17. Phytepsin

Databanks

MEROPS name: Phytepsin
MEROPS classification: clan AA, family A1, peptidase A01.020
IUBMB: EC 3.4.23.40
CAS registry: 78169-47-8
Species distribution: Magnoliophyta
Sequence known from: *Arabidopsis thaliana, Brassica napus, Brassica oleracea, Centaurea calcitrapa, Cicer arietinum, Cucurbita pepo, Cynara cardunculus, Glycine max, Helianthus annuus, Hemerocallis sp., Hordeum vulgare, Ipomoea batatas, Lycopersicon esculentum, Manihot esculenta, Nepenthes alata, Oryza sativa, Pyrus pyrifolia, Theobroma cacao, Vigna unguiculata*
Tertiary structure: Available

Name and History

Aspartic proteinases (APs) are ubiquitously expressed throughout the plant kingdom. All plants appear to contain the enzyme since its presence is observed in at least some tissues of every plant so far tested. In 1997, the name ***phytepsin*** (L. *phyto*, plant) was introduced to denote all related plant APs and was adopted by NC-IUBMB. However, in the literature, no universal name is commonly used to represent all plant APs. Instead, the names tend to indicate the plant species or tissue from which the specific enzyme is purified. In this review, the name phytepsin is used to denote plant APs in general and specific names are given where applicable.

Phytepsins have been especially well characterized from cardoon (*Cynara cardunculus* L.) and barley (*Hordeum vulgare* L.) plant tissues. Flowers of the cardoon plant contain a milk-clotting activity that has been exploited in traditional cheese-making in Portugal and Spain for centuries. This activity was eventually found to be due to a family of APs. Three phytepsins, ***cynarases*** 1–3, were purified from cardoon flowers and partially characterized (Heimgartner *et al.*, 1990; Cordeiro *et al.*, 1992). Later these phytepsins were renamed ***cyprosins*** 1, 2 and 3 (cynara protease with the ending -sin) and the primary structure of cyprosin was determined (Cordeiro *et al.*, 1994). Two other phytepsins, named cardosin A and cardosin B (*cardo* is the Portuguese name of the plant), have subsequently been extensively studied from cardoon flowers (Veríssimo *et al.*, 1996; Ramalho-Santos *et al.*, 1997, 1998; Faro *et al.*, 1999; Egas *et al.*, 2000; Vieira *et al.*, 2001; Barros & Malcata, 2002) and the crystal structure of cardosin A has been solved (Frazão *et al.*, 1999). Barley grains contain a prominent phytepsin activity (Morris *et al.*, 1985; Sarkkinen *et al.*, 1992; Wrobel & Jones, 1992) suggested to initiate the hydrolysis of grain storage proteins at the onset of germination (Mikola, 1987). Barley grain AP was the first phytepsin for which the primary structure was determined (Runeberg-Roos *et al.*, 1991). The corresponding enzyme, abbreviated HvAP, was purified from resting barley grains and characterized (Sarkkinen *et al.*, 1992; Kervinen *et al.*, 1993; Runeberg-Roos *et al.*, 1994; Törmäkangas *et al.*, 1994; Costa *et al.*, 1997). Its modification along the route to vacuoles has been studied (Glathe *et al.*, 1998) and the crystal structure of the zymogen form has been solved (Kervinen *et al.*, 1999).

Based on these and other studies, it became apparent that phytepsins share high sequence and structure homology, as well as general biochemical properties, with the animal and microbial APs. However, many phytepsins contain an extra protein sequence of about 100 amino acids, referred to as a plant-specific insert (PSI) or domain (e.g. Mutlu & Gal, 1999). This region has no sequence homology to mammalian or microbial APs and thus the PSI region can be regarded as a unique structural feature of phytepsins. Interestingly, this structure exhibits close homology to members of the saposin-like family of proteins, a group of lipid-binding proteins found in various organisms (Guruprasad *et al.*, 1994; Liepinsh *et al.*, 1997; Schuette *et al.*, 2001). Several theories have been proposed regarding the function of the PSI region (see below).

Activity and Specificity

The activity of phytepsin isolated from various barley tissues has been measured both as a pepstatin-sensitive protein cleavage at pH 3.7, using hemoglobin as a substrate (Sarkkinen *et al.*, 1992; Törmäkangas *et al.*, 1994), and by an electrophoretic method using a native gel with immobilized edestin (Wrobel & Jones, 1992; Zhang & Jones, 1995). Purified phytepsin hydrolyzes hemoglobin and a chromophoric substrate, Pro-Thr-Glu-Phe┼Nph-Arg-Leu (NovaBiochem; see Appendix 2 for full names and addresses of suppliers), optimally at pH 3.5–4.1 (Sarkkinen *et al.*, 1992; Kervinen *et al.*, 1993). Insulin B chain, glucagon and melittin have been used for characterization of the hydrolytic specificity of phytepsin. The cleavage of insulin B chain is as follows:

FVNQHLCGSHL┼VEA┼L┼YLVCGERGF┼F┼YTPKA

The cleavage typically occurred between two residues with hydrophobic side chains (Leu, Ile, Val, Phe) or next to one hydrophobic residue. In glucagon, the Asp┼Tyr bond was also readily cleaved (Kervinen *et al.*, 1993):

HSQGTFTSD┼YSKY┼L┼DS┼RR┼AQDF┼VQW┼L┼MNT

Similar hydrolytic specificity has been observed for several other phytepsins, although some slight variation occurs between individual enzymes (Polanowski *et al.*, 1985; Faro *et al.*, 1995; Ramalho-Santos *et al.*, 1996; Bleukx *et al.*, 1998). Besides hemoglobin, other *in vitro* substrates of phytepsins include albumin, gliadin, chromophoric peptides, casein (Mutlu & Gal, 1999), and firefly luciferase (Amidon *et al.*, 1999).

Barley phytepsin is inhibited by pepstatin and several other substrate–analog inhibitors which are effective against the mammalian lysosomal AP cathepsin D (Chapter 8) (Sarkkinen *et al.*, 1992). A systematic series of synthetic inhibitors developed against mammalian and retroviral APs have been tested against the recombinantly produced cardoon enzyme (cyprosin); several inhibited cyprosin with measured K_i values in the nanomolar range (White *et al.*, 1999). Endogenous AP inhibitors have been detected or purified from several plant species including potato (Mareš *et al.*, 1989; Maganja *et al.*, 1992), tomato (Hansen & Hannapel, 1992; Cater *et al.*, 2002), squash (Christeller *et al.*, 1998) and *Anchusa strigosa* (Abuereish, 1998). These inhibitors were effective against mammalian APs such as cathepsin D, pepsin and yeast proteinase A. Interestingly, however, AP inhibitors from potato and tomato did not inhibit the cardoon phytepsins cyprosin (White *et al.*, 1999) and cardosin B (Cater *et al.*, 2002), respectively. The efficacy of these inhibitors against other phytepsins is unknown.

Structural Chemistry

Primary Structure

The overall primary translation product of most phytepsins is similar to their mammalian and microbial counterparts, consisting of the endoplasmic reticulum (ER) signal sequence followed by a self-inhibition peptide of around 40 residues (propeptide) and the mature enzyme sequence (e.g. Runeberg-Roos *et al.*, 1991; Cordeiro *et al.*, 1994; Hiraiwa *et al.*, 1997; Faro *et al.*, 1999; Domingos *et al.*, 2000; Vieira *et al.*, 2001). The mature enzyme can be roughly divided into N-terminal and C-terminal sequence domains, based on the bilobal tertiary structure of the enzyme. The primary structure of phytepsins contains several typical sequence domains found in all APs. These include the conserved catalytic active-site residues Asp-Thr-Gly and Asp-Ser-Gly in the N-terminal and C-terminal regions, respectively, although most mammalian and microbial enzymes contain the Asp-Thr-Gly sequence in both domains. Whether the Asp-Thr/Ser-Gly variation confers any biological significance remains to be determined. A unique feature among many phytepsins is an extra sequence of around 100 residues (plant-specific insert; PSI) that is located in the C-terminal region of the protein. Due to the PSI sequence, the primary translation product of phytepsins (~500 amino acids) is significantly larger than those of most mammalian enzymes. A single conserved and functional *N*-glycosylation site is also present in the PSI region (Costa *et al.*, 1997; Frazão *et al.*, 1999).

Although most phytepsins studied in detail contain close sequence homology, some recent results reveal that APs represent a very diverse class of enzymes in the plant kingdom. For instance, tobacco chloroplast nucleoids contain a phytepsin-like enzyme with a fully conserved active-site region, but the cDNA encoding the protein does not include either a PSI region or the propeptide sequence (Nakano *et al.*, 1997). Likewise, Chen & Foolad (1997) have reported the cloning and characterization of a phytepsin-like proteinase of around 45 kDa (nucellin) from barley ovaries, which is abundantly expressed after pollination. The cDNA contains a likely signal sequence, but lacks both the propeptide and the PSI region. A seven-member family of glycosylphosphatidylinositol (GPI)-anchored APs has been identified from the *Arabidopsis thaliana* genome (Borner *et al.*, 2002). In all, the *Arabidopsis* genome contains at least 66 putative phytepsin genes. Sixty of them contain a fully conserved active-site sequence, whereas six members are probably inactive due to a mutated active-site Asp residue. In terms of their primary structures, *Arabidopsis* phytepsins form a diverse group of enzymes. Interestingly, only four of the putative *Arabidopsis* phytepsins contain a PSI region (Yiji Xia, personal communication).

Tertiary Structure

Crystal structures have only been solved for cardoon (Frazão *et al.*, 1999) and barley (Kervinen *et al.*, 1999) phytepsins; both enzymes exhibit structures typical for the AP family. The crystal structure of a zymogen form of barley phytepsin, solved at 2.3 Å resolution, is presented in Figure 17.1. The model includes residues 6–26 out of the 41 residue sequence of the propeptide, the mature protein, and most of the PSI region. The structure of the mature enzyme consists of two similar β barrel domains with the two active-site residues Asp36 and Asp223 (illustrated in ball-and-stick in Figure 17.1) located in the interdomain cleft. The bottom of the active-site cleft is covered with a six strand β sheet shielding the internal hydrophobic core from the exterior.

The self-inactivation mechanism used by barley prophytepsin, and apparently also by all closely related phytepsins,

varies from that of the well-known gastric AP zymogen pepsinogen (Bernstein & James, 1999; Kervinen *et al.*, 1999). The phytepsin propeptide wraps around the mature protein and its N-terminal strand is involved in the formation of the inter-domain six-stranded β sheet. The helical part of the propeptide is located in the vicinity of the active site and, together with the first 13 residues of the N-terminus of the mature enzyme, fully blocks it. This conformation is anchored to the active site both by an ionic interaction of Lys11 (see Figure 17.1) with the active-site Asp residues and by the interactions of other residues in the S1′ and S1 sites.

The crystal structure of cardoon enzyme (cardosin A) has been determined to a 1.7 Å resolution (Frazão *et al.*, 1999). Comparison of the mature cardosin A structure to a zymogen form of barley phytepsin suggests that only minor conformational changes occur during activation of plant phytepsins (Kervinen *et al.*, 1999). Cardosin A has two *N*-glycosylation sites (Asn67 and Asn257) and the crystal structure reveals a new glycan of the plant complex type. Due to their localization away from the active site, the glycans are likely to be important for the stability and proper processing of the precursor rather than for activity. The unique feature of cardosin A among phytepsins is the presence of the Arg-Gly-Asp (RGD) cell-attachment motif. The crystal structure shows that the RGD motif is located at the base of the molecule, opposite to the active site, and projects itself out of the molecular surface. This may indicate a protein–protein interaction motif for this area.

Barley phytepsin structure revealed the fold and orientation of the PSI in relation to the mature part of the protein (Figure 17.1). It is attached to the C-terminal domain of the enzyme by two flexible polypeptide linkages, permitting some freedom in the positioning of PSI relative to the main body of the enzyme. The PSI domain comprises five amphipathic helices forming a helical bundle with a large internal hydrophobic core. In addition to being stabilized by hydrophobic interactions, the tertiary structure of PSI is maintained by three disulfide bridges (Kervinen *et al.*, 1999).

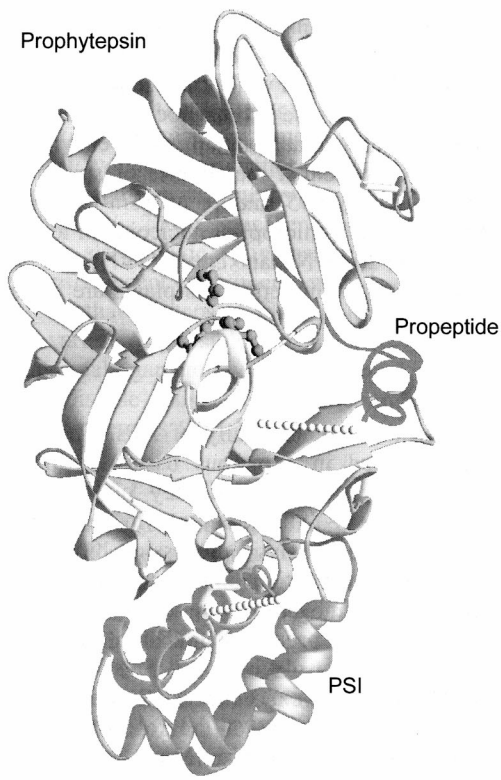

Figure 17.1 The crystal structure of a zymogen form of barley phytepsin. Propeptide and PSI region are shown in dark gray and a mature enzyme in gray. Lys11 as well as the active-site residues Asp36 and Asp223 are shown in ball-and-stick.

Preparation

In most cases, phytepsins have been purified from crude extracts by affinity chromatography using immobilized pepstatin as an affinity matrix (Kervinen *et al.*, 1995; Mutlu & Gal, 1999). However, the expression level of phytepsins in plant tissues is generally low and in most cases purification procedures have yielded at most a milligram amount of pure protein. In addition, the preparations often contain several enzyme forms or processing intermediates, further hindering detailed analyses. Bacterial and eukaryotic cell expression systems have been tested for some phytepsins. For example, the proform of the rice enzyme, oryzasin 1, has been expressed in *E. coli* as a fusion protein linked to glutathione-*S*-transferase (GST). The resulting purified enzyme underwent autocatalytic activation and showed proteolytic activity using hemoglobin as a substrate (Asakura *et al.*, 2000). However, a general problem with bacterial expression methods has been the very low yield of the correctly folded product (Glathe *et al.*, 1998; White *et al.*, 1999). Expression systems utilizing eukaryotic cells for the production of phytepsins have been used more successfully. For example, the cDNA encoding the precursor of a phytepsin (cyprosin) from cardoon flowers was expressed in *Pichia pastoris* cells and active enzyme was purified from the culture media (White *et al.*, 1999). The baculovirus-infected insect cell expression method was used to obtain barley prophytepsin for crystallographic and other purposes (Glathe *et al.*, 1998; Kervinen *et al.*, 1999). It is interesting to note that cyprosin was purified from the cell culture media in its fully processed and active form, whereas barley phytepsin was produced in its zymogen form (without the ER-signal sequence). It is not known if the observed difference in the processing state of cardoon and barley phytepsins was due to the use of different expression systems or is indicative of the divergent processing mechanisms these enzymes undergo during activation. Nevertheless, these two examples demonstrate the usefulness of eukaryotic expression systems for obtaining correctly folded and fully functional phytepsins in reasonable quantities. Hopefully, these methods will be used in the future to yield protein for the study of other phytepsin family members.

Biological Aspects

Proenzyme Activation and Transport

Phytepsins are expressed in their zymogen form and undergo N-terminal processing during activation. Some phytepsins

undergo further internal processing to achieve the mature enzyme. The PSI region is often removed during this process, resulting in the formation of a two-chain enzyme (Veríssimo *et al.*, 1996; Glathe *et al.*, 1998; Ramalho-Santos *et al.*, 1998; White *et al.*, 1999; Park *et al.*, 2001). The reason some phytepsins exist as monomers and the others as two-chain forms is not known. Both autocatalytic and heterocatalytic steps are needed for the activation of phytepsins. These mechanisms presumably depend both on the prevailing pH as well as on processing proteases present in the particular intracellular compartments traversed by the phytepsin precursor. Recombinantly produced barley phytepsin zymogen undergoes autoproteolytic activation at pH $\leq$ 4.5. However, a comparison of the *in vitro*/*in vivo* processing sites revealed slight differences, indicating that additional proteases are needed for completion of maturation *in vivo* (Glathe *et al.*, 1998). Cardoon phytepsin (procardosin A) undergoes proteolytic processing as the flower matures. *In vitro* studies on procardosin A have shown that the PSI region is probably removed through an autocatalytic mechanism, whereas processing of the propeptide results from an action of a cysteine or serine proteinase (Veríssimo *et al.*, 1996; Ramalho-Santos *et al.*, 1998).

The intracellular transport route has been studied for barley phytepsin (Glathe *et al.*, 1998; Törmäkangas *et al.*, 2001). The primary translation product is synthesized and then is translocated into the rough ER, where an ER-signal sequence (prepeptide) is removed and the proenzyme becomes *N*-glycosylated. Phytepsin apparently uses a COPII vesicle-mediated transport mechanism to exit the ER on the way toward the Golgi apparatus (Törmäkangas *et al.*, 2001). At the Golgi, the glycans are modified to a final complex type and the proenzyme is packed into a transport vesicle destined for vacuoles. Upon arrival in the vacuole, the acid pH probably triggers autocatalytic activation of prophytepsin that includes proteolytic cleavage of the propeptide and removal of the PSI region. The protein is also slightly modified by other proteases to yield the final mature enzyme (Glathe *et al.*, 1998).

Several sequence regions may play a role as targeting signals to direct phytepsin family members to vacuoles or to exit the cell. For barley phytepsin, the suggested sequence regions include the NPLR-peptide in the self-inhibitory peptide immediately prior to the N-terminus of the mature enzyme and some areas in the PSI domain (Guruprasad *et al.*, 1994; Paris *et al.*, 1996; Kervinen *et al.*, 1999; Törmäkangas *et al.*, 2001). The C-terminal peptide region has also been suggested to function as an intracellular targeting signal sequence for phytepsins (Ramalho-Santos *et al.*, 1998). However, the exact intracellular targeting signal sequences and possible interactions of the phytepsin proenzymes with other cellular proteins during transport are still largely hypothetical.

Distribution in Organisms and Tissues

Several phytepsins have been purified from seed material. However, phytepsins are widely expressed in various tissues and organs. Most of the enzymes accumulate in vacuoles or are secreted into the extracellular space. In cardoon, cynarases as well as cardosins A and B are present in the floral organs, pistils. These proteinases seem to be strictly floral specific and their expression is developmentally regulated. Appearance of mRNA encoding cyprosin starts in early stages of floral development and switches off at maturation of the flower (Cordeiro *et al.*, 1994). Closely related cardosin A accumulates in protein storage vacuoles of the stigmatic epidermal papillae and in vacuoles of the epidermal cells of the style (Ramalho-Santos *et al.*, 1997), whereas cardosin B accumulates in the cell wall and in the extracellular matrix of the transmitting tissue (Vieira *et al.*, 2001). This indicates that although cardoon enzymes are homologous in their primary structures, they occupy distinct cellular locations in the pistils.

Barley phytepsin is expressed in several tissues of the grain (mainly in embryo, scutellum and aleurone layer), stem, leaves, flowers, and roots (Törmäkangas *et al.*, 1994). The enzyme has been localized by immunocytochemistry to the scutellar and aleuronal protein storage vacuoles in grains (Marttila *et al.*, 1995; Bethke *et al.*, 1996) as well as to leaf and root cell vacuoles (Runeberg-Roos *et al.*, 1994; Paris *et al.*, 1996). Within barley roots, phytepsin is abundantly present in both the developing tracheary elements and sieve cells (Runeberg-Roos & Saarma, 1998). Electrophoretically, four AP activities with different mobilities have been detected from grain extracts (Zhang & Jones, 1995) although it is not known if these enzymes are derived from differential processing of the same precursor. A putative viral-type AP is also encoded by part of the BARE-1 retroelement in the barley genome (Vicient *et al.*, 1999).

Two secreted phytepsins have been purified from tobacco and tomato leaves (Rodrigo *et al.*, 1989, 1991). An AP has also been located in the cell wall of maize pollen (Radlowski *et al.*, 1996). Other diverse cellular locations for phytepsins include the observation that a phytepsin-like enzyme has been localized to photosystem II membranes of spinach plastids (Kuwabara & Suzuki, 1995) and a phytepsin-like enzyme has been purified from the chloroplast nucleoids of tobacco cells (Murakami *et al.*, 2000; Nakano *et al.*, 1997). APs present in the trapping organs of insectivorous plants are described in Chapter 18.

Function

Phytepsins form a diverse family of APs. These enzymes have been observed in numerous tissues of many plant types and thus it is likely that they play an important role in multiple proteolytic processes in cells and in the extracellular space. Proteolytic processing by phytepsins is likely to include modification of storage proteins, cleavage of proproteins and zymogens, tissue degradation to facilitate pollen growth, cellular protein degradation during apoptotic processes, as well as playing a defensive role against pathogens. However, the exact function of phytepsins is still obscure. A thorough review of the role of phytepsins in various organisms can be found in Mutlu & Gal (1999) and so only some specific functions are briefly described here.

Phytepsins have been found in large quantities in various germinating seeds and several purified enzymes have been demonstrated to be capable of degrading seed storage proteins *in vitro* (D'Hondt *et al.*, 1993; Voigt *et al.*, 1995; Hiraiwa *et al.*, 1997; Bleukx *et al.*, 1998). Thus these

enzymes may directly participate in storage protein modification and degradation or they may process and activate other hydrolases during seed development and germination. In barley, widespread tissue distribution of phytepsin throughout vegetative and reproductive tissues as well as the enzyme's intracellular location in hydrolytic organelles, vacuoles, suggest that phytepsin performs numerous functions in protein processing and turnover in different stages of the plant life cycle (Runeberg-Roos *et al.*, 1994; Törmäkangas *et al.*, 1994; Paris *et al.*, 1996). Moreover, phytepsins have recently been associated with programmed cell death (apoptosis) in the developing tracheary elements of barley roots (Runeberg-Roos & Saarma, 1998; Lindholm *et al.*, 2000) and in the degenerating nucellar cells of the barley ovule (Chen & Foolad, 1997).

Phytepsins in cardoon floral organs likely have several roles during flower development and reproductive events. For example, cardosin A contains an RGD cell attachment motif and it has been suggested that cardosin A has a role in adhesion-mediated proteolytic mechanisms involved in pollen recognition and growth (Faro *et al.*, 1999). Cardosin B accumulates in the cell wall and in the extracellular matrix of the transmitting tissue and thus it may have a role in defensive mechanisms against pathogens (Vieira *et al.*, 2001). Likewise, an extracellular phytepsin purified from the leaves of tomato and tobacco degrades secreted pathogenesis-related proteins *in vitro*. The sensitivity of pathogenesis-related proteins to phytepsin activity and the co-distribution of the enzyme together with pathogenesis-related proteins indicate that the biological action of these defensive proteins might be regulated by phytepsins (Rodrigo *et al.*, 1989, 1991). Phytepsin activity has been found to be highly elevated in certain stress situations. For example, expression of the tomato enzyme is systemically induced in leaves by wounding (Schaller & Ryan, 1996), as well as expression of bean enzyme by drought conditions (Cruz de Carvalho *et al.*, 2001). Senescence-associated genes from daylily petals also include a phytepsin (Panavas *et al.*, 1999). In *Arabidopsis*, recent experiments with developed knockout lines suggest that phytepsins play multiple important roles in a wide variety of physiological processes. These include embryo and male gametophyte development, pollen–stigma interaction, disease resistance response, light response and chloroplast biogenesis (Yiji Xia, personal communication).

Distinguishing Features

Many phytepsins contain an internal region of about 100 residues (plant-specific insert; PSI) that is curiously similar to saposins, sphingolipid-activating proteins found in mammalian cells. Six conserved Cys residues, a glycosylation site, a hydrophobicity pattern, and conserved Pro and Tyr/Phe residues found in the PSI domain are analogous to those of saposins (Guruprasad *et al.*, 1994; Ponting & Russell, 1995; Schuette *et al.*, 2001). The tertiary structure of the PSI region (Kervinen *et al.*, 1999) also closely resembles that of the saposin-like proteins NK-lysin (Liepinsh *et al.*, 1997) and lung surfactant peptide SP-B (Kurutz & Lee, 2002). Interestingly, the sequence of the PSI region in phytepsins is equivalent to the sequence of interchanged N- and C-terminal domains in saposin. However, their circular permutation (swapping) does not affect the orientation of helices or overall globular structure of the PSI region (Kervinen *et al.*, 1999). A recent report describes the first circularly permutated saposin motif from an animal source, J3-crystallin of the jellyfish lens (Piatigorsky *et al.*, 2001).

Saposins interact with lipids and promote sphingolipid degradation in animal cells, among other functions (e.g. Vaccaro *et al.*, 1999; Qi & Grabowski, 2001; Schuette *et al.*, 2001). Similarly, a putative membrane-binding area has been identified in the PSI region based on the crystal structure of barley phytepsin (Kervinen *et al.*, 1999). Thus, the PSI region has been proposed to function as a vacuolar targeting signal by bringing prophytepsin into contact with membranes and possibly also with a membrane-bound receptor in the Golgi apparatus. This would promote a formation of phytepsin-containing transport vesicles to release from Golgi destined to vacuoles. It remains to be determined if this hypothesis is correct. However, it has been shown that deletion of the PSI region led to the secretion of truncated barley phytepsin in tobacco cells, whereas the intact protein was transported into vacuoles (Törmäkangas *et al.*, 2001).

The PSI region may play a defensive role against pathogenic attack. Egas *et al.* (2000) demonstrated that the proform of cardosin A is able to interact with phospholipid-containing membrane vesicles *in vitro* whereas mutant cardosin A lacking the PSI region did not have this ability. This interaction of the PSI region with vesicular membranes induced leakage of vesicular contents. Thus, the authors suggested that the PSI region may function as a defensive weapon by causing membrane damage to the invading pathogen. Additionally, phytepsins are highly expressed in some senescing or apoptotic tissues. That phytepsins can induce membrane damage to their own cells during late autolysis and thereby promote programmed cell death is another interesting hypothesis suggested by Egas *et al.* (2000) that remains to be tested.

Some studies point to the importance of the PSI region in proper processing or folding of the precursor protein. Expression of the rice phytepsin, oryzasin 1, in *E. coli* with or without the PSI region did not affect the enzymatic activity of the processed enzyme. However, the pattern of processing was changed (Asakura *et al.*, 2000). When the cardoon phytepsin cynarase was expressed in *Pichia pastoris* cells without the PSI region, unprocessed and misfolded precursor protein accumulated, suggesting that the PSI region plays a role in ensuring the correct folding and processing of the enzyme (White *et al.*, 1999).

Related Peptidases

Phytepsins have been detected or purified from a diverse array of common and exotic plants. For instance, these enzymes have been characterized from wheat (Belozersky *et al.*, 1989; Galleschi & Felicioli, 1994; Bleukx *et al.*, 1998), *Arabidopsis* (Mutlu *et al.*, 1998, 1999), rice (Doi *et al.*, 1980; Asakura *et al.*, 1995, 2000), sorghum (Garg & Virupaksha, 1970), hemp (St. Angelo & Ory, 1970), buckwheat (Elpidina *et al.*, 1990), lotus (Shinano & Fukushima, 1971), cucumber and squash (Polanowski *et al.*, 1985), rape (D'Hondt *et al.*, 1993), cauliflower (Fujikura & Karssen, 1995), sunflower (Park *et al.*, 2001), cocoa (Voigt *et al.*, 1995), potato (Brierley *et al.*, 1996; Guevara *et al.*, 2001) and pine (Sälmia,

1981; Bourgeois & Malek, 1991). For a more complete list of purified enzymes, see a phytepsin review (Mutlu & Gal, 1999).

Acknowledgments

We thank Yiji Xia (Donald Danforth Plant Research Center, St. Louis, MO, USA) for allowing us to cite his unpublished results on *Arabidopsis* phytepsins. A. Jeannine Lincoln (Thomas Jefferson University, Philadelphia, PA, USA) is warmly thanked for her many valuable comments on the manuscript.

Further Reading

A thorough phytepsin review by Mutlu & Gal (1999) is available. Structural comparison of the activation mechanism of mammalian, microbial and plant APs is presented by Bernstein & James (1999).

References

Abuereish, G.M. (1998) Pepsin inhibitor from roots of *Anchusa strigosa*. *Phytochemistry* **48**, 217–221.

Amidon, W.J., Pfeil, J.E. & Gal, S. (1999) Modification of luciferase to be a substrate for plant aspartic proteinase. *Biochem. J.* **343**, 425–433.

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.

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.

Barros, R.M. & Malcata, F.X. (2002) Modeling the kinetics of whey protein hydrolysis brought about by enzymes from *Cynara cardunculus*. *J. Agric. Food Chem.* **50**, 4347–4356.

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.

Bernstein, N.K. & James, M.N.G. (1999) Novel ways to prevent proteolysis - prophytepsin and proplasmepsin II. *Curr. Opin. Struct. Biol.* **9**, 684–689.

Bethke, P.C., Hillmer, S. & Jones, R.L. (1996) Isolation of intact protein storage vacuoles from barley aleurone. *Plant. Physiol.* **110**, 521–529.

Bleukx, W., Brijs, K., Torrekens, S., Van Leuven, F. & Delcour, J.A. (1998) Specificity of a wheat gluten aspartic proteinase. *Biochim. Biophys. Acta* **1387**, 317–324.

Borner, G.H., Sherrier, D.J., Stevens, T.J., Arkin, I.T. & Dupree, P. (2002) Prediction of glycosylphosphatidylinositol-anchored proteins in *Arabidopsis*. A genomic analysis. *Plant Physiol.* **129**, 486–499.

Bourgeois, J. & Malek, L. (1991) Purification and characterization of an aspartyl proteinase from dry jack pine seeds. *Seed Sci. Res.* **1**, 139–147.

Brierley, E.R., Bonner, P.L.R. & Cobb, A.H. (1996) Factors influencing the free amino acid content of potato (*Solanum tuberosum* L) tubers during prolonged storage. *J. Sci. Food Agric.* **70**, 515–525.

Cater, S.A., Lees, W.E., Hill, J., Brzin, J., Kay, J. & Phylip, L.H. (2002) Aspartic proteinase inhibitors from tomato and potato are more potent against yeast proteinase A than cathepsin D. *Biochim. Biophys. Acta* **1596**, 76–82.

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.

Christeller, J.T., Farley, P.C., Ramsay, R.J., Sullivan, P.A. & Laing, W.A. (1998) Purification, characterization and cloning of an aspartic proteinase inhibitor from squash phloem exudate. *Eur. J. Biochem.* **254**, 160–167.

Cordeiro, M., Jakob, E., Puhan, Z., Pais, M.S. & Brodelius, P.E. (1992) Milk clotting and proteolytic activities of purified cynarases from *Cynara cardunculus* – a comparison to chymosin. *Milchwissenschaft* **47**, 683–687.

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.

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.

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.

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.

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.

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

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.

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.

Faro, C., Veríssimo, P., Lin, Y., Tang, J. & Pires, E. (1995) Cardosin A and B, aspartic proteases from the flowers of cardoon. *Adv. Exp. Med. Biol.* **362**, 373–377.

Faro, C., Ramalho-Santos, M., Vieira, M., Mendes, A., Simoes, I., Andrade, R., Veríssimo, 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.

Frazão, C., Bento, I., Costa, J., Soares, C.M., Veríssimo, 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.

Fujikura, 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.

Galleschi, L. & Felicioli, F. (1994) Purification, characterization and activation by anions of an aspartic proteinase isolated from bran of soft wheat. *Plant Sci.* **98**, 15–24.

Garg, G.K. & Virupaksha, T.K. (1970) Acid protease from germinated sorghum. 2. Substrate specificity with synthetic peptides and ribonuclease A. *Eur. J. Biochem.* **17**, 13–18.

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.

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

Guruprasad, K., Törmäkangas, 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.

Hansen, J.D. & Hannapel, D.J. (1992) A wound-inducible potato proteinase inhibitor gene expressed in non-tuber-bearing species is not sucrose inducible. *Plant Physiol.* **100**, 164–169.

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

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.

Kervinen, J., Sarkkinen, P., Kalkkinen, N., Mikola, L. & Saarma, M. (1993) Hydrolytic specificity of the barley grain aspartic proteinase. *Phytochemistry* **32**, 799–803.

Kervinen, J., Törmäkangas, K., Runeberg-Roos, P., Guruprasad, K., Blundell, T. & Teeri, T.H. (1995) Structure and possible function of aspartic proteinases in barley and other plants. *Adv. Exp. Med. Biol.* **362**, 241–254.

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.

Kurutz, J.W. & Lee, K.Y. (2002) NMR structure of lung surfactant peptide SP-B 11–25. *Biochemistry* **41**, 9627–9636.

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.

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

Lindholm, P., Kuittinen, T., Sorri, O., Guo, D., Merits, A., Törmäkangas, 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.

Maganja, D.B., Strukelj, B., Pungercar, J., Gubensek, F., Turk, V. & Kregar, I. (1992) Isolation and sequence analysis of the genomic DNA fragment encoding an aspartic proteinase inhibitor homologue from potato (*Solanum tuberosum* L.). *Plant Mol. Biol.* **20**, 311–313.

Mareš, M., Meloun, B., Pavlík, M., Kostka, V. & Baudyš, M. (1989) Primary structure of cathepsin D inhibitor from potatoes and its structure relationship to soybean trypsin inhibitor family. *FEBS Lett.* **251**, 94–98.

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

Mikola, J. (1987) Proteinases and peptidases in germinating cereal grains. In: *Fourth International Symposium on Pre-harvest Sprouting in Cereals* (Mares, D.J., ed.). Boulder, CO: Westview Press, pp. 463–473.

Morris, P.C., Miller, R.C. & Bowles, D.J. (1985) Endopeptidase activity in dry harvest-ripe wheat and barley grains. *Plant Sci.* **39**, 121–124.

Murakami, S., Kondo, Y., Nakano, T. & Sato, F. (2000) Protease activity of CND41, a chloroplast nucleoid DNA-binding protein, isolated from cultured tobacco cells. *FEBS Lett.* **468**, 15–18.

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

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

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.

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.

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.

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

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

Piatigorsky, J., Norman, B., Dishaw, L.J., Kos, L., Horwitz, J., Steinbach, P.J. & Kozmik, Z. (2001) J3-crystallin of the jellyfish lens: similarity to saposins. *Proc. Natl. Acad. Sci. USA* **98**, 12362–12367.

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

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

Qi, X. & Grabowski, G.A. (2001) Differential membrane interactions of saposins A and C: implications for the functional specificity. *J. Biol. Chem.* **276**, 27010–27017.

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

Ramalho-Santos, M., Veríssimo, P., Faro, C. & Pires, E. (1996) Action on bovine alpha s1-casein of cardosins A and B, aspartic

proteinases from the flowers of the cardoon *Cynara cardunculus* L. *Biochim. Biophys. Acta* **1297**, 83–89.

Ramalho-Santos, M., Pissarra, J., Veríssimo, 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.

Ramalho-Santos, M., Veríssimo, 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.

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.

Rodrigo, I., Vera, P., Van Loon, L.C. & Conejero, V. (1991) Degradation of tobacco pathogenesis-related proteins. *Plant Physiol.* **95**, 616–622.

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.

Runeberg-Roos, P., Törmäkangas, K. & Östman, A. (1991) Primary structure of a barley-grain aspartic proteinase. A plant aspartic proteinase resembling mammalian cathepsin D. *Eur. J. Biochem.* **202**, 1021–1027.

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.

St. Angelo, A.J. & Ory, R.L. (1970) Properties of a purified proteinase from hempseed. *Phytochemistry* **9**, 1933–1938.

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

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.

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.

Schuette, C.G., Pierstorff, B., Huettler, S. & Sandhoff, K. (2001) Sphingolipid activator proteins: proteins with complex functions in lipid degradation and skin biogenesis. *Glycobiology* **11**, 81R–90R.

Shinano, S. & Fukushima, K. (1971) Studies on lotus protease. III. Some physicochemical and enzymic properties. *Agric. Biol. Chem.* **35**, 1488–1494.

Törmäkangas, K., Kervinen, J., Östman, A. & Teeri, T.H. (1994) Tissue-specific localization of aspartic proteinase in developing and germinating barley grains *Planta* **195**, 116–125.

Törmäkangas, 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.

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

Veríssimo, 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.

Vicient, C.M., Suoniemi, A., Anamthawat-Jonsson, K., Tanskanen, J., Beharav, A., Nevo, E. & Schulman, A.H. (1999) Retrotransposon BARE-1 and its role in genome evolution in the genus *Hordeum*. *Plant Cell* **11**, 1769–1784.

Vieira, M., Pissarra, J., Veríssimo, 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.

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* L. seeds. *J. Plant Physiol.* **145**, 299–307.

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.

Wrobel, R. & Jones, B.L. (1992) Appearance of endoproteolytic enzymes during the germination of barley. *Plant Physiol.* **100**, 1508–1516.

Zhang, N. & Jones, B.L. (1995) Development of proteolytic activities during barley malting and their localization in the green malt kernel. *J. Cereal Sci.* **22**, 147–155.

Jukka Kervinen
3-Dimensional Pharmaceuticals, Inc., Johnson & Johnson
665 Stockton Drive,
Exton, PA 19341, USA
Email: jukka.kervinen@3dp.com

Alexander Wlodawer
Macromolecular Crystallography Laboratory,
NCI-Frederick, National Institutes of Health,
Frederick, MD 21702, USA
Email: wlodawer@ncifcrf.gov

Alexander Zdanov
Macromolecular Crystallography Laboratory,
NCI-Frederick, National Institutes of Health,
Frederick, MD 21702, USA
Email: zdanov@ncifcrf.gov

Handbook of Proteolytic Enzymes 2nd Edn
ISBN 0-12-079610-4