

REVIEW

Mobilization of seed protein reserves

Anna L. Tan-Wilson* and Karl A. Wilson

Department of Biological Sciences, State University of New York at Binghamton, Binghamton, NY 13902, USA

Correspondence

*Corresponding author,
e-mail: annatan@binghamton.edu

Received 4 May 2011;
revised 7 July 2011

doi:10.1111/j.1399-3054.2011.01535.x

The mobilization of seed storage proteins upon seed imbibition and germination is a crucial process in the establishment of the seedling. Storage proteins fold compactly, presenting only a few vulnerable regions for initial proteolytic digestion. Evolutionarily related storage proteins have similar three-dimensional structure, and thus tend to be initially cleaved at similar sites. The initial cleavage makes possible subsequent rapid and extensive breakdown catalyzed by endo- and exopeptidases. The proteolytic enzymes that degrade the storage proteins during mobilization identified so far are mostly cysteine proteases, but also include serine, aspartic and metalloproteases. Plants often ensure early initiation of storage protein mobilization by depositing active proteases during seed maturation, in the very compartments where storage proteins are sequestered. Various means are used in such cases to prevent proteolytic attack until after imbibition of the seed with water. This constraint, however, is not always enforced as the dry seeds of some plant species contain proteolytic intermediates as a result of limited proteolysis of some storage proteins. Besides addressing fundamental questions in plant protein metabolism, studies of the mobilization of storage proteins will point out proteolytic events to avoid in large-scale production of cloned products in seeds. Conversely, proteolytic enzymes may be applied toward reduction of food allergens, many of which are seed storage proteins.

Introduction

Imbibition of the seed with water marks a major change in protein metabolism. Proteins that were synthesized and stored in abundance in the seed are broken down to free amino acids for biosynthesis and energy generation. In this review of the pathways and regulation of seed storage protein (SSP) mobilization, we will show that the turnaround in protein metabolism between seed maturation and post-germination is not as sharp as presumed. Some plants continue processes associated with storage protein precursor processing in seedling growth. Others begin processes usually associated with post-germination as early as seed maturation. These two situations bracket a continuum of strategies that

plants use to balance the need to store reserves in the seed and the need to mobilize those protein stores soon after water imbibition. In addition to providing a different perspective into a fundamental question in plant metabolic regulation, this review is relevant to ‘molecular farming’ and food allergies. Seeds are well suited for industrial-scale production of peptides and proteins because they accumulate a small number of proteins in abundance, and are easy to transport and to store. Seeds, however, also express high levels of active proteolytic enzymes. Failure to anticipate their action often results in hydrolysis of cloned products (Boothe et al. 2010, Cunha et al. 2010). Thus, information with respect to seed proteases and conditions necessary for catalysis will be increasingly important. As more food

Abbreviations – CPPh1, cysteine protease 1 of *Phaseolus vulgaris*; DAI, days after imbibition; GA, gibberellic acid; PSV, protein storage vacuole; SSP, seed storage protein; VPE, vacuolar processing enzyme.

allergies are being traced to SSPs, plant biologists and immunologists have come to realize that the very same characteristics that make for good storage of proteins in the seed are also the ones that predispose a protein to be immunogenic. Attempts to reduce the synthesis of specific storage proteins often result in plants shifting toward making other proteins that are also allergenic (Kawakatsu and Takaiwa 2010). The use of endogenously or exogenously applied proteolytic enzymes to reduce the content of allergenic proteins in foods is a logical strategy that has not been well explored. Understanding the structure and function of proteolytic enzymes that plants evolved expressly to mobilize the storage proteins is a good starting point for such endeavors.

Seed storage proteins

Familiarity with the structures of the major groups of storage proteins is fundamental to understanding their proteolysis; but the nomenclature sometimes makes this difficult. Many storage protein names relate back to Osborne (1924) who classified seed proteins as albumins, globulins, prolamins and glutelins by their successive extractability in water, dilute salt solution, alcohol/water mixture and dilute acid/alkali, respectively. As amino acid sequence and structural information have become available, it is clear that solubility does not always reflect evolutionary relationships. For instance, rice (*Oryza sativa*) glutelins are related in sequence to the group of 11S globulins, while 2S albumins are considered to be in the prolamins superfamily. Because structure, not solubility, is more relevant to how storage proteins are processed for packing and how they are later degraded, this review is organized according to a number of major SSP gene families that have been studied in the context of storage protein mobilization. Table 1 lists a few salient structural properties of these gene families.

The storage protein families described in Table 1 have major structural features in common. They are folded compactly and often form multimeric structures. (1) The 11S and 7S globulins: The monomeric units of these evolutionarily related storage proteins are folded into two structurally equivalent, compactly folded, barrel-shaped cupin domains (Dunwell et al. 2001) joined by a short linker (Maruyama et al. 2001, Adachi et al. 2003), and assembled as hexamers and trimers, respectively. The 11S globulins (also known as legumins) are synthesized as a single chain that is then processed at one point on the C-terminal side of an Asn residue by asparaginyl endopeptidases [also called legumains or vacuolar processing enzymes (VPEs)] to acidic and basic chains.

Correct processing is followed by assembly into the hexameric 11S globulins, after which the storage protein is resistant to further proteolysis by the processing enzyme. When the proglobulin is misfolded, and the processing enzyme cleaves at an incorrect Asn residue, the legumin does not assemble properly and the VPE continues to digest the storage protein. Accumulation of misfolded SSPs, which might interfere with the proper packing into the protein storage vacuole (PSV), is thus prevented (Jung et al. 1998). Mutation of all the VPEs results in alternative processing by other seed proteases (Gruis et al. 2004). The 7S globulins are typically homo- or heterotrimers of evolutionarily related vicilins and the convicilins, the latter being larger because of an insertion in the vicilin N-terminal domain (Bown et al. 1988). (2) The 2S albumins: These proteins are also synthesized as a single chain. In many species, the protein is processed into two chains still linked by a disulfide bond. In *Arabidopsis* (*Arabidopsis thaliana*) processing of the napin precursor is catalyzed by an aspartic protease (D'Hondt et al. 1993). These proteins present a tight arrangement of four to five α -helices in a right-handed super-helix (Pantoja-Uceda et al. 2004, Lehmann et al. 2006). (3) The prolamins: Spectroscopic studies of the proline- and glutamine-rich sequences of these storage proteins indicate folding into compact rod-shaped molecules in which the residues form β -reverse turns organized into a spiral structure (reviewed in Shewry and Halford 2002). Their packaging into protein bodies involves extensive hydrogen bonding in addition to intermolecular disulfide bonds, accounting for the insolubility of these storage proteins in aqueous solutions. (4) The Bowman-Birk trypsin inhibitors: In dicots, seven disulfide bridges stabilize two protease inhibitor domains, each domain consisting of three β -strands (reviewed in Qi et al. 2005). Some inhibitors further dimerize, this being attributed to electrostatic attraction between an Arg/Lys residue at a trypsin inhibitor reactive site and Asp/Glu residues at the C-terminus of the other monomer (Kumar et al. 2004).

Looking at the group of storage proteins as a whole, compactness seems to be a desirable attribute for a SSP. Large amounts of amino acids are stored in a small volume and the storage protein is not easily digested during seed maturation (Shutov et al. 2003). While this feature appears counterproductive to allowing the plant to make use of the stored reserves during seedling growth, the compactness can also be seen as a means of pacing the use of those reserves.

SSPs, of course, also serve as major sources of amino acids in our food supply. It is precisely the compact nature of these abundant SSPs that makes these proteins more resistant to proteolysis by proteases

Table 1. The major storage protein families considered in this review. BBSTI, Bowman–Birk soybean trypsin inhibitor; HMW-GS, high-molecular-weight glutenin subunit; LMW-GS, low-molecular-weight glutenin subunit; MBTI, mung bean trypsin inhibitor; PDB, protein data bank.

Group	Description	Common names, PDB or GenBank accession
11S globulin (legumins)	Heterohexamers of single-chain precursors processed into disulfide-bonded acidic (approximately 30–40 kDa) and basic (approximately 17–20 kDa) chains (Jung et al. 1998)	Glycinin (<i>Glycine max</i>) PDB:1FXZ, Ara h 3 (<i>Arachis hypogaea</i>) PDB: 3C3V, Cruciferins (<i>Arabidopsis thaliana</i>) CAB81440, 13S globulin (<i>Fagopyrum esculentum</i>) Q9XFM4, Rice glutelins (<i>Oryza sativa</i>) CAA38110
7S globulin (vicilins)	Homo-, heterotrimers of two types of homologous subunits: Vicilin (approximately 45–50 kDa), Convicilin (approximately 60–70 kDa) has additional hydrophilic, mostly acidic residues in N-terminal region (Bown et al. 1988)	Vicilin-type subunits: Phaseolin (<i>Phaseolus vulgaris</i>) PDB:2PHL, β -conglycinin β subunit (<i>G. max</i>) PDB:1UIJ, Vicilin (<i>Pisum sativum</i>) CAA32239 Convicilin-type subunits: α and α' subunits of β -conglycinin PDB:1UIK, Convicilin (<i>P. sativum</i>) CAB82855, Ara H 1 (<i>A. hypogaea</i>) P43238, Beta-conglutin (<i>Lupinus albus</i>) AAS97865
2S albumins	Approximately 18–21 kDa, characteristic sequence with eight cysteine residues forming four intramolecular disulfide bonds; some processed to form two chains (e.g. napin, 4 and 9 kDa) (D'Hondt et al. 1993)	Single-chain type: Ara h 2 & Ara h 6 (<i>A. hypogaea</i>) PDB:1W2Q Two-chain type: Napin (<i>Brassica napus</i>) PDB:1PNB, Conglutin δ_2 (<i>L. albus</i>) CAJ42100
Sulfur-rich prolamins	Repetitive Pro- and Gln-rich domain followed by Cys-rich domain homologous to 2S albumins; Cys residues in intramolecular disulfide bonds. LMW-glutenin subunits have additional Cys residues in intermolecular S–S bonds (D'Ovidio and Masci 2004)	α/β -gliadin (<i>Triticum</i> sp.) A27319, γ -gliadin (<i>Triticum</i> sp.) CAB75404, γ -hordein (<i>Hordeum vulgare</i>) P17990, LMW-GS (<i>Triticum</i> sp.) ACK44494, B-hordein (<i>H. vulgare</i>) ABB82614
Sulfur-poor prolamins	Repetitive Pro- and Gln-rich domain, but no Cys-rich domain (Qi et al. 2006)	ω -gliadin (<i>Triticum</i> sp.) AAG17702 C-hordein (<i>H. vulgare</i>) AAA92333
High-molecular-weight prolamins	Highly repetitive central domain of Pro-, Gly- and Glu-rich tri- hexa- and nona-peptide repeats, non-repetitive N- and C-terminal domains with the disulfide-bonded cysteine residues (Anjum et al. 2007)	HMW-GS (<i>Triticum</i> sp.) AAS67322 D-hordein (<i>H. vulgare</i>) BAA11642
Bowman–Birk inhibitor	Dicots: two reactive sites, 8 kDa; monocots: one or two reactive sites (8 or 16 kDa). Highly conserved disulfide bonds (Qi et al. 2005)	BBSTI (<i>G. max</i>) PDB 1K9B, MBTI (<i>Vigna radiata</i>) PDB 3MYW, HGI-III inhibitor (<i>Dolichos biflorus</i>) AAB47193

of the mammalian gut, leading to food allergies when undigested proteins cross the intestinal mucosal barrier (Mills et al. 2004). Allergenicity stems from the ability of the protein to induce the production of immunoglobulin E (IgE) antibodies that are displayed on mast cells. When the protein is resistant to proteolysis, more of the protein fragments contain multiple epitopes that bind IgE and cross-link mast cell receptors, leading to degranulation and the symptoms of the allergic reaction (Matsuda et al. 2006). This was demonstrated in the epitope mapping and proteolysis of Ara h 1 (the convicilin of peanut also known as conarachin) by mammalian gastrointestinal proteases. Under conditions that reduce other proteins to peptides <10 kDa in

less than an hour, 3 h of digestion of Ara h 1 still left peptides with multiple IgE epitopes (Maleki et al. 2000).

Storage protein compartments

In dicots, storage proteins are sequestered in membrane-bound PSVs that are especially prominent in the parenchyma cells of the cotyledons as well as in the embryonic axis (Jiang et al. 2001). Targeting of storage proteins and proteases to these compartments involves multiple pathways originating from the endoplasmic reticulum (ER) or the Golgi with targeting based on aggregation and/or the presence of specific sorting

determinants (Vitale and Hinz 2005). In monocots, storage proteins are found in the dead starchy endosperm, cells that undergo programmed cell death after reserve deposition. Where the proteins are located within this space depends on the major type of storage protein accumulated by that plant species (reviewed in Shewry and Halford 2002). Within the endosperm in maize (*Zea mays*), prolamins are stored in membrane-bound protein bodies originating from the ER, with the proportions of the different prolamins changing with the size of protein body and distance from the aleurone. In wheat (*Triticum* sp.) and barley (*Hordeum vulgare*), the prolamins accumulate in vacuoles. Oats (*Avena sativa*) and rice tend to store more 11S globulins than prolamins, and sequester these in biphasic protein bodies or in vacuoles. Storage proteins are also stored in the aleurone, live cells that, after germination, secrete hydrolases into the adjacent starchy endosperm to break down the reserves stored

there. Presumably the storage proteins in the aleurone provide the resources for de novo synthesis of those enzymes. The type of protein stored in the aleurone differs from those in the endosperm. In maize, for instance, the aleurone has less of the prolamins and more of the globulin type of storage proteins (Reyes et al. 2011).

The proteolytic enzymes

To be considered a protease that plays a role in mobilization of SSPs, a number of criteria are usually sought. These include demonstration in vitro of the ability of the purified enzyme to catalyze a step in storage protein breakdown, correspondence of substrate specificity to sequences of proteolytic intermediates/products observed in the plant, localization or eventual delivery to the compartment in which the storage proteins are stored, timing of appearance of active protease and

Table 2. Proteolytic enzymes discussed in this review.

Protease	GenBank	MEROPS Clan/family (Rawlings et al. 2010)	Enzyme Commission
Cysteine endopeptidase			
Vignains (papain-like)			
CPR1 (<i>Vicia sativa</i>)	X75749	Clan CA, Family C1	EC 3.4.22.-
CPR2 (<i>V. sativa</i>)	Z30338		
CPR4 (<i>V. sativa</i>)	Z99172		
EP-C1/CPPH1 (<i>Phaseolus vulgaris</i>)	X56753		
EP-B (<i>Hordeum vulgare</i>)	P25249		
Gliadin/Triticain (<i>Triticum aestivum</i>)	BAE96008		
HvPap-6 (<i>H. vulgare</i>)	AM941120		
HvPap-10 (<i>H. vulgare</i>)	U19384		
Protease C2 (<i>Glycine max</i>)	AI900531		
Pumpkin protease (<i>Cucurbita</i> sp.)	Not Sequenced		
REP-1 (<i>Oryza sativa</i>)	AB004819		
Legumains		Clan CD, Family C13	EC 3.4.22.34
LLP (<i>P. vulgaris</i>)	Z99956		
Proteinase B (<i>Vicia sativa</i>)	Z34899		
VsPB2 (<i>V. sativa</i>)	AJ007743		
Serine endopeptidases		Clan SB, Family S8	EC 3.4.21.25
Protease C1 (<i>G. max</i>)	AY136509		
Protease F (<i>Vigna radiata</i>)	Not sequenced		
Aspartic endopeptidases		Clan AA, Family A1	EC 3.4.23.-
Aspartic proteinase (<i>Arabidopsis thaliana</i>)	AAC49730		
Metallo-endopeptidases		Not identified	EC 3.4.24.-
Metalloproteinase (<i>Fagopyrum esculentum</i>)	Not sequenced		
Endopeptidase – uncertain class		Not identified	Not identified
Thiocalsin (<i>H. vulgare</i>)	Not sequenced		
Exopeptidases			
Serine carboxypeptidase		Clan SC, Family S10	EC 3.4.16.5
Aminopeptidases		Clan MF, Family M17	EC 3.4.11.1
Cytosolic dipeptidases		Not identified	EC 3.4.13.18

environmental conditions in the storage compartment that coincide with the reaction happening in the live plant. As more plant genomes are sequenced, we will see more studies that involve comparative quantitative profiling of native storage proteins and their proteolytic intermediates as a result of loss-of-function mutation or over-expression of specific proteases. We focus this review on the families of proteases that help us illustrate pathways of storage protein mobilization and modes of regulation (Table 2). So far, proteases that have been associated with reserve protein utilization belong to only a few gene families, notably the C1 and C13 cysteine proteases, the S8 and S10 serine proteases, the A1 aspartic proteases, and as yet unidentified metalloproteases and dipeptidases.

Pathways of 7S globulin breakdown

Mobilization of easily accessible regions of storage proteins

Sodium dodecyl sulfate polyacrylamide gel electrophoresis analysis of soy proteins extracted from seedling cotyledons at periodic intervals after imbibition of the seed with water illustrates the stepwise progression by which the 7S globulins are broken down (Wilson et al. 1986). We will refer to these initial steps in the proteolytic degradation of SSPs as 'mobilization'. In some cases, amino acids begin to be available to the plant. In other cases, the proteolytic enzyme(s) catalyze the hydrolysis of a peptide bond on a polypeptide chain of the storage protein resulting in a change in storage protein structure without release of amino acid. In the soybean, the convicilin type of 7S subunits, namely α' and α subunits of β -conglycinin, are first to be digested. These reactions are catalyzed by protease C1, so named because it catalyzes the first step in conglycinin degradation. Protease C1 targets the N-terminal regions that have no homologous region on the β subunit of the β -conglycinin. Consistent with this, the enzyme does not cleave the β subunit. For the same reason, the 50 and 48 kDa proteins derived from α' and α subunits, respectively, are the enzyme's final products (Qi et al. 1992). Thus, this early acting protease leaves intact the double-barreled region of the cupin domains common to 7S globulin subunits. Protease C1 was later found to be a subtilisin-like serine endoprotease (Liu et al. 2001, Barnaby et al. 2004) that has preference for regions with strings of glutamic acid and glutamine residues as shown with both native protein and synthetic peptide substrates (Tan-Wilson et al. 1996, Boyd et al. 2002). Such sequences are present in the α' and α subunit N-terminal regions, outside of their core cupin

domains. It appears that the substrate specificity of the protease that acts soon after seed imbibition allows it to mobilize the most accessible regions of the 7S storage proteins and at the same time preserves the majority of the amino acids in the core of the 7S globulins for later use.

Separation of the cupin domains

Later in soybean seedling growth, a protease that separates the cupin domains of the 7S globulins appears. This is evident *in vivo* as the disappearance of the β -conglycinin β subunit as well as the 50 and 48 kDa products of protease C1 action on β -conglycinin, and the concomitant appearance of proteolytic intermediates of about half their sizes (Wilson et al. 1986). This is followed by rapid and extensive proteolysis of the 26–29 kDa intermediates. The idea of limited cleavage of the storage protein by an initiating endopeptidase that makes the protein vulnerable to further extensive proteolysis by one or more proteases, sometimes including the initiating protease, was first suggested by Shutov and Vaintraub (1987). Indeed, the temporal succession of rapid extensive degradation following initiating proteolytic cleavage of the 7S globulin was observed in several species including the common bean (*Phaseolus vulgaris*) (Senyuk et al. 1998), soybean (Wilson et al. 1986) and vetch (*Vicia sativa*) (Schlereth et al. 2000). These early reports showed something in common: the initial cleavage of the vicilin subunits gives rise to proteolytic intermediates of about half their original size (Freitas et al. 2007). The site of this initiating cleavage has been mapped to the linker region between the two barrel structures that make up the cupin structures of each monomeric unit of the 7S globulins (Senyuk et al. 1998, Seo et al. 2001). Such a pathway is consistent with the known resistance of the cupin barrels to proteolysis, and suggests that separation would allow for more avenues by which proteolytic enzymes can attack the barrel structures.

This hypothesis was experimentally tested in *P. vulgaris*. In this system, the enzyme responsible for cleaving the phaseolin polypeptide chains in half is a legumain, LLP, that targets an Asn–Thr bond in the linker between the cupin domains. The resulting modified phaseolin, consisting of two non-covalently associated domains of about 26 kDa each, is hydrolyzed later in seedling growth by CPPH1, a papain-like cysteine protease (Rotari et al. 1997). The importance of the initiating action of the LLP was demonstrated in the observation that extensive degradation of phaseolin occurred when the storage protein was incubated first with LLP, then with CPPH1, but not when the phaseolin was subjected

to the enzymes in reverse order (Zakharov et al. 2004). In the soybean, the vulnerable sites for the initiating cleavage of the α' , α and β subunits of β -conglycinin was also in the linker region between the cupin domains, at Lys–Thr and Thr–Ile bonds. Such a reaction is catalyzed by soybean protease C2, a papain-like cysteine protease rather than a legumain (Seo et al. 2001). The specificity of cleavage at these bonds is not intrinsic to protease C2, because this enzyme cleaves the oxidized insulin B chain at many sites, but rather to the exposure of few vulnerable sites in the 7S globulin subunits. Protease C2, not protease C1, is the true initiating protease as referred to in the Shutov and Vaintraub model of storage protein mobilization. This is best illustrated by the fact that protease C2 does not require the prior action of protease C1. When presented in vitro with the intact 7S globulin, protease C2 cleaves the linker regions between the cupin domains of the α' , α and β subunits, and, in fact, also digests the accessible N-terminal regions of these subunits (Seo et al. 2001).

Timing of steps in 7S globulin breakdown

Synthesis of the initiating protease only after germination ensures that the storage proteins will remain intact through seed maturation. The *P. vulgaris* LLP mRNA first appears 1 DAI (day after imbibition of the seed with water), with both LLP mRNA and protein peaking at 5–6 DAI (Senyuk et al. 1998, Müntz et al. 2001). Soybean protease C2 activity appears in cotyledons 6 DAI (Seo et al. 2001), protease C1 appearing much earlier (1 DAI) (Qi et al. 1992) to provide amino acids from the most easily accessible regions of the stored globulins. Surprisingly, protease C1 is found in active form in developing seeds (Liu et al. 2001), localized to the PSVs where its protein substrate is located (He et al. 2007). Müntz and coworkers have demonstrated the presence of papain-like proteases CPR2 and CPR4, as well as the legumain VsPB2, in the PSVs of both the embryonic axis and cotyledons of vetch seed (Fischer et al. 2000, Müntz et al. 2001). Purified CPR2 can attack vetch vicilin and the acidic chains of vetch legumin in vitro, and may thus be involved in the initiation of degradation of these SSPs upon seed imbibition. Stored proteases may also be involved in the initiation of storage protein degradation in the cowpea, *Vigna unguiculata* (Minamikawa et al. 1983).

This early deposition of protease in the compartment with the storage protein eliminates lag time due to new enzyme synthesis, ensuring an early supply of amino acids. This is probably especially important in the axis during the early growth of the radicle, where

mobilization of the SSPs occurs earlier than in the cotyledon (Minamikawa et al. 1983, Schlereth et al. 2000, Tiedemann et al. 2000). The preponderance of Asp, Glu, Asn and Gln in those easily accessible regions of the SSPs, degraded by stored proteases such as the soybean protease C1, would provide a supply of intermediates in the Krebs and glyoxylate cycles, as well as amino acids for protein synthesis. Based on these examples, we present a generalized schematic diagram for the early steps in the mobilization of legume 7S globulin reserves in Fig. 1.

Pathways of breakdown of other storage proteins

Proteolysis of the 11S globulins

Breakdown of the 11S globulins follows a similar pattern of limited proteolysis followed by extensive degradation. A 34 kDa Zn-metalloproteinase (EC 3.4.24.-) initiates mobilization of the 11S globulin (actually sedimenting as a 13S globulin) in buckwheat (*Fagopyrum esculentum*), cleaving the 48 and 45 kDa acidic chains into 23 kDa intermediates. In vitro reconstitution of the process shows that this initiating action renders the globulin susceptible to the combined action of a 24 kDa cysteine

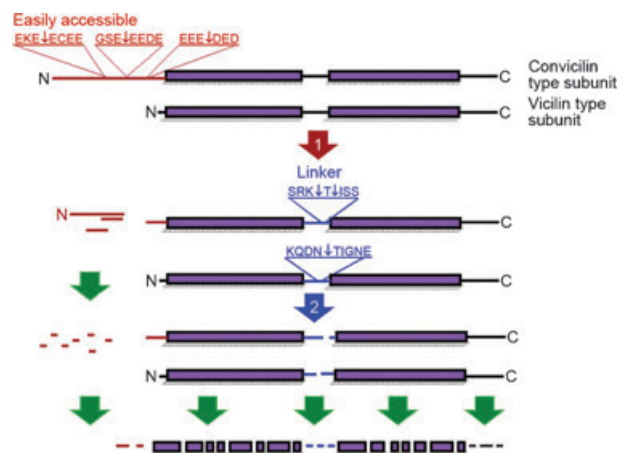


Fig. 1. Schematic diagram of post-germination 7S globulin breakdown in legumes. A generalized sequence of proteolytic events in the breakdown of 7S globulins based on specific examples: the α subunit of soybean β -conglycinin (convicilin type) and the β subunit of *Phaseolus vulgaris* phaseolin (vicilin type). In step 1 (red arrow), easily accessible regions in the N-terminal regions of convicilin type of subunits are cleaved to provide an early supply of amino acids. Products resemble the vicilin type of subunits. In step 2 (blue arrow), another protease cleaves at the linker region between the cupin domains (purple). This separates the cupin domains, allowing access by other proteases. Subsequent steps (green arrows) catalyzed by other endoproteases and possibly those involved in steps 1 and 2 cleave the cupin domains. Exopeptidases (not shown) complete the breakdown to free amino acids.

protease and a 78.5 kDa serine carboxypeptidase that appear only after germination. Neither enzyme can act on the storage protein without prior action of the metalloproteinase (Dunaevsky et al. 1993). The same pattern was reported for pumpkin (*Cucurbita* sp.). The acidic, not the basic, chain of its globulin is cleaved first. During the first 4 days after imbibition, a cysteine protease with pH optimum of 6.8 converts the insoluble protein to water-soluble 43 kDa $F\alpha\beta$, consisting of 30 and 10 kDa chains (Hara and Matsubara 1980). Proteolysis of the 11S globulins of soybean follows a similar pattern. The 40 kDa acidic chains are cleaved to 21 kDa proteolytic intermediates still disulfide-bonded to the 20 kDa basic chains (Wilson et al. 1986). Prior to that, around 1 DAI, soybean protease C1 catalyzes limited proteolysis of the acidic subunits, reducing them in size by 1 or 2 kDa at the C-termini (Qi et al. 1994). Like the regions on the 7S globulins attacked by this enzyme, these C-termini are also accessible and rich in acidic amino acid residues. Proteolysis of the 11S globulins in monocots follows a similar pattern. In the mobilization of glutelin, the 11S globulin of rice, initiating limited proteolysis, was confined to the acidic chain (Yano et al. 2001).

Mobilization of the prolamins

Prolamins, the major storage proteins in many monocot species, tend to form large insoluble complexes by virtue of hydrogen bonding, and in some cases intermolecular disulfide bonding, among its monomeric units. The pathway for storage protein mobilization in monocots starts with solubilization of storage proteins. Metalloproteases and aspartic proteases have been implicated in such process (Jones 2005). Past that, pathways of degradation are simpler than described for the globulins. In barley, several studies led to the finding of the same or very similar papain-like cysteine proteases, variously referred to as EP-B, the 30 kDa or the 31 kDa proteinases of green malt (Jones 2005), which are responsible for digesting C-hordein, a sulfur-poor prolamins in barley. When incubated with hordein (produced in pure form by expression of the cloned gene), EP-B attacks first at five primary cleavage sites. Notably, in this proline-rich protein substrate, none of the initially cleaved bonds involved proline residues. The same enzyme, presumably with the assistance of other enzymes in the endosperm, subsequently participates in digesting the rest of the storage protein (Davy et al. 2000). In wheat, another papain-like cysteine protease named gliadain digests the α/β - and γ -gliadins (Kiyosaki et al. 2007).

Mobilization of the Bowman–Birk trypsin inhibitors

Bowman–Birk trypsin inhibitors are not usually listed among the SSPs. However, with more than 10% of their residues being cysteine, compared to the low proportion of Cys and Met in the abundant storage proteins in legumes (Black et al. 2006), the post-germination degradation of the Bowman–Birk inhibitors provides a significant percentage of the sulfur-containing amino acids available to the plant (Tan-Wilson et al. 1982). Removal of acidic residues at the C-termini starts the degradation of the Bowman–Birk inhibitors. This was shown to be the case in the Bowman–Birk inhibitor HGI-III of horsegram (*Dolichos biflorus*), where the C-terminal residues–SHDD are removed in post-germination growth (Muricken and Gowda 2010). Similarly, the initial cleavage of soybean and mung bean (*Vigna radiata*) Wilczek Bowman–Birk trypsin inhibitors (Bowman–Birk trypsin soybean inhibitor and mung bean trypsin inhibitor, respectively) involves the removal of their respective C-terminal sequences (–EN and –KDDD) (Wilson and Chen 1983, Madden et al. 1985). Initial cleavage is followed by rapid and extensive degradation. These native inhibitors are dimeric (Millar et al. 1969), a result of the interaction of the negatively charged C-terminal residues with a positively charged residue at the inhibitor reactive site (Muricken and Gowda 2010) so that removal of the C-terminal residues would result in dissociation of the dimer. Although achieving the same overall result, different types of proteases are responsible. In soybean, the initial cleavage is carried out by a papain-like cysteine protease (Papastoitis and Wilson 1991); in mung bean, the cleavage is carried out by protease F, a serine protease (Wilson and Tan-Wilson 1987).

Completion of storage protein breakdown

Storage protein breakdown past the initial events is often the result of action of many proteolytic enzymes catalyzing a multiplicity of pathways, working very rapidly. Four papain-like cysteine proteases potentially participate in the degradation of the 7S and 11S globulins of the vetch. Two proteases, CPR2 and CPR4, are synthesized and deposited along with the storage proteins in the PSVs of both cotyledon and embryonic axis during seed development, but do not work at this stage. Later in seedling growth, a papain-like protease CPR1 and a legumain, Proteinase B, are newly synthesized in the same tissues (Fischer et al. 2000, Müntz et al. 2001). The fact that storage proteins are degraded even when detached embryonic axes are treated with cycloheximide, and CPR2 but

not CPR4 attacks the intact storage proteins, points to the role of CPR2 in the initiation of storage protein degradation (Schlereth et al. 2001). Many monocot systems studied also involve multiple cysteine proteases working together in storage protein degradation as illustrated in wheat (Shi and Xu 2009) and barley (Martinez et al. 2009).

Serine carboxypeptidases are common constituents of the PSVs in dicots (Wilson and Tan-Wilson 1987, Dunaevsky et al. 1993, Jones et al. 1996), and in the dead starchy endosperm of monocots (Ranki et al. 1994), strongly suggesting a role for this type of enzyme in helping to complete SSP breakdown. The final products of the action of the endopeptidases and serine carboxypeptidases are presumed to be a mixture of free amino acids and small oligopeptides. In dicots, these would then be transported out of the PSV into the cytosol, where the peptides would be degraded further to amino acids by the action of aminopeptidases (Abdala et al. 1999). In monocots, membrane transporters in the epidermal cells of the scutellum allow the uptake of oligopeptides and amino acids from the endosperm into the embryo (Waterworth and Bray 2006).

Regulation of storage protein mobilization

Like other major processes in seed development and post-germination growth, phytohormones control storage protein accumulation during seed fill and storage protein breakdown upon germination. There are other levels of regulation of storage protein mobilization. In this section, we will also describe situations in which proteases are synthesized during seed development, delivered to the compartment where seed proteins are stored, but prevented from acting due to environmental conditions in that compartment.

Hormonal control

Phytohormones orchestrate the turnaround in metabolism from protein synthesis in the developing seed to proteolysis post-germination. The clearest picture has emerged in the seeds of the monocots such as barley where gibberellic acids (GAs) secreted by the embryo trigger the early mobilization of the reserves within the aleurone cells. In barley, papain-like cysteine endopeptidases (Davy et al. 1998) are synthesized in the aleurone and secreted to the starchy endosperm to mobilize the bulk of the stored reserves there. GA was found to induce the expression during the first 72 h after water imbibition of a cysteine proteinase (REP-1) associated with the degradation of glutelin in rice, while abscisic acid antagonizes it (Kato and Minamikawa 1996, Yang

et al. 2007). Promoter deletion analysis in a transient expression system using rice aleurone tissue established the involvement of GA-responsive promoter elements and the need for transactivation by OsGAMyb transcription factor. The same elements are found in the promoter of the barley EPB1 proteinase gene (Sutoh and Yamauchi 2003). The transcription of three wheat cysteine proteases in the aleurone, triticain α , β and γ , was also determined to be controlled by GA (Kiyosaki et al. 2009). In contrast, hormonal control of storage protein mobilization in dicots is not resolved. Although the action of GA and brassinosteroids in germination and seedling development has been described in *A. thaliana* (reviewed in Holdsworth et al. 2008), their role in protein mobilization is not yet clear in this system because the proteases responsible for protein reserve mobilization have not yet been identified.

Source–sink regulation

In legumes, studies in various species (Yamauchi 2007) suggest possible feedback inhibition by storage protein breakdown products, demonstrated when separation of the axis from the cotyledons disrupts the normal source–sink relationship in the seedling (Davies and Slack 1981). Prevented from being exported to the axis, free amino acids accumulate in the detached cotyledon and presumably inhibit storage protein degradation.

Subcompartmentation and solubility

Besides being compact and providing very limited access to proteases, some native SSPs, such as the 11S globulins of the pumpkin seed cotyledon, are rendered further inaccessible to proteases because they are initially insoluble. A pumpkin (*Cucurbita* sp.) cysteine protease, which in vitro catalyzes the limited specific proteolysis of the 11S globulin acidic chain, is synthesized during seed development. The storage protein is protected from proteolytic action as the two are separated – storage protein in the crystalloid portion of the PSV, protease in the matrix surrounding the crystalloid. It is only when the seed is hydrated that the protease gains access to the storage protein (Hara-Nishimura et al. 1982).

Control by protease inhibitors

The metalloprotease that initiates the cleavage of the 13S globulin in buckwheat is already present in the dry seed, albeit with very low activity. The protease is localized in the PSV along with its 13S globulin protein substrate. The low activity of the metalloproteinase has been ascribed to the existence of a proteinase inhibitor (distinct from the

serine protease inhibitors also in buckwheat) that forms a complex with the proteolytic enzyme. Protease activity begins to increase 2 days after imbibition of buckwheat seed. It has been suggested that Zn^{2+} , released during germination by the hydrolysis of phytins (salts of inositol hexaphosphate and various metal cations), dissociates the metalloprotease-inhibitor complex, releasing the active enzyme (Elpidina et al. 1991).

The phytocystatins are small plant proteins that bind very tightly to cysteine proteases with K_d of 10^{-9} to $10^{-10} M$. They have been described from many plant species, and are especially prominent in monocot seeds (reviewed in Oliveira et al. 2003). Because they interact with endogenous cysteine proteases, it has been suggested that these inhibitors regulate the activity of these hydrolases. In barley, two cathepsin-like cysteine proteases that fully or partially degrade barley hordein in vitro, HvPap-6 and HvPap-10, are exported from the aleurone into the endosperm to catalyze SSP degradation (Martinez et al. 2009). Several observations support the idea that these proteases are regulated by barley phytocystatins: inhibition in vitro, interaction of the cystatins with cysteine proteases demonstrated by bimolecular fluorescence complementation in onion epidermal cells, and decreased cystatin transcript levels upon treatment with GA. Similarly, wheat cystatins WC1 and WC2 inhibit gliadain, the gibberellin-induced cysteine protease that digests the gliadins in vitro. The sites of gliadain and cystatin expression coincide, but the elimination of cystatin to release the protease to act has not been documented (Kiyosaki et al. 2007). Overexpression of phytocystatin in *Arabidopsis* was shown to retard seed germination and proteolysis of SSP; however, no direct evidence for interaction of the phytocystatin with an endogenous cysteine protease in this species has been presented (Hwang et al. 2009). These phytocystatins should not be confused with others that do not control endogenous proteases. In legumes, the phytocystatins are localized in the cytosol, not in the PSV, and are thought to serve as protection in the event that the vacuole containing active protease is ruptured (Flores et al. 2001), or to protect the seed from the proteases of insects, fungi or other pathogens (Dubey et al. 2007).

Redox state

Intermolecular and intramolecular disulfide bonds are a common feature of several families of storage proteins. It is evident from the study of their structures that disulfide bond reduction would reduce the compact folding of these storage proteins. In fact, the redox state within the PSV contributes to the regulation of storage protein degradation, illustrated in wheat (Wong et al. 2004),

barley (Shahpiri et al. 2008) and rice (Yano and Kuroda 2008). In rice, limited proteolysis of the acidic chains of the glutelins is followed by reduction of the disulfide bonds linking the acidic to the basic chain, making the modified storage protein more accessible to subsequent rapid extensive proteolysis. Reduction is carried out by thioredoxin h (Kobrehel et al. 1992, Yano et al. 2001), which is restored to its reduced state by reaction with NADPH, catalyzed by NADPH-thioredoxin reductase (EC 1.6.4.5, GenBank EU314717). The thioredoxin system is also required for activating or maintaining the activity of some proteases. Besse et al. (1996) described the cleavage of wheat gliadin and glutenin by thiolcalsin, a 14 kDa calcium-dependent protease from germinated wheat seeds. For this to occur, the storage proteins as well as thiolcalsin must first be reduced by thioredoxin. The involvement of the thioredoxin system in monocot SSP degradation has been supported by antisense mRNA knock-down studies in wheat (Guo and Yin 2007). Analogous thioredoxin h systems have been described in pea and *Medicago truncatula* (Renard et al. 2011), although their importance in SSP degradation have not yet been demonstrated.

pH

Early post-imbibition proteolysis of soybean β -conglycinin is ensured by the presence of protease C1, already in active form (Liu et al. 2001), already deposited in the PSV of developing seed where its storage protein substrate is sequestered (He et al. 2007). Judging by the protein profiles of the proteins in the dry seed, proteolysis of β -conglycinin in developing seed is minimal despite the coincidence of protease and substrate in time and location. Regulation of enzyme activity in the maturing seed has been ascribed to the fact that the enzyme is active only between pH 3.0 and 5.5, with an optimum at pH 4.5 (Qi et al. 1992), and the demonstration by acridine orange staining of acidic conditions in the PSVs only after germination (He et al. 2007). Acidification of the starchy endosperm in monocots is also an important component of storage protein mobilization in that tissue. Upon imbibition, the aleurone layer of wheat regulates the pH of the starchy endosperm at pH close to 4.2 by the secretion and uptake of malic acid in a GA-dependent process (Dominguez and Cejudo 1999, Martinez-Camacho et al. 2004). A similar acidification of the starchy endosperm has been described in barley, although acidification in this case begins late in seed maturation, rather than after seed imbibition (Macnicol and Jacobsen 1992). This acidification produces an environment conducive to protease activity and also increases the solubility of SSPs.

Cross-over proteolytic events

We have described proteolytic events that are usually associated in the literature with either processing of storage protein precursors to the 'mature' forms, or storage protein mobilization, ultimately to free amino acids. In the discussion so far, the processing occurs in developing seed; the mobilization occurs post-imbibition. Elaborate regulatory strategies ranging from hormonal control of proteolytic enzyme synthesis to switching on of proteolytic activity only upon acidification ensure that the proteolytic events associated with mobilization do not occur during seed fill. This is not always so. As noted above, acidification in the barley endosperm begins during seed maturation. In this section, we will describe cross-over proteolytic events, when processing of storage protein precursors continues in the seedling, or when partial degradation of the storage proteins, such as is seen at the start of mobilization, occurs in the developing seed.

Storage protein processing occurring post-germination

As noted earlier, plant species vary in their processing of the 2S albumin precursors to the disulfide-bonded two-chain structure. In some species, the 2S albumin remains as a single chain. Therefore, it is not clear whether, as commonly believed, processing of 2S albumin is necessary for close packing. The processing may even be a preliminary to mobilization of the proteins in seedling growth. This was shown in castor bean where not all of the 2S albumin precursors were completely processed to the two-chain 'mature' structure during seed development. After germination the processing resumes, followed by degradation of the subunits of the processed 2S albumins (Ahn and Chen 2007).

Storage protein proteolysis occurring in the developing seed

Not all plants exhibit tight control of SSP proteolysis during seed maturation. This is evident in the mature seed of pea (*Pisum sativum*), in the form of vicilin polypeptides with N-termini usually associated with the proteolytic reactions that occur only after germination. The full-length polypeptide chains of mature pea vicilins consists of three consecutive regions, α - β - γ . The cleavage products observed in mature peas consist of polypeptides with regions α , α - β , β , β - γ and γ (Gatehouse et al. 1983). Pea vicilins are encoded by six genes. Detailed proteomic analysis (Bourgeois et al. 2009) of pea seeds showed polypeptide chain cleavage

only in those vicilins that bear a -KEN (not -EED) sequence at the β - γ junction. The limited proteolysis of the pea vicilins is well known. For this review, we have collected several more examples of such 'precocious' storage protein proteolysis. N-terminal sequencing and proteomic analysis of lentil (*Lens culinaris*) seeds show extensive preliminary modification of its 50 kDa vicilin to 26 and 16 kDa chains. Proteolysis occurred at cleavage sites analogous to those described for pea vicilin (Lopez-Torreon et al. 2003, Scippa et al. 2010). Similarly, most of the 58 and 54 kDa vicilins of lima bean (*Phaseolus lunatus*) in the dry seed are already cleaved. As with the other species, the products were about half the size of the parent chain, with N- and C-termini that indicate cleavage into N-terminal 25 and 21 kDa and C-terminal 35 kDa regions (Sparvoli et al. 1996). Storage protein proteolysis was also observed in lupin (*Lupinus angustifolius*). A 2D-PAGE proteomic analysis conducted on lupin seeds harvested at different stages during seed development identified gel spots in the molecular weight range from 20 to 26 kDa ascribed to the beta-conglutin storage protein, a convicilin type of 7S globulin subunit that is 49 kDa in its mature form (Foley et al. 2011). These examples suggest that proteolysis of storage proteins leading to peptide bond cleavage usually associated with events that occur after water imbibition of the seed may occur during seed development more often than originally thought.

It is interesting to note that immunoblots probed with IgE antibodies from patients with food allergies (Sanchez-Monge et al. 2004) showed weaker binding to the <16 kDa chains of pea vicilin compared to those polypeptides that have at least one complete cupin domain. This is consistent with the idea proposed by Mills et al. (2002) and Barre et al. (2005) that the stability of the cupin domain, including resistance to proteolysis, contributes to the immunogenicity of the globulin storage proteins. Considering that some plant species already produce proteases and conditions that allow for initiation of storage protein breakdown during seed maturation, it is possible to think that increasing the expression of such proteases would be a possible strategy for decreasing the allergenicity due to vicilins.

Concluding remarks

At first glance, the proteolytic degradation of SSPs appears to be a relatively simple process, avoided during seed development, and commenced upon imbibition and germination in all plant species. As should be apparent from this review, the dividing line between seed maturation, when storage proteins are synthesized, and seedling growth, when the stored seed proteins

are degraded to free amino acids, is not as clear cut as thought. Pathways of storage protein mobilization among different plant species show more commonality with regard to the vulnerable cleavage sites on the storage proteins and more variation with respect to the proteolytic enzymes recruited to carry out those steps. Understanding of the subtleties of storage protein proteolysis has become more important with efforts to genetically modify plants to produce seeds with augmented nutritional value, to use seeds for the production of high-value proteins, or to reduce allergenicity. Understandably, much of the work done in this research area has been carried out on plant species that are important as food crops. Genetic and proteomic approaches will enhance this area of research as the genomes of more of these plant species are sequenced and annotated. However, when working with plant species that have been cultivated for so long, one must exercise caution in making too many interpretations in light of evolution. Thus, expansion of the study of the mobilization of storage proteins to species that have not been subjected to domestication can be of great value.

References

- Abdala AP, Takeda LH, Freitas Jr JO, Alves KB (1999) Purification and partial characterization of *Phaseolus vulgaris* seed aminopeptidase. *Braz J Med Biol Res* 32: 1489–1492
- Adachi M, Kanamori J, Masuda T, Yagasaki K, Kitamura K, Mikami B, Utsumi S (2003) Crystal structure of soybean 11S globulin: glycinin A3B4 homohexamer. *Proc Natl Acad Sci USA* 100: 7395–7400
- Ahn Y-J, Chen GQ (2007) Temporal and spatial expression of 2S albumin in castor (*Ricinus communis* L.). *J Agric Food Chem* 55: 10043–10049
- Anjum FH, Khan MR, Din A, Saeed M, Pasha I, Arshad MU (2007) Wheat gluten: high molecular weight glutenin subunits – structure, genetics, and relation to dough elasticity. *J Food Sci* 72: R56–R63
- Barnaby NG, He F, Liu X, Wilson KA, Wilson KA, Tan-Wilson A (2004) Light-responsive subtilisin-related protease in soybean seedling leaves. *Plant Physiol Biochem* 42: 125–134
- Barre A, Borges JP, Roug   P (2005) Molecular modeling of the major peanut allergen Ara h 1 and other homotrimeric allergens of the cupin superfamily: a structural basis for their IgE-binding cross-reactivity. *Biochimie* 87: 499–505
- Besse I, Wong JH, Kobrehel K, Buchanan BB (1996) Thiolcalsin: a thioredoxin-linked, substrate-specific protease dependent upon calcium. *Proc Natl Acad Sci USA* 93: 3169–3175
- Black M, Bewley JD, Halman P (eds) (2006) *The Encyclopedia of Seeds. Science, Technology and Uses*. CAB International, Wallingford, pp 226
- Boothe J, Nykiforuk C, Shen Y, Zaplachinsky S, Szarka S, Kuhlman P, Murray E, Morck D, Moloney MM (2010) Seed-based expression systems for plant molecular farming. *Plant Biotechnol J* 8: 588–606
- Bourgeois M, Jacqu  n F, Savoie V, Sommerer N, Labas V, Henry C, Burstin J (2009) Dissecting the proteome of pea mature seeds reveals the phenotypic plasticity of seed protein composition. *Proteomics* 9: 254–271
- Bown D, Ellis THN, Gatehouse JA (1988) The sequence of a gene encoding convicilin from pea (*Pisum sativum* L.) shows that convicilin differs from vicilin by an insertion near the N-terminus. *Biochem J* 251: 717–726
- Boyd PM, Barnaby N, Tan-Wilson A, Wilson KA (2002) Cleavage specificity of the subtilisin-like protease C1 from soybean. *Biochim Biophys Acta* 1596: 269–282
- Cunha NB, Araujo ACG, Leite A, Murad AM, Vianna GR, Rech EL (2010) Correct targeting of proinsulin in protein storage vacuoles of transgenic soybean seeds. *Genet Mol Res* 9: 1163–1170
- Davies HV, Slack PT (1981) The control of food mobilization in seeds of dicotyledonous plants. *New Phytol* 88: 41–51
- Davy A, Svendsen I, Sorensen SO, Sorensen MB, Rouster J, Meldal M, Simpson DJ, Cameron-Mills V (1998) Substrate specificity of barley endoproteases EP-A and EP-B. *Plant Physiol* 117: 255–261
- Davy A, Sorensen B, Svendsen I, Cameron-Mills V, Simpson DJ (2000) Prediction of protein cleavage sites by the barley cysteine endoproteases EP-A and EP-B based on the kinetics of synthetic peptide hydrolysis. *Plant Physiol* 122: 137–145
- D’Hondt K, Bosch D, Damme JV, Goethals M, Vandekerckhove J, Krebbers E (1993) An aspartic proteinase present in seeds cleaves *Arabidopsis* 2 S albumin precursors *in vitro*. *J Biol Chem* 268: 20884–20891
- D’Ovidio R, Masci S (2004) The low-molecular weight glutenin subunits of wheat gluten. *J Cereal Sci* 39: 321–339
- Dominguez F, Cejudo FJ (1999) Patterns of starch endosperm acidification and protease gene expression in wheat grains following germination. *Plant Physiol* 119: 81–87
- Dubey VK, Pande M, Singh BK, Jagannadham MV (2007) Papain-like proteases: applications of their inhibitors. *Afr J Biotechnol* 6: 1077–1086
- Dunaevsky YE, Belozersky MA, Voskoboinikova NE (1993) In vitro reconstitution of the proteolytic system responsible for hydrolysis of the 13 S globulin in buckwheat seeds. *FEBS Lett* 324: 216–218
- Dunwell JM, Culham A, Carter CE, Sosa-Aguirre CR, Goodenough PW (2001) Evolution of functional

- diversity in the cupin superfamily. *Trends Biochem Sci* 26: 740–746
- Elpidina EN, Voskoboinikova NE, Belozersky MA, Dunaevsky YE (1991) Localization of a metalloproteinase and its inhibitor in protein bodies of buckwheat seeds. *Planta* 185: 46–52
- Fischer J, Becker C, Hillmer S, Horstmann C, Neubohn B, Schlereth A, Senyuk V, Shutov A, Müntz K (2000) The families of papain- and legumain-like cysteine proteinases from embryonic axes and cotyledons of *Vicia* seeds: developmental patterns, intracellular localization and functions in globulin proteolysis. *Plant Mol Biol* 43: 83–101
- Flores VMQ, Louro RP, Xavier-Filho J, Barratt DHP, Shewry PR, Fernandes KVS (2001) Temporal and tissue localization in the cowpea (*Vigna unguiculata*) cystatin. *Physiol Plant* 112: 195–201
- Foley RC, Gao L, Spriggs A, Soo LYC, Goggin D, Smith PMC, Atkins CA, Singh KB (2011) Identification and characterization of seed storage protein transcripts from *Lupinus angustifolius*. *BMC Plant Biol* 11: 59
- Freitas RL, Teixeira AR, Ferreira RB (2007) Vicilin-type globulins follow distinct patterns of degradation in different species of germinating legume seeds. *Food Chem* 102: 323–329
- Gatehouse JA, Lycett GW, Delauney AJ, Croy RRD, Boulter D (1983) Sequence specificity of the post-translational proteolytic cleavage of vicilin, a seed storage protein of pea (*Pisum sativum* L.). *Biochem J* 212: 427–432
- Gruis DF, Schulze J, Jung R (2004) Storage protein accumulation in the absence of the vacuolar processing enzyme family of cysteine proteases. *Plant Cell* 16: 270–290
- Guo HX, Yin J (2007) Effects of antisense TRX gene on hydrolysis of storage proteins in germination of transgenic wheat (*Triticum aestivum* L.). *Acta Agron Sin* 33: 885–890
- Hara I, Matsubara H (1980) Pumpkin (*Cucurbita* sp.) seed globulin V. Proteolytic activities involved in globulin degradation in ungerminated seeds. *Plant Cell Physiol* 21: 219–232
- Hara-Nishimura I, Nishimura M, Matsubara H, Akazawa T (1982) Subcellular localization of proteinase catalyzing the limited hydrolysis of pumpkin globulin. *Plant Physiol* 70: 699–703
- He F, Huang F, Wilson KA, Tan-Wilson A (2007) Protein storage vacuole acidification as a control of storage protein degradation in soybeans. *J Exp Bot* 58: 1059–1070
- Holdsworth MJ, Bentsink L, Soppe WJJ (2008) Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. *New Phytol* 179: 33–54
- Hwang JE, Hong JK, Je JH, Lee KO, Kim DY, Lee SY, Lim CO (2009) Regulation of seed germination and seedling growth by an *Arabidopsis* phytocystatin isoform, AtCYS6. *Plant Cell Rep* 28: 1623–1632
- Jiang L, Phillips TE, Hamm CA, Drozdowicz YM, Rea PA, Maeshima M, Rogers SW, Rogers JC (2001) The protein storage vacuole: a unique compound organelle. *J Cell Biol* 155: 991–1002
- Jones BL (2005) Endoproteases of barley and malt. *J Cereal Sci* 42: 139–156
- Jones CG, Lycett GW, Tucker GA (1996) Protease inhibitor studies and cloning of a serine carboxypeptidase cDNA from germinating seeds of pea (*Pisum sativum* L.). *Eur J Biochem* 235: 574–578
- Jung R, Scott MP, Nam Y-W, Beaman TW, Bassüner R, Saalbach I, Müntz K, Nielsen NC (1998) The role of proteolysis in the processing and assembly of 11S seed globulins. *Plant Cell* 10: 343–357
- Kato H, Minamikawa T (1996) Identification and characterization of a rice cysteine endopeptidase that digests glutelin. *Eur J Biochem* 239: 310–316
- Kawakatsu T, Takaiwa F (2010) Cereal seed storage protein synthesis: fundamental processes for recombinant protein production in cereal grains. *Plant Biotechnol J* 8: 939–953
- Kiyosaki T, Matsumoto I, Asakura T, Funaki J, Kuroda M, Misaka T, Arai S, Abe K (2007) Gliadain, a gibberellin-inducible cysteine proteinase occurring in germinating seeds of wheat, *Triticum aestivum* L., specifically digests gliadin and is regulated by intrinsic cystatins. *FEBS J* 274: 1908–1917
- Kiyosaki T, Asakura T, Matsumoto I, Tamura T, Terauchi K, Funaki J, Kuroda M, Misaka T, Abe K (2009) Wheat cysteine proteases triticain α , β and γ exhibit mutually distinct responses to gibberellin in germinating seeds. *J Plant Physiol* 166: 101–106
- Kobrehel K, Wong JH, Balogh A, Kiss F, Yee BC, Buchanan BB (1992) Specific reduction of wheat storage proteins by thioredoxin h. *Plant Physiol* 99: 919–924
- Kumar P, Rao AGA, Haritharaputran S, Chandra N, Gowda LR (2004) Molecular mechanism of dimerization of Bowman-Birk inhibitors. *J Biol Chem* 279: 30425–30432
- Lehmann K, Schweimer K, Rese G, Randow S, Suhr M, Becker W-M, Vieths S, Rösch P (2006) Structure and stability of 2S albumin-type peanut allergens: implications for the severity of peanut allergic reactions. *Biochem J* 395: 463–472
- Liu X, Zhang Z, Barnaby N, Wilson KA, Tan-Wilson A (2001) Soybean subtilisin-like protease involved in initiating storage protein degradation. *Seed Sci Res* 11: 55–68
- Lopez-Torreon G, Salcedo G, Martin-Esteban M, Diaz-Perales A, Pascual CY, Sanchez-Monge R (2003) Len c 1, a major allergen and vicilin from lentil seeds:

- protein isolation and cDNA cloning. *J Allergy Clin Immunol* 112: 1208–1215
- Macnicol PK, Jacobsen JV (1992) Endosperm acidification and related metabolic changes in the developing barley grain. *Plant Physiol* 98: 1098–1104
- Madden MA, Tan-Wilson AL, Wilson KA (1985) Proteolysis of soybean Bowman-Birk trypsin inhibitor during germination. *Phytochemistry* 24: 2811–2815
- Maleki SJ, Kopper RA, Shin DS, Chun-Wook P, Compadre CM, Sampson H, Burks AW, Bannon GA (2000) Structure of the major peanut allergen Ara h 1 may protect IgE-binding epitopes from degradation. *J Immunol* 164: 5844–5849
- Martinez M, Cambra I, Carrillo L, Diaz-Mendoza M, Diaz I (2009) Characterization of the entire cystatin gene family in barley and their target cathepsin L-like cysteine-proteases, partners in the hordein mobilization during seed germination. *Plant Physiol* 151: 1531–1545
- Martinez-Camacho JL, Gonzalez-de la Vara L, Hamabata A, Mora-Escobedo R, Calderon-Salinas V (2004) A pH-stating mechanism in isolated wheat (*Triticum aestivum*) aleurone layers involves malic acid metabolism. *J Plant Physiol* 161: 1289–1298
- Maruyama N, Adachi M, Takabashi K, Yagasaki A, Kohno M, Takenaka Y, Okuda E, Nakagawa S, Mikami B, Utsumi S (2001) Crystal structures of recombinant and native soybean β -conglycinin β homotrimers. *Eur J Biochem* 268: 3595–3604
- Matsuda T, Matsubara T, Hino S (2006) Immunogenic and allergenic potentials of natural and recombinant innocuous proteins. *J Biosci Bioeng* 101: 203–211
- Millar DBS, Willick GE, Steiner RF, Frattali V (1969) Soybean inhibitors IV. The reversible self-association of a soybean proteinase inhibitor. *J Biol Chem* 244: 281–284
- Mills ENC, Jenkins J, Marigheto N, Belton PS, Gunning AP, Morris VJ (2002) Allergens of the cupin superfamily. *Biochem Soc Trans* 30: 925–929
- Mills ENC, Jenkins JA, Alcocer MJ, Shewry PR (2004) Structural, biological, and evolutionary relationships of plant food allergens sensitizing via the gastrointestinal tract. *Crit Rev Food Sci Nutr* 44: 379–407
- Minamikawa T, Suzuki Y, Koshiba T (1983) Metabolic changes in axes of germinating *Vigna unguiculata* seeds as related to the effects of removal of cotyledons. *Plant Cell Physiol* 24: 1361–1369
- Müntz K, Belozersky MA, Dunaevsky YE, Schlereth A, Tiedermann J (2001) Stored proteinases and the initiation of storage protein mobilization in seeds during germination and seedling growth. *J Exp Bot* 52: 1741–1752
- Muricken DG, Gowda LR (2010) Functional expression of horse gram (*Dolichos biflorus*) Bowman-Birk inhibitor and its self-association. *Biochim Biophys Acta* 1804: 1413–1423
- Oliveira AS, Xavier-Filho J, Sales MP (2003) Cysteine proteinases and cystatins. *Braz Arch Biol Technol* 46: 91–104
- Osborne TB (1924) *The Vegetable Proteins*, 2nd Edn. Longmans Green and Co., London
- Pantoja-Uceda D, Palomares O, Bruix M, Villaba M, Rodriguez R, Rico M, Santoro J (2004) Solution structure and stability against digestion of rpoBnIb, a recombinant 2S albumin from rapeseed: relationship to its allergenic properties. *Biochemistry* 43: 16036–16045
- Papadimitrakou F, Wilson KA (1991) Initiation of the degradation of the soybean Kunitz and Bowman-Birk trypsin inhibitors by a cysteine protease. *Plant Physiol* 96: 1086–1092
- Qi X, Wilson KA, Tan-Wilson AL (1992) Characterization of the major protease involved in the soybean beta-conglycinin storage protein mobilization. *Plant Physiol* 99: 725–733
- Qi X, Chen R, Wilson KA, Tan-Wilson AL (1994) Characterization of a soybean beta-conglycinin degrading protease cleavage site. *Plant Physiol* 104: 127–135
- Qi RF, Song ZW, Chi CW (2005) Structural features and molecular evolution of Bowman-Birk protease inhibitors and their potential application. *Acta Biochim Biophys Sin* 37: 283–292
- Qi PF, Wei YM, Yue YW, Yan ZH, Zheng YL (2006) Biochemical and molecular characterization of gliadins. *Mol Biol (Russ)* 40: 713–723
- Ranki H, Mendez-Lozano J, Sopanen T (1994) Three carboxypeptidases occurring in the starchy endosperm of germinating barley grain have different sites of synthesis. *Physiol Plant* 91: 90–96
- Rawlings ND, Barrett AJ, Bateman A (2010) MEROPS: the peptidase database. *Nucleic Acids Res* 38: D227–D233. Available at <http://merops.sanger.ac.uk/> (accessed 2 November 2011)
- Renard M, Alkhalfouri F, Schmidt-Keichinger C, Ritzenthaler C, Montrichard F (2011) Identification and characterization of thioredoxin h isoforms differentially expressed by germinating seeds of the model legume *Medicago truncatula*. *Plant Physiol* 55: 1113–1126
- Reyes FC, Chung T, Holding D, Jung R, Vierstra R, Otegui MS (2011) Delivery of prolamins to the storage vacuole in maize aleurone cells. *Plant Cell* 23: 769–784
- Rotari V, Senyuk V, Horstmann C, Jivotovskaya A, Vaintraub I (1997) Proteinase A-like enzyme from germinated kidney bean seeds. Its action on phaseolin and vicilin. *Physiol Plant* 100: 171–177
- Sanchez-Monge R, Lopez-Torres G, Pascual CY, Varela J, Martin-Esteban M, Salcedo G (2004) Vicilin and convicilin are potential major allergens from pea. *Clin Exp Allergy* 34: 1747–1753
- Schlereth A, Becker C, Horstmann C, Tiedermann J, Müntz K (2000) Comparison of globulin mobilization and

- cysteine proteinases in embryonic axes and cotyledons during germination and seedling growth of vetch (*Vicia sativa* L.). *J Exp Bot* 51: 1423–1433
- Schlereth A, Standhardt D, Mock H-P, Müntz K (2001) Stored cysteine proteinases start globulin mobilization in protein bodies of embryonic axis and cotyledons during vetch (*Vicia sativa* L.) seed germination. *Planta* 212: 718–727
- Scippa GS, Rocco M, Iallicco M, Trupiano D, Viacosi V, Michele MD, Arena S, Chiatante D, Scaloni A (2010) The proteome of lentil (*Lens culinaris* Medik.) seeds: discriminating between landraces. *Electrophoresis* 31: 497–506
- Senyuk V, Rotari V, Becker C, Zakharov A, Horstmann C, Müntz K, Vaintraub IA (1998) Does an asparaginyl-specific cysteine endopeptidase trigger phaseolin degradation in cotyledons of kidney bean seedlings? *Eur J Biochem* 258: 546–558
- Seo SB, Tan-Wilson A, Wilson KA (2001) Protease C2 in the continuing mobilization of soybean β -conglycinin seed proteins. *Biochim Biophys Acta* 1545: 192–206
- Shahpiri A, Svensson B, Finnie C (2008) The NADPH-dependent thioredoxin reductase/thioredoxin system in germinating barley seeds: gene expression, protein profiles, and interactions between isoforms of thioredoxin h and thioredoxin reductase. *Plant Physiol* 146: 789–799
- Shewry PR, Halford NG (2002) Cereal seed storage proteins: structure, properties and role in grain utilization. *J Exp Bot* 53: 947–958
- Shi C, Xu LL (2009) Characters of cysteine endopeptidases in wheat endosperms during seed germination and subsequent seedling growth. *J Integr Plant Biol* 51: 52–57
- Shutov AI, Vaintraub IA (1987) Degradation of storage proteins in germinating seeds. *Phytochemistry* 26: 1557–1566
- Shutov AD, Bäumlein H, Blattner FR, Müntz K (2003) Storage and mobilization as antagonistic functional constraints on seed storage globulin evolution. *J Exp Bot* 54: 1645–1654
- Sparvoli F, Daminati MG, Lioi L, Bollini R (1996) In vivo endoproteolytically cleaved phaseolin is stable and accumulates in developing *Phaseolus lunatus* L. seeds. *Biochim Biophys Acta* 1292: 15–22
- Sutoh K, Yamauchi D (2003) Two cis-acting elements necessary and sufficient for gibberellins-upregulated proteinase expression in rice seeds. *Plant J* 34: 635–645
- Tan-Wilson AL, Rightmire BR, Wilson KA (1982) Different rates of metabolism of soybean proteinase inhibitors during germination. *Plant Physiol* 70: 493–497
- Tan-Wilson AL, Liu X, Chen R, Qi X, Wilson KA (1996) An acidic amino acid-specific protease from germinating soybeans. *Phytochemistry* 42: 313–319
- Tiedemann J, Neubohn B, Müntz K (2000) Different functions of vicilin and legumin are reflected in the histopattern of globulin mobilization during germination of vetch (*Vicia sativa* L.). *Planta* 211: 1–12
- Vitale A, Hinz G (2005) Sorting of proteins to storage vacuoles: how many mechanisms? *Trends Plant Sci* 10: 316–323
- Waterworth WM, Bray CM (2006) Enigma variations for peptides and their transporters in higher plants. *Ann Bot* 98: 1–8
- Wilson KA, Chen JC (1983) The amino acid sequence of mung bean trypsin inhibitor and its modified forms appearing during germination. *Plant Physiol* 71: 341–349
- Wilson KA, Tan-Wilson AL (1987) Characterization of the proteinase that initiates the degradation of the trypsin inhibitor in germinating mung beans (*Vigna radiata*). *Plant Physiol* 84: 93–98
- Wilson KA, Rightmire BR, Chen JC, Tan-Wilson AL (1986) Differential proteolysis of glycinin and beta-conglycinin polypeptides during soybean germination and seedling growth. *Plant Physiol* 82: 71–76
- Wong JH, Cai N, Tanaka CK, Vensel WH, Hurkman WJ, Buchanan BB (2004) Thioredoxin reduction alters the solubility of proteins of wheat starchy endosperm: an early event in cereal germination. *Plant Cell Physiol* 45: 407–415
- Yamauchi D (2007) Hormonal regulation of the expression of cysteine proteinase genes in germinated cotyledons of common bean seeds. *Plant Biotechnol* 24: 231–235
- Yang P, Li X, Wang X, Chen H, Chen F, Shen S (2007) Proteomic analysis of rice (*Oryza sativa*) seeds during germination. *Proteomics* 7: 3358–3368
- Yano H, Kuroda S (2008) Introduction of the disulfide proteome: application of a technique for the analysis of plant storage proteins as well as allergens. *J Proteome Res* 7: 3071–3079
- Yano H, Wong JH, Cho M-J, Buchanan BB (2001) Redox changes accompanying the degradation of seed storage proteins in germinating rice. *Plant Cell Physiol* 42: 879–883
- Zakharov A, Carchilan M, Stepurina T, Wilson K, Vaintraub I (2004) A comparative study of the role of the major proteinases of germinated common bean (*Phaseolus vulgaris* L.) and soybean (*Glycine max* (L.) Merrill seeds in the degradation of their storage proteins. *J Exp Bot* 55: 2241–2249