REVIEW ARTICLE

Structure and function of plant aspartic proteinases

Isaura Simões and Carlos Faro

Departamento de Biologia Molecular e Biotecnologia, Centro de Neurociências e Biologia Celular, Universidade de Coimbra and Departamento de Bioquímica, Faculdade de Ciências e Tecnologia, Universidade de Coimbra, Portugal

Aspartic proteinases of the A1 family are widely distributed among plant species and have been purified from a variety of tissues. They are most active at acidic pH, are specifically inhibited by pepstatin A and contain two aspartic residues indispensible for catalytic activity. The three-dimensional structure of two plant aspartic proteinases has been determined, sharing significant structural similarity with other known structures of mammalian aspartic proteinases. With a few exceptions, the majority of plant aspartic proteinases identified so far are synthesized with a prepro-domain and subsequently converted to mature two-chain enzymes. A characteristic feature of the majority of plant aspartic proteinase precursors is the presence of an extra protein domain of about 100 amino acids known as the plant-specific insert, which is highly similar both in sequence and structure to saposin-like proteins. This insert is usually removed during processing and is absent from the mature form of the enzyme. Its functions are still unclear but a role in the vacuolar targeting of the precursors has been proposed. The biological role of plant aspartic proteinases is also not completely established. Nevertheless, their involvement in protein processing or degradation under different conditions and in different stages of plant development suggests some functional specialization. Based on the recent findings on the diversity of A1 family members in *Arabidopsis thaliana*, new questions concerning novel structure–function relationships among plant aspartic proteinases are now starting to be addressed.

Keywords: aspartic proteinases; cardosin; phytepsin; programmed cell death; stress response.

Introduction

Aspartic proteinases (APs; EC 3.4.23) have been extensively studied and characterized and are widely distributed among vertebrates, plants, yeast, nematodes, parasites, fungi and viruses [1,2]. AP activity has also been detected in recombinant proteins from bacterial origin [3]. According to the MEROPS database (http://www.merops.ac.uk), created by Rawlings & Barrett [4], APs are now grouped into 14 different families, on the basis of their amino acid sequence homology, which in turn are assembled into six different clans based on their evolutionary relationship and tertiary structure. Plant APs have been distributed among families A1, A3, A11 and A12 of clan AA, and family A22 of clan AD. The majority of plant APs belongs to the A1 family, together with pepsin-like enzymes from many different origins.

In common with other members of the A1 family, plant APs are active at acidic pH, are specifically inhibited by pepstatin and have two aspartic acid residues responsible for the catalytic activity [2,5]. However, there are several structural and functional features that make plant APs unique among aspartic proteinases. These aspects will be highlighted throughout the present review article which aims to provide an overview of the current knowledge about plant aspartic proteinases in terms of their structure, processing, inactivation, localization, proposed biological functions and genomic diversity.

Primary structure organization

The majority of plant APs identified so far are synthesized as single-chain preproenzymes and subsequently converted to mature enzymes that can be either single- or two-chain enzymes. The cDNA derived amino acid sequences of several plant APs revealed that the primary structures of their precursors are quite similar [6–15]. These precursors are characterized by the presence of a hydrophobic N-terminal signal sequence, responsible for translocation into the ER, followed by a prosegment of about 40 amino acids, and a N-terminal domain and a C-terminal domain separated by an insertion comprising approximately 100 amino acids, named as plant-specific insert (PSI) (Fig. 1).

While the prosegment is present in all APs and is involved either in the inactivation or in the correct folding, stability and intracellular sorting of several zymogens [16], the PSI is an insertion only identified in plant APs, which is highly similar to saposins and saposin-like proteins and whose biological function has not been completely established [8,13,17–21].

Correspondence to C. Faro, Departamento de Bioquímica, Universidade de Coimbra, Apt. 3126, 3000 Coimbra, Portugal.
Fax: + 351 239 480208, Tel.: + 351 239 480210,
E-mail: cfaro@imagem.ibili.uc.pt

```
                Signal peptide          Prosegment
Cardosin A  MGTSIKANVLALFLFYLLSPT-VFSVSD-IGLIRIGLKKRKVDRIDQLRGRRAIMEGNAR  58
Cardosin B  MGTPIKASLLALFLFFLLSPT-AFSVSN-GGLLRVGLKKRKVDRLDQLRAHGVHMLGNAR  58
Phytepsin   MGT--RGLALALLAAVLLQTVLPAASEAEGLVRIALKKRPIDRNSRVATGLSGGEEQPL   58
AtAsp1      MKIYSRTVAVSLIVSFLLCFS-AFAERN-DGTFRVGLKKLKLDSKNRLAASRVESKQEKPL 58
AtAsp2      MGVYSRAVAFSVFVSFLLFFT-AYSKRN-DGTFRVGLKKLKLDPNNRLATRFGSKQEEAL  58
AtAsp3      MGTRFQSFLLVFLLSCLLISTASCERNGDGTIRIGLKKKHRDRSNRLASCLPLKNRGSH   60
Chlapsin    ---MARSYVFALIALAAVEALLGVAAEQQAGMLRVTLRKTEMLTTLGRFRFYLLGEQGLL  57

Cardosin A  K----DFGFRGTVRDSGSAVVALTNDRDTSYFGEIGIGTPPQKFTVIFDTGSSVLWVPSS 114
Cardosin B  K----DFGFRRTLSDSGSGIVALTNDRDTAYYGEIGIGTPPQNFAVIFDTGSDLWVPST  114
Phytepsin   L----SGANPLR-SEEEGDIVALKNYMNAQYFGEIGVGTPPQKFTVIFDTGSSNLWVPSA 111
AtAsp1      R----AYR--LG-DSGDADVVVLKNYLDAQYYGEIAIGTPPQKFTVIFDTGSSNLWVPSS 111
AtAsp2      RSSLRSYNNNLGGDSGDADIVFLKNYLDAQYGDITIGTPPQKFTVVFDTGSSNLWVPSG  116
AtAsp3      W----SPKHYFRLNDENADMVPLKNYLDAQYYGDITIGTPPQKFTVIFDTGSSNLWIPST 116
Chlapsin    G------------SSDQGQVTLKNMDAQYYGEIGLGTPPQLFNVIFDTGSANLWVPSS   104

Cardosin A  KCINS-KACRAHSMYESSDSSTYKENGTPGAIIYGTGSITGFFSQDSVTIGDLVVKEQDF 173
Cardosin B  KCDTS-LACVIHPRYDSGDSSTYKNNGPAAIQYGTGSIAGYFSEDSVEVGDLVVKQEF   172
Phytepsin   KCYFS-IACYLHSRYKAGASSTYKKNGQKPAAIQYGTGSIAGYFSEDSVTIGDLVVKDQEF 172
AtAsp1      KCYFS-LACLLHPKYKSSRSSTYKENGKAAAIHYGTGAIAGFFSNDAVTVGDLVVKDQEF 172
AtAsp2      KCFFS-LSCYFHAKYKSSRSSTYKKSGKRAAIHYGSGSISGFFSYDAVTVGDLVVKDQEF 177
AtAsp3      KCYLS-VACYFHSKYKASQSSSYRKNGKPASIRYGTGAISGYFSNDDVKVGDIVVKEQQF 175
Chlapsin    KCALFNIACRLHRKYNAAKSKTYKANGTEFAIEYGTGSLDGYISQDVLTWGGLTIKDQGF 164

Cardosin A  IEATDEADNVFLHRLFDGILGLSFQTIS----VPVWYNMLNQGLVKERRPSFWLNRNVDE 229
Cardosin B  IETTEEDDTVFLKSEFDGILGLGFQEISAGKAVPVWYNMVNQGLVEEAVFSFWLNRNVDE 231
Phytepsin   IEATKEPGITFLVAKFDGILGLGFKEISVGKAVPVWYKMIEQGLVSDPVFSFWLNRHVDE 232
AtAsp1      IEATKEPGITFVVAKFDGILGLGFQEISVGKAAPVWYNMLKQGLIKEPVFSFWLNRNADE 231
AtAsp2      IETTSSEPGLTFLVAKFDGLLGLGFQEIAVGNATPVWYNMLKQGLIKRPVFSFWLNRPFKS 237
AtAsp3      IEATSEPGITFLLAKFDGILGLGFQEISVGNSTPVWYNMVEKGLVKEPIFSFWLNRNPKD 237
Chlapsin    AEAINEPGLTFVAAKFDGILGMGFPAISVQHVPPPFTRLVEEGGLAAPVFSFWLNRDPNA 224

Cardosin A  EEGGELVFGGLDPNHFRGDHTYVPVTYQYYMQFGIGDVLIGDKSTGFCAPGCQAFDSGT  289
Cardosin B  EEGGELVFGGVDPNHFRGNHTYVPVTRKGYMQFEMGDVLIGDKSSGFCAGGCAAIDSGT  292
Phytepsin   GEGGEIIFGGMDPKHYVGEHTYVPVTQKGYMQFDMGDVLVGGKSTGFCAGGCAAIDSGT  292
AtAsp1      EEGGELVFGGVDPKHFKGEHTYVPVTQKGYWQFDMGDVLIGGAPTGFCESGCSAIDSGT  291
AtAsp2      EEGGELVFGGVDPNHFKGEHTFVPVTHKGYWQFDMGDVLIGGKPTGFCAAGCSAIDSGT  297
AtAsp3      PEGGEIVFGGVDPNHFKGEHITFVPVTHKGYWQFLMGDLQLAGKPTGFCAPGCQAIDSGT 297
Chlapsin    PNGGELVLGGIDPTHFTGEHTWVFVTRQGYWQFTMEGLDLGPGSQKMCAKGCAAIDTGT  284

Cardosin A  SLLSGPTAIVTQINHAIGANGVMNQQCKTVVSKYGRDIIEMLRSKIQPDKICSHMKLCTF 349
Cardosin B  SFFAGPTAIIQINQAIGAKGVLNQQCKTLVGQYGHMNIQMLTSEVQFDKICSHMKLCTF  353
Phytepsin   SLLAGPTAIIEINEKIGAAGVVSQQCKTIVSQYGQQILDLLLSETQFPKKICSQVGLCTF 350
AtAsp1      SLLAGPTTIITMINHAIGAAGVVSQQCKTVVDQYGQTILDLLLSETQFPKKICSQIGLCTF 350
AtAsp2      SLLAGPTAVVAMINKAIGASGVVSDQYGQITLDLLLAETQFPKKICSQIGLCAY      354
AtAsp3      SLLTGPSTVITMINHAIGAQGIVSRECKAVVDQYGKTMLNSLLAQEDPKKVCSQIGVCAY 355
Chlapsin    SLIAGPSDEVAALNHAIGATSALSAQCRQLVRDYLPQIIAQLHD-LPLDQVCASIGLCFM 343

Cardosin A  DGAR----------------------------------------------------    353
Cardosin B  DGAH----------------------------------------------------    357
Phytepsin   DGTR----------------------------------------------------    356
AtAsp1      DGTR----------------------------------------------------    354
AtAsp2      DGTH----------------------------------------------------    361
AtAsp3      DGTQ----------------------------------------------------    355
Chlapsin    AAABIIHPABRLLATTTAAGTHSIRTSSGAAAVADEAAAGDASDVDAAVAAVVKAQLANLL 403

ACardosin   -------------DVSSIIESVVDKONNDKSSGGIHD-------EMCTFCEMAVVW    388
BCardosin   -------------DVRSMIESVVDKNNDKSSGG--------EICTFCEMALVR      389
Phytepsin   -------------GVSAGIRSVVDKDEPVKSNGLRAD-------PMCSACEMAVVW    391
AtAsp1      -------------GVSMGIESVVDKENAKLSNGVGD--------AACSACEMAVVW    389
AtAsp2      -------------GVSMGIVSVVDKENTRSSSGLAD--------AGCPACEMAVVW    396
AtAsp3      -------------SVSMGIQSVVDDG---TSGLLNQ--------AMCSACEMAAVW    391
Chlapsin    GHARAAGAITTNGRGAAASDGGVISKLVGEAAAAKAQGSKAESAGGSVVCSFCQTAVAY  463

Cardosin A  MQNEIKQSETEDNIINYANELCEHLLTSEEELQVDCNTLSSMPNVSFTIGGKKFGLTFEQ 449
Cardosin B  MQNEIKRNETEDNIINHVNEVCDQLFTSSAAESIVDCNGISSMPNIAFTIGSKLFEVTFEQ 449
Phytepsin   MQNLAQNMTQDLLILDYVNQLCNRLLPSFMGESAVDCGSLGSMPDIEFTIGGKKFALKFEB 449
AtAsp1      IQSQLPQNMTQERILNYVNHELCERLLPSFMGESAVDCAQLSTMPTVSLTIGGKVFDLAFEB 449
AtAsp2      IQSQLSQNMTQERIVNYINEICERMPSFMGESAVDCSQLSKMFTVSFTIGGKVFDLAFEB 456
AtAsp3      MESELTQMQTQERILAYAAELCDHIPTQNQQSAVDCGRVSSMFIVTFSIGGRSFDLTPQD 451
Chlapsin    IKIALQSNSTIEQIADAVGGLCDQVSFG-GPSVVDCDKISTLFVISFNIGGRVFFLRPEQ 522

Cardosin A  YILKVGHGE-ATQCISGFTAMDAT-LLGPLWILGDVFMRFYHTVYDYGKLGVGFAEAA   504
Cardosin B  YIYKVGEGE-AATCISGFTALDIMSPQGPIWILGIMFMGPYHTVFDYGKLRVGFAEAV   506
Phytepsin   YILKVGEGSA-AAQCISGFTAMDIPPRGPLWILGDVFMGFYHTVFDYGKLRIGFAEAA   508
AtAsp1      YVLKVGEGP-VAQCISGFIALLVVAPPRGPLWILGDVFMGKYHTVFDFGNEQVGFAEAA  506
AtAsp2      YVLKIGEGP-VAQCISGFTALDIPPPRGPLWILGDVFMGKYHTVVFDVGNEQVGFAEAV  513
AtAsp3      YIFKIGEGV-ESQCTSGFTAMDIAPPRGPLWILGDIFMGPYHTVFDYGKGRVGFANAA  508
Chlapsin    YVLQLDAGGGEMOCISGFMGLLDVP--AGPLWILGDIFLGAYHTVFDYGAARLGFANAA 570
```
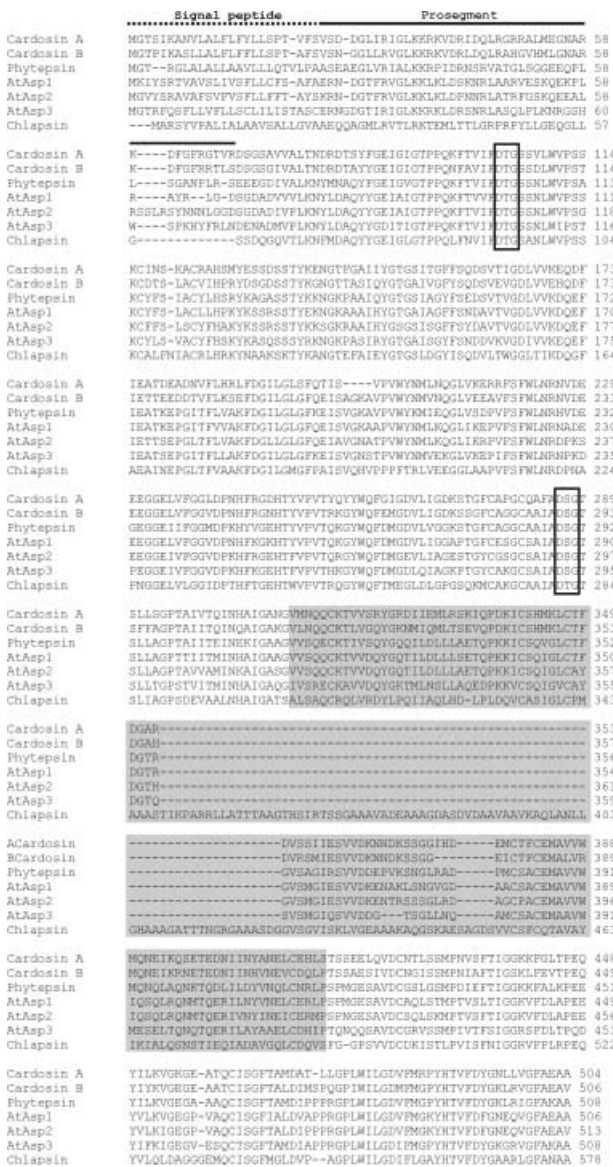
Fig. 1. Plant aspartic proteinase precursors. Comparison of the amino acid sequences of representative members of the A1 family of plant aspartic proteinases. The regions corresponding to the signal peptide (dotted line), the prosegment (solid line) and the plant specific insert (shaded grey) are highlighted. The catalytic aspartic acid residues are boxed. Cardosin A and cardosin B were purified from *C. cardunculus* L. (accession numbers: AJ132884 and AJ237674, respectively – EBI Data Bank); phytepsin was purified from barley (*H. vulgare*) (accession number: X56136); AtAsp1, AtAsp2 and AtAsp3 are *A. thaliana* aspartic proteinases (accession numbers: U51036, AY070453 and AF076243, respectively); chlapsin was purified from *Chlamydomonas reinhardtii* (accession number: AJ579366).

To date, the only exceptions to this primary structure organization are nucellin, specifically expressed in barley ovule nucellar cells [22], an AP-like protein from tobacco chloroplasts [23] and an AP encoded by the *cdr-1* gene involved in disease resistance [24].

In general, plant APs share high amino acid sequence similarity in their N- and C-terminal domains (over 60% identity), and about 45% identity with cathepsin D, the closest AP of nonplant origin. The two catalytic sequence motifs are Asp-Thr-Gly (DTG) and Asp-Ser-Gly (DSG) in all plant APs belonging to the A1 family with the exception of chlapsin, the AP from *Chlamydomonas reinhardtii* that contains DTG/DTG (accession number: AJ579366) (C. M. Almeida and C. Faro, unpublished results). In some APs from fungi and from protozoa, the catalytic Asp residues also occur within the DTG/DSG motifs (http://www.merops.ac.uk). The evolutionary or biological significance of this variation observed in APs of different kingdoms has not been established.

Three-dimensional structure

The three-dimensional structures of several members of the A1 family have been determined and they share significant structural similarity [2]. Regarding plant APs, only two crystal structures have been determined – mature cardosin A (PDB code: 1B5F) [17] and prophytepsin, the precursor form of barley AP containing the prosegment and the PSI (PDB code: 1QDM) [25] (Fig. 2). Both APs are two-chain polypeptides in their mature forms and present a very similar fold to what was found for other APs. The overall secondary structure consists essentially of β-strands with very little α-helix. The molecules are bilobal with the active site located in a large cleft between the two similar β-barrel-like domains, each contributing one of the catalytic sequence motifs (DTG/DSG). The catalytic aspartic residues are located at the base of this large cleft. Three conserved disulfide bridges stabilize the structure and both polypeptide chains are held together by hydrophobic interactions and hydrogen bonds. As in the other AP structures, there is a flexible region known as the flap which projects out over the cleft and encloses substrates and inhibitors in the active site [5].

Besides the common pepsin-like topology for the main body of mature phytepsin, the structural characterization of the enzyme precursor also gave new insights about the prosegment and the PSI [25]. Although part of the prosegment was not traced due to a disordered structure, it was shown that its N-terminal part is involved in the formation of the six-stranded β-sheet, while the helical portion of this prosegment approaches the active site and partially covers it. As will be discussed below, the authors propose an inactivation mechanism based on the interactions found between the prosegment and the active site.

The PSI forms an independent subunit in the prophytepsin structure that is inserted into the C-terminal domain. Structurally, the PSI comprises five amphipathic α-helices folded into a compact globular domain and linked with each other by three disulfide bridges. A quite similar structure was described for NK-lysin, which is also a saposin-like protein [26].

Processing and inactivation mechanism

Plant AP precursors undergo several proteolytic cleavages to produce mature single-chain or two-chain form of the enzymes. Proteolytic processing of plant APs starts with removal of the signal sequence upon translocation to the ER lumen. The following conversion steps include cleavage of the prosegment and total or partial removal of the internal
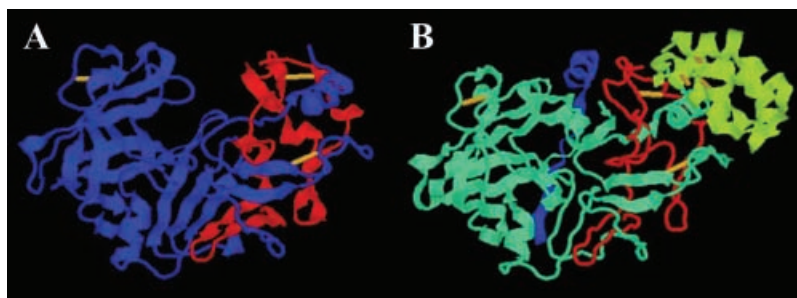
Fig. 2. Ribbon representation of the crystal structures of cardosin A (A) and prophytepsin (B). (A) Structure of mature cardosin A from *C. cardunculus* L. (PDB code: 1B5F) [17]. The heavy chain is shown in blue, the light chain in red and disulfide bridges in yellow. (B) Structure of prophytepsin from *H. vulgare* L. (PDB code: 1QDM) [25]. The propeptide is shown in blue, the mature protein is shown in cyan (heavy chain) and red (light chain), the plant specific insert (PSI) in green and disulfide bridges in yellow. Prepared with the program PROTEIN EXPLORER http://www.proteinexplorer.org.

PSI. Proteolytic removal of the prosegment is an important step in generation of active protease from inactive zymogen [1]. Zymogen conversion generally occurs by limited proteolysis and removal of the 'activation segment'. It may involve accessory molecules that trigger activation or the process may be autocatalytic requiring only a drop in pH [27] as is described for the gastric APs [28].

In general, processing of plant aspartic proteinase precursors involves removal of the prosegment and the PSI domain [18,20,21,29–33]. Nevertheless there are some variations on the mechanism and order by which each segment is removed from the precursor.

Procardosin A, the precursor of cardosin A, undergoes proteolytic processing as the flower matures and during this process the PSI is totally removed, probably by an aspartic proteinase, before the prosegment. Its conversion into an active form is likely to occur inside the vacuoles where the protein is accumulated [20]. Processing by a similar autocatalytic mechanism has also been proposed for cenprosin, the AP from *Centaurea calcitrapa* [30] and for recombinant oryzasin 1, the rice AP [29].

A slightly different picture has emerged for prophytepsin. Using metabolic labeling and immunoprecipitation it was shown that prophytepsin in barley roots is sequentially processed into two different two-chain forms by cleavage of the prosegment and partial removal of the PSI (and not completely like in procardosin A) [18]. Although it was not clearly established which is removed first, whether the prosegment or the PSI, a recent paper proposed a model in which the prosegment is removed prior to the PSI [33]. As the intermediate forms and final products obtained *in vitro* are slightly different from those detected *in vivo*, it was suggested that complete maturation of the protein probably requires the presence of other proteinases/exopeptidases besides the autoactivation mechanism [18].

The activation of recombinant cyprosin produced in *Pichia pastoris* has given us a third processing scheme. Like prophytepsin, the precursor form of cyprosin was processed in different isoforms by the excision of the prosegment and of most of the PSI [21]. Conversely to what has been found *in vivo* [31], heavy and light chains of the processed forms of recombinant cyprosin are held together by disulfide bonds. It has been suggested that this different processing is caused by the action of host cell proteinases and not by auto-

activation [21]. A similar processing mechanism has been suggested for the sunflower seed AP. The precursor is sequentially cleaved into different intermediate forms, whose chains remain associated by disulfide bridges. However, and in contrast to recombinant cyprosin, the PSI is finally removed to completion in order to generate the mature form of the sunflower AP in which the chains are no longer held together by disulfide bridges [32].

In any case, processing of plant AP precursors leads ultimately to the formation of a two-chain enzyme, without the prosegment and the PSI domain, with a domain organization similar to that of mammalian or microbial APs.

An inactivation mechanism for plant APs has been proposed by Kervinen *et al.* based on the three-dimensional structure of phytepsin precursor [25]. The inactivation mechanism proposed for prophytepsin resembles the mechanism accepted for mammalian gastric APs zymogens, progastricsin and pepsinogen, with a preformed active site blocked by the prosegment [34,35]. In prophytepsin the active site is blocked not only by the prosegment, but also by the 13 residues of the N-terminal of the mature enzyme and by the 'flap'. The anchorage of the prosegment and of part of the N-terminus in the active site cleft is made by ionic interactions established between Lys11/Tyr13 of the mature enzyme sequence and the catalytic aspartic acids at the bottom of the cleft. In fact, these two residues replace the characteristic Lys36p/Tyr37p (where p stands for prosegment) found in mammalian APs zymogens and known to be responsible for the ionic interactions with the Asp residues of the active site.

Most plant APs contains a Lys/Tyr sequence in a position equivalent to Lys11/Tyr13 of prophytepsin suggesting a similar inactivation mechanism. However, cardosin A, cardosin B and two rice APs do not contain this sequence either in the prosegment or in the N-terminus of the mature enzyme. Biochemical studies with recombinant precursors of cardosins revealed that, conversely to other zymogens, procardosins are active (M. Vieira & C. Faro, unpublished results). These evidences suggest that procardosins probably do not share the inactivation mechanism described above. Most likely, the interactions between the prosegment and the active site render the prosegment more flexible and enable the substrate to enter the catalytic cleft. Nevertheless, only the structural characterization of procardosins and

other precursors will give new clues about the different modes of inactivation in plant APs.

The plant-specific insert

Except for the barley nucellin [22], an AP-like protein from tobacco chloroplasts [23] and the product of *cdr-1* gene from *Arabidopsis* [24], all plant APs identified so far are characterized by the presence of an extra protein domain of approximately 100 amino acids known as the plant specific insert (PSI). This segment, inserted into the C-terminal domain of the plant APs precursors, is usually removed during the proteolytic maturation of the proteinases. The PSI sequence shows no homology with mammalian or microbial APs, but is highly similar to that of saposin-like proteins (SAPLIPs) [36]. This protein family includes saposins, which are lysosomal sphingolipid-activator proteins [37], NK-lysin, granulysin, surfactant protein B, amoebapores and domains of acid sphingomyelinase and acyloxyacyl hydrolase [38–40]. Like other members of this family, the PSI contains six conserved cysteines, several hydrophobic residues and a consensus glycosylation site. In the particular case of *Chlamydomonas reinhardtii* AP, and besides these common features, the PSI domain comprises an extra region of approximately 80 amino acids rich in alanine triplets whose function is still unknown (C. M. Almeida & C. Faro, unpublished results) (Fig. 1).

The structural characterization of prophytepsin's PSI revealed the same 'saposin fold' [25] as first determined for NK-lysin [26] and recently for granulysin [41]. In fact, the proteins belonging to this SAPLIPs family all share a closely related compact globular structure comprising five amphipathic α-helices linked with each other by three disulfide bridges. A unique feature of the PSI is the swap of the N- and C-terminal portions of the saposin-like domain, where the C-terminal portion of one saposin is linked to the N-terminal portion of the other saposin. Hence, the PSI is not a true saposin but a swaposin [25,38,42] (Fig. 3).

The functions of the PSI are still unclear, however, an important role in vacuolar targeting of plant AP precursors has been proposed. Besides its possible direct interaction with lipid bilayers, as described for other SAPLIPs members [38], the structural characterization of phytepsin PSI revealed a putative membrane-binding region comprising the PSI and an adjacent area of the mature enzyme [25]. Thus, the authors suggest that this saposin-like domain in plant APs may be responsible for bringing AP precursors into contact with membranes or membrane-bound receptor proteins mediating the sorting of enzyme precursors during Golgi-mediated intracellular transport to the vacuoles. In fact, the role of the PSI in protein sorting to vacuoles has also been demonstrated in transient expression experiments in tobacco protoplasts [33] where it was shown that deletion of the PSI from phytepsin results in secretion of the truncated phytepsin, whereas the wild-type phytepsin still accumulates inside the vacuoles. In addition to this role of the PSI as a vacuolar sorting signal it is also suggested that this domain may have a strong influence on how phytepsin leaves the ER, implying that the vacuolar sorting may not be restricted to the Golgi apparatus but can start as early as the ER [33]. The proposed role of the PSI in the targeting of plant APs to the vacuole resembles what has been described for mammalian saposin C and cathepsin D. It has been suggested that the association of saposin C with cathepsin D may be responsible for the mannose-6-phosphate independent targeting of the latter to the lysosome [43,44]. An important difference between both targeting mechanisms is that in plants, APs and the PSI sorting domain are encoded in the same precursor molecule, whereas in mammalian cells different genes encode cathepsin D and saposin C. However, and similarly to what has been described for saposin C [38], intracellular protein targeting may not be the only function of the PSI. In fact, Egas *et al.* demonstrated that besides its ability to interact with membranes, the PSI of cardosin A is a potent inducer of vesicle leakage [45]. The results described either with procardosin A or with recombinant PSI support the idea that plant AP precursors are bifunctional molecules containing a membrane-destabilizing domain in addition to their protease domain. Thus, the authors suggest that the PSI may take part in defensive mechanisms against pathogens and/or as an effector of cell death. Based on these results it was also suggested that the PSI from carnivorous plants may contribute to prey digestion by destroying prey cell membranes [6].

Distribution and localization

Plant APs are widely distributed in the plant kingdom and have been detected or purified from monocotyledonous and dicotyledonous species as well as gymnosperms. Recently, the cDNA of an AP was cloned from *Chlamydomonas*

Fig. 3. The 'saposin fold'. (A) Ribbon representation of the structure of NK-lysin, a saposin-like protein [26]. The N-terminal domain is shown in blue and C-terminal domain in red. (B) Ribbon representation of the structure of the PSI domain of barley prophytepsin [25] (N-terminal domain, blue; C-terminal domain, red). (C) Model structure of the PSI domain of cardosin A based on the crystal structure of prophytepsin PSI (N-terminal domain, blue; C-terminal domain, red). Prepared with the program PROTEIN EXPLORER http://www.proteinexplorer.org.

reinhardtii indicating therefore that the A1 family of AP is also represented in the unicellular green algae which are the closest ancestral precursors of vascular plants (C. M. Almeida & C. Faro, unpublished results).

In gymnosperms, AP activity has been detected in the seeds of two pine species [46], whereas in angiosperms APs have been detected or purified in monocotyledonous plants such as barley, rice, wheat, sorghum and maize [7,47–54] and in dicotyledonous plants like cucumber, squash, figleaf gourd, castor bean, sunflower, cacao, *Arabidopsis*, *Brassica*, spinach, potato, tobacco, tomato, cardoon, *Centaurea calcitrapa* and carnivorous plants such as *Nepenthes* [12,30,31,55–69].

Plant APs are either single-chain (cucumber, squash, spinach, potato, sorghum, *Brassica*, rice, wheat, tomato and tobacco) or two-chain (barley, figleaf gourd, castor bean, sunflower, cacao, *Centaurea*, cardoon, *Arabidopsis* and maize) enzymes. However, it has not been established what determines the additional processing step of converting a single-chain inactive enzyme into a two-chain active form. Some authors suggest that these processing differences may be caused by the presence or absence of protein-processing enzymes responsible for the conversion because, in terms of primary structure organization, plant APs precursors are, in general, very similar.

Like for monocotyledonous plants, AP expression or activity in some dicotyledonous plants has been detected in other tissues besides those where the protein was first purified [6,8,10,14,70–75]. However, tissue-specific localization has been described for some plant APs and revealed that these enzymes are not randomly distributed throughout the organs. Moreover, it is now clear that some plant species have multiple genes for APs. In fact, the differential expression observed for these AP homologs in *Cynara cardunculus* L., *Arabidopsis*, barley and *Nepenthes* clearly suggests some functional specialization and imply the potential involvement of the different APs in a wide variety of cellular processes [6,8,15,22,54,76,77].

In barley, two independent studies demonstrated that in developing grains and during seed germination the localization of the AP (phytepsin) was very specific. Immunohistochemical studies in barley roots have also revealed that phytepsin is specifically expressed in developing tracheary elements and sieve cells [77]. Castor bean AP was localized in the endosperm of maturing seeds [56] and in *Nepenthes alata*, transcripts of two of the five AP homologues were detected, by *in situ* hybridization, in the digestive glands of the pitchers, the trapping organs of the plant [6]. Using immunohistochemistry and immunogold transmission EM, APs purified from the flowers of the cardoon *Cynara cardunculus* L. have been specifically localized in the floral transmitting tissue (cardosin B) [15], in the stigmatic papillae (cardosin A) [76] and in the epidermal cells of the style (cardosin A and cyprosins) [76,78]. In a recently published report, Chen *et al.* demonstrated, by *in situ* hybridization studies, the differential expression of the three typical aspartic proteinases of *Arabidopsis* [8] and confirmed previously published results on the AP localization in seed tissues [79]. In the recently published paper, the authors showed that transcripts of these three APs are detected in all seed cell types, in the outer cell layers of the anthers early in flower development and in the guard cells of the sepals. The

mRNA of one of the APs (AtPaspA2) was also weakly detected in the transmitting tract of the flowers [8].

The great majority of the purified plant APs are intracellular, and subcellular localization studies revealed that they accumulate essentially inside protein storage vacuoles. Biochemical and immunocytochemistry analysis of barley roots and leaves showed that phytepsin was localized to the vacuoles of these cells [80] and, in a different study, phytepsin was also shown to accumulate in protein bodies and large vacuoles of barley seeds [81]. The same vacuolar localization was found for the APs present in the seeds of castor bean [56], buckwheat [72] and *Arabidopsis* [79]. Cardosin A, one of the APs purified from the flowers of *C. cardunculus* L. also accumulates in protein storage vacuoles in the stigmatic papillae [76].

The exceptions to this intracellular location are the secreted APs found in the extracellular matrix of tobacco [64] and tomato leaves [63], cardosin B found in the extracellular matrix of the floral transmitting tissue in *C. cardunculus* L. [15], the APs from *Nepenthes* that are secreted into the pitchers [66] and the AP encoded by the *Arabidopsis cdr-1* gene [24]. The AP purified from maize pollen is believed to be in the cell wall [51] and, surprisingly, the AP from spinach has been localized to the plastids [62].

Biological functions

Plant APs have been detected and purified from many different plant species. However, their biological functions are not as well assigned or characterized as those of their mammalian, microbial or viral counterparts that were shown to perform many different and diverse functions, including specific protein processing (e.g. rennin, cathepsin D and yapsins), protein degradation (e.g. gastric enzymes such as chymosin, pepsin and gastricsin) or viral polyprotein processing (human immunodeficiency virus AP) [1,5,19]. For the great majority of plant APs no definitive role has been assigned and the biological functions are still hypothetical. Actually, much of our knowledge about plant AP functions arises from colocalization studies with putative protein substrates, experimental evidences for the processing or degradation of those substrates *in vitro* and/or specific expression in certain tissues or under specific conditions. In general, plant APs have been implicated in protein processing and/or degradation in different plant organs, as well as in plant senescence, stress responses, programmed cell death and reproduction.

Protein processing and/or degradation as nitrogen source

In citrus leaf extracts, an AP has been implicated in the proteolysis of the photosynthetic enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase which plays a significant role as a nitrogen source during the growth of new organs [70]. In carnivorous plants like *Nepenthes* or *Drosera*, APs secreted into the pitchers may participate in the degradation of insect proteins suggesting that these plants may use insect proteins as nitrogen sources [6,66]. Participation of plant APs in storage protein degradation during the mobilization of reserve proteins in seed germination has been proposed for rice and wheat. In rice seeds it was

proposed that the AP could be involved in the hydrolysis of γ-globulin during the initial stage of germination because both proteins are similarly distributed in the seeds [50]. A similar role was suggested for wheat seeds AP based on its ability to hydrolyse, *in vitro*, the main wheat storage protein, gliadin. Localization of both proteins in the wheat seed endosperm also favored this hypothesis [48].

The involvement of plant APs in proteolytic processing and maturation of storage proteins has also been proposed. The AP present in castor bean seeds was able to partially process the propeptide of pro2S albumin *in vitro*. Because both proteins colocalized in the matrix of protein storage vacuoles it was suggested that castor bean AP may be involved in the maturation process of proprotein precursors in concert with the vacuolar processing enzyme [56]. Similar results were also obtained with AP from *Brassica napus* seeds [82]. In the case of phytepsin from barley seeds and the AP purified from *A. thaliana* seeds, it was shown that both enzymes partially processed probarley lectin *in vitro*, suggesting also the concerted participation of these plant APs and other proteases in protein-storage processing mechanisms [57,80].

Stress responses

In tobacco and tomato leaves, an extracellular AP has been implicated in the degradation of pathogenesis-related (PR) proteins. These proteins are synthesized and accumulated in the intercellular spaces as a response of plants to different biotic or abiotic stress situations, and APs may have a role in a conserved mechanism for PR-protein turnover, preventing overaccumulation and thereby regulating the biological functions of these stress induced proteins. These APs were shown to be constitutively expressed either in healthy or infected leaves but their functions in uninfected tissues remain unknown [63,64].

Induction of AP gene expression has also been detected in tomato leaves by wounding or treatments with systemin and methyl jasmonate. It has been suggested that this AP may be involved in intracellular protein turnover to increase amino acid pools for the synthesis of specific defense-related proteins and/or in the defense against pathogens by hydrolyzing proteins secreted by the invasive pathogens [14].

In drought-susceptible common bean cultivars subjected to abiotic stress conditions of water deficit, the expression of an AP gene was also shown to be transcriptionally upregulated and AP activity was significantly increased, as well. This enhanced AP activity may indicate the involvement of the enzyme in nitrogen remobilization for other parts of the plant [74]. In potato, AP activity was shown to be induced by

wounding both in leaves and tubers and by aeration only in the latter. However, the purified enzymes had different properties suggesting that they may have different physiological roles. For the AP from leaves, a role in defense responses of potato plants against pathogens or insects has been suggested. The inhibition of the potato tuber AP by a PR-protein may also suggest its involvement in plant stress responses [61,65]. In a recently published report, Xia *et al.* have also showed the accumulation of the AP encoded by the *cdr-1* gene in response to pathogen attacks, suggesting its potential involvement in the activation of inducible resistance mechanisms by generating a peptide signal [24].

Plant senescence and programmed cell death (PCD)

The specific accumulation and maximum activity of cyprosins in mature flowers of *C. cardunculus* L. may indicate a role of these APs in the senescence of this organ [78]. A similar function was also proposed for cardosin A, though in this case the authors suggest that this is probably a secondary function of the enzyme [76].

In a recently published paper, Bhalerao *et al.* studied gene expression during autumn leaf senescence in aspen trees and found, among others, two genes encoding similar APs that were upregulated in the autumn leaves. According to their data, APs may play an important role during chloroplast degradation [83].

In barley, two independent studies have shown that phytepsin and nucellin were highly expressed during autolysis of different tissues suggesting the involvement of these APs in developmentally regulated PCD. Nucellin gene expression was shown to be synchronized with degeneration of barley nucellar cells by PCD after pollination, and it was suggested that this AP-like protein lacking the prosegment and the PSI may function as an apoptotic protease triggering nucellar cell death or as a hydrolytic protease that converts cell death proteins into new nutrients for embryo and endosperm development [22]. In the case of phytepsin, gene expression was upregulated in developing tracheary elements and sieve cells of barley roots. However, the specific function of the enzyme in these tissues undergoing autolysis is still undetermined [77]. In another study phytepsin was detected during the onset of DNA fragmentation in germinating barley scutella [84].

During daylily petal senescence and cell death, several genes were shown to be strongly upregulated. One of these daylily senescence-associated genes encodes an AP and it was suggested that this protein may be involved in cell death by hydrolyzing important cell components and/or activating other proteinases [85].

Table 1. Distinguishing sequence features of *Arabidopsis thaliana* aspartic proteinases.

AP Type	Active site motif	Other sequence features
Typical	Hydrophobic-hydrophobic-DTG-serine-serine	Saposin-like sequence HTVFD at the C-terminal region
Nucellin-like	Acidic-hydrophobic-DTG-serine-acidic	Cys-rich sequence between Asp32 and Tyr75 QCYDE before Tyr75
Atypical	Hydrophobic-hydrophobic-DTG-serine-acidic	Cys-rich sequence between Asp32 and Tyr75 Cysteine at the C-terminus of the protease domain

Plant sexual reproduction

Because cardosin A was shown to be highly expressed in the stigmatic papillae and to contain an Arg-Gly-Asp (RGD) motif which is a well known integrin-binding sequence it was suggested that this enzyme may participate in an RGD-dependent proteolytic mechanism in pollen–pistil interaction [10,76]. The specific localization of cardosin B in the stylar transmitting tissue also suggests that this enzyme may be involved in the remodeling and/or degradation of the pistil extracellular matrix during pollen tube growth [15]. It is thus possible that both cardosins may fulfill important roles during the sexual reproduction of the plant.

As mentioned above, most of the biological functions proposed for plant APs are still speculative and more work will be needed to confirm the proposed hypotheses or elucidate new *in vivo* functions for plant APs.

The *Arabidopsis thaliana* genome reveals more than the typical plant APs

In what concerns sequence, three-dimensional structure, localization, proteolytic specificity or function, most knowledge about plant APs of the A1 family comes from studies with typical APs, which are characterized by the presence of signal peptide, prosegment and PSI. Thus far, barley nucellin, tobacco CND41 and *cdr-1* constitute the exceptions showing a different structural arrangement [22–24]. However, and with the completion of the *Arabidopsis* genome, new perspectives have arisen concerning the diversity of plant A1 family members. A bioinformatic analysis of the *Arabidopsis* genome sequence revealed more than 50 AP-like proteins of this family [86,87]. In a first analysis based on their gene structure organization and evolutionary relationship, *Arabidopsis* A1 proteases were compiled in five groups [86]. Besides the variability in the intron sizes, the typical and nucellin-like *Arabidopsis* APs exhibit conservation of exon/intron organization when compared to the coding region of AP genes from *Brassica*, rice [7,8,55,86] and cardoon (C. Pimentel & C. Faro, unpublished results). The intron insertion sites follow a pattern significantly different from that observed in animal AP genes. While the majority of the typical plant AP genes comprise 12 introns and 13 exons in the coding sequence (two of the three *Arabidopsis* typical-AP genes comprise 11 introns), the nucellin-like APs have a distinct exon/intron organization with eight exons and seven introns.

In a second study the deduced amino acid sequences were grouped into three classes – typical plant aspartic proteinase, nucellin-like and atypical aspartic proteinase sequences, depending on their putative domain organizations and their active site sequence motifs [87] (Table 1). From this study it emerges that most plant aspartic proteinases have remained elusive, most likely because their enzymatic properties are atypical and their localizations are unexpected.

New questions concerning functional significance and specialization of this multigene family of proteases are now starting to be addressed and will definitely give new insights regarding the roles of plant APs.

References

1. Davies, D.R. (1990) The structure and function of the aspartic proteinases. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 189–215.
2. Dunn, B.M. (2002) Structure and mechanism of the pepsin-like family of aspartic peptidases. *Chem. Rev.* **102**, 4431–4458.
3. Hill, J. & Phylip, L.H. (1997) Bacterial aspartic proteinases. *FEBS Lett.* **409**, 357–360.
4. Rawlings, N.D. & Barrett, A.J. (1999) MEROPS: the peptidase database. *Nucleic Acids Res.* **27**, 325–331.
5. Rawlings, N.D. & Barrett, A.J. (1995) Families of aspartic peptidases, and those of unknown catalytic mechanism. *Methods Enzymol.* **248**, 105–120.
6. An, C.I., Fukusaki, E. & Kobayashi, A. (2002) Aspartic proteinases are expressed in pitchers of the carnivorous plant *Nepenthes alata* Blanco. *Planta* **214**, 661–667.
7. 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.
8. Chen, X., Pfeil, J.E. & Gal, S. (2002) The three typical aspartic proteinase genes of *Arabidopsis thaliana* are differentially expressed. *Eur. J. Biochem.* **269**, 4675–4684.
9. Cordeiro, M.C., Xue, Z.T., Pietrzak, M., Pais, M.S. & Brodelius, P.E. (1994) Isolation and characterization of a cDNA from flowers of *Cynara cardunculus* encoding cyprosin (an aspartic proteinase) and its use to study the organ-specific expression of cyprosin. *Plant Mol. Biol.* **24**, 733–741.
10. Faro, C., Ramalho-Santos, M., Vieira, M., Mendes, A., Simões, 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.
11. Laloi, M., McCarthy, J., Morandi, O., Gysler, C. & Bucheli, P. (2002) Molecular and biochemical characterisation of two aspartic proteinases TcAP1 and TcAP2 from *Theobroma cacao* seeds. *Planta* **215**, 754–762.
12. 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.
13. 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.
14. 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.
15. Vieira, M., Pissarra, J., Verissimo, P., Castanheira, P., Costa, Y., Pires, E. & Faro, C. (2001) Molecular cloning and characterization of cDNA encoding cardosin B, an aspartic proteinase accumulating extracellularly in the transmitting tissue of *Cynara cardunculus* L. *Plant Mol. Biol.* **45**, 529–539.
16. Koelsch, G., Mares, M., Metcalf, P. & Fusek, M. (1994) Multiple functions of pro-parts of aspartic proteinase zymogens. *FEBS Lett.* **343**, 6–10.
17. 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.

18. 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.

19. Mutlu, A. & Gal, S. (1999) Plant aspartic proteinases: enzymes on the way to a function. *Physiol. Plantarum* **105**, 569–576.

20. Ramalho-Santos, M., Verissimo, P., Cortes, L., Samyn, B., Van Beeumen, J., Pires, E. & Faro, C. (1998) Identification and proteolytic processing of procardosin A. *Eur. J. Biochem.* **255**, 133–138.

21. 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.

22. Chen, F. & Foolad, M.R. (1997) Molecular organization of a gene in barley which encodes a protein similar to aspartic protease and its specific expression in nucellar cells during degeneration. *Plant Mol. Biol.* **35**, 821–831.

23. Nakano, T., Murakami, S., Shoji, T., Yoshida, S., Yamada, Y. & Sato, F. (1997) A novel protein with DNA binding activity from tobacco chloroplast nucleoids. *Plant Cell* **9**, 1673–1682.

24. Xia, Y., Suzuki, H., Borevitz, J., Blount, J., Guo, Z., Patel, K., Dixon, R.A. & Lamb, C. (2004) An extracellular aspartic protease functions in *Arabidopsis* disease resistance signaling. *EMBO J.* **23**, 980–988.

25. 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.

26. Liepinsh, E., Andersson, M., Ruysschaert, J.M. & Otting, G. (1997) Saposin fold revealed by the NMR structure of NK-lysin. *Nat. Struct. Biol.* **4**, 793–795.

27. Khan, A.R. & James, M.N. (1998) Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. *Protein Sci.* **7**, 815–836.

28. Khan, A.R., Khazanovich-Bernstein, N., Bergmann, E.M. & James, M.N. (1999) Structural aspects of activation pathways of aspartic protease zymogens and viral 3C protease precursors. *Proc. Natl Acad. Sci. USA* **96**, 10968–10975.

29. Asakura, T., Matsumoto, I., Funaki, J., Arai, S. & Abe, K. (2000) The plant aspartic proteinase-specific polypeptide insert is not directly related to the activity of oryzasin 1. *Eur. J. Biochem.* **267**, 5115–5122.

30. Domingos, A., Cardoso, P.C., Xue, Z., 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.

31. Heimgartner, U., Pietrzak, M., Geertsen, R., Brodelius, P., Figueiredo, A.C.D. & Pais, M.S.S. (1990) Purification and partial characterization of milk clotting proteases from flowers of *Cynara cardunculus*. *Phytochemistry* **29**, 1405–1410.

32. Park, H., Kusakabe, I., Sakakibara, Y. & Kobayashi, H. (2001) Autoproteolytic processing of aspartic proteinase from sunflower seeds. *Biosci. Biotechnol. Biochem.* **65**, 702–705.

33. 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.

34. James, M.N.G. & Sielecki, A.R. (1986) Molecular-structure of an aspartic proteinase zymogen, porcine pepsinogen, at 1.8 Å resolution. *Nature* **319**, 33–38.

35. Richter, C., Tanaka, T. & Yada, R.Y. (1998) Mechanism of activation of the gastric aspartic proteinases: pepsinogen, progastricsin and prochymosin. *Biochem. J.* **335**, 481–490.

36. Guruprasad, K., Tormakangas, K., Kervinen, J. & Blundell, T.L. (1994) Comparative modelling of barley-grain aspartic proteinase: a structural rationale for observed hydrolytic specificity. *FEBS Lett.* **352**, 131–136.

37. O'Brien, J.S. & Kishimoto, Y. (1991) Saposin proteins: structure, function, and role in human lysosomal storage disorders. *FASEB J.* **5**, 301–308.

38. Munford, R.S., Sheppard, P.O. & O'Hara, P.J. (1995) Saposin-like proteins (SAPLIP) carry out diverse functions on a common backbone structure. *J. Lipid Res.* **36**, 1653–1663.

39. Vaccaro, A.M., Salvioli, R., Tatti, M. & Ciaffoni, F. (1999) Saposins and their interaction with lipids. *Neurochem. Res.* **24**, 307–314.

40. Stenger, S., Hanson, D.A., Teitelbaum, R., Dewan, P., Niazi, K.R., Froelich, C.J., Ganz, T., Thoma-Uszynski, S., Melian, A., Bogdan, C., Porcelli, S.A., Bloom, B.R., Krensky, A.M. & Modlin, R.L. (1998) An antimicrobial activity of cytolytic T cells mediated by granulysin. *Science* **282**, 121–125.

41. Anderson, D.H., Sawaya, M.R., Cascio, D., Ernst, W., Modlin, R., Krensky, A. & Eisenberg, D. (2003) Granulysin crystal structure and a structure-derived lytic mechanism. *J. Mol. Biol.* **325**, 355–365.

42. Ponting, C.P. & Russell, R.B. (1995) Swaposins: circular permutations within genes encoding saposin homologues. *Trends Biochem. Sci.* **20**, 179–180.

43. Glickman, J.N. & Kornfeld, S. (1993) Mannose 6-phosphate-independent targeting of lysosomal enzymes in I-cell disease B lymphoblasts. *J. Cell Biol.* **123**, 99–108.

44. Zhu, Y. & Conner, G.E. (1994) Intermolecular association of lysosomal protein precursors during biosynthesis. *J. Biol. Chem.* **269**, 3846–3851.

45. Egas, C., Lavoura, N., Resende, R., Brito, R.M., Pires, E., de Lima, M.C. & Faro, C. (2000) The saposin-like domain of the plant aspartic proteinase precursor is a potent inducer of vesicle leakage. *J. Biol. Chem.* **275**, 38190–38196.

46. Salmia, M.A. (1981) Proteinase activities in resting and germinating-seeds of Scots pine, *Pinus sylvestris*. *Physiol. Plantarum* **53**, 39–47.

47. Sarkkinen, P., Kalkkinen, N., Tilgmann, C., Siuro, J., Kervinen, J. & Mikola, L. (1992) Aspartic proteinase from barley grains is related to mammalian lysosomal cathepsin-D. *Planta* **186**, 317–323.

48. Belozersky, M.A., Sarbakanova, S.T. & Dunaevsky, Y.E. (1989) Aspartic proteinase from wheat seeds – isolation, properties and action on gliadin. *Planta* **177**, 321–326.

49. Asakura, T., Watanabe, H., Abe, K. & Arai, S. (1997) Oryzasin as an aspartic proteinase occurring in rice seeds: purification, characterization, and application to milk clotting. *J. Agric. Food Chem.* **45**, 1070–1075.

50. Doi, E., Shibata, D., Matoba, T. & Yonezawa, D. (1980) Characterization of pepstatin-sensitive acid protease in resting rice seeds. *Agric. Biol. Chem.* **44**, 741–747.

51. Radlowski, M., Kalinowski, A., Adamczyk, J., Krolikowski, Z. & Bartkowiak, S. (1996) Proteolytic activity in the maize pollen wall. *Physiol. Plantarum* **98**, 172–178.

52. Garg, G.K. & Virupaks, T.K. (1970) Acid protease from germinated Sorghum. 1. Purification and characterization of enzyme. *Eur. J. Biochem.* **17**, 4–12.

53. Segundo, B.S., Casacuberta, J.M. & Puigdomenech, P. (1990) Sequential expression and differential hormonal regulation of proteolytic activities during germination in *Zea mays* L. *Planta* **181**, 467–474.

54. Tormakangas, K., Kervinen, J., Ostman, A. & Teeri, T. (1994) Tissue-specific localization of aspartic proteinase in developing and germinating Barley grains. *Planta* **195**, 116–125.

55. 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.

56. 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.

57. Mutlu, A., Pfeil, J.E. & Gal, S. (1998) A probarley lectin processing enzyme purified from *Arabidopsis thaliana* seeds. *Phytochemistry* **47**, 1453–1459.

58. Polanowski, A., Wilusz, T., Kolaczkowska, M.K., Wieczorek, M. & Wilimowska-Pelc, A. (1985) Purification and characterization of aspartic proteinases from *Cucumis sativus* and *Cucurbita maxima* seeds. In *Aspartic Proteinases and their Inhibitors*. (Kostka, V. ed) pp. 49–52. Walter de Gruyter Co., Berlin.

59. Stachowiak, D., Wilimowska-Pelc, A., Kolaczkowska, M., Polanowski, A., Wilusz, T. & Larsen, L.B. (1994) Aspartic proteinase from the seeds of figleaf gourd (*Cucurbita ficifolia*)*. *Acta Biochim. Pol.* **41**, 181–182.

60. Voigt, J., Kamaruddin, S., Heinrichs, H., Wrann, D., Senyuk, V. & Biehl, B. (1995) Developmental stage-dependent variation of the levels of globular storage protein and aspartic endoprotease during ripening and germination of *Theobroma cacao* 1 seeds. *J. Plant Physiol.* **145**, 299–307.

61. Guevara, M.G., Daleo, G.R. & Oliva, C.R. (2001) Purification and characterization of an aspartic protease from potato leaves. *Physiol. Plantarum* **112**, 321–326.

62. Kuwabara, T. & Suzuki, K. (1995) Reversible changes in conformation of the 23-kDa protein of photosystem-II and their relationship to the susceptibility of the protein to a proteinase from photosystem-II membranes. *Plant Cell Physiol.* **36**, 495–504.

63. Rodrigo, I., Vera, P. & Conejero, V. (1989) Degradation of tomato pathogenesis-related proteins by an endogenous 37-kDa aspartyl endoproteinase. *Eur. J. Biochem.* **184**, 663–669.

64. Rodrigo, I., Vera, P., Vanloon, L.C. & Conejero, V. (1991) Degradation of tobacco pathogenesis-related proteins – evidence for conserved mechanisms of degradation of pathogenesis-related proteins in plants. *Plant Physiol.* **95**, 616–622.

65. Guevara, M.G., Oliva, C.R., Machinandiarena, M. & Daleo, G.R. (1999) Purification and properties of an aspartic protease from potato tuber that is inhibited by a basic chitinase. *Physiol. Plantarum* **106**, 164–169.

66. Tokes, Z.A., Woon, W.C. & Chambers, S.M. (1974) Digestive enzymes secreted by carnivorous plant *Nepenthes macferlanei*-1. *Planta* **119**, 39–46.

67. Verissimo, P., Faro, C., Moir, A.J., Lin, Y., Tang, J. & Pires, E. (1996) Purification, characterization and partial amino acid sequencing of two new aspartic proteinases from fresh flowers of *Cynara cardunculus* L. *Eur. J. Biochem.* **235**, 762–768.

68. Chen, S., Zhao, J. & Agboola, S. (2003) Isolation and partial characterization of rennet-like proteases from Australian cardoon (*Cynara cardunculus* L.). *J. Agric. Food Chem.* **51**, 3127–3134.

69. Timotijevic, G.S., Radovic, S.R. & Maksimovic, V.R. (2003) Characterization of an aspartic proteinase activity in buckwheat (*Fagopyrum esculentum* Moench) seeds. *J. Agric. Food Chem.* **51**, 2100–2104.

70. Garciamartinez, J.L. & Moreno, J. (1986) Proteolysis of ribulose-1,5-bisphosphate carboxylase oxygenase in citrus leaf extracts. *Physiol. Plantarum* **66**, 377–383.

71. Mikkonen, A. (1986) Activities of some peptidases and proteinases in germinating kidney bean, *Phaseolus vulgaris*. *Physiol. Plantarum* **68**, 282–286.

72. Elpidina, E.N., Dunaevsky, Y.E. & Belozersky, M.A. (1990) Protein bodies from Buckwheat seed cotyledons – Isolation and characteristics. *J. Exp. Bot.* **41**, 969–977.

73. Bond, H.M. & Bowles, D.J. (1983) Characterization of soybean endopeptidase activity using exogenous and endogenous substrates. *Plant Physiol.* **72**, 345–350.

74. Cruz de Carvalho, M.H., d'Arcy-Lameta, A., Roy-Macauley, H., Gareil, M., El Maarouf, H., Pham-Thi, A.T. & Zuily-Fodil, Y. (2001) Aspartic protease in leaves of common bean (*Phaseolus vulgaris* L.) and cowpea (*Vigna unguiculata* L. Walp): enzymatic activity, gene expression and relation to drought susceptibility. *FEBS Lett.* **492**, 242–246.

75. Voigt, G., Biehl, B., Heinrichs, H. & Voigt, J. (1997) Aspartic proteinase levels in seeds of different angiosperms. *Phytochemistry* **44**, 389–392.

76. Ramalho-Santos, M., Pissarra, J., Verissimo, P., Pereira, S., Salema, R., Pires, E. & Faro, C.J. (1997) Cardosin A, an abundant aspartic proteinase, accumulates in protein storage vacuoles in the stigmatic papillae of *Cynara cardunculus* L. *Planta* **203**, 204–212.

77. Runeberg-Roos, P. & Saarma, M. (1998) Phytepsin, a barley vacuolar aspartic proteinase, is highly expressed during autolysis of developing tracheary elements and sieve cells. *Plant J.* **15**, 139–145.

78. Cordeiro, M.C., Pais, M.S. & Brodelius, P.E. (1994) Tissue-specific expression of multiple forms of cyprosin (aspartic proteinase) in flowers of *Cynara cardunculus*. *Physiol. Plantarum* **92**, 645–653.

79. Mutlu, A., Chen, X., Reddy, S.M. & Gal, S. (1999) The aspartic proteinase is expressed in *Arabidopsis thaliana* seeds and localized in the protein bodies. *Seed Sci. Res.* **9**, 75–84.

80. 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.

81. Marttila, S., Jones, B.L. & Mikkonen, A. (1995) Differential localization of 2 acid proteinases in germinating barley (*Hordeum vulgare*) seed. *Physiol. Plantarum* **93**, 317–327.

82. D'Hondt, K., Bosch, D., Van Damme, J., Goethals, M., Vandekerckhove, J. & Krebbers, E. (1993) An aspartic proteinase present in seeds cleaves *Arabidopsis* 2 S albumin precursors *in vitro*. *J. Biol. Chem.* **268**, 20884–20891.

83. Bhalerao, R., Keskitalo, J., Sterky, F., Erlandsson, R., Bjorkbacka, H., Birve, S.J., Karlsson, J., Gardestrom, P., Gustafsson, P., Lundeberg, J. & Jansson, S. (2003) Gene expression in autumn leaves. *Plant Physiol.* **131**, 430–442.

84. Lindholm, P., Kuittinen, T., Sorri, O., Guo, D.Y., Merits, A., Tormakangas, K. & Runeberg-Roos, P. (2000) Glycosylation of phytepsin and expression of dad1, dad2 and ost1 during onset of cell death in germinating barley scutella. *Mech. Dev.* **93**, 169–173.

85. Panavas, T., Pikula, A., Reid, P.D., Rubinstein, B. & Walker, E.L. (1999) Identification of senescence-associated genes from daylily petals. *Plant Mol. Biol.* **40**, 237–248.

86. Beers, E.P., Jones, A.M. & Dickerman, A.W. (2004) The S8 serine, C1A cysteine and A1 aspartic protease families in *Arabidopsis*. *Phytochemistry* **65**, 43–58.

87. Faro, C. & Gal, S. (2004) Aspartic proteinase content of the *Arabidopsis* genome. *Curr. Protein Pept. Sci.* in press.