January 2005

Molecular Characterization of a isoenzyme of the targeting peptide degrading protease, PreP2-catalysis, subcellular localization, expression and evolution

S. Bhushan
A. Stahl
S. Nilsson
B. Lefebvre
D. McWilliams

See next page for additional authors

Follow this and additional works at: http://trace.tennessee.edu/utk_biocpubs

Part of the Biochemistry Commons, Cell Biology Commons, and the Molecular Biology Commons

Recommended Citation


This Article is brought to you for free and open access by the Biochemistry, Cellular and Molecular Biology at Trace: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Biochemistry, Cellular and Molecular Biology Publications and Other Works by an authorized administrator of Trace: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.
Authors
S. Bhushan, A. Stahl, S. Nilsson, B. Lefebvre, D. McWilliams, S.J. Wright, M. Seki, D.A. Liberles, K. Shinozaki, Barry D. Bruce, M. Boutry, and E. Glaser
Catalysis, Subcellular Localization, Expression and Evolution of the Targeting Peptides Degrading Protease, AtPreP2

Shashi Bhushan 1, Annelie Ståhl 1, Stefan Nilsson 1, Benoit Lefebvre 2, Motoaki Seki 3, Christian Roth 4, David McWilliam 5, Sarah J. Wright 6, David A. Liberles 4, Kazuo Shinozaki 3, Barry D. Bruce 5, 6, Marc Boutry 2 and Elzbieta Glaser 1,7

1 Department of Biochemistry and Biophysics, Arrhenius Laboratories for Natural Sciences, Stockholm University, 10691 Stockholm, Sweden
2 Unité de Biochimie Physiologique, Institut des sciences de la vie, Université Catholique de Louvain de Louvain, Croix du Sud, 2-20, B-1348 Louvain-la-Neuve, Belgium
3 Plant Functional Genomic Research Group, RIKEN GSC, Japan
4 Computational Biology Unit, BCCS, University of Bergen, 5020 Bergen, Norway
5 Graduate Program in Genome Science and Technology, University of Tennessee, Knoxville, TN 37996, USA
6 Center of Excellence in Structural Biology, Department of Biochemistry, Cellular and Molecular Biology, University of Tennessee, Knoxville, TN 37996 USA

We have previously identified a zinc metalloprotease involved in the degradation of mitochondrial and chloroplast targeting peptides, the presequence protease (PreP). In the Arabidopsis thaliana genomic database, there are two genes that correspond to the protease, the zinc metalloprotease (AAL90904) and the putative zinc metalloprotease (AAG13049). We have named the corresponding proteins AtPreP1 and AtPreP2, respectively. AtPreP1 and AtPreP2 show significant differences in their targeting peptides and the proteins are predicted to be localized in different compartments. AtPreP1 was shown to degrade both mitochondrial and chloroplast targeting peptides and to be dual targeted to both organelles using an ambiguous targeting peptide. Here, we have overexpressed, purified and characterized proteolytic and targeting properties of AtPreP2. AtPreP2 exhibits different proteolytic subsite specificity from AtPreP1 when used for degradation of organellar targeting peptides and their mutants. Interestingly, AtPreP2 precursor protein was also found to be dual targeted to both mitochondria and chloroplasts in a single and dual in vitro import system. Furthermore, targeting peptide of the AtPreP2 dually targeted green fluorescent protein (GFP) to both mitochondria and chloroplasts in tobacco protoplasts and leaves using an in vivo transient expression system. The targeting of both AtPreP1 and AtPreP2 proteases to chloroplasts in A. thaliana in vivo was confirmed via a shotgun mass spectrometric analysis of highly purified chloroplasts. Reverse transcription–polymerase chain reaction (RT–PCR) analysis revealed that AtPreP1 and AtPreP2 are differentially expressed in mature A. thaliana plants. Phylogenetic evidence indicated that AtPreP1 and AtPreP2 are recent gene duplicates that may have diverged through subfunctionalization.

Keywords: Chloroplasts — Dual targeting — Mitochondria — Presequence protease — Protein import — Zinc metalloprotease.

Abbreviations: ESI, electrospray ionization; EST, expressed sequence tag; GB, grinding buffer; GFP, green fluorescent protein; GST, glutathione-S-transferase; HPLC, high-performance liquid chromatography; IP, import buffer; MPP, mitochondrial processing peptidase; MS, mass spectrometry; PK, proteinase K; PreP, protease protease; RT–PCR, reverse transcription–polymerase chain reaction; SPP, stromal processing peptidease.

Introduction

The majority of mitochondrial and chloroplastic proteins are encoded in the nucleus and synthesized on polyribosomes in the cytosol as precursor proteins carrying a cleavable N-terminal extension known as a signal or targeting peptide; in some cases, targeting information is stored in the mature part of the protein (Pfanner and Geissler 2001). The mitochondrial and chloroplastic targeting peptides display similar physiochemical properties; they have a high content of hydrophobic, basic and hydroxylated amino acid residues and a very low content of acidic amino acids. Bioinformatic analysis of mitochondrial and chloroplastic targeting peptides revealed structural differences as mitochondrial presequences are predicted to fold into a positively charged amphiphilic α-helix (von Heijne 1986), whereas chloroplastic transit peptides display a random coil structure (Schmidt et al. 1979). Nuclear magnetic resonance (NMR) studies confirmed that the mitochondrial presequences adopt an amphiphilic α-helix structure (Abe et al. 2000, Moberg et al. 2004). However, a helical structure for chloroplast transit peptides in membrane mimetic environments has also been reported (Bruce 2000). Protein import into mitochondria and chloroplasts is believed to be highly specific in vivo, despite the fact that mistargeting of mitochondrial proteins to chloroplasts has been observed in vitro (Whelan et al. 1990, Cleary et al. 2002). Recent advances in the yeast mitochondrial protein import indicate the presence of a sorting...
mechanism of proteins at the level of mRNA (Ginsberg et al. 2003); however, no such sorting mechanism has been identified so far in plants. After import of precursor proteins to the mitochondrial matrix, mitochondrial presequences are proteolytically cleaved off by the mitochondrial processing peptidase (MPP). In plants, MPP is integrated into the cytochrome bc1 complex of the respiratory chain, whereas in mammals and yeast, MPP is a soluble matrix protease (Glaser et al. 1998, Pfanner and Geissler 2001). The proteolytic processing of precursor proteins results in the production of mature proteins that fold into their native functional conformation and free presequences. After import in chloroplast stroma, precursor proteins are processed by a proteolytic action of the soluble stromal processing peptidase (SPP) also resulting in the production of mature proteins (Richter and Lamppa 1998, Richter and Lamppa 2002) and free transit peptides.

Targeting peptides are potentially harmful for the structure and function of mitochondria and chloroplasts. They can perturb natural and artificial lipid bilayers. Addition of presequences to mitochondria results in membrane lysis, uncoupling of respiration and dissipation of the membrane potential (Roise et al. 1986, Glaser and Cumsky 1990a, Glaser and Cumsky 1990b, Nicolay et al. 1994). The mechanism of action of presequences on the mitochondrial membrane is not clear, but it has been proposed that the presequence peptides induce channel opening (Lu and Beavis 1997) or that the peptides themselves form a pore (Matsuzaki et al. 1996). Furthermore, mitochondrial presequences have been shown to possess antimicrobial activity (Hugosson et al. 1994). Therefore, free targeting peptides generated inside the mitochondria and chloroplasts have to be rapidly removed, e.g. by proteolytic degradation. Both ATP-dependent and ATP-independent proteases have been reported in mitochondria and chloroplasts (Dyck et al. 1994, Sagarra et al. 1999, Adam et al. 2001, Halperin et al. 2001, Arnold and Langer 2002). ATP-dependent proteases are usually involved in degradation of misfolded proteins and in maintaining the stoichiometric amounts of the protein complexes, while ATP-independent proteases are responsible for the degradation of short unfolded polypeptides.

We have shown rapid proteolytic degradation of the mitochondrial presequences after import into the mitochondria by an ATP-independent protease (Stähl et al. 2000). A protease, presequence protease (PreP), responsible for this degradation, was isolated from potato tuber mitochondrial matrix and identified by mass spectrometric (MS) analysis, electrospray ionization (ESI) and tandem MS (MS/MS). The peptide sequence obtained after MS matched two proteins in the Arabidopsis thaliana genomic database, zinc metalloprotease (AAL90904) and a putative zinc metalloprotease (AAG13049) (Stähl et al. 2002) that we have named AtPreP1 (previously referred to as AtZnMP, Stähl et al. 2002) and AtPreP2, respectively. Both proteases display high amino acid sequence similarity, with
most differences occurring in their predicted organelar targeting peptides. Both AtPreP1 and AtPreP2 harbor a characteristic inverted zinc-binding motif, HILEHX₁₁E, and are classified to the pitrilysin protease subfamily A. We have shown that the PreP1 functions as a signal peptide-degrading protease in both mitochondria and chloroplasts and is dual targeted to both organelles by an ambiguous targeting peptide (Bhushan et al. 2003). Here we have overexpressed and purified AtPreP2 and investigated its proteolytic properties against targeting peptides and their mutants. Furthermore, we have studied targeting properties of AtPreP2 in vitro as well as in vivo using green fluorescent protein (GFP) fusion constructs. A random shotgun approach followed by MS analysis was used to study the occurrence of AtPreP1 and AtPreP2 in chloroplasts, and reverse transcription–polymerase chain reaction (RT–PCR) analysis to study tissue-dependent expression of the proteins. Phylogenetic methods were used to characterize the timing of the duplication event and subsequent sequence divergence.

Results

Overexpression, purification and proteolytic activity of the recombinant AtPreP2

The AtPreP2 cDNA encodes a protein of 1080 amino acids (AAG13049). Single organelle intracellular prediction programs, MitoProt (http://mips.gsf.de/cgi-bin/proj/medgen/mitofilter) and ChloroP (http://www.cbs.dtu.dk/services/ChloroP/), predicted that the targeting peptide of AtPreP2 is 85 amino acid residues long and the mature protein 995 residues long. The mature portion of AtPreP2 was cloned as a fusion protein with glutathione-S-transferase (GST) and the fusion protein (GST–PreP2) was overexpressed in *Escherichia coli* (Fig. 1a). The fusion protein was purified on a GSTrap FF column, and the AtPreP2 was eluted after cleavage of the fusion protein with PreScission protease (Fig. 1a, lane 3). The recombinant AtPreP2 protein (99% purity) had an estimated molecular mass of 110 kDa. The proteolytic action of the recombinant AtPreP2 was tested against a mitochondrial targeting peptide derived from the ATP synthase F<sub>1</sub>β subunit from *Nicotiana plumbaginifolia* (AtpF<sub>β</sub>(2–54)), and a chloroplastic targeting peptide derived from the small subunit of RuBP carboxylase/oxygenase (Rubisco) from *Nicotiana tabacum*, SStpNt. Incubation of both targeting peptides with AtPreP2 resulted in complete degradation of the targeting peptides (Fig. 1b, lanes 2 and 5). Addition of α-phenanthroline, a specific inhibitor of the metalloprotease, completely abolished the proteolytic activity of AtPreP2 (Fig. 1b, lanes 3 and 6). These results show that AtPreP2 has a dual proteolytic function against both mitochondrial and chloroplastic targeting peptides.

Differences in cleavage specificity of AtPreP1 and AtPreP2

The cleavage specificity of AtPreP1 and AtPreP2 was studied by investigating cleavage of a specific fluorescent peptide P<sub>1</sub>, the mitochondrial presequence peptide N<sub>5.7</sub>pF<sub>β</sub>(2–54)–hsl peptide and analysed on Tris-tricine gels as described in Materials and Methods. (b) Generation of proteolytic fragments of the F<sub>1</sub>β presequence by AtPreP1 and AtPreP2. AtPreP1 and AtPreP2 were incubated with the N5.7pF<sub>β</sub>(2–54)–hsl peptide and analysed on Tris-tricine gels as described in Materials and Methods.

![Fig. 2 Differences in cleavage specificity of AtPreP1 and AtPreP2. Degradation of peptides by AtPreP1 and AtPreP2. (a) The fluorescent peptide P<sub>1</sub> incubated with AtPreP1 or AtPreP2. The degradation assays were performed as described in Materials and Methods. (b) Generation of proteolytic fragments of the F<sub>1</sub>β presequence by AtPreP1 and AtPreP2. AtPreP1 and AtPreP2 were incubated with the N5.7pF<sub>β</sub>(2–54)–hsl peptide and analysed on Tris-tricine gels as described in Materials and Methods.](image-url)
Alternative processing has been detected with these mutants (S. J. Wright and B. D. Bruce, unpublished). *At*PreP2 had the capacity to degrade both SStpNt and all the mutants, whereas *At*PreP1 could not degrade the SStpNtP36A mutant with decreased flexibility as proline has been changed to alanine (Fig. 3b). These results additionally confirm the different cleavage specificity of *At*PreP1 and *At*PreP2.

In vitro single and dual import of *At*PreP2 into mitochondria and chloroplasts

Most sequence differences between *At*PreP2 and *At*PreP1 were found in their targeting peptides. Whereas *At*PreP1 was predicted to be a mitochondrial protein by both TargetP (http://www.cbs.dtu.dk/services/TargetP/) and Predotar (http://www.inra.fr/predotar/), programs designed to predict intracellular organelar targeting ability, *At*PreP2 was suggested to be a chloroplast protein. In order to study the subcellular localization of *At*PreP2, in vitro import of the *At*PreP2 precursor was studied into isolated chloroplasts and mitochondria. Incubation of the *At*PreP2 precursor protein with isolated chloroplasts resulted in import and processing of the precursor form of the protein to the mature sized product (Fig. 4a, lane 2). The mature form was resistant to thermolysin treatment, evidence that *At*PreP2 was fully translocated and processed in the chloroplasts (Fig. 4a, lane 3). The additional protein band below *At*PreP2 (in the precursor lane) is clearly different from that of the mature sized protein produced after import and processing, and is not imported nor bound to mitochondria or chloroplasts as it disappears after pre-incubation and washing of organelles. Incubation of the *At*PreP2 precursor protein with isolated mitochondria also resulted in import and processing of the precursor to the mature size product (Fig. 4b, lane 2). The imported precursor and mature form of the *At*PreP2 protein were resistant to proteinase K (PK) treatment, showing import of *At*PreP2 into mitochondria (Fig. 4b, lane 3). Furthermore, import of the *At*PreP2 precursor into mitochondria was dependent on membrane potential, as neither the PK-resistant form of the precursor nor the mature protein was detected in the presence of an ionophore, valinomycin (Fig. 4b, lanes 4 and 5). These results indicate that *At*PreP2 might be a dually targeted protein to both chloroplasts and mitochondria.

Import of the *At*PreP2 precursor was also tested using the dual in vitro import system in which the precursor is simultaneously incubated with both isolated organelles (Rudhe et al. 2002). Under these conditions *At*PreP2 was imported and processed in both mitochondria and chloroplasts (Fig. 4c, lanes 2 and 4). The mature form of *At*PreP2 generated after processing of the precursor was resistant in both organelles to externally

---

**Table 1**

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Peptide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SStpNt</td>
<td>MASVLSDAVAATSVAGQNNYAPPTQKGKAKKFGVQ30QG3</td>
</tr>
<tr>
<td>SStpNcP37Y</td>
<td>MASVLSDAVAATSVAGQNNYAPPTQKGKAKKFGVQ30QG3</td>
</tr>
<tr>
<td>SStpNcP36A</td>
<td>MASVLSDAVAATSVAGQNNYAPPTQKGKAKKFGVQ30QG3</td>
</tr>
<tr>
<td>SStpNcP28A</td>
<td>MASVLSDAVAATSVAGQNNYAPPTQKGKAKKFGVQ30QG3</td>
</tr>
</tbody>
</table>

---

**Fig. 3** Degradation of SStpNt and its mutants by *At*PreP1 and *At*PreP2. (a) Amino acid sequences of wild-type SStpNt and its mutants. Mutations in SStpNt are underlined. (b) Degradation of SStpNt peptide and its mutants by *At*PreP1 and *At*PreP2. The degradation assays were performed as described in Materials and Methods. *At*PreP1 and *At*PreP2 were incubated with the indicated peptide and analyzed on Tris-tricine gels as described in Materials and Methods. The lower band in the peptide represents a degradation product as marked by an asterisk.
Dual targeting and function of the *AtPreP2*

In vitro import of *AtPreP2* into mitochondria and chloroplasts.

(a) Chloroplastic in vitro import of *AtPreP2*. Thermolysin (5 µg µl⁻¹) was added after import where indicated. (b) Mitochondrial in vitro import of *AtPreP2*. Proteinase K (10 µg/µl) was added after import where indicated. Valinomycin (1 µM) was added prior to import where indicated. (c) Simultaneous dual in vitro import of *AtPreP2* in mitochondria and chloroplasts. *AtPreP2* was incubated with the isolated mitochondria and chloroplasts in the same reaction mixture. Mitochondria and chloroplasts were reisolated on a 4% Percoll gradient after import as described in Materials and Methods. Thermolysin (5 µg µl⁻¹) was added after reisolation of the mitochondria and chloroplasts as indicated. P, precursor; m, mature.

### Table 1

<table>
<thead>
<tr>
<th>Lane</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AtPreP2</strong></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>mitochondria</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>chloroplasts</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>thermolysine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>kDa</td>
<td>220</td>
<td>97</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 4** In vitro import of *AtPreP2* into mitochondria and chloroplasts. (a) Chloroplastic in vitro import of *AtPreP2*. Thermolysin (5 µg µl⁻¹) was added after import where indicated. (b) Mitochondrial in vitro import of *AtPreP2*. Proteinase K (10 µg/µl) was added after import where indicated. Valinomycin (1 µM) was added prior to import where indicated. (c) Simultaneous dual in vitro import of *AtPreP2* in mitochondria and chloroplasts. *AtPreP2* was incubated with the isolated mitochondria and chloroplasts in the same reaction mixture. Mitochondria and chloroplasts were reisolated on a 4% Percoll gradient after import as described in Materials and Methods. Thermolysin (5 µg µl⁻¹) was added after reisolation of the mitochondria and chloroplasts as indicated. P, precursor; m, mature.

**In vivo dual import of the *AtPreP2*–GFP fusion protein into tobacco protoplasts**

Dual import of the *AtPreP2* into mitochondria and chloroplasts was investigated further in vivo using GFP fusion and transient expression in tobacco protoplasts and leaves. The targeting peptide and 70 amino acid residues from the mature part of the *AtPreP2* protein were fused to GFP (Fig. 5B) under the strong plant transcription promoter EN50PMA4 (Zhao et al. 1999). Seventy amino acids from the mature part of the *AtPreP2* were fused in order to preserve the native processing site. A known mitochondrial targeting peptide of the ATP synthase F₁β-subunit from *N. plumaginifolia* fused to GFP (F₁β–GFP) was used as a control for mitochondrial targeting (Duby et al. 2001). Transient expression of the *AtPreP2*–GFP fusion protein was performed in tobacco protoplasts and targeting was analyzed by confocal microscopy. Protoplasts transformed with the F₁β–GFP construct targeted GFP to small, punctuated shape structures (Fig. 5C, a). These small punctuated structures were also labeled by the red fluorescence of Mitotracker (Fig. 5C, b). Co-localization of GFP with Mitotracker is shown in the yellow merged image in Fig. 5C, c. GFP alone was found to be present all over in the cytosol (Fig. 5C, f–j). When protoplasts were transformed with the *AtPreP2*–GFP construct, fluorescence was found to be localized in two different locations, the punctuated structures and large, round structures (Fig. 5C, k). Fluorescence in the punctuated structures co-localized with Mitotracker (Fig. 4C, l and m), while fluorescence in the large, round structures co-localized with chloroplast autofluorescence (Fig. 5C, n and o). These results showed that the targeting peptide of *AtPreP2* can target GFP to both mitochondria and chloroplasts in vivo and support the dual localization of *AtPreP2*.

**Mass spectrometric identification of *AtPreP1* and *AtPreP2* in *Arabidopsis* chloroplasts**

We have shown earlier that the PreP is present in both mitochondria and chloroplasts in different plant species using antibodies raised against an 18 amino acid residue peptide in the C-terminal portion of the enzyme (Moberg et al. 2003). However, these antibodies do not distinguish between *AtPreP1* and *AtPreP2* as the antigenic peptide is identical in both enzymes. We also tested cross-reactivity of the *AtPreP1* antibodies raised against the full-length *AtPreP1* and found that these antibodies also recognized both *AtPreP1* and *AtPreP2* (data not shown).

A random, "bottom-up" analysis of isolated intact *A. thaliana* chloroplasts was carried out in order to identify the proteases. One peptide corresponding to *AtPreP1* was repeatedly found with the highest score; however, several other unique peptides with lower confidence scores were also identified for both the *AtPreP1* and *AtPreP2* proteins (Table 1). These results suggest that under these growth conditions, both *AtPreP1* and *AtPreP2* are present in the *Arabidopsis* chloroplasts isolated under the highly stringent conditions. Unfortunately, this methodology cannot be used to give quantitative information, although abundant proteins, such as Rubisco, certainly provide more abundant trypsin-derived peptides (data not shown).
Dual targeting and function of the AtPreP2

Organ-specific differential expression of AtPreP1 and AtPreP2 transcripts

Expression of the AtPreP1 and AtPreP2 transcripts was studied in 3-week-old seedlings and in different organs in mature A. thaliana plants by a semi-quantitative RT–PCR method. PCR conditions were carefully optimized for the quantitative measurements of the transcripts level. Both the AtPreP1 and AtPreP2 transcripts were detected in the 3-week-old seedlings employing 30 cycles of PCR (Fig. 6a). For the quantitative measurements of the AtPreP1 and AtPreP2 transcripts, it was found that 125 ng of total RNA with 25 cycles of PCR yielded quantitative measurements of the transcripts. Constitutively expressed actin was used as an internal control (Fig. 6b, lower panel). The AtPreP1 transcript was detected only in sili- que and flower although the transcript level was much higher in the inflorescence (Fig. 6b, middle panel). In contrast to AtPreP1, the AtPreP2 transcript was detected in leaf, flower and root with no transcript detected in silique and shoot (Fig. 6b, upper panel). These results suggest that AtPreP1 and AtPreP2 are expressed in a differential, organ-specific manner.

Evolutionary analysis

A gene family consisting of close and long distance homologs of AtPreP1 and AtPreP2 was built by combining information from three sources. The gene families from The Adaptive Evolution Database (TAED) (Roth et al. 2005) were supplemented with longer distance BLAST hits from GenBank and near-full length homologs that could be manually assem-
bled from a BLAST search against the NCBI expressed sequence tag (EST) collection. The sequences in this gene family were used to calculate a multiple sequence alignment using POA (Lee et al. 2002) and to build a phylogenetic tree using MRBAYES (Huelsenbeck and Ronquist 2001) (Fig. 7). This tree shows AtPreP1 and AtPreP2 to be close relatives and suggests the importance of searching for close homologs of these proteins in other Brassicaceae species.

Discussion

In the present study, we have investigated proteolytic function, intracellular localization, expression and evolution of AtPreP2, a homolog of the targeting peptide degrading AtPreP1. The purified recombinant AtPreP2 retained the proteolytic activity against the mitochondrial and chloroplastic targeting peptides, showing that AtPreP1 and AtPreP2 in Arabidopsis thaliana are paralogs with the same cellular function. There is a single AtPreP homolog in yeast, Ydr430cp and human, hMP1. The Ydr430cp gene is non-essential for yeast survival. Arabidopsis thaliana T-DNA insertion knock-out mutants of AtPreP1 and AtPreP2 are also available (University of Wisconsin Biotechnology Center USA, Nottingham Arabidopsis Stock Center, UK or Salk Institute Genomic Analysis Laboratory, USA). No Arabidopsis thaliana phenotype has been observed, indicating that when one of the proteases is absent, the isoenzyme metalloprotease or another protease takes over degradation of the targeting peptides. It will be interesting to investigate this possibility in the future by generating a double knockout mutant of AtPreP1 and AtPreP2. Comparison of the AtPreP2 and AtPreP1 proteolytic activities shows that the proteases have different proteolytic subsite preference. Whereas incubation of the substrate P1 and N_iF_iβ(2–54) peptides with AtPreP2 resulted in generation of an easily detected intermediate product (see Fig. 2), no such intermediate was observed upon degradation with AtPreP1. Also, degradation of the chloroplastic transit peptide SStpNt and its mutants has confirmed different cleavage specificity of AtPreP1 and AtPreP2. The SStpNtP36A mutant has decreased flexibility in comparison with the wild-type SStpNt transit peptide where proline has been changed to alanine; SStpNtP36A could not be degraded by AtPreP1, whereas it was completely degraded by AtPreP2. Thorough substrate specificity studies of the two proteases using MS analysis of degradation products, mitochondrial presequence mutant peptides as well as a number of other synthetic peptides show differences in amino acid recognition and cleavage efficiency (Ståhl et al. 2005). These studies indicate that AtPreP1 and AtPreP2 may have overlapping but complementary proteolytic specificity allowing great variety of targeting peptides being rapidly degraded.

As both TargetP and Predotar clearly suggest that AtPreP2 is localized to chloroplasts, it was of interest to investigate organellar targeting of AtPreP2. We have investigated targeting properties of AtPreP2 both in vitro and in vivo. In vitro import studies of the full-length AtPreP2 precursor in both a single and a dual import system (in the presence of a competing organelle, Rudhe et al. 2002) showed that AtPreP2 could be targeted to both mitochondria and chloroplasts, giving a strong indication of the dual localization of AtPreP2. There are a few reports indicating in vitro mistargeting of chloroplastic proteins to mitochondria, but no reports showing mistargeting of a mitochondrial precursor protein into chloroplasts are available. As
Dual targeting and function of the *AtPreP2*

peptides with respect to the amino acid content but is predicted to be unstructured (Jpred; http://www.compbio.dundee.ac.uk/~www-jpred/submit.html), which is a typical feature of chloroplastic targeting peptides. So far there are 25 proteins, each of them encoded by a single gene, that are reported to be dually targeted to both mitochondria and chloroplasts. These proteins are called dual-targeted proteins (Peeters and Small 2001, Silva-Filho 2003). The majority of these proteins are involved in gene expression, e.g. aminoacyl-tRNA synthetases, RNA polymerase, methionine aminopeptidases and a peptidyl deformylase. Other dual-targeted enzymes are related to protection against oxidative stress, e.g. glutathione reductase (Rudhe et al. 2002) and ascorbate peroxidase (Chew et al. 2003). *AtPreP1* (Bhushan et al. 2003, Moberg et al. 2003) and *AtPreP2* involved in cellular protein turnover are new representatives of this group.

Intact chloroplasts have been analysed by ‘bottom-up’ or shotgun proteomics in order to identify the proteases (see Table 1). In this technique, isolated chloroplasts were treated with trypsin to produce peptides which are separated by high-performance liquid chromatography (HPLC) and then fed on-line to the mass spectrometer where they are ionized and separated by electric fields. The peptide ions are resolved by their mass to charge ratio and subjected a collision-induced dissociation (CID) MS/MS stage. A computer algorithm (Mascot, Matrix Science, Ltd; Perkins et al. 1999) was used to compare the experimental spectrum with a theoretical spectrum calculated from an in silico tryptic digest of the proteome. Proteins that have a protein score >30 can be considered as certain; however, the presence of several fragments of a protein with a lower score gives a significant indication of its presence. Four separate, double chloroplast isolation and proteomic analysis experiments were done. In three separate experiments, a single peptide was repeatedly detected that was unique to *AtPreP1*. Several other peptides as indicated in Table 1 for both *AtPreP1* and *AtPreP2* were also identified, indicating the presence of both *AtPreP1* and *AtPreP2* in chloroplasts. The chloroplast preparations used were very pure. Only several cytosolic ribosomal proteins were identified, but no proteins with mitochondrial, endoplasmic reticulum (ER), Golgi apparatus, peroxisomal or nuclear annotation were detected.

Expression analysis of the *A. thaliana* mitochondrial protein import apparatus TOM and TIM isoforms in various organs showed that although they were present in small multigene families, only one member was prominently expressed (Lister et al. 2004). As the *A. thaliana* genome harbors both the *AtPreP1* and the *AtPreP2* genes and both the proteases are dually targeted to both mitochondria and chloroplasts, we next studied whether both of these proteases are expressed in *A. thaliana* and if so are they expressed constitutively or in an organ-specific manner. Expression of the *AtPreP1* and *AtPreP2* transcripts was studied using semi-quantitative RT–PCR under carefully optimized conditions for the quantitative measurements of the transcripts. Both the *AtPreP1* and *AtPreP2* tran-
scripts were detected in young seedlings; however, in varying amounts. The AtPreP1 transcript was detected to be present in silique and in flower although the transcript level was much higher in flower. In contrast to the AtPreP1 transcript, the AtPreP2 transcript was found to be present in leaf, flower and in root, with no transcript detected in shoot and silique. These results showed that both AtPreP1 and AtPreP2 are expressed in an organ-specific manner in A. thaliana plants. It will be interesting to investigate the functional importance of higher transcript levels of AtPreP1 present during flower development.

AtPreP1 and AtPreP2 are classified as belonging to the pitrilysin protease subfamily A. This subfamily contains oligopeptidases of 100 kDa, such as the insulin-degrading enzyme (IDE) and the bacterial homolog protease III (Rawling and Barrett 1991) that degrade small peptides in a metal-dependent manner (Duckworth et al. 1998). In contrast, the MPP and the SPP belong to subfamily B and they mediate single proteolytic cleavage of precursor proteins. Systematic analysis of gene duplication in both A. thaliana (Moore and Purugganan 2003) and in Caenorhanditis elegans (Katju and Lynch 2003) has indicated that recent gene duplications are subject to selective pressures different from those of genes that have not been duplicated. For duplicates where both copies have been retained, both neofunctionalization (positive selection) and subfunctionalization (a largely neutral process) play roles in divergently shaping gene function. It appears that AtPreP1 and AtPreP2 have subfunctionalized, and analysis of DNA sequences from Arabidopsis and other Brassicaceae will enable testing for an entirely neutral process. Other examples of subfunctionalization through divergence of expression patterns have been observed (Force et al. 1999). It is clear that truly redundant gene functions are transitory in evolutionary history. AtPreP1 and AtPreP2 are recent duplicates in the process of diverging in function, an ongoing process that generates evolutionary novelty in genomes.

### Materials and Methods

#### Cloning of AtPreP2

Full-length cDNA of AtPreP2 (RAFL09993-D24) was originally obtained from RIKEN Genomic Sciences, Japan (Seki et al. 2002). The region of the AtPreP2 precursor encoding the mature portion of the protein was amplified using Pfu DNA polymerase (Stratagene, La Jolla, CA, USA) that degrade small peptides in a metal-dependent manner (Duckworth et al. 1998). In contrast, the MPP and the SPP belong to subfamily B and they mediate single proteolytic cleavage of precursor proteins. Systematic analysis of gene duplication in both A. thaliana (Moore and Purugganan 2003) and in Caenorhanditis elegans (Katju and Lynch 2003) has indicated that recent gene duplications are subject to selective pressures different from those of genes that have not been duplicated. For duplicates where both copies have been retained, both neofunctionalization (positive selection) and subfunctionalization (a largely neutral process) play roles in divergently shaping gene function. It appears that AtPreP1 and AtPreP2 have subfunctionalized, and analysis of DNA sequences from Arabidopsis and other Brassicaceae will enable testing for an entirely neutral process. Other examples of subfunctionalization through divergence of expression patterns have been observed (Force et al. 1999). It is clear that truly redundant gene functions are transitory in evolutionary history. AtPreP1 and AtPreP2 are recent duplicates in the process of diverging in function, an ongoing process that generates evolutionary novelty in genomes.

### Table 1

<table>
<thead>
<tr>
<th>Peptides identified for AtPreP1</th>
<th>Stringency score</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADDLFNLNMLCLQEVQTCDQR</td>
<td>41.7</td>
</tr>
<tr>
<td>SLTNVEKSVAFLDLPPENPSGGLVTWDGR</td>
<td>14.9</td>
</tr>
<tr>
<td>GVSEENVQVEEIMDLTK</td>
<td>14.3</td>
</tr>
<tr>
<td>GVSEENVQVEEIMDLTK</td>
<td>13.9</td>
</tr>
<tr>
<td>DEPCSK</td>
<td>13.0</td>
</tr>
<tr>
<td>NGCIVNMTADGKSLTNVEK</td>
<td>11.9</td>
</tr>
<tr>
<td>DKGTVAVASAEDAAN</td>
<td>10.8</td>
</tr>
<tr>
<td>KKHMCWNLLSEKLPLDQTQLALTFLDMLGTPASPLR</td>
<td>10.8</td>
</tr>
<tr>
<td>DEPCSK</td>
<td>10.0</td>
</tr>
</tbody>
</table>

Tandem mass spectroscopy of the total chloroplast protein tryptic digest identified 20 peptides corresponding to both AtPreP1 and AtPreP2 using Mascot as described in Materials and Methods. For the stringency score, the higher the score, the better and more confident the assignment.

### Overexpression and purification of the recombinant AtPreP2

The E. coli overexpression strain, BL21 (DE3), was transformed with the pGEX-6P-2 vector containing the mature portion of AtPreP2 fused to GST and grown at 30°C in LB-medium. After 4 h, 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added to the culture and the incubation was continued for another 6 h. Cells were resuspended in PBS buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄,
1.8 mM KH$_2$PO$_4$, pH 7.3) and lysed by addition of 0.5 mg ml$^{-1}$ lysosome and 10 µg ml$^{-1}$ DNase I followed by sonication for 3×30 s. The lysate was centrifuged for 20 min at 15,000 × g and filtered through a 0.2 µm membrane. The supernatant was loaded onto a GSTrap™ FF 1 ml column (Amersham Biosciences) equilibrated with PBS buffer. AtPreP2 was eluted after on-column cleavage with PreScission™ Protease according to the manufacturer’s instructions (Amersham Biosciences). The eluted AtPreP2 was applied to a Superdex™ 200 10/30 column (Amersham Biosciences) equilibrated with 20 mM HEPES-KOH, 10 mM MgCl$_2$ (pH 8.0). The protein content of eluted fractions from GSTrap FF and Superdex 200 HR10/30 was analyzed by 12% SDS–PAGE in the presence of 4 M urea (Laemmli 1970) and stained with silver.

**Proteolytic activity of the recombinant AtPreP2**

The proteolytic activity of the recombinant mature AtPreP2 was investigated against the mitochondrial presequence derived from the F$_{1}$β-subunit of the *N. plumaginifolia* ATP synthase, N$_{26}$βF$_{2}$β(2–54)-hsl (Ståhl et al. 2002), and a chloroplastic transit peptide derived from the small subunit of Rubisco from *N. tabacum*, StspNt (Moberg et al. 2003). The proteolytic reaction contained 1.0 µg of the recombinant mature AtPreP2 and 1 µg of either N$_{26}$βF$_{2}$β(2–54)-hsl or StspNt in the reaction buffer containing 20 mM HEPES-KOH (pH 8.0) and 10 mM MnCl$_2$. Degradation was performed for 50 min at 30°C and the reaction was stopped by the addition of Laemmli sample buffer: o-Phenanthroline (10 mM) was added and pre-incubated with the reaction mixture for inhibition studies. For the detection of N$_{26}$βF$_{2}$β(2–54)-hsl, samples were analyzed on 10–20% Tris-tricine gels (Bio-Rad, Sweden) and stained with Coomassie brilliant blue. For the detection of StspNt, samples were subjected to 12–20% SDS–PAGE in the presence of 4 M urea and immunological cross-reactivity was analyzed by Western blotting with an antibody raised against StspNt followed by detection with horseradish peroxidase (HRP)-conjugated secondary antibodies. Point mutations in wild-type StspNt were generated by Quick change site-directed mutagenesis (Stratagene), and overexpression and purification of mutants were carried out as described earlier by Moberg et al. (2003). For the detection of StspNt and its mutant in Fig. 3, samples were analyzed on 10–20% Tris-tricine gels (Bio-Rad) and stained with Coomassie brilliant blue.

In order to identify intermediate proteolytic fragments of the F$_{1}$β presequences generated by AtPreP1 and AtPreP2, 2 µg of the N$_{26}$βF$_{2}$β(2–54)-hsl peptide was incubated with 0.3 µg of the respective protease in degradation buffer for 30 min at 30°C. Recombinant AtPreP1 was purified as described earlier by Moberg et al. (2003). The samples were analysed as described below. We also investigated cleavage of a fluorescent peptide, P1 (Ståhl et al. 2002). The degradation assay contained 0.3 µg of AtPreP1 or AtPreP2 and 20 µM of P1 (Pep Tag Protease assay) in degradation buffer. Degradation was carried out at 30°C for 1 h before 80% glycerol was added. The samples were visualized by UV light.

**In vitro import of the AtPreP2 precursor into mitochondria and chloroplasts**

The putative AtPreP2 precursor protein was synthesized in a coupled transcription/translation reticulocyte system (Promega, SDS Biosciences, Falkenberg, Sweden) in the presence of [35S]methionine. Potato mitochondria were isolated and import experiments were performed as described by Von Stedingk et al. (1997). For the import reactions, the AtPreP2 precursor was incubated with isolated mitochondria for 20 min at 25°C. Chloroplasts were isolated from spinach leaves and import experiments were performed according to Bruce et al. (1994). For the import reactions, the AtPreP2 precursor was incubated for 25 min at 25°C. Dual import was carried out with isolated spinach mitochondria and chloroplasts as described by Rudhe et al. (2002).

**Transient expression of AtPreP2-GFP fusion constructs in tobacco protoplasts**

The in vivo targeting properties of the targeting peptide of AtPreP2 were investigated using GFP as a reporter gene. One hundred and fifty-five N-terminal amino acids of AtPreP2 were fused in-frame to GFP in the AtPreP2–GFP fusion construct using the primers 5′-pre2p2 (5′-agatctCCGAGGAGTACCTTGCGTG-3′) and 3′-pre2p2gfp (5′-ggtacc-ATCCTTCGAGGAGTACCTTGCGTG-3′). The PCR product was directly cloned into a TOPO zero-blunt vector (Invitrogen) and digested with BglII and KpnI. The digested fragment was gel extracted and cloned into the pTZ19U-derived vector between the EN50PMA4 promoter and GFP (Duby et al. 2001). The expression cassette was transferred into a pBI101-derived vector (Clontech, CA, USA) for Agrobacterium-mediated transient expression. All constructs were verified by sequencing.

Protoplasts were prepared from leaves of *N. tabacum* cv. SR1 (Maliga et al. 1973) and transformed by electroporation as described by Lukaszewicz et al. (1998). Protoplasts were incubated in culture medium with 400 mM Mitotracker Red CM-H2Xros (Molecular Probes, Eugene, OR, USA) for 40 min and washed three times before confocal analysis. Confocal microscopy was performed with a Bio-Rad MRC-1024 laser-scanning confocal imaging system. For the detection of GFP, excitation was at 488 nm and detection was between 506 and 538 nm. Mitotracker staining was detected between 589 and 621 nm with excitation at 568 nm. Chloroplast autofluorescence was detected between 664 and 696 nm with excitation at 488 nm.

**Double chloroplast isolation for mass spectrometric analysis**

*Arabidopsis thaliana* was grown on 0.8% agar in Magenta boxes at 18°C and 170 µE fluorescent with incandescent lights. Plants were harvested before flowering at 21–22 d by cutting the green leaf tissue with scissors above the level of the agar, thus avoiding the roots. Chloroplasts were isolated using a method adapted from Arousson and Jarvis (2002). In brief, the plants were chopped with a razor blade in an ice-cold dish containing grinding buffer (GB; 50 mM HEPES, 330 mM sorbitol, 1 mM MgCl$_2$). The chopped plants were treated with a Polytron using a 13 mm rotor for 5–3 s. The brie was filtered through a layer of Miracloth (Calbiochem, San Diego, CA, USA) and stored at –80°C.

Frozen chloroplast pellets were resuspended in ice-cold 90% acetone and returned to –20°C for 1 h. The precipitated protein was collected by centrifugation at 20,000 × g, dissolved in 6 M guanidine and incubated at 60°C for 1 h. The solution was diluted to 1.0 M guanidine-HCl with 50 mM Tris, 1 mM CaCl$_2$, and sequencing grade, modified trypsin (Promega) was added at a level of 1:50 trypsin.
protein. The sample was rotated at 37°C overnight. The following day, a second aliquot of trypsin was added and digestion continued for another 4 h at 37°C. The digest was reduced by addition of 10 mM dithiothreitol (DTT) and heated at 60°C for 1 h. This trypsin digest was then clarified by centrifugation at 20,000 ×g to remove insoluble protein. The peptides were desalted with a C18 Sep-Pak (Waters, Milford, MA, USA), eluted with acetonitrile, and concentrated in a SpeedVac to a final volume of 100–200 μl.

One- and two-dimensional LCQ tandem mass spectrometric analysis

The purified protein digest was subjected to either one- or two-dimensional capillary/nano HPLC ESI-MS/MS equipped with an electrospray source LCQ™Deca XP (Thermo Finnigan, NJ, USA). In the one-dimensional analysis, the experiments were performed on a Ultimate HPLC (LC Packings, CA, USA) using a 150 μm i.d. × 15 cm, 300 A 5 μm C18 packing (VYDAC, Southboro, MA, USA) for separation using a linear gradient from 100% solvent A [water: acetonitrile: formic acid, 95: 0.5: 0.5 (by vol.)] to 100% solvent B (water: acetonitrile: acetic acid, 30: 70: 0.5). The eluate was analyzed by a LCQ Deca XP mass spectrometer with dynamic exclusion and data-dependent MS/MS enabled. Typically, multiple injections were made with scanning over narrow mass ranges to increase peptide detection.

In the two-dimensional analysis, the purified protein digest was injected onto a strong cation exchange column. Fractions were eluted from the SCX column using NH4OAc to a C18 trap. The trap was then backflushed to a 150 μm i.d. × 75 mm, C18 column (VYDAC) where the reverse phase separation was conducted with a linear gradient from 100% A to 100% B. The eluate was analyzed by electrospray MS/MS with dynamic exclusion and data-dependent MS/MS.

Protein assignment of DTA files

The MS/MS spectra were analyzed using the commercial algorithm, Mascot (Matrix Science, UK) using the most up to date Arabidopsis database (release date: January 22, 2004 from TIGR, Rockville, MD, USA).

Semi-quantitative RT–PCR

Total cellular RNA was extracted from 100 mg of 3-week-old seedlings and leaves, shoots, roots, flowers and silicles of mature Arabidopsis plants using plant RNA extraction reagent (Invitrogen) and was treated with RNase-free DNase (Amersham) using an RNAeasy column (Qiagen, West Sussex, UK). Cross-reactivity of the AtPreP1-specific and AtPreP2-specific primers was tested using AtPreP1 and AtPreP2 cDNAs. No PCR product was detected when AtPreP1-specific primers were applied on the cDNA of AtPreP1, and 3′-GGGAATCGAGCACAATACCGG-3′. Primers were designed with scanning over narrow mass ranges to increase peptide detection.

This work was supported by grants from The Swedish Research Council to E.G., from the Interuniversity Poles of Attraction Program-Belgian State, the Prime Minister’s Office for Scientific, Technical and Cultural Affairs and the Belgian Fund for Scientific Research to M.B. and from the National Science Foundation (MCB-0344601) to B.D.B.

References


(Rceived February 23, 2005; Accepted April 7, 2005)