



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](http://trace.tennessee.edu/utk_biocpubs)

 Part of the [Biochemistry Commons](#), [Cell Biology Commons](#), and the [Molecular Biology Commons](#)

## Recommended Citation

Bhushan, S.; Stahl, A.; Nilsson, S.; Lefebvre, B.; McWilliams, D.; Wright, S.J.; Seki, M.; Liberles, D.A.; Shinozaki, K.; Bruce, Barry D.; Boutry, M.; and Glaser, E., "Molecular Characterization of a isoenzyme of the targeting peptide degrading protease, PreP2-catalysis, subcellular localization, expression and evolution" (2005). *Biochemistry, Cellular and Molecular Biology Publications and Other Works*. [http://trace.tennessee.edu/utk\\_biocpubs/54](http://trace.tennessee.edu/utk_biocpubs/54)

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](mailto: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<sup>1</sup>, Annelie Ståhl<sup>1</sup>, Stefan Nilsson<sup>1</sup>, Benoit Lefebvre<sup>2</sup>, Motoaki Seki<sup>3</sup>, Christian Roth<sup>4</sup>, David McWilliam<sup>5</sup>, Sarah J. Wright<sup>6</sup>, David A. Liberles<sup>4</sup>, Kazuo Shinozaki<sup>3</sup>, Barry D. Bruce<sup>5,6</sup>, Marc Boutry<sup>2</sup> and Elzbieta Glaser<sup>1,7</sup>

<sup>1</sup> Department of Biochemistry and Biophysics, Arrhenius Laboratories for Natural Sciences, Stockholm University, 10691 Stockholm, Sweden

<sup>2</sup> 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

<sup>3</sup> Plant Functional Genomic Research Group, RIKEN GSC, Japan

<sup>4</sup> Computational Biology Unit, BCCS, University of Bergen, 5020 Bergen, Norway

<sup>5</sup> Graduate Program in Genome Science and Technology, University of Tennessee, Knoxville, TN 37996, USA

<sup>6</sup> 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.

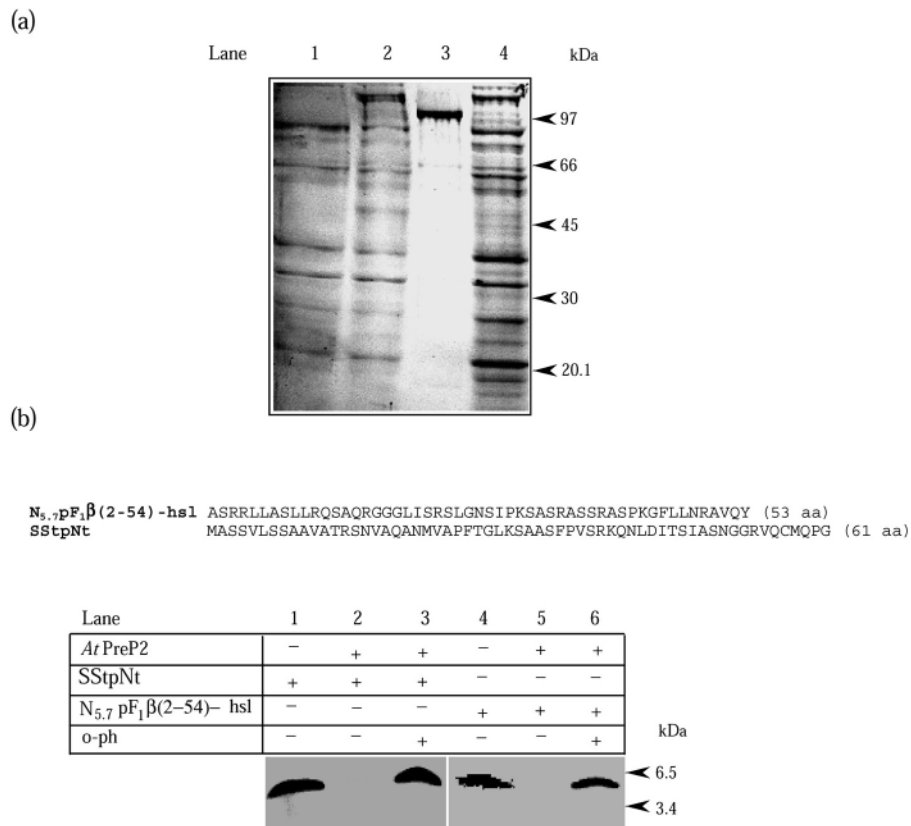
**Keywords:** Chloroplasts — Dual targeting — Mitochondria — Presequence protease — Protein import — Zinc metalloprotease.

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, presequence protease; RT–PCR, reverse transcription–polymerase chain reaction; SPP, stromal processing peptidase.

## 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 physicochemical 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

<sup>7</sup> Corresponding author: E-mail, e\_glaser@dbb.su.se; Fax, +46-81-53679.



**Fig. 1** Overexpression, purification and catalytic activity of *AtPreP2*. The mature (m) *AtPreP2* was cloned as a fusion protein with GST and the overexpression of the fusion protein was carried out in *E. coli*. (a) Expression and purity of the m*AtPreP2* analyzed by 12% SDS-PAGE. Total cell lysate from uninduced *E. coli* (lane 1) and cell lysate induced after addition of IPTG (lane 2). Total cell lysate from overexpressed *E. coli* was applied to a GSTrap FF 1-ml column, and the m*AtPreP2* was eluted after cleavage with PreScission protease (lane 3). Flow through after binding of the GST-*AtPreP2* fusion protein to the column (lane 4). (b) Mitochondrial targeting peptide N<sub>5.7</sub>pF<sub>1</sub>β(2-54)-hsl and chloroplastic targeting peptide SStpNt were incubated in the absence (lanes 1 and 4) or presence (lanes 2 and 5) of m*AtPreP2*, or in the presence of m*AtPreP2* and *o*-phenanthroline (lanes 3 and 6), as described in Materials and Methods. The amino acid sequences of both peptides are depicted at the top.

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<sub>1</sub> 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.

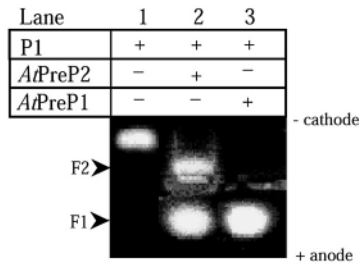
Targeting peptides are potentially harmful for the structure and function of mitochondria and chloroplasts. They can perturb natural and artificial lipid bilayers. Addition of presequences to mitochondria results in membrane lysis, uncoupling of respiration and dissipation of the membrane potential (Roise 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.

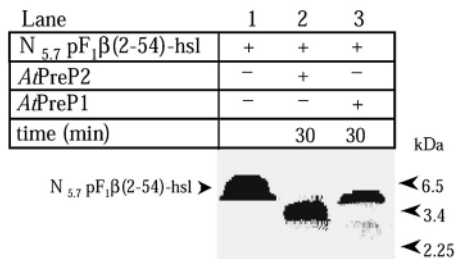
We have shown rapid proteolytic degradation of the mitochondrial presequences after import into the mitochondria by an ATP-independent protease (Ståhl et al. 2000). A protease, presequence protease (PreP), responsible for this degradation, was isolated from potato tuber mitochondrial matrix and identified 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 *AtPreP1* (previously referred to as *AtZnMP*, Ståhl et al. 2002) and *AtPreP2*, respectively. Both proteases display high amino acid sequence similarity, with

(a)

P1: Dye-Pro-Leu-Ser-Arg-Thr-Leu-Ser-Val-Ala-Ala-Lys



(b)



**Fig. 2** Differences in cleavage specificity of *AtPreP1* and *AtPreP2*. Degradation of peptides by *AtPreP1* and *AtPreP2*. (a) The fluorescent peptide P1 incubated with *AtPreP1* or *AtPreP2*. The degradation assays were performed as described in Materials and Methods. (b) Generation of proteolytic fragments of the  $F_1\beta$  presequence by *AtPreP1* and *AtPreP2*. *AtPreP1* and *AtPreP2* were incubated with the  $N_{5.7}pF_1\beta(2-54)$ -hsl peptide and analysed on Tris-tricine gels as described in Materials and Methods.

most differences occurring in their predicted organellar targeting peptides. Both *AtPreP1* and *AtPreP2* harbor a characteristic inverted zinc-binding motif, HILEHX<sub>7</sub>E, and are classified to the ptilysin 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, Moberg 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  $F_1\beta$  subunit from *Nicotiana plumbaginifolia*,  $N_{5.7}pF_1\beta(2-54)$ , 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 *AtPreP2* resulted in complete degradation of the targeting peptides (Fig. 1b, lanes 2 and 5). Addition of *o*-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.

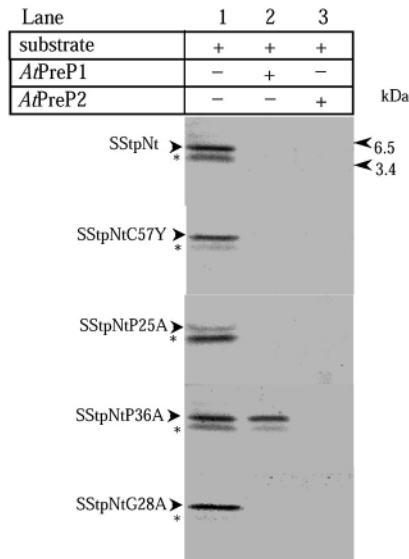
### Differences in cleavage specificity of *AtPreP1* and *AtPreP2*

The cleavage specificity of *AtPreP1* and *AtPreP2* was studied by investigating cleavage of a specific fluorescent peptide P1, the mitochondrial presequence peptide  $N_{5.7}pF_1\beta(2-54)$  (Stahl et al. 2002, Moberg et al. 2003) and the chloroplastic transit peptide of *N. tabacum* small subunit of Rubisco, SStpNt (Fig. 2). Incubation periods were shortened in comparison with those used in Fig. 1 in order to detect formation of degradation intermediates. After incubation of the peptides for 30 min at 30°C, an intermediate product was produced by *AtPreP2* with both the P1 and  $N_{5.7}pF_1\beta(2-54)$  peptide (Fig. 2a, lane 2; and b, lane 2), while no such intermediate product was seen when these two peptides were incubated with *AtPreP1* (Fig. 2a, lane 3; and b, lane 3). The intermediate could not be detected upon degradation with *AtPreP1* even when much shorter incubation times and different concentrations of *AtPreP1* were used (not shown). Accumulation of intermediate products of the peptides after incubation with *AtPreP2* shows that *AtPreP1* and *AtPreP2* have different cleavage specificity. Degradation by *AtPreP1* and *AtPreP2* has also been investigated with the chloroplastic transit peptide SStpNt and its mutants (Fig. 3). Mutants were designed to study the effect of changed flexibility of the transit peptide on import and process-

(a)

peptide	peptide sequence
SStpNt	MASSVLSSAAVATRSNVAQANMVAF <u>FTGLKSAASFPVSRKQ</u> NLDITSIASNGGRVQCMQPG
SStpNtC57Y	MASSVLSSAAVATRSNVAQANMVAF <u>FTGLKSAASFPVSRKQ</u> NLDITSIASNGGRVQCMQPG
SStpNtP25A	MASSVLSSAAVATRSNVAQANMVAF <u>FTGLKSAASFPVSRKQ</u> NLDITSIASNGGRVQCMQPG
SStpNtP36A	MASSVLSSAAVATRSNVAQANMVAF <u>FTGLKSAASFPVSRKQ</u> NLDITSIASNGGRVQCMQPG
SStpNtG28A	MASSVLSSAAVATRSNVAQANMVAF <u>FTGLKSAASFPVSRKQ</u> NLDITSIASNGGRVQCMQPG

(b)



**Fig. 3** Degradation of SStpNt and its mutants by *AtPreP1* and *AtPreP2*. (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 *AtPreP1* and *AtPreP2*. The degradation assays were performed as described in Materials and Methods. *AtPreP1* and *AtPreP2* 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.

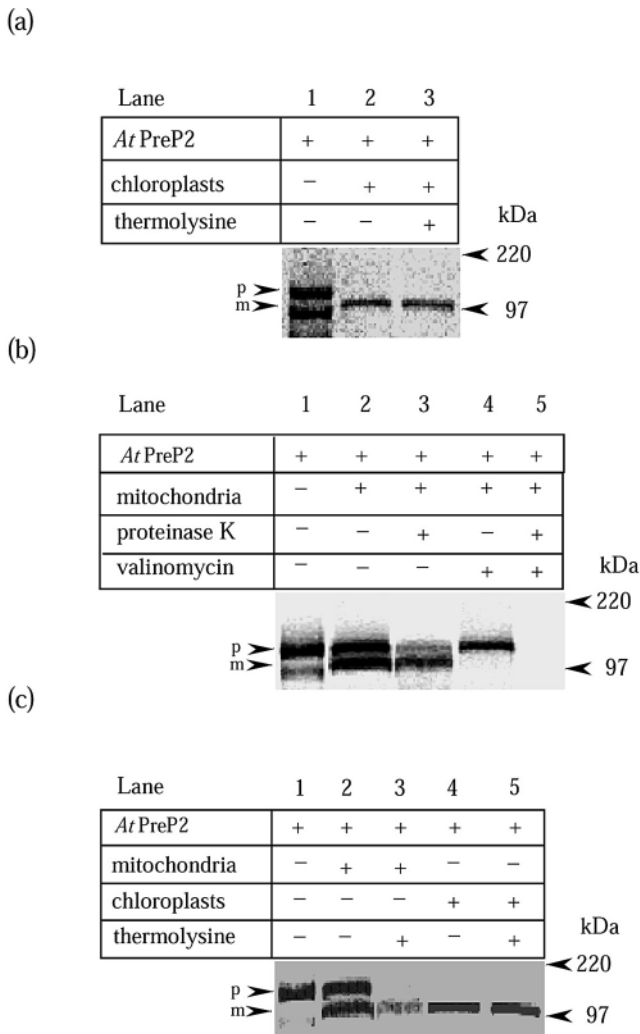
ing. Alternative processing has been detected with these mutants (S. J. Wright and B. D. Bruce, unpublished). *AtPreP2* had the capacity to degrade both SStpNt and all the mutants, whereas *AtPreP1* could not degrade the SStpNtP36A mutant with decreased flexibility as proline has been changed to alanine (Fig. 3b). These results additionally confirm the different cleavage specificity of *AtPreP1* and *AtPreP2*.

#### *In vitro* single and dual import of *AtPreP2* into mitochondria and chloroplasts

Most sequence differences between *AtPreP1* and *AtPreP2* were found in their targeting peptides. Whereas *AtPreP1* 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 organellar targeting ability, *AtPreP2* was suggested to be a chloroplastic protein. In order to study the subcellular localization of *AtPreP2*, *in vitro* import of the *AtPreP2* precursor was studied into isolated chloroplasts and mitochondria. Incubation of the *AtPreP2* 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 *AtPreP2* was fully translocated and processed in the

chloroplasts (Fig. 4a, lane 3). The additional protein band below *AtPreP2* (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 *AtPreP2* 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 *AtPreP2* protein were resistant to proteinase K (PK) treatment, showing import of *AtPreP2* into mitochondria (Fig. 4b, lane 3). Furthermore, import of the *AtPreP2* 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 *AtPreP2* might be a dually targeted protein to both chloroplasts and mitochondria.

Import of the *AtPreP2* 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 *AtPreP2* was imported and processed in both mitochondria and chloroplasts (Fig. 4c, lanes 2 and 4). The mature form of *AtPreP2* generated after processing of the precursor was resistant in both organelles to externally



**Fig. 4** In vitro import of *AtPreP2* into mitochondria and chloroplasts. (a) Chloroplastic in vitro import of *AtPreP2*. Thermolysin ( $5 \mu\text{g} \mu\text{l}^{-1}$ ) was added after import where indicated. (b) Mitochondrial in vitro import of *AtPreP2*. Proteinase K ( $10 \mu\text{g}/\mu\text{l}$ ) was added after import where indicated. Valinomycin ( $1 \mu\text{M}$ ) 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 ( $5 \mu\text{g} \mu\text{l}^{-1}$ ) was added after reisolation of the mitochondria and chloroplasts as indicated. P, precursor; m, mature.

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

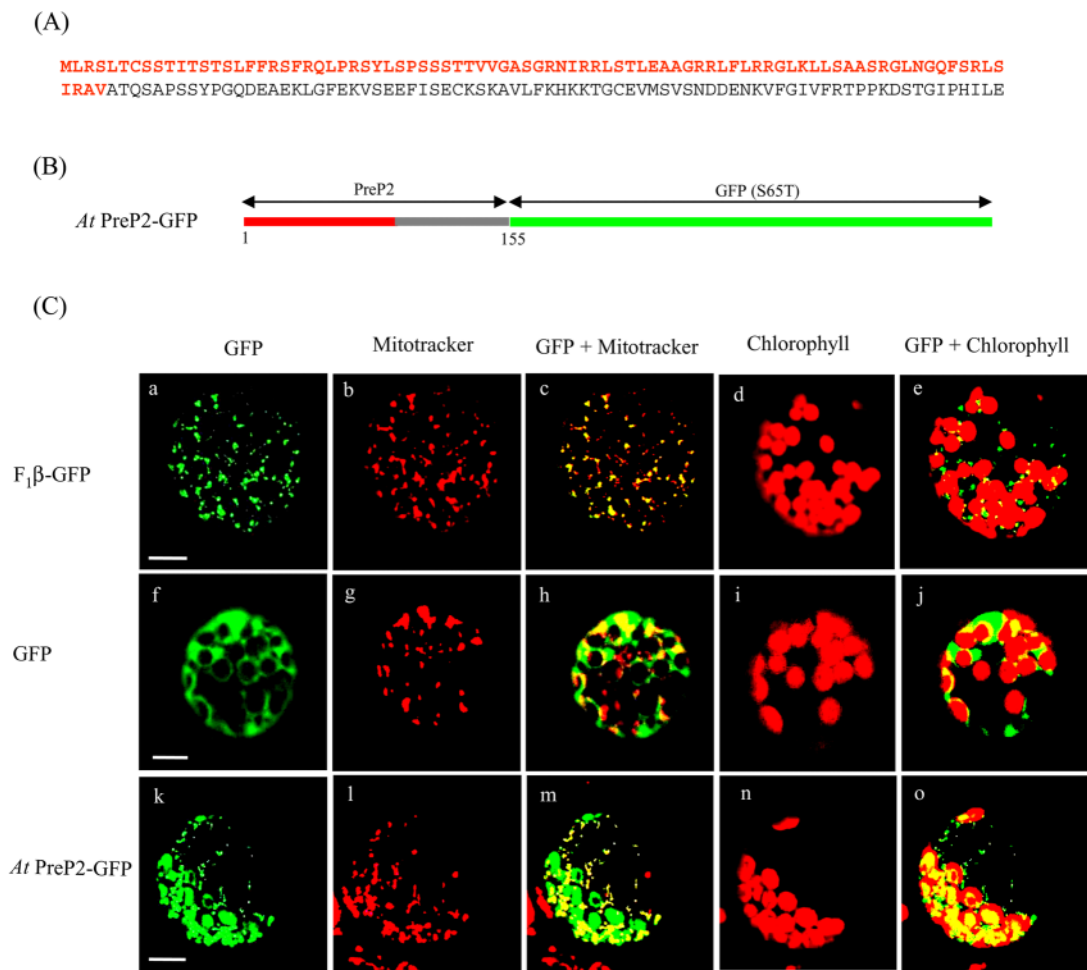
Dual import of the *AtPreP2* into mitochondria and chloroplasts was investigated further in vivo using GFP fusion and transient expression in tobacco protoplasts and leaves. The tar-

geting 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  $F_1\beta$ -subunit from *N. plumbaginifolia* fused to GFP ( $F_1\beta$ -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  $F_1\beta$ -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).



**Fig. 5** In vivo transient expression of the *AtPreP2*–GFP fusion protein in tobacco protoplasts. (A) The 155 N-terminal amino acid residues of *AtPreP2*. The predicted targeting sequence is shown in red. (B) Schematic presentation of the *AtPreP2*–GFP fusion construct used in the transient expression experiments. (C) Transient expression of the GFP fusion constructs in *N. tabacum* protoplasts: F<sub>1</sub>β-GFP (a–e), GFP alone (f–j) and *AtPreP2*–GFP (k–o), as described in Materials and Methods. The GFP column shows the signal detected in the green channel; the Mitotracker column shows the signal detected in the red channel; the GFP + Mitotracker column corresponds to the merging of the two previous columns, where yellow represents the superpositions of green and red; the chlorophyll column corresponds to the chloroplast autofluorescence signal detected in the far-red channel; and the GFP + chlorophyll column corresponds to the merging of the green channel and the chlorophyll signal detected in the far-red channel. Scale bars, 10 μm.

#### Organ-specific differential expression of *AtPreP1* and *AtPreP2* transcripts

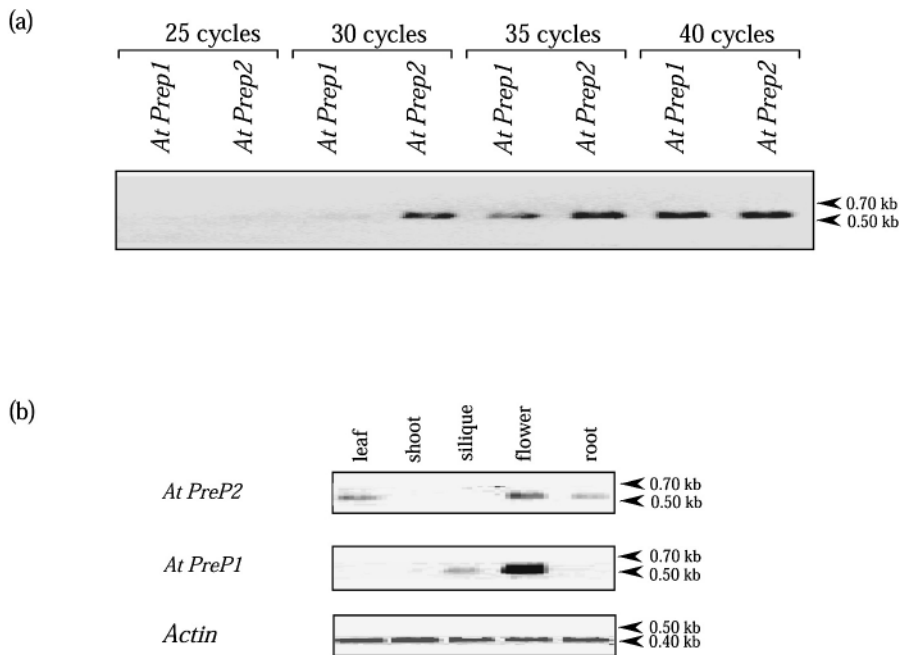
Expression of the *AtPreP1* and *AtPreP2* transcripts was studied in 3-week-old seedlings and in different organs in mature *A. thaliana* plants by a semi-quantitative RT–PCR method. PCR conditions were carefully optimized for the quantitative measurements of the transcripts level. Both the *AtPreP1* 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 sil-

ique 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-





**Fig. 6** Organ-specific differential expression of *AtPreP1* and *AtPreP2* transcripts in *A. thaliana* plants. (a) RT-PCR was performed on the total cellular RNA isolated from 3-week-old *A. thaliana* seedlings. Different numbers of cycles were employed after reverse transcriptase reactions as indicated. (b) RT-PCR was performed on the total cellular RNA isolated from leaves, shoot, siliques, flowers and roots of mature *A. thaliana* plants as described in Materials and Methods. The RT-PCR products were analysed using 2% agarose gel electrophoresis in the presence of ethidium bromide.

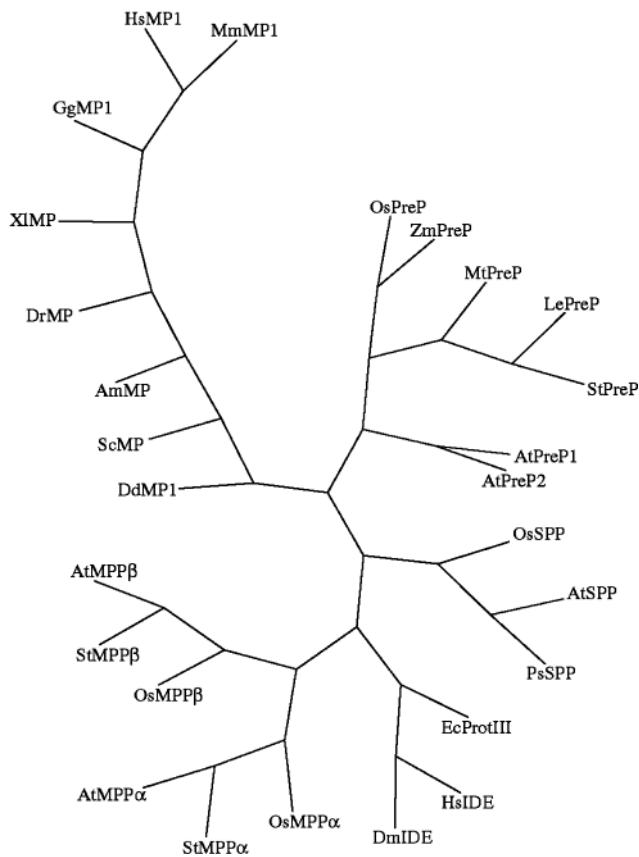
bled from a BLAST search against the NCBI expressed sequence tag (EST) collection. The sequences in this gene family were used to calculate a multiple sequence alignment using POA (Lee et al. 2002) and to build a phylogenetic tree using MRBAYES (Huelsenbeck and Ronquist 2001) (Fig. 7). This tree shows *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 *A. thaliana* are paralogs with the same cellular function. There is a single *AtPreP* homolog in yeast, Ydr430cp and human, hMPI. 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 *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 *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  $N_{5,7}P_{F1}\beta(2-54)$  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 (Stahl 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



**Fig. 7** Evolutionary analysis of *AtPreP1* and *AtPreP2* together with other homologs. An unrooted phylogenetic tree of *AtPreP1* and *AtPreP2* together with homologs from GenBank and reconstructed full-length transcripts from the EST database was calculated using Mr. Bayes. This tree shows *AtPreP1* and *AtPreP2* to be products of a recent gene duplication event, given the likely rooting. Hs, *Homo sapiens* (human); Mm, *Mus musculus* (mouse); Gg, *Gallus gallus* (chicken); Xl, *Xenopus laevis* (african clawed frog); Dr, *Danio rerio* (zebrafish); Am, *Apis mellifera* (honey bee); Sc, *Saccharomyces cerevisiae* (yeast); Dd, *Dictyostelium discoideum* (amoeba); At, *Arabidopsis thaliana* (thale cress); Mt, *Medicago truncatula* (barrel medic); Le, *Lycopersicon esculentum* (tomato); St, *Solanum tuberosum* (potato); Os, *Oryza sativa* (rice); Zm, *Zea mays* (corn); Dm, *Drosophila melanogaster* (fruit fly); Ec, *Escherichia coli* (bacterium); Ps, *Pisum sativum* (pea); SPP, stromal processing peptidase; MPP, mitochondrial processing peptidase; *AtPreP*: presequence protease; MP, metalloprotease.

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 in both systems, showing that it has ambiguous targeting properties. It has classical features of both the mitochondrial and chloroplastic targeting

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 multi-gene 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-

**Table 1** Mass spectrometric identification of *AtPreP1* and *AtPreP2* in *Arabidopsis* chloroplasts

	Stringency score
Peptides identified for <i>AtPreP1</i>	
ADDLFNL MNCLLQEVQFTDQQR	41.7
SLTNVEKSVAKFLDLLPENPSGGLVTWDGR	14.9
GVSEENVQKVEELIMDTLK	14.3
GVSEENVQKVEELIMDTLK	13.9
DEPCSK	13.0
NGCIVNMTADGKSLTNVEK	11.9
DKGVAVAVASAEDIDAANNER	10.8
KKHMLCVNWLLSEKPLDLQTLALGFLDHLMLGTPASPLR	10.8
DEPCSK	10.0
Peptides identified for <i>AtPreP2</i>	
VLSEYLD MF DASPARDSSK	17.6
VRVSGGAYGGSCDFD SHSGVFSFLSYRDPNLLK	12.7
VSEEFISECKSK	11.9
VEELVMNTRK	11.6
EPIYVPTEVGDINGVKVLR	11.2
LLSAASRGLNGQFSRLSIR	10.4
EPIYVPTEVGDINGVKVLR	10.1

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.

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 *Caenorhabditis 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 ena-

ble 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) and the primers 5'prep2m (5'-cccgggGTCGCTACACAATCCGCACC-3') and 3'prep2 (5'-gcggccgcGAGAGCTGCTTCTTACCTCGAAAA-3'). The PCR product was cloned into a zero-blunt vector (Invitrogen, Stockholm, Sweden). The insert was cleaved with *NotI* and *XhoI* and subcloned into a pGEX-6P-2 vector downstream of GST (Amersham Biosciences). In-frame cloning of the mature *PreP2* was verified by DNA sequencing using a DYEnamic sequencing kit (Amersham Biosciences, Uppsala, Stockholm, Sweden).

### 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 Na<sub>2</sub>HPO<sub>4</sub>,

1.8 mM  $\text{KH}_2\text{PO}_4$ , pH 7.3) and lysed by addition of 0.5 mg  $\text{ml}^{-1}$  lysozyme and 10  $\mu\text{g ml}^{-1}$  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  $\mu\text{m}$  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 h 10/30 column (Amersham Biosciences) equilibrated with 20 mM HEPES-KOH, 10 mM  $\text{MgCl}_2$  (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  $F_1\beta$ -subunit of the *N. plumbaginifolia* ATP synthase,  $N_{5,7}pF_1\beta(2-54)$ -hsl (Stahl et al. 2002), and a chloroplastic transit peptide derived from the small subunit of Rubisco from *N. tabacum*, SStpNt (Moberg et al. 2003). The proteolytic reaction contained 1.0  $\mu\text{g}$  of the recombinant mature *AtPreP2* and 1  $\mu\text{g}$  of either  $N_{5,7}pF_1\beta(2-54)$ -hsl or SStpNt in the reaction buffer containing 20 mM HEPES-KOH (pH 8.0) and 10 mM  $\text{MnCl}_2$ . 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  $N_{5,7}pF_1\beta(2-54)$ -hsl, samples were analyzed on 10–20% Tris-tricine gels (Bio-Rad, Sundryberg, Sweden) and stained with Coomassie brilliant blue. For the detection of SStpNt, 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 SStpNt followed by detection with horseradish peroxidase (HRP)-conjugated secondary antibodies. Point mutations in wild-type SStpNt 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 SStpNt 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  $F_1\beta$  presequences generated by *AtPreP1* and *AtPreP2*, 2  $\mu\text{g}$  of the  $N_{5,7}pF_1\beta(2-54)$ -hsl peptide was incubated with 0.3  $\mu\text{g}$  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 before. We also investigated cleavage of a fluorescent peptide, P1 (Stahl et al. 2002). The degradation assay contained 0.3  $\mu\text{g}$  of *AtPreP1* or *AtPreP2* and 20  $\mu\text{M}$  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% agarose 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 [ $^{35}\text{S}$ ]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 incu-

bated 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'prep2 (5'-agatctGCCGGAGAAAATATTCCG-3') and 3'prep2gfp (5'-ggtagc-ATCCTTCGGAGGAGTCCCTG-3'). The PCR product was directly cloned into a TOPO zero-blunt vector (Invitrogen) and digested with *Bgl*II and *Kpn*I. 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 nM 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  $\mu\text{E}$  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 Aronsson 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  $\text{MgCl}_2$ , 1 mM  $\text{MnCl}_2$ , 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 Mira cloth (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  $\text{ml}^{-1}$  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  $\text{CaCl}_2$  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 an Ultimate HPLC (LC Packings, CA, USA) using a 150 µm i.d.×15 cm, 300 Å 5 µm C<sub>18</sub> 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 NH<sub>4</sub>OAc to a C<sub>18</sub> trap. The trap was then backflushed to a 150 µm i.d.×75 mm. C<sub>18</sub> 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 RNaseasy 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′-AGATCTCAGTAGCAGTCTCCG-CCG-3′ and 3′rtrep1, 5′-GTGGAATCCTTCGGAGGAGTT-3′; for *AtPreP2* it was 5′rtrep2, 5′-AGATCTGCCGGAGAAAATATTCCG-3′ and 3′rtrep2, 5′-TGGAATCCTTCGGAGGAGTC-3′; and for actin it was 5′rtactin, 5′-GGAATGGAATGGTGAAGGCT-3′ and 3′rtactin, 5′-CAGAATCGAGACAATACCGG-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 (Huelsenbeck 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

- Abe, Y., Shodai, T., Muto, T., Mihara, K., Torii, H., Nishikawa, S., Endo, T. and Kohda, D. (2000) Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. *Cell*. 3: 551–560.
- Adam, Z., Adamska, I., Nakabayashi, K., Ostersetzer, O., Haussuhl, K., Manuell, A., Zheng, B., Vallon, O., Rodermeil, S., Shinozaki, K. and Clarke, A. (2001) Chloroplast and mitochondrial proteases in *Arabidopsis*. A proposed nomenclature. *Plant Physiol.* 125: 1912–1918.
- Arnold, I. and Langer, T. (2002) Membrane protein degradation by AAA proteases in mitochondria. *Biochim. Biophys. Acta.* 1592: 89–96.
- Aronsson, H. and Jarvis, P. (2002) A simple method for isolating import-competent *Arabidopsis* chloroplasts. *FEBS Lett.* 529: 215–220.
- Bhushan, S., Lefebvre, B., Ståhl, K.A., Wright, S.J., Mullineaux, P.M., and Glaser, E. (2003) Dual targeting and function of a protease in mitochondria and chloroplasts. *EMBO Rep.* 4: 1073–1078.
- Bruce, B.D. (2000) Chloroplastic transit peptides: structure, function and evolution. *Trends Cell Biol.* 10: 440–447.
- Bruce, B.D., Perry, S., Froelich, J. and Keegstra, K. (1994) In vitro import of proteins into chloroplasts. In *Plant Molecular Biology Manual*, 2nd edn. Edited by Gelvin, S.B. and Schilferoot, R.A. Kluwer Academic Publishers, Dordrecht, The Netherlands. pp. J1–J15.
- Chew, O., Whelan, J. and Millar, A.H. (2003) Molecular definition of the ascorbate-glutathione cycle in *Arabidopsis* mitochondria reveals dual targeting of antioxidant defenses in plants. *J. Biol. Chem.* 278: 46869–46877.
- Cleary, S.P., Tan, F.C., Nakrieko, K.A., Thompson, S.J., Mullineaux, P.M., Creissen, G.P., von Stedingk, E., Glaser, E., Smith, A.G. and Robinson, C. (2002) Isolated plant mitochondria import chloroplast precursor proteins in vitro with the same efficiency as chloroplasts. *J. Biol. Chem.* 277: 5562–5569.
- Duby, G., Oufattole, M. and Boutry, M. (2001) Hydrophobic residues within the predicted N-terminal amphiphilic alpha-helix of a plant mitochondrial targeting presequence play a major role in in vivo import. *Plant J.* 27: 539–549.
- Duckworth, W.C., Bennett, R.G. and Hamel, F.G. (1998) Insulin degradation: progress and potential. *Endocr. Rev.* 19: 608–624.
- Dyck, L.V., Pearce, D. and Sherman, F. (1994) PIM1 encodes a mitochondrial ATP-dependent protease that is required for mitochondrial function in the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* 269: 238–242.
- Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L. and Postlethwait, J. (1999) Preservation of duplicate genes by complementary degenerative mutations. *Genetics* 151: 1531–1545.
- Ginsberg, M.D., Feliciello, A., Jones, J.K., Avvedimento, E.V. and Gottesman, M.E. (2003) PKA-dependent binding of mRNA to the mitochondrial AKAP121 protein. *J. Mol. Biol.* 327: 885–897.
- Glaser, S. and Cumsky, M. (1990a) Localization of a synthetic presequence that blocks protein import into mitochondria. *J. Biol. Chem.* 265: 8817–8822.
- Glaser, S. and Cumsky, M. (1990b) A synthetic presequence reversibly inhibits protein import into yeast mitochondria. *J. Biol. Chem.* 265: 8808–8816.

- Glaser, E., Sjöling, S., Tanudji, M. and Whelan, J. (1998) Mitochondrial protein import in plants. Signals, sorting, targeting and regulation. *Plant Mol. Biol.* 38: 311–338.
- Halperin, T., Zheng, B., Itzhaki, H., Clarke, A. and Adam, Z. (2001) Plant mitochondria contain proteolytic and regulatory subunits of the ATP-dependent Clp protease. *Plant Mol. Biol.* 45: 461–468.
- Huelsenbeck, J.P. and Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.
- Hugosson, M., Andreu, D., Boman, H. and Glaser, E. (1994) Antibacterial peptides and mitochondrial presequences affect mitochondrial coupling, respiration and protein import. *Eur. J. Biochem.* 223: 1027–1033.
- Katju, V. and Lynch, M. (2003) The structure and early evolution of recently arisen gene duplicates in the *Caenorhabditis elegans* genome. *Genetics* 165: 1793–1803.
- Laemmli, U. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 15: 680–685.
- Lee, C., Grasso, C. and Sharlow, M.F. (2002) Multiple sequence alignment using partial order graphs. *Bioinformatics* 18: 452–464.
- Liberles, D.A., Schreiber, D.R., Govindarajan, S., Chamberlin, S.G. and Benner, S.A. (2001) The Adaptive Evolution Database (AED). *Genome Biol.* 2: 0028.1–0028.6.
- Lister, R., Chew, O., Lee, M.N., Heazlewood, J.L., Clifton, R., Parker, K.L., Millar, A.H. and Whelan, J. (2004) A transcriptomic and proteomic characterization of the *Arabidopsis* mitochondrial protein import apparatus and its response to mitochondrial dysfunction. *Plant Physiol.* 134: 777–789.
- Lu, Y. and Beavis, A.D. (1997) Effect of leader peptides on the permeability of mitochondria. *J. Biol. Chem.* 272: 13555–13561.
- Lukaszewicz, M., Jerouville, B. and Boutry, M. (1998) Signs of translational regulation within the transcript leader of a plant plasma membrane H<sup>+</sup>-ATPase gene. *Plant J.* 14: 413–423.
- Maliga, P., Sz-Breznovits, A. and Marton, L. (1973) Streptomycin-resistant plants from callus culture of haploid tobacco. *Nat New Biol.* 244: 29–30.
- Matsuzaki, K., Murase, O., Fujii, N. and Miyajima, K. (1996) An antimicrobial peptide, magainin 2, induced rapid flip–flop of phospholipids coupled with pore formation and peptide translocation. *Biochemistry* 35: 11361–11368.
- Moberg, P., Ståhl, A., Bhushan, S., Wright, S.J., Eriksson, A., Bruce, B.D. and Glaser, E. (2003) Characterization of a novel zinc metalloprotease involved in degrading targeting peptides in mitochondria and chloroplasts. *Plant J.* 36: 616–628.
- Moberg, P., Nilsson, S., Ståhl, A., Eriksson, A.C., Glaser, E. and Maler, L. (2004) NMR solution structure of the mitochondrial F<sub>1</sub>beta presequence from *Nicotiana glauca*. *J. Mol. Biol.* 336: 1129–1140.
- Moore, R.C. and Purugganan, M.D. (2003) The early stages of duplicate gene evolution. *Proc. Natl Acad. Sci. USA* 100: 15682–15687.
- Nicolay, K., Laterveer, F. and Heerde, W. (1994) Effects of amphipathic peptides, including presequences, on the functional integrity of rat liver mitochondrial membranes. *J. Bioenerg. Biomembr.* 26: 327–334.
- Peeters, N. and Small, I. (2001) Dual targeting to mitochondria and chloroplasts. *Biochim. Biophys. Acta* 12: 54–63.
- Perkins, D.N., Pappin, D.J., Creasy, D.M. and Cottrell, J.S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20: 3551–3567.
- Pfanner, N. and Geissler, A. (2001) Versatility of the mitochondrial protein import machinery. *Nat. Rev. Mol. Cell. Biol.* 2: 339–349.
- Rawling, N.D. and Barrett, A.J. (1991) Homologues of insulinase, a new superfamily of metalloendopeptidases. *Biochem. J.* 275: 389–391.
- Rehling, P., Wiedemann, N., Pfanner, N. and Truscott, K. (2001) The mitochondrial import machinery for preproteins. *Crit. Rev. Biochem. Mol. Biol.* 36: 291–336.
- Richter, S. and Lamppa, G.K. (1998) A chloroplast processing enzyme functions as the general stromal processing peptidase. *Proc. Natl Acad. Sci. USA* 95: 7463–7468.
- Richter, S. and Lamppa, G.K. (2002) Determinants for removal and degradation of transit peptides of chloroplast precursor proteins. *J. Biol. Chem.* 277: 43888–43894.
- Roise, D., Horvath, S., Tomich, J., Richards, J. and Schatz, G. (1986) A chemically synthesized pre-sequence of an imported mitochondrial protein can form an amphiphilic helix and perturb natural and artificial phospholipid bilayers. *EMBO J.* 5: 1327–1334.
- Roth, R.A. (1998) Introduction: clan ME containing ptilin and its relatives. In *Handbook of Proteolytic Enzymes*. Edited by Barrett, A.J., Rawlings, N.D. and Woessner, J.F. pp. 1360–1362. Academic Press, San Diego.
- Roth, C., Betts, M.J., Steffansson, P., Saelensminde, G. and Liberles, D.A. (2005) The Adaptive Evolution Database (AED): a phylogeny based tool for comparative genomics. *Nucleic Acids Res.* 33 (Database issue): D495–497.
- Rudhe, C., Chew, O., Whelan, J. and Glaser, E. (2002) A novel in vitro import system for simultaneous import of precursor proteins into mitochondria and chloroplasts. *Plant J.* 30: 213–220.
- Sagarra, M.D., Mayo, I., Marco, S., Rodriguez-Vilarino, S., Oliva, J., Carrasco, J. and Castan, J. (1999) Mitochondrial localization and oligomeric structure of HClpP, the human homologue of *E. coli* ClpP. *J. Mol. Biol.* 1: 819–825.
- Schmidt, G.W., Devillers-Thiery, A., Desruisseaux, H., Blobel, G. and Chua, N.H. (1979) NH<sub>2</sub>-terminal amino acid sequences of precursor and mature forms of the ribulose-1, 5-bisphosphate carboxylase small subunit from *Chlamydomonas reinhardtii*. *J. Cell Biol.* 83: 615–622.
- Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T., Nakajima, M., Enju, A., Akiyama, K., Oono, Y., Muramatsu, M., Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A. and Shinozaki, K. (2002) Functional annotation of a full-length *Arabidopsis* cDNA collection. *Science* 296: 141–145.
- Silva-Filho, M.C. (2003) One ticket for multiple destination: dual targeting of proteins to distinct subcellular location. *Curr. Opin. Plant Biol.* 6: 589–595.
- Ståhl, A., Moberg, P., Ytterberg, J., Panfilov, O., Brockenhuus Von Lowenhielm, H., Nilsson, F. and Glaser, E. (2002) Isolation and identification of a novel mitochondrial metalloprotease (PreP) that degrades targeting presequences in plants. *J. Biol. Chem.* 277: 41931–41939.
- Ståhl, A., Pavlov, P., Szigyarto, C. and Glaser, E. (2000) Rapid degradation of the presequence of the F1beta precursor of the ATP synthase inside mitochondria. *Biochem. J.* 349: 703–707.
- Ståhl, A., Nilsson, S., Lundberg, P., Bhushan, S., Biverstahl, H., et al. (2005) Two Novel Targeting Peptide Degrading Proteases, PrePs, in Mitochondria and Chloroplasts, so Similar and Still Different. *J. Mol. Biol.* 349: 847–860.
- Tabb, D.L., McDonald, W.H. and Yates, J.R. 3rd (2002) DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomic. *Proteome Res.* 1: 21–26.
- von Heijne, G. (1986) Mitochondrial targeting sequences may form amphiphilic helices. *EMBO J.* 5: 1335–1342.
- Von Stedingk, E.M., Pavlov, P.F., Grinkevich, V.A. and Glaser, E. (1997) Mitochondrial protein import: modification of sulfhydryl groups of the inner mitochondrial membrane import machinery in *Solanum tuberosum* inhibits protein import. *Plant Mol. Biol.* 35: 809–820.
- Whelan, J., Knorr, C. and Glaser, E. (1990) Sorting of precursor proteins between isolated spinach leaf mitochondria and chloroplasts. *Plant Mol. Biol.* 14: 977–982.
- Zhao, R.M., Moriau, L. and Boutry, M. (1999) Expression analysis of the plasma membrane H<sup>+</sup>-ATPase *pma4* transcription promoter from *Nicotiana glauca* activated by the CaMV 35S promoter enhancer. *Plant Sci.* 149: 157–165.

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