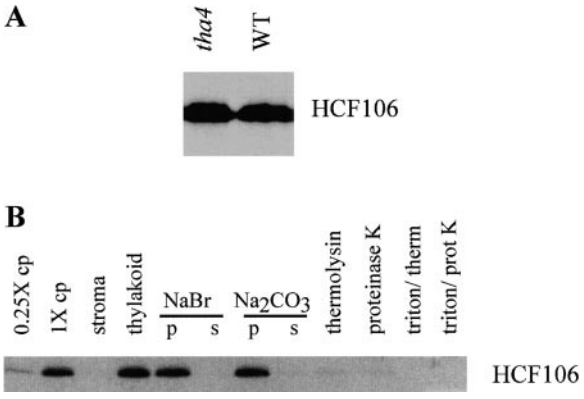

Figure 3. Immunoblot analysis of HCF106 in *tha4-m1* mutants. (A) 5 μg of *tha4-m1* and wild-type (WT) leaf protein were fractionated and probed with antiserum specific for HCF106. (B) The immunoblot shown in Fig. 2 was stripped and reprobed with antiserum specific for HCF106.

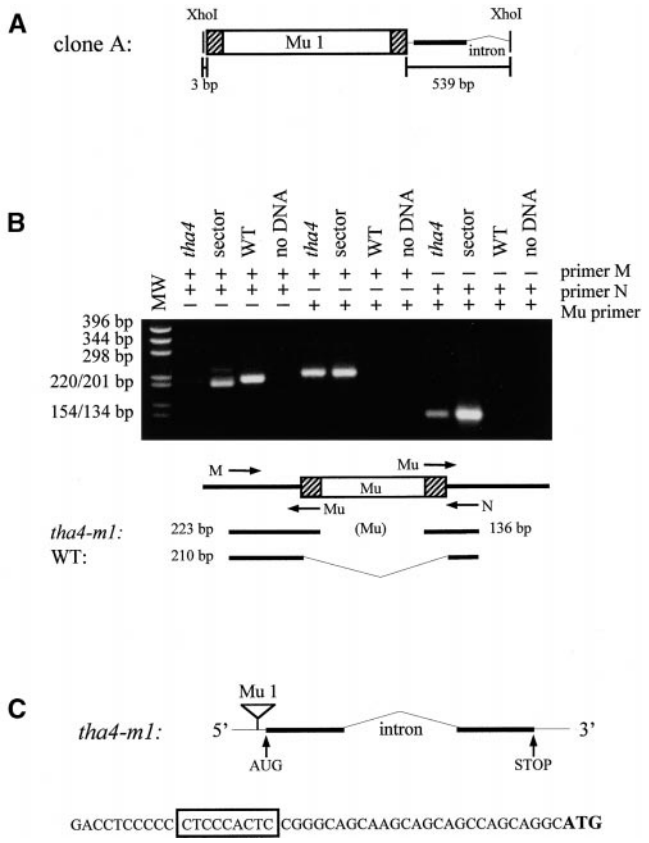

Figure 4. Molecular cloning of the *tha4* gene. (A) Map of the original genomic clone of *tha4-m1*. Clone A was a 1.9-kb XhoI fragment that included the 1.4-kb *Mu1* transposon and 542 bp of flanking sequence. Comparison with cDNA sequence identified 129 bp of intron sequence within this clone. The bold line represents the translated sequence. The XhoI site on the left side of the clone was created upon insertion of the *Mu* transposon. (B) PCR analysis of the cloned genomic region in a somatic revertant sector. DNA was extracted from a revertant sector (sector), from mutant tissue on the same leaf (*tha4*), and from a +/+ sibling (WT). Control reactions lacked template DNA (no DNA). Arrows represent primer binding sites. The sizes of the amplification products predicted for each primer pair are illustrated. (C) Site of *Mu1* insertion in *tha4-m1*. Untranslated and translated regions are represented by fine and bold lines, respectively. The *Mu1* insertion is 35 bp upstream of sequences encoding the putative start codon. The known intron is diagrammed; other introns may be encoded in sequences downstream of the genomic clone. The sequence surrounding the *Mu* insertion site is shown beneath the map, with the box indicating the 9-bp that were duplicated upon *Mu* insertion. The putative start codon is shown in bold.

tively amplify either the mutant or wild-type allele (Fig. 4 B).

Control reactions first established that the predicted amplification products were obtained with homozygous mutant and wild-type DNA samples. As expected, DNA from homozygous wild-type tissue (WT) gave no amplification products when the *Mu* primer was used in conjunction with either of the gene-specific primers, M or N. Primers M and N together, however, gave rise to an amplification product of 210 bp with a WT template DNA, the size predicted for the wild-type allele. Amplification of DNA from the mutant tissue on the sectored leaf (Fig. 4 B, *tha4*) with the *Mu* primer in conjunction with primers M or N resulted in the predicted DNA fragments of 223 and 136 bp, respectively. As expected, the *tha4* mutant DNA did not yield an abundant product when the gene-specific primers M and N were paired because PCR fails to amplify across intact *Mu* elements.

Revertant sectors are expected to be heterozygous and should, therefore, give rise to the products representing both alleles. With revertant DNA as a template, primer M or N paired with the *Mu* primer gave rise to the 223- or 136-bp fragments expected for the mutant allele (Fig. 4 B). The key finding was that the revertant DNA also contained an allele lacking the cloned *Mu* insertion: primer M paired with primer N yielded a robust amplification product. This product was slightly smaller than that resulting from a wild-type DNA template, indicating that it did not result from contamination with wild-type DNA. These results suggest that imprecise excision of *Mu1* caused a small deletion of flanking genomic sequences, and that this excision, nonetheless, restored *tha4* gene function. As described below, the *Mu* insertion in *tha4-m1* disrupts the untranslated sequence in the 5′ portion of the *tha4* gene, such that excision accompanied by a small deletion could well restore gene function. These results indicate that excision of the *Mu1* insertion represented by clone A correlates with reversion to a wild-type phenotype, providing strong evidence that the clone contains a portion of the *tha4* gene.

The genomic sequence flanking the cloned *Mu1* insertion was used to obtain *tha4* cDNAs. The *tha4* cDNA encoded a continuous open reading frame of 170 amino acids. The *Mu1* insertion disrupted sequences mapping 35 bp upstream of those encoding the predicted start codon (Fig. 4 C). A probe prepared from the unique 3′ UTR sequences of the *tha4* cDNA detected a leaf mRNA of ~900 nucleotides that accumulated normally in *hcf106* mutants, but was barely detectable in *tha4-m1* mutants (Fig. 5). The longest *tha4* cDNA obtained began at the predicted start codon and included 740 nucleotides between the start codon and the beginning of the poly(A) tail. Given that

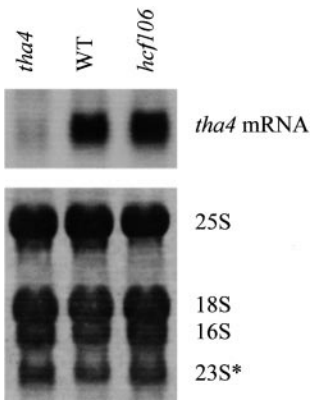

Figure 5. RNA gel blot showing loss of *tha4* mRNA in *tha4-m1* mutant. 20 μg of total leaf RNA from *tha4-m1*, wild-type (WT), or *hcf106* seedlings was gel-fractionated, transferred to nylon membrane, and probed with a radiolabeled 190-bp DNA fragment that hybridizes specifically to the 3′ UTR of the *tha4* mRNA (top). The same filter was stained with methylene blue to visualize the bound rRNAs (bottom). 23S* is a fragment of the chloroplast 23S rRNA.

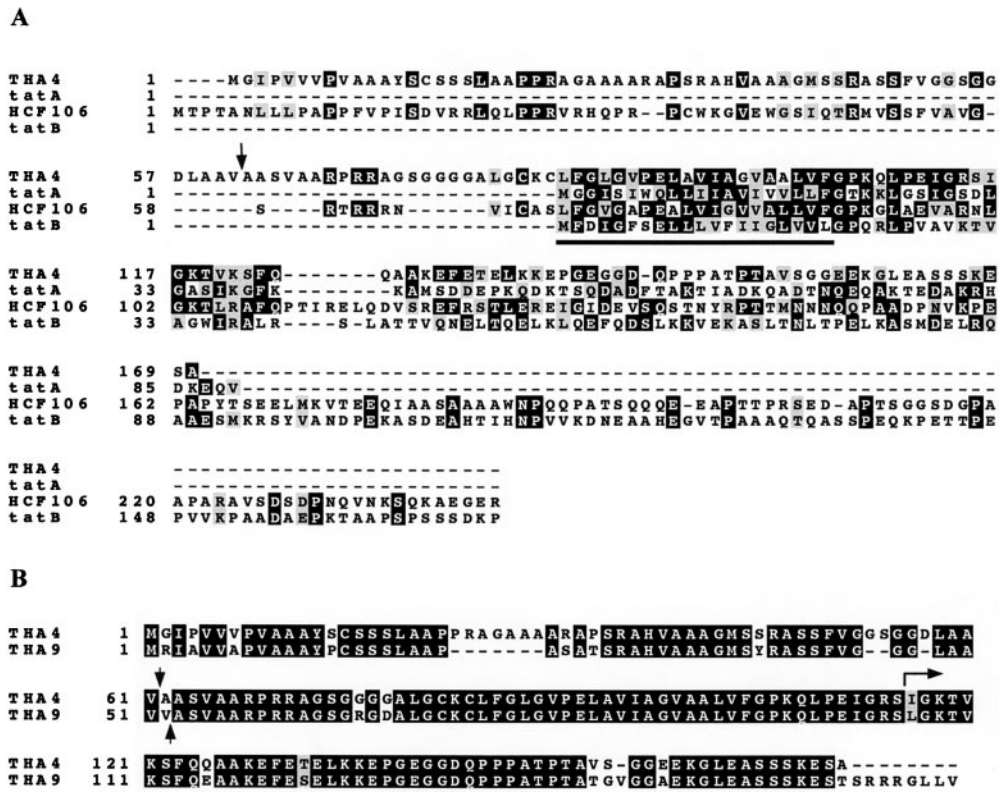

Figure 6. Sequence alignments, calculated with ClustalW 1.7 (Thompson et al., 1994) and BoxShade. The vertical arrows indicate the locations of transit peptide cleavage in THA4 and THA9, as predicted by the ChloroP algorithm (Emanuelsson et al., 1999). (A) Sequence alignment showing the relationship between maize THA4, *E. coli* TatA, maize HCF106, and *E. coli* TatB. The predicted membrane spanning domain is underlined. (B) Sequence alignment showing the relationship between maize THA4 and maize THA9. The horizontal arrow illustrates the portion of THA9 used as an antigen for the production of polyclonal antisera.

the RNA gel blot analysis indicated a length of ~900 nucleotides for the polyadenylated mRNA and that poly(A) tails commonly contain >100 residues in plants, we predict that the 5′ UTR of the *tha4* mRNA contains 50–100 nucleotides. No in-frame ATGs are found in 300 bp of genomic sequence upstream of those encoding the predicted start codon. The *Mu1* insertion, mapping 35 bp upstream of the predicted start codon, therefore, likely disrupts the 5′ UTR of the *tha4* gene.

THA4 Resembles HCF106 and Bacterial Proteins Implicated in Sec-Independent Protein Export

The deduced *tha4* gene product (THA4) is 170 amino acids in length and has a single predicted membrane spanning domain. The ChloroP algorithm (Emanuelsson et al., 1999) predicts that THA4 is a chloroplast-localized protein. THA4 is related to maize HCF106 and to the products of the bacterial *tatA* and *tatB* genes (Fig. 6 A). TatA and TatB are bacterial proteins implicated in Sec-independent protein export (Sargent et al., 1998; Weiner et al., 1998) and *hcf106* functions in the thylakoid Δ pH mechanism (Voelker and Barkan, 1995; Settles et al., 1997). The four proteins are closely related in their membrane-spanning domains (underlined in Fig. 6 A) and in the adjacent amphipathic helical domain (see Settles et al., 1997). The predicted mature form of THA4 is similar in size to TatA, whereas the predicted mature form of HCF106 is similar in size to TatB. THA4 and TatA lack the extended COOH-terminal acidic region found in both HCF106 and TatB. Previously, it was proposed that *hcf106* is more closely related to *tatA* than to *tatB* (Sargent et al., 1998). However, these results suggest that *hcf106* is more closely related to *tatB* and that *tha4* is more closely related to *tatA*. In addition, an *Arabidopsis* cDNA sequence (accession number H37534) proposed previously to represent an *hcf106* homologue (Settles et al., 1997) is, in fact, much more similar to *tha4* (data not shown) and likely represents a *tha4* ortholog.

cDNA library screens yielded two classes of cDNA, representing *tha4* and a closely related gene, which we named *tha9* (see Materials and Methods). Both genes were represented as cDNAs in a seedling leaf cDNA library, indicating that both are transcribed in seedling leaf tissue. The *tha9* cDNA encodes a protein that is very closely related to THA4 (Fig. 6 B), and that is predicted by the ChloroP algorithm to be chloroplast-localized. The predicted mature form of THA9 is nearly identical to that of THA4 (sequence downstream of vertical arrows in Fig. 6 B); even the predicted transit peptides diverge to only a small degree. This degree of identity strongly suggests that these two proteins are localized similarly in the cell and that they have similar or identical functions.

THA4 Is an Integral Thylakoid Membrane Protein with Its COOH Terminus Exposed to the Stroma

The near identity of the mature regions of THA4 and THA9 precluded the generation of a THA4-specific antiserum. A polyclonal antiserum that would detect both THA4 and THA9 was generated to the COOH-terminal region of THA9 (Fig. 6 B, horizontal arrow). The antiserum detected a protein in wild-type leaf tissue that migrated during SDS-PAGE with an apparent mass of 16 kD and accumulated to much reduced levels in *tha4-m1* mutant leaf tissue (Fig. 7 A). The residual protein in the mu-

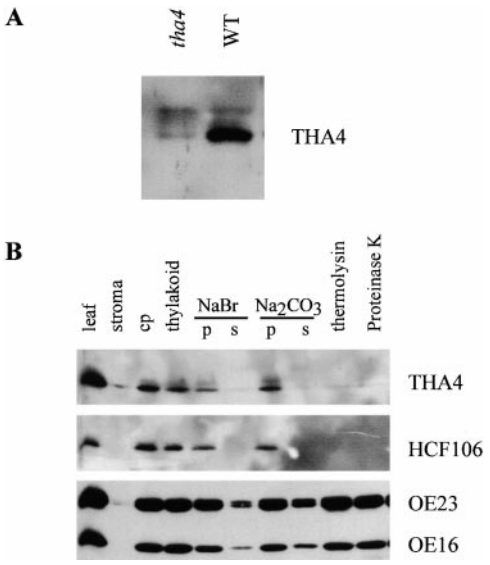

Figure 7. (A) Immunoblot showing loss of THA4 in *tha4-m1* mutants. 5 μg of *tha4-m1* and wild-type (WT) leaf protein were fractionated and probed with antiserum that detects THA4. The doublet detected with this antibody was observed only intermittently. The upper band may represent an alternative protein conformation or, possibly, the *tha9* gene product. (B) Immunoblot showing that THA4 and HCF106 are integrated similarly into the thylakoid membrane. Wild-type chloroplasts were fractionated as in Fig. 2. An immunoblot was probed with antisera that detect THA4 or HCF106. The blot was reprobed with antisera for OE23 and OE16 (bottom) to illustrate the fractionation behavior of lumenal proteins.

tant sample could result from incomplete disruption of *tha4* gene expression by the *Mu1* insertion and/or could be the product of the *tha9* gene. In either case, these findings strongly suggest that the *tha9* gene contributes no more than a small percentage of the immunoreactive protein in wild-type seedling leaves. However, it is also possible that the *tha9* and *tha4* gene products interact in such a way that THA9 is destabilized in the absence of THA4.

These antibodies can be used to localize THA4 in the cell because the bulk of the immunoreactive signal in wild-type leaves most likely represents THA4 (Fig. 7 A). The near identity of THA9 and THA4 suggests that the two proteins are, in any case, localized similarly. For simplicity, the immunoreactive protein in leaf tissue will be referred to as THA4. A substantial proportion of THA4 copurified with intact chloroplasts during sedimentation through Percoll gradients (Fig. 7 B, compare leaf and cp lanes). THA4 was largely recovered in the thylakoid membrane fraction of chloroplasts (Fig. 7 B, stroma versus thylakoid lanes). THA4 is tightly associated with the membrane, as it was not removed by carbonate or NaBr washes. Treatment of thylakoid membrane vesicles with the proteases thermolysin and proteinase K eliminated immunologically reactive material. Because the antiserum was raised to the COOH terminus of the protein and because the THA4 amino acid sequence predicts a short hydrophilic NH_2-terminal domain but a long hydrophilic COOH-terminal domain, these results suggest that the COOH terminus of THA4 is exposed to the stroma. The fractionation behavior of HCF106 was similar to that of THA4 (Fig. 7 B), consistent with the previous report that HCF106 is tightly associated with the membrane and is susceptible to proteases applied to the stromal face (Settles et al., 1997). Thus, THA4 and HCF106 are likely to be oriented similarly in the membrane, with their COOH-terminal acidic tails in the stroma.

The tha4 and tha9 mRNAs Accumulate in Different Ratios in Different Tissue Types

Duplicate gene pairs analogous to *tha4/tha9* are common in maize as a consequence of the tetraploidy of the ancestral maize genome (Ahn and Tanksley, 1993). The near identity of the THA4 and THA9 amino acid sequences suggests that the two proteins have retained similar or identical biochemical functions. Nonetheless, *tha9* function cannot fully compensate for the absence of *tha4* since mutations in *tha4* disrupt the Δ pH–dependent translocation mechanism in seedling leaves. To address the possibility that the two genes have acquired different patterns of regulation, their mRNAs were quantified in green seedling leaf tissue, etiolated seedling leaf tissue, basal leaf tissue, roots, endosperm, and immature ears (Fig. 8). These tissues differ with regard to their plastid populations: green seedling leaf contains mature chloroplasts; basal leaf is enriched in proplastids, the chloroplast progenitors found in undifferentiated cells; etiolated leaf contains etioplasts, which differentiate into chloroplasts upon exposure to light; and root, endosperm, and immature ear contain a variety of nonphotosynthetic plastid forms.

RNase protection assays were used to distinguish the *tha4* and *tha9* mRNAs. The probe generated from the *tha4* cDNA was nearly completely protected by *tha4* mRNA, as expected (Fig. 8 A, *tha4* band). The *tha9* mRNA protected smaller but still substantial fragments of the *tha4* probe, because of their high degree of sequence complementarity (Fig. 8 A, *tha9* bands); this permitted the simultaneous detection of both mRNAs in each lane. In addition, the *tha9* mRNA was assayed with a probe generated from the *tha9* cDNA (Fig. 8 B). The results of these two *tha9* mRNA assays were consistent and served to reinforce one another.

The level of the *tha4* mRNA was substantially reduced in *tha4-m1* mutants (Fig. 8 A, compare *tha4* leaf and WT leaf samples), whereas the accumulation of *tha9* mRNA was unaltered in *tha4* mutants (Fig. 8, A and B). The *tha4* mRNA accumulated to similar levels in green seedling leaves, etiolated seedling leaves, basal leaf, and immature ear (Fig. 8 A). Its level was somewhat lower in endosperm and it was barely detectable in the root (Fig. 8 A).

The profile of *tha9* mRNA accumulation differed significantly from that of *tha4*. The *tha9* mRNA accumulated to the highest level in green leaf tissue, to slightly lower levels in etiolated leaf tissue, and to much lower levels in leaf base, endosperm, immature ear, and root. Most notably, the ratio of *tha4* to *tha9* mRNA was much higher in leaf base and in immature ear than in green leaf (Fig. 8 A and B). The fact that *tha4* mRNA is predominant in basal leaf tissue suggests that *tha4* function is required early in the proplastid to chloroplast transition, when the elaboration of thylakoid membranes is initiated. The *tha9* gene may

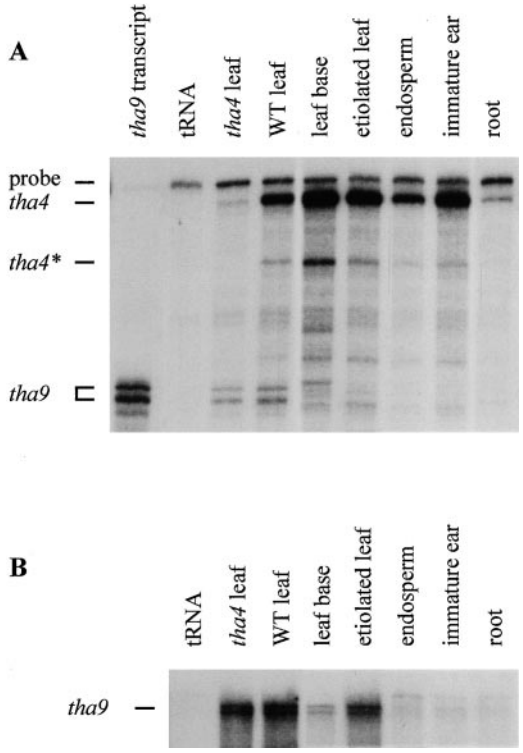

Figure 8. RNase protection analysis of *tha4* and *tha9* mRNA in different maize tissues. (A) Assays using a *tha4* probe. 15 μg of total RNA from the indicated maize tissues was analyzed with an antisense probe generated from the *tha4* cDNA. Control reactions contained 15 μg of tRNA plus 1 pg of a sense *tha9* transcript generated by in vitro transcription of the *tha9* cDNA (first lane) or 15 μg of tRNA alone (second lane). The band designated *tha4* represents the *tha4* mRNA; it is slightly smaller than the full-length probe, which included some vector sequence. The band designated *tha4** also results from protection by the *tha4* mRNA (note its absence in the *tha4* mutant leaf sample); this band may result from a sequence polymorphism between the *tha4* probe and *tha4* mRNA, resulting in inefficient cleavage of the duplexed probe. The cluster of bands designated *tha9* result from protection by the *tha9* mRNA. (B) Assays using a *tha9* probe. 10 μg of total RNA from the indicated maize tissues or 10 μg of tRNA (first lane) was analyzed with an antisense probe generated from the *tha9* cDNA. The band designated *tha9* represents the *tha9* mRNA.

then provide supplemental protein in young and mature chloroplasts to maintain optimal function of the Δ pH–dependent translocation machinery.

Discussion

We have presented evidence that the maize *tha4* gene functions in the Δ pH–dependent mechanism for the translocation of proteins across the chloroplast thylakoid membrane. THA4 is the second plant protein to be identified that participates in this mechanism, the first being the related protein HCF106 (Voelker and Barkan, 1995; Settles et al., 1997). Both THA4 and HCF106 are tightly associated with the thylakoid membrane; sequence analysis and protease sensitivity studies suggest they both have a single transmembrane domain and an acidic COOH-terminal tail on the stromal side of the membrane. However, the stromal domain of HCF106 is considerably longer and more acidic than that of THA4.

THA4 and HCF106 are related to the *tatA* and *tatB* genes found in all sequenced eubacterial genomes. *tatA* and *tatB* were recently implicated in a novel sec-independent mechanism for the export of periplasmic proteins that bind redox cofactors (for review see Dalbey and Robinson, 1999). Alignment of the most conserved regions of these proteins (the transmembrane and amphipathic helical domains) revealed a comparable degree of similarity between all protein pairs. However, these proteins fall into two classes based upon the length of their COOH-terminal hydrophilic tails: HCF106 and TatB have a long tail of nearly identical length, and THA4 and TatA have a much shorter tail, also of nearly identical length (Fig. 6). On this basis, we propose that *tha4* is more closely related to *tatA* and *hcf106* is more closely related to *tatB*.

The similarity between the THA4/TatA and HCF106/TatB sequences raises questions about the functional relationship between these genes. As is true of *tha4* and *hcf106*, mutations in either *tatA* or *tatB* disrupt protein export (Sargent et al., 1998; Weiner et al., 1998), indicating that these genes have at least some distinct functions. The situation is still more complex because many bacterial genomes include a second *tatA*-like gene, termed *tatE* in *E. coli*, and mutations in *tatE* also disrupt export to some extent (Sargent et al., 1998). In maize, both *tha4* and *hcf106* are themselves members of duplicate gene pairs: from a maize leaf cDNA library, we recovered cDNAs encoding a protein with >90% identity to HCF106 (Barkan, A., unpublished results) as well as the *tha9* cDNA described here, which is very closely related to *tha4*. That mutant phenotypes result from disruption of either *hcf106* or *tha4* is consistent with the notion that the members of each gene pair may not be completely redundant in their biochemical function. However, the near identity of the proteins in each pair leads us to favor the notion that the members of each pair play identical biochemical roles, but are subject to different patterns of regulation. This is supported by our finding that the *tha4* and *tha9* mRNAs accumulate differentially in different plant tissues. In any case, a picture is emerging of a group of related proteins involved in this novel protein translocation mechanism. The biochemical role of these proteins is not known, although it has been proposed based upon their orientation in the membrane that they function as receptors (Settles et al., 1997). It will be fascinating to learn how these proteins are organized in the membrane and how they relate in a functional sense to one another.

Mutant phenotypes originally implicated THA4 and HCF106 in the Δ pH–dependent translocation mechanism. Recently, conclusive evidence that these proteins function directly in translocation was obtained from in vitro import assays performed in the presence of anti–THA4 or anti–HCF106 antibodies (Mori et al., 1999). Thus far, THA4 and HCF106 are the only proteins known to participate in the Δ pH–dependent mechanism in chloroplasts. However, by comparison to the bacterial system, it seems likely that a plant homologue of the bacterial TatC protein is also involved. TatC is predicted to be a polytopic membrane protein and is encoded in the same operon as TatA

and TatB in *E. coli* (Sargent et al., 1998). A *tatC* deletion mutant has a severe defect in the export of tat substrates (Bogsch et al., 1998). Genes with similarity to *tatC* are found in the chloroplast genomes of certain algae and plant cDNAs encoding *tatC* homologues have recently appeared in the databases. However, direct evidence for the role of these *tatC* homologues has not been reported. We recently recovered a maize mutant with a more severe defect in the Δ pH–dependent pathway than either *hcf106* or *tha4* mutants, and we have established that the mutation defines a third gene involved in this process (Pedersen, R., M. Walker, and A. Barkan, unpublished results). Whether this new mutation disrupts a *tatC* homologue or defines a novel component of this interesting protein translocation mechanism remains to be determined.

We would like to thank Devin Oglesbee for help in propagating *tha4-m1*, Rob Martienssen (Cold Spring Harbor Laboratory) for antibodies to HCF106 used in the early phases of this work, and Ken Cline (University of Florida) for communicating unpublished data and for helpful discussions. We are extremely grateful to Doug Rice (Pioneer Hi-Bred) for providing the B73 genomic library. We also thank Bethany Jenkins and Dennis McCormac (both from University of Oregon) for providing comments on the manuscript.

This work was supported by a grant from the National Institutes of Health (RO1 GM48179).

Submitted: 29 April 1999
Revised: 7 September 1999
Accepted: 10 September 1999

References

Ahn, S., and S.D. Tanksley. 1993. Comparative linkage maps of the rice and maize genomes. *Proc. Natl. Acad. Sci. USA*. 90:7980–7984.

Barkan, A. 1998. Approaches to investigating nuclear genes that function in chloroplast biogenesis in land plants. *Methods Enzymol.* 297:38–57.

Beckett, J.B. 1978. B-A translocations in maize. *J. Heredity*. 69:27–36.

Bennetzen, J. 1996. The Mutator transposable element system of maize. *Curr. Top. Microbiol. Immunol.* 204:195–229.

Berks, B.C. 1996. A common export pathway for proteins binding complex redox cofactors. *Mol. Microbiol.* 22:393–404.

Bogsch, E.G., F. Sargent, N.R. Stanley, B.C. Berks, C. Robinson, and T. Palmer. 1998. An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J. Biol. Chem*. 273: 18003–18006.

Chaddock, A.M., A. Mant, I. Karnauchov, S. Brink, R.G. Herrmann, R.B. Kloesgen, and C. Robinson. 1995. A new type of signal peptide: central role of a twin-arginine motif in transfer signals for the Δ pH -dependent thylakoidal protein translocase. *EMBO (Eur. Mol. Biol. Organ.) J.* 14:2715–2722.

Clark, S.A., and S.M. Theg. 1997. A folded protein can be transported across the chloroplast envelope and thylakoid membranes. *Mol. Biol. Cell.* 8:923–934.

Cline, K., W.F. Ettinger, and S.M. Theg. 1992. Protein-specific energy requirements for protein transport across or into thylakoid membranes. Two lumenal proteins are transported in the absence of ATP. *J. Biol. Chem*. 26: 2688–2696.

Creighton, A.M., A. Hulford, A. Mant, D. Robinson, and C. Robinson. 1995. A monomeric, tightly folded stromal intermediate on the delta pH-dependent thylakoidal protein transport pathway. *J. Biol. Chem*. 270:1663–1669.

Dalbey, R.E., and C. Robinson. 1999. Protein translocation into and across the bacterial plasma membrane and the plant thylakoid membrane. *Trends Biochem Sci.* 24:17–22.

Emanuelsson, O., H. Nielsen, and G. von Heijne. 1999. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Prot. Science.* 8:978–984.

Fisk, D.G., M.B. Walker, and A. Barkan. 1999. Molecular cloning of the maize gene *crp1* reveals similarity between regulators of mitochondrial and chloroplast gene expression. *EMBO (Eur. Mol. Biol. Organ.) J.* 18:2621–2630.

Henry, R., M. Carrigan, M. McCaffery, X. Ma, and K. Cline. 1997. Targeting determinants and proposed evolutionary basis for the Sec and Δ pH protein transport systems in chloroplast thylakoid membranes. *J. Cell. Biol.* 136:823–832.

Hynds, P.J., D. Robinson, and C. Robinson. 1998. The sec-independent twin-arginine translocation system can transport both tightly folded and malfolded proteins across the thylakoid membrane. *J. Biol. Chem.* 273:34868–34874.

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

Mori, H., and K. Cline. 1998. A signal peptide that directs non-Sec transport in bacteria also directs efficient and exclusive transport on the thylakoid delta pH pathway. *J. Biol. Chem*. 273:11405–11408.

Mori, H., E.J. Summer, X. Ma, and K. Cline. 1999. Component specificity for the thylakoidal Sec and Δ pH–dependent protein transport pathways. *J. Cell Biol.* 146:45–56.

Roy, L.M., and A. Barkan. 1998. A secY homologue is required for the elaboration of the chloroplast thylakoid membrane and for normal chloroplast gene expression. *J. Cell Biol.* 141:385–395.

Santini, C.-L., B. Ize, A. Chanal, M. Mueller, G. Giordano, and L.-F. Wu. 1998. A novel Sec-independent periplasmic protein translocation pathway in *Escherichia coli*. *EMBO (Eur. Mol. Biol. Organ.) J.* 17:101–112.

Sargent, F., E.G. Bogsch, N.R. Stanley, M. Wexler, C. Robinson, B.C. Berks, and T. Palmer. 1998. Overlapping functions of components of a bacterial Sec-independent protein export pathway. *EMBO (Eur. Mol. Biol. Organ.) J.* 13:3640–3650.

Schnell, D.J. 1998. Protein targeting to the thylakoid membrane. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49:97–126.

Settles, A.M., A. Yonetani, A. Baron, D.R. Bush, K. Cline, and R. Martienssen. 1997. Sec-independent protein translocation by the maize Hcf106 protein. *Science.* 278:1467–1470.

Settles, A.M., and R. Martienssen. 1998. Old and new pathways of protein export in chloroplasts and bacteria. *Trends Cell Biol.* 8:494–501.

Teter, S.A., and S.M. Theg. 1998. Energy-transducing thylakoid membranes remain highly impermeable to ions during protein translocation. *Proc. Natl. Acad. Sci. USA*. 95:1590–1594.

Thompson, J., D. Higgins, and T. Gibson. 1994. ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673–4680.

Voelker, R., and A. Barkan. 1995. Two nuclear mutations disrupt distinct pathways for targeting proteins to the chloroplast thylakoid. *EMBO (Eur. Mol. Biol. Organ.) J.* 14:3905–3914.

Voelker, R., J. Mendel-Hartvig, and A. Barkan. 1997. Transposon-disruption of a maize nuclear gene, *tha1*, encoding a chloroplast SecA homolog: in vivo role of cp-SecA in thylakoid protein targeting. *Genetics.* 145:467–478.

Weiner, J.H., P.T. Bilous, G.M. Shaw, S.P. Lubitz, L. Frost, G.H. Thomas, J.A. Cole, and R.J. Turner. 1998. A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell.* 93:93–101.

Wexler, M., E.G. Bogsch, R.B. Klosgen, T. Palmer, C. Robinson, and B.C. Berks. 1998. Targeting signals for a bacterial Sec-independent export system direct plant thylakoid import by the delta pH pathway. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 431:339–342.