Characterization of a cDNA Encoding the Thylakoidal Processing Peptidase from Arabidopsis thaliana

IMPLICATIONS FOR THE ORIGIN AND CATALYTIC MECHANISM OF THE ENZYME*

(Received for publication, September 25, 1997, and in revised form, October 22, 1997)

Balbir K. Chaal‡, Ruth M. Mould§, Adrian C. Barbrook‡, John C. Gray§, and Christopher J. Howe‡*

From the Cambridge Center for Molecular Recognition, ‡Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, United Kingdom and the §Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, United Kingdom

We have identified and sequenced a cDNA containing a complete open reading frame for a putative 340-amino acid precursor of the thylakoidal processing peptidase from Arabidopsis thaliana. The predicted amino acid sequence of the protein includes regions highly conserved among Type I leader peptidases and indicates that the enzyme uses a serine-lysine catalytic dyad mechanism. Phylogenetic analysis indicated a common ancestry of the enzyme with those from oxygenic photosynthetic prokaryotes, suggesting that the cDNA encoded the chloroplast enzyme. The catayltic domain was overexpressed in Escherichia coli, generating a product capable of cleaving the thylakoid-transfer domain from a chloroplast protein. Antibodies to the overexpressed polypeptide cross-reacted with a 30-kDa thylakoid membrane protein.

In plants, chloroplast proteins that function in the thylakoid lumen are generally initially synthesized in the cytoplasm as precursors with amino-terminal presequences, which are responsible for the transport of the precursors into the thylakoid (1). The presequence consists of two distinct domains (2). The first, the envelope-transfer domain, is responsible for the transport of the precursors into the chloroplast membrane into the stroma, where a soluble stromal processing peptidase usually removes this domain, producing an intermediate form of the protein. The second domain of the presequence then directs the transport of the intermediate form across the thylakoid membrane into the lumen where the thylakoidal processing peptidase (TPP)1 removes the thylakoid-transfer domain to produce the mature form of the protein (3). Thylakoid lumen proteins processed by TPP include plastocyanin and the 33-, 23-, and 16-kDa extrinsic proteins of Photosystem II (4). TPP is a thylakoid membrane protein with its active site located at the luminal face (5).

Since TPP cleaves a signal or leader peptide from a precursor protein, it can be referred to as a signal or leader peptidase and belongs to the family of Type I leader peptidases (6). Members of this family include the leader peptidases from Escherichia coli and the cyanobacterium Phormidium laminosum, where the enzyme also processes polypeptides of the cyanobacterial thylakoid lumen (7, 8). An inhibitor of the E. coli enzyme has recently been shown to inhibit the cyanobacterial and chloroplast enzymes (9). Substrate specificity has been shown to be very similar between the plant TPP and the cyanobacterial enzyme (10), and the specificity of plant TPP has been shown to be virtually identical to that of the E. coli leader peptidase, as certain thylakoid lumen precursors have been found to be cleaved at the same sites by the E. coli leader peptidase and TPP, and certain bacterial leader peptides are also cleaved at the same sites by TPP and the E. coli leader peptidase (11). The structures of the bacterial leader peptides and the thylakoid-transfer domain have similar features. These include a positively charged amino-terminal region, a hydrophobic central core and a polar carboxyl-terminal region in which the –3 and –1 residues relative to the cleavage site are short uncharged amino acids (12). Comparison of amino acid sequences among the Type I leader peptidases shows the presence of a few highly conserved regions, and site-directed mutagenesis of the E. coli leader peptidase shows particular serine and lysine residues to be involved in the catalytic activity (13–16). Other prokaryotic and mitochondrial leader peptidases are believed to contain such a catalytic dyad, although the homologous subunits of the endoplasmic reticulum signal peptidase contain a histidine residue at the equivalent position to the lysine (reviewed in Ref. 16).

A major hindrance to studies of the plant TPP has been the lack of a cloned gene or cDNA, which has been largely due to the difficulty of purifying the protein (17). We report the isolation and sequence analysis of a cDNA for a TPP precursor from Arabidopsis thaliana. Phylogenetic analysis indicated that the cDNA encodes the chloroplast enzyme. A region encoding the catalytic domain of the enzyme was expressed in E. coli, and catalytic activity of the expressed protein was demonstrated using a precursor of a thylakoid lumen protein. Antibodies to the overexpressed domain cross-reacted with a thylakoid membrane protein, confirming the identification of the cDNA as encoding the chloroplast protein.

EXPERIMENTAL PROCEDURES

Sequencing of the A. thaliana TPP cDNA—A partial cDNA clone 177H22T7 that showed regions of sequence similarity to well conserved blocks of amino acid sequence found among Type I leader peptidases from different organisms was identified from the Arabidopsis Biological Resource Center at Ohio State University (18). The insert of 1081 base pairs was isolated by digestion with Sal1 and Not1 of the pZL1 construct generated by excision in vivo from the Alziplox library vector, and an internal HindIII fragment of 437 base pairs was used as a probe to IPTG, isopropyl-thio-β-D-galactopyranoside; iOE23, stromal intermediate of the 23-kDa extrinsic polypeptide of Photosystem II from wheat; PAGE, polyacrylamide gel electrophoresis.
positively hybridizing phage was isolated and excision instructions for the random prime labeling system (Amersham). A system (Amersham). DNA preparation, digestion, and library screening probe was labeled with fluorescein using a random prime labeling obtained from the Arabidopsis Biological Resource Center (18). The Cambridge, United Kingdom), making the construct pET-TPP558.

resistance cassette added (Carl Webster, Department of Plant Sciences, dependent expression vector pET16b (Novagen) with a kanamycin kan, a vector comprising the IPTG-inducible T7 RNA polymerase- 5 AAAGGATCCCTGGTTTTGCACAATCTTGTGC to amplify the cata-

screen the PRL2A. thaliana var. Columbia cDNA library in Aziplix also obtained from the Arabidopsis Biological Resource Center (18). The probe was labeled with fluorescein using a random prime labeling system (Amersham). DNA preparation, digestion, and library screening were carried out as described in Ref. 19 and by the manufacturer by instructions for the random prime labeling system (Amersham). A positively hybridizing phage was isolated and excision in vivo used to generate a plasmid pBKC1 containing the cDNA insert. Automated sequencing was carried out using an Applied Biosystems/Perkin-Elmer 373 DNA sequencing system.

Phylogenetic Analysis—Amino acid sequences were aligned by CLUSTALW (20) and edited by eye to exclude unreliably aligned residues. Phylogenetic trees were generated using SEQBOOT, PROTPARS, PROTDIST, FITCH, and CONSENSE programs from the PHYLIP3.5 package (21) and SPLITSTREE version 1.0 (22). With PHYLIP tree-

100-fold dilution of this culture was grown at 37 °C until an A600 between 0.4 and 0.5 (measured in a Hitachi U-2000 spectrophotometer) was reached before adding IPTG to a final concentration of 2 mM, followed by further cell growth at 37 °C for 5 h. From a 10-ml culture, cells were pelleted, washed in sterile distilled water, resuspended in 500 μl of 50 mM Tris-HCl, pH 7.0, and then lysed by sonication. To obtain soluble material, the cell extract was centrifuged at 12,000 × g for 20 min. The protein concentration in the supernatant was determined using the Bio-Rad DC protein assay, and 10-μg samples were analyzed by electrophoresis in 15% polyacrylamide gels containing SDS (23, 24). The overexpressed protein was purified on nickel-nitrilotriacetic acid-agarose resin (Qiagen) and eluted with an imidazole step gradient from 0.1 to 0.5M according to the manufacturer’s instructions. Purity of eluted fractions was monitored by SDS-PAGE.

Synthesis of Precursor Protein in Vitro—A cDNA clone for the stro-

Preparation of Chloroplast TPP—Pea (Pisum sativum var. Feltham First) seedlings were grown for 8–10 days in Levington compost (Fisons, Ipswich, United Kingdom) in a greenhouse with supplemen-
living photosynthetic active radiation of 125 μmol photons m⁻² s⁻¹ for 16 h per day. Intact chloroplasts were prepared as described in Ref. 26 and lysed by resuspension in 10 mM Tris-HCl, pH 7.0, at a final chlorophyll concentration of 2 mg/ml. Membranes were pelleted by centrifugation at 12,000 × g for 10 min at 4 °C and the supernatant discarded. The pellet was washed with 50 mM Tris-HCl, pH 7.0, resuspended in 50 mM Tris-HCl, pH 7.0, 0.15% v/v Triton X-100 to a concentration of 2.5 mg of detergent/mg of chlorophyll, incubation for 30 min at 4 °C and centrifuged to a final chlorophyll concentration of 1 mg/ml. Membranes were then resuspended in 50 mM Tris-HCl, pH 7.0, resuspended in 50 mM Tris-HCl, pH 7.0, 0.15% v/v Triton X-100 at a final chlorophyll concentration of 1 mg/ml and incubated on a rotating mixer for 2 h at 4 °C. After centrifugation at 12,000 × g for 20 min at 4 °C, the supernatant was stored at -80 °C and used within 1 week.

Immunological Detection of TPP—Intact pea chloroplasts were prepared as described in Ref. 26, except that a 35% Percoll cushion was used. Chloroplasts were then resuspended in lysis buffer (5 mM MgCl₂, 10 mM Heps-KOH, pH 8.0) followed by centrifugation at 12,000 × g for 10 min at 4 °C. The stroma was separated from the thylakoid pellet, which was washed three times in lysis buffer. A Triton X-100 extract was prepared from the thylakoids by addition of 0.25% v/v Triton X-100 to a concentration of 2.5 mg of detergent/mg of chlorophyll, incubation for 30 min at 4 °C on a rotating mixer, and centrifugation at 30,000 × g for 30 min at 4 °C. Samples corresponding to 20 μg of chlorophyll were analyzed by 15% polyacrylamide gels transferred to Hybond™ nitrocellulose membrane (Amersham) using a Bio-Rad semidry electrophoretic transfer cell. Immunological detection was carried out with antibodies to the overexpressed catalytic domain of TPP raised by UltraClone Ltd. (Wellow, Isle of Wight) at a dilution of 1:3000 and donkey anti-rabbit Ig horseradish peroxidase conjugate and ECL (Amersham) according to the manufacturer’s instructions.

RESULTS

Identification of a Full-length cDNA Encoding the A. thaliana TPP—A cDNA clone, 177H22T7, was identified in the AIMS Arabidopsis EST data base as showing sequence similarity to the leader peptidase of the cyanobacterium P. laminosum. The cDNA clone was obtained from the Arabidopsis Biological Resource Center, and the insert was fully sequenced. The predicted amino acid sequence showed regions that are highly conserved among other Type I signal peptidases. However, the insert did not encode a full-length cDNA, as judged by the lack of a potential initiation codon. The insert was excised from the vector and used to screen an A. thaliana var. Columbia library (PRL2, obtained from the Arabidopsis Biological Resource Center) containing cDNA derived from mixed tissues inserted into ZIPlox (18). One positive clone was identified from 50,000 plaques screened. Excision in vivo generated a plasmid designated pBRC1 with an insert of 1784 base pairs, which was fully sequenced (GenBank™/EBI accession number Y10477). From the sequence, an open reading frame was identified that could encode a protein of 340 amino acids and a molecular mass of 37.9 kDa. The amino acid sequence of the putative A. thaliana TPP was aligned with those of the leader peptidases from E. coli and the cyanobacteria P. laminosum and Synechocystis sp. PCC6803 (Fig. 1). Blocks of conserved amino acid sequence were found, with the A. thaliana sequence showing 34, 38, 39, and 42% identity to the catalytic domains of the E. coli, P. laminosum, and the two Synechocystis sp. PCC6803 leader peptidases, respectively. These regions included the conserved serine and lysine residues believed to function in the catalytic mechanism of the enzyme.

Phylogenetic Analysis—The phylogenetic relationship of the predicted sequence from A. thaliana to other Type I signal peptidase sequences was examined. A distance matrix method indicated a specific grouping of the plant TPP sequence with the cyanobacterial sequences with high (95%) bootstrap significance (Fig. 2). When other tree selection criteria such as parsimony (21) and Split-decomposition (22) were used, congruent results were obtained. Given the proposed origin of chloroplasts from oxygenic photosynthetic prokaroytes, this strongly indicates that the cDNA obtained is for the chloroplast enzyme.

Expression of the Catalytic Domain of the A. thaliana TPP in E. coli—A 558-base pair DNA fragment encoding the putative catalytic domain of the A. thaliana TPP (from residues 177–340) was amplified by polymerase chain reaction and subcloned into the expression vector pET16b-kan as described above, making the construct pET-TPP558. E. coli BL21(DE3)pLysS cells were transformed with this plasmid or pET16b-kan as a control and induced with IPTG as described above. After induction, cells were lysed and the polypeptide profiles compared by SDS-PAGE (Fig. 3). A polypeptide with a mobility corresponding closely to the expected size (~23 kDa) was present in the lysate from cells containing pET-TPP558 (lane 1) but not from those containing pET16b-kan (lane 2), indicating that expression of the domain had occurred.
The catalytic activity of the overexpressed TPP protein from Arabidopsis thaliana was studied using stromal intermediate of Photosystem II from wheat. Radiolabeled protein was incubated with extracts containing 20 μg of total protein prepared from induced E. coli BL21(DE3)pLysS cultures containing pET-TPP558 or pET16b-kan, or a TPP preparation made from pea chloroplasts, and the products analyzed by SDS-PAGE and fluorography (Fig. 4). Incubation with the extract from cells containing pET-TPP558 generated a polypeptide of lower molecular mass (lane 2), which had the same mobility as that generated by incubation with TPP prepared from chloroplasts (lane 4). The extract from control cells containing pET16b-kan did not produce this polypeptide (lane 3). The overexpressed domain was therefore shown to be catalytically active.

**Immunological Localization of the TPP**—Western blots of chloroplast fractions were probed with antibodies to the overexpressed catalytic domain to test the predicted thylakoid location of the TPP and to determine the size of the mature protein. The antibodies detected a protein with a mobility corresponding to a size of approximately 30 kDa in a chloroplast preparation (Fig. 5, lane 1). This protein was absent from the stromal fraction (lane 2), but present in the thylakoid fraction (lane 3). Treatment with Triton X-100 solubilized some, but not all, of the protein (lanes 4 and 5).

**Discussion**

A cDNA encoding a putative A. thaliana TPP was identified and shown to encode highly conserved regions found among other Type I leader peptides. Phylogenetic analysis of the putative TPP indicated a common ancestry with cyanobacterial TPP to the exclusion of those from other bacteria and mitochondria. This is consistent with the putative TPP being a chloroplast enzyme and the belief that important elements of the protein translocation machinery of chloroplast thylakoids are derived from the oxygenic prokaryote that gave rise to chloroplasts by endosymbiosis (27). The amino-terminal domain of the predicted translation product is not conserved in other leader peptides and is presumably an envelope-transfer domain for chloroplast import. It shows characteristic features of such sequences, including a large number of serine and threonine residues, although the cleavage site cannot be deduced from the sequence alone, as such sites have rather relaxed sequence requirements (28, 29). A Kyte-Doolittle hydroathy plot (30) (not shown) identified a putative transmembrane region, which is marked in Fig. 1. This is consistent with the demonstration that TPP is an integral membrane protein (5) and indicates that the chloroplast TPP belongs to the class of enzymes (including those of cyanobacteria) with a single transmembrane span, rather than two as in E. coli. Twenty residues NH2-terminal to the putative transmembrane domain is a stretch (residues 121–139) very rich in charged (mainly acidic) amino acids. The rest of the protein COOH-terminal to the transmembrane domain contains the catalytic site and presumably protrudes into the thylakoid lumen. Serine and lysine residues corresponding to those at positions 90 and 145, respectively, in the E. coli leader peptide sequence are conserved. This strongly indicates that the A. thaliana enzyme uses the serine-lysine catalytic dyad mechanism believed to operate in the enzymes from prokaryotes and mitochondria, but not the endoplasmic reticulum (16). The operation of the serine-lysine mechanism is consistent with our recent report that an inhibitor of the E. coli enzyme also inhibits the cyanobacterial and chloroplast enzymes (9). The -GDNR- motif of uncertain function is also conserved (residues 288–291) in A. thaliana.

The catalytic domain of the enzyme was expressed in E. coli, and shown to retain catalytic activity, by the ability to process the intermediate form of the 23-kDa extrinsic polypeptide of Photosystem II. It is not clear why the catalytic activity of the overexpressed domain was lost. It may be that in vivo the translocation pore serves to hold the substrate in the correct configuration for peptidease action. Antibodies to the overexpressed protein bound to a 30-kDa thylakoid protein, which could be partially solubilized with Triton X-100, as with the bona fide TPP (17), supporting the identification of the cDNA described here. The size of this protein is large enough to include the acidic stretch, which is presumably exposed in the stroma. The function of this region remains to be determined. It may be involved in interactions with other components of the protein translocation machinery.

**Acknowledgment**—We thank Dr. J. C. L. Packer for useful discussions.

**References**

22. Huson, D. H. (1994) *SPLITSTREE Version 1.0 FSP Mathematisierung*, Universität Bielefeld (E-mail: huson@mathematik.uni-bielefeld.de)
23. Bielefeld, Germany
Characterization of a cDNA Encoding the Thylakoidal Processing Peptidase from *Arabidopsis thaliana*: IMPLICATIONS FOR THE ORIGIN AND CATALYTIC MECHANISM OF THE ENZYME
Balbir K. Chaal, Ruth M. Mould, Adrian C. Barbrook, John C. Gray and Christopher J. Howe

doi: 10.1074/jbc.273.2.689

Access the most updated version of this article at http://www.jbc.org/content/273/2/689

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 28 references, 11 of which can be accessed free at http://www.jbc.org/content/273/2/689.full.html#ref-list-1