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: [https://trace.tennessee.edu/utk_biocpubs](https://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 Faculty Publications and Other Works -- Biochemistry, Cellular and Molecular Biology by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.
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 Boutilier 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 sub-functionalization.

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

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 $\alpha$-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 $\alpha$-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
Dual targeting and function of the AtPreP2

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 bc\textsubscript{1} 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 (Ståhl 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) (Ståhl et al. 2002) that we have named *At*\textsuperscript{PreP1} (previously referred to as *AtZnMP*, Ståhl et al. 2002) and *At*\textsuperscript{PreP2}, respectively. Both proteases display high amino acid sequence similarity, with...
most differences occurring in their predicted organelar targeting peptides. Both \( \text{AtPreP1} \) and \( \text{AtPreP2} \) harbor a characteristic inverted zinc-binding motif, \( \text{HILEHX} \), and are classified to the pitrilsin protease subfamily \( \text{A} \). We have shown that the \( \text{PreP1} \) 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). Furthermore, we have studied targetting peptides with this protease. 

Here we have overexpressed and purified \( \text{AtPreP2} \) and investigated its proteolytic properties against targeting peptides and their mutants. Furthermore, we have studied targetting properties of \( \text{AtPreP2} \) 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 \( \text{AtPreP1} \) and \( \text{AtPreP2} \) 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 \( \text{AtPreP2} \)**

The \( \text{AtPreP2} \) 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 \( \text{AtPreP2} \) is 85 amino acid residues long and the mature protein 995 residues long. The mature portion of \( \text{AtPreP2} \) 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 \( \text{AtPreP2} \) was eluted after cleavage of the fusion protein with PreScission protease (Fig. 1a, lane 3). The recombinant \( \text{AtPreP2} \) protein (99% purity) had an estimated molecular mass of 110 kDa.

The proteolytic action of the recombinant \( \text{AtPreP2} \) was tested against a mitochondrial targeting peptide derived from the ATP synthase \( \beta \) subunit from *Nicotiana plumbaginifolia*, \( \text{N}_{5.7} \beta (2–54) \), and a chloroplastic targeting peptide derived from the small subunit of biphosphate carboxylase/oxygenase (Rubisco) from *Nicotiana tabacum*, SStpNt. Incubation of both targeting peptides with \( \text{AtPreP2} \) resulted in complete degradation of the targeting peptides (Fig. 1b, lanes 2 and 5). Addition of \( \alpha \)-phenanthroline, a specific inhibitor of the metalloprotease, completely abolished the proteolytic activity of \( \text{AtPreP2} \) (Fig. 1b, lanes 3 and 6). These results show that \( \text{AtPreP2} \) has a dual proteolytic function against both mitochondrial and chloroplastic targeting peptides.

**Differences in cleavage specificity of \( \text{AtPreP1} \) and \( \text{AtPreP2} \)**

The cleavage specificity of \( \text{AtPreP1} \) and \( \text{AtPreP2} \) was studied by investigating cleavage of a specific fluorescent peptide \( \text{P1} \), the mitochondrial presequence peptide \( \text{N}_{5.7} \beta (2–54) \)-hsl peptide and analysed on Tris-tricine gels as described in Materials and Methods.

Fig. 2 Differences in cleavage specificity of \( \text{AtPreP1} \) and \( \text{AtPreP2} \). Degradation of peptides by \( \text{AtPreP1} \) and \( \text{AtPreP2} \). (a) The fluorescent peptide \( \text{P1} \) incubated with \( \text{AtPreP1} \) or \( \text{AtPreP2} \). The degradation assays were performed as described in Materials and Methods. (b) Generation of proteolytic fragments of the \( \beta \)-presequence by \( \text{AtPreP1} \) and \( \text{AtPreP2} \). \( \text{AtPreP1} \) and \( \text{AtPreP2} \) were incubated with the \( \text{N}_{5.7} \beta (2–54) \)-hsl peptide and analysed on Tris-tricine gels as described in Materials and Methods.

**Fig. 2** Differences in cleavage specificity of \( \text{AtPreP1} \) and \( \text{AtPreP2} \). Degradation of peptides by \( \text{AtPreP1} \) and \( \text{AtPreP2} \). (a) The fluorescent peptide \( \text{P1} \) incubated with \( \text{AtPreP1} \) or \( \text{AtPreP2} \). The degradation assays were performed as described in Materials and Methods. (b) Generation of proteolytic fragments of the \( \beta \)-presequence by \( \text{AtPreP1} \) and \( \text{AtPreP2} \). \( \text{AtPreP1} \) and \( \text{AtPreP2} \) were incubated with the \( \text{N}_{5.7} \beta (2–54) \)-hsl peptide and analysed on Tris-tricine gels as described in Materials and Methods.
Alternative processing has been detected with these mutants (S. J. Wright and B. D. Bruce, unpublished). *At*PreP2 had the capacity to degrade both *S*Stp*Nt* and all the mutants, whereas *At*PreP1 could not degrade the *S*Stp*Nt*P36A 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 AtPreP2 into mitochondria and chloroplasts**

Most sequence differences between *At*PreP1 and *At*PreP2 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>
<thead>
<tr>
<th>peptide</th>
<th>peptide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S</em>Stp<em>Nt</em></td>
<td>MASIVLGDAATVRSGNAYAFTGKLGSAAFFYSQKLD LTTSSASNGGRVQQMQG</td>
</tr>
<tr>
<td><em>S</em>Stp<em>Nt</em>P36A</td>
<td>MASIVLGDAATVRSGNAYAFTGKLGSAAFFYSQKLD LTTSSASNGGRVQQMQG</td>
</tr>
<tr>
<td><em>S</em>Stp<em>Nt</em>P29A</td>
<td>MASIVLGDAATVRSGNAYAFTGKLGSAAFFYSQKLD LTTSSASNGGRVQQMQG</td>
</tr>
<tr>
<td><em>S</em>Stp<em>Nt</em>G28A</td>
<td>MASIVLGDAATVRSGNAYAFTGKLGSAAFFYSQKLD LTTSSASNGGRVQQMQG</td>
</tr>
</tbody>
</table>
Dual targeting and function of the AtPreP2

added thermolysin, showing dual import of AtPreP2 (Fig. 4c, lanes 3 and 5) and additionally supporting dual localization of AtPreP2.

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 Fβ-subunit from N. plumbaginifolia fused to GFP (Fβ–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 Fβ–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 was 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. 5C, 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 AtPreP 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).

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.
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 siliques 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 siliques 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 \textit{At}PreP1 and \textit{At}PreP2 to be close relatives and suggests the importance of searching for close homologs of these proteins in other \textit{Brassicaceae} species.

**Discussion**

In the present study, we have investigated proteolytic function, intracellular localization, expression and evolution of \textit{At}PreP2, a homolog of the targeting peptide degrading \textit{At}PreP1. The purified recombinant \textit{At}PreP2 retained the proteolytic activity against the mitochondrial and chloroplastic targeting peptides, showing that \textit{At}PreP1 and \textit{At}PreP2 in \textit{A. thaliana} are paralogs with the same cellular function. There is a single \textit{At}PreP homolog in yeast, Ydr430cp and human, hMPI. The Ydr430cp gene is non-essential for yeast survival. \textit{Arabidopsis thaliana} T-DNA insertion knock-out mutants of \textit{At}PreP1 and \textit{At}PreP2 are also available (University of Wisconsin Biotechnology Center USA, Nottingham Arabidopsis Stock Center, UK or Salk Institute Genomic Analysis Laboratory, USA). No \textit{A. 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 \textit{At}PreP1 and \textit{At}PreP2. Comparison of the \textit{At}PreP2 and \textit{At}PreP1 proteolytic activities shows that the proteases have different proteolytic subsite preference. Whereas incubation of the substrate P1 and N\textsubscript{5.7}F\textsubscript{1}β(2–54) peptides with \textit{At}PreP2 resulted in generation of an easily detected intermediate product (see Fig. 2), no such intermediate was observed upon degradation with \textit{At}PreP1. Also, degradation of the chloroplastic transit peptide SStpNt and its mutants has confirmed different cleavage specificity of \textit{At}PreP1 and \textit{At}PreP2. 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 \textit{At}PreP1, whereas it was completely degraded by \textit{At}PreP2. 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 \textit{At}PreP1 and \textit{At}PreP2 may have overlapping but complementary proteolytic specificity allowing great variety of targeting peptides being rapidly degraded.

As both TargetP and Predotar clearly suggest that \textit{At}PreP2 is localized to chloroplasts, it was of interest to investigate organellar targeting of \textit{At}PreP2. We have investigated targeting properties of \textit{At}PreP2 both in vitro and in vivo. In vitro import studies of the full-length \textit{At}PreP2 precursor in both a single and a dual import system (in the presence of a competing organelle, Rudhe et al. 2002) showed that \textit{At}PreP2 could be targeted to both mitochondria and chloroplasts, giving a strong indication of the dual localization of \textit{At}PreP2. 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
in vitro import approaches might have limitations due to a lack of an intact cellular system, we have also investigated the targeting ability of the AtPreP2 targeting peptide in vivo using GFP constructs. One hundred and fifty-five N-terminal amino acid residues of AtPreP2 were fused to GFP and transient expression of the fusion construct was performed in tobacco protoplasts (see Fig. 5) and intact tobacco leaves (data not shown). The targeting peptide of AtPreP2 dually targeted GFP 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 \textit{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 \textit{AtPreP1} transcript, the \textit{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 \textit{AtPreP1} and \textit{AtPreP2} are expressed in an organ-specific manner in \textit{A. thaliana} plants. It will be interesting to investigate the functional importance of higher transcript levels of \textit{AtPreP1} present during flower development.

\textit{AtPreP1} and \textit{AtPreP2} are classified as belonging to the \textit{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 \textit{A. thaliana} (Moore and Purugganan 2003) and in \textit{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 \textit{AtPreP1} and \textit{AtPreP2} have subfunctionalized, and analysis of DNA sequences from \textit{Arabidopsis} and other \textit{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. \textit{AtPreP1} and \textit{AtPreP2} are recent duplicates in the process of diverging in function, an ongoing process that generates evolutionary novelty in genomes.

\section*{Materials and Methods}

\subsection*{Cloning of \textit{AtPreP2}}

Full-length cDNA of \textit{AtPreP2} (RAFL09993-D24) was originally obtained from RIKEN Genomic Sciences, Japan (Seki et al. 2002). The region of the \textit{AtPreP2} precursor encoding the mature portion of the protein was amplified using \textit{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 \textit{A. thaliana} (Moore and Purugganan 2003) and in \textit{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 \textit{AtPreP1} and \textit{AtPreP2} have subfunctionalized, and analysis of DNA sequences from \textit{Arabidopsis} and other \textit{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. \textit{AtPreP1} and \textit{AtPreP2} are recent duplicates in the process of diverging in function, an ongoing process that generates evolutionary novelty in genomes.

\section*{Materials and Methods}

\subsection*{Cloning of \textit{AtPreP2}}

Full-length cDNA of \textit{AtPreP2} (RAFL09993-D24) was originally obtained from RIKEN Genomic Sciences, Japan (Seki et al. 2002). The region of the \textit{AtPreP2} precursor encoding the mature portion of the protein was amplified using \textit{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 \textit{A. thaliana} (Moore and Purugganan 2003) and in \textit{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 \textit{AtPreP1} and \textit{AtPreP2} have subfunctionalized, and analysis of DNA sequences from \textit{Arabidopsis} and other \textit{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. \textit{AtPreP1} and \textit{AtPreP2} are recent duplicates in the process of diverging in function, an ongoing process that generates evolutionary novelty in genomes.

\begin{table}
\centering
\caption{Mass spectrometric identification of \textit{AtPreP1} and \textit{AtPreP2} in \textit{Arabidopsis} chloroplasts}
\begin{tabular}{ll}
\hline
Peptides identified for \textit{AtPreP1} & Stringency score \\
ADDLFNLMLCNCLQEVQFDQQR & 41.7 \\
SLTNVEKSVAKFLDLLPENPSGGLVTWDR & 14.9 \\
GVSEENVQKVEELIMDLK & 14.3 \\
GVSEENVQKVEELIMDLK & 13.9 \\
DEPCSK & 13.0 \\
NGCIVNMTADGKSLTNVEK & 11.9 \\
DKGVAVASAEDIDAANNER & 10.8 \\
KKHMCLVNWLLSEKPLDLQTQLALGFLDHLMGTPASPLR & 10.8 \\
DEPCSK & 10.0 \\
Peptides identified for \textit{AtPreP2} & \\
VLSEYLDMFDASPDSSK & 17.6 \\
VRSGGAYGSGCDFDSHISGVFSFLSYPDNLK & 12.7 \\
VSEFISECKSK & 11.9 \\
VEELVMNTRLK & 11.6 \\
EPIYVPTEVDINGVKVLR & 11.2 \\
LLSAASRGLNGQFSRSLIR & 10.4 \\
EPIYVPTEVDINGVKVLR & 10.1 \\
\hline
\end{tabular}
\end{table}
1.8 mM KH₂PO₄, pH 7.3) and lysed by addition of 0.5 mg ml⁻¹ lysozyme 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 MgCl₂, pH 8.0. The protein content of eluted fractions from GSTrap FF and Superdex 200 HR/1030 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 F₁β-subunit of the *N. plumbaginifolia* ATP synthase, N₁pF₁β(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 N₁pF₁β(2–54)-hsl or StSPnT in the reaction buffer containing 20 mM HEPES-KOH (pH 8.0) and 10 mM MgCl₂. 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 N₁pF₁β(2–54)-hsl, samples were analyzed on 10–20% Tris-tricine gels (Bio-Rad, Sundbyberg, 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 F₁β presequences generated by AtPreP1 and AtPreP2, 2 µg of the N₁pF₁β(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 above. 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 analyzed directly on a 1% agar gel, and the fluorescent peptides 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′-agactGCCGAGAATATTCG-3′) and 3′pre2p2fp (5′-ggtacc-ATCCCTCGAGGAGTCCGT-3′). The PCR product was directly cloned into a TOPO zero-blunt vector between the EN50PM4 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 Lukasiewicz 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 Aroonsson and Jarvis (2002). In brief, the plants were chopped with a razor blade in an ice-cold dish containing gridding buffer (GB; 50 mM HEPES, 330 mM sorbitol, 1 mM MgCl₂, 1 mM MnCl₂, 2 mM EDTA, pH 7.3). The chopped plants were treated with a Polytron using a 13 mm rotor for 2–3 s. The brie was filtered through a layer of Miracloth (Calbiochem, CA, USA) supported by two layers of cheesecloth. The brie was returned to the Polytron with fresh GB and the process repeated a total of five times. The combined filtrates were centrifuged at 250–g for 6 min. The supernatant was removed and the pellets resuspended in import buffer (IB; 50 mM HEPES, 330 mM sorbitol, pH 8.0) using a soft natural bristle paint brush. The suspended chloroplasts were layered on top of a 50% Percoll gradient and centrifuged at 3,500×g for 15 min. The lower intact chloroplast band was collected, diluted 3-fold with IB, and centrifuged at 250–g for 6 min. The pellets were resuspended in fresh IB and layered on top of a second 50% Percoll gradient. The bottom layer was isolated as before and made up to a final chlorophyll concentration of approximately 1 mg ml⁻¹ in IB, and stored at –80°C.

**Total chloroplast protein tryptic digestion**

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 CaCl₂, 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 20,000g 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 an 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: 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 NH4H2O to a C4 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 analysed 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 A. thaliana 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 siliques of mature A. thaliana 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 AtPreP2, and AtPreP2-specific primers were applied on the cDNA of AtPreP1, indicating the specificity of the PCR primers. The number of cycles for quantitative RT–PCR was carefully optimized with actin-specific primers as an internal control using 15, 20 and 25 cycles. Primers were designed from the exon–exon boundary encompassing an intron in order to discriminate genomic contamination. RT–PCR, including reverse transcriptase and 25 cycles, was carried out using the Titan One Tube RT–PCR System (Roche) on 125 ng of total RNA. The primer set used for AtPreP1 was 5′rtprep1, 5′-AGATCTCTAGAGCTCTCCGCG-3′ and 3′rtprep1, 5′-GTGGAATTCCTCGAGGAGATGT-3′; for AtPreP2 it was 5′rtprep2, 5′-AGATCTCGAGGAGATATATCCGG-3′ and 3′rtprep2, 5′-TGGGAATTCCTCGAGGAGATGT-3′; and for actin it was 5′rtactin, 5′-GGACTGGAATTGTTGAGGCTT-3′ and 3′rtactin, 5′-CAGAAATCCGACACATACCCGG-3′. No PCR product was detected when reverse transcriptase was eliminated from the reaction. The RT–PCR products were analyzed by 2% agarose gel electrophoresis in the presence of ethidium bromide.

Evolutionary studies

AtPreP1 and AtPreP2 were BLASTed against embryophyte gene families from the TAED database (Liberles et al. 2001, Roth et al. 2005) and against protein sequences in GenBank. The EST collection in GenBank was then subjected to BLAST searching and 164 unique ESTs from 42 species were identified. These ESTs were manually combined into species-specific ESTs where no ambiguity existed, a multiple sequence alignment calculated using POA (Lee et al. 2002) and a phylogenetic tree calculated using Mr. Bayes (Husonbeck and Ronquist 2001) (Fig. 7). From this tree, AtPreP1 and AtPreP2 clustered in a Brassicaceae-specific group.

Acknowledgments

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)