Transport of Proteins into Chloroplasts

THE THYLAKOIDAL PROCESSING PEPTIDASE IS A SIGNAL-TYPE PEPTIDASE WITH STRINGENT SUBSTRATE REQUIREMENTS AT THE -3 AND -1 POSITIONS*

(Received for publication, December 4, 1990)

Jamie B. Shackleton and Colin Robinson‡
From the Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, United Kingdom

The transport of proteins across the thylakoid membrane in higher plant chloroplasts is usually mediated by an amino-terminal peptide extension which is subsequently removed by a specific thylakoid processing peptidase. We have previously shown that the reaction specificity of this enzyme is very similar to those of signal peptidases located in the endoplasmic reticulum and bacterial plasma membrane. In the present report, the reaction mechanism of the thylakoid peptidase has been investigated by substituting a variety of amino acids for the alanine residues at the -3 and -1 positions of a thylakoid lumen protein precursor. Small neutral side chains are known to be essential at these positions for cleavage by signal peptidases, and we find that these residues likewise play a critical role in defining the thylakoid processing peptidase cleavage site. However, the requirements of the thylakoidal enzyme at these sites are significantly more restrictive than those of the bacterial or endoplasmic reticulum peptidases. Whereas leucine at the -3 position in the substrate is tolerated by the latter two enzymes, cleavage by the thylakoidal peptidase is almost completely inhibited. At the -1 position the presence of alanine appears to be critical; substitution of this residue by glycine, serine, threonine, leucine, lysine, or glutamate leads to either substantial or complete inhibition of cleavage at this site. Substitutions at either -3 or -1 which blocked cleavage at the correct site led to cleavage taking place at an alternative site, probably after the -21 residue.

The biogenesis of a significant proportion of prokaryotic and eukaryotic proteins involves translocation across one or more cellular membranes. In many cases, the transported proteins are initially synthesized with amino-terminal extensions which initiate translocation across the appropriate membrane. The best studied presequences fall into two broad categories. "Signal" peptides, which direct the transport of proteins across the endoplasmic reticulum and bacterial plasma membrane, are usually between 15 and 30 residues in length and are rather hydrophobic. These peptides are removed during or shortly after translocation by membrane-bound processing peptidases (Jackson and Blobel, 1977; Zimmermann et al., 1982; Wolfe et al., 1983; Zimmermann and Mollay, 1986). In contrast, the presequences which direct proteins across the chloroplast envelope or mitochondrial membranes are more hydrophilic, enriched in hydroxylated and basic residues, and are often longer. These sequences are removed by soluble processing peptidases in the stroma or matrix (Robinson and Ellis, 1984; Hawlitschek et al., 1988).

The biogenesis of cytosolically synthesized thylakoid lumen proteins is of particular interest because these precursor proteins contain both types of signal in tandem. Initially precursor proteins are targeted across the chloroplast envelope by a hydrophilic "envelope transfer" domain, after which they are cleaved to intermediate forms by the stromal processing peptidase (SPP). The intermediate forms are then targeted into the thylakoid lumen by a "thylakoid transfer" domain and processed to the mature size by a thylakoidal processing peptidase, TPP (Smeebns et al., 1986; Hageman et al., 1986; James et al., 1989). The thylakoid transfer domains of luminal proteins have features in common with signal sequences, in particular a hydrophobic "core" region and the presence of short-chain residues at the -3 and -1 positions, relative to the terminal processing site (von Heijne et al., 1989). These observations prompted Halpin et al. (1989) to compare TPP with other processing enzymes. It was found that pea TPP and Escherichia coli leader (signal) peptidase (LEP) have identical reaction specificities, although the two enzymes appear to differ structurally. TPP was also found to cleave a eukaryotic signal peptide, suggesting similarities in reaction specificity with signal peptidase in the endoplasmic reticulum.

Various studies with E. coli have shown that the -3 and -1 residues of the signal peptide play an important role in the LEP cleavage mechanism. A statistical comparison of numerous prokaryotic signal peptides showed that small neutral side chains were found almost exclusively at these positions (von Heijne, 1986). Biochemical evidence confirming the importance of these residues has come from a number of reports in which other types of side chains were introduced by mutagenesis (Koshland et al., 1982; Kuhn and Wickner, 1985; Fikes and Bassford, 1986). The most comprehensive of these studies was by Fikes et al. (1990), who generated numerous amino acid substitutions at these positions in the maltose-binding protein precursor to probe the requirements for cleavage by LEP. In the present study, we have carried out a similar type of analysis of the substrate specificity of TPP by site-specific mutagenesis of the -3 and -1 codons of a thylakoid lumen protein precursor. The results indicate that these residues play a critical role in defining the cleavage site, and that TPP has far more stringent requirements at these positions than

* This work was supported by Science and Engineering Research Council Grant GR/P 78309. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 44-0203-525557; Fax: 44-0203-525701.

The abbreviations used are: SPP, stromal processing peptidase; TPP, thylakoidal processing peptidase; 33K, 33-kDa protein of the wheat oxygen-evolving complex; pre-33K, full precursor of 33K; 36K, 36-kDa cleavage product of pre-33K; LEP, Escherichia coli leader peptidase.
corresponding signal peptidases in the E. coli plasma membrane or in the endoplasmic reticulum.

**EXPERIMENTAL PROCEDURES**

**Materials**—Intact chloroplasts were isolated from pea seedlings (*Pisum sativum*, var. Feltham First) as described (Robinson and Barnett, 1988). Radioactive materials were obtained from Amersham International (United Kingdom).

**Generation and Analysis of Pre-33K Mutants**—Oligodeoxynucleotide-directed mutagenesis of wheat pre-33K was as follows. The -3 and -1 codons of a full-length wheat pre-33K cDNA, p33K-2 (Kirwin et al., 1989), were converted to encode 10 different amino acids as outlined in Fig. 1. Complementary 24-base deoxynucleotides containing -1 and -1 codons of a full-length wheat pre-33K cDNA, p33K-2 (Kirwin et al., 1989), were annealed to a 150-base fragment containing the presequence-mature protein junction, which had been ligated into M13mp18. Mutagenesis was carried out using an Amersham site-specific mutagenesis kit according to the manufacturer’s protocol. Plaques containing the required mutations were identified by nucleotide sequencing, after which the 150-base fragment was isolated by KpnI/PstI digestion of minilysis DNA and re-ligated into p33K-2. Wild-type and mutated cDNAs were transcribed using SP6 RNA polymerase, and capped transcripts were translated in a wheat germ cell-free translation system in the presence of [35S]methionine (Kirwin et al., 1989). In vitro processing assays were carried out using partially purified SFP (James et al., 1989) involved incubating 1-µl translation products with 20 µl of partially purified SPP for 60 min at 20 °C. Chloroplast import experiments were carried out as detailed in Robinson and Barnett (1988).

**RESULTS**

The precursor used in this study was that of the 33-kDa protein of the photosynthetic oxygen-evolving complex (pre-33K) from wheat. This protein is located in the thylakoid lumen and synthesized with a bipartite presequence consisting of an amino-terminal envelope transfer domain followed by a thylakoid transfer domain (Tyagi et al., 1987; Kirwin et al., 1989). The -3 and -1 residues of pre-33K (alanine in each case) were replaced by a variety of amino acids as shown in Fig. 1. Oligonucleotide-directed mutagenesis was used to alter the -3 and -1 codons of cDNA encoding pre-33K, and the mutant precursors were expressed by in vitro transcription and translation.

Initial tests were carried out to confirm that the substitutions did not affect the gross properties of the presequence upstream of the terminal processing site. Fig. 2 shows the results of assays in which wild-type and mutant precursors were incubated with partially purified SPP, which cleaves roughly in the middle of the presequence. All of the precursors are processed to the intermediate form with reasonable efficiency. The different precursors are synthesized with varying efficiencies in Fig. 2, but counting of the [35S]label present in the precursor and intermediate forms indicates that each precursor is processed by SFP to a very similar extent. Since SFP is believed to recognize structural features of the presequence (Robinson and Ellis, 1985), this result suggests that none of the mutations has significantly altered the conformation of the middle section of the presequence.

**Effects of Amino Acid Substitutions at the -3 Residue of the 33K Presequence**—Each of the -3 mutants was incubated with intact chloroplasts in a standard import assay to monitor the effects of the amino acid substitution on import and processing. Fig. 3 shows that all of the mutants are targeted into the thylakoid lumen with essentially equal efficiency; in each case, the imported protein is located in the thylakoid fraction and is resistant to added protease, indicating a location within the luminal phase. The mutations do not, therefore, alter the membrane-translocation properties of the precursors in any detectable manner. However, an analysis of the sizes of the imported proteins indicates that the substitutions can be divided into two groups based on their effects on maturation by TPP. -3 Val is the sole member of the Group I mutations, which are defined as having no detectable effect on cleavage by TPP. This precursor is imported and converted to mature size 33K with kinetics that are indistin-

![Fig. 2. Cleavage of the pre-33K mutants by stromal processing peptidase.](image)

Wild-type and mutated pre-33Ks were synthesized by in vitro transcription-translation as described under “Experimental Procedures.” 1-µl aliquots of translation mixture were incubated with partially purified SPP for 60 min at 27 °C. Lanes T and I: wild-type pre-33K before and after incubation with SPP. Lanes 2–11, incubation of SPP with -3 Val, Glu, Lys, Leu, and -1 Gly, Ser, Glu, Lys, Leu, and Thr, respectively. p33K and 33K, precursor and intermediate forms of 33K, respectively.

![Fig. 3. Import of -3 mutants into isolated chloroplasts.](image)

Wild-type pre-33K and four -3 mutants (~3 Val, Lys, Glu, and Leu) were synthesized by transcription-translation; the translation products are shown in lanes T. Each translation product was incubated with isolated pea chloroplasts for 20 min, after which samples were analyzed immediately (lanes 1), after protease treatment of the chloroplasts (lanes 2) and after fractionation of protease-treated chloroplasts into stromal (lanes 3) and thylakoid (lanes 4) samples. Lanes 5, thylakoid samples were treated with protease K before analysis. 33K denotes mobility of authentic mature 33K protein; 36K denotes polypeptide of apparent mass 36 kDa.
guishable from those of wild-type pre-33K (Fig. 3).

The other three -3 substitutions (−3 Lys, Leu, and Glu) have dramatic effects on the maturation of the precursors, and all three fall into the Group II category. In this group, the imported protein is imported and converted to two different polypeptides. A proportion is processed to mature size 33K, and the remainder to a polypeptide of molecular mass 36 kDa; both polypeptides are located in the thylakoid lumen (Fig. 3). The 36-kDa form (36K) migrates on sodium dodecyl sulfate-polyacrylamide gels approximately midway between mature size 33K and the stromal intermediate form which is generated by SPP (data not shown).

Different -3 substitutions give rise to different ratios of 33K and 36K. In the cases of −3 Lys and −3 Glu, virtually equimolar amounts of the two products are generated. However, −3 Leu is processed almost exclusively to 36K (Fig. 3). These results indicate that the substitution at the −3 residue by either lysine or glutamate leads to a substantial inhibition of TPP cleavage at the correct processing site, and that substitution by leucine almost completely blocks processing. In each case, the precursor is cleaved instead at a position upstream of the correct processing site.

Effects of Amino Acid Substitutions at the −1 Residue of the 33K Presequence—The six −1 mutants were analyzed by the same procedures as used for the −3 mutants. As with the −3 mutants, all of the −1 mutants are imported into the thylakoid lumen of isolated chloroplasts, indicating that translocation across the envelope and thylakoid membranes is not affected by any of the substitutions (Fig. 4). However, processing by TPP is dramatically affected in every case, and therefore none of the mutants fall into the Group I wild-type category. Two of the mutants, −1 Gly and −1 Ser fall into the Group II class, in which both 33K and 36K are generated. In the case of −1 Gly, approximately 45% of imported protein is converted to 33K; with −1 Ser, only 14% of imported protein is found as 33K. The other four −1 mutants (−1 Leu, Lys, Gly, and Thr) fall into a third (Group III) category: all of the imported protein is converted to the 36K form, and no 33K can be detected (Fig. 4). Clearly, substitution at the −1 position by any of these residues completely blocks cleavage at this site by TPP.

The data shown in Figs. 3 and 4 are useful for comparing the effects on cleavage of various substitutions at the −3 and −1 positions, and allow different side chains to be ranked to some extent, according to their effects on TPP. However, the data do not necessarily give an accurate indication of the extent to which a given substitution inhibits cleavage by TPP at the correct site, especially in the Group II mutants. This is because all of the data shown in Figs. 3 and 4 were generated from chloroplast import incubations of 20 min, a period which was chosen so that multiple incubations could be carried out with little sampling error. Fig. 5 shows the results of a more kinetic analysis of the maturation of the −1 Gly mutant, which was carried out by taking time point samples during an import incubation and rapidly processing these samples (by shortening the protease treatment and omitting the fractionation procedure). The data clearly show that, at early time points, imported −1 Gly is present almost entirely as the 36K polypeptide, and that this form is gradually converted to 33K during the import incubation. The impression given in Fig. 4 for the −1 Gly mutant thus grossly underestimates the extent to which this substitution inhibits cleavage at the correct site. Rather than cleaving at the correct and alternative (38K) sites at essentially equal rates, it is clear that TPP displays a massive preference for the alternative site. In order to quantitatively compare the effects of the mutation on cleavage by TPP, similar rapid analysis time courses were carried out using each of the Group I and II mutants, and the levels of mature size 33K were determined by scanning the autoradiographs. Fig. 6 shows that at earlier time points, the levels of mature size −3 Val are similar to those of wild-type 33K, but that the rates of maturation of the Group II mutants are drastically reduced. At the earliest time point (5 min) low levels of mature size −3 Ser and −3 Lys are present, but no mature size −3 Gly, −3 Glu, or −3 Leu can be detected. These data serve to reinforce the conclusion that all of the mutations tested in this study, with the exception of −3 Val, inhibit correct maturation by TPP either completely or very substantially.

Significance of the 36K Polypeptide—During import into chloroplasts most of the mutants tested in this work are

![Fig. 4. Import of −1 mutants into isolated chloroplasts. The −1 mutants (−1 Ser, Gly, Lys, Glu, Leu, and Thr) were synthesized by transcription-translation, incubated with chloroplasts, and analyzed as with −3 mutants in Fig. 3. Lanes 1–5 represent total import mixture after incubation, protease-treated chloroplasts, stromal fraction, thylakoid fraction, and protease-treated thylakoids, respectively. Other symbols as described in the legend to Fig. 3.](image)

![Fig. 5. Time course analysis of import of mutant −1 Gly into chloroplasts. The mutant precursor was incubated with chloroplasts for the times given above the lanes (in min). After incubation, the chloroplasts were treated with trypsin (200 μg ml⁻¹) for 10 min at 4 °C, after which samples were analyzed immediately.](image)

![Fig. 6. Time course analysis of the maturation of Group I and Group II mutants. Wild-type and mutant pre-33Ks were incubated with chloroplasts for the times indicated, after which the chloroplasts were trypsin-treated as described in the legend to Fig. 5. The levels of mature size 33K were quantified by laser densitometry (Molecular Dynamics Imagequant v.3.0) of fluorographs of the dried gel. To take into account differences in translation efficiency, corrections were made to give equal values for total imported protein (i.e. 36K plus mature size 33K) for each species of 33K at a given time point. The figure shows corrected values for mature size 33K in each case, after import of wild-type 33K (●), −3 Val (○), −3 Leu (●), −3 Lys (□), −3 Glu (▲), −1 Ser (■), and −1 Glu (●).](image)
converted, at least partially, to the 36K form. In no case do we observe the stromal intermediate size form in the thylakoid lumen indicating that this form is efficiently cleaved to 33K, 36K, or a mixture of both. These findings strongly suggest that the 36K form is generated by TPP and not by a nonspecific protease, because it is unlikely that nonspecific proteolysis could account for such an efficient conversion of the 33K stromal intermediate to a single product. If this is the case, the 33K presequence must contain an alternative TPP cleavage site which is recognized when the correct site is blocked. This possibility was examined further by applying an algorithm developed by von Heijne (1986) to the thylakoid transfer domain of the 33K presequence. This algorithm was developed for predicting processing sites within prokaryotic signal sequences, and, given the similarities in reaction specificity exhibited by TPP and LEP, might also be expected to predict sites of cleavage by TPP. Fig. 7 shows that within the carboxyl-terminal 35 residues of the 33K presequence (the approximate extent of the thylakoid transfer domain) several potential cleavage sites are predicted. The algorithm correctly predicts the bona fide cleavage site, since a high score is given for cleavage after the -1 alanine. However, a high score is also given for cleavage after the -21 residue (also alanine) and this would be consistent with a product size of 36 kDa. These data therefore strengthen the proposal that the 36K form is generated by TPP.

The physiological significance of the 36K form is unclear, since we do not know whether this polypeptide is generated during the import of wild-type 33K into thylakoids. We cannot discount the possibility that this polypeptide represents a genuine processing intermediate, but at present we consider this to be unlikely. Although pre-33K is a poor substrate for TPP in *in vitro* processing assays, the available evidence indicates that the partially purified enzyme processes this, and other lumenal protein precursors, only to the mature sizes and not to intermediate forms (James et al., 1989). The intermediate (36K) cleavage site may therefore represent an example of redundant information within the 36K presequence.

**Localization of 36K**—Although the data shown in Figs. 3 and 4 clearly show that, after import, the 36K polypeptide is located in the thylakoid lumen, we carried out a further experiment to analyze the location of this polypeptide more precisely. In particular, we were interested in finding out whether 36K is correctly assembled into photosystem II, by testing whether 36K and 33K are released from the thylakoids at similar rates during sonication. With this aim in mind, the -1 Gly mutant was imported into chloroplasts, since this mutant is converted to essentially equimolar levels of 36K and 33K (Fig. 8, lane 1). The thylakoids were then sonicated and subsequently pelleted. After 5 s sonication, both 33K and 36K remain almost exclusively bound to the luminal face of the thylakoid membrane (lane 2). The result with 33K is consistent with previous studies on this protein, which have shown that it is fairly tightly bound to the membrane and released only after extended sonication (Kirwin et al., 1988). Accordingly, after a 20-s sonication, a significant proportion of 33K is released from photosystem II and detected in the soluble phase (Fig. 8, lane 5). However, 36K remains associated with the thylakoids, indicating that this species is much more tightly associated with the thylakoid membrane than is 33K. This result suggests that 36K is probably not correctly assembled into photosystem II, and that the amino-terminal extension on this polypeptide may have inserted into the membrane bilayer.

**DISCUSSION**

Previous work on the substrate specificities of bacterial and eukaryotic signal peptides has clearly shown that the -3 and -1 residues of the presequence play a critical role in defining the cleavage site (Fikes et al., 1990; Folz et al., 1988; Kosherl et al., 1982; Kuhn and Wickner, 1985). Given the similarities in reaction specificity between *E. coli* leader peptidase and pea TPP (Halpin et al., 1989) we were interested in determining whether TPP likewise requires the presence of certain types of residue at these positions in substrate presequences. Our data indicate that this is very much the case; the replacement of the -3 or -1 alanine residues by any of a variety of residues has profound effects on the efficiency of cleavage by TPP at the normal site, and causes the enzyme to cleave instead at a site nearer the amino terminus of the presequence.

Perhaps more interestingly, however, our data indicate that, although all three types of peptidase appear to carry out similar reactions, TPP is far more restrictive in its tolerance of residues at the -3 and -1 residues than either *E. coli* leader peptidase or eukaryotic signal peptidase. In a comparison of prokaryotic signal peptides, von Heijne (1986) noted that only alanine, glycine, serine, threonine, leucine, and valine were found at the -3 position, and only alanine, glycine, serine, and threonine at -1. *In vitro* mutagenesis experiments by Fikes et al. (1990) on -3 and -1 sites of the maltose-binding protein precursor generally agreed with the predictions of the statistically based proposals. It was found that the substitution of the -3 alanine by glycine, serine, cysteine, threonine, leucine, or isoleucine had little or no effect on cleavage by leader peptidase; the incorporation of aspartate, arginine,
histidine, or tyrosine at -3 led to very inefficient processing. At the -1 site, the substitution of alanine by glycine, serine, or cysteine had no detectable consequences, whereas the presence of other residues significantly inhibited cleavage at the correct site. Interestingly, in these substitutions, the presence of undesirable residues at the -1 site led to processing taking place at an alternative site (after the -3 residue). Similar types of study on eukaryotic signal peptidase indicate that this enzyme has very similar requirements to E. coli leader peptidase (von Heijne, 1983; Folz et al., 1988). In particular, alanine, serine, and glycine are all acceptable at the -1 position.

Our results contrast strongly with these findings, particularly with respect to the importance of the -1 residue in thylakoid transfer domains. Efficient cleavage by TPP appears to be absolutely dependent on the presence of alanine at this position, and the presence of even glycine or serine (which most closely resemble alanine in structural terms) leads to a very marked inhibition of processing. The presence of bulky side chains, such as those of threonine, leucine, glutamate, or lysine, completely blocks cleavage, and it appears extremely unlikely that any other amino acid at the -1 position would be conducive to processing. These conclusions are consistent with the known sequences of thylakoid transfer domains; relatively few have been published, but all contain alanine at the -1 position (von Heijne et al., 1989).

As with signal peptides, our data indicate that the -3 position of thylakoid transfer domains is less restrictive than the -1 position, since substitution of the -3 alanine by valine has no detectable effects on cleavage by TPP. Given this result, it appears likely that substitutions by other small chain residues for example, serine) would be similarly tolerated by TPP. However, our data clearly indicate that the bulky side chains of leucine, glutamate, and lysine are very undesirable at this position. Again, these findings are consistent with the available sequence data; the -3 residue of thylakoid transfer domains is usually alanine, but this position is occupied by valine in the presequence of the spinach 16 kDa oxygen-evolving protein (Jansen et al., 1987) and by serine in spinach CFo subunit 2 (von Heijne et al., 1989).

Although relatively few -3 mutants were analyzed, our data suggest that the -3 position in thylakoid transfer domains may be more restrictive than the corresponding residue in signal peptides. In the latter peptides, bulky charged side chains and aromatic side chains are unfavorable at the -3 position, but leucine and isoleucine are generally acceptable, although less common than smaller neutral side chains. However, the presence of leucine at this position in the 33K presequence almost completely blocks cleavage, which may indicate that the -3 side chain must remain within a critical length. Our data also suggest that the valine side chain may fit such a critical length criterion, since this residue is acceptable at -3, whereas leucine, the side chain of which is longer by 1 (-CH,-) unit but which is otherwise identical is unacceptable.

On the basis of these data, the reaction specificities of TPP and bacterial leader peptidase may not be completely identical, as has been proposed (Halpin et al., 1989). The two enzymes certainly exhibit many similarities in reaction specificity, and it is likely that the bacterial enzyme will be able to cleave most, if not all, precursors of thylakoid lumen proteins. However, it is entirely possible that TPP may not be able to recognize signal peptides which contain residues other than alanine at the -1 position (the bacterial and eukaryotic signal peptides cleaved by TPP in the study by Halpin et al. (1989) contained alanine at this position in each case).

In conclusion, the -3 and -1 residues of the 33K presequence play a critical role in the TPP cleavage mechanism. At the -3 position, a short chain residue is required, with valine possibly representing the maximum tolerated length of side chain. The branched nature of the valine side chain does not inhibit cleavage by TPP. The importance of the -2 residue was not investigated in this study, but the available evidence suggests that the side chain of this residue is insignificant, since methionine, threonine, aspartate, and leucine occupy this position in various thylakoid transfer domains (von Heijne et al., 1989). The -1 residue, however, is of particular significance since TPP appears to specifically require alanine at this position for efficient cleavage to take place. Finally, it should be emphasized that, since TPP does not cleave after a variety of alanine-X-alanine (where X is any amino acid) combinations within presequences and mature proteins, the reaction mechanism of this enzyme must involve the recognition of other, as yet obscure, signals within target presequences.

REFERENCES