ELSEVIER

Eur. J. Cell Biol. 83 (2004); 145–152
http://www.elsevier.de/ejcb

EJCB

A novel aspartic proteinase is targeted to the secretory pathway and to the vacuole in the moss *Physcomitrella patens*

Andreas Schaaf, Ralf Reski, Eva L. Decker[1)]

Plant Biotechnology, University of Freiburg, Freiburg, Germany

Received October 29, 2003
Received in revised version March 8, 2004
Accepted March 16, 2004

Aspartic protease; Intracellular sorting; Signal peptide; Confocal microscopy; Physcomitrella patens

In seed plants aspartic proteases (APs) are known to reside in storage vacuoles. Targeting to this compartment is provoked by a secretory signal peptide and the plant-specific insert (PSI). In order to study secretory and vacuolar targeting in a seedless plant, the moss *Physcomitrella patens*, we isolated a cDNA encoding PpAP1, a novel aspartic proteinase. Sequence alignment with other members of the family of plant APs (EC 3.4.23) revealed a high overall identity and the Pfam motifs for aspartic proteinase and PSI were clearly recognised. In phylogenetic analysis PpAP1 was placed at a very basal position outside of the bigger clusters. Protoplasts transiently expressing the PpAP1 signal peptide fused to GFP showed fluorescence in a well-developed ER-Golgi network. A C-terminal fusion of GFP to the entire PpAP1 protein showed vacuolar fluorescence in transiently transfected protoplasts. Therefore, the vacuole is apparently the in-vivo target for PpAP1. In this study the three-dimensional peculiarity of the endomembrane continuum of ER and Golgi was visualised in a seedless plant for the first time. Above all the functionality of the secretory and the vacuolar targeting signals make them become useful tools for biotechnological approaches.

Introduction

The sorting of proteins to different cell compartments is an essential process in all eukaryotes. The presence of additional organelles such as plastids or the vacuole in plant cells makes the targeting machinery even more complex than in animal cells. In higher plants, the majority of the sorting mechanisms are well studied nowadays. In seedless plants, in contrast, only little is known about protein sorting and the respective signals. Being situated at the basis of land plant evolution, the moss *Physcomitrella patens* offers a perfect model for the analysis of protein targeting in seedless plants. In recent studies the isolation and functionality of plastidal and mitochondrial signal peptides from *Physcomitrella* has been shown (Kiessling et al., 2000; Richter et al., 2002). For biotechnological purposes we aimed to isolate secretory and vacuolar signals from the moss. As proteins from the family of plant aspartic proteases (plant APs) contain both signals, we decided to search for such an AP in *Physcomitrella*.

Genes for APs are found in eukaryotes, bacteria and some viruses. Like most proteinases, they are synthesised as inactive proproteins, also called zymogens. Common features of APs include an active-site cleft containing two catalytic aspartic acid residues, acidic pH optima for enzymatic activity and the presence of an inhibitory propeptide which is removed upon activation of the zymogen. Well-known representatives of the family are e.g. human pepsin (Fujinaga et al., 1995) and HIV-1 protease (Silva et al., 1996). In plants the presence of APs has been shown for several species. Special attention has been paid to cardosins and cyprosin from the thistle *Cynara cardunculus*. Due to the fact that these proteinases are abundant in pistils and flowers, extracts of the dried plant are good milk clotting agents. In Portugal and western Spain they are used in the production of sheep milk cheese (Faro et al., 1999; Frazao et al., 1999; White et al., 1999). Other plant APs have been functionally described in seeds of cocoa (Voigt et al., 1997), castor beans (Hiraiwa et al., 1997), sunflowers (Park et al., 2000), tomato leaves (Schaller and Ryan, 1996) and others.

Phytepsin from barley, which is probably the best-studied plant AP, was first extracted from resting grains (Runeberg-Roos et al., 1991, 1994). Like all members of the plant AP family, it contains a secretory signal peptide and an additional domain of about 100 residues which is not present in APs from non-plant sources and is therefore called the plant-specific insert (PSI). To reach its intracellular destination, the vacuole, the nascent protein is cotranslationally inserted into the lumen

[1)] Dr. Eva L. Decker, Plant Biotechnology, University of Freiburg, Schänzlestr. 1, D-79104 Freiburg, Germany, e-mail: eva.decker@biologie.uni-freiburg.de, Fax: +49 761 203 6945.

0171-9335/04/83/04-145 $30.00/0

of the ER. This process is mediated by its N-terminal hydrophobic signal peptide (Blobel et al., 1979). For targeting the proprotein to the vacuole an additional signal is required to conduct it away from the bulk flow of the secretory pathway. It was shown that Δ-phytepsin, in which the PSI had been removed, failed to be sorted to the vacuole and was efficiently secreted. Therefore, the domain is assumed to contain the sorting information which is essential for the collection of the proprotein in trans-Golgi vesicles (Kervinen et al., 1999; Tormakangas et al., 2001). These vesicles are subsequently transported to the vacuole where phytepsin is processed to its active two-chain form by removal of the PSI (Glathe et al., 1998).

In order to study secretory and vacuolar targeting in the moss we isolated the full-length cDNA of PpAP1, a novel aspartic proteinase. We demonstrated that the primary structure of the moss AP is similar to that of APs from higher plants. In addition, we were able to visualise the three-dimensional peculiarity of a well-developed ER-Golgi network by fusion of the PpAP1 signal peptide to GFP. Fusion of the entire PpAP1 protein to the same reporter finally revealed the vacuole as in-vivo target for the novel proteinase.

Material and methods

Plant material and growth conditions

Physcomitrella patens (Hedw.) B.S.G. was grown axenically under standard conditions (agitated liquid Knop medium, 250 mg/l KH_2PO_4, 250 mg/l $MgSO_4 \times 7H_2O$, 250 mg/l KCl, 1000 mg/l $Ca(NO_3)_2 \times 4H_2O$, 12.5 mg/l $FeSO_4 \times 7H_2O$, pH 5.8) in a growth chamber ($25 \pm 1\,°C$; light provided from above by two fluorescent tubes, Philips TL-D 36W/25; light flux of 55 $\mu mol\ s^{-1}\ m^{-2}$ outside the flasks, light-dark regime of 16:8 h). Plants were sub-cultured in 7-day intervals. Bioreactor cultivation was performed as described previously (Hohe et al., 2002).

RNA isolation

RNA was extracted from ground plant material under liquid nitrogen as described recently (Bierfreund et al., 2003).

RT-PCR

First strand synthesis was carried out with 500 ng of total RNA from wild-type protonema. SuperscriptII Reverse Transcriptase (Invitrogen GmbH, Karlsruhe, Germany) was used for first strand synthesis according to the manufacturer's protocol. PCR reactions were performed according to standard procedures (Sambrook et al., 2001).

Sequence analysis

Sequences of proteins of the aspartic protease family (EC 3.4.23) were aligned with that of PpAP1 using ClustalW (Thompson et al., 1994). Phytepsin (*Hordeum vulgare*) served as a structural template. TreeCon (Van de Peer and De Wachter, 1997) was used for constructing the neighbour-joining tree with Tajima and Nei correction taking into account all aligned positions (548 amino acids). Bootstrap analysis with 100 replicates was performed to test the significance of the nodes. Putative secretory signal sequences were analysed using the programme SignalP (Nielsen et al., 1997).

Construction of targeting vectors

For PCR amplification of the coding sequence of the PpAP1 signal peptide from cDNA, the primers PL-f (5′-GTCGACATGGGGG-CATCGAGGAGTGTT-3′) and PL-r (5′-AGATCTGCTTGCCT-CAGCTAAGGC-3′) were deduced from a published EST (acc. no. BJ199022). The amplification product was directly cloned into the pCR4-TOPO vector (Invitrogen). Introduced Sal I and Bgl II restriction sites were used for subcloning the sequence into the corresponding sites of the GFP expression vector mAV4 (Kircher et al., 1999). The resulting plasmid was named pSP-GFP.

For the construction of pSP-GFP-KDEL, amplification of GFP from mAV4 with the primers KDEL-f (5′-GAGCTCTTACAGCTCATCCT-TACCGGTACCTTTGTATAGTTCATCCATGC-3′) and KDEL-r (5′-GATCTCCCGGGATGAGTAAA-3′) led to the C-terminal fusion of the KDEL signal to the GFP-coding sequence. After cloning the PCR product into pCR4-TOPO the newly introduced restriction sites for Sma I and Sac I were used for replacing the original GFP of mAV4 by the newly generated GFP-KDEL.

The isolation of a putative AP homolog from *Physcomitrella* cDNA was performed by PCR with the primers PL-f and PhyR (5′-GGCTGCCTCGGCAAATCCTA-3′). The resulting PCR product was cloned into pCR4-TOPO. The PpAP1 cDNA was inserted into mAV4 via Xho I and Bgl II restriction sites which were introduced to the cDNA by a subsequent PCR reaction.

The plasmid pJL517 was used as a positive control for ER-targeted GFP. It carries the Hind III-Sac I fragment of pBIN m-gfp5-ER (Haseloff et al., 1997) inserted into pUC18.

Transfection of *Physcomitrella* protoplasts

The DNA for transfection was prepared with the Qiafilter Plasmid Maxi kit (Qiagen). Protoplasts (3×10^5) were transiently transfected with 25 μg of circular plasmid DNA. Protoplasts were isolated from protonema grown in bioreactors. Isolation and regeneration of protoplasts in the absence of any antibiotic drug were carried out as described previously (Rother et al., 1994).

Staining of acidic vacuoles with neutral red

Acidic vacuoles were stained with a 0.1% solution of neutral red (NR) in water. Protoplasts in 1.5 ml regeneration medium were centrifuged in an Eppendorf tube (500 rpm, 20 °C, 10 min). Approximately 1.4 ml of the supernatant were removed and discarded. Six μl of 0.1% NR were added directly to the protoplast suspension. The pH of the medium was adjusted by adding 1 μl 1 N NaOH. Protoplasts were observed after 5 min incubation.

Detection of GFP by CLSM and fluorescence microscopy

Localisation of GFP and its fusions was analysed in transfected protoplasts by confocal laser scanning microscopy (CLSM) (TCS 4D; Leica Microsystems, Heidelberg, Germany) using 488-nm excitation and two-channel measurement of emission from 510–580 nm (green/GFP) as well as 590 nm (red/chlorophyll). Cellular distribution of the PpAP1-GFP fusion protein was additionally visualised by fluorescence microscopy (Olympus BX41; filter U-M61002 and Zeiss Axioplan; filterset 09).

Results

Isolation of PpAP1 cDNA and sequence analysis

TBLASTN searches with the amino acid sequence of barley phytepsin as query against *Physcomitrella* ESTs yielded several hits for both termini of the peptide sequence. Using this information primers were deduced to amplify the coding sequence of the putative AP from randomly transcribed *Physcomitrella* cDNA. Sequencing of the cloned PCR product revealed an open reading frame of 1512 bp. The deduced protein sequence of 504 amino acids shows a high degree of similarity to several plant APs (Fig. 1a). The putative protein was termed PpAP1 (EMBL acc. no. AJ586914). To confirm the presence of characteristic domains and structural elements the sequence of PpAP1 was scanned for conserved domains using

a

```
                        PpAP1 signal peptide

Horvu Phytepsin     MGTRGLALALLAAVLLLQTVLPAASEAEGLVRIALKKRPIDRNSRVATGLSG--GEEQPLLSGANPLRSEE-EGDI-VALKNYMNAQYFGEIGVGTPPQK
Orysa Oryzasin 1    MGTRSVALVLLAAVLL-QALLP-ASAEEGLVRIALKKRPIDENSRVAARLSGEEGARRLGLRGANSLGGGGGEGDI-VALKNYMNAQYFGEIGVGTPPQK
Arath AP            --------------------GRFAERNDGTFRVGLKKLKLDSKNRLAARVES----KQEKPLRAYRLGDSG-DADV-VVLKNYLDAQYYGEIAIGTPPQK
Cynca Cyprosin      ----------------------------------LKKFKVNILNHFGEHAGSNDANARRKYGVRGNFRDSD--GEL-IALKNYMDAQYFGEIGIGTPPQK
Cynca Cardosin A    MGTSIKANVLALFLFYLLSPTVFSVSDDGLIRIGLKKRKVDRIDQLRGRRALMEGNARKDFGFRGTVRDSG--SAV-VALTNDRDTSYFGEIGIGTPPQK
Phypa AP1           ---MGASRSVRLAFFLVVLVVLAALAEASPRRIALKKRPVTLQSVRNAASRT-IQRAKTFTRSEDELRDGE---DI-VALNNYLDAQYFGEIGIGSPPQP
Homsa Cathepsin D   ---------MQPSSLLPLALCLLAAPASALVRIPLHKFTSIRRTMSEVGGSVEDLIAKGPVSKYSQAVPAVTEGPIPEVLKNYMDAQYYGEIGIGTPPQC
Neucr AP            --MKGALLTAAMLLGSAQAGVHTMKLKKVPLADELESVPIDVQVQHLGQKYTG-LRTESHTQAMFKATDAQVSGNHPVPITNFMNAQYFSEITIGTPPQT

                       as 1

Horvu Phytepsin     FTVIFDTGSSNLWVPSAKC-YFSIACYLHSRYKAGASSTYKKNGKPAAIQYGTGSIAGYFSEDSVTV-----------GDLVVKDQEFIEATKEPGITFL
Orysa Oryzasin 1    FTVIFDTGSSNLWVPSAKC-YFSIACFFHSRYKSGQSSTYQKNGKPAAIQYGTGSIAGFFSEDSVTV-----------GDLVVKDQEFIEATKEPGLTFM
Arath AP            FTVVFDTGSSNLWVPSSKC-YFSLACLLHPKYKSSRSSTYEKNGKAAAIHYGTGAIAGFFSNDAVTV-----------GDLVVKDQEFIEATKEPGITFV
Cynca Cyprosin      FTVIFDTGSSNLWVPSSKC-YFSVACLFHSKYRSTDSTTYKKNGKSAAIQYGTGSISGFFSQDSVKL-----------GDLLVKEQDFIEATKEPGITFL
Cynca Cardosin A    FTVIFDTGSSVLWVPSSKC-INSKACRAHSMYESSDSSTYKENGTFGAIIYGTGSITGFFSQDSVTI-----------GDLVVKEQDFIEATDEADNVFL
Phypa AP1           FAVIFDTGSSNLWVPSAKC-YLSIACYFHHRYKSGKSSTYKEDGTSFAIQYGTGSMEGFLSQDHVTL-----------GDLTVKGQVFAEATKEPGLTFV
Homsa Cathepsin D   FTVVFDTGSSNLWVPSIHCKLLDIACWIHHKYNSDKSSTYVKNGTSFDIHYGSGSLSGYLSQDTVSVPCQSASSASALGGVKVERQVFGEATKQPGITFI
Neucr AP            FKVVIDTGSSNLWVPSSQC--GSIACYLHNKYESSESSTYKKNGTSFKIEYGSGSLSGFVSQDRMTI-----------GDITINDQLFAEATSEPGLAFA

                                                        euk. asp. protease

Horvu Phytepsin     VAKFDGILGLGFKEISVGKAVPVWYKMIEQGLVSDPVFSFWLNRHVDEGEGGEIIFGGMDPKHYVGEHTYVPVTQKGYWQFDMGDVLVGGKSTGFCAGGC
Orysa Oryzasin 1    VAKFDGILGLGFQEISVGDAVPVWYKMVEQGLVSEPVFSFWFNRHSDEGEGGEIVFGGMDPSHYKGNHTYVPVSQKGYWQFEMGDVLIGGKTTGFCASGC
Arath AP            VAKFDGILGLGFQEISVGKAAPVWYNMLKQGLIKEPVFSFWLNRNADEEEGGELVFGGVDPNHFKGKHTYVPVTQKGYWQFDMGDVLIGGAPTGFCESGC
Cynca Cyprosin      AAKFDGILGLGFQEISVGDAVPVWYTMLNQGLVQEPVFSFWLNRNADEQEGGELVFGGVDPNHFKGEHTYVPVTQKGYWQFEMGDVLIGDKTTGFCASGC
Cynca Cardosin A    HRLFDGILGLSFQTI----SVPVWYNMLNQGLVKERRFSFWLNRNVDEEEGGELVFGGLDPNHFRGDHTYVPVTYQYYWQFGIGDVLIGDKSTGFCAPGC
Phypa AP1           VAKFDGILGLGFKEISVNRVTPPWYNMLDQGLVKEPVFSFWLNRNPDESSGGELVLGGVDPKHFKGEHVYTPVTRKGYWQFDLGDVTINGRTTGFCANGC
Homsa Cathepsin D   AAKFDGILGMAYPRISVNNVLPVFDNLMQQKLVDQNIFSFYLSRDPDAQPGGELMLGGTDSKYYKGSLSYLNVTRKAYWQVHLDQVEVASGLT-LCKEGC
Neucr AP            FGRFDGILGLGYDRLAVPGITPPFYKMVEQKLVDEPVFSFYL---ADQDGESEVVFGGVNKDRYTGKITTIPLRRKAYWEVDFDAIGYGKDFAEL--EGH

                      as 2                               Sap B2

Horvu Phytepsin     AAIADSGTSLLAGPTAIITEINEKIGAAGVVSQECKTIVSQYGQQILDLLLAETQPKKICSQVGLCTFDGTRGVSAGIRSVVDDEPVKSNGLRADPMCSA
Orysa Oryzasin 1    SAIADSGTSLLAGPTAIITEINEKIGATGVVSQECKTVVSQYGQQILDLLLAETQPSKICSQVGLCTFDGKHGVSAGIKSVVDDEAGESNGLQSGPMCNA
Arath AP            SAIADSGTSLLAGPTTIITMINHAIGAAGVVSQQCKTVVDQYGQTILDLLLSETQPKKICSQIGLCTFDGTRGVSMGIESVVDKENAKLSNGVGDAACSA
Cynca Cyprosin      AAIADSGTSLLAGTTTIVTQINQAIGAAGVMSQQCKSLVDQYGKSMIEMLLSEEQPEKICSQMKLCSFDGSHDTSMIIESVVDKSKGKSSGLP--MRCVP
Cynca Cardosin A    QAFADSGTSLLSGPTAIVTQINHAIGANGVMNQQCKTVVSRYGRDIIEMLRSKIQPDKICSHMKLCTFDGARDVSSIIESVVDKNNDKSSGGIHDEMCTF
Phypa AP1           TAIADSGTSLLAGPSGIVAEINQAIGATGVVSQQCKMVVQQYGDQIVEMLLAQMNPGKVCTTLGLCNFGAG---EPGIASVVEKD--QSHSLREDPLCTV
Homsa Cathepsin D   EAIVDTGTSLMVGPVDEVRELQKAIGA-------------------------------------------------------------------------
Neucr AP            GVIDDTGTSLIALPSQLAEMLNAQIGA-------------------------------------------------------------------------

                             Sap B1                                           euk. asp. protease

Horvu Phytepsin     CEMAVVWMQNQLAQNKTQDLILDYVNQLCNRLPSPMGESAVDCGSLGSMPDIEFTIGGKKFALKPEEYILKVGEGAAAQCISGFTAMDIPPPRGPLWILG
Orysa Oryzasin 1    CEMAVVWMQNQLAQNKTQDLILNYINQLCDKLPSPMGESSVDCGSLASMPEISFTIGAKKFALKPEEYILKVGEGAAAQCISGFTAMDIPPPRGPLWILG
Arath AP            CEMAVVWIQSQLRQNMTQERILNYVNELCERLPSPMGESAVDCAQLSTMPTVSLTIGGKVFDLAPEEYVLKVGEGPVAQCISGFIALDVAPPRGPLWILG
Cynca Cyprosin      CARWVVWMQNQIRQNETEENIINYVDKLCERLPSPMGESAVDCGSLSSMPNIAFTVGGKTFNLSPEQYVLKVGEGATAQCISGFTAMDVAPPHGPLWILG
Cynca Cardosin A    CEMAVVWMQNEIKQSETEDNIINYANELCEHLSTSSEELQVDCNTLSSMPNVSFTIGGKKFGLTPEQYILKVGKGEATQCISGFTAMD-ATLLGPLWILG
Phypa AP1           CGMAVVWAQNQLSQNRTKEQIDAYLNQLCERLPSPNGESAVDCNSLSSMPNVAFTISNKTFELKPEEYILKIGEGAEAQCISGFLGLDVPPPAGPLWILG
Homsa Cathepsin D   ------------------------------VPLIQGEYMIPCEKVSTLPAITLKLGGKGYKLSPEDYTLKVSQAGKTLCLSGFMGMDIPPPSGPLWILG
Neucr AP            ------------------------------KKSWNGQFTIDCGKKSSLEDVTFTLAGYNFTLGPEDYILEAS----GSCLSTFMGMDMPAPVGPLAILG

Horvu Phytepsin     DVFMGPYHTVFDYGKLRIGFAKAA-----
Orysa Oryzasin 1    DVFMGAYHTVFDYGKMRVGFAKSA-----
Arath AP            DVFMGKYHTVFDFGNEQVGFAEAA-----
Cynca Cyprosin      DVFMGQYHTVFDYGNLRVGFAEAA-----
Cynca Cardosin A    DVFMRPYHTVFDYGNLLVGFAEAA-----
Phypa AP1           DVFMGVYHTVFDFGNTRLGFAEAAKGEFV
Homsa Cathepsin D   DVFIGRYYTVFDRDNNRVGFAEAARL---
Neucr AP            DAFLRKYYSIYDLGADTVGIATAKR----
```

b

Fig. 1. a: Sequence alignment. The amino acid sequence of PpAP1 was aligned with those of other APs from plant and non-plant sources using ClustalW. Barley phytepsin served as a structural template. as1, as2: active sites of aspartic proteases. Domains (as recognised by the programme Pfam), euk. asp. protease: eukaryotic aspartyl protease; SapB1, SapB2: saposin-like domains B1 and B2 of the PSI. b: Schematic domain structure of PpAP1. Domains of PpAP1 were predicted with the programmes SignalP, Pfam and by ClustalW alignment with other APs. sp: secretory signal peptide; pro: propeptide; PSI: plant-specific insert.

the Pfam database. Significant hits were found for the asp motif (E-value $9.9 \times$ E-175) and the two saposin-like type B domains (SapB 1, SapB 2; $4.5 \times$ E-16 and $8.1 \times$ E-16, respectively). The asp motif represents the conserved domains of eukaryotic aspartyl proteases including the two active sites which are also present in the moss sequence. The recognised SapB motifs mark the PSI which is known to be structurally related to saposins and has therefore been termed saposin-like domain in

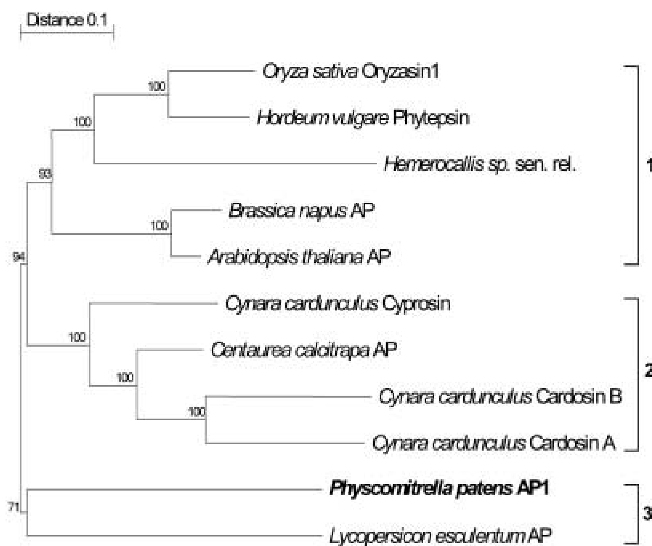

Fig. 2. Unrooted neighbour-joining tree of the plant aspartic proteases with Tajima and Nei correction for distance calculation. Brackets mark the monophyletic clusters 1–3. Bootstrap support (small numbers) was calculated with 100 samples.

previous publications. In human and fungal AP sequences the PSI is missing. Further analysis with the programme SignalP clearly showed the presence of a secretory signal peptide at the N-terminus of PpAP1 (mean S-value: 0.941; max. C-value: 1.000 at pos. 25). In Figure 1b a schematic domain structure based on the results of the sequence analysis is presented.

Phylogenetic analysis

Within the family of plant APs only few enzymes have been functionally characterised. In order to get evidence for a putative function of PpAP1 we aligned its peptide sequence to those of the characterised APs (Table 1) and calculated their phylogenetic distance. The resulting tree (Fig. 2) splits into three monophyletic clusters which were assigned the numbers 1–3. Cluster 1 contains the proteinases from barley, rice and rape, which have been described to be involved in mobilisation of storage proteins and metabolic turnover (Asakura et al., 1995; D'Hondt et al., 1997; Glathe et al., 1998). Cluster 2 includes all thistle APs. Within this group the *Centaurea* proteinase has been assigned a role in senescence of flowers (Domingos et al., 2000). For cardosin A a function in pollen recognition or germination was proposed (Frazao et al., 1999). Cluster 3 is built up by the *Physcomitrella* AP and a proteinase from tomato. The expression pattern of the latter was described

Table. 1. Aspartic proteases used in phylogenetic tree analysis

Species	Protein	Acc.-No.
A. thaliana	AP	AAC49730
Hemerocallis sp.	senescence associated protein 4	AAC34854
O. sativa	Oryzasin 1	Q42456
C. calcitrapa	AP	CAA70340
C. cardunculus	Preprocardosin A	CAB40134
C. cardunculus	Preprocardosin B	CAB40349
C. cardunculus	Cyprosin	P40782
H. vulgare	Phytepsin	P42210
L. esculentum	AP	S71591
B. napus	AP	AAB03108

to be regulated like that of systemic wound response proteins. Therefore, a role of the tomato AP in defence response was proposed (Schaller and Ryan, 1996).

Intracellular localisation of signal peptide-GFP fusion proteins

To investigate the functionality of the predicted signal peptide of PpAP1 two different GFP fusion constructs were made. pSP-GFP contains the coding sequence for the 26 amino acids of the signal fused to the coding sequence for GFP (Fig. 3d). In pSP-GFP-KDEL the ER-retention signal was added to the C-terminus of GFP (Fig. 3h). pGFP (Fig. 3m) lacking a signal peptide served as a control. Protoplasts were transiently transfected with the different plasmids and GFP fluorescence was detected using confocal microscopy.

Protoplasts expressing GFP without any signal peptide showed fluorescence distributed in a diffuse manner all over the cytosol (Fig. 3i). Once synthesised, part of the cytoplasmic GFP entered the nucleus which could therefore be detected as a filled globular structure (Fig. 3l). The tendency to accumulate in the nucleus is characteristic for cytosolic GFP and has been described before (Haseloff et al., 1997).

In contrast, cells transfected with pSP-GFP clearly showed GFP targeted to the endomembrane system forming a finely ramified network pervading the whole cell (Fig. 3a). The peripheral section (Fig. 3b) demonstrates the close meshes of this network probably representing the cortical ER. Additional evidence for GFP being efficiently targeted to the ER is provided by the shape of the nucleus (Fig. 3c). In contrast to the leader-less control, the reporter protein accumulated in the nuclear envelope whereas no fluorescence could be detected within the nucleus (Fig. 3c).

Transfection with pSP-GFP-KDEL slightly changed the distribution pattern of GFP compared to pSP-GFP. The fluorescence pattern also resembled a network but – compared with the non-retarded construct – GFP accumulation in the nuclear envelope was more intense here. Additionally, areas of bright fluorescence both on the cell surface as well as in the inner cytosol were observed. Localisation of these areas at the branching points of the network suggest the retention of GFP within the endomembrane system (Fig. 3e–g).

As a positive control for ER-retained GFP, protoplasts were transformed with the plasmid pJL517 where GFP is targeted to the ER by the secretory leader peptide of *Arabidopsis* basic chitinase (Fig. 3o). Efficient targeting of GFP to the ER network with this construct was shown recently (Haseloff et al., 1997). In transformed *Physcomitrella* protoplasts GFP was targeted efficiently to the ER which appeared as the same network-like structures that were detected upon transformation with pSP-GFP and pSP-GFP-KDEL (Fig. 3n).

Taken together these results clearly prove the functionality of the predicted secretory signal peptide of PpAP1.

Localisation of the PpAP1-GFP fusion protein

As most APs from higher plants are described to be targeted to the central vacuole by the PSI we aimed to study vacuolar targeting of PpAP1. For this reason, the GFP-coding sequence was fused to the 3′-end of the entire PpAP1-coding sequence. The resulting construct PpAP1-GFP (Fig. 4a) was used for transient transformation of *Physcomitrella* protoplasts. Seven days after the transformation GFP fluorescence could be observed. Fig. 4 b and 4e show fluorescence microscopic images of transiently transfected protoplasts. For comparison,

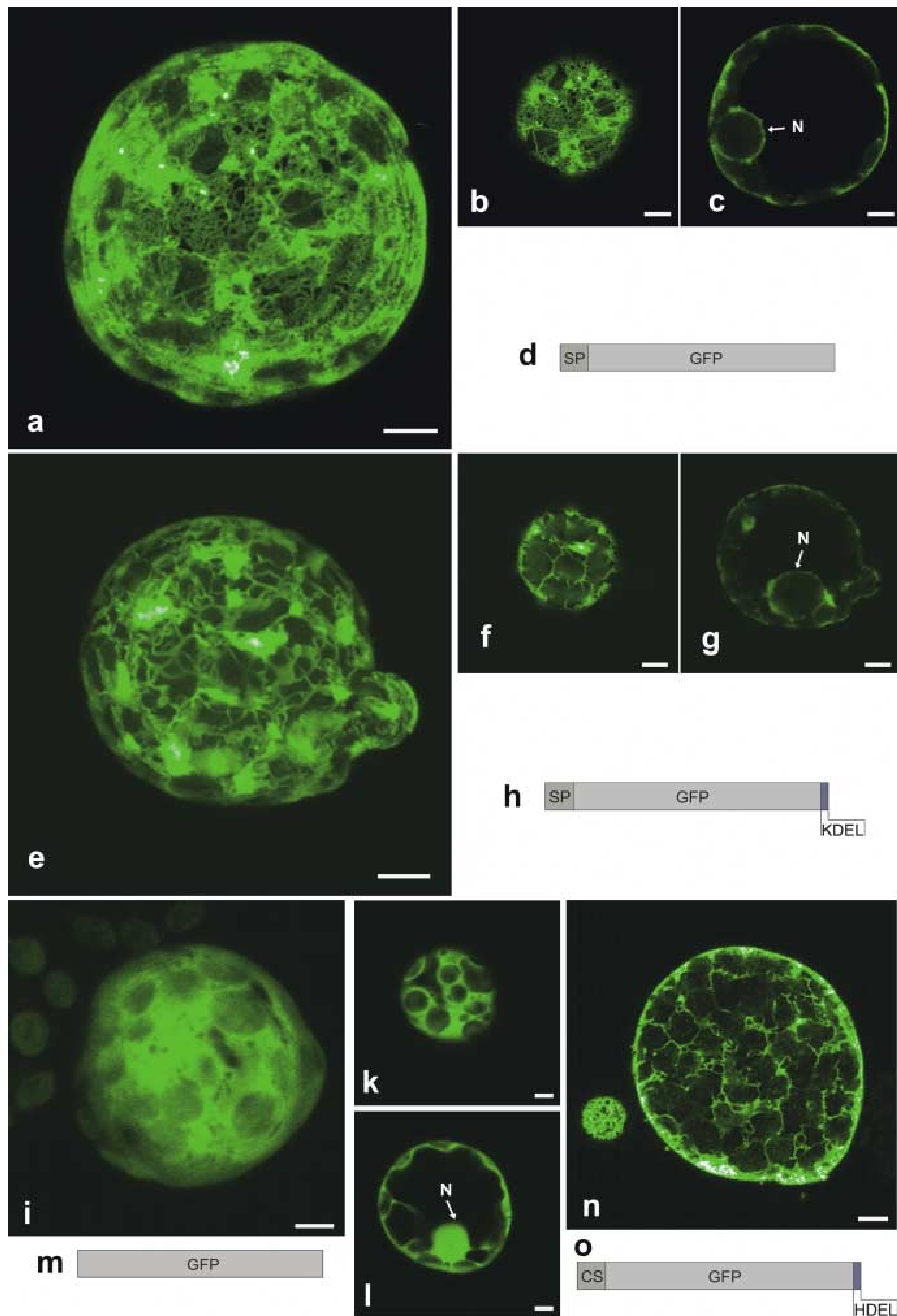

Fig. 3. Localisation of SP-GFP fusion proteins. The secretory signal peptide of PpAP1 was fused to GFP. Fluorescence of transiently transfected protoplasts was observed one to two days after transfection by confocal microscopy. a–c: Fluorescence pattern of pSP-GFP (construct d). e–g: Fluorescence pattern of pSP-GFP-KDEL (construct h). i–l: Fluorescence pattern of pGFP without any signal peptide (construct m). n: Fluorescence pattern of pJL517 (positive control for ER-targeted GFP; construct o). a, e, i: Three-dimensional reconstructions from confocal sections. b, c, f, g, k, l, n: Confocal sections. N: nucleus. Bars: 5 µm.

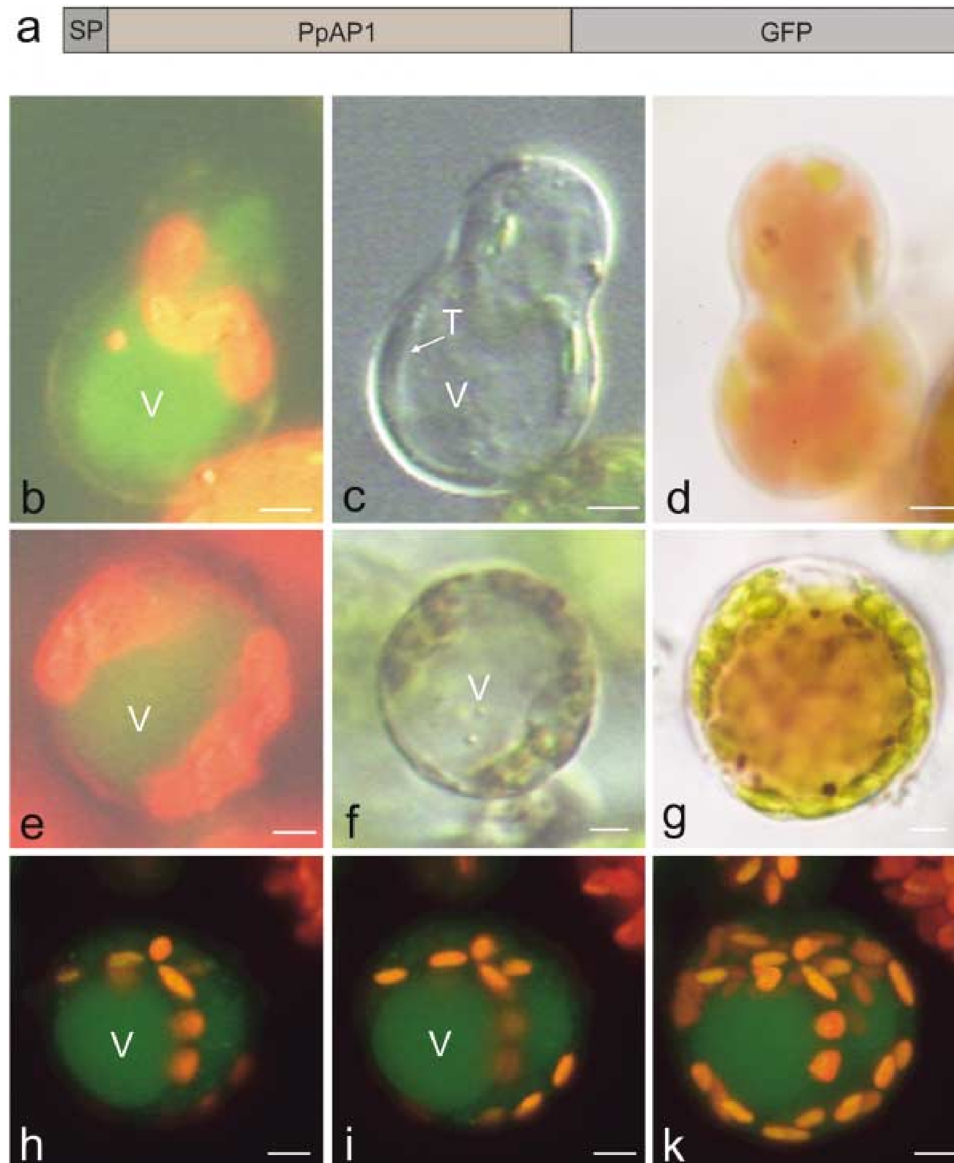

Fig. 4. Localisation of the PpAP1-GFP fusion protein. GFP was fused to the C-terminus of the entire PpAP1 protein. Fluorescence patterns of transiently transfected protoplasts were observed seven days after transformation. a: PpAP1-GFP fusion construct. b, e: Green fluorescence of GFP localised by fluorescence microscopy (chloroplasts appear red due to autofluorescence of chlorophyll). c, f: DIC images of the same protoplasts. d, g: Brightfield images of protoplasts stained with neutral red which accumulates in acidic vacuoles. h–k: Confocal images of a transfected protoplast expressing PpAP1-GFP. h, i: Single-layer scans. k: Three-dimensional reconstruction of the whole protoplast. V: vacuole; T: tonoplast. Bars: 5 µm.

differential interference contrast (DIC) images of the same protoplasts (Fig. 4c and f) are shown. Figures 4d and g show protoplasts stained with NR which specifically accumulates in acidic vacuoles where it becomes protonated and therefore changes its colour to far-red. Comparison of the fluorescence microscopic and the DIC images reveals that the green fluorescence of the GFP marker is clearly localised within the contour of the vacuole which is visible in the DIC images (marked V). Additional evidence is provided by NR stains of the central vacuole. Figures 4d and g reveal that the big central compartment which encloses the GFP fluorescence can be specifically stained with NR. As there is no other acidic compartment of that size known in plant cells we conclude that the central acidic vacuole is the in-vivo target of PpAP1. In Figure 4h–k confocal images of a protoplast expressing PpAP1-GFP are shown. In spite of faint fluorescence due to the low pH (for details see Discussion) GFP could again be detected within the vacuole. The three-dimensional reconstruction of the protoplast (Fig. 4k) localises the chloroplasts outside the green area of accumulated GFP. These findings corroborate the conclusion that PpAP1-GFP is in the vacuole, a comparably voluminous central compartment which presses all other organelles to a peripheral place near the plasmalemma.

Therefore, the central vacuole can be assumed to be the target compartment of PpAP1 as it is for most of the described APs of higher plants.

Discussion

Protein targeting to plant vacuoles is not a stereotypical process. It was shown to be dependent on a variety of different signals and their respective receptors. Not all of them have been fully characterised yet (for review see (Hadlington and Denecke, 2000; Neumann et al., 2003)). In the case of plant APs several studies found the PSI to contain information that is essential for the vacuolar targeting of proteins (Kervinen et al., 1999; Tormakangas et al., 2001). The proposed targeting route for these proteins starts by cotranslational insertion of their precursors into the ER. This process is provoked by the presence of a hydrophobic N-terminal signal peptide (Walter and Blobel, 1981). Once arrived in the ER, the signal peptide is cleaved off and the proproteins are collected in COPII vesicles that subsequently mediate the transport to the Golgi cisternae. After passing through the Golgi stack the proteins are sorted to distinct microdomains of the trans side. Here, they are packed in vesicles which then transport their cargo to the protein storage vacuoles. At this step, the PSI is assumed to interact with the membrane and to enable vacuolar sorting this way (Faro et al., 1999; Kervinen et al., 1999). However, it must be noted that other studies found barley phytepsin to reside in both, the protein storage vacuole and the lytic vacuole (Paris et al., 1996). This finding would require a different targeting mechanism for the lytic compartment.

With the aim of studying secretory and vacuolar targeting in the moss *Physcomitrella patens*, we isolated a cDNA showing a high degree of identity to all other plant aspartic proteinases. The characteristic domains of aspartic proteinases including active sites and the plant-specific saposin-like domain are present (Pfam motifs).

As in-silico analyses of the *Physcomitrella* sequence predicted the presence of an N-terminal signal peptide, we fused it to GFP in order to trace the secretory pathway in a bryophyte. The results obtained by confocal microscopy visualise a well-developed endomembrane network comparable to that of higher plants (Boevink et al., 1998). The observation that SP-GFP clearly accumulates in the nuclear membrane but does neither appear in the nucleoplasm nor in the cytosol additionally proves functionality of the PpAp1 signal peptide.

Addition of the ER-retention signal KDEL (Munro and Pelham, 1987) to the C-terminus of the SP-GFP fusion protein slightly changed the fluorescence pattern of transfected protoplasts. Larger areas of fluorescence at the branching points of the network suggest an accumulation of the reporter within the ER. Nevertheless, structures originating from the nuclear envelope branching out in the cytosol combined with a GFP accumulation in the envelope itself indicate an efficient secretory targeting of the reporter fusion in all observed cases.

In addition to the secretory signal, the presence of the plant-specific insert could be shown in sequence analyses. As previously discussed, this domain is thought to contain vacuolar sorting information. In order to prove the presence of this signal in the moss AP and to reveal its in-vivo target, we fused GFP to the C-terminus of the entire PpAP1 protein. A general problem of GFP as an in-vivo marker is the sensitivity of its fluorescence properties to pH conditions. For the S65T variant which was used in this study it was reported that fluorescence decreases to about one tenth when pH drops from 7.5 to 5.0 (Kneen et al., 1998). As plant vacuoles have a pH of around 5 to 6 (Swanson et al., 1998; Buchanan et al., 2000) we expected a much weaker fluorescence in case of the PpAP1 fusion being located to the central vacuole. Nonetheless, GFP accumulation in this compartment could be detected by confocal microscopy and light microscopy which was strengthened by NR staining that is specific for the central vacuole. These findings together with the in silico predicted PSI demonstrate that PpAP1 is targeted to the vacuole.

To get an idea of a putative function of PpAP1, we calculated phylogenetic distances of the protein to a set of functionally described plant APs. In the resulting tree three monophyletic clusters are delimited with rather good bootstrap support. Cluster 3 is built up of the moss AP and an aspartic proteinase from tomato which has been proposed to be functionally involved in systemic wound response. However, this proposal was not based on biochemical experiments but on expression data (Schaller and Ryan, 1996). On account of this and due to the fact that physiological roles of plant APs seem to be diverse and have not been fully investigated yet, we cannot conclude a functional role for PpAP1 from the phylogenetic data.

Taken together, our results demonstrate presence and functionality of the two signals required for secretory and vacuolar sorting within the sequence of PpAP1 as they have been described for APs of higher plants. It can therefore be assumed that both targeting mechanisms are similar in seedless plants and in seed plants and have probably evolved as early as the occurrence of land plants.

In addition to this, the results provide valuable tools for the use of *Physcomitrella* in biotechnological approaches. One prerequisite for the production of heterologous proteins is the availability of a set of targeting signals. The fusion of such signals enables the accumulation of the produced protein in distinct cell compartments or in the extracellular space and thereby facilitates their downstream processing. PpAP1 provides two of those sequences allowing secretory and vacuolar targeting.

Acknowledgements. We would like to thank Agnes Kinal for help with cloning of the PpAP1 sequence, Antje Schaaf for correction of the manuscript and Dr. V. Speth for help with the Axioplan microscope. The vector pJL517 was kindly provided by Jan Lucht. This work was funded by the German Federal Ministry of Education and Research (BMB+F).

References

Asakura, T., Watanabe, H., Abe, K., Arai, S., 1995. Rice aspartic proteinase, oryzasin, expressed during seed ripening and germination, has a gene organization distinct from those of animal and microbial aspartic proteinases. Eur. J. Biochem. 232, 77–83.

Bierfreund, N. M., Reski, R., Decker, E. L., 2003. Use of an inducible reporter gene system for the analysis of auxin distribution in the moss *Physcomitrella patens*. Plant Cell Rep. 21, 1143–1152.

Blobel, G., Walter, P., Chang, C. N., Goldman, B. M., Erickson, A. H., Lingappa, V. R., 1979. Translocation of proteins across membranes: the signal hypothesis and beyond. Symp. Soc. Exp. Biol. 33, 9–36.

Boevink, P., Oparka, K., Cruz, S. S., Martin, B., Betteridge, A., Hawes, C., 1998. Stacks on tracks: the plant Golgi apparatus traffics on an actin/ER network. Plant J. 15, 441–447.

Buchanan, B. B., Gruissem, W., Jones, R. L., 2000. Biochemistry and molecular biology of plants. John Wiley & Sons, New York.

D'Hondt, K., Stack, S., Gutteridge, S., Vandekerckhove, J., Krebbers, E., Gal, S., 1997. Aspartic proteinase genes in the Brassicaceae *Arabidopsis thaliana* and *Brassica napus*. Plant Mol. Biol. 33, 187–192.

Domingos, A., Cardoso, P. C., Xue, Z. T., Clemente, A., Brodelius, P. E., Pais, M. S., 2000. Purification, cloning and autoproteolytic processing of an aspartic proteinase from *Centaurea calcitrapa*. Eur. J. Biochem. 267, 6824 – 6831.

Faro, C., Ramalho-Santos, M., Vieira, M., Mendes, A., Simoes, I., Andrade, R., Verissimo, P., Lin, X., Tang, J., Pires, E., 1999. Cloning and characterization of cDNA encoding cardosin A, an RGD-containing plant aspartic proteinase. J. Biol. Chem. 274, 28724 – 28729.

Frazao, C., Bento, I., Costa, J., Soares, C. M., Verissimo, P., Faro, C., Pires, E., Cooper, J., Carrondo, M. A., 1999. Crystal structure of cardosin A, a glycosylated and Arg-Gly-Asp-containing aspartic proteinase from the flowers of *Cynara cardunculus* L. J. Biol. Chem. 274, 27694 – 27701.

Fujinaga, M., Chernaia, M. M., Tarasova, N. I., Mosimann, S. C., James, M. N., 1995. Crystal structure of human pepsin and its complex with pepstatin. Protein Sci. 4, 960 – 972.

Glathe, S., Kervinen, J., Nimtz, M., Li, G. H., Tobin, G. J., Copeland, T. D., Ashford, D. A., Wlodawer, A., Costa, J., 1998. Transport and activation of the vacuolar aspartic proteinase phytepsin in barley (*Hordeum vulgare* L.). J. Biol. Chem. 273, 31230 – 31236.

Hadlington, J. L., Denecke, J., 2000. Sorting of soluble proteins in the secretory pathway of plants. Curr. Opin. Plant Biol. 3, 461 – 468.

Haseloff, J., Siemering, K. R., Prasher, D. C., Hodge, S., 1997. Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. Proc. Natl. Acad. Sci. USA 94, 2122 – 2127.

Hiraiwa, N., Kondo, M., Nishimura, M., Hara-Nishimura, I., 1997. An aspartic endopeptidase is involved in the breakdown of propeptides of storage proteins in protein-storage vacuoles of plants. Eur. J. Biochem. 246, 133 – 141.

Hohe, A., Decker, E. L., Gorr, G., Schween, G., Reski, R., 2002. Tight control of growth and cell differentiation in photoautotrophically growing moss (*Physcomitrella patens*) bioreactor cultures. Plant Cell Rep. 20, 1135 – 1140.

Kervinen, J., Tobin, G. J., Costa, J., Waugh, D. S., Wlodawer, A., Zdanov, A., 1999. Crystal structure of plant aspartic proteinase prophytepsin: inactivation and vacuolar targeting. EMBO J. 18, 3947 – 3955.

Kiessling, J., Kruse, S., Rensing, S. A., Harter, K., Decker, E. L., Reski, R., 2000. Visualization of a cytoskeleton-like FtsZ network in chloroplasts. J. Cell Biol. 151, 945 – 950.

Kircher, S., Wellmer, F., Nick, P., Rugner, A., Schafer, E., Harter, K., 1999. Nuclear import of the parsley bZIP transcription factor CPRF2 is regulated by phytochrome photoreceptors. J. Cell Biol. 144, 201 – 211.

Kneen, M., Farinas, J., Li, Y., Verkman, A. S., 1998. Green fluorescent protein as a noninvasive intracellular pH indicator. Biophys. J. 74, 1591 – 1599.

Munro, S., Pelham, H. R., 1987. A C-terminal signal prevents secretion of luminal ER proteins. Cell 48, 899 – 907.

Neumann, U., Brandizzi, F., Hawes, C., 2003. Protein transport in plant cells: in and out of the Golgi. Ann. Bot. 92, 167 – 180.

Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Int. J. Neural Syst. 8, 581 – 599.

Paris, N., Stanley, C. M., Jones, R. L., Rogers, J. C., 1996. Plant cells contain two functionally distinct vacuolar compartments. Cell 85, 563 – 572.

Park, H., Yamanaka, N., Mikkonen, A., Kusakabe, I., Kobayashi, H., 2000. Purification and characterization of aspartic proteinase from sunflower seeds. Biosci. Biotechnol. Biochem. 64, 931 – 939.

Richter, U., Kiessling, J., Hedtke, B., Decker, E., Reski, R., Borner, T., Weihe, A., 2002. Two RpoT genes of *Physcomitrella patens* encode phage-type RNA polymerases with dual targeting to mitochondria and plastids. Gene 290, 95 – 105.

Rother S., Hadeler B., Orsini M., Abel W. O., Reski R., 1994. Fate of a mutant macrochloroplast in somatic hybrids. J. Plant Physiol. 143, 72 – 77.

Runeberg-Roos, P., Kervinen, J., Kovaleva, V., Raikhel, N. V., Gal, S., 1994. The aspartic proteinase of barley is a vacuolar enzyme that processes probarley lectin in vitro. Plant Physiol. 105, 321 – 329.

Runeberg-Roos, P., Tormakangas, K., Ostman, A., 1991. Primary structure of a barley-grain aspartic proteinase. A plant aspartic proteinase resembling mammalian cathepsin D. Eur. J. Biochem. 202, 1021 – 1027.

Sambrook, J., Fritsch, E. F., Maniatis, T., Russell, D. W., 2001. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schaller, A., Ryan, C. A., 1996. Molecular cloning of a tomato leaf cDNA encoding an aspartic protease, a systemic wound response protein. Plant Mol. Biol. 31, 1073 – 1077.

Silva, A. M., Cachau, R. E., Sham, H. L., Erickson, J. W., 1996. Inhibition and catalytic mechanism of HIV-1 aspartic protease. J. Mol. Biol. 255, 321 – 346.

Swanson, S. J., Bethke, P. C., Jones, R. L., 1998. Barley aleurone cells contain two types of vacuoles: characterization of lytic organelles by use of fluorescent probes. Plant Cell 10, 685 – 698.

Thompson, J. D., Higgins, D. G., Gibson, T. J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673 – 4680.

Tormakangas, K., Hadlington, J. L., Pimpl, P., Hillmer, S., Brandizzi, F., Teeri, T. H., Denecke, J., 2001. A vacuolar sorting domain may also influence the way in which proteins leave the endoplasmic reticulum. Plant Cell 13, 2021 – 2032.

Van de Peer, Y., De Wachter, R., 1997. Construction of evolutionary distance trees with TREECON for Windows: accounting for variation in nucleotide substitution rate among sites. Comput. Appl. Biosci. 13, 227 – 230.

Voigt, G., Biehla, B., Heinrichs, H., Voigt, G., 1997. Aspartic proteinase levels in seeds of different angiosperms. Phytochemistry 44, 389 – 392.

Walter, P., Blobel, G., 1981. Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. J. Cell Biol. 91, 557 – 561.

White, P. C., Cordeiro, M. C., Arnold, D., Brodelius, P. E., Kay, J., 1999. Processing, activity, and inhibition of recombinant cyprosin, an aspartic proteinase from cardoon (*Cynara cardunculus*). J. Biol. Chem. 274, 16685 – 16693.