The three typical aspartic proteinase genes of *Arabidopsis thaliana* are differentially expressed

Xia Chen, Joanne E. Pfeil and Susannah Gal

Department of Biological Sciences, The State University of New York at Binghamton, Binghamton, NY, USA

Genomic sequencing has identified three different typical plant aspartic proteinases in the genome of *Arabidopsis thaliana*, named Pasp-A1, A2 and A3. A1 is identical to a cDNA we had previously isolated and the two others produce proteins 81 and 63% identical to that predicted protein. Sequencing of the aspartic proteinase protein purified from *Arabidopsis* seeds showed that the peptides are derived from two of these genes, A1 and A2. Using gene specific probes, we have analyzed RNA from different tissues and found these three genes are differentially expressed. A1 mRNA is detected in all tissues analyzed and more abundant in leaves during the light phase of growth. The other two genes are expressed either primarily in flowers (A3) or in seeds (A2). *In situ* hybridization demonstrated that all three genes are expressed in many cells of the seeds and developing seed pods. The A1 and A3 genes are expressed in the sepals and petals of flowers as well as the outer layer of the style, but are not expressed in the transmitting tract or on the stigmatal surface. The A2 gene is weakly expressed only in the transmitting tissue of the style. All three genes are also expressed in the guard cells of sepals. These data suggest multiple roles for aspartic proteinases besides those proposed in seeds.

Keywords: *Arabidopsis*; guard cells; *in situ* hybridization; protein glycosylation; proteinase.

We have been studying the aspartic proteinase genes in *Arabidopsis* with the goal of understanding their function in the whole plant. Aspartic proteinases (EC 3.4.23) are one of the major classes of proteolytic enzymes found in plant and animal viruses, microbes and plant and animal cells (reviewed in [1–3]). They are a relatively simple class of enzymes that contain two aspartic acid residues at the active site. Most of the aspartic proteinases are active at acidic pH and specifically inhibited by pepstatin A. These enzymes have been purified from a variety of different monocotyledonous and dicotyledonous plants including *Arabidopsis*, barley, *B. napus*, castor bean, figleaf gourd, maize, potato, rice, spinach, thistle, tobacco, tomato and wheat [4] (reviewed in [3]). These enzymes have been primarily isolated from seeds, but aspartic proteinases have also been purified from flowers of thistles, leaves of spinach, tobacco and tomato, tubers of potato and from pollen of maize. Although these enzymes have been associated with cell death and with plant defense [5–8], a clear link with those physiological changes and the proteolytic activity is still lacking.

A number of gene sequences have been published or deposited in the databases for aspartic proteinases from different plants including *Arabidopsis*, barley, *B. napus*, *B. oleracea*, *C. calcitrapa,* cowpea, daylily, pumpkin, and thistles (*C. cardunculus*) (reviewed in [3]) The typical plant sequences predict preproproteins similar to the animal and fungal aspartic proteinases with a signal peptide and a proregion at the amino-terminus of the mature protein. In contrast, nearly all of the genes from plants contain an extra region in the latter third of the sequence called the plant specific sequence (PSS). This approximately 100 amino acid sequence has homology to the precursor of mammalian saposins with six conserved cysteine residues and the potential glycosylation site [9,10]. The PSS region is unlikely to be critical for enzymatic activity of the aspartic proteinases, however, as it is processed out of some plant enzymes [11] and is not encoded in animal or fungal genes [2]. Recently, a bacterially expressed form of the rice aspartic proteinase lacking this region was shown to be active [12]. The PSS sequence may play a role in protein targeting to the vacuole as is proposed for the homologous protein, saposin with some lysosomal enzymes [13] or in proper folding of the plant aspartic proteinases as suggested by expression of these enzymes in heterologous systems [14]. Egas and colleagues [15] have shown that the PSS of the cardosin A precursor containing this region can integrate into membranes and cause leakage. This process was pH and lipid composition dependent suggesting it may involve some cellular membranes more than others. Törmäkangas and colleagues [16] recently provided evidence that this sequence is the vacuolar sorting determinant for the barley aspartic proteinase and influences the way the protein leaves the endoplasmic reticulum. Although the vast majority of plant aspartic proteinase genes and proteins characterized to date contain both the pro region and the PSS, a few sequences have been identified which do not. Chen and Foolad [17] isolated a sequence specifically expressed in the degenerating nucellar cells of the barley embryo. This gene, called nucellin, predicts a protein with aspartic acid residues in the active site context for aspartic proteinases as well as other homologous regions, but appears to lack most of the pro

Correspondence to S. Gal, Department of Biological Sciences, The State University of New York at Binghamton, Binghamton, NY 13902-6000, Fax: + 1 607 777 6521, Tel.: + 1 607 777 4448, E-mail: sgal@binghamton.edu

Abbreviations: AtPasp, *Arabidopsis thaliana* aspartic proteinase; BAC, bacterial artificial chromosome; DIG, digoxigenin; PSS, plant specific sequence.

(Received 16 April 2002, revised 4 July 2002, accepted 5 August 2002)

region and the PSS. At this point, there is no evidence for protein production from this gene or that the protein produced is an active aspartic proteinase. Thus, while the genes for plant aspartic proteinases have a similar structure, others appear to show little identity outside the active site domain.

The aspartic proteinases isolated from plants occur as single chain or two chain enzymes. Single-chain enzymes vary in size from 30 to 65 kDa while the two-chain enzymes contain peptides of 9–16 kDa and ≈ 30 kDa molecular mass (reviewed in [3]). The genes encoding these different enzymes are remarkably similar and provide no indication of the mechanism and the signal that result in differential processing of the same type of preproprotein to either a single-chain or two-chain enzyme. Protein sequence, when it has been obtained, suggests that the peptides in the two-chain enzymes are derived from the same gene [11,18–20].

We have been characterizing the aspartic proteinases of *Arabidopsis* at both biochemical and molecular levels. We initially isolated an active two-chain aspartic proteinase from *Arabidopsis* seeds using affinity chromatography [21]. We localized this enzyme to the protein storage body in dry seeds using biochemical fractionation and immunocytochemistry [22]. A nearly full-length cDNA for the first *Arabidopsis* aspartic proteinase gene was isolated and displayed high percentage identity to several aspartic proteinase genes from other plants [23]. Southern blotting with this cDNA as a probe detected a single band under moderate stringency hybridization and with several different restriction enzymes suggesting there is a single gene for this enzyme in *Arabidopsis*. But since that work was published, we have found multiple genes homologous to this first clone in the *Arabidopsis* genome. Here, we characterize the genomic sequences of the three typical aspartic proteinase genes from *Arabidopsis*, those having the common propeptide and PSS. One of these sequences encodes the published cDNA while the other two are predicted to be 81 and 63% identical at the amino acid level. We found that these genes are differentially expressed in *Arabidopsis* plants and the active aspartic proteinase isolated from seeds is derived from two of these genes.

MATERIALS AND METHODS

Plant growth conditions

Arabidopsis thaliana plants, ecotype RLD were grown under constant light conditions in soil at 24 °C in an AR75L or an AR36L incubator (Percival Scientific, Boone, IA USA). To test the effect of light cycling on the aspartic proteinase gene expression, tissue samples were taken from plants grown under a regime of 16 h light 21 °C/8 h dark 15 °C with the light sample taken 4 h into the day and the dark sample taken 3 h into the night.

Purification and analysis of aspartic proteinase

The *Arabidopsis thaliana* aspartic proteinase was purified from dry seeds using the protocol described previously [21]. The protein sequence was obtained from three of the peptides after separation on an SDS polyacrylamide gel, and transfer to a poly(vinylidene difluoride) membrane as described [21]. To analyze the carbohydrates attached to the proteins, approximately 0.5 μg of purified protein was separated on a 12.5% homogeneous PhastSystem gel using SDS buffer strips as described by the manufacturer (Amersham Pharmacia Biotech, Piscataway, NJ, USA). One portion of the gel was stained with Coomassie blue while the other was transferred to nitrocellulose using the semidry method and the PhastSystem (Amersham Pharmacia Biotech). The glycosylated proteins were detected using concanavalin A linked to alkaline phosphatase according to the supplier's instructions (EY Laboratories, San Mateo, CA, USA) followed by visualization of the alkaline phosphatase using SigmaFast tablets containing nitroblue tetrazolium chloride and 5-bromo-4-chloro-3-indolyl phosphate (Sigma Chemical Company, St. Louis, MO, USA).

Isolation and analysis of *Arabidopsis* genome sequences

Genomic sequences from *Arabidopsis thaliana*, ecotype Columbia, were obtained from the GenBank. Reanalysis of the intron junctions was made using the splice predictor software from the MAIZE Genome Database at Iowa State University available at http://www.zmdb.iastate.edu [24]. The phylogenetic comparisons were made using the PHYLIP software version 3.573c using bootstrap software obtained from Joseph Felsenstein at the University of Washington (available through the web site http://evolution.genetics.washington.edu/phylip.html). To confirm a sequence within the *AtPasp A1* gene, we amplified a PCR product from RLD genomic DNA using gene specific primers then sequenced the PCR product directly using the Ampli-Taq ready reaction mix (Applied Biosystems, Perkin-Elmer, Foster City, CA, USA) on an ABI Prism Genetic Analyzer Model 310 sequencer (Applied Biosystems). To confirm the sequence of one intron/exon border in the *AtPasp A2* gene, we performed RT-PCR using RNA isolated as below and first strand cDNA using Ready-To-Go beads (Amersham Pharmacia Biotech). The specific region of the *AtPasp A2* gene was then amplified using primers and sequenced as above.

Northern blot and *in situ* hybridization

To make the gene specific probes for the Northern blots, we used oligonucleotides and DNA from *Arabidopsis* plants, RLD ecotype in a PCR with digoxigenin (DIG) labeling mix (Roche Biochemicals, Indianapolis, IN, USA) to produce the specific DIG-labeled fragment. The oligonucleotides for amplification of the *AtPasp A1* specific probe were (5′→3′) GTTGTCAATGAATAGGTAAAATG and CAGAATCTCCAAGTCTGTAAG; for the *AtPasp A2* gene-specific probe, the oligonucleotides were TGCTTTG ATTTTGTAGGTCA and CATCTCCAGAATCACC ACCAAG; and for the *AtPasp A3* gene-specific probe, the oligonucleotides were TGATGACAGCTAAAAAT GGGAACTAGG and CCATATCCGCATTTTCATC GTTCAGG. To generate strand-specific probes for *in situ* hybridization, these PCR fragments were cloned using the AdvanTAge system (Clontech Laboratories, Palo Alto, CA, USA) and then subcloned into the pBluescript II vector (Stratagene, La Jolla, CA, USA) which contains the T3 and T7 promoters for RNA synthesis.

For Northern blot analysis, total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen Inc., Valencia, CA, USA) according to the manufacturer's instructions from various plant tissues. Fifteen micrograms of total RNA was separated on a 1.0% formaldehyde-containing agarose gel with an RNA molecular mass marker (Promega Corp., Madison, WI, USA) (equal loading of RNA was observed using ethidium bromide staining of the gel prior to transfer), and then blotted to a nylon membrane (Roche Biochemicals) overnight at room temperature. After baking the nylon membrane at 80 °C for 1 h in a vacuum oven, the blot was prehybridized at 55 °C. Hybridization was done with DIG-labeled *AtPasp A1*, *A2* or *A3* gene specific probes at 55 °C overnight. After hybridization, the membrane was washed twice in 2 × NaCl/Cit plus 0.1% SDS for 5 min and twice in 0.1 × NaCl/Cit plus 0.1% SDS for 15 min at 55 °C (10 × NaCl/Cit contains 1.5 M NaCl, 0.15 M sodium citrate). The DIG label was detected by addition of the anti-DIG Ig followed by chemiluminescence using disodium 3-(4-methoxyspiro [1,2-dioxetane-2,3′-[5′chloro],tricycleo [3.3.1.1^{3,7}]decan]-4-y1) phenyl phosphate (CSPD) according to the manufacturer (Roche Biochemicals). Developed blots were scanned and analyzed using IMAGE MASTER VDS Software (Amersham Pharmacia Biotech).

In situ hybridization was carried out based on the protocols described previously [22]. Sense and antisense DIG-labeled RNA probes were generated by *in vitro* transcription from the cloned gene specific regions of *AtPasp A1*, *A2* or *A3* genes (see above) using the DIG RNA labeling kit (Roche Biochemicals). The labeled probe was purified by precipitating the RNA using LiCl and ethanol, and the optimum final length of the RNA probe of 50–100 bases was generated by alkaline hydrolysis as described by Drews and Okamuro [25]. Hybridization and washing were performed at 55 °C [26] and after development, the dehydrated sections were sealed with permount (Fisher Scientific, Atlanta, GA, USA) and examined with an Olympus system microscope model BH-2.

RESULTS AND DISCUSSION

The genome of *Arabidopsis* contains three typical plant aspartic proteinase genes

We are characterizing the aspartic proteinases in *Arabidopsis thaliana* and have isolated both the enzyme from seeds and a cDNA (*AtPasp A1*) (accession no. U51036) [21,23]. Sequencing of the *Arabidopsis* genome has now identified several putative aspartic proteinase genes with significant homology to the *AtPasp A1* cDNA clone (Fig. 1). One of these genomic clones is identical to *AtPasp A1* (from BAC F12F1.24 also called At1g11910) while two other genes were observed in the genome database with significant homology to this gene particularly in the putative active site regions containing the Asp-Thr-Gly (DTG) and Asp-Ser-Gly (DSG) sequences. We have called these genes *AtPasp A2* and *AtPasp A3* as they are the two other genes with the most homology to our original cDNA (the *AtPasp A2* gene is from BAC F19K23.21, also called At1g62290 and the *AtPasp A3* gene is from BAC T26N6.7, also called At4g04460). The genes are on three different chromosomal regions, the *AtPasp A1* and *A2* genes being on different regions of chromosome I while the *A3* gene is on chromosome IV. The original *AtPasp A2* gene in the annotated genomic sequence predicted a slightly different protein (sequence not shown but see Δ in Fig. 1). These differences were found at the assignments of intron/exon borders, so this region of the BAC sequence was reanalyzed using another splice site predicting program at the Maize Genome Database at lowa State University, USA [24]. This program did find introns at the sites expected based on the first cDNA sequence and when retranslated, the new AtPasp A2 predicted protein with the changes at residues 304 and 416 had a higher identity to the AtPasp A1 predicted protein (Fig. 1). This altered splicing pattern at the first site in the *AtPasp A2* mRNA was confirmed by sequencing an RT-PCR product from the *AtPasp A2* gene (data not shown), while the splicing at the second site was supported by the size of the mRNA on Northern blots and sequence of the protein (see below). Thus, we propose another annotation of this BAC clone in these regions to reflect our data and analysis. The accession no. of this modified sequence is TrEMBLO04593.

These three genes encode proteins that contain the commonly observed arrangement of structural regions found in most plant aspartic proteinases (reviewed in [3]) and so are termed typical for this class of enzymes. These include a predicted signal peptide (assessed by the algorithm of [27]) a pro-region, the mature large subunit containing both active site aspartic acid residues, and the mature small subunit interrupted by the PSS of approximately 100 amino acids (Fig. 1). The AtPasp A1 and A2 proteins have overall 81% identity while the protein derived from *A3* gene is 63 and 64% identical to the predicted A1 and A2 proteins, respectively. Unsurprisingly, the predicted proteins have highest identity in the mature protein regions, the mature heavy and light chains (Fig. 1). There is still significant homology in the PSS and the proregions among the sequences, regions presumed to be under less selective pressure than the regions involved in the activity of the proteinase. A phylogenetic comparison of 15 of the known plant aspartic proteinases proteins using the region between the two active site aspartic acid residues revealed relationships within this family of genes (Fig. 2). The AtPasp A1 protein is highly related to the aspartic proteinase sequences from *B. napus* and *B. oleracea*. The two other *Arabidopsis* protein sequences, A2 and A3, appear to be on a different subbranch from the A1 protein but on the same larger branch as the aspartic proteinases from cowpea and pumpkin (Fig. 2). The three predicted proteins from monocots are grouped together while the protein from a tomato gene appears in a distinct branch. Interestingly, the three predicted proteins from *C. cardunculus* are not all on the same branch. Two of the proteins appear in a cluster together with the protein from *C. calcitrapa*, while the third enzyme from *Cynara* called cardosin A appears on a distinct branch of the phylogenetic tree (Fig. 2).

Comparison of the intron insertion sites of these three *Arabidopsis* genes confirms that plants have a pattern significantly different from those observed in animal aspartic proteinase genes [23,28] (Fig. 3). The *Arabidopsis AtPasp A1* gene, like the genes from *B. napus* and rice, has 12 introns within the coding region (Fig. 3). The *AtPasp A2* and *AtPasp A3* genes are each missing one of these introns (in different places), but all other introns are

 © FEBS 2002

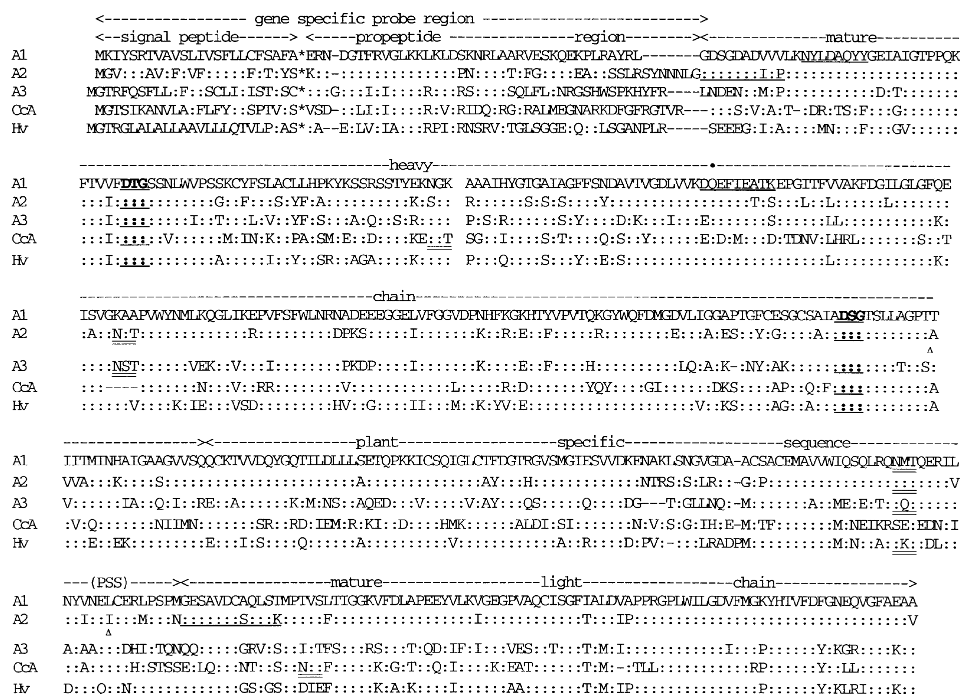

Fig. 1. Comparison of three aspartic proteinase sequences from the *Arabidopsis* genome with other plant aspartic proteinases. The amino acid sequences from *Arabidopsis* are deduced from our aspartic proteinase cDNA and the corresponding genomic sequence (*AtPaspA1* (A1), accession no. U51036 for cDNA [23] and genomic from BAC F12F1.24, also known as At1g11910), and two related genomic sequences (*AtPaspA2* (A2) from F19K23.21 also known as At1g62290 and *AtPaspA3* (A3) from T26N6.7 also known as At4g04460). These sequences are compared to the cardosin A from *C. cardunculus* (CcA; accession no. AJ132884 [20]) and the barley aspartic proteinase (Hv; accession no. X56136 [18]). The regions of the sequence are identified above the top sequence using nomenclature from previous publications (reviewed in [3]). Residues identical to the AtPaspA1 protein were given ':' and gaps '-' were inserted to improve alignment. The '*' indicates the end of the predicted signal peptide [27], confirmed using the barley cDNA [59]. DSG and DTG shown as underlined and bold type are the active site aspartic acid residues. The potential (in the *Arabidopsis* genes) and actual N-glycosylation sites (in the cardosin A and barley enzymes [32]) are double-underlined and italicized. Underlined peptides represent sequences from the protein isolated from *Arabidopsis* seeds. The sequences of the amino terminal peptides were GDSGDA DIVPL from the 31 kDa subunit and GESAVD?SQL?K from the 6 kDa subunit, and NYLDAQYY and DGEFIEATK from two internal peptides of the 28 kDa protein. The '●' above the 'Q' in the second underlined sequence indicates a difference in the peptide sequence and that predicted by the *AtPAspA1* gene. The two places in the A2 sequence with Δ indicate positions with distinct differences from the annotated sequence in BAC F19K23.21 due to differences in intron assignments. The new assignment of this sequence is accession no. TrEMBL O04593. The region representing that used to probe Northern blots specifically for each of the genes is indicated above the top line.

in identical positions. The sizes of the introns in the *Arabidopsis* genes are all small, ranging from 72 to 184 bp, similar in size to those found in the gene from *B. napus*, but smaller than those found in the rice gene [23,28]. The positions of these introns contrast with the positions of introns in the other barley aspartic proteinase-like gene, nucellin and the animal aspartic proteinases typified by human cathepsin D [17,29] (Fig. 3). This would support the hypothesis that the plant genes for the typical aspartic proteinases were derived from a common ancestor, distinct from the predecessor of the nucellin gene, which contained these introns, but only gained them after the separation of the plant and animal kingdoms. The fact that two genes are each missing one of the commonly found introns suggests that the presence of introns in these genes is unstable.

It is clear that the three genes described in this work are the only typical plant aspartic proteinases in the *Arabidopsis* genome, those with the well characterized arrangement of propeptide and PSS. However, there are at least 37 other genomic sequences which encode potential aspartic proteinases (S. Gal & C. J. Faro, unpublished results). It is not known whether these genes produce active aspartic proteinases; some predict significantly different proteins from those previously characterized. Thus, there are nearly 40 different aspartic proteinase-like sequences in the *Arabidopsis* genome. Several other plants have multiple aspartic proteinases (reviewed in [3]). There are at least two distinct ESTs from *B. oleracea*, while the other close relative of *Arabidopsis*, *B. napus* appears to have at least four genes. Another distantly related dicotyledonous plant, *C. cardunculus* has at least six, while the monocotyledonous plants rice and barley have so far three and two, respectively. If one can extrapolate from the *Arabidopsis* genome, these other plants should have many more as yet unidentified aspartic proteinase-like sequences in their genomes. A recent report from the nematode worm, *Caenorhabditis elegans* indicates as many as 12 aspartic proteinases in the genome of this simple organism [30] while many new aspartic proteinase sequences are also being detected in the human genome [31] (J. Kay, Cardiff University, Wales, UK, personal communication).

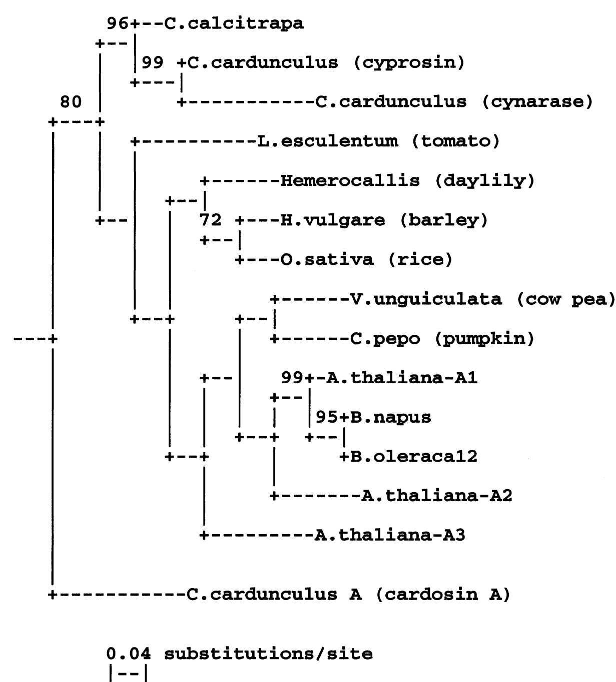

Fig. 2. Phylogenetic tree of several plant aspartic proteinases. The sequence containing the two active sites as well as the region between these sites for 15 different plant aspartic proteinases was used to build a phylogenetic tree using the neighbor-joining method as described in Materials and methods. The numbers along some of the branches represent bootstrap values for proportion of 100 trees showing the indicated grouping with values below 60 not shown. The sequences used are as follows for *Arabidopsis thaliana* the *A1* gene from BAC F12F1.4 (accession no. AC002131 from chromosome 1), *A2* gene from BAC F19K23.21 (AC000375 also from chromosome 1), and *A3* gene from BAC T26N6.7 (AF076243 from chromosome 4), *B. napus* (U55032), *B. oleracea* is a combination between two sequences (X80067 and X77260), *C. calcitrapa* (*C. calcitrapa* Y09123) *C. cardunculus* (cynarase X69193, cyprosin X81984 and cardosin A AJ132884), *Cucurbita pepo* (AB002695), *Hemerocallis* (AF082029), *Hordeum vulgare* (phytepsin X56136), *Lycopersicon esculentum* (L46681), *Oryza sativa* (D32165) and *Vigna unguiculata* (U61396).

The seed proteinase is derived from two of these genes

The aspartic proteinase from *Arabidopsis* seeds has four polypeptide components of molecular mass 31, 28, 16 and 6 kDa [21]. The sizes of these peptides are similar to those noted for the aspartic proteinase purified from barley seeds [19]. In recent experiments with different seed lots, we have only isolated the two larger forms and the smallest polypeptide; we have not reproducibly isolated the 16 kDa form (J. E. Pfeil & S. Gal, unpublished data). The reason for this is unclear. We previously confirmed the identity of our peptides relative to the first cDNA clone using amino terminal sequencing [21]. However, as the sequenced regions share identity with the other aspartic proteinase genes, we obtained more sequence information from those same three peptides. As shown in Fig. 1, the extended peptide sequences corresponded to different genes. The extended amino terminal sequences from the 31 and 6 kDa peptides corresponded to the AtPasp A2 protein (GDSGDADIVPL from the 31 kDa peptide and GE SAVD?SQL?K from the 6 kDa peptide), while the amino acid sequence from another internal peptide from the 28 kDa protein corresponded to the AtPasp A1 protein (sequence DGEFIEATK). The unidentified amino acid residues in the peptide from the 6 kDa subunit correspond to a cysteine and a serine in the predicted sequence (Fig. 1). The fact that these residues were not confidently determined suggest that they are modified either as a disulfide bond in the case of the cysteine or with a modification on the hydroxyl of the serine. This peptide sequence that corresponds to the AtPasp A2 protein is confirmatory evidence for the alternative splicing of the gene from the original annotated sequence that would not have contained this peptide (Fig. 1). Interestingly, the sequence from the internal peptide of the 28 kDa protein indicates a G at amino acid number 105 relative to the predicted start of the mature protein rather than the Q predicted by the cDNA and genomic clones from the *AtPasp A1* gene (marked with a '●' in Fig. 1). This difference could have been due to a polymorphism between the different cultivars of *Arabidopsis* used (Columbia for the nucleotide sequence and RLD for the protein). However, amplification of this region using RLD genomic DNA as

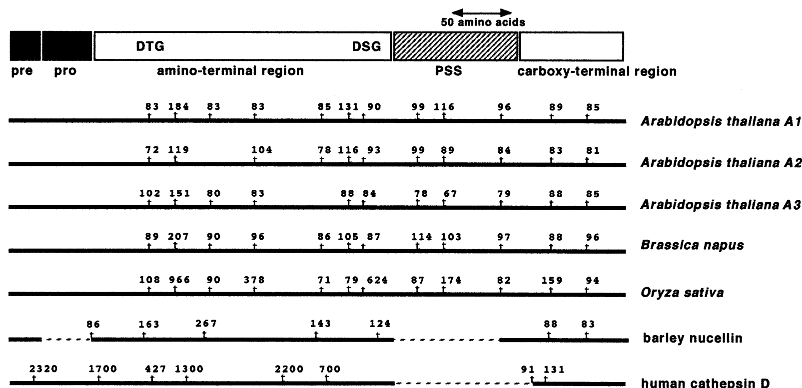

Fig. 3. Intron comparison of several plant aspartic proteinase genes with the human cathepsin D gene. The top line of the figure shows the arrangement of the functional regions of the plant aspartic proteinase genes like in Fig. 1. Below that are the representative sequences of three *Arabidopsis* (this work), the *B. napus* [23], the *Oryza sativa* [28] and the barley nucellin [17] aspartic proteinase gene sequences showing the positions (site of arrow head) and sizes of the introns (number above arrowhead, in bp). This is compared with the intron arrangement in the mammalian aspartic proteinase gene for human cathepsin D [29].

the template and sequencing of the resulting fragment encoded a Q in this position. Thus, it is not clear whether peptide sequencing error occurred, or a post-translational modification could explain this difference.

Potential N-glycosylation sites differ among the three predicted *Arabidopsis* aspartic proteinases (Fig. 1). Binding of concanavalin A was observed only to the 31 kDa peptide of the purified aspartic proteinase indicating that this peptide is likely glycosylated with a high mannose chain (data not shown). The binding was blocked by the addition of methyl-mannose as expected. Thus, seeds contain proteins from both the *AtAspP-A1* and *AtAspP-A2* aspartic proteinase genes, with the N138 on the mature heavy chain of the latter proteinase being glycosylated (see Fig. 1). In comparing the sequences of these five proteinases, we see differences in their predicated glycosylation sites. All of the predicted *Arabidopsis* proteins and the barley sequence have a potential N-glycosylation site within the PSS, but cardosin A does not contain that site (Fig. 1). This region has been shown to be glycosylated on the 16 kDa peptide from the barley seed purified protein [32]. We do not have this peptide in our present preparations; it is probably removed during processing in a manner similar to that which occurs with cardosin A [11]. The cardosin A protein is glycosylated on two other asparagines (N70 and N363) in the mature heavy and light peptides, respectively [32]. Neither of these sites are potential glycosylation sites in any of the predicted proteins from the barley gene or these genes from *Arabidopsis*. Thus, it appears that N-glycosylation does not play an essential role for the activity of the plant aspartic proteinases as the sites of addition of N-glycans are not conserved in the different proteins. Glycosylation of human renin expressed in monkey cells improved the stability of the protein [33] suggesting another potential function of this modification in the plant enzymes. Thus, we obtained protein sequence from the aspartic proteinase purified from seeds [21] that showed that this enzyme was derived from the *AtPasp A1* and *A2* genes. This is consistent with both these genes being expressed in the seeds and seed pods of *Arabidopsis* plants (see below). We have not yet detected a protein derived from the *AtPasp A3* gene that is primarily expressed in flowers (see below), but polyclonal antibodies made to the seed proteinase detect peptides in this tissue [34].

The three genes are differentially expressed

Although, we have not detected the AtPasp A3 protein, ESTs for this gene are present in the data base (accession numbers T75975 and Z37495) suggesting the gene is expressed. Thus, we suspected differential expression of these three aspartic proteinases genes in *Arabidopsis*. In our original Northern blots [22,35], we used the first cDNA isolated from the *AtPasp A1* gene as a probe but this also detected cloned products of the *AtPasp A2* and *A3* genes (data not shown). Thus, we developed probes that were gene specific using the 5′ end of the coding region and higher stringency hybridization and wash conditions to distinguish expression of these three genes. The region chosen encodes the signal sequence and proregions of these three genes (Fig. 1). Using these probes and hybridization conditions, DNA from the three genes could be distinguished (data not shown). We then used these probes to monitor levels of expression of the *AtPasp A1, A2* and *A3* genes using Northern blot hybridization of total RNA from different tissues. These results confirmed that the three genes are differentially expressed in the tissues examined (Fig. 4). All three genes produce a 1.9 kb mRNA, as expected (Fig. 1). The previous annotation for the *AtPasp A2* gene predicts a stop codon near the end of the PSS in the protein and a mRNA that is only ≈ 1.3 kb (see Δ in Fig. 1). When that sequence was reanalyzed, a strong donor site was predicted and an intron assigned in that region changed the predicted mRNA moving the frame of the gene. This predicts a longer message more consistent with the size we observed on Northern blots. The *AtPasp A1* gene is expressed in multiple tissues including dry seeds, flowers, stems/bolts and roots (Fig. 4). Both the *AtPasp A2* and *A3* genes exhibit more restricted expression patterns at the organ level. The *AtPasp A3* gene is strongly expressed in flowers and more weakly in seed pods while the *AtPasp A2* gene message was found in seed pods and dry seeds, but not detected in any of the other tissues (Fig. 4).

Desprez and colleagues [36] using cDNA microarrays to analyze gene expression changes caused by light, tested the *AtPasp A1* cDNA (accession no. U51036) and found that it was induced by light. Because hybridization of the whole cDNA is not gene specific, we examined whether the *AtPasp A1* gene was regulated by light. We isolated leaf

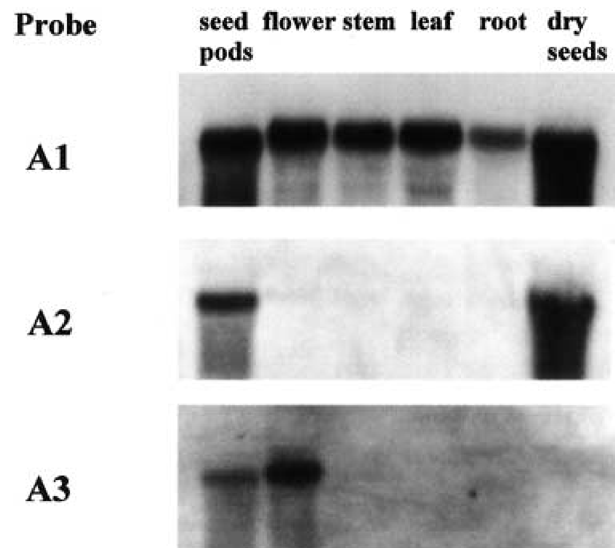

Fig. 4. Northern blot hybridization showing tissue specific expression of *Arabidopsis* aspartic proteinases. Fifteen micrograms of total RNA isolated from seed pods, flowers, stems (bolts), leaves, roots and dry seeds was separated on a 1.0% formaldehyde agarose gel and blotted to a nylon membrane. Equal loading of RNA was observed using ethidium bromide staining of the gel prior to transfer (not shown). Hybridization was performed with a digoxigenin labeled *AtPasp A1, A2* or *A3* gene-specific probe at 55 °C overnight followed by washing at the same temperature with $0.1 \times$ NaCl/Cit containing 0.1% SDS. The probe was then detected using anti-DIG Ig conjugated to alkaline phosphatase as described in Materials and methods.

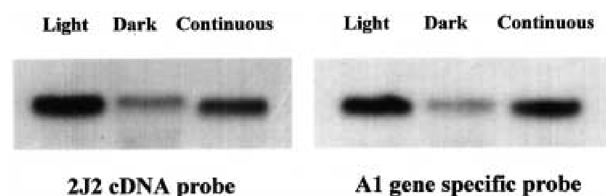

Fig. 5. Northern blot hybridization showing effect of light on *Arabidopsis* aspartic proteinase *A1* gene. RNA was isolated from leaves of plants grown in continuous light or from light cycling plants taken either during the day or during the night. Fifteen micrograms of total RNA was separated on a 1.0% formaldehyde agarose gel and blotted to a nylon membrane. Equal loading of RNA was observed using ethidium bromide staining of the gel prior to transfer (not shown). Hybridization was performed with a digoxigenin labeled *AtPasp A1* gene-specific probe or with the entire *AtPasp A1* gene cDNA (2J2) at 55 °C overnight followed by washing at the same temperature with 0.1 × NaCl/Cit containing 0.1% SDS. The probe was then detected using anti-DIG Ig conjugated to alkaline phosphatase as described in Materials and methods.

total RNA from plants grown in continuous light, and from plants grown under a light cycling regime either at night or during the day and performed Northern blots with the *AtPasp A1* gene specific probe (Fig. 5). The *AtPasp A1* gene-specific probe detected a band that was nearly four times stronger in the tissue from the light when compared to the tissue from the dark. This was the same order of induction detected by the whole cDNA with the same RNA (2J2 in Fig. 5). Thus, the *AtPasp A1* gene is expressed in leaves and to a larger extent in leaves taken from plants during the light phase of growth than those taken from the dark phase.

Messages for most of the aspartic proteinases in other plants have also been found in a wide variety of tissues. The barley aspartic proteinase mRNA was found in developing seeds, mature grains and leaves [18], and in flowers, leaves, leaf internodes, pericarp and testa [37]. The other barley aspartic proteinase-like gene, nucellin was found only in pollinated ovaries, and not in leaves or anthers [17]. In rice, the message for one aspartic proteinase gene was found in developing seeds, seedlings up to 5 days after germination and in roots at all times [38]. The tomato aspartic proteinase gene is expressed in roots, stems, flowers and green fruit, but not in leaves or in red fruit [39]. One of the genes from *Cynara* is expressed predominantly in flowers and bracts and not in leaves [40]. Thus it appears that there is a complicated regulation of these genes in plants and that other tissues besides seeds contain aspartic proteinases. Using the antibody to the seed protein, we have detected antigenic species in nonseed tissues but have not completely characterized these peptides (J. E. Pfeil, A. Mutlu & S. Gal, unpublished data). The *AtPasp A1* gene appears also to be regulated by light in leaves. Aspartic proteinase genes have been shown to be induced by wounding in tomato leaves [39], by senescence in daylily petals [7] and by low osmoticum in *B. oleracea* [41]. But to our knowledge, no other laboratories have demonstrated the response to light by other plant aspartic proteinases. Thus, we have found both tissue-specific and light-regulated expression of the aspartic proteinase sequences from *Arabidopsis*.

The aspartic proteinase genes show different distribution of expression in flowers

As at least two organs appear to express more than one of these three genes, we were interested in determining if the genes show tissue or cell type specificity by *in situ* hybridization using the gene specific probes. The *AtPasp A1* and *A3* genes are both expressed in flowers, while all three genes are expressed in seed pods. In seeds, we found expression of all three genes in all seed cell types (Fig. 6A and data not shown). Although little labeling was visualized with the Northern blots of seed tissues using the AtPasp A3 gene specific probe, we were able to see some labeling in these tissues with the *in situ* hybridization experiments. It appears this latter technique may be better for seeing small amounts of label in a few cell types than the Northern blots which necessarily take large amounts of cell tissue together. Little or no label was detected in the sense controls of any of the genes (Fig. 6B and data not shown). In seed pods, the three genes showed similar expression (Fig. 6I and data not shown). The genes were expressed in the inner and outer cells in both the central and outer layers of the seed pod. Although some labeling was detected on the outer layer of the developing seed, this labeling was also seen in the sense controls (data not shown). These results corroborated previously published work using the cDNA [22]. The genes are expressed in many cell types in the seed overlapping the expression of the two seed storage proteins of *Arabidopsis*, 12S globulin and 2S albumin [42,43]. The expression of our proteinase genes in the same cell types as the seed storage proteins would be consistent with the proposed role of these enzymes in the processing and degradation of the storage proteins [21,44,45] (A.T. Corcoran, S.M. Reddy & S. Gal, unpublished results).

The localization of the messages in flowers does show some gene-specific differences. In flowers, the *AtPasp A1* and *A3* gene messages were strongly detected in the petals and carpel tissues, but not in the transmitting tract and not on the stigmatic surface (Fig. 6D,F). The messages for these genes were also visualized in the outer cell layers of the anther early in flower development (Fig. 6F), but not later after the flower opened (Fig. 6D). The mRNA from the *AtPaspA2* gene was not detected in flowers by Northern analysis, but was found to be weakly detected in the transmitting tract of the flowers using *in situ* hybridization (Fig. 6C). Thus some of our genes are expressed in different parts of the flower. The localization of cardosin A protein on the stigmatal surface [46], while the cardosin B protein is found in the transmitting tract of the thistle flower [47], suggests other aspartic proteinases may have similar differential localization in this tissue. Unfortunately we can not make direct comparisons as our work has primarily involved the detection of messages for the different genes while the work in thistles involves detection of protein. Interestingly the cardosin A enzyme has an RGD motif which has been proposed to play a role in an adhesion-mediated proteolytic process in pollen recognition and growth at the stigmatic surface [48]. The *Arabidopsis* gene expressed in the transmitting tract, *AtPaspA2*, has an RGE motif in that same position which would probably bind to similar components. The fact that these potentially related proteins may be localized in different parts of the flowers may reflect different rates of pollen tube growth or the size

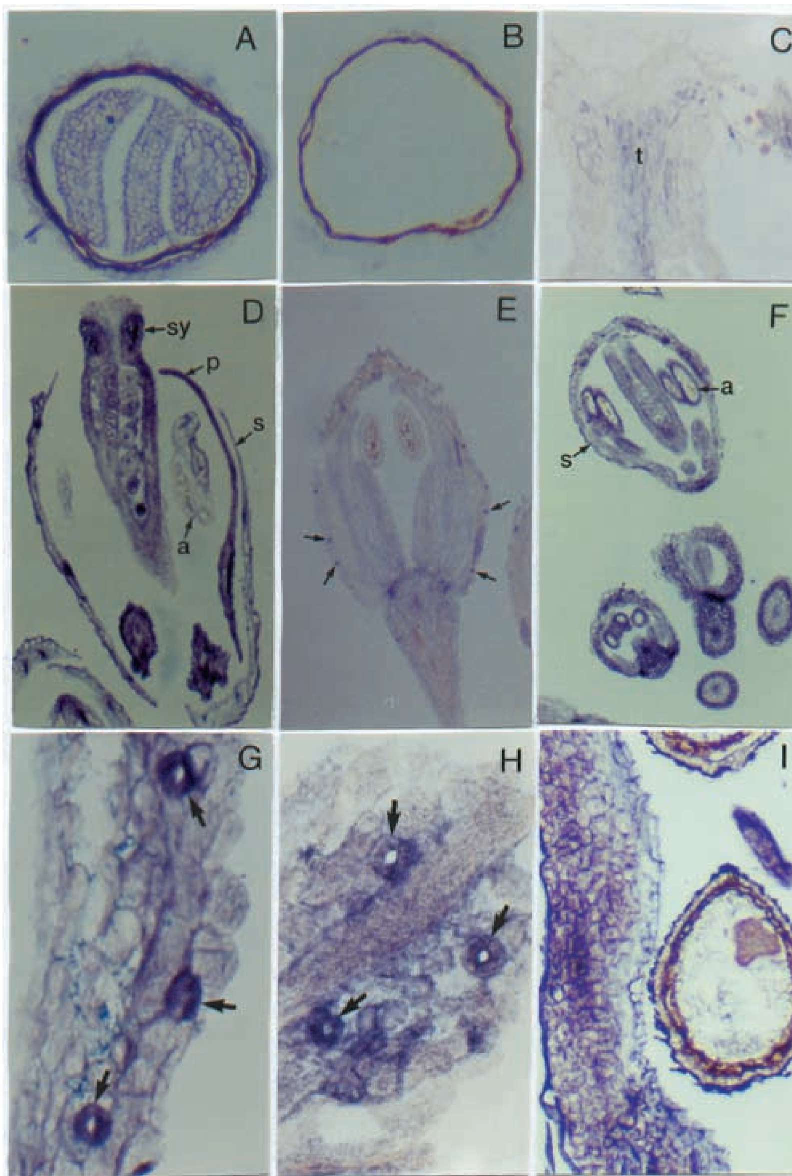

Fig. 6. *In situ* hybridization of the *Arabidopsis* aspartic proteinases to seed and flower tissues. Tissue from dry seeds (panels A and B), flowers (panels C-H) or seed pods (panel I) were prepared, hybridized and detected as described in Materials and methods section. Panels A, D, G and I, antisense probe for *AtPasp A1* gene; panel B, sense probe for *AtPasp A1* gene; panels C and E, antisense *AtPasp A2*-specific probe; panels F and H antisense *AtPasp A3*-gene specific probe. The magnifications for the panels are as follows: panels A, B, C and I: 100×, panels D, E and F: 40×, and panels G and H: 400×. The tissues are labeled as follows: a, anther; p, petal; s, sepal; sy, style; t, transmitting tract, while guard cells are indicated in panels E, G and H with arrows.

of the style in the two plants, thistles and *Arabidopsis*. Interestingly, the cardosin B protein is found in the extracellular matrix of the transmitting tract [47], while we have only found our enzymes inside cells although we have only looked in seeds.

The mRNA from the *AtPaspA2* gene was also detected in a distinct punctate pattern on the sepals using *in situ* hybridization (Fig. 6E). Upon higher magnification, this expression was localized to guard cells (data not shown). When flowers hybridized with the *AtPasp A1* or *A3* gene antisense probe were examined at a higher magnification, expression in guard cells of the sepals was also seen (Fig. 6G,H). Guard cells of flowers are not as well characterized as those on leaves [49] but they have been analyzed in lily, avocado and apple inflorescences [50–52]. The distribution of the stomates in flowers is significantly lower than in leaves, but the tissue has been shown to be photosynthetically active. As far as we are aware, this is the first indication of a protease predominantly expressed in guard cells of flowers. An *Arabidopsis* mutant with an increased stomatal density in leaves was found to be disrupted in a putative subtilisin-like serine proteinase [53]. Whether this enzyme plays a role inside the guard cell or is involved in the modulation of some external developmental factor is not yet clear. At present, we do not know whether the aspartic proteinase genes are also expressed in the guard cells of leaves; if so, this could explain the greater expression of the *AtPaspA1* gene observed in the light.

The aspartic proteinases in seeds are believed to be involved in storage protein processing during seed development and in storage protein breakdown during germination (reviewed in [3]), but what role could they play in guard cells? The appearance of the aspartic proteinase in the seive cells of barley stems [6], in degenerating lily petals [7] and senescing leaves of *B. napus* [5] suggest an expanded role of these enzymes in processing and degradation of other substrates. We have obtained some experimental evidence for this role *in vitro* [54]. The aspartic proteinases therefore

could play a role in processing and degrading proteins in the guard cells, similar to that proposed for other tissues. A model for the opening of guard cells, which occurs during the light phase of growth, involves the fusion of small vesicles to reform the large central vacuole [55]. As the aspartic proteinases are found in the vacuoles of seeds [22], and at least one of them is induced by light (this work), the appearance of the messages for these genes in guard cells may be consistent with a general increase in vacuolar enzymes when guard cells open during the light phase of growth. These enzymes may be needed to breakdown specific proteins for the appropriate closure of the stomata.

Thus, it is clear there are multiple genes for aspartic proteinases in *Arabidopsis*, with some tissues expressing one gene and others expressing multiple genes. But, it is not yet clear why some tissues require multiple aspartic proteinases in the same cells. One reason a cell might contain multiple proteolytic enzymes of the same class is to segregate them in different compartments. Plant aspartic proteinases have been found in extracellular compartments and in storage and lytic vacuoles in plants [22,56] (reviewed in [3]). Another possible explanation for the appearance of multiple aspartic proteinases in the same cell is that these enzymes have different substrate or amino acid bond specificities that would alter the action of the enzyme on different proteins. Cardosins A and B are 73% identical yet have significant differences in the cleavage of the same protein and peptide substrates [47,57,58]. Our future research will focus on any biochemical or localization differences between the three *Arabidopsis* aspartic proteinases and the identification of protein products from the 37 other aspartic-proteinase-like genes that may reveal why such a simple plant would have so many similar genes.

ACKNOWLEDGEMENTS

Financial support for this research was obtained through a US National Science Foundation CAREER award to S. G. (IBN 9506195). The authors wish to thank Mr Matthew Nichols for assistance with the plant material, Dr Matthew Parker for help with the phylogenetic analysis, Drs Anna Tan-Wilson and Karl Wilson for the loan of a plant incubator, and Dr Lawrence Smart for critical reading of the manuscript.

REFERENCES

1. Davies, D.R. (1990) The structure and function of the aspartic proteinases. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 189–215.
2. Rawlings, N.D. & Barret, A.J. (1995) Families of aspartic peptidases, and those of unknown catalytic mechanism. *Methods Enzymol.* **248**, 105–120.
3. Mutlu, A. & Gal, S. (1999) Plant aspartic proteinases: enzymes on the way to a function. *Physiol. Plant.* **105**, 569–576.
4. 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. Plant.* **106**, 164–169.
5. Buchanan-Wollaston, V. & Ainsworth, C. (1997) Leaf senescence in *Brassica napus*: cloning of senescence related genes by subtractive hybridisation. *Plant Mol. Biol.* **33**, 821–834.
6. 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.
7. 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.
8. Xia, Y., Borevitz, J., Blount, J., Zejian, G., Dixon, R. & Lamb, C. (1999) Activation tagging approach for identification of genes involved in plant disease resistance response. Talk presented at the meeting of the American Society of Plant Biologists, Baltimore, MD USA.
9. Vaccaro, A.M., Tatti, M., Ciaffoni, F., Salvioli, R., Maras, B. & Barca, A. (1993) Function of saposin C in the reconstitution of glucosylceramidase by phosphatidylserine liposomes. *FEBS Lett.* **336**, 159–161.
10. Guruprasad, K., Törmäkangas, K., Kervinen, J. & Blundell, T.L. (1994) Comparative modelling of barley-grain aspartic proteinase: a structural nationale for observed hydrolytic specificity. *FEBS Lett.* **352**, 131–136.
11. Ramalho-Santos, M., Veríssimo, P., Cortes, L., Samyn, B., Van Beeumen, J., Pires, E. & Faro, C.J. (1998) Identification and proteolytic processing of procardosin A. *Eur. J. Biochem.* **255**, 133–138.
12. 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. *Eur J. Biochem.* **267**, 5115–5122.
13. Zhu, Y. & Conner, G.E. (1994) Intermolecular association of lysosomal protein precursors during biosynthesis. *J. Biol. Chem.* **269**, 3846–3851.
14. 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.
15. 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.
16. Törmäkangas, K., Handlington, J.L., Pimpl, P., Hillmer, S., Brandizzi, F., Teeri, T. & Denecke, J. (2001) A vacuolar sorting domain may also influence the way in which proteins leave the endoplasmic reticulum. *Plant Cell* **13**, 2021–2032.
17. 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.
18. Runeberg-Roos, P., Törmäkangas, K. & Östman, A. (1991) Primary structure of a barley grain aspartic proteinase resembling mammalian cathepsin D. *Eur. J. Biochem.* **202**, 1021–1027.
19. Sarkkinen, P., Kalkkinen, N., Tilgmann, C., Siuro, J., Kervinen, J. & Micola, L. (1992) Aspartic proteinase from barley grains is related to mammalian lysosomal cathepsin D. *Planta* **186**, 317–323.
20. Faro, C., Ramalho-Santos, M., Veríssimo, P., Pissarra, J., Frazão, C., Costa, J., Lin, X.-L., Tang, J. & Pires, E. (1998) Structural and functional aspects of cardosins. *Adv. Exp. Med. Biol.* **436**, 423–433.
21. Mutlu, A., Pfeil, J.E. & Gal, S. (1998) A probarley lectin processing enzyme purified from *Arabidopsis thaliana* seeds. *Phytochem.* **47**, 1453–1459.
22. Mutlu, A., Chen, X., Reddy, S.R. & Gal, S. (1999) The seed aspartic proteinase is expressed in many cell types and the protein is found in the vacuoles. *Seed Sci. Res.* **9**, 75–84.
23. 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.
24. Brendel, V. & Kleffe, J. (1998) Prediction of locally optimal splice sites in plant pre-mRNA with applications to gene identification in

Arabidopsis thaliana genomic DNA. *Nuc. Acids Res.* **26**, 4748–4757 (http://www.zmdb.iastate.edu).

25. Drews, G.N. & Okamuro, J.K. (1996) *In situ* hybridisation with nonradioactive probes. Cold Spring Harbor *Arabidopsis* Molecular Genetics Course, Cold Spring Harbor, NY USA.
26. Yu, L., Nasrallah, J., Valenta, R. & Parthasarathy, M.V. (1998) Molecular cloning and mRNA localization of tomato pollen profilin. *Plant Mol. Biol.* **36**, 699–707.
27. Schultz, J., Milpetz, F., Bork, P. & Ponting, C.P. (1998) SMART, a simple modular architecture research tool: Identification of signalling domains. *Proc. Natl Acad. Sci. USA* **95**, 5857–5864.
28. 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.
29. Redecker, B., Heckendorf, B., Grosch, H.-W., Mersmann, G. & Hasilik, A. (1991) Molecular organization of the human cathepsin D gene. *DNA Cell. Biol.* **10**, 423–431.
30. Geier, G., Banaj, H.-J., Heid, H., Bini, L., Pallini, V. & Zwilling, R. (1999) Aspartyl proteases in *Caenorhabditis elegans*: Isolation, identification and characterization by a combined use of affinity chromatography, two-dimensional gel electrophoresis, microsequencing and databank analysis. *Eur. J. Biochem.* **264**, 872–879.
31. Vassar, R., Bennett, B.D., Babu-Khan, S., Kahn, S., Mendiaz, E.A., Denis, P., Teplow, D.B., Ross, S., Amarante, P., Loeloff, R., Luo, Y., Fisher, S., Fuller, J., Edenson, S., Lile, J., Jarosinski, M.A., Biere, A.L., Curran, E., Burgess, T., Louis, J.-C., Collins, F., Treanor, J., Rogers, G. & Citron, M. (1999) β-Secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* **286**, 735–741.
32. Costa, J., Ashford, D.A., Nimtz, M., Bento, I., Frazão, C., Esteves, C.L., Faro, C.J., Kervinen, J., Pires, E., Veríssimo, P., Wlodawer, A. & Carrondo, M.A. (1997) The glycosylation of the aspartic proteinases from barley (*Hordeum vulgare* L.) and cardoon (*Cynara cardunculus* L.). *Eur. J. Biochem.* **243**, 695–700.
33. Hori, H., Yoshino, T., Ishizuka, Y., Yamauchi, T. & Murakami, K. (1988) Role of N-linked oligosaccharides attached to human renin expressed in COS cells. *FEBS Lett.* **232**, 391–394.
34. Mutlu, A. (1997) Purification, characterization and localization of a proBL processing enzyme from *Arabidopsis thaliana*, PhD Thesis, Binghamton University, Binghamton, NY USA.
35. Chen, X. (1998) Aspartic proteinase gene expression in *Arabidopsis thaliana.* Master's Thesis, State University of New York – Binghamton, Binghamton, NY USA.
36. Desprez, T., Amselem, J., Caboche, M. & Höfte, H. (1998) Differential gene expression in Arabidopsis monitored using cDNA arrays. *Plant J.* **14**, 643–652.
37. Törmäkangas, K., Kervinen, J., Östman, A. & Teeri, T. (1994) Tissue-specific localization of aspartic proteiase in developing and germinating barley grains. *Planta* **195**, 116–125.
38. 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.
39. Schaller, A. & Ryan, C. (1996) Molecular cloning of a tomato leaf cDNA encoing an aspartic protease, a systemic wound response protein. *Plant Mol. Biol.* **31**, 1073–1077.
40. 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.
41. Fujikara, Y. & Karssen, C.M. (1995) Molecular studies on osmoprimed seeds of cauliflower: a partial amino acid sequence of a vigour related protein and osmopriming enhanced expression of putative aspartic protease. *Seed Sci. Res.* **5**, 177–181.
42. Pang, P.P., Pruitt, R.E. & Meyerowitz, E.M. (1988) Molecular cloning, genomic organization, expression and evolution of 12S seed storage protein genes of *Arabidopsis thaliana*. *Plant Mol. Biol.* **11**, 805–820.
43. Conceição, A.D.S. & Krebbers, E. (1994) A cotyledon regulatory region is responsible for the different spatial expression patterns of Arabidopsis 2S albumin genes. *Plant J.* **5**, 493–505.
44. 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 plant. *Eur. J. Biochem.* **246**, 133–141.
45. Runeberg-Roos, P., Kervinen, J., Kovaleva, V., Raikel, 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.
46. Ramalho-Santos, M., Pissarra, J., Veríssimo, P., Pereira, S., Salema, R., Pires, E. & Faro, C. (1997) Cardosin A, an abundant aspartic proteinase, accumulates in protein storage vacuoles in the stigmatic papillae of *Cynara cardunculus* L. *Planta* **203**, 204–212.
47. Vieira, M., Pissarra, J., Veríssimo, Castanheira, P., Costa, Y., Pires, E. & Faro, C. (2001) Molecular cloning and characterization of cDNA encoding cardosin B, an aspartic proteinase accumlating extracellularly in the transmitting tissue of *Cynara cardunculus* L. *Plant Mol. Biol.* **45**, 529–539.
48. Faro, C., Ramalho-Santos, M., Vieira, M., Mendes, A., Simões, I., Andrade, R., Veríssimo, P., Lin, W., 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.
49. Assmann, S.M. & Shimazaki, K. (1999) The multisensory guard cell: stomatal responses to blue light and abscisic acid. *Plant Physiol.* **119**, 809–815.
50. Blanke, M.M. & Lovatt, C.J. (1993) Anatomy and transpiration of the avocado inflorescence. *Ann. Bot.* **71**, 543–547.
51. Vemmos, S.N. & Goldwin, G.K. (1993) Stomatal and chlorophyll distribution of Cox's orange Pippin apple flowers relative to other cluster parts. *Ann. Bot.* **71**, 245–250.
52. Clément, C., Mischler, P., Burrus, M. & Audran, J.-C. (1997) Characteristics of the photosynthetic apparatus and CO_2- fixation in the flower bud of *Lilium*. I. corolla. *Int. J. Plant Sci.* **158**, 794–800.
53. Berger, D. & Altmann, T. (2000) A subtilisin-like serine protease involved in regulation of stomatal density and distribution in *Arabidopsis thaliana*. *Genes Dev.* **14**, 1119–1131.
54. Amidon, W.L., Pfeil, J.E. & Gal, S. (1999) Modification of luciferase to be a plant aspartic proteinase substrate. *Biochem. J.* **343**, 425–433.
55. Blatt, M.R. (2000) Cellular signaling and volume control in stomatal movements in plants. *Annu. Rev. Cell Dev Biol.* **16**, 221–241.
56. Paris, N., Stanley, C.M., Jones, R.L. & Rogers, J.C. (1996) Plant cells contain two functionally distinct vacuolar compartments. *Cell* **85**, 563–572.
57. Ramalho-Santos, M., Veríssimo, P., Faro, C. & Pires, E. (1996) Action on bovine α_{S1}-casein of cardosins A and B, aspartic proteinases from the flowers of the cardoon *Cynara cardunculus* L. *Biochim. Biophys. Acta.* **1297**, 83–89.
58. Veríssimo, P., Faro, C., Moir, A.J.G., Lin, Y., Tang, J. & Pires, E. (1996) Purification, characterization and partial amino sequencing of two new aspartic proteinases from fresh flowers of *Cynara cardunculus* L. *Eur. J. Biochem.* **235**, 762–768.
59. 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.