PURIFICATION OF A UNIQUE MATURATION ENZYME INVOLVED IN THE PROCESSING OF PROALEURAIN

By

CORALIE EMILIE HALLS

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy in Molecular Plant Sciences

WASHINGTON STATE UNIVERSITY

Program in Molecular plant Sciences

May 2004

To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation/thesis of CORALIE EMILIE HALLS find it satisfactory and recommend that it be accepted.

_____

_____

_____

_____

PURIFICATION OF A MATURATION ENZYME INVOLVED IN THE

PROCESSING OF PROALEURAIN

Abstract

by Coralie Emile Halls, Ph.D.

Washington State University

May 2004

Chair: John C. Rogers

Proaleurain is a cysteine aminopeptidase proenzyme expressed in most plant tissues. It trafficks through the secretory pathway from the Golgi apparatus to a prevacuolar compartment (PVC) for the lytic vacuole. It is processed to the active form by another protease, a unique maturation mechanism for cysteine proteases. Our data indicate that this "proaleurain maturation enzyme" is specifically localized to the PVC, and its identification could provide a biochemical marker for that compartment. Therefore we have substantially purified the maturation enzyme from cauliflower florets. The central problem has been progressive loss of activity with increasing purity through ammonium sulfate precipitation, Mono S cation exchange and bacitracin affinity chromatography. Surprisingly, enzyme activity after the bacitracin step could be increased 50-100 fold by treatment with 2% SDS followed by 10x dilution into 0.2% NP-40, at pH 4.5. Renaturation of the maturation enzyme from SDS-PAGE gel slices demonstrated a size of 25-32 kDa. Proteins comprising the five bands in this size range have been identified by Q-TOF mass spectrometry. Two proteases are present, a serine carboxypeptidase and a cysteine

protease containing a granulin domain. Also present is a previously uncharacterized ~30 kDa Künitz-type protease inhibitor. We postulate that SDS-activation could result either from dissociation of the Künitz inhibitor from the maturation protease, or from unfolding of an inhibitory domain, either the N-terminal propeptide or the granulin domain, away from the maturation protease active site. Either mechanism could limit maturation protease activity to the acidic environment of the PVC. We demonstrated that a recombinant cysteine protease containing the granulin domain processed proaleurain to a size indistinguishable on SDS-PAGE from mature barley aleurain, and that the Künitz inhibitor inhibits the cauliflower maturation enzyme activity as well as the activity of the recombinant granulin domain cysteine protease. Further studies are underway to define the amino acid sequence in proaleurain that is cleaved by the cysteine protease with the granulin domain, and to determine the localization of that enzyme and of the Künitz inhibitor in the vacuole and/or prevacuolar compartment.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

1.1 Vacuolar proteins and the secretory pathway

Proteins reach the plant vacuole (equivalent to the mammalian lysosome or the yeast vacuole) by vesicle trafficking through the secretory pathway. The secretory pathway is an endomembrane system comprised of separate organelles, each of which has unique functions that depend on the environment created by the membrane boundary of the organelle. This endomembrane system is composed of the endoplasmic reticulum, the Golgi apparatus, prevacuolar compartments and vacuoles. Proteins move between the separate organelles in transport vesicles. As proteins move through the organelles they are modified in ways that are specific for each organelle. In this introduction we will concentrate on soluble protein targeting from the Golgi apparatus to the plant vacuole, but first must consider the role of the endoplasmic reticulum in preparing proteins for their trip through the secretory pathway.

1.1.1 Endoplasmic reticulum

ER is present in all eukaryotic cells. Its membrane forms a long tubular network in the cell with a common internal space, called the ER lumen. The ER plays a central role in lipid synthesis and protein biosynthesis. Almost all the proteins destined for secretion to the cell exterior or destined for the lumen of the ER itself, Golgi, or vacuole are initially synthesized on the ER membrane.

- **The signal peptide; a pass for entry to the ER lumen**

A newly synthesized protein enters the secretory pathway through the membrane of the endoplasmic reticulum as an unfolded molecule. Once inside the ER lumen, a soluble protein

will not have to cross any other membranes; movement from one organelle to another occurs when the protein is packaged into transport vesicles.

Proteins that enter the secretory pathway are synthesized on ribosomes associated with the ER membranes and are targeted to enter the ER by their cleavable amino-terminal signal peptide. A functional signal sequence has three conserved domains: a positively charged amino terminal, a central hydrophobic domain and a carboxy-terminal hydrophilic domain. These signal peptides have no sequence homology; their important criterion is the hydrophobic character (Cioffi et al., 1989). Kaiser et al. demonstrated that many artificial peptides can work as a ER targeting signal in yeast as long as the sequences feature hydrophobicity above a certain threshold (Kaiser et al., 1987).

The signal peptides are recognized by a signal recognition particle that is in most cases required for targeting the nascent protein to the ER membrane. After co-translational translocation of the soluble protein into the ER lumen through a translocation channel, the signal peptide is cleaved by a signal peptidase.

- **Reticuloplasmins help in proper protein folding**

Proteins enter the ER in an unfolded stage. The ER lumen is an oxidizing environment that promotes disulfide bond formation (Hwang et al., 1992), and ATP is required for the process by which the protein folds into a proper three dimensional structure (Braakman et al., 1992). Enzymes and molecular chaperones residing in the ER lumen have been termed reticuloplasmins.

It is believed that molecular chaperones increase the yield but not the rate of properly folded proteins (Hartl, 1996). They transiently bind to hydrophobic domains of the protein to be folded until these domains are buried in the folded structure thus decreasing the probability of

non-wanted reaction that might give rise to non-functional proteins. BiP is the best characterized chaperone (Vitale and Denecke, 1999) and is commonly used as an ER marker.

Folding enzymes such as protein disulfide isomerase help protein folding by catalyzing the formation of disulfide bonds, while the calnexin/calreticulum/glycosyl transferase system interacts specifically with nascent glycans. Glycans are added to the protein by glycosyl transferases as the nascent polypeptide is translocated into the ER lumen. The process is known as glycosylation and for most proteins the folding process is facilitated via the interaction of calnexins and calreticulums with these glycans (Vitale and Denecke, 1999). Glycosylation is important for the proper folding and stability of a protein (Lerouge et al., 1998; Lowe and Marth, 2003).

The reticuloplasmins seem to be important factors in the proper folding of proteins. Presumably they remain attached to the protein until the proper structure has been reached. They have also been shown to stay attached to misfolded proteins that remain in the ER before being degraded. The chaperones retain immature and misfolded proteins in the ER (Hammond and Helenius, 1995; Vitale and Denecke, 1999), preventing them from reaching their destination where they could negatively interfere with cell metabolism. This concept is known as quality control (Hammond and Helenius, 1995). The quality control of misfolded proteins and their degradation enables the cell to recycle the amino acids. The degradation site is not clear (Vitale and Denecke, 1999). The possible degradation sites are the cytosol, the ER itself or the vacuole (Vitale and Denecke, 1999).

- **Retention of reticuloplasmins in the ER**

Reticuloplasmins are retained in the ER by a carboxy-terminal sorting signal (i.e. HDEL, KDEL : using the single letter code for amino acids). If they should escape the ER, KDEL

receptor proteins on the *cis* Golgi will transport the soluble ER-resident proteins back to the ER lumen (Vitale and Denecke, 1999). The KDEL receptor is a membrane protein with a lumenal part that recognizes the sorting signal and a cytosolic tail that interacts with a specific set of proteins involved in the budding of the ER membrane into a transport vesicle (Vitale and Denecke, 1999). These proteins will form a "coat" around the transport vesicle, known as the COPI vesicle (Robinson et al., 1998).

- **Export from the ER**

One commonly accepted process by which proteins leave the ER to go to the Golgi apparatus is explained by the bulk flow model. In a bulk flow process, proteins leaving the ER have no particular signal or particular membrane receptor, to concentrate these proteins on the export site. As described earlier, the properly folded proteins are free of reticuloplasmins that are thought to prevent the escape of non-folded proteins (Vitale and Denecke, 1999). The process by which proteins destined to be exported to the Golgi accumulate on an exit site where they are packed into transport vesicles is not yet understood. More information is known about formation of the transport vesicles. On the cytosolic side of the ER membrane, proteins will associate and form a coat that will enable the membrane to bud out and form the transport vesicles (Robinson et al., 1998). That process requires the hydrolysis of GTP (Robinson et al., 1998). The vesicles leaving the ER are called COPII vesicles and the vesicles involved in the recycling from the Golgi of escaped proteins from the ER are called COPI vesicles (Robinson et al., 1998). The coats, as indicated by the different names of these two types of vesicles are composed of different coat proteins (Robinson et al., 1998).

1.1.2 Golgi apparatus

The plant Golgi apparatus is involved in glycosylation of proteins on serine and threonine residues, modification of asparagine-linked glycans to so called complex "forms", and synthesis of cell wall polysaccharides (hemicellulose and pectins, but not cellulose). The Golgi apparatus is also the place where vacuolar proteins are sorted from proteins that will be secreted from the cell. It is composed of 5 to 7 flatterned sacs, depending on the cell type and cell development stage (Andreeva et al., 1998). These Golgi apparatus sacs are different in morphology and enzymatic content (Andreeva et al., 1998). The secretory proteins enter the *cis*-face of the Golgi apparatus and leave at the *trans*-face (Neumann et al., 2003).

- **Golgi residents and their retention in the Golgi**

Enzymes remaining in the Golgi sacs are termed Golgi residents. The Golgi residents are involved in the modification of asparagine-linked (Asn-linked) glycans of glycoproteins synthesized in the ER (Neumann et al., 2003). There are no known Golgi retention sequences for Golgi resident proteins. Most of them are membrane anchored proteins, and there is speculation that this property has something to do with their distribution among the different sacs and their retention in the Golgi (Neumann et al., 2003).

- **Vesicular shuttle and progression/maturation models**

The mechanism for transport of proteins between the sacs of the Golgi apparatus is not clear. Two models have been proposed: the vesicular shuttle model and the progression/maturation model (Nebenfuhr and Staehelin, 2001).

In the vesicular shuttle model, the sacs of the golgi apparatus are considered as stable units containing sac-specific Golgi residents. A secretory protein is shuttled from one sac to the other via transport vesicles.

In the progression/maturation model, the sacs move through the Golgi apparatus. The *cis*-sac is formed by the fusion of the transport vesicles from the ER to the Golgi. That sac will mature as a new *cis*-sac is formed. The enzymes specific to each sac are transported back from the old *cis*-sac to the new *cis*-sac by transport vesicles (and so on for the remaining sacs). The *trans*-sac, the oldest sac, is fragmented into post-Golgi transport vesicles.

- **Sorting proteins from the Golgi to the vacuole in plants**

The mammalian mannose 6-phosphate receptor recognizes mannose-6-phosphate groups on protein glycans, and delivers proteins to the lysosome or prelysosomal compartment via clathrin coated transport vesicles (Kornfeld and Mellman, 1989). The role of glycans in plant protein targeting to the vacuolar compartment was analysed in transgenic tobacco, where the targeting of mutated vacuolar proteins like patatin (Sonnewald et al., 1990) and barley lectin (Wilkins et al., 1990) that could no longer be glycosylated was studied. These unglycosylated mutants were still targeted to the vacuole, demonstrating that in plant cells, Asn-linked glycans are not involved in vacuolar sorting (Sonnewald et al., 1990; Wilkins et al., 1990). The idea that plant targeting information is contained in the protein sequence and not in a glycan arose from the study of the pathogenesis related proteins, β-1,3-glucanase and chitinase (Hadlington and Denecke, 2000; Neuhaus et al., 1991; Shinshi et al., 1988). These proteins are found in tobacco vacuoles and are expressed as preproproteins, with an amino-terminal signal peptide and a carboxy-terminal propeptide (Melchers et al., 1993). In both cases the carboxy-terminal peptide is removed post-Golgi and is no longer present in the vacuolar located protein (Melchers et al., 1993). In tobacco, expression of chimeric genes of β-1,3-glucanase or chitinase with or without the carboxy-terminal propeptide revealed a different location of these two forms in the cell (Melchers et al., 1993). The full length protein was targeted to the vacuolar compartment,

whereas the carboxy-peptide deleted form was secreted, indicating the importance of β-1,3-glucanase propeptide in the proper cellular targeting (Melchers et al., 1993). Confirmation of the concept of a vacuolar sorting signal in plants being dependent on the protein sequence arose from the discovery of additional sorting sequences: barley lectin contains a vacuolar carboxy-terminal sorting sequence (Bednarek et al., 1990) wherease sweet potato sporamin (Matsuoka and Nakamura, 1991) and barley aleurain (Holwerda et al., 1992) contain an amino-terminal sorting sequence. These plant sorting determinants encoded in the protein sequences, are known as vacuolar sorting determinants (VSDs).

Three different groups of VSDs have been described in plants (Neuhaus and Rogers, 1998): a sequence-dependent VSD, a C-terminal VSD and a protein structure-dependent VSD. Only the sequence-dependent VSD has a conserved sequence. That sequence can be located anywhere in the protein sequence. Such a VSD can be found at the amino-terminal sequence of certain cysteine proteases (for exemple: barley aleurain (Holwerda et al., 1992; Holwerda and Rogers, 1993)), sweet potato sporamin, a protease inhibitor (Nakamura et al., 1993) and Künitz type protease inhibitors (Matsumoto and Neuhaus, 1999).

1.1.3 Sorting of proaleurain to the vacuole

My dissertation is focused on the transport of the barley cysteine protease, aleurain, to the vacuole. The proaleurain VSD was identified by incorporating different sequences from proaleurain into a secreted thiol protease, proendoproteinase B (ProEP-B) (Holwerda et al., 1992). The chimeric proteases were then transiently expressed in electroporated tobacco protoplasts and analyzed for their location in or out of the cell (Holwerda et al., 1992). The targeting determinant SSSSFADSNPIR of the aleurain propeptide, substituted into the propeptide of EP-B caused vacuolar targeting of the resulting chimeric protein (Holwerda et al.,

1992). These results emphasized that plant proteins contain targeting information within their polypeptide sequences (Hille-Rehfeld, 1995). Proteins lacking a sorting determinant were secreted because of the absence of targeting information. Thus a receptor that recognizes VSD sequences and pulls vacuolar proteins away from the bulk flow of secreted proteins must exist (Hille-Rehfeld, 1995).

The discovery of a pea vacuolar receptor, BP-80 (Humair et al., 2001; Kirsch et al., 1994), confirmed the parallel vacuolar targeting mechanisms between the plant and mammalian systems. Proteins extracted from the membranes of clathrin coated vesicles purified from developing pea cotyledons were run on an affinity column with aleurain targeting sequence peptide as bait (Kirsch et al., 1994). BP-80 bound the column at neutral pH and was eluted at pH 4 (Kirsch et al., 1994). BP-80 is a type I membrane protein with a lumenal amino-terminal and a cytosolic carboxy-terminal domain. The luminal domain recognizes specific sequences that contain the floor amino acid motif, NPIR. BP-80 binds to the motif at neutral pH (Kirsch et al., 1994). Results from *in vitro* experiments (Cao et al., 2000) and *in vivo* experiments in yeast (Humair et al., 2001) indicate a specific recognition between the receptor BP-80 and the NPIR containing VSD of barley proaleurain. The receptor's carboxy-terminal cytoplasmic tail is sufficient for protein targeting to a prevacuolar compartment (Jiang and Rogers, 1998b). Because the receptor was not detected on the vacuolar membrane it is believed that it recycles back from the pre-vacuolar compartment to the *trans*-Golgi (Jiang and Rogers, 1998b), in a manner similar to the recycling of mammalian mannose 6-phosphate receptor (Hille-Rehfeld, 1995).

In mammalian cells clathrin coated vesicles contain the mannose 6-phosphate receptor and hydrolytic enzymes destined for the acidic lysosome (Hille-Rehfeld, 1995). Plant cells also

have clathrin coated vesicles (CCVs) (Hohl et al., 1996). CCVs purified from developing pea cotyledons through successive sucrose gradients contained the receptor BP-80 but didn't contain the storage proteins, vicilin and legumin (Hohl et al., 1996). These storage proteins were detected in a different type of vesicle, designated as dense vesicles (Hoh et al., 1995; Hohl et al., 1996). These results established the concept that storage protein and lytic proteins exit the Golgi in different types of vesicles. Because each vesicle type is specific for an individual pathway to an organelle, two different vesicles indicated two pathways from Golgi to vacuoles and raised the question of whether there were two separate types of vacuoles, lytic and storage.

1.1.4 The concept of two vacuoles in plant cells

A vacuole is a compartment that can occupy up to 90% of the volume of a mature plant cell. The vacuolar lumen is surrounded by a single membrane known as the tonoplast. The vacuolar lumen contains water and many different types of other molecules, for example ions, hydrolytic enzymes, storage proteins, pigments and defense molecules, harmful chemicals. The types of components in the vacuole depends greatly on the plant type, cell tissue type and cell developmental stage.

The concept of separate compartments for proteases and storage proteins was developed when the cysteine protease, aleurain, was localized to small vacuoles that were morphologically distinct from vacuoles containing storage proteins in barley aleurone layers. Localization was determined from immunogold EM studies using an affinity-purified polyclonal antibody raised to recombinant aleurain (Holwerda et al., 1990; Holwerda and Rogers, 1992).

Subsequently experiments where sucrose gradient separation of developing pea cotyledon cell organelles was employed revealed the presence of two distinct vacuolar compartments (Hoh et al., 1995). These vacuoles contained distinctive tonoplast intrinsic proteins (TIPs) (Hoh et al.,

1995): α-TIP (Johnson et al., 1990) and TIP-Ma27 (Marty-Mazars et al., 1995). These observations were confirmed when confocal immunofluorescence experiments in pea and barley root tip cells showed α-TIP colocalising with the storage protein barley lectin and TIP-Ma27 colocalising with the cysteine protease aleurain (Paris et al., 1996). Barley lectin and aleurain were in separate vacuoles. This established the concept that plant cells have separate lytic and storage vacuoles.

There has been much interest in understanding how plant cells generate two separate types of tonoplast and sort proteins specifically to one or the other vacuolar compartment. One approach is to identify proteins that are specific for one type of vacuole and use them to expand our knowledge of targeting mechanisms. Aleurain is an exemple of a key protein that has been used as a tool to characterize protein sorting mechanisms.

1.2 Cysteine proteases

Cysteine proteases belonging to the papain family (C1 family) are found in prokaryotes as well as eukaryotes and by definition share significant similarity in amino acid sequence either to the type example of the family, papain, a cysteine protease from *Carica papaya* latex. The activity of these enzymes is involved in the hydrolysis of a peptide amide bond in the protein substrate. After the enzyme has bound the substrate the basic mechanism of action involves attack of the nucleophilic thiol of the enzyme's cysteine residue on the carbonyl of the scissible amide bond of the protein substrate to form a transition state complex known as tetrahedral intermediate. The imidazole ring of the enzyme histidine residue takes up the liberated proton, forming an imidazolium ion. The tetrahedral intermediate will form an acyl-enzyme intermediate under the driving force of the proton donation from the histidine residue. In other terms you have the amide bond that has been cut leaving the substrate in two parts, the amino-

portion and the carboxy-portion. The carboxy-part is free to leave the active site and the amino-part is covalently linked to the enzyme cysteine residue forming the acyl-enzyme intermediate. This intermediate is very unstable. The amino-portion of the substrate is released from the enzyme by hydrolysis of a water molecule thereby regenerating the active enzyme and forming the new carboxy-terminal of the cleaved substrate amino-portion. The enzyme cysteine and histidine residues involved in the hydrolysis of the substrate peptide amide bond are referred as the catalytic residues (Lecaille et al., 2002).

The members of the papain family also share a similar tertiary structure (Turk et al., 2002). The protease is folded into two domains (L and R) of comparable size with the active site located in a groove formed between the two domains (Turk et al., 2002).

The primary translation product of a eukaryotic cysteine protease is called a preproenzyme comprised of an N-terminal signal peptide (between 10-20 amino-acids), followed by a propeptide (between 38 and 250 amino-acids) and finally the sequence for the mature enzyme (between 220-260 amino-acids) (Wiederanders, 2003). After removal of the signal peptide in the ER, the protein is called a proenzyme or zymogen

- **Activation of the zymogen**

In the zymogen tertiary structure the propeptide sits in the active groove in the opposite direction of the natural substrate (McGrath, 1999). The removal of the propeptide is necessary for activation of the enzyme. This process is called maturation or processing.

The maturation of zymogens into active enzymes includes one or more limited proteolytic cleavages. Two mechanisms are known: intramolecular, if catalyzed by the same enzyme, or intermolecular, if catalyzed by a different enzyme(s). Most of the cysteine proteases studied showed the ability *in vitro* of intramolecular processing. For example, human

procathepsin K expressed in baculovirus-infected cells, shows spontaneous *in vitro* activation at pH 4 (Bossard et al., 1996; McQueney et al., 1997). However, it is believed that intermolecular processing greatly facilities proenzyme activation (Wiederanders, 2003).

The propeptide, besides being an inhibitor (Wiederanders, 2003), assists proper folding of the protease (Wiederanders, 2003) and participates, as described earlier, in the correct intracellular targeting in plant cells (Holwerda et al., 1992).

- **Carboxy-terminal extension**

Some papain cysteine proteases of plants, for exemple RD21 (Yamada et al., 2001a), and parasite cysteine proteases of intracellular parasites like *Leishmania* CPB2.8, have a C-terminal extension (CTE) (Brooks et al., 2000). This extension is not found in mammalian cysteine proteases. In plants the extension peptide sequence is related to a mammalian protein, granulin (http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00396) wherease in parasite cysteine proteases the extension sequence is not related to granulin. In all cases the CTE sequence is composed of a proline rich repeat followed by a cysteine rich domain with a conserved motif :

- $PPX_2CX_5CX_5CCX_8CCX_6CCX_5CCX_5CX_5C$ **Mammalian Granulin**

- $PPX_2CX_5CX_5CCCX_7CCX_4CCX_6CCX_5CCX_6CX_6C$ **RD21 *Arabidopsis thaliana***

- $PX_7CX_4CX_3CX_9CX_{12}CX_7CX_6CX_5CX_7C$ **CPB2.8 *Leishmania mexicana***

Mammalian granulins are approximately 6 kD proteins that stimulate or inhibit the growth of animal cells (Bateman and Bennett, 1998). The CTE domain of parasite or plant papain-like cysteine proteases has an unknown function and gets cut from the protease after reaching the parasite lysosome (Brooks et al., 2000) or plant vacuole (Yamada et al., 2001a). In *Leishmania mexicana* the expression of CPB 2.8 without its CTE did not affect the targeting of

the mutated protein to the lysosome (Brooks et al., 2000). The functions of cysteine protease CTEs are unknown.

- **Relation between activity and structure**

The majority of the papain like proteases have endopeptidase activity, cutting inside instead of at the ends of the substrate proteins . Observed activities of the remaining proteases are aminopeptidase activities, cutting the amino-end of the substrate and carboxypeptidase activity, cutting the carboxy-end of the substrate. These activities are explained by structural differences in the active site groove. For example mammalian cathepsin B carboxypeptidase has an occluding loop which favors its peptidyl activity, cutting the C-terminal dipeptide of the substrate (Podobnik et al., 1997). Mammalian aminopeptidase cathepsin H has a mini-chain in the active groove blocking partial access to the substrate (Figure 1.1) (Guncar et al., 1998). Plant aleurain is also an aminopeptidase (Holwerda and Rogers, 1992) and the presence of a mini-chain in aleurain has been determined (Rothe et al., 1994).

1.2.1 Cathepsin H and aleurain

Plant aleurain and mammalian cathepsin H share more than 60% identity on the amino acid sequence level. Both enzymes are members of the papain family or C1 family. In contrast to most of the cysteine proteases from this family which are endoproteases, aleurain and cathepsin H show aminopeptidase activity. The presence of a mini-chain explains that difference in activity (Figure 1.1).

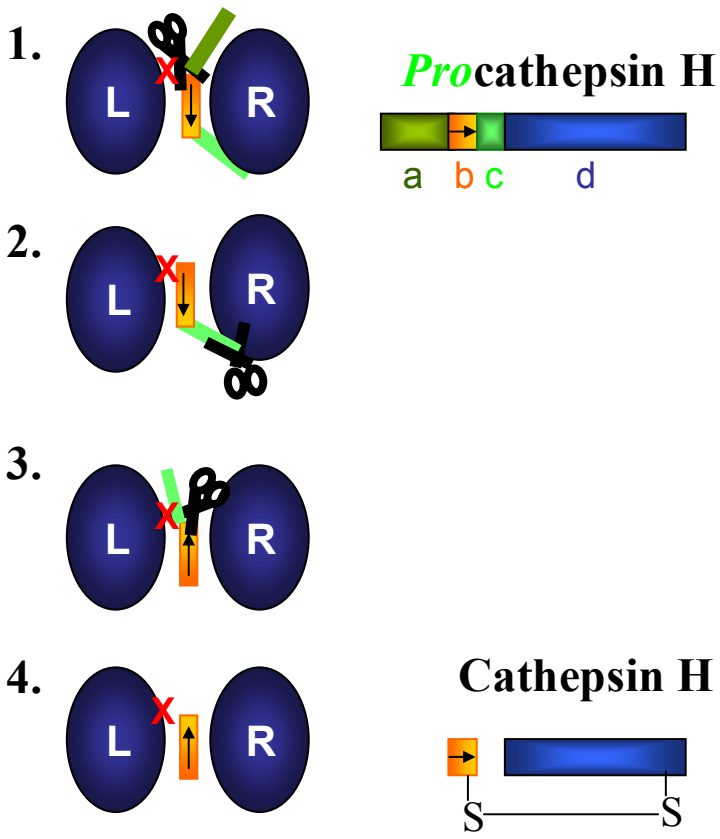


Figure 1.1. Hypothetical maturation mechanism of cathepsin H

Cathepsin H sequence has been devided in 4 domains: the propeptide #1 (**a**), the mini-chain (**b**), the propeptide #2 (**c**) and the cysteine protease domain (**d**). The propeptide #1 gets cut away (**1**), then the propeptide #2 and the mini-chain are released from the cysteine protease domain (**2**), then somehow that sequence gets flipped around in the active site groove (**3**)**,** now facing the same direction as the protein substrate so that the processing of the propeptide #2 from the mini-chain could be an autocatalytic mechanism (**3**). The mini-chain is sitting in the active site groove, blocking some of the space to the substrate (**4**). Position of the active cysteine residue **(X).**

It had been speculated that the mini-chain sits in the active site groove, blocking access to the protein substrate (Baudys et al., 1991). The determination of cathepsin H crystal structure confirmed the hypothesis of the link between the presence of a mini-chain in the active groove and the aminopeptidase activity (Figure 1.1). Cathepsin H structure is composed of two globular domains: the right domain (R) and the left domain (L) (Figure 1.1)(Guncar et al., 1998). The space between the domains forms a groove that accepts the substrate. The top of the groove contains the active site cysteine and histidine residues. In the groove a small peptide of 8 amino acids sits on the bottom blocking the substrate from further entering the site (Guncar et al., 1998). The mini-chain is a small fragment from the pro-peptide. The mini-chain is linked to the mature cysteine protease, or main body sequence by a disulfide bond (Baudys et al., 1991; Guncar et al., 1998)

Crystal structures of cysteine pro-proteases showed the presence of the propeptide in the opposite direction of the peptide substrate in the active site groove (McGrath, 1999). In contrast, the cathepsin H crystal structure showed the mini-chain in the active site groove in the same direction as the substrate. This led the authors to form the following hypothesis to explain the cathepsin H maturation mechanism (Figure 1.1). The maturation process has four steps. In the first step the propeptide attached to the N-terminal end of the mini-chain is cut away, then the remaining propeptide including the mini-chain is cut away from the main body. In the third step, the propeptide flips from the opposite direction to the same direction as the substrate in the active groove. In the last step the small remaining propeptide is cut away from the mini-chain (see picture). Because at that step the propeptide would be in the same orientation as the substrate, removal of the remaining amino acids from the mini-chain could involve an autocatalytic or intramolecular process. Results from use of group specific protease inhibitors indicate that the

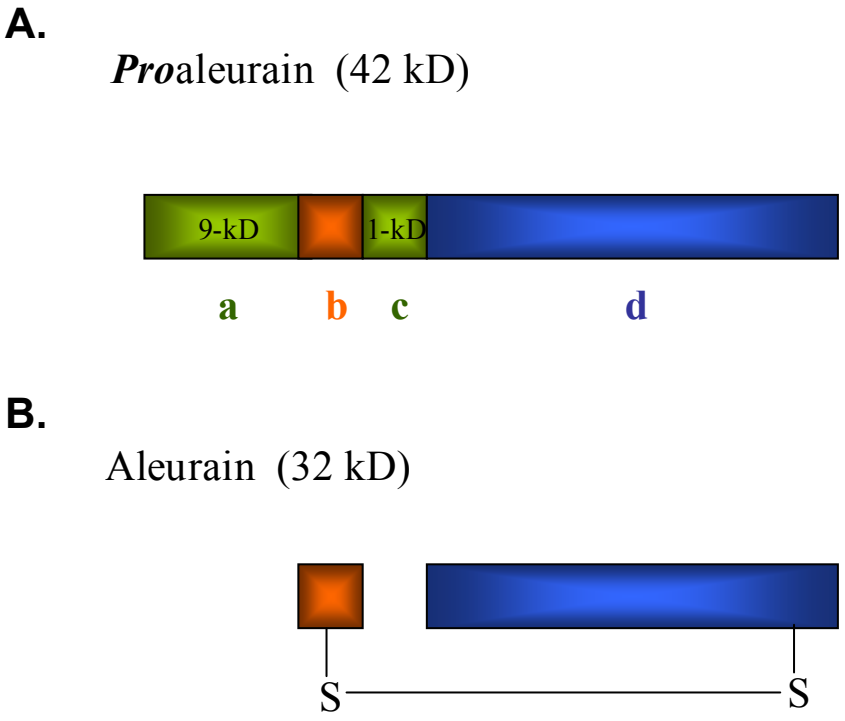


Figure 1.2. Aleurain proform and mature form

A. Representation of 42 kDa proaleurain sequence that has been divided in 4 domains: the 9 kDa propeptide (**a**), the mini-chain hexapeptide (**b**), the 1 kDa propeptide (**c**) and the cysteine protease domain (**d**). **B.** Representation of 32 kDa mature aleurain with its mini-chain and cysteine protease domain.

mechanism for processing procathepsin H to mature cathepsin H involves a pepstatin-sensitive aspartic protease(s) (Nishimura and Kato, 1988). Pulse-chase experiments indicate that the processing of cathepsin H takes place in the lysosome, equivalent of the plant lytic vacuole (Nishimura and Kato, 1987).

1.2.2 Aleurain maturation

Aleurain is also expressed as a proenzyme, proaleurain (Figure 1.2A). The limited experimental data available indicate that proaleurain is one of the few cysteine protease proenzymes that are not autocatalyticaly activated. Results from *in vitro* and *in vivo* studies indicate that proaleurain maturation involves two major steps (Holwerda et al., 1990): a clipping step in which the protein size diminishes from 42 kDa to 33 kDa, and a trimming step in which it diminishes from 33 kDa to 32 kDa (Figure 1.3). The 42 kDa proform is first clipped to 33 kDa as demonstrated by the following in vivo maturation experiment. Gibberellic acid (GA)-treated barley aleurone cells were pulse labelled with [^{35}S] labeled cysteine and methionine, then chased for various time points; aleurain was immunoprecipitated and fractionated by SDS-PAGE. The corresponding fluorograph showed that the proaleurain precursor size was 42 kDa, which was converted after 1 hour of chase to a 33 kDa band. After more then 2 hours of chase a stable 32 kDa form appeared; the latter change corresponded to the trimming step. Sequence comparison of the predicted aleurain with the mature sequences of cathepsin H and papain (Rogers et al., 1985) indicated that maturation of aleurain results from the removal of an amino-terminal propeptide. Amino-terminal sequencing of the mature aleurain confirmed the maturation hypothesis (Rothe et al., 1994).

The clipping and trimming steps require acidic conditions, as demonstrated by the maturation of correctly folded, radiolabeled *Xenopus* oocyte-synthesized proaleurain by enzymes

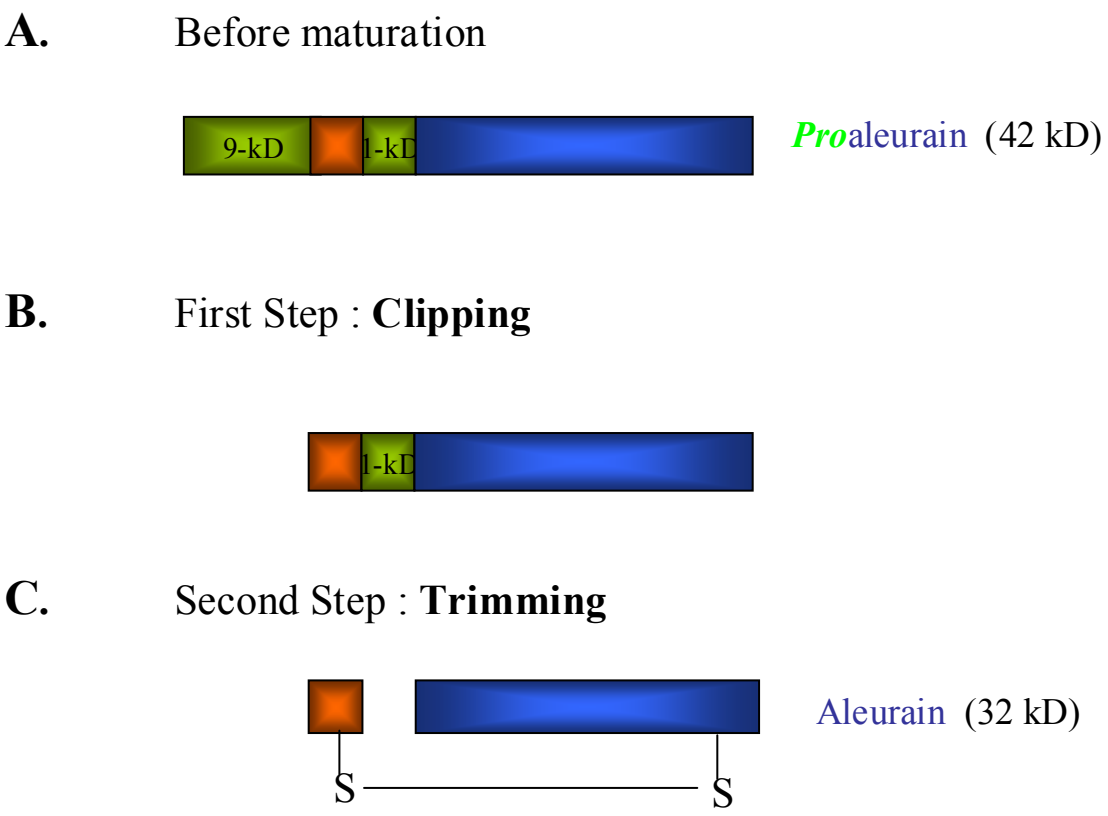


Figure 1.3. Proposed mechanism of proaleurain maturation

In the following order the 42-kD proaleurain (A) gets its 9-kD propeptide clipped (B) and then the 1-kD propeptide gets trimmed (C) leaving the mini-chain linked by a disulfide band to the cysteine protease domain.

of barley aleurone extracts *in vitro* (Holwerda et al., 1990). Because, in plant cells, aleurain travels through the secretory pathway, it was thought to be important to choose an expression system that would be able to modify the protein in the same manner as the plant cell. The secretory pathway in *Xenopus* oocyte cells would provide the secretory pathway modifications described earlier, leading to a correctly folded substrate. The importance of the proper folding of the substrate is significant in the comparison of *in vivo* processing versus *in vitro* processing mechanisms (where you can control the parameters of the maturation environment).Only aleurone extracts prepared in buffers of pH values of 5 or 6 were able to proteolytically process the zymogen, and the processed forms were indistinguishable in size from those produced *in vivo* (Holwerda et al., 1990).

The SDS-PAGE radiographs from *in vitro* assays incubated in the presence of various protease inhibitors showed that the clipping step was not affected by the cysteine protease inhibitors E-64 and leupeptin, a serine protease inhibitor phenylmethylsulfonyl fluoride, an aspartic protease inhibitor pepstatin, or metalloprotease inhibitors EDTA and 1,10-phenanthroline. In contrast the trimming step was inhibited by E-64 (Holwerda et al., 1990). Therefore the maturation of proaleurain involves two different protease activities and the inhibitor studies indicated that only the trimming step was mediated by a cysteine protease. The known orientation of the mini-chain in the same direction as the protein substrate in cathepsin H crystal structure (Guncar et al., 1998) and the demonstrated involvement of a cysteine protease in the trimming step led to the hypothesis that the trimming step is an autocatalytic step.

Immunoblot analysis of extracts from aleurone layers incubated with gibberellic acid showed the presence of endogenous mature aleurain wherease extracts from abscisic acid treated aleurone layers showed no endogenous mature aleurain (Holwerda et al., 1990). However, both

extracts had similar amounts of proaleurain clipping activities. These results seem to indicate that mature aleurain is not involved in proaleurain maturation (Holwerda et al., 1990). Thus, the authors proposed that the clipping step involves a maturation enzyme distinct from aleurain (Holwerda et al., 1990). The maturation of proaleurain occurs in a post-Golgi compartment (Figure 1.4) as shown in a pulse-chase experiment on GA treated aleurone layers in the presence or absence of the drug monensin which blocks the secretory pathway at the medial Golgi (Tartakoff, 1983). The processing of 42 kDa proaleurain to the 32 kDa form was blocked in the presence of the monensin, while in its absence the 32 kDa form was observed.

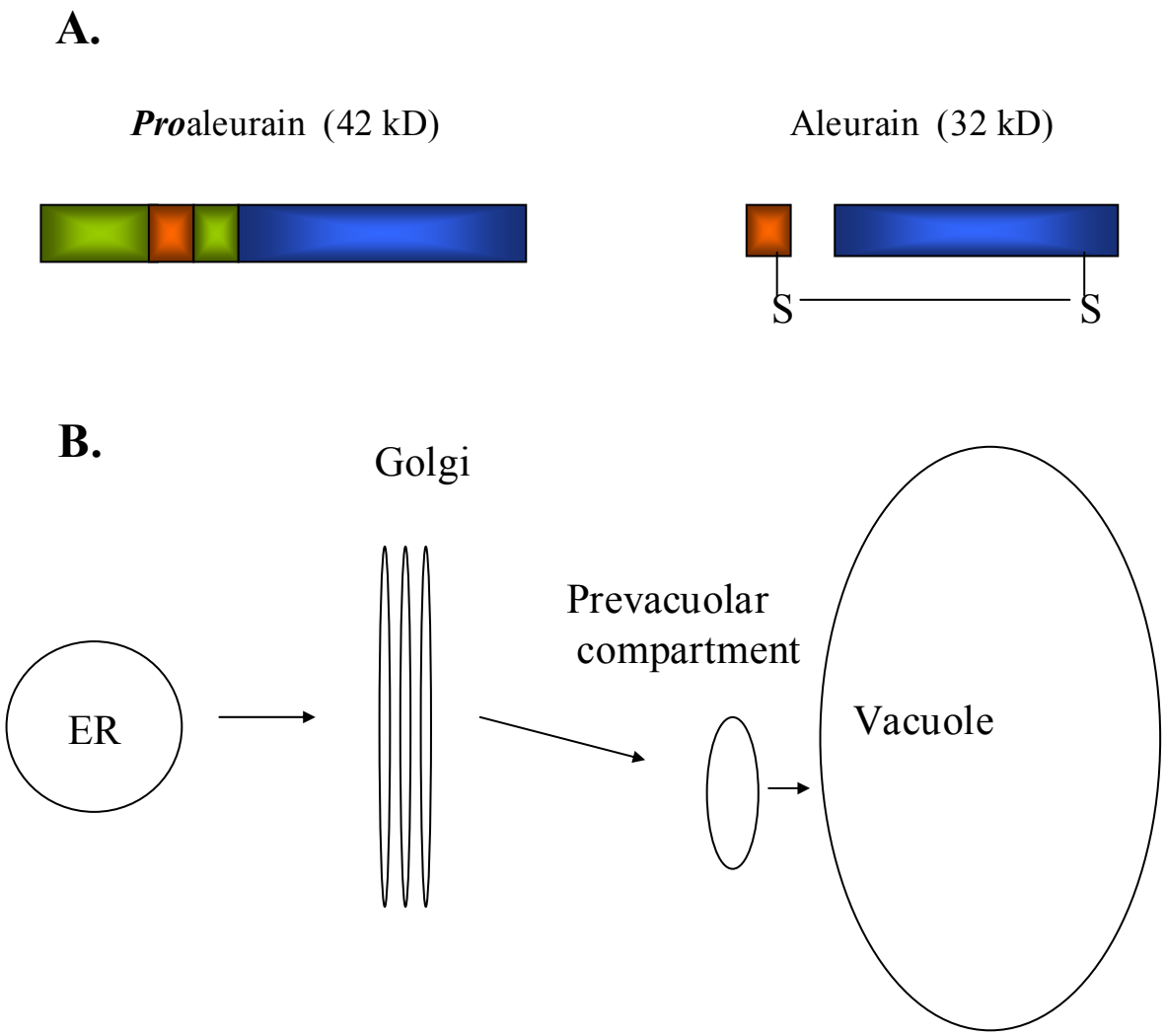


Figure 1.4. The secretory pathway and aleurain

A. Representation of proaleurain and aleurain sequences with their respective molecular weight as determined by SDS-PAGE. **B**. Proaleurain is cotranslationally inserted into the endoplasmic reticulum (ER), then is transported via transport vesicles to the Golgi apparatus, then to the prevacuolar compartment and finally to the lytic vacuole . The processing site of proaleurain to aleurain is thought to occur either in the prevacuolar compartment or the lytic vacuole.

CHAPTER 2 MATERIALS AND METHODS

2.1 Expression of Proaleurain in yeast cells

A construct with the full length proaleurain cDNA has been described previously (Rogers et al., 1985). Proaleurain cDNA was excised from the plasmid vector with *HindIII* and *EcoRI* and was ligated into the *EcoRI* and *HindIII* interval of plasmid Bluescript pKS(+). From this new construct the proaleurain cDNA was excised with *EcoRI* and *KpnI*, was ligated into *EcoRI* and *KpnI* interval of plasmid pPICZA (Invitrogen, Carlsbad, CA).

This plasmid was linearized with *BstX1* and then transfected into *Pichia* host strain KM71 (EasySelect™*Pichia* expression kit, Invitrogen) as described by the EasyComb™ transformation protocol supplied with the EasySelect™*Pichia* Expression kit (Invitrogen). Proaleurain expression was induced with methanol as described by the Expression protocol supplied with the EasySelect™*Pichia* Expression kit (Invitrogen). Expression analysis was done by SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) as described by protocol supplied with the EasySelect™*Pichia* Expression kit (Invitrogen).

2.2 Expression of Proaleurain in Drosophila cells

A construct with the full length proaleurain cDNA has been described previously (Rogers et al., 1985). Proaleurain cDNA excised from the plasmid vector with *EcoR1* and *HindIII* (blunted) was ligated into the *EcoR1* and *XhoI* (blunted) interval of plasmid pMT/VT-His A (Invitrogen, Carlsbad, CA). The proaleurain sequence included a stop codon, its coding sequence was not in frame with a C-terminal six His residues.

This plasmid was transfected into *Drosophila* S2 cells (*Drosophila* expression system, Invitrogen) using calcium phosphate transfection as described by protocols supplied with the

Drosophila Expression System kit (Invitrogen). First, transfected cells were grown in complete Schneider's *Drosophila* medium without hygromycin for two days at 24ºC. After that, the cells were transferred into medium containing hygromycin for selection for 2 – 4 weeks. Finally, the cells were maintained in long-term culture in the presence of hygromycin.

For production of the protein, the cell culture was centrifuged (1000g 3 min) and washed with serum-free and hygromycin-free medium, then resuspended in the same medium at a cell density of $4x10^6$ cells/ml. Synthesis of proaleurain protein was induced by addition of $CuSO_4$ to a final concentration of 500 μM and the medium was harvested two days later.

For concentration and change of buffer, the recombinant proaleurain protein was also dialyzed into 50 mM succinic acid pH 4.5, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 1 mM dithiothreitol (DTT) using a Micro-ProDicon® Dialysis/Concentration system (Spectrum, Houston, TX) with a Micro-ProDicon™ Membrane with a molecular weight cut off of 10,000 (Spectrum, Houston, TX).

2.3 Plant material

Cauliflower heads were bought at the local supermarket. The florettes were cut away, frozen with liquid nitrogen and stored at -80ºC until use.

2.4 Protein Gel Blots and Coomassie Staining

Methods for SDS-PAGE, Coomassie staining of gel, transfer of proteins to nitrocellulose membranes, incubation with primary antibodies, and detection with horseradish peroxidase–linked secondary antibodies have been described elsewhere (Cao et al., 2000).

2.5 Maturation enzyme activity assay and SDS-activation

Enzymatic activity was monitored on immunoblots using polyclonal anti-aleurain antibodies; maturation enzyme activity would cause disappearance of the 42 kDa proaleurain band, with the corresponding appearance of a 32 kDa band representing mature aleurain. As a standard, a sample of barley leaf extract containing authentic aleurain was analysed on the same gel with assay samples. We subsequently found that the most purified preparation of maturation enzyme needed to be SDS activated before optimal activity could be obtained (see Results). SDS activation was achieved by incubating 50µl of the enzyme preparation with 5 µl of 20% SDS in a 1.5 ml plastic centrifuge tube at room temperature for 1 hour. It was important that the tube contents were mixed gently by flicking the tube every 15 min during the 1 hour incubation at room temperature. To form mixed micelles with the SDS, 49.5 µl 10% v:v diluted Nonidet-P-40 (NP-40; IGEPAL-CA 630, Sigma, St Louis, MO) in water was added with 445.5 µl Buffer B (50mM succinic acid pH 4.5, 1 mM $CaCl_2$, 1 mM DTT, 0.2 % NP-40). The SDS-activated enzyme preparation was incubated 5 min at room temperature before being added to reaction mixture containing proaleurain substrate at a ratio of 50µl/2µl or as otherwise described, and incubated at room temperature for the necessary duration.

2.6 Preparation of a crude extract

The cauliflower florettes were converted into a fine powder using a Waring® blender and liquid nitrogen. To extract the soluble proteins Buffer A (75 mM sodium acetate pH 5.0, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 1 mM DTT, 0.5 mM EDTA, 0.2 µg/ml Phenylmethylsulfonyl Fluoride) was added to the powdered florettes (1 liter /500g of cauliflower). The extract was stirred at 4ºC for 30 min and then it was passed through 2 layers of cheesecloth.

2.7 Ammonium sulfate precipitation

Solid $(NH_4)_2SO_4$ was added to the cauliflower extract to 50% saturation, and the resulting precipitate was removed by centrifugation at 10,000 x g for 15 min. More $(NH_4)_2SO_4$ was added to the supernatant to 60% saturation, and this precipitate was collected by centrifugation and resuspended in 20 ml/500g starting material of buffer B (50 mM succinic acid pH 4.5, 1 mM $CaCl_2$, 1 mM DTT, 0.2 % NP-40). After extensive dialysis against several changes of buffer B, the preparation was clarified by filtration using a vaccum driven disposable filtration system (0.22 μm, Stericup™, Millipore corporation, Bedford, MA). All these steps were done at 4º C.

2.8 Mono S chromatography

The 50 to 60% $(NH_4)_2SO_4$ fraction was loaded onto a cation exchange column (10x100mm) containing Mono S (Amersham Biosciences, Piscataway, NJ) equilibrated with buffer B followed by washing with the same buffer. For the batch elution, buffer C was used (buffer B + 1 M NaCl). For change of buffer, the sample was dialysed extensively against several changes of buffer B and was clarified by filtration (Millex®GP 0.22 μm, Millipore)

2.9 Bacitracin column

An affinity column was produced as described (van Noort et al., 1991). The matrix for the affinity column was 400 mg of Cyanogen bromide-activated Sepharose 4B (Sigma, St Louis, MO) that was washed and reswelled with 1 mM HCl using a scintered glass funnel. The ligand was 40 mg of Bacitracin (Sigma, St Louis, MO) dissolved in 3ml of 0.1 M $NaHCO_3$/ 0.5 M NaCl, pH 8.3. The matrix and the ligand were mixed overnight at 4º C by rotation. The next day the bacitracin solution was drained from the resin. To block unreacted cyanogens bromide sites

the matrix was incubated with 1M Tris-HCl, pH 8 for 2 hours at room temperature then washed with 0.1 M $NaHCO_3$/ 0.5 M NaCl, pH 8.3.

For use in enzyme purification the sample from the Mono S column was extensively dialysed against buffer B at 4º C and then applied to the bacitracin column previously equilibrated with the same buffer. The bacitracin column was washed with buffer B until the A_{280} of the eluent was <0.01. Finally the bound proteins were eluted with buffer C + 20% ethanol and then extensively dialysed against buffer B at 4º C.

2.10 In gel trypsin digestion

The protease fraction from the bacitracin column (50µg) was separated on a 4%-20% acrylamide gradient Express gel (ISC BioExpress, Kaysville, UT). The gel was incubated for at least 20 min in a fixing solution (45 ml methanol, 1ml acetic acid and 54 ml water) and then stained with 17% (w/v) ammonium sulfate, 34 % methanol, 0.5% acetic acid, 0.1% (w/v) Coomassie G250 (Sigma). Protein bands were visualized by destaining in water.

For trypsin digestions, the protein bands were cut out from the stained gel and prepared as described (http://donatello.ucsf.edu/ingel.html). The gel fragments were diced into small pieces and placed in pyrogen safe 1.5 ml plastic centrifuge tubes (Phenix Research Products). To remove excess SDS the gel slices were washed three times with 100 µl (or enough to cover) of 25 mM NH_4HCO_3/50% acetonitrile. The supernatant was discarded and the gel pieces were speed vacuumed to complete dryness (around 20 min). For tryptic digestion of the proteins in the gel slices ,a trypsin solution (12.5 ng/µl trypsin in 25mM NH_4HCO_3) was added to barely cover the gel slices (5-25µl). After an incubation of 10 min on ice, enough 25mM NH_4HCO_3 was added to cover the gel slices to insure even reswelling of the gel slices. The tubes were centrifuged briefly, followed by overnight incubation at 37 ºC. To recover the tryptic peptides

from the gel slices, the trypsin solutions were transferred to new tubes, then the slices were covered with 30 µl of 50% acetonitrile / 5% formic acid, vortexed 20-30 min, pelleted by centrifugation and then sonicated 5 min (repeated once) before MS analysis, 5 µl of 5% formic acid was added to the sample.

2.11 Mass spectroscopy sequencing

Tryptic digests were applied to a nano-flow POROS R2 (C-18) packed capillary using a Famos/Ultimate HPLC system (Dionex, Belerica MA) with 75 nL pre-column splitter. The elution was then monitored by mass spectrometry employing nano-electrospray using a Q-TOF (Micromass, Manchester, UK). Data dependent MS conditions were used to examine the sequence of the peptides by utilizing an initial survey scan of all the eluting peptides (MS), followed by a selection of significant ions for subsequent individual tandem (MS/MS) analysis which enabled fragmentation of the selected ion. Data base comparisons of the peptide mass fragments and MS/MS fragmentation patterns with the Arabidopsis genome were made with the aid of Mascot software, MS/MS Ion Search (http://www.matrixscience.com). These procedures were performed at EMSEL, Tri-cities, Washington. Additionally, manual sequencing was performed by identifying and matching specific *b* and *y* ions from select MS/MS spectra.

2.12 Expression of the cysteine protease with granulin domain in the *E. coli* periplasmic space

The following nucleotide primers were used to amplify by means of the polymerase chain reaction (PCR) a DNA fragment encoding the coding sequence of a cysteine protease with a granulin domain (template used was the cDNA U16700 ordered from Stanford DNA Sequence and Technology Center)

(1) 5’ primer (*Xho* I):

5’ GGG *CTC GAG* GTC ACA GCG ACG GAG ACT A 3’

(2) 3’ primer (FLAG epitope and *Sal* I)

5’ GGG *GTC GAC* CTT ATC GTC GTC ATC CTT GTA ATC AAC AAC CGG TGA AGG CGC 3’

The resulting PCR fragment contained the coding region for the propeptide and the cysteine protease domain (excluded was the C-terminal cysteine rich sequence of the granulin domain). This fragment was inserted into the *Xho* I to *Sal* I interval of plasmid pBAD/gIII (Invitrogen, Carlsbad, CA) so that its coding sequence was in frame with a terminal six His residues, and with a phage signal peptide sequence at its N-terminus.

For production of the protein, overnight cultures (grown in Luria Broth and 50 μg/ml carbenicillin) were diluted 100-fold in the same medium and were allowed to grow at 37°C until early-log phase ($A_{600} \simeq 0.5$). Expression of U16700 was induced by the addition of 0.002 % of L-arabinose. The induced cell culture was transferred to a 29.5ºC incubator and kept growing for 4.5 hours.

2.13 Genomic DNA isolation and cloning

Preparation of genomic DNA from cauliflower florettes or *Arabidopsis* was performed as described by Asemota (Asemota, 1995).

Cauliflower full length Künitz inhibitor sequence was obtained by amplifying the cauliflower genomic DNA with the following nucleotide primers

(1) 5’ primer

5’ GAT GTC ATC ATT CCC ATT GGT C 3’

(2) 3’ primer

5’ AAT CCA GGT ACA TGC ATG AGG 3’

Arabidopsis full length Künitz inhibitor sequence (gi|27530934) was obtained by amplifying the *Arabidopsis* genomic DNA (graciously provided by Marianne Poxleitner) with the following nucleotide primers

(1) 5’ primer

5’ GAT CAG AAT GTG GCA TGA GG 3’

(2) 3’ primer

5’ AAC GGG AAC ATC CTT AGA CC 3’

The procedure was successful in both instances because the two genes lack introns.

2.14 Expression of a Künitz inhibitor in the *E. coli* M15 cells (construct # 1)

The following nucleotide primers were used to amplify by PCR a fragment encoding the coding sequence of a Künitz type protease inhibitor (template used was the *Arabidopsis* genomic DNA PCR product described above)

(1) 5’ primer (*Bam* HI):

5’ GG GGA TCC CAC GGA AAT GAA CCG 3’

(2) 3’ primer (*Hind* III)

5’ GG AAG CTT TCA ACC CGG GAA GTA 3’

The resulting PCR fragment contained the coding region for the propeptide and the Künitz inhibitor domain. This fragment was inserted into the *Bam* HI to *Hind* III interval of plasmid pQE-30 (Qiagen GmbH, Germany) so that its coding sequence was in frame with a terminal six His residues.

For production and harvest of the protein, we followed the protocols given by the manufacturer (The QIA expressionist, Qiagen, Growing large scale expression culture and Denaturing purification of insoluble proteins).

2.15 Refolding of E.coli M15 expressed the Künitz inhibitor

The refolding procedure was performed as described (http://www.nwfsc.noaa.gov/protocols/inclusion.html). The principle was to remove the denaturing agent by dialysis in the presence of 3(1-pyridinio)-1-propane sulfonate (Sigma).

2.16 Expression of a Künitz inhibitor in the *E. coli* periplasmic space (construct # 2)

The following nucleotide primers were used to amplify by PCR a fragment encoding the coding sequence of a Künitz type protease inhibitor (template used was Arabidopsis genomic DNA PCR product as for construct # 1)

(1) 5' primer (*Bam* HI):

5' GG GGA TCC CAC GGA AAT GAA CCG 3'

(2) 3' primer (*Sal* I)

5' GG GTC GAC ACC CGG GAA GTA TAA 3'

The resulting PCR fragment contained the coding region for the propeptide and the Künitz inhibitor domain. This fragment was inserted into the *Bgl* II to *Sal* I interval of plasmid pBAD/gIII B (Invitrogen, Carlsbad, CA) so that its coding sequence was in frame with a terminal six His residues, and with a phagic signal peptide sequence at its N-terminus.

This plasmid was transformed into TOP 10 cells (*E.coli* periplasmic expression system, Invitrogen) using calcium chloride (Sambrook and Russell, 2001). The transformed cells were selected for their resistance to 50 μg/ml carbenicillin. For production of the protein, overnight

cultures (grown in LB and 50 μg/ml carbenicillin) were diluted 100-fold in the same medium and were allowed to grow at 37°C until early-log phase ($A_{600} \simeq 0.5$). Expression of Künitz inhibitor was induced by the addition of 0.002 % of L-arabinose. The induced cell culture was transferred to 29.5ºC incubator and kept growing for 4.5 hours.

2.17 Purification of proteins expressed in the *E. coli* periplasmic space

Periplasmic proteins were isolated based on the protocols provided with The QIAexpressionist kit (Qiagen, Valencia CA, 2nd edition). For extracting the periplasmic space-localized proteins the cells were pelleted by centrifugation and resuspended in osmotic buffer (20 mM Tris-HCl pH 8, 2.5 mM EDTA, 20% sucrose, 200 μg/ml lysosyme), and incubated 15 min at room temperature. The cells were then pelleted at 2000 g for 15 min and the supernatant was stored at -20º C until use. This periplasmic fraction was dialyzed extensively against 50 mM NaH_2PO_4, 300 mM NaCl pH 7.8 and then purified by passing the dialyzed and filtered sample (Millex®GP 0.22 μm, Millipore) through a 1.5 ml Polu Prep® chromatography column (BioRad, Hercules, CA) containing Ni^{2+}-NTA agarose (Qiagen, Valencia CA). The column was washed with 50 mM NaH_2PO_4, 300 mM NaCl pH 7.8 and 10% glycerol pH 6.0 and then eluted with 50 mM NaH_2PO_4, 300 mM NaCl pH 7.8, 10% glycerol and 100 mM imidazole pH 6.0. The eluate was dialyzed extensively against buffer B.

2.18 Silver staining

Siver staining of SDS-PAGE gels was performed as described (http://www4.amershambiosciences.com).

CHAPTER 3 RESULTS

3.1 Purification of the clipping enzyme

3.1.1 Proteolytic assay: visualization of proaleurain conversion to aleurain on immunoblot.

We first had to establish an assay for the proaleurain clipping enzyme. The principle of the assay was to incubate properly folded recombinant barley proaleurain with the sample to be tested, then the incubation mixture would be separated by SDS-PAGE, transferred to a membrane and analyzed by immunoblot using anti-aleurain antibodies. The activity was monitored by observing the shift of mass size of proaleurain (42 kDa) to mature aleurain (32 kDa) as was observed by Holwerda *et al.* (1990). We reasoned as did Holwerda *et al.* (Holwerda et al., 1990) that such an assay would require correctly folded proaleurain and thus would require expression of recombinant proaleurain in a eukaryote cell. In that study, proaleurain was expressed in frog oocytes, a system not available to us.

We first tried to express proaleurain in the yeast strain *Pichia pastoris* using a methanol-inducible promoter. We checked for proaleurain expression in extracts from the transgenic yeast by SDS-PAGE and immunoblot (Figure 3.1, lane 3) along with a positive control, barley leaf extract (Figure 3.1, lane 2). A band detected by anti-aleurain antibodies was present in the yeast extract but its mass size corresponded to that of mature barley aleurain (compare lane 2 and 3 in Figure 3.1), and no band of the expected mass size of 42 kDa for proaleurain was present (Figure 3.1, lane 3). Additionally, no proaleurain was present in the yeast cell medium (data not presented). From this result we concluded that yeast expressed proaleurain, but the proaleurain

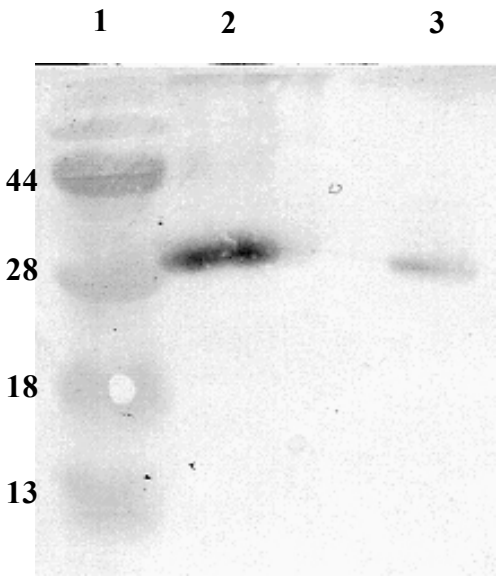


Figure 3.1. Expression of Proaleurain in *Pichia pastoris* .

Proteins were separated by SDS-PAGE and aleurain was visualized on an immunoblot using anti-aleurain antibodies. Lane 1, molecular marker with corresponding mass size in kilodaltons; Lane 2, barley leaf extract; Lane 3, extract of induced yeast cells transformed with proaleurain.

was not secreted and was instead processed by yeast enzymes to its mature size. Thus this expression system was not suitable for our purposes.

We then stably transformed *Drosophila* S2 suspension culture cells to express recombinant proaleurain. The cells synthesized the proenzyme and secreted proaleurain into the culture medium as shown by the immunoblot in Figure 3.2, lane 2. Induction of proaleurain synthesis and its secretion was accompanied by growth of the *Drosophila* cells, with no evidence for cell death. Therefore, secretion of proaleurain would require that it had undergone proper folding in the ER and it would thus be a correct substrate for the clipping enzyme. The medium was then dialyzed and concentrated, and was used without further purification in the proteolytic assay.

Our assay measures conversion of recombinant barley proaleurain to a molecular mass indistinguishable in size from aleurain present in barley leaf extracts. Although we would expect that a protease capable of this activity present in other plant tissues might be similar to the clipping enzyme defined by Holwerda et al. (1990), our expectation might not be correct. For that reason the protease activity that converts proaleurain to an aleurain-sized molecule will be referred as a "maturation" enzyme.

To determine optimum conditions for the proaleurain maturation assay we referred to the previously established protocol (Holwerda et al., 1990). Those authors established that the clipping enzyme was present in barley aleurone extracts. We therefore tested for maturation of the *Drosophila*-expressed proaleurain by an extract from barley aleurone layers, and demonstrated proteolytic processing of the recombinant proenzyme to a size similar to that for mature barley aleurain (Figure 3.3). The maturation reaction was initiated by mixing an

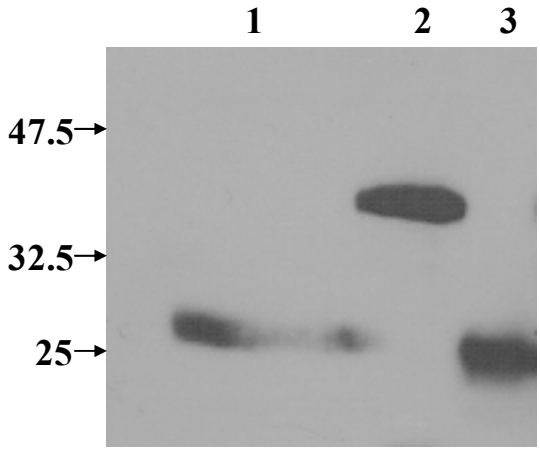


Figure 3.2. Expression of proaleurain in *Drosophila* S2 suspension culture cells.

Proteins were separated by SDS-PAGE and visualized on immunoblot using antialeurain antibodies. Numbers at left indicate the positions of molecular mass markers in kilodaltons. Lane 1, barley leaf extract; Lane 2, *Drosophila* expressed proaleurain; Lane 3, cauliflower extract.

aliquot of the concentrated culture medium containing recombinant proaleurain with the barley aleurone extract and, after incubation at room temperature, the reaction mixture was separated by SDS-PAGE and the shift in mass size of proaleurain (Figure 3.3, lane 1) to mature aleurain (Figure 3.3, lane 3) was observed on immunoblot. This assay allowed us to evaluate for the presence of maturation enzyme activity, so that an optimal source for purification of a maturation enzyme could be established.

3.1.2 Cauliflower florettes used as source for proaleurain maturation enzyme.

The preparation of barley aleurone layers was extremely tedious and would not be suitable for large-scale preparations necessary in enzyme purification, so we therefore searched for an alternate source of the proaleurain maturation enzyme. Although aleurain was first identified and purified from barley (Holwerda and Rogers, 1992; Rogers et al., 1985), all plants appear to express aleurain (Rogers et al., 1997). Therefore, proaleurain maturation enzymes must also be widely present in plants. We had to take into account several factors. Our assay would not be effective if large amounts of endogenous mature aleurain were present. We had to be able to monitor the conversion of *Drosophila* cell-expressed proaleurain into aleurain; if endogenous mature aleurain were present we wouldn't be able to differentiate the endogenous mature aleurain from that produced from the recombinant proenzyme substrate. Additionally anti-barley aleurain antibodies cross react with aleurain homologs from other plants (Rogers et al., 1997).

Broccoli and cauliflower are in the genus *Brassicacea* and therefore are closely related to *Arabidopsis,* the genome of which has been sequenced. This relationship would facilitate identification of the maturation enzyme. Also, broccoli or cauliflower florets are frequently an

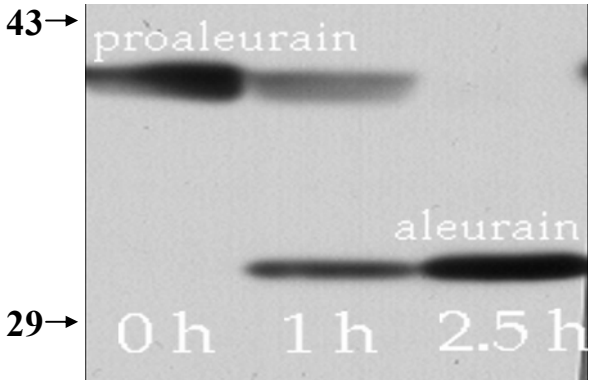


Figure 3.3. Maturation enzyme assay

Drosophila expressed proaleurain was incubated with aleurone layer extract at room temperature for times from 0 to 2.5 hours as indicated at bottom. The samples were then separated by SDS-PAGE and maturation of proaleurain to aleurain was visualized by immunoblot using aleurain antibodies. Numbers at left indicate the positions of molecular mass markers in kilodaltons.

optimal starting source for enzyme purification because they lack cells with large vacuoles, and therefore present fewer problems with protein degradation in initial extracts. When tested for their endogenous content of aleurain on an immunoblot using antialeurain antibodies; an extract from broccoli florettes had a large amount of broccoli aleurain (Figure 3.4, lanes 1a and 2a, indicated by arrows), while cauliflower florettes had a weaker background (Figure 3.4,compare lane 1a versus lane 6a). In an effort to minimize the amount of endogenous aleurain that would appear on the maturation enzyme assay immunoblot, the cauliflower and broccoli samples were diluted ten fold before being incubated with recombinant proaleurain (Figure 2.4, lanes 2ab, 4ab, 7ab, 9ab). Proteins present in a 60-80% (Figure 3.4, lanes 6 and 7) and a 40-60% (Figure 3.4, lanes 8 and 9) ammonium sulfate precipitation from an extract of cauliflower florettes showed maturation activity towards the *Drosophila* expressed proaleurain (Figure 3.4, lanes 6-9b). Moreover the mature aleurain obtained with cauliflower maturation enzyme(s) was indistinguishable in size on immunoblots from the barley leaf mature aleurain (data not shown). Hence, cauliflower florettes were used as the source of proaleurain maturation enzyme.

We noticed the presence of a band that was situated between the proaleurain band and aleurain band on the maturation assay immunoblot when cauliflower extract was used (Figure 3.4, open arrow). This band was absent when aleurone extract (Figure 3.3) was used for the maturation enzyme assay. An explanation will be given in section 3.2.3.

The maturation enzyme assay as presented so far is qualitative. It allowed the identification of active fractions in various purification steps as described below, and it allowed comparison of different buffers to establish conditions under which the enzyme would be more stable. A central limitation is the fact that production of mature aleurain by the activity of a

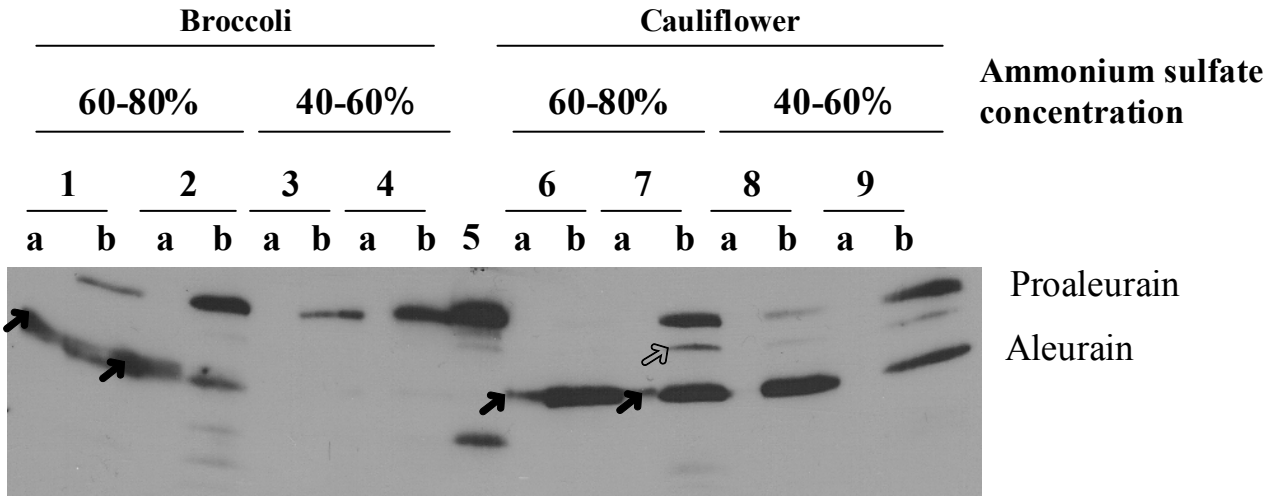


Figure 3.4. Comparison of broccoli or cauliflower extract activity on proaleurain maturation

Proteins were purified from a broccoli extract with ammonium sulfate at 60 to 80% saturation (lanes 1, 2) or 40 to 60% (lanes 3, 4). Proteins were precipitated from a cauliflower extract with ammonium sulfate at 60 to 80% saturation (lanes 6, 7) or 40 to 60% saturation (lanes 8, 9). Resuspended proteins were used without dilution (lanes 1, 3, 6, 8) or were diluted ten fold (lanes 2, 4, 7, 9) before being incubated for 1 hour at room temperature with *Drosophila* expressed proaleurain (lanes indicated with b) or without proaleurain (lanes indicated with a) . Lane 5, *Drosophila* expressed proaleurain was incubated without any extract for 1 hour at room temperature. Endogenous aleurain and aleurain produced from *Drosophila* expressed proaleurain by maturation activity were visualized on immunoblot using anti-aleurain antibodies. The presence of endogenous brocoli or cauliflower aleurain in the extracts is indicated by solid arrows. The open arrow indicates a separate proaleurain protolytic fragment that is larger than mature aleurain. Names at right indicate the positions of the corresponding proteins.

correct maturation enzyme competes for substrate with the proteases that nonspecifically degrade proaleurain.

For that reason, qualitative and then quantitative results, will be presented for individual purification steps. The latter results will indicate that little purification was achieved, a conclusion that we believe is an artifact caused by unusual features of the maturation enzyme and/or the co-purification of an inhibitor of the protease.

3.1.3 Purification of proaleurain maturation enzyme from cauliflower florettes

The general strategy was to prepare an extract of powdered cauliflower florettes that would be our starting material and referred as the crude extract. We found that a pH of 4.5 and the presence of DTT were important initially in stabilizing the maturation enzyme activity.

- **Ammonium sulfate fractionation**

We started our purification with an ammonium sulfate precipitation step. Ammonium sulfate saturated pellets were resuspended and dialyzed to remove of the ammonium sulfate salt. To check which sample contained the maturation enzyme, these samples were incubated with recombinant proaleurain, resulting in the processing to produce a mature aleurain was observed on an immunoblot (Figure 3.4). No activity was observed in 0-40% ammonium sulfate precipitated samples from either broccoli or cauliflower (data not presented). To differentiate possible endogenous aleurain (Figure 3.4, lanes 6a, 7a) from aleurain that resulted from the processing of recombinant proaleurain (Figure 3.4, lanes 6b, 7b), each assay incorporated a pair of samples: one represented the enzyme source incubated in the absence of recombinant proaleurain (Figure 3.4, lanes designated a), while the other represented enzyme plus substrate proaleurain (Figure 3.4, lanes designated b).

The best result was that originating from the 40-60% ammonium sulfate precipitation of proteins from cauliflower extracts: there was maturation enzyme activity (Figure 3.4, lanes 8b, 9b) but there was little amount of endogenous aleurain present (Figure 3.4, lane 8a).

- **Mono S chromatography**

Protein from the 40-60% ammonium sulfate fraction were separated by chromatography on a Mono S cation exchange column. Proteins retained on the column were eluted with a linear 0-1 M sodium chloride gradient. The fractions with activity were 4-10 of the 20 collected (Figure 3.5). No endogenous aleurain was observed in any fraction (Figure 3.5, lanes -P) which facilitated visualization of the mature aleurain (Figure 3.5, lanes +P) resulting from processing of the recombinant proaleurain substrate. Because the number of fractions containing the maturation activity was nearly half of the collected fractions, for subsequent purification efforts we decided to do a batch elution of the column with 1 M sodium chloride instead of a gradient elution. The eluant was then dialyzed to get rid of the NaCl and to change the buffer from pH 4.5 to 5. The latter step was necessary for optimal use of the Poros HQ anion exchange column.

- **Poros HQ chromatography**

The sample from the Mono S chromatography was then applied to a Poros HQ column and eluted with a gradient of 0-1 M NaCl. The fractions were assayed for activity as before. Maturation activity was found in fraction 1 as well as in fractions 8-10 (Figure 3.6). Fraction 1 had a larger amount of mature aleurain, while fractions 8-10 had a much lower amount, as concluded from the intensity difference of the aleurain bands on the immunoblot. On the immunoblot for fractions 8-10 (Figure 3.6), the intensity of the signals for recombinant proaleurain was approximately equal to

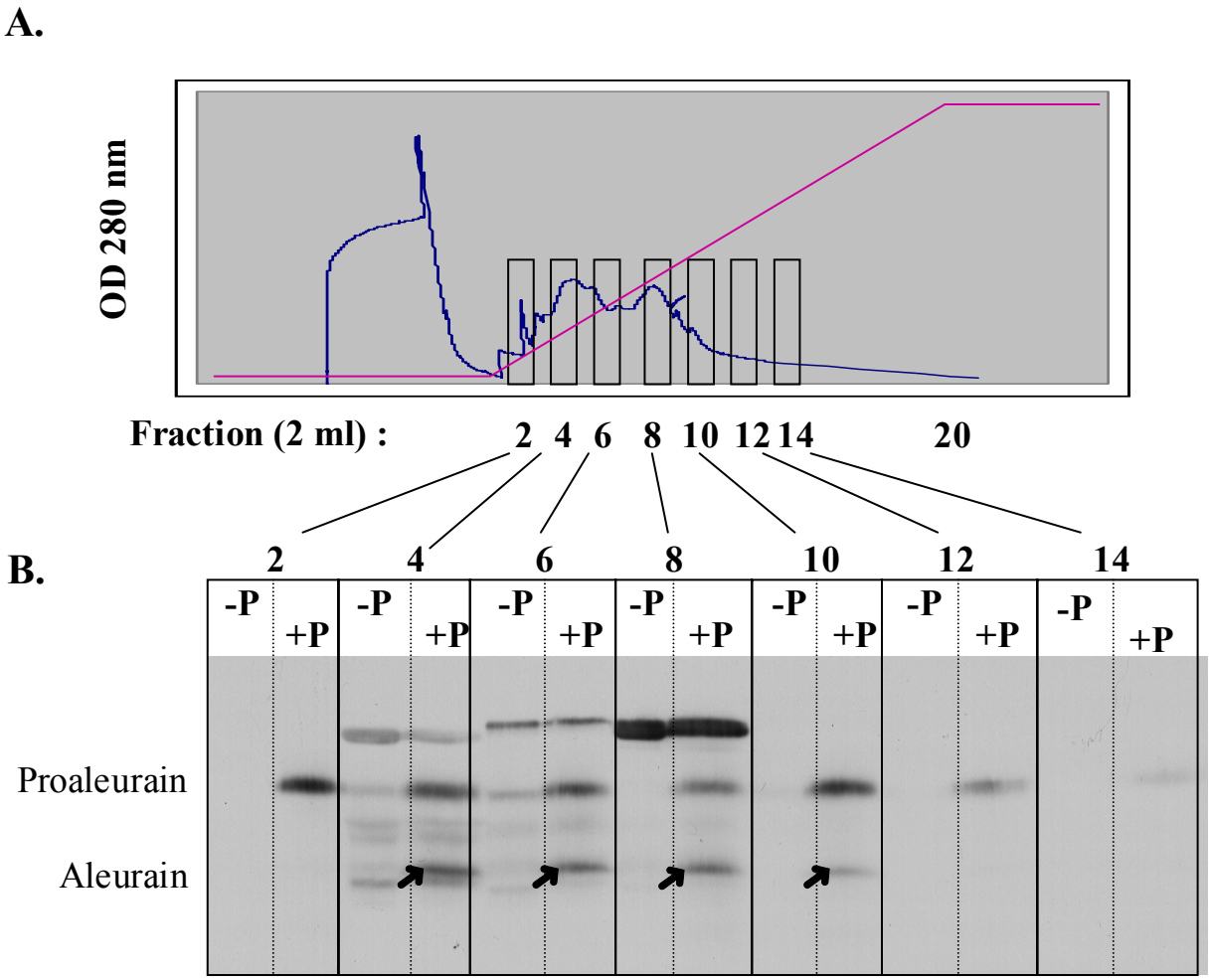


Figure 3.5. Mono S chromatography

(A) Chromatogram of the protein elution profile. The A280 is shown by the black line with the column wash step on the far left. The red line represents the linear elution gradient from 0 to 1 M sodium chloride. Fractions that were assayed in (B) are indicated by rectangles. (B) Aliquots of fractions eluted from the Mono S column were incubated at room temperature for 15 min without (-P) or with *Drosophila* expressed proaleurain (+P). The Mono S samples with activity converted proaleurain to aleurain as visualized by immunoblot as indicated by the arrows. Lanes marked "-P" had no substrate and therefore gave us an indication of the small amount of endogenous cauliflower aleurain present as background in each fraction. We concluded that the peak of maturation enzyme activity eluted in fractions 4-10.

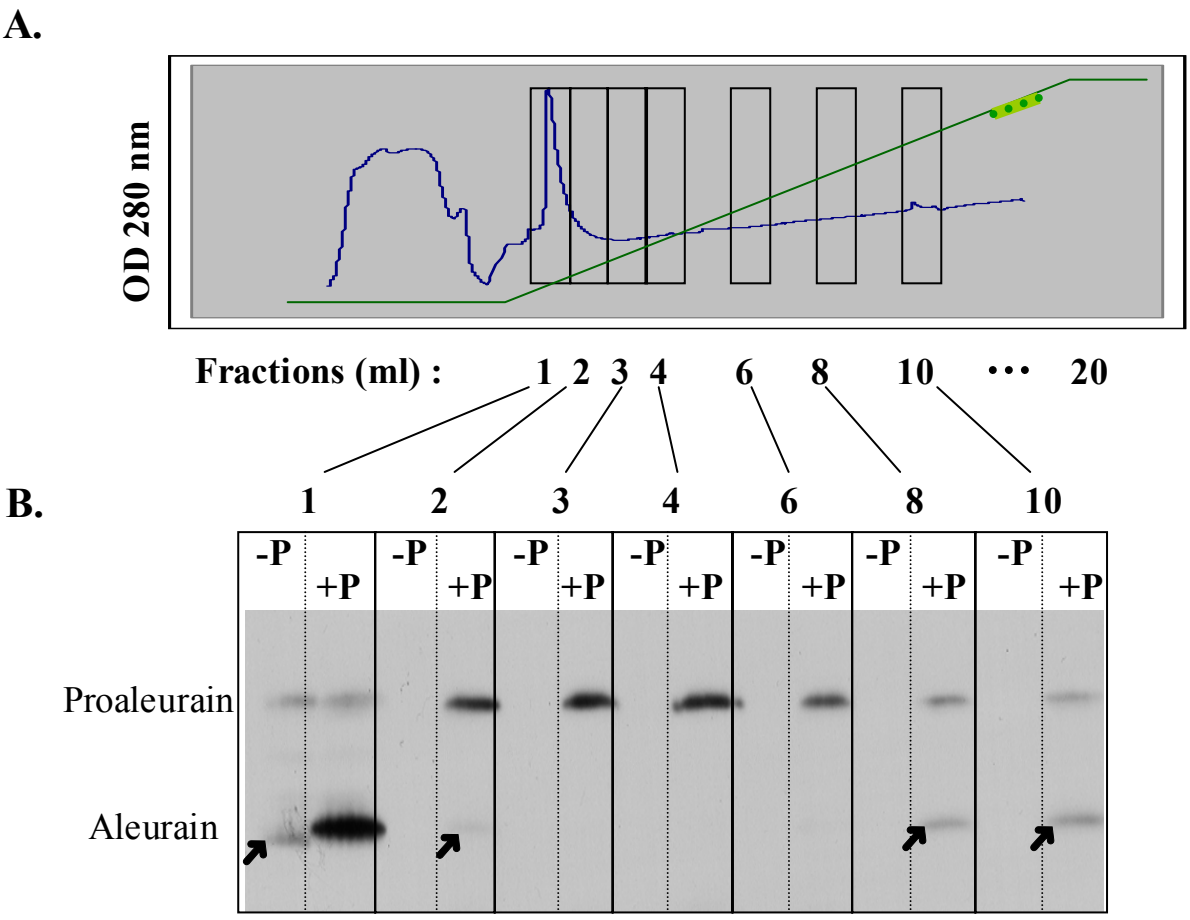


Figure 3.6 Poros HQ chromatography

(A) FPLC (280 nm) chromatograms of the protein elution profile. The A280 is shown by the black line with the column wash step on the far left. The green line represents the linear elution gradient from 0 to 1 M sodium chloride. Fractions that were assayed in (B) are indicated by rectangles. (B) Aliquots from fractions eluted from the Poros HQ column were incubated at room temperature for 15 min without (-P) or with *Drosophila* expressed proaleurain (+P). The fractions with activity converted proaleurain to aleurain as visualized by immunoblot indicated by arrows. Lanes marked "-P" had no substrate and therefore gave us an indication of the amount of endogenous cauliflower aleurain present as background in each fraction. We concluded that the peaks of maturation enzyme activity eluted in fractions 1 and 8-10.

the intensity of recombinant proaleurain in fraction 1 (Figure 3.6), whereas the intensity of the signal for mature product, aleurain, in fraction 1 was much more intense then the signal observed for aleurain in fractions 8-10. The likely explanation is that proaleurain, in fractions 8-10, was partially degraded by another protease, and that only a fraction was converted to aleurain by the maturation enzyme.

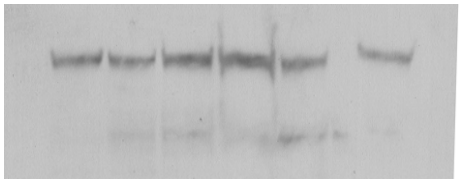
As we gained experience in these purification steps, we found that the presence of 0.2% NP-40, a non-ionic detergent, helped stabilize the maturation enzyme activity. The steps presented here were performed in the presence of NP-40.

- **Purification parameters**

The nature of our assay limits our ability to define purification parameters for the following reasons. (1) Nonspecific protease activity may degrade the proaleurain substrate, particularly in the tissue extract and ammonium sulfate precipitated fractions. With further purification this activity was minimized. (2) In the cauliflower tissue extract, a large but variable amount of cauliflower aleurain was present. These two features limited our ability to quantify the product produced in the assay. In an attempt to estimate the purification levels we had achieved we used densitometry of signals from immunoblots from the assays to quantify the amount of proaleurain consumed relative to aleurain generated. We compared the signals from known amounts of aleurain expressed as a recombinant protein in *E. coli* and purified to homogeneity (Rogers et al., 1997)to those from our recombinant proaleurain to estimate its concentration in our standard assay. We could then define a unit of activity as that generating one ng of aleurain per minute.

For initial analyses, arbitrary densitometry units were graphed on a y-axis, relative to the time of incubation on the x-axis. As shown in Figure 3.7, when the 40-60% ammonium sulfate

A.



B.

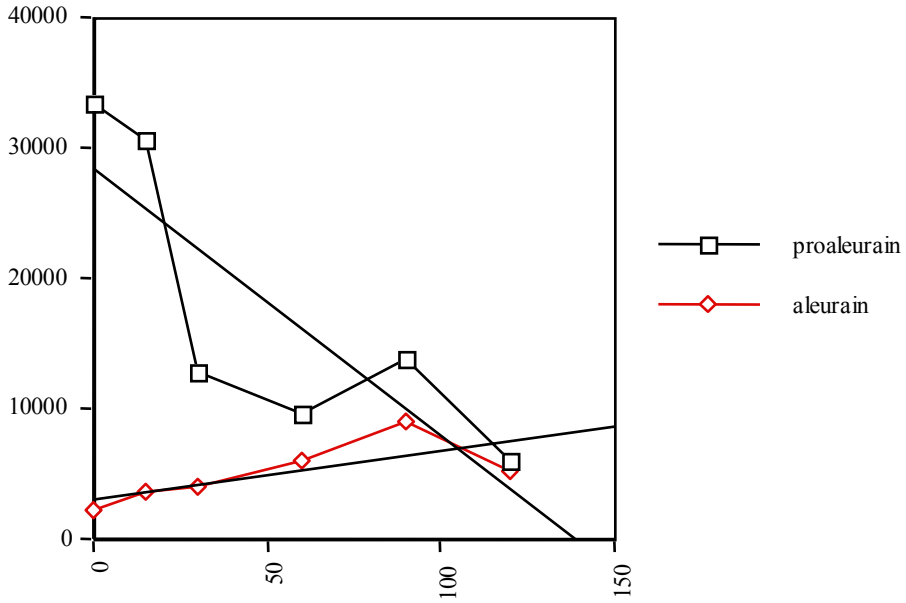
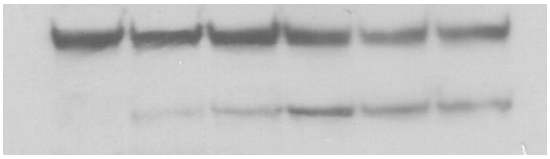


Figure 3.7. Quantification of maturation enzyme activity in the 40-60% ammonium sulfate precipitate.

(A) Immunoblot visualizing maturation enzyme activity of the 40-60% ammonium sulfate saturated fraction incubated with *Drosophila* expressed proaleurain for the indicated minutes at room temperature. (B) Analysis by densitometry of the immunoblot signals from proaleurain (squares) and aleurain (diamonds) in arbitory units (y axis) versus the time of incubation (minute, x axis). Linear best fit analysis of each set of data points gave us the value of the corresponding slope used in Table 3. 1.

A.



B.

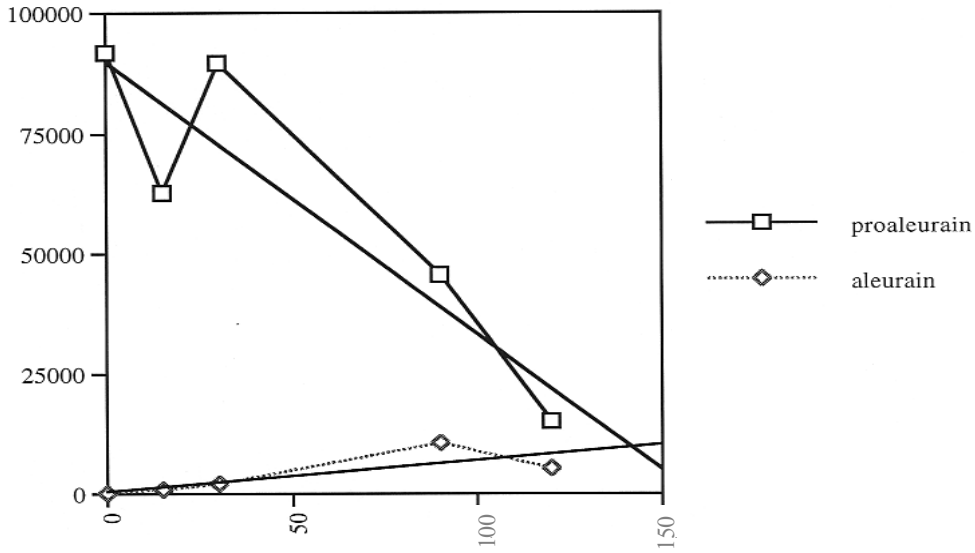


Figure 3.8. Quantification of maturation enzyme activity in pooled fractions eluted from the MonoS column.

(A) Immunoblot visualizing maturation enzyme activity of the pooled fractions eluted from the MonoS column incubated with *Drosophila* expressed proaleurain for the indicated minutes at room temperature. (B) Analysis by densitometry of the immunoblot signals from proaleurain (squares) and aleurain (diamonds) in arbitory units (y axis) versus the time of incubation (minute, x axis). Linear best fit analysis of each set of data points gave us the value of the corresponding slope used in Table 3. 1.

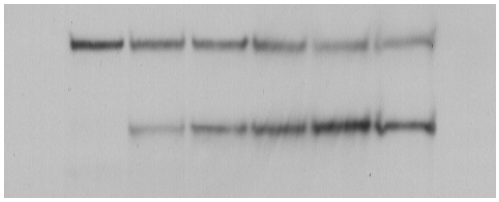
precipitate was assayed for maturation enzyme activity, proaleurain substrate disappeared rapidly while the amount of aleurain generated amounted only about 1/3 of that expected. As the antibodies used for detection recognize only the cysteine protease segment of proaleurain and not the prosegment (Rogers et al., 1997), the signal for proaleurain consumed versus the signal for aleurain produced should be equivalent during the maturation process. A similar discrepancy was observed for assays using fractions eluted from the Mono S column (Figure 3.8). Only when the Poros HQ chromatography step was reached did we observe a relatively proportional amount of aleurain generated to substrate consumed (Figure 3.9).

By projecting a linear fit onto the data points from each densitometry analysis for aleurain produced, we attempted to estimate units of maturation enzyme activity present at each purification step. These numbers were summarized in Table 3.1. We recovered only 14% of our enzyme activity between the ammonium sulfate and Mono S steps, with no purification. There was only an 11% recovery of activity from the Poros HQ column relative to activity present in the ammonium sulfate fraction, with only a 9-fold purification. While it is recognized that there are considerable uncertainties inherent in these numbers, we concluded that we lost most of our enzyme activity in the Mono S step. One possible explanation was that the maturation enzyme was unstable. We therefore sought an affinity chromatography approach to try to speed up the purification process as well as achieve a larger purification increment in one step.

- **Bacitracin column chromatography**

Bacitracin, a peptide antibiotic, is also a nonspecific inhibitor of many different proteases and has been used as an affinity ligand to purify proteases from all classes (van Noort et al., 1991). We coupled bacitracin to cyanogen bromide-activated sepharose as described (van Noort et al., 1991) and applied active fractions from MonoS chromatography. After extensive

A.



B.

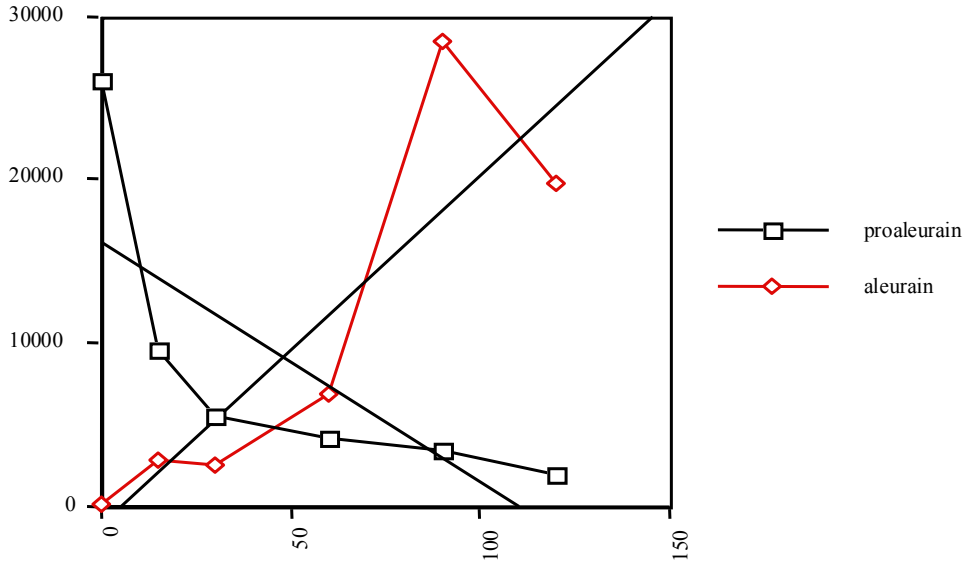


Figure 3.9. Quantification maturation enzyme activity in fraction eluted from the Poros HQ column. (A) Immunoblot visualizing maturation enzyme activity of the fraction eluted from the Poros HQ column incubated with *Drosophila* expressed proaleurain for the indicated minutes at room temperature. (B) Analysis by densitometry of the immunoblot signals from proaleurain (squares) and aleurain (diamonds) in arbitory units (y axis) versus the time of incubation (minutes, x axis). Linear best fit analysis of each set of data points gave us the value of the corresponding slope used in Table 3. 1.

	Slope Units (ng/min)	Total † Units	Protein mg	Specific‡ Activity	Purification fold
Extract	?	?	800	?	?
Ammonium Sulfate	36	3,600	43	83,700	1.00
Mono S	65	510	12	50,000	0.60
Poros HQ	214	400	0.6	775,600	9.30

Table 3.1. Purification of maturation enzyme activity : Quantification of specific activities present at each purification step.

† ng Aleurain produced min^{-1} ml^{-1}

‡ ng Aleurain produced min^{-1} ml^{-1} mg^{-1} proteins

Slope values were obtained from Figures 3.7-9. The crude extract had too much endogenous aleurain present so that it was not possible to estimate the amount of maturation enzyme activity in a quantitative manner.

wash the column was eluted and the fractions tested for maturation activity. The eluate contained maturation enzyme activity but the activity rapidly was lost after storage for a day on ice (data not shown). The proteins present were visualized by coomassie blue stain after SDS-PAGE (Figure 3.10 B), and there were fewer bands present compared to the sample purified from the Mono S column (Figure 3.10 A, Mono S).

The first time we tried the bacitracin column, no activity was obtained in the elution fraction when 1 M sodium chloride was applied for elution. We therefore decided to strip the column by using SDS running buffer to learn if some proteins were still retained on the column. Proteins were present in the SDS eluent as observed on a silver stained gel of the sample separated by SDS-PAGE (data not shown) but when that sample containing 2% SDS was assayed for the maturation enzyme no activity was detected. We then decided to counteract the SDS effect by diluting the sample ten fold with buffer B containing O.2% NP-40. When this mixture was assayed, abundant maturation activity was present. We had rescued the activity, and later used that "SDS treatment" to boost activity in samples that with time lost their activity (Figure 3.11).

3.2 Characterization of a proaleurain maturation enzyme

3.2.1 Influence of SDS on the activity

The sample eluted from the bacitracin column lost its weak activity after 1 day. That activity could be rescued (Figure 3.11) by incubation of the sample in 2% SDS solution followed by ten fold dilution with a solution containing 0.2% NP-40. This SDS activation phenomenon further complicated our evaluation of the specific activity of the maturation enzyme present at the various purification steps. However, it was clear that at least a part of the enzyme activity

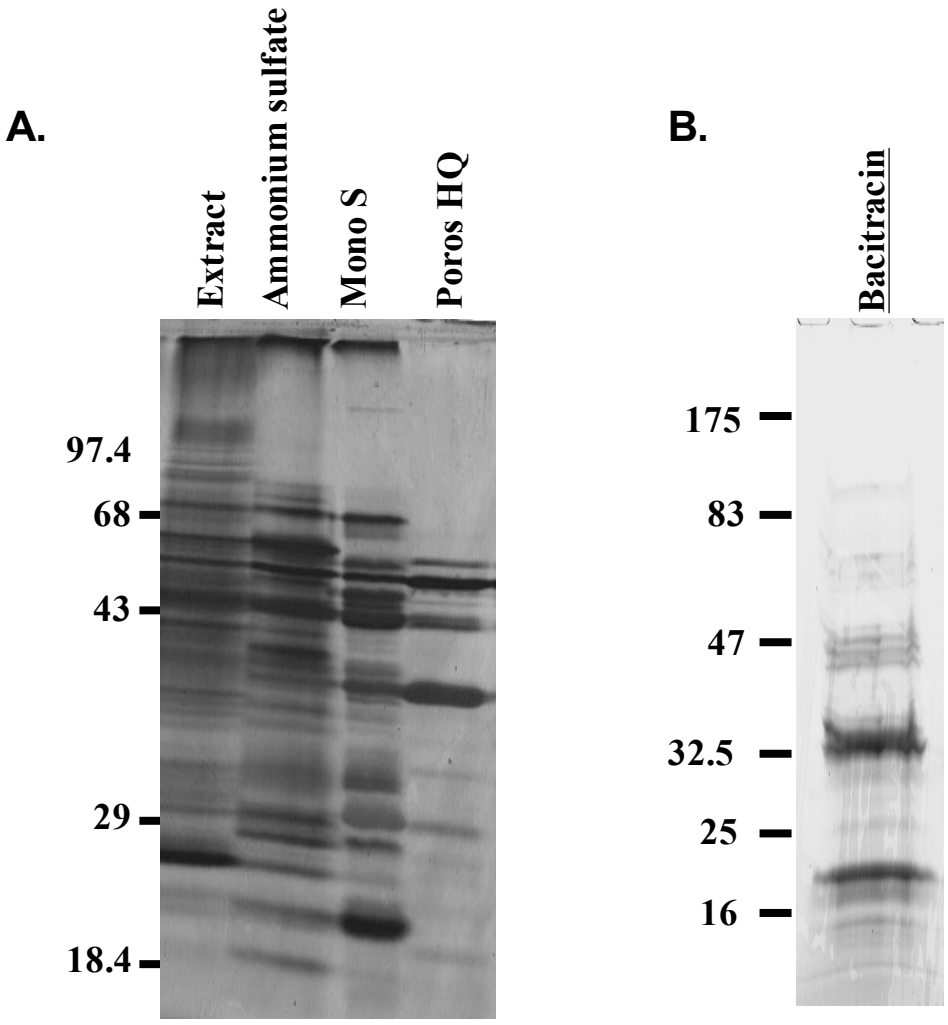


Fig 3.10. Analysis of maturation enzyme purification

SDS-PAGE analysis of fractions from each step of the maturation enzyme purification. (A) Analysis of 10 μg of the cauliflower crude extract; 6.5 μg of protein from the 50 to 60% saturated ammonium sulfate fraction; 6.6 μg of protein from fractions after Mono S chromatography; μg of proteins from fraction after Poros HQ chromatography. Gel was silver stained. (B) 30 μg of protein from the fraction eluted from the bacitracin column. Gel was stained with colloidal coomassie blue. Numbers at left in A and B indicate the positions of molecular mass markers in kilodaltons.

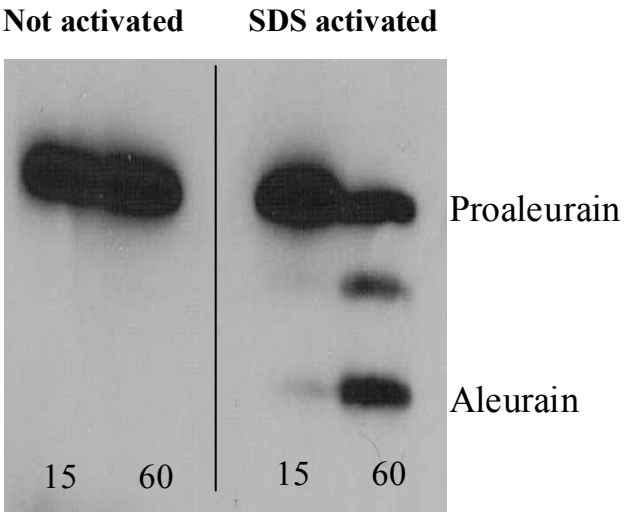


Figure 3.11. In vitro effect of SDS treatment on maturation enzyme activity

Drosophila expressed proaleurain was incubated for 15 or 60 min (numbers at bottom) at room temperature with Cauliflower extract eluted from the bacitracin column that either had been SDS activated or not (as indicated at top). The conversion of proaleurain to aleurain was detected by immunoblot using aleurain antibodies. Names at right indicate the positions of the corresponding proteins.

lost during purification could be explained by inactivation that could be reversed by SDS treatment.

This SDS activation could be explained either by potential denaturation of the proenzyme to cause removal of the propeptide from the active site (detergent-induced activation), or by dissociation of an inhibitor bound to the mature protease active site, restoring activity (detergent-induced loss of inhibition). These possibilities are discussed in detail in Chapter 4.

3.2.2 Effect of pH on maturation activity

When the bacitracin elution sample was SDS treated under various pH conditions and then assayed for processing of proaleurain to aleurain activity was observed at pH 4.5 but not at pH 5 or above (Figure 3.12).

3.2.3 Molecular size

We used two methods to estimate the size of proaleurain maturation enzyme. First, using a size exclusion column we estimated a mass size range of 29-66 kDa, for the maturation activity found in aleurone layers (Figure 3.13). For the cauliflower enzyme, we had the advantage of knowing that its activity was not affected by the presence of SDS. Protein eluted from the bacitracin column was suspended in SDS-PAGE sample buffer at room temperature in the absence of reducing agent, and then separated by SDS-PAGE (Figure 3.14A). The gel was cut into slices using the molecular weight markers to define the fragments. Proteins were eluted from the gel slices and assayed for maturation activity. Activity yielding a product the size of aleurain was found in the sample from the gel slice cut between the molecular markers of 32 and 25 kDa (Figure 3.14B, lane 5), thereby providing an estimate for the size of the cauliflower

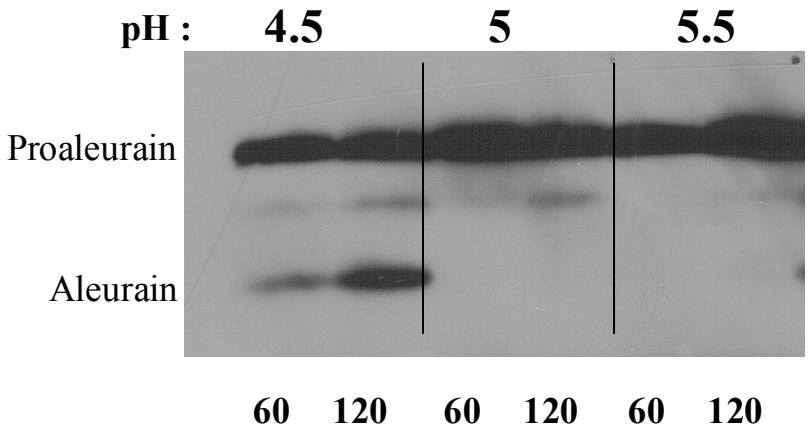


Figure 3.12. pH influence on maturation activity

Sample eluted from the bacitracin column was dialysed against buffers with pH values as indicated at top, then was SDS treated before assay for activity at 60 or 120 min (numbers at bottom). The maturation activity was monitored on an immunoblot using aleurain antibodies. Names at left indicate the position of the corresponding protein.

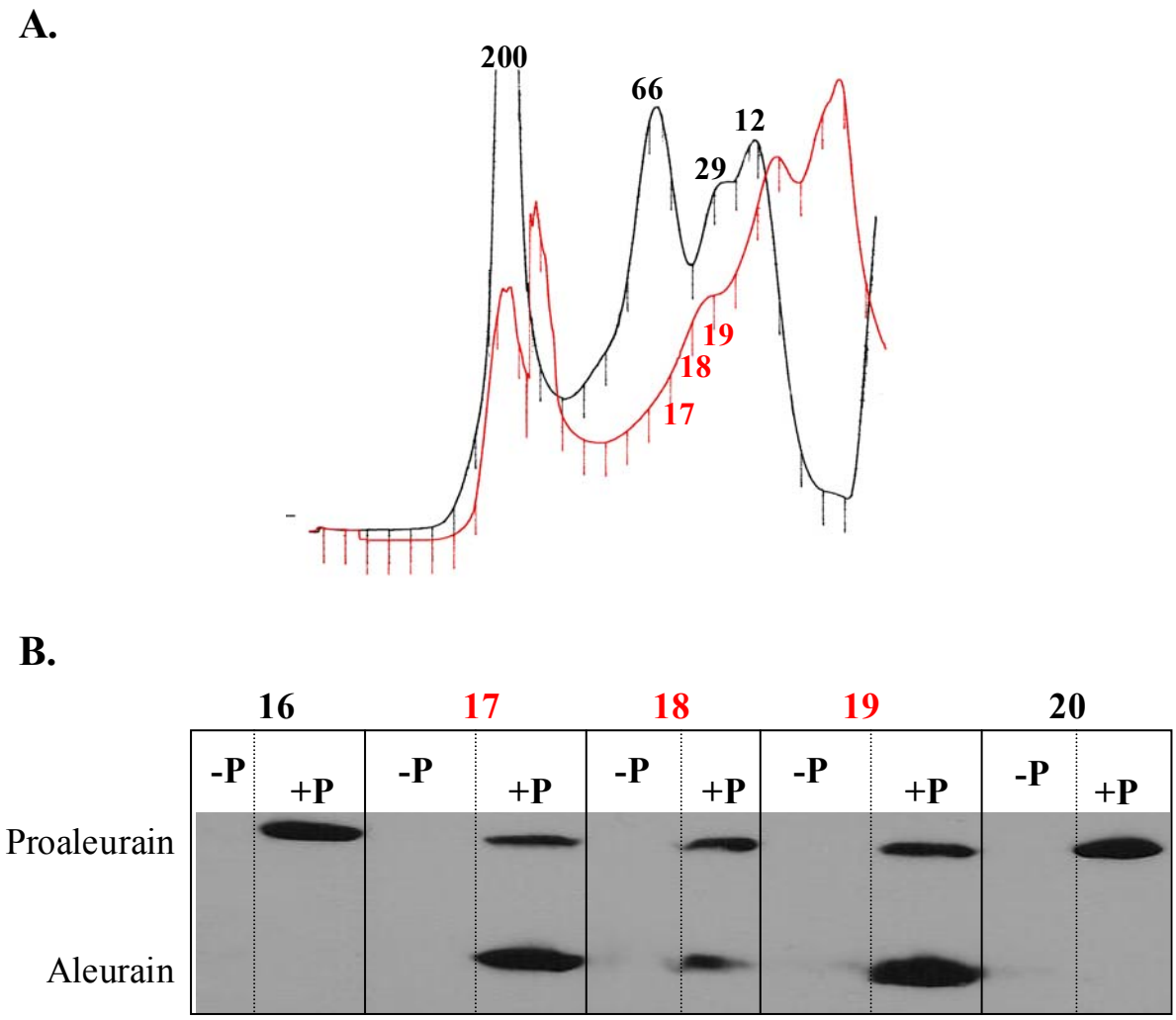


Figure 3.13. Gel filtration chromatography of maturation enzyme activity in barley aleurone layer extract.

(A) Superimposed A_{280} elution profiles from the aleurone extract (red) and of a mixture of standard proteins (black) with their corresponding masses indicated in kilodaltons. (B) Aliquots from the indicated fractions were assayed for maturation enzyme activity. The samples were incubated with *Drosophila* expressed proaleurain for 1 hour at room temperature before being separated on SDS-PAGE and visualized on immunoblot using antialeurain antibodies. Names at left indicate the positions of the corresponding proteins.

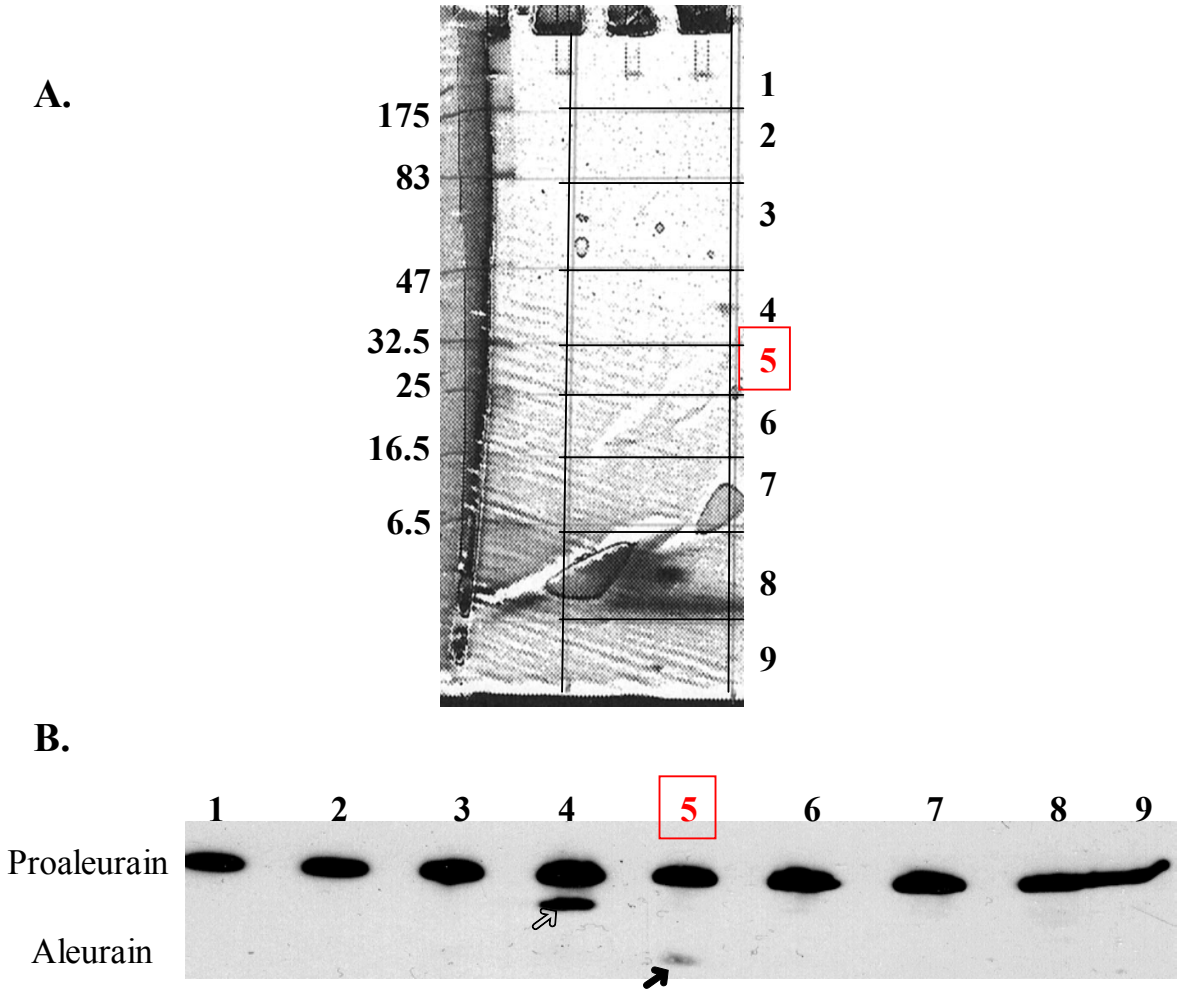


Figure 3.14. Determination of the cauliflower maturation enzyme molecular weight.

(A) Cauliflower maturation enzyme eluted from the bacitracin column was separated by SDS-PAGE. The gel was cut into sections as indicated by the boxes, and each slice was given a number for identification as indicated to the right. Numbers at left indicate the positions of molecular mass markers in kilodaltons. (B) Proteins were extracted from the gel slices optained in (A) and were tested for maturation activity. The presence of the proper aleurain maturation product produced by the slice 5 extract is indicated by a solid arrow. In contrast, a larger maturation product produced by the slice 4 extract is indicated with an open arrow. Numbers on the top correspond to gel slices (A). Names at left indicate the positions of the corresponding proteins.

maturation enzyme. Notice that the larger band representing a processing product that would appear between proaleurain and aleurain was not present in lane 5, but only in lane 4 of Figure 3.14. We concluded that the production of that middle band was due to the activity of a protease that was different from the protease activity converting proaleurain to aleurain.

3.3 Sequencing of the potential maturation enzyme

We separated the fraction eluted from the bacitracin column by SDS-PAGE and stained the gel with colloidal Coomassie blue (Figure 3.10B). Because the maturation enzyme activity was found between the molecular markers 25 and 32.5 kDa (Figure 3.14) we cut out the bands from the stained gel (as indicated, Figure 3.15) and performed an in-gel tryptic, digest on them. Each digested band was separated by capillary C-18 nano-HPLC system prior to MS/MS analysis, thereby enabling MS/MS spectra from individual peptides to be obtained as they elute. The data obtained were analyzed using the program Mascot, which matches MS/MS spectra from each band with an *in silico* (i.e. theoretical) digestion and MS/MS fragmentation of the published *Arabidopsis* data base (Table 3.2). In addition to the algorithm search I manually sequenced the MS/MS spectra of several significant bands to obtain a more precise identification of each gel band. These specific sequences were searched against the NCBI viridiplantae data base (http://www.ncbi.nlm.nih.gov) for nearly exact peptide matches (Table 3.3). As is typical for tandem MS sequencing, the isobaric amino acids, leucine and isoleucine, could not be distinguished, and insufficient resolution existed in the tandem MS mode for the resolution of the nominally isobaric pair: lysine and glutamic acid.

We concluded that the only protease identified from this approach that was likely to have maturation activity was a specific member of the cysteine protease family which has a C-terminal granulin domain extension. All the cysteine proteases that are present in Tables 3.2 and

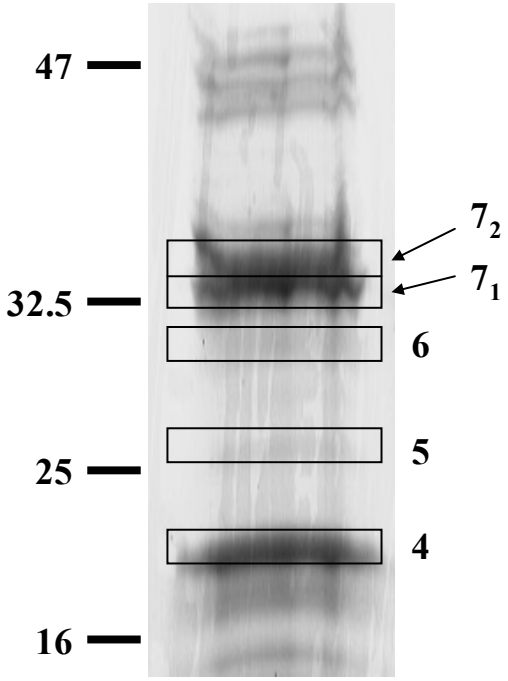


Figure 3.15. MS/MS analysis on bacitracin eluate separated by SDS-PAGE

Part of the gel image from Figure 3.10 B is presented. The boxes represent the gel slices used to perform tryptic digest that were then analysed on a liquid chromatography quadrupole time of flight spectroscope (LC Q-TOF).

Table 3.2. MASCOT search results.

MASCOT search results of *Arabidopsis* proteins with high probability of identification for the given SDS-Page gel bands as identified by comparison of the *in silico* tryptic peptide masses with tandem MS spectra. Ions score is -10*Log(P), where P is the probability that the observed match is a random event. Protein scores greater than 59 are significant ($p<0.05$). Column 1, numbers refer to the original gel band (Figure 3.16) from which selected tandem MS spectra were analyzed. Column 2, accession number for the matched protein. Column 3, mass in Daltons of the identified protein. Column 4, score value for the identified peptide. Column 5, description of the protein which sequence contained the sequence of the matched tryptic peptide.

1	2	3	4	5
4	**gi|15237348**	18094	366	expressed protein
	gi|21537415	18163	366	S-adenosylmethionine:2-demethylmenaquinone methyltransferase-like protein
	gi|15232963	18110	277	expressed protein
	gi|25289792	47658	189	SERINE CARBOXYPEPTIDASE II-like protein
	gi|18417667	53536	189	SERINE CARBOXYPEPTIDASE II - like protein
	gi|15241047	204836	79	unknown protein
	gi|22330143	281237	71	unknown protein
	gi|25405349	284902	71	hypothetical protein F2J10.9
	gi|20197603	391406	66	putative SNF2 subfamily transcription regulator
	gi|22325952	386673	66	putative SNF2 subfamily transcription regulator
5	**gi|15232963**	18110	70	expressed protein
	gi|22328529	150408	59	similar to putative retroelement pol polyprotein
6	**gi|14600257**	48111	260	papain-like cysteine peptidase XBCP3
	gi|18391078	48042	251	cysteine protease XBCP3
	gi|25289988	45655	251	cysteine proteinase
	gi|18402225	49276	179	cysteine proteinase
	gi|22325401	178658	125	similar to putative retroelement pol polyprotein
	gi|18422289	51171	118	cysteine proteinase RD21A
	gi|12382276	132488	114	SMC2-1
	gi|16314	25705	113	histone H1-1
	gi|15241831	132518	111	chromosomal protein - like
	gi|15230484	137847	107	putative protein
7_1	**gi|15227813**	166390	84	putative retroelement pol polyprotein
	gi|15237319	142601	77	putative protein
	gi|18402225	49276	76	cysteine proteinase
	gi|15221524	127126	70	putative nuclear matrix constituent protein 1(NMCP1)
	gi|15225604	123294	69	putative retroelement pol polyprotein
	gi|8778583	362270	69	F28C11.19
	gi|15234644	219944	68	hypothetical protein
	gi|15222859	276942	65	unknown protein
	gi|2443880	53944	65	Hypothetical protein
	gi|8569097	282735	65	(AC015445) Contains strong similarity to FRAP2

Table 3.3. Manual sequencing of selected tandem MS spectra.

Cauliflower sequences marked in red deviated from the plant sequence identified through homology. Column 1, numbers refer to the original gel band (Figure 3.16) from which selected tandem MS spectra were analyzed. Column 2, peptide sequence determined by manual analysis of tandem MS spectra sequenced manually. Column 3, accession number of the protein which contained the peptide sequence of column 2. Column 4, Description of matched protein. Column 5, organism origin of the matched protein.

1	2	3	4	5
6	NSGGGLLPVPVK	**gi|387907**	Künitz type protease inhibitor	*Brassica Rapa*
	ASSAGK	**gi|18141285**	Senescence associated cysteine protease	*Brassica Rapa*
	NIAEPTGK	**gi|18141283**	Senescence associated cysteine protease	*Brassica Rapa*
	NSWGGSWWSGYIK	**gi|18141283**	Senescence associated cysteine protease	*Brassica Rapa*
	ALANQPISVAIEAGGR	**gi|17065297**	Cysteine protease	*Arabidopsis thaliana*
	AFQLYSK	**gi|2160175**	Strong similarity to Dianthus cysteine proteinase	*Arabidopsis thaliana*
	VVT**MP**SYAGVK	**gi|2160175**	Strong similarity to Dianthus cysteine proteinase	*Arabidopsis thaliana*
	VVT**MP**SYAGVK	**gi|18391078**	Cysteine protease XBCP3	*Arabidopsis thaliana*
	AVSLPR	**gi|14334663**	Expressed protein	*Arabidopsis thaliana*
	CPPR	**gi|26451199**	Unknown protein	*Arabidopsis thaliana*
	GVVDYELNPK	**gi|15235335**	Putative protein	*Arabidopsis thaliana*
	SPATLNSP	**gi|136406**	Trypsin	*Canis Familiaris*
	TVQPVQEGK		NO MATCH	
	DSSSATDDSR		NO MATCH	
	DPLVAMWR		NO MATCH	
	HNQASWNTLNLK		NO MATCH	
	VA?SLPR		NO MATCH	
	RTVPNVC		NO MATCH	
	NDLAK		NO MATCH	
	NLAEPTGK		NO MATCH	
	GVHWCK		NO MATCH	
	HNQASLPIK		NO MATCH	
7_1	YPEYK	**gi|18411376**	Vacuolar protein sorting receptor homolog	*Arabidopsis thaliana*
	LSDQTYTLK	**gi|15222375**	Hypothetical protein	*Arabidopsis thaliana*
	ISDQTY**TLK**	**gi|18417667**	Serine carboxypeptidase II	*Arabidopsis thaliana*
	TA**Q**DSL**T**MLLK	**gi|18417667**	Serine carboxypeptidase II	*Arabidopsis thaliana*
	LSSPATL**D**SR	**gi| 136406**	trypsin	*Canis*

				Familiaris
	TTLNLNQQVR	**gi|1769968**	Myrosine-associated protein	*Brassica napus*
	NCLAVEPSYPLK	**gi|6682829**	Cysteine protease	*Zea mays*
	NCLAVEPSYPLK	**gi|129231**	Oryzain alpha chain precursor	*Oryza sativa*
	NCLAVEPSYPLK	**gi|7435806**	Cysteine proteinase mir3	*Zea mays*
	ALA**D**QPISVAIEA**N**R	**gi|18402225**	Cysteine proteinase	*Arabidopsis thaliana*
	EQMLLW	**gi|18395013**	Expressed protein	*Arabidopsis thaliana*
	HYSRFT**TQDK**	**gi|15220936**	Hypothetical protein	*Arabidopsis thaliana*
	YQVTTNE**K**	**gi|38424039**	Hypothetical protein	*Oryza sativa*
	VAT**V**SLPR	**gi|20160574**	Putative laccase	*Oryza sativa*
	HNQASWNTLNLK		NO MATCH	
	TAQDSLTFPK		NO MATCH	
7_2	TA**Q**DSL**T**FLLK	**gi|15293048**	Serine carboxypeptidase II	*Arabidopsis thaliana*
	PEYK	**gi|15293048**	Serine carboxypeptidase II	*Arabidopsis thaliana*
	WVVR	**gi|196209**	Glucuronosyl transferase like protein	*Arabidopsis thaliana*

3.3 have a C-terminal granulin domain extension. The tryptic peptides matched only the cysteine protease domain sequence; no peptides were identified that matched either the propeptide sequence or the granulin domain sequence.

The sequence NSGGGLLPVPVK, determined for a tryptic peptide from band 6 (Table 3.3), exactly matched a sequence from a protein identified as a Künitz-type protease inhibitor. This observation raised the question of wether such an inhibitor could play a role in inhibiting a maturation protease. In this regard it was interesting that the Künitz inhibitor propeptide sequence contains an NPLR motif that is known to function as a targeting sequence for proteins sorted to the lytic vacuole (Matsumoto and Neuhaus, 1999). The NPLR motif is similar to the NPIR motif found in proaleurain (see Introduction), and would predict that the Künitz inhibitor would also traffic to the destination where proaleurain processing occurred. Therefore we investigated the possible role of granulin domain-containing cysteine proteases in proaleurain maturation, as well as the possibility that this Künitz-type protease inhibitor could inhibit their maturation activity.

3.4 Cysteine protease with granulin domain

To test the possibility that the maturation enzyme was a cysteine protease with a granulin domain extension, we selected an *Arabidopsis* granulin domain protease (gi|18402225) whose sequence was most similar to the cauliflower peptides. The protease was expressed in *E. coli* under conditions which would secrete the protein to the periplasmic space. In the oxidizing environment of the periplasmic space, proteins undergo formation of disulfide bonds (Malamy and Horecker, 1964; Neu and Heppel, 1965) to yield a stable active protease. Additionally, since none of our peptide sequences matched those expected for granulin domains, and because the

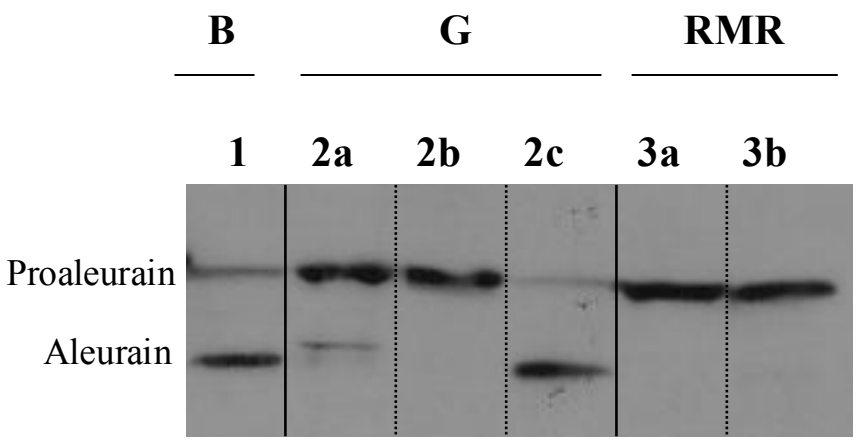


Figure 3.16. Maturation activity of cysteine protease with granulin domain

B, sample eluted from bacitracin column . G, *E. coli* expressed cysteine protease with granulin domain. RMR, *E. coli* expressed recombinant protein to serve as control. Lane 1, B incubated with proaleurain; lane 2a, G incubated with proaleurain; lane 2b, G diluted ten fold and incubated with proaleurain; lane 2c, SDS treated G incubated with proaleurain; lane 3a, RMR ten fold diluted incubated with proaleurain; lane 3b, SDS treated RMR incubated with proaleurain. All incubation were done at room temperarture for 2 hours. The maturation activity was visualised by immunoblot using anti-aleurain antibodies. Names at left indicate the position of the corresponding protein.

cysteine residue-rich granulin domain might cause problems with folding and disulfide bond formation in the *E. coli* system, we chose to express only the portion of the protein sequence encoding the enzyme propeptide plus mature enzyme with a C-terminal His-tag to facilitate purification. The protease was then partially purified by Ni^{2+}-agarose chromatography. When assayed the protease showed the same activity towards proaleurain as did the cauliflower maturation enzyme eluted from the bacitracin column (Figure 3.16, lanes 1, and 2c). The sizes of the aleurain products generated by the two enzymes were indistinguishable by SDS-PAGE (Figure 3.16, compare lanes 1, 2c). Both the enzyme from the bacitracin column and the *E. coli* expressed protease had to be SDS treated before maturation activity towards proaleurain could be observed (compare Figure 3.16, lanes 2a and 2b with 2c). This observation would argue that SDS activation of the maturation protease probably involves displacement of the propeptide from the enzyme active site.

The activity observed was not due to contaminating *E. coli* proteases because no activity was observed when a control RMR protein expressed in the *E. coli* periplasmic space and purified under the same conditions was assayed (Figure 3.16, lanes 3a and 3b).

3.5 Künitz-type protease inhibitor

To test the possibility that the maturation enzyme was inhibited by the specific Künitz inhibitor annotated as GenBank # 387907, we cloned the gene for the protein cauliflower genomic DNA; the homologue from *Arabidopsis* was similarly cloned. Neither gene contains introns. The cloned sequences were very similar (Figure 3.17). The cauliflower clone was nearly identical to *Brassica oleracea* gi|27530934 sequence (data not shown) and very similar to *Brassica rapa* gi|387907 (Table 3.3 and Figure 3.17).

The cloned *Arabidopsis* Künitz inhibitor was expressed in *E.coli* without its signal peptide. We first used the QIAexpress system (Qiagen GmbH, Germany). The Künitz inhibitor coding sequence was cloned into a pQE-30 vector and expressed in *E. coli* host strain M15. The protein was extracted under denaturing conditions and partially purified on a Ni^{2+}-agarose chromatography. When the denaturing agent (guanidine hydrochloride) was removed by dialysis, the protein precipitated out of solution. We then added a detergent before dialysis that would assist in the protein refolding so that the protein would not precipitate out of solution. The inhibitor effect of that Künitz inhibitor was tested on the cauliflower maturation enzyme eluted from the bacitracin column (Figure 3.18). At a concentration of 100 μg/ml (Figure 3.18, lane 100) the maturation of proaleurain to aleurain was considerably lower compared to the sample that didn't have any inhibitor present. This effect appeared to be specific because the presence of two control proteins, recombinant phytate embedded protein expressed in *E.coli* and purified in the same manner, and commercial soybean Trypsin inhibitor, did not appreciably affect proaleurain processing to mature aleurain.

Because we couldn't evaluate the amount of inhibitor that had been refolded properly from the refolded pool, we decided to use a second system that would express the protein in the periplasmic space of *E.coli*. The *Arabidopsis* Künitz inhibitor was expressed in *E. coli* under conditions which would secrete the protein to the periplasmic space in a manner similar to that used for expression of the granulin domain cysteine protease (see section 3.4). The inhibitor was then partially purified by Ni^{2+}-agarose chromatography. The inhibitory effect on the *E.Coli* expressed cysteine protease (section 3.4) with a granulin domain was tested (Figure 3.18). The inhibitor slowed down the maturation of proaleurain to aleurain at concentration as low as 10 μg/ml (Figure 3.18, lane Künitz). The activity observed was not due to contaminating *E. coli*

protein because no inhibitory effect was observed when a control protein expressed in the *E. coli* periplasmic space and purified under the same conditions was assayed (Figure 3.19, RMR).

We concluded that the recombinant Künitz inhibitor could inhibit the recombinant maturation enzyme activity.

Figure 3.17. Alignment of Künitz inhibitor sequences from various organisms

Sequence 1, *Brassica rapa* sequence gi|387907; Sequence 2, Cauliflower cloned sequence; Sequence 3, *Arabidopsis* sequence gi|27530934. The dots represent identity and the lines represent amino acid similarity at the indicated positions in all three sequences.

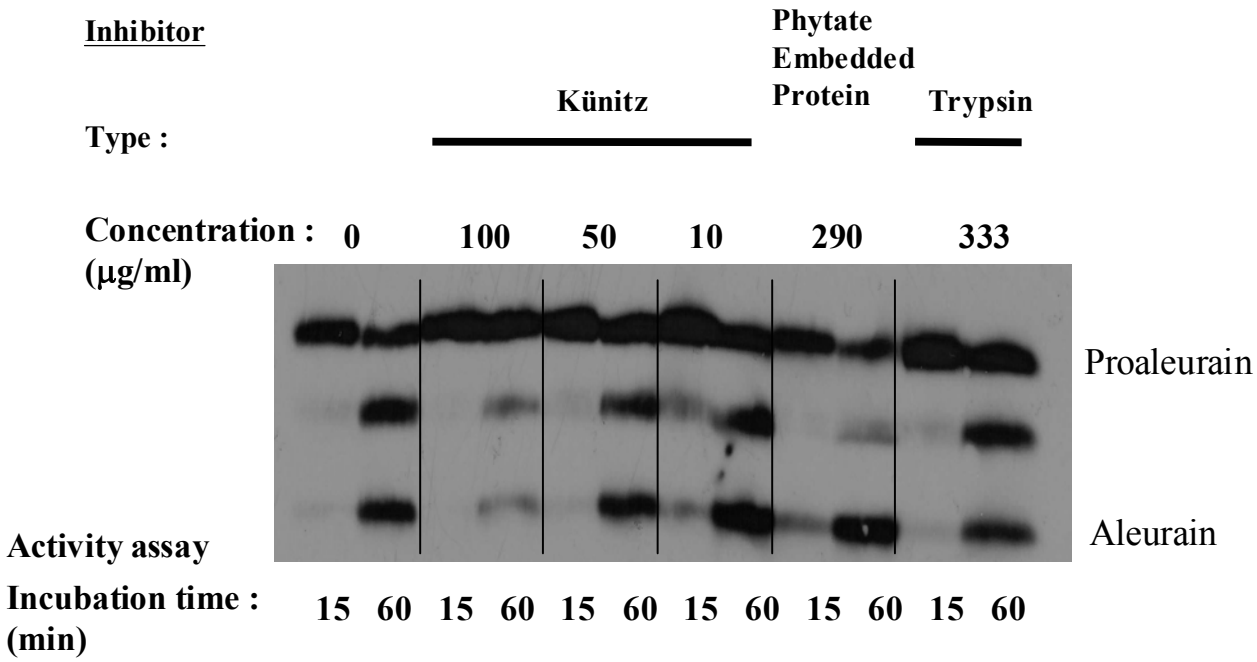


Figure 3.18. Effect of recombinant Künitz inhibitor on maturation activity.

Cauliflower maturation enzyme eluted from the bacitracin column was SDS activated and then incubated for 1 hour at room temperature with 0 to 100 mg/ml of the Künitz inhibitor expressed in *E. coli* and partially purified by Ni^{2+}-agarose chromatography. This mixture was then incubated with *Drosophila* expressed proaleurain for 15 or 60 min at room temperature. As controls, commercial soybean trypsin inhibitor and a "phytate-embeded protein" (J. Gillespie and J.C. Rogers, unpublished data) expressed in *E. coli* and purified by Ni^{2+}-agarose chromatography under the same conditions as the Künitz inhibitor inhibitor were incubated and assayed in the same manner. The assay samples were fractionated by SDS-PAGE and visualized by immunoblot using anti-aleurain antibody. Names at right indicate the positions of the corresponding proteins.

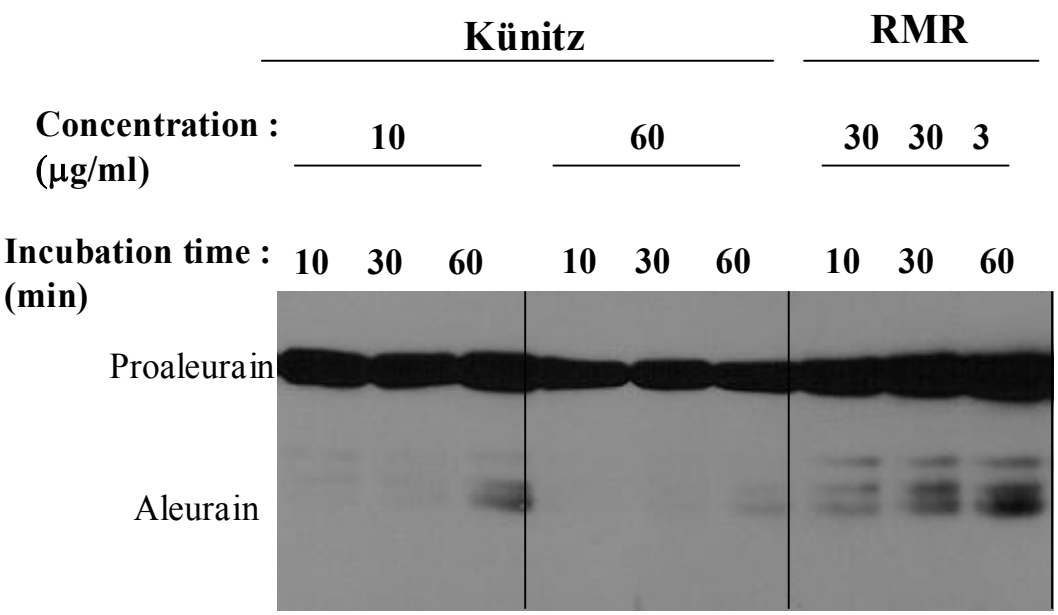


Figure 3.19. Effect of the Künitz inhitor on the maturation activity of the recombinant granulin domain cysteine protease.

The granulin domain cysteine protease expressed in *E. coli* and partially purified by Ni^{2+}-agarose chromatography SDS treated and then incubated with recombinant Künitz inhibitor at the indicated concentrations or with a control recombinant protein (RMR). These samples were then incubated with proaleurain between 10 and 60 min then maturation activity was visualized by immunoblot using anti-aleurain antibodies. Names at left indicate the position of the corresponding protein.

CHAPTER 4 DISCUSSION

Our goal in this study was to purify the maturation enzyme responsible for the processing of barley proaleurain to mature aleurain. This goal was of interest for several reasons. First, the observation that proaleurain maturation was not autocatalytic but required the activity of another protease (Holwerda et al., 1990) indicated that this maturation protease could potentially regulate aleurain activity through its activation. The role aleurain plays in plant cells is not known (beyond its enzymatic activity), but the fact that probably all plants have an aleurain indicates its function may be essential for some process important in growth or development. Thus the possibility of regulating aleurain activation would be of some intrest. Second, proaleurain maturation has been used as a biochemical marker for a compartment close to, or perhaps equivalent with the lytic vacuole (Holwerda et al., 1992; Jiang and Rogers, 1998a). Little is known of the biochemical composition of prevacuolar compartments and of specialized types of vacuoles. Most data in the literature (Wink, 1993)was compiled from studies where it was assumed that only one type of vacuole existed, and thus the composition of specific functional types of vacuoles cannot be inferred. Thus, identification of a proaleurain maturation enzyme would provide a new a marker which could enable a more precise definition of the organelles involved in protein trafficking from the Golgi to the lytic vacuole.

We started our purification attempts by using extracts from barley aleurone layers, the material in which aleurain was first identified. However, the tedious preparation process involved in obtaining the starting material from germinating barley grains prompted us to search for a more amenable source for the maturation enzyme, and ultimately led to the use of cauliflower florettes. An additional advantage for the use of cauliflower is that it shares the same

family with *Arabidopsis*, which has a completely sequenced genome. This second advantage was of great assistance in identification of the maturation enzyme from protein sequence data. Additionally cauliflower florettes have a low background of endogenous aleurain, which improves, but does not alleviate the problem contributed, by endogenous aleurain in the assay.

In this study, we partially purified a protease, from cauliflower florettes extracts, demonstrating proaleurain to aleurain maturation activity. In the purification process we learned that the maturation enzyme activity was more stable when the non-ionic detergent, NP-40, was added to the sample. We do not know the reason for its beneficial effects, but speculate that the detergent either prevented protein aggregation or prevented adsorption to the plastic tubes used in the purification process. Interestingly, we demonstrated that the activity of the maturation enzyme could be increased greatly by treating the most purified preparation with SDS. SDS activation has been previously reported, for example a prometalloprotease (Duncan et al., 1998) where it was proposed that the detergent-mediated the release of a protease inhibitor (Morodomi et al., 1992). Additionally, SDS treatment has been shown to activate a maize cysteine protease with C-terminal granulin domain extension (Yamada et al., 1998). This maize protease termed CPPIC for cysteine protease component of maize was originally isolated as a component of a trimeric complex comprised of the protease together with two subunits of the cysteine protease inhibitor, cystatin (Yamada et al., 1998; Yamada et al., 2000). Protease activity was assayed using the synthetic peptide substrate Boc-Val-Leu-Lys-MCA and optimal activity required pH 5 and pre-treatment with 0.1% SDS. Subsequent studies separated the protease from the cystatin inhibitor by gel filtration chromatography in the presence 3 M guanidine-HCl (Yamada et al., 2001b). The purified protease demonstrated the same SDS-activation profile as that shown by the protease-inhibitor complex. The authors concluded that SDS treatment induced a

conformational change necessary for protease activity. It is important that the purified protease had a molecular mass assessed on SDS-PAGE of ~ 40 kDa. This size would correspond to that expected for a molecule lacking the propeptide but comprised of the active protease domain plus the granulin domain CTE (Yamada et al., 2001b). Thus it is possible that the presence of the granulin domain was inhibitory, and that SDS treatment caused displacement of the granulin domain to allow activity.

It was difficult to evaluate the progress of the purification process due to the presence of endogenous aleurain, in the early steps, as an assay contaminant. Additionally, non-specific proteolytic degradation was in competition with the maturation enzyme for the proaleurain substrate. These two complications caused difficulties in evaluating optimal maturation activity as well as in properly quantifying the specific activities of the enzyme at various purification steps. Nevertheless, the assay was sufficient to allow partial purification and identification of the maturation enzyme by using LC-Q TOF tandem MS sequencing techniques.

From the tandem MS spectra of some partially purified fractions, two families of proteases were identified. One was a serine carboxypeptidase, which likely lacked endoprotease activity (gi|25289792 and gi|18417667). However the maturation enzyme required acidic conditions for activity, and thus differed from the serine carboxypeptidase. The second family, cysteine proteases with a granulin domain extension, seemed to be the strongest candidate. We chose a representative member of this family of enzymes, the product of *Arabidopsis* gene At3g19390, for further study because sequences of tryptic peptides obtained from the cauliflower maturation enzyme preparation matched its sequence best. At3g19390 was expressed in the E. coli periplasmic space in a form comprised of propeptide plus mature protease domain (but lacking the C-terminal granulin domain) and partially purified. This recombinant enzyme, in

contrast to a control or recombinant protein expressed and purified similarly, processed proaleurain to aleurain in a manner indistinguishable from the most purified preparation from cauliflower. The recombinant enzyme was activated by SDS treatment similarly to the cauliflower enzyme. This latter finding indicates that SDS may cause displacement of the propeptide from the active side group. The fact that the cauliflower enzyme eluted from SDS-PAGE in a size fraction between 24-32 kDa indicates that it lacked both N-terminal propeptide and c-terminal C-terminal granulin domain (Yamada et al., 2001b). We would speculate that the SDS treatment associated with SDS-PAGE caused self cleavage of the cauliflower enzyme propeptide, resulting in the mature enzyme with a mass in the range of 32 kDa that was identified by elution from the gel. Relatively little is known about the function of plant granulin domain cysteine proteases. Three have been studied previously. Transcripts for Oryzains α and β were shown to be induced in germinating rice seeds (Watanabe et al., 1991). The maze CCPIP was discussed above. In Arabidopsis in the gene for granulin domain cysteine protease, RD21 is upregulated during dehydration (Koizumi et al., 1993) and senescence (Kinoshita et al., 1999; Weaver et al., 1998). The protein was studied in Arabidopsis leaves, where the authors used antibodies specific for the full length proenzyme or the granulin domain to define the structures of different sized molecules (Yamada et al., 2001a). As assessed on immunoblots, only a 38 kDa form lacking the N-terminal propeptide, but comprised of the protease plus granulin domain, or a 33 kDa form comprised of the protease domain alone were detected. Both forms were exclusively present in the vacuole fraction when studied in extracts of *Arabidopsis* leaves, 38 kDa granulin domain-containing molecule formed aggregate at pH 5.5, an aggregate formation was dependant upon the presence of the granulin domain (Yamada et al., 2001a). The authors did not measure protease activity, but speculated that the presence of the granulin domain might

inhibit the protease, and that activity could be regulated by granulin domain-mediated trapping of the protein in aggregates in vacuoles.

Our results allow us to conclude that the cauliflower granulin domain cysteine protease is a bona fide maturation protease for proaleurain, analogous to the clipping enzyme activity identified in barley enzyme extracts. The fact that Arabidopsis RD21 is a vacuolar enzyme (Yamada et al., 2001a) would suggest that granulin domain proteases could be in the proper compartment for proaleurain compartment *in vivo*. The primary limitation of our study is that we have yet to confirm the correct N-terminal sequence of the processed aleurain from the cysteine protease from the granulin domain.

The most purified cauliflower maturation enzyme preparation contained, in addition to the maturation protease, a protein predicted to be a Künitz-type protease inhibitor. The prototype for the protease inhibitor family is the soy bean trypsin inhibitor, and the proteins typically contain ~200 amino acid residues and one or two intramolecular disulfide bonds (Rawlings et al., 2004). Künitz-type inhibitors may inhibit serine, aspartic, and cysteine proteases. An example of the latter is PCPI 8.3, purified from potato (Krizaj et al., 1993).

When the sequence of the cauliflower Künitz inhibitor predicted from the PCR amplified cloned cauliflower amplified gene (sequence in Figure 3.17) was aligned to that of PCPIP 8.3, little sequence similarity was observed with two exceptions: the two conserved cysteine residues, and the sequence DGNPLR that is near the N-terminus of the proteins after cleavage of the signal peptides (data not shown). A protein closely related to the cauliflower inhibitor is induced by water stress of leaves of *Brassica napus* (Downing et al., 1992). The *Arabidopsis* equivalent of the cauliflower inhibitor (sequence in Figure 3.17) was expressed in *E. coli*, partially purified, and tested for its ability to inhibit proaleurain maturation from the cauliflower extract (Figure

3.18) and recombinant Arabidopsis granulin domain protease (Figure 3.19). Inhibition was documented in both instances. And thus it is possible that the cauliflower inhibitor co-purified with the cauliflower protease, and was released from the protease by SDS treatment. This also could have contributed to the activation of maturation enzyme activity that was observed after SDS treatment.

We have not yet purified the recombinant inhibitor to homogeneity and tested its ability to inhibit proteases of different classes; also our results to date indicate it has inhibitory activity for cysteine proteases. The Arabidopsis inhibitor was recently transiently expressed in tobacco suspension culture protoplasts and its localization was compared to that of BP80, the receptor protein that sorts ligands to the prevacuolar compartment for lytic vacuoles. The inhibitor in BP80 showed essentially complete co-localization (M.Oufattole, unpublished data). At steady state, BP80 concentrates in prevacuolar compartments (Li et al., 2002). This result indicates that the inhibitor probably traffics in the same intracellular pathway as proaleurain and that its ability to inhibit the proaleurain maturation enzyme may be functionally significant.

In future work we will determine the pathways by which both the granulin domain cysteine protease and the Künitz inhibitor reach their vacuolar destinations. It will be particularly important to learn if the protease is concentrated in prevacuolar compartments, because results from other studies have indicated that proaleurain processing occurs in those organelles (Jiang and Rogers, 1998b). Initially, it should be possible to test the potential role of the inhibitor in blocking maturation enzyme activity using transient expression assays in protoplasts. One would predict that over-expression of the inhibitor would block maturation of barley proaleurain that was co-expressed in the same cells. Thus my studies of proaleurain

maturation may lead to a more detailed understanding of the traffic of proteins to the lytic vacuole, and of the biochemical composition of the prevacuolar compartment.

REFERENCES

Andreeva, A.V., M.A. Kutuzov, D.E. Evans, and C.R. Hawes. 1998. The structure and function of the Golgi apparatus: a hundred years of questions. *Journal of experimental botany*. 49:1281-1291.

Asemota, H.N. 1995. A fast, simple and efficient miniscale method for the preparation of DNA from tissue of yam (Discorea ssp.). *Plant Mol Biol Rep*. 13:214-218.

Bateman, A., and H.P. Bennett. 1998. Granulins: the structure and function of an emerging family of growth factors. *J Endocrinol*. 158:145-51.

Baudys, M., B. Meloun, T. Gan-Erdene, M. Fusek, M. Mares, V. Kostka, J. Pohl, and C.C. Blake. 1991. S-S bridges of cathepsin B and H from bovine spleen: a basis for cathepsin B model building and possible functional implications for discrimination between exo- and endopeptidase activities among cathepsins B, H and L. *Biomed Biochim Acta*. 50:569-77.

Bednarek, S.Y., T.A. Wilkins, J.E. Dombrowski, and N.V. Raikhel. 1990. A carboxyl-terminal propeptide is necessary for proper sorting of barley lectin to vacuoles of tobacco. *Plant Cell*. 2:1145-55.

Bossard, M.J., T.A. Tomaszek, S.K. Thompson, B.Y. Amegadzie, C.R. Hanning, C. Jones, J.T. Kurdyla, D.E. McNulty, F.H. Drake, M. Gowen, and M.A. Levy. 1996. Proteolytic activity of human osteoclast cathepsin K. Expression, purification, activation, and substrate identification. *J Biol Chem*. 271:12517-24.

Braakman, I., J. Helenius, and A. Helenius. 1992. Role of ATP and disulphide bonds during protein folding in the endoplasmic reticulum. *Nature*. 356:260-2.

Brooks, D.R., L. Tetley, G.H. Coombs, and J.C. Mottram. 2000. Processing and trafficking of cysteine proteases in Leishmania mexicana. *J Cell Sci*. 113 (22):4035-41.

Cao, X., S.W. Rogers, J. Butler, L. Beevers, and J.C. Rogers. 2000. Structural requirements for ligand binding by a probable plant vacuolar sorting receptor. *The Plant cell.* 12(4):493-506.

Cioffi, J.A., K.L. Allen, M.O. Lively, and B. Kemper. 1989. Parallel effects of signal peptide hydrophobic core modifications on co-translational translocation and post-translational cleavage by purified signal peptidase. *J Biol Chem*. 264:15052-8.

Downing, W.L., F. Mauxion, M.O. Fauvarque, M.P. Reviron, D. de Vienne, N. Vartanian, J. Giraudat, and Curbs, Y.F. 1992. A Brassica napus transcript encoding a protein related to the Kèunitz protease inhibitor family accumulates upon water stress in leaves, not in seeds. *The Plant journal : for cell and molecular biology.* 2(5):685-93.

Duncan, M.E., J.P. Richardson, G.I. Murray, W.T. Melvin, J.E. Fothergill, M. Department of, and U.o.A.U.K. Cell Biology. 1998. Human matrix metalloproteinase-9: activation by limited trypsin treatment and generation of monoclonal antibodies specific for the activated form. *European journal of biochemistry / FEBS.* 258(1):37-43.

Guncar, G., M. Podobnik, J. Pungercar, B. Strukelj, V. Turk, and D. Turk. 1998. Crystal structure of porcine cathepsin H determined at 2.1 A resolution: location of the mini-chain C-terminal carboxyl group defines cathepsin H aminopeptidase function. *Structure*. 6:51-61.

Hadlington, J.L., and J. Denecke. 2000. Sorting of soluble proteins in the secretory pathway of plants. *Curr Opin Plant Biol*. 3:461-8.

Hammond, C., and A. Helenius. 1995. Quality control in the secretory pathway. *Curr Opin Cell Biol*. 7:523-9.
Hartl, F.U. 1996. Molecular chaperones in cellular protein folding. *Nature*. 381:571-9.
Hille-Rehfeld, A. 1995. Mannose 6-phosphate receptors in sorting and transport of lysosomal enzymes. *Biochim Biophys Acta*. 1241:177-94.
Hoh, B., G. Hinz, B.K. Jeong, and D.G. Robinson. 1995. Protein storage vacuoles form de novo during pea cotyledon development. *J Cell Sci*. 108 (Pt 1):299-310.
Hohl, I., D.G. Robinson, M.J. Chrispeels, and G. Hinz. 1996. Transport of storage proteins to the vacuole is mediated by vesicles without a clathrin coat. *J Cell Sci*. 109 (Pt 10):2539-50.
Holwerda, B.C., N.J. Galvin, T.J. Baranski, and J.C. Rogers. 1990. In vitro processing of aleurain, a barley vacuolar thiol protease. *The Plant cell*. 2:1091-1106.
Holwerda, B.C., H.S. Padgett, and J.C. Rogers. 1992. Proaleurain vacuolar targeting is mediated by short contiguous peptide interactions. *The Plant cell*. 4:307-318.
Holwerda, B.C., and J.C. Rogers. 1992. Purification and characterization of aleurain. A plant thiol protease functionally homologous to mammalian cathepsin H. *Plant physiology*. 99:848-855.
Holwerda, B.C., and J.C. Rogers. 1993. Structure, functional properties and vacuolar targeting of the barley thiol protease, aleurain. *Journal of experimental botany*. 44:321-329.
Humair, D., D. Hernandez Felipe, J.M. Neuhaus, and N. Paris. 2001. Demonstration in yeast of the function of BP-80, a putative plant vacuolar sorting receptor. *Plant Cell*. 13:781-92.
Hwang, C., A.J. Sinskey, and H.F. Lodish. 1992. Oxidized redox state of glutathione in the endoplasmic reticulum. *Science*. 257:1496-502.
Jiang, L., and J.C. Rogers. 1998a. Integral membrane protein sorting to vacuoles in plant cells: evidence for two pathways. *The Journal of cell biology*. 143:1183-1199.
Jiang, L., and J.C. Rogers. 1998b. Integral membrane protein sorting to vacuoles in plant cells: evidence for two pathways. *J Cell Biol*. 143:1183-99.
Johnson, K.D., H. Hofte, and M.J. Chrispeels. 1990. An intrinsic tonoplast protein of protein storage vacuoles in seeds is structurally related to a bacterial solute transporter (GIpF). *Plant Cell*. 2:525-32.
Kaiser, C.A., D. Preuss, P. Grisafi, and D. Botstein. 1987. Many random sequences functionally replace the secretion signal sequence of yeast invertase. *Science*. 235:312-7.
Kinoshita, T., K. Yamada, N. Hiraiwa, M. Kondo, M. Nishimura, I. Hara-Nishimura, and N.I.f.B.B.O.J. Department of Cell Biology. 1999. Vacuolar processing enzyme is up-regulated in the lytic vacuoles of vegetative tissues during senescence and under various stressed conditions. *The Plant journal : for cell and molecular biology.* 19(1):43-53.
Kirsch, T., N. Paris, J.M. Butler, L. Beevers, and J.C. Rogers. 1994. Purification and initial characterization of a potential plant vacuolar targeting receptor. *Proc Natl Acad Sci U S A*. 91:3403-7.
Koizumi, M., K. Yamaguchi-Shinozaki, H. Tsuji, K. Shinozaki, T.L.S.C.I.o.P. Laboratory of Plant Molecular Biology, and I.J. Chemical Research. 1993. Structure and expression of two genes that encode distinct drought-inducible cysteine proteinases in Arabidopsis thaliana. *Gene.* 129(2):175-82.
Kornfeld, S., and I. Mellman. 1989. The biogenesis of lysosomes. *Annu Rev Cell Biol*. 5:483-525.

Krizaj, I., M. Drobnic-Kosorok, J. Brzin, R. Jerala, V. Turk, B. Department of, and J.S.I.L.S. Molecular Biology. 1993. The primary structure of inhibitor of cysteine proteinases from potato. *FEBS letters.* 333(1-2):15-20.

Lecaille, F., J. Kaleta, and D. Bromme. 2002. Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. *Chem Rev*. 102:4459-88.

Lerouge, P., M. Cabanes-Macheteau, C. Rayon, A.C. Fischette-Laine, V. Gomord, and L. Faye. 1998. N-glycoprotein biosynthesis in plants: recent developments and future trends. *Plant Mol Biol*. 38:31-48.

Li, Y.B., S.W. Rogers, Y.C. Tse, S.W. Lo, S.S. Sun, G.Y. Jauh, and L. Jiang. 2002. BP-80 and homologs are concentrated on post-Golgi, probable lytic prevacuolar compartments. *Plant Cell Physiol*. 43:726-42.

Lowe, J.B., and J.D. Marth. 2003. A genetic approach to Mammalian glycan function. *Annu Rev Biochem*. 72:643-91.

Malamy, M.H., and B.L. Horecker. 1964. Release of Alkaline Phosphatase from Cells of Escherichia Coli Upon Lysozyme Spheroplast Formation. *Biochemistry*. 19:1889-93.

Marty-Mazars, D., M.C. Clemencet, P. Dozolme, and F. Marty. 1995. Antibodies to the tonoplast from the storage parenchyma cells of beetroot recognize a major intrinsic protein related to TIPs. *Eur J Cell Biol*. 66:106-18.

Matsumoto, K., and J.M. Neuhaus. 1999. Cis-elements of protein transport to the plant vacuoles. *Journal of Experimental botany*. 50:165-174.

Matsuoka, K., and K. Nakamura. 1991. Propeptide of a precursor to a plant vacuolar protein required for vacuolar targeting. *Proc Natl Acad Sci U S A*. 88:834-8.

McGrath, M.E. 1999. The lysosomal cysteine proteases. *Annu Rev Biophys Biomol Struct*. 28:181-204.

McQueney, M.S., B.Y. Amegadzie, K. D'Alessio, C.R. Hanning, M.M. McLaughlin, D. McNulty, S.A. Carr, C. Ijames, J. Kurdyla, and C.S. Jones. 1997. Autocatalytic activation of human cathepsin K. *J Biol Chem*. 272:13955-60.

Melchers, L.S., M.B. Sela-Buurlage, A.S. Vloemans, C.P. Woloshuk, J.S.C.v. Roekel, J. Pen, P.J.M.v.d. Elzen, and B.J.C. Cornelissen. 1993. Extracellular targeting of the vacuolar tobacco proteins AP24, chitinase and beta-1,3-glucanase in transgenic plants. *Plant molecular biology : an international journal on molecular biology, biochemistry and genetic engineering*. 21:583-593.

Morodomi, T., Y. Ogata, Y. Sasaguri, M. Morimatsu, H. Nagase, B. Department of, and U.o.K.M.C.K.C. Molecular Biology. 1992. Purification and characterization of matrix metalloproteinase 9 from U937 monocytic leukaemia and HT1080 fibrosarcoma cells. *The Biochemical journal.* 285:603-11.

Nakamura, K., K. Matsuoka, F. Mukumoto, and N. Watanabe. 1993. Processing and transport to the vacuole of a precursor to sweet potato sporamin in transformed tobacco cell line BY-2. *Journal of experimental botany*. 44:331-338.

Nebenfuhr, A., and L.A. Staehelin. 2001. Mobile factories: Golgi dynamics in plant cells. *Trends Plant Sci*. 6:160-7.

Neu, H.C., and L.A. Heppel. 1965. The release of enzymes from Escherichia coli by osmotic shock and during the formation of spheroplasts. *J Biol Chem*. 240:3685-92.

Neuhaus, J.M., and J.C. Rogers. 1998. Sorting of proteins to vacuoles in plant cells. *Plant Mol Biol*. 38:127-44.

Neuhaus, J.M., L. Sticher, F. Meins, Jr., and T. Boller. 1991. A short C-terminal sequence is necessary and sufficient for the targeting of chitinases to the plant vacuole. *Proceedings of the National Academy of Sciences of the United States of America*. 88:10362-10366.
Neumann, U., F. Brandizzi, and C. Hawes. 2003. Protein transport in plant cells: in and out of the Golgi. *Ann Bot (Lond)*. 92:167-80.
Nishimura, Y., and K. Kato. 1987. Intracellular transport and processing of lysosomal cathepsin H. *Biochem Biophys Res Commun*. 148:329-34.
Nishimura, Y., and K. Kato. 1988. Identification of latent procathepsin H in microsomal lumen: characterization of proteolytic processing and enzyme activation. *Arch Biochem Biophys*. 260:712-8.
Paris, N., C.M. Stanley, R.L. Jones, and J.C. Rogers. 1996. Plant cells contain two functionally distinct vacuolar compartments. *Cell*. 85:563-72.
Podobnik, M., R. Kuhelj, V. Turk, and D. Turk. 1997. Crystal structure of the wild-type human procathepsin B at 2.5 A resolution reveals the native active site of a papain-like cysteine protease zymogen. *J Mol Biol*. 271:774-88.
Rawlings, N.D., D.P. Tolle, A.J. Barrett, and W.T.G.C.H.C.C.B.S.A.U.K.n.s.a.u. The Wellcome Trust Sanger Institute. 2004. Evolutionary families of peptidase inhibitors. *The Biochemical journal.* 378(Pt) 3:705-16.
Robinson, D.G., G. Hinz, and S.E. Holstein. 1998. The molecular characterization of transport vesicles. *Plant Mol Biol*. 38:49-76.
Rogers, J.C., D. Dean, and G.R. Heck. 1985. Aleurain: a barley thiol protease closely related to mammalian cathepsin H. *Proceedings of the National Academy of Sciences of the United States of America*. 82:6512-6516.
Rogers, S.W., M. Burks, and J.C. Rogers. 1997. Monoclonal antibodies to barley aleurain and homologs from other plants. *The Plant journal : for cell and molecular biology*. 11:1359-1368.
Rothe, M., A. Zichner, E.A. Auerswald, and J. Dodt. 1994. Structure/function implications for the aminopeptidase specificity of aleurain. *Eur J Biochem*. 224:559-65.
Sambrook, J., and D.W. Russell. 2001. Molecular Cloning. 1:116-118.
Shinshi, H., H. Wenzler, J.M. Neuhaus, G. Felix, J. Hofsteenge, and F. Meins, Jr. 1988. Evidence for N- and C-terminal processing of plant defense-related enzyme: primary structure of tobacco prepro-beta-1,3-glucanase. *Proceedings of the National Academy of Sciences of the United States of America*. 85:5541-5545.
Sonnewald, U., A. von Schaewen, and L. Willmitzer. 1990. Expression of mutant patatin protein in transgenic tobacco plants: role of glycans and intracellular location. *Plant Cell*. 2:345-55.
Tartakoff, A.M. 1983. Perturbation of vesicular traffic with the carboxylic ionophore monensin. *Cell*. 32:1026-8.
Turk, V., B. Turk, G. Guncar, D. Turk, and J. Kos. 2002. Lysosomal cathepsins: structure, role in antigen processing and presentation, and cancer. *Adv Enzyme Regul*. 42:285-303.
van Noort, J.M., P. van den Berg, and I.E. Mattern. 1991. Visualization of proteases within a complex sample following their selective retention on immobilized bacitracin, a peptide antibiotic. *Anal Biochem*. 198:385-90.
Vitale, A., and J. Denecke. 1999. The endoplasmic reticulum-gateway of the secretory pathway. *Plant Cell*. 11:615-28.

Watanabe, H., K. Abe, Y. Emori, H. Hosoyama, and S. Arai. 1991. Molecular cloning and gibberellin-induced expression of multiple cysteine proteinases of rice seeds (oryzains). *J Biol Chem*. 266:16897-902.

Weaver, L.M., S. Gan, B. Quirino, R.M. Amasino, and U.o.W.a.M.U.S.A. Dept of Biochemistry. 1998. A comparison of the expression patterns of several senescence-associated genes in response to stress and hormone treatment. *Plant molecular biology*. 37(3):455-69.

Wiederanders, B. 2003. Structure-function relationships in class CA1 cysteine peptidase propeptides. *Acta Biochim Pol*. 50:691-713.

Wilkins, T.A., S.Y. Bednarek, and N.V. Raikhel. 1990. Role of propeptide glycan in post-translational processing and transport of barley lectin to vacuoles in transgenic tobacco. *Plant Cell*. 2:301-13.

Wink, M. 1993. The plant vacuole: A mutifunctional compartment. *Journal of Experimental botany*. 44:231-46.

Yamada, K., R. Matsushima, M. Nishimura, and I. Hara-Nishimura. 2001a. A slow maturation of a cysteine protease with a granulin domain in the vacuoles of senescing Arabidopsis leaves. *Plant Physiol*. 127:1626-34.

Yamada, T., A. Kondo, H. Ohta, T. Masuda, H. Shimada, and K. Takamiya. 2001b. Isolation of the protease component of maize cysteine protease-cystatin complex: release of cystatin is not crucial for the activation of the cysteine protease. *Plant Cell Physiol*. 42:710-6.

Yamada, T., H. Ohta, T. Masuda, M. Ikeda, N. Tomita, A. Ozawa, Y. Shioi, and K. Takamiya. 1998. Purification of a novel type of SDS-dependent protease in maize using a monoclonal antibody. *Plant Cell Physiol*. 39:106-14.

Yamada, T., H. Ohta, A. Shinohara, A. Iwamatsu, H. Shimada, T. Tsuchiya, T. Masuda, and K. Takamiya. 2000. A cysteine protease from maize isolated in a complex with cystatin. *Plant Cell Physiol*. 41:185-91.