Title: Proteolysis of the Peanut Allergen Ara h 1 by an Endogenous Aspartic Protease

Authors: Karl A. Wilson and Anna Tan-Wilson*, Department of Biological Sciences, State University of New York at Binghamton, Binghamton, NY 13902-6000 *corresponding author: E-mail address: annatan@binghamton.edu

ABBREVIATIONS: DAI, days after imbibition; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; PSV, protein storage vacuole

1. INTRODUCTION

Legume seeds, valued as foods with high protein content, are well known to have abundant stores of storage protein. One such protein in the peanut (*Arachis hypogaea* L.) is Ara h 1, also known as conarachin, identified by its sequence to be a 7S globulin. Like other 7S globulins, cloned and expressed Ara h 1 assembles into homotrimers, within which each Ara h 1 monomer assumes a bicupin fold, a structure with two β -barrel core domains connected by a linker region (Chruszcz et al. 2011). The structures of cupin domains are highly compact and stable, and are thought to have evolved to maximize amino acid storage capabilities for later use by the seedling, and at the same time to resist premature proteolysis while the proteins are synthesized and accumulated in the developing seed (Shutov et al. 2003).

Reserve protein mobilization is a vital process in germination (Tan-Wilson and Wilson 2012). Yet, except for a study carried out on peanut seedlings ten days after seed imbibition (Mahaboob Basha and Cherry 1978), not much is known about the enzymes that mobilize the peanut storage globulins shortly after imbibition of the seed with water. The pathways of storage protein mobilization in soybean (*Glycine max* L. Merrill) and *Phaseolus vulgaris* L. (variously named, e.g. common bean, kidney bean, string bean) demonstrate how plants have evolved to catalyze the hydrolysis of 7S globulins efficiently, strategies that may help us map out the proteolysis of storage proteins, specifically of the 7S globulin Ara h 1, in germinating peanut. The soybean initiates digestion of two of its 7S globulins (the β -conglycinin α and α' subunits) by successive cleavage at the N-terminal region outside of the bicupin domain. Proteolysis is catalyzed by a subtilisin-like serine protease (protease C1) (Qi et al. 1994; Tan-Wilson et al. 1996) that leaves the bicupin structure intact. Onset of the initial breakdown is observed within one day of seed imbibition because protease C1 is already present in active form in the dry seed, in the same membrane-bound protein storage vacuoles (PSVs) where the storage proteins are sequestered. Protease C1 is prevented

1 from acting during seed development because it acts only at low pH, a condition that is met
2 in the protein storage vacuoles only after seed imbibition (He et al. 2007). Its action is
3 followed by a papain-like cysteine protease (protease C2) that cleaves at the linker region
4 between the two cupin domains (Seo et al. 2001), followed by rapid, extensive proteolysis.
5 The breakdown of phaseolin, the 7S globulin of *P. vulgaris* that lacks an extensive N-
6 terminal region outside of its bicupin structure, starts with cleavage at the linker between the
7 two cupin domains. Proteolysis is catalyzed by a legumain-like cysteine protease (LLP).
8 Subsequent action of a papain-like protease (CPPh) working in concert with the LLP leads
9 to extensive degradation of the modified phaseolin (Zakharov et al. 2004). If present in
10 developing seeds and left unchecked, the action of these cysteine proteases would prevent
11 accumulation of seed protein reserves (Müntz et al. 2001; Tan-Wilson and Wilson 2012).
12 Accordingly, soybean protease C2 and *P. vulgaris* LLP and CPPh are synthesized only after
13 germination (Senyuk et al. 1998; Seo et al. 2001); thus cleavage into the bicupin structure
14 occurs post-germination.

15 This is not always the case, as illustrated in the proteolytic cleavage of pea (*Pisum*
16 *sativum* L.) vicilin, also a 7S globulin. This storage protein was found to undergo cleavage
17 even as the protein was accumulating in developing seed, nicking but not disassembling the
18 trimer, and not proceeding into extensive proteolysis. One site of cleavage is, as in soybean
19 and *P. vulgaris*, in the linker between the two cupin domains; a second site is within an
20 exposed loop of the C-terminal cupin, the second cupin domain counting from N-terminus to
21 C-terminus (Gatehouse et al. 1983; Shutov et al. 2003).

22 The study of proteolysis of Ara h 1 is also important in another context. Ara h 1 is
23 one of the major food allergens in the peanut, a property that is thought to stem partly from
24 its resistance to proteolysis by mammalian gut enzymes (Chassaigne et al. 2009; Moreno
25 2007). Proteolysis using exogenous proteases already used in food fermentation or other
26 preparation has been suggested as a strategy for reducing allergenicity (Cabanillas et al.

2012). The use of proteolytic enzymes that evolved in the peanut plant expressly to break down Ara h 1 during germination could be used for this purpose. We have characterized the enzyme activity in peanut seedling cotyledons responsible for the initial breakdown of the storage globulins with respect to pH optimum and mechanistic class, and found the participation of a class of proteolytic enzyme that has been more often associated with protein processing in developing seed (Simoes and Faro 2004) rather than with storage protein breakdown in the seedling. Focusing on the proteolysis of the 7S globulin, Ara h 1, we suggest a possible principal pathway of Ara h 1 breakdown that is based on the sizes and LC-MS/MS characterization of principal intermediates generated *in vitro*, and building upon information on the proteolysis of other 7S globulins. Remarkably, these findings explain a food preparation method that reduces peanut allergy due to Ara h 1.

2. MATERIALS AND METHODS

2.1. Plant materials

Peanut (*Arachis hypogaea* L., cv Jumbo Virginia) seeds were obtained from Gurney's Seed & Nursery Company (Greendale, IN). For germination studies, seeds were planted in moist vermiculite at a depth of 1 cm and grown in a dark growth chamber with a 16 h 25°C / 8 h 20°C cycle. Developing seeds were harvested from plants grown in natural light in the greenhouse from seeds that were planted in ProMix Professional soil mix (Premier Horticulture, Ltd., Quakerstown, PA) and washed quartz sand (4:1 by volume). Harvested cotyledons and seeds were rinsed with distilled water, blotted dry and stored at -20°C until needed.

2.2. Chemical reagents

AEBSF (4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride), and E-64 (N-[N-(L-3-transcarboxyirane-2-carbonyl)-L-leucyl]-agmatine) were obtained from Research Products International (Mt. Prospect, IL); pepstatin A, and disodium EDTA were from Sigma-Aldrich (St. Louis, MO). Reagents used for electrophoresis were from BioRad Laboratories (Hercules, CA).

2.3. Preparation of seed and cotyledon extracts

Partially thawed seeds or cotyledons were minced to 1 mm fragments, suspended in either Extraction Buffer A (50 mM Tris-HCl, 0.1M NaCl, pH 8.0) or Extraction Buffer B (0.05 M sodium acetate, 0.1 M NaCl, 0.01 M 2-mercaptoethanol, pH 4.7), using 1 ml buffer per cotyledon, and homogenized on ice with a Biohomogenizer (Biospec Products Inc., Bartlesville, OK). After holding on ice for 1 h, the homogenate was centrifuged at 38,700 x g at 4°C, and the clear supernatant frozen in aliquots at -20°C. Alternatively, seeds or cotyledons were ground to a very fine powder in liquid nitrogen and suspended in extraction buffer with stirring for 4 h at 4°C, and centrifuged as described above. To make extract to serve as protein substrate, the procedure was carried out on dry seeds with the addition of "Halt Protease Inhibitor Cocktail" (Thermo Fisher Scientific, Rockford, IL) to Extraction Buffer A according to the manufacturer's instructions, and the clarified extract dialyzed against extraction buffer to remove the residual protease inhibitor reagents. This will be referred to as "dry seed protein extract".

2.4. Electrophoresis

SDS-PAGE under reducing conditions was done on either precast (4-20% w/v) Precise Protein gels (Fisher Scientific, Pittsburgh, PA), or gels cast with 12.5 % (w/v) total acrylamide monomer and 2.6% cross-linking. Electrophoresis was carried out in the mini-gel format using the Protean 3 apparatus (BioRad Laboratories, Inc., Hercules, CA). Gels were

1 stained overnight with 0.125% (w/v) Coomassie Blue G-250 in 40% (v/v) methanol, 10%
2 (w/v) acetic acid, and de-stained in 12.5% (v/v) 2-propanol, 10% (v/v) acetic acid.

4 **2.5. SDS-PAGE analysis of protein extracts**

5 Varying loads of protein extract were applied to SDS-PAGE gel lanes, resolved by
6 electrophoresis and stained with Coomassie Blue to determine the range of protein extract
7 concentration within which stain intensity for bands of interest was still linear. For analysis of
8 the main proteins in peanut seedling cotyledons harvested 1 DAI, stain intensity was linear
9 up to a load representing 2×10^{-4} cotyledons, i.e. a 1 in 10 dilution of the protein extracts
10 prepared as described in Section 2.3. Therefore, for SDS-PAGE analysis of seedling
11 cotyledon extracts, loads on gel lanes were normalized to this number of cotyledons.

13 **2.6. Proteolytic enzyme assays**

14 The autolysis assay mix consisted of 1 volume of extract, 2 volumes of McIlvaine
15 citrate/phosphate buffer of the desired pH (McIlvaine 1921) and 5 volumes of water.
16 Proteolysis reaction mixtures consisted of 0.5 μ l of extract of appropriate dilution incubated
17 with 6.5 μ l of dry seed protein extract (using an amount from the equivalent of 2×10^{-4}
18 cotyledons) in McIlvaine citrate/phosphate buffer of appropriate pH. Reaction mixtures were
19 incubated at 37°C and stopped after the desired time interval (or immediately for the zero-
20 hour control) by addition of an equal volume of Laemmli SDS-PAGE sample buffer to the
21 reaction mix, and frozen immediately. For assays carried out at pH 3 and 4, the final reaction
22 mix was titrated with a few microliters of 1 N NaOH until the bromophenol blue in the
23 Laemmli sample buffer turned from yellow back to blue. All the samples were then heated at
24 95°C for 10 min, and analyzed by SDS-PAGE (Sections 2.4 and 2.8).

26 **2.7. Assay of inhibition of proteolysis**

Protein extract (in an amount calculated to deliver 4×10^{-6} cotyledons per gel lane) was incubated with inhibitor in final concentrations of 4 mM AEBSF, 10 mM EDTA, 16 μ M E64 or 16 μ M pepstatin A in treatment buffer for 1 h on ice. To serve as solvent control for E64 and pepstatin A treatments, respectively, protein extracts were also treated with 0.1% (v/v) ethanol and 0.2% (v/v) methanol. Water served as solvent control for AEBSF and EDTA. For treatment with AEBSF, E64 and pepstatin A, reagents that are expected to work as a result of enzyme action, the treatment buffer was citrate-phosphate buffer (McIlvaine, 1921) at the pH optimum for activity. For treatment with EDTA, the treatment buffer was Extraction Buffer A (pH 8) to allow for chelation of divalent metal ions. The treated protein extracts were assayed for proteolytic activity as described in Section 2.6, and the percentage of inhibition calculated as described in Section 2.8.

2.8. Data analysis

Stained gels were scanned, and the band stain intensities quantified using Kodak 1D Image Analysis software (Eastman Kodak Co., Rochester, NY) or Quantity One (Bio-Rad Laboratories Inc.). Calculation of the extent of proteolysis of specific proteins was based on: A, the stain intensity of the band representing the protein of interest at time 0; and B, the stain intensity of the same band at time t, according to the formula: % proteolysis in time t = $100 (A - B) / A$.

The percentage of inhibition of the proteolytic enzyme activity was calculated based on: C, the % loss of band intensity of the protein substrate in a given time period when the enzyme was pre-treated with inhibitor; and D, the % loss of band intensity of the protein substrate in the same time period when the enzyme was pre-treated with solvent that the inhibitor was dissolved in, according to the formula: % inhibition = $100 (D - C) / D$.

When bands gain in intensity, the percent gain is calculated based on: E, the band intensity at time t; and F, the band intensity at time 0, according to the formula: % gain in

1 time $t = 100$ (F/E).

2 Expression of proteolytic activity in terms of the amount of protein substrate digested
3 per unit time is based on an estimation of the amount of the specific substrate protein in the
4 assay. This estimate was calculated from the protein content of the extract determined by
5 the Bradford method for determining protein concentration (P) and by determination of the
6 fraction of stain intensity of the band of interest relative to all other bands on SDS-PAGE gel
7 analysis (F). The proteolytic activity would then be: (P) (F) (% digested / unit time)

8

9 **2.9. Proteomic analysis**

10 In-gel trypsin digestion and LC-MS/MS were done as described previously (Pan et al.
11 2011), using the Uniprot database of curated protein sequences and the full database of
12 available sequences for peanut proteins.

13

14 **2.10. Confocal laser scanning fluorescence microscopy**

15 Transverse sections of cotyledons approximately 0.5 mm thick were cut by hand with
16 a razor blade, and incubated with 50 μ M acridine orange with and without 30 mM
17 methylamine as described previously (He et al. 2007). Confocal microscopy was performed
18 with a Zeiss LSM 510 confocal microscope with a Plan-Apochromat 5X objective (NA 0.16).
19 Excitation was at 488 nm using an argon laser, with fluorescence emission detected at 636-
20 668 nm, to visualize accumulation of acridine orange. The plane of focus was set at
21 approximately one-half the thickness of the slice.

22

23

24 **3. RESULTS**

25 **3.1. Proteolysis of storage proteins in peanut seedlings**

Protein extracts of peanut seeds and of cotyledons that were harvested from seedlings during the first six days after imbibition (DAI) of the seed with water, were analyzed by SDS-PAGE (Figure 1A). The two isoforms of Ara h 1 (P43237 and P43238 in the curated Uniprot database) in the dry seed (Day 0) appear as two closely spaced bands at 69 and 66 kD. These sizes are consistent with the published sequences, after accounting for the removal of a 25-residue signal peptide, further removal of an N-terminal propeptide (Wichers et al. 2004), and inclusion of 1 to 4 kD for the oligosaccharide moiety (Kolarich and Altmann 2000). The acidic and basic chains of the arachins are components of 11S globulins, another type of storage protein commonly found in legume seeds, and evolutionarily related to the 7S globulins. In addition to having multiple isoforms (Chassaing et al. 2009), there is heterogeneity due to additional proteolytic cleavage of some of the arachin acidic chains during seed development (Piersma et al. 2005), but not of the 42, 39 and 35 kD species (Figure 1). Also indicated on the gel are two isoforms of Ara h 2, the 2S albumins present in the seed. Decline in the intensity of bands representing Ara h 1 and the acidic chains of arachin are evident two days after water imbibition, coinciding with radicle emergence (Figure 1B). With loss of the storage proteins, new bands appear, with the main bands corresponding to 59, 52, 30-34, 25, 18, 16, 13, 11 and 10 kD (Figure 1A).

3.2. pH Dependence and mechanistic class of the protease activity directed toward storage globulins in cotyledons harvested 1 DAI

Extracts of peanut seedling cotyledons harvested one day after imbibition (1 DAI) of the seed with water were prepared in Extraction Buffer A. Autolysis was conducted for 20 h, varying pH from 3 to 8. The best condition for loss of Ara h 1 was at pH 3 (Figure 2A, unfilled symbols). The proteolytic activity at pH 3 was characterized for sensitivity to AEBSF, E-64, EDTA, and pepstatin A, reagents that inhibit serine, cysteine, metallo- and aspartic proteases, respectively (Table 1). Proteolysis was inhibited only by pepstatin A, indicating

1 that proteolysis was due to an aspartic protease. This is consistent with the preference of
2 aspartic proteases for very low pH, attributed in part to the condition required for activation
3 of the zymogen (Szecsi 1992). An apparent activation of Ara h 1 cleavage was observed in
4 the presence of EDTA. The effect that divalent calcium and magnesium ions have on the
5 equilibrium between aggregation and disaggregation of 7S globulins has been documented
6 in lupin (Santos et al. 2012). The effect of EDTA on Ara h 1 proteolysis may be due to a shift
7 in the equilibrium toward disaggregation of Ara h 1 by chelation of divalent cations by EDTA,
8 thus facilitating Ara h 1 proteolysis. Alternatively, the apparent activation by EDTA might be
9 due to inhibition of a metalloprotease that would otherwise have digested the aspartic
10 protease responsible for Ara h 1 proteolysis.

11 Furthermore, we observed during autolysis, proteolytic activity directed against the
12 acidic and basic chains of the arachins. We measured the level of proteolytic activity
13 directed toward the non-truncated acidic chains in the pH range from 3 to 8, and noted that
14 the activity was highest at pH 3 (Figure 3). The same was found for proteolytic activity
15 directed toward the basic chains. When tested with the inhibitor reagents listed in Table 1,
16 only pepstatin A treatment of the protein extract resulted in inhibition ($91.8 \pm 6.9\%$).

17 On the chance that cotyledons of seedlings 1 DAI also had cysteine proteases
18 directed at Ara h 1 and the arachin chains, that required activation by reducing agent, 0.7
19 mM DTT was added to the enzyme assay mix throughout the 20 hour incubation period
20 (Figures 2A, 3A and 3B, filled symbols). DTT addition had no effect throughout the pH range
21 examined. Additionally, protein extracts of cotyledons harvested 1DAI were prepared in
22 Extraction Buffer B, containing 10 mM 2-mercaptoethanol as reducing agent. When
23 examined for proteolytic activity directed toward Ara h 1 and the arachin chains at pH 3, 5
24 and 7, these extracts showed proteolysis at pH 3, but not at pH 5 or 7. Proteolysis at pH 3
25 was inhibited by pepstatin A, and was not affected by prior treatment of protein extract with
26 5 mM sodium iodoacetate, a reagent that typically inhibits cysteine proteases.

3.3. Proteolytic activity directed toward storage proteins in later seedling growth

The continued presence of the aspartic protease that breaks down Ara h 1 in seedling cotyledons past 1 DAI would support a role for the enzyme in post-germination storage protein breakdown. Thus, cotyledon extracts of seedlings harvested 3 DAI were examined for proteolytic activity directed against Ara h 1. As found in 1 DAI seedlings, Ara h 1 proteolysis by enzyme in 3 DAI extracts also exhibited maximum activity at pH 3 (Figure 2B). As at 1 DAI, the proteolytic activity was inhibited essentially only by treatment with pepstatin A; and apparent activation by EDTA was also observed (Table 1).

The level of proteolytic activity directed toward Ara h 1 at pH 3 in protein extracts prepared with Extraction Buffer B from the equivalent number of seedling cotyledons in mature seed and in cotyledons harvested at different days after seed imbibition remained relatively constant at $67 \pm 4\%$ of Ara h 1 lost in 20 h at 37°C (Figure 4). Similarly, the level of proteolytic activity directed toward the non-truncated arachin acidic chains remained relatively constant at $25 \pm 5\%$ of the arachin acidic chains lost in 20 h at 37°C (Figure 4). We estimate the level of proteolytic activity in protein extracted from the equivalent of 1 g of peanut to be 1.2 mg of Ara h 1 and 0.9 mg of the arachin acidic chains digested per hour under our assay conditions. It was difficult to carry out similar analysis with respect to the proteolytic activity directed toward the basic chains because at some time points, the degradation of the basic chains is obscured by the accumulation of intermediates of sizes similar to those of the basic chains, arising from the proteolysis of proteins of higher mol wt.

3.3. Proteolytic intermediates generated during *in vitro* proteolysis of Ara h 1

Dry seed protein extract (diluted 1 in 10) was incubated for 0.5, 1 and 18 h with extract from cotyledons of seedlings harvested 1 DAI (diluted 1 in 50). Incubation was in different pH conditions ranging from pH 2.6 to pH 6. The resulting proteins were resolved by

1 SDS-PAGE and visualized by Coomassie blue staining. A gel representing the reaction at
2 pH 2.6 is shown in Figure 5. Under these conditions, Ara h 1 and the acidic chains of
3 arachin are digested more rapidly than the arachin basic chains. This reflects the situation
4 seen *in vivo* (Figure 1). The rate of breakdown of Ara h 1 during the first hour of incubation
5 was most rapid at pH 2.6-3.0 (Figure 6). New bands appeared at 33 and 34 kD within the
6 first hour, with the same pH dependence for the gain in intensity of these bands as was
7 observed for loss of Ara h 1. During the subsequent 17 h, the 33/34 kD bands decreased in
8 intensity, accompanied by increase in intensities of bands at 18, 13, 11 and 10 kD; with all of
9 these changes being maximal also at pH 2.6-3.0 (Figure 6). When the reaction at pH 2.6
10 was carried out in the presence of inhibitors of aspartic, cysteine and serine proteases
11 (enzymes that are known to break down storage proteins optimally in the acidic pH range),
12 treatment of enzyme extract with sodium iodoacetate, E-64 and AEBSF had no effect on the
13 storage protein degradation. The loss of Ara h 1 as well as the gain in intensity of the 33/34
14 kD band, were inhibited only in the presence of pepstatin A.

15 The tryptic peptides in the 33/34 kD band that matched to sequences in the Ara h 1
16 isoforms had the same sequence coverage as the tryptic peptides in the band
17 corresponding to intact Ara h 1 (66 and 69 kD) at 0 h (Table 2 and Figure 7A), even though
18 33/34 kD is only half the size. Such a result suggests that cleavage occurred in the region
19 that links the two cupin domains of Ara h 1 to produce two different 33/34 kD fragments. The
20 33/34 kD fragments would represent the N- and C- terminal halves of Ara h 1, with each
21 fragment encompassing an intact cupin domain.

22 When the 18 kD band was analyzed, the tryptic peptides that matched to Ara h 1
23 corresponded to sequences near the N- and the C- termini. Analysis of the 10 kD band
24 yielded tryptic peptides that matched to the C-terminal portion of the first cupin domain of
25 Ara h 1 and a portion of the linker close to that domain; while analysis of the 11 kD band
26 yielded tryptic peptides that matched to a portion of the linker close to the second cupin

1 domain (Table 2, Figure 7A,B). From these results, combined with the observation of the
2 transient nature of the increase in intensity of the 33/34 kD fragments and the later
3 appearance of the 18, 11 and 10 kD bands (Figure 5), we deduce that these lower mol wt
4 fragments were produced by cleavage of each of the 33/34 kD fragments into two fragments
5 of unequal sizes, one (18 kD) nearly twice the size of the other (10/11 kD). The further
6 reduction in size (the sum of 18 and 10/11 kD being about 4-6 kD lower than 33/34 kD), may
7 be due to additional proteolysis in regions outside of the cupin domains, possibly the regions
8 N-terminal to the first cupin domain, C- terminal to the second cupin domain, or part of the
9 linker region. The size discrepancy may also arise from loss of sugar residues from the one
10 glycosylation site in the second cupin domain (Kolarich and Altmann 2000).

11 We noted that storage protein breakdown *in vivo* (Figure 1) generates more
12 intermediates than observed *in vitro* (Figure 5). For instance, in addition to the cluster of
13 transient intermediates in the 30-34 kD range, a transient 25 kD band was also prominent.
14 LC-MS/MS analysis of a trypsin digest of the band also yielded a sequence coverage similar
15 to that of intact Ara h 1 (Table 2) that could be explained also by cleavage of Ara h 1 in the
16 inter-cupin linker region with additional cleavage reactions accounting for the 8-9 kD lower
17 mass compared to the 33/34 kD fragments.

18 19 **3.4. Aspartic protease activity directed against Ara h 1 in developing peanut seeds**

20 The proteolytic cleavage, during seed development, of the monomeric precursors of
21 the arachins into acidic and basic chains, as well as further proteolysis of the acidic chains
22 of some arachin isoforms, are well documented (Piersma et al. 2005). This stands in
23 contrast to the absence of evidence from 2D-PAGE (Chassaigne et al. 2009) of proteolytic
24 cleavage of Ara h 1 in seeds, other than the N-terminal processing described by Wichers et
25 al (2004). We verified the absence of Ara h 1 cleavage in developing seeds that we
26 harvested from greenhouse-grown peanuts (Figure 8). The early initiation of Ara h 1

1 proteolysis (Figure 1) and the presence of aspartic protease activity in mature seed (Figure
2 4) indicated that the protease was already present in developing seed. This was verified by
3 assays that were conducted on extracts of developing seeds from pH 3 to 8 for Ara h 1
4 proteolytic activity (Figure 2C). Indeed, in the pH range tested, the proteolysis was most
5 rapid at pH 3, with no difference when the autolysis was conducted in the presence or
6 absence of DTT. As was found for seedling cotyledons, the proteolytic activity directed
7 against Ara h 1 at pH 3 was inhibited only by pepstatin A (Table 1). Some activation by
8 EDTA was also observed.

9 10 **3.5. pH condition in protein storage vacuoles of peanut cotyledons**

11 In past work on the proteolysis of the soybean 7S globulins where degradation is
12 initiated by an enzyme with a sharp pH optimum of 4 (He et al. 2007), we saw evidence of
13 control by pH. Acridine orange was seen to accumulate in the PSVs of soybean seedling
14 cotyledons but not in developing seeds at mid-maturation stage. Such dye accumulation in a
15 membrane-bound compartment was demonstrated to occur when there is at least a ten-fold
16 difference in proton concentration across the vacuolar membrane, the inside being more
17 acidic (Millot et al. 1997). In peanut, the optimum pH for the aspartic protease activity
18 directed toward Ara h 1 at pH 2.6-3 makes the requirement for acidic conditions even more
19 stringent. To test the hypothesis that the pH condition for Ara h 1 proteolysis is not met in
20 developing seeds, slices of freshly harvested developing seeds of peanut were examined for
21 accumulation of acridine orange in PSVs (light gray areas on the micrographs). Confocal
22 microscopy images show that although acridine orange accumulated in the cell wall region,
23 there was no fluorescence within the cells indicating that the dye did not accumulate in the
24 PSVs (Figure 9A). In contrast, we observe fluorescent PSVs within the cells of seedling
25 cotyledons harvested 3DAI (Figure 9B), indicating that the timing of PSV acidification is
26 subsequent to germination. To ensure that dye accumulation in the PSVs of the seedling

1 cotyledons is due to acidity and not to the dye binding to cellular structures, control
2 experiments were carried out on the 3DAI cotyledon slices. These were incubated with
3 acridine orange and methylamine, the latter being a membrane-permeable base that would
4 neutralize vacuolar acidic contents, preventing dye accumulation that was due to acidity.
5 This was indeed the result (Figure 9C),
6

7 **4. DISCUSSION**

8 We have demonstrated the mobilization of the main storage globulins in peanut
9 seedlings, the 7S conarachins and 11S arachins, to be due to an aspartic protease with a
10 pH optimum of 2.6 - 3. Focusing on the 7S globulin, Ara h 1, we have determined the region
11 of initial cleavage to be in the region between the two cupin domains, with subsequent
12 cleavage in an accessible loop within each of the two cupin domains. The aspartic protease
13 was present in active form during seed maturation, but is prevented from exerting its action
14 with one contributing factor being that acidification of the protein storage vacuoles occurs
15 only after germination.

16 The role of aspartic proteases in protein precursor processing during seed
17 maturation has been described in other plant species (Simoes and Faro 2004). Evidence of
18 involvement of aspartic proteases in seed storage protein breakdown after germination in
19 sunflower (*Helianthus annuus* L.), star-thistle (*Centaurea calcitrapa* L.), buckwheat
20 (*Fagopyrum esculentum* Moench), and mung bean (*Vigna radiata* L. Wilczek) (Park et al.
21 2000; Salvador et al. 2006; Timotijevic et al. 2003; Kulkarni and Rao 2009) have been
22 indirect, based on activities directed against heterologous protein substrates such as
23 hemoglobin and bovine serum albumin. A more direct demonstration has been the *in vitro*
24 digestion of the 7S globulin of cacao (*Theobroma cacao* L.) by an aspartic protease isolated
25 from cacao seeds (Voigt et al. 1994). Our report establishes a link between the aspartic
26 protease activity directed against the 7S globulin in peanut seedling cotyledons, and maps

1 out a pathway of Ara h 1 mobilization that parallels but is not identical to those delineated
2 earlier for the 7S globulins of other legume species.

3 Our suggested site for initiation of proteolysis in the linker region between the two
4 cupin domains was also a cleavage site for or early mobilization of 7S globulins of soybean
5 (Seo et al. 2001), common bean (Zakharov et al. 2004), pea (Gatehouse et al. 1983), and
6 lupin (Magni et al. 2012). The similarities in cleavage sites, however, do not extend to the
7 type of proteolytic enzyme that catalyzes the reactions, which were legumain-like or papain-
8 like cysteine proteases in soybean and common bean and a serine protease in the lupin.
9 Our work indicates that yet another class of proteolytic enzyme, an aspartic protease,
10 accomplishes this task in the peanut. It would appear that, as in other legume species
11 (Shutov et al. 2003), it is the location of the vulnerable sites on the otherwise protease-
12 resistant bicupin structure that dictates a pathway of initial 7S vicilin breakdown. Evidently,
13 different mechanistic classes of proteolytic enzymes can cleave at the linker region,
14 although there are enough constraints so that not every protease that is present can carry
15 out the reaction. For instance, although soybean protease C1 is able to process the N-
16 terminal region of the α and α' subunits of soybean β -conglycinin, it does not cut into the
17 linker region. This is thought to be due to a need for protease C1 to interact with nine
18 contiguous residues of the protein substrate (Boyd et al. 2002; Tan-Wilson et al. 2012), a
19 requirement that is not met perhaps because the linker region is flanked by the two cupin
20 structures. We have no comparable structural information for the peanut aspartic protease;
21 however, a model of the plant aspartic protease cardosin shows interaction with six
22 contiguous amino acid residues of a substrate (Frazão et al. 1999). In addition to a possibly
23 smaller region of interaction between protease and substrate, the crystal structure
24 (PDB:3S71_A) (Chruszcz et al. 2011) shows the region between the two cupin domains of
25 of the P43238 isoform of Ara h 1 to be 16-20 residues longer than the corresponding
26 regions in the soybean 7S globulins (Maruyama et al. 2004), allowing for greater accessibility

1 to the protease. Furthermore, the linker region in Ara h 1 contains 14 hydrophobic
2 (branched-chain) amino acid residues (Figure 5A), consistent with the preference of plant
3 aspartic proteases for cleavage at bonds involving hydrophobic amino acid residues
4 (Mazorra-Manzano et al. 2010).

5 The limited nature of the proteolysis of the 33/34 kD intermediates indicates that
6 even after the initial cleavage at the linker region, the cupin structures of Ara h 1 are not yet
7 subjected to extensive proteolysis. This suggests that only specific sites within the cupin
8 domains are accessible to the aspartic protease. Legume 7S vicilins are similar in having an
9 exposed loop in their C-terminal cupin domain, as demonstrated by cleavage of vicilin in
10 developing pea seeds and *in vitro* trypsin digestion of the soybean β -conglycinin α' subunit
11 (Shutov et al. 1996). On the P43238 isoform of Ara h 1, the location of that region from
12 residues 473-493 (Chruszcz et al. 2011) (Figure 5A) could generate an 18 kD fragment from
13 the C-terminal region of Ara h 1. Although composed mostly of hydrophilic residues, the loop
14 in Ara h 1 contains an E-G sequence that would be similar to an E-A cleavage site for a
15 wheat aspartic protease (Bleukx et al. 1998). A vulnerable site in the N-terminal cupin
16 domain of phaseolin was revealed by *in vivo* and *in vitro* digestion by the legumain-like
17 cysteine protease, LLP (Senyuk et al. 1998). The positions of the β -strands and loops on the
18 N-terminal cupin domain of the P43238 isoform of Ara h 1 do not exactly parallel those of
19 phaseolin (Chruszcz et al. 2011); however, there is a 6-residue loop starting at residue 293
20 with the sequence (V)NTPGQF(E) that includes a large hydrophobic residue at F298. The
21 sequence of the P43237 isoform is identical or very close to the sequence of the P43238
22 isoform at all of the above suggested regions of cleavage (Figure 5A). Furthermore, the
23 extended loop of the second cupin domain, where the only non-hydrophilic residue in
24 P43238 was a glycine, the P43237 isomer includes a hydrophobic tryptophan residue.

25 The situation where enzyme capable of digesting the 7S globulin is present in
26 developing seeds, but does not actually work, is similar to that in the soybean where the β -

1 conglycinin is not digested in developing seed despite the presence of protease C1, and
2 stands in contrast to the pea, where cleavage of the vicilin during seed development is
3 manifested as a set of lower mol wt bands on SDS-PAGE analysis of immature seeds
4 (Gatehouse et al. 1983). One explanation, based on our finding of PSV acidification
5 occurring only after germination, is that in developing peanut seeds the aspartic protease is
6 inactive because the low pH requirement is not met. We could observe the activity in our
7 assays only because we provided the low pH condition. There could be other possible
8 reasons. Although aspartic proteases have been localized in protein storage vacuoles in
9 cardoon (*Cynara cardunculus* L.) and thale cress (*Arabidopsis thaliana* L. Heynh.) (Pereira
10 et al. 2008; Mutlu et al. 1999), the localization of the peanut aspartic protease has not been
11 determined. Thus it may be possible that the protease in the developing peanut seed may
12 not be in contact with Ara h 1 until after germination. Alternatively, the aspartic protease in
13 developing seed may be controlled by an endogenous aspartic protease inhibitor as in mung
14 bean (Kulkarni and Rao 2009), where such inhibition of a hemoglobin-degrading aspartic
15 protease was lifted after germination.

16 We had noted earlier that information on how peanut seedlings degrade Ara h 1
17 might suggest ways to reduce allergenicity due to Ara h 1. The present work suggests that
18 imbibing raw peanuts in an acidic solution may result in cleavage of Ara h 1 that may lead to
19 faster and more extensive gastrointestinal digestion. Indeed, in a report exploring a link
20 between regional foods and the low incidence of peanut allergy in Korea, Kim et al. (2012)
21 found that a local practice of soaking raw peanuts overnight in vinegar (pH 2.3) or soaking in
22 a pH 1.0 solution of acetic acid resulted in almost complete loss of the Ara h 1 band in SDS-
23 PAGE. Using acetic acid solutions at higher pH had no effect. We note that the newly
24 appearing bands are at 34, 16, 11 and 10 kD, coinciding approximately in size to the
25 intermediates we observed in *in vitro* proteolysis of peanut seed proteins with 1DAI seedling
26 cotyledon extracts. Interestingly, Kim et al. (2012) observed that the 34 and 16 kD bands

1 were barely detected in immunoblots of IgE from pooled patients with peanut
2 hypersensitivity, and that the 11 and 10 kD bands were not detected. It is possible that
3 imbibition of the seeds in acidic solution accelerated the onset of Ara h 1 proteolysis, most
4 likely by acidifying the PSVs and activating the aspartic protease that we describe in this
5 report, which then digested the Ara h 1 to fragments that were no longer antigenic.

6

7 **ACKNOWLEDGEMENTS**

8 We thank Dr. Karin Sauer for assistance with the LC-MS/MS analysis. We also thank
9 the undergraduate students who helped in various stages of this project: Breanna Bridges,
10 Jianzhong Chen, Hau Chieng, Zemelia Miller, and Yiting Ou. Portions of this work were
11 supported by NSF DUE0737203 and by the SUNY Upstate Bridges to the Baccalaureate
12 program of the State University of New York at Binghamton.

REFERENCES

- Bleukx W., Brijs K., Torrekens S., Van Leuven F., Delcour J.A., 1998. Specificity of a wheat gluten aspartic proteinase. *Biochim Biophys Acta* 1387, 317-324.
doi:10.1016/S0167-4838(98)00146-0
- Boyd P.M., Barnaby N., Tan-Wilson A., Wilson K.A., 2002. Cleavage specificity of the subtilisin-like protease C1 from soybean. *Biochim Biophys Acta* 1596, 269-282.
doi:10.1016/S0167-4838(02)00228-5
- Cabanillas B., Pedrosa M.M., Rodriguez J., Muzquiz M., Maleki S.J., Cuadrado C., Burbano C., Crespo J.F., 2012. Influence of enzymatic hydrolysis on the allergenicity of roasted peanut protein extract. *Int Arch Allergy Immunol* 157, 41-50.
doi:10.1159/000324681
- Chassaing H., Tregoat V., Norgaard J.V., Maleki S.J., van Hengel A.J., 2009. Resolution and identification of major peanut allergens using a combination of fluorescence two-dimensional differential gel electrophoresis, Western blotting and Q-TOF mass spectrometry. *J Proteomics* 72, 511-526. doi:10.1016/j.jprot.2009.02.002
- Chruszcz M., Maleki S.J., Majorek K.A., Demas M., Bublin M., Solberg R., Hurlburt B.K., Ruan S., Mattison C.P., Breiteneder H., Minor W., 2011. Structural and immunologic characterization of Ara h 1, a major peanut allergen. *J Biol Chem* 286, 39318-39327.
doi:10.1074/jbc.M111.270132
- Frazão C., Bento I., Costa J., Soares C.M., Veríssimo P., Faro C., Pires E., Cooper J., Carrondo M.A., 1999. Crystal structure of cardosin A, a glycosylated and Arg-Gly-Asp-containing aspartic proteinase from the flowers of *Cynara cardunculus* L. *J Biol Chem* 274, 27694-27701. doi:10.1074/jbc.274.39.27694
- Gatehouse J.A., Lycett G.W., Delauney A.J., Croy R.R., 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

- He F., Huang F., Wilson K.A., Tan-Wilson A., 2007. Protein storage vacuole acidification as a control of storage protein mobilization in soybeans. *J Exp Bot* 58, 1059-1070. doi:10.1093/jxb/erl267
- Kim J., Lee J., Seo W.H., Han Y., Ahn K., Lee S.I., 2012. Changes in major peanut allergens under different pH conditions. *Allergy Asthma Immunol Res* 4, 157-160. doi:10.4168/aaair.2012.4.3.157
- Kolarich D., Altmann F., 2000. N-Glycan analysis by matrix-assisted laser desorption/ionization mass spectrometry of electrophoretically separated nonmammalian proteins: application to peanut allergen Ara h 1 and olive pollen allergen Ole e 1. *Anal Biochem* 285, 64-75. doi:10.1006/abio.2000.4737
- Kulkarni A., Rao M., 2009. Differential elicitation of an aspartic protease inhibitor: Regulation of endogenous protease and initial events in germination in seeds of *Vigna radiata*. *Peptides* 30, 2118-2126. doi:10.1016/j.peptides.2009.08.024
- Magni C., Sessa F., Tedeschi G., Negri A., Scarafoni A., Consonni A., Duranti M., 2012. Identification in lupin seed of a serine-endopeptidase activity cleaving between twin arginine pairs and causing limited proteolysis of seed storage proteins. *Mol Plant* 5, 1011-1019. doi:10.1093/mp/ssr116
- Mahaboob Basha S.M., Cherry J.P., 1978. Proteolytic enzyme activity and storage protein degradation in cotyledons of germinating peanut (*Arachis hypogaea* L.) seeds. *J Agric Food Chem* 26, 229-234. doi: 10.1021/jf60215a045
- Maruyama Y., Maruyama N., Mikami B., Utsumi S., 2004. Structure of the core region of the soybean beta-conglycinin alpha' subunit. *Acta Crystallogr D Biol Crystallogr*, D60, 289-297. doi:10.1107/S0907444903027367
- Mazorra-Manzano M.A., Tanaka T., Dee D.R., Yada R.Y., 2010. Structure-function characterization of the recombinant aspartic proteinase A1 from *Arabidopsis thaliana*. *Phytochemistry* 71, 515-523. doi:10.1016/j.phytochem.2009.12.005

- McIlvaine T.C., 1921. A buffer solution for colorimetric comparison. *J Biol Chem* 49, 183-186
- Millot C., Millot J.M., Morjani H., Desplaces A., Manfait M., 1997. Characterization of acidic vesicles in multidrug-resistant and sensitive cancer cells by acridine orange staining and confocal microspectrofluorometry. *J Histochem Cytochem* 45, 1255-1264.
doi: 10.1177/002215549704500909
- Moreno F.J., 2007. Gastrointestinal digestion of food allergens: Effect on their allergenicity. *Biomed Pharmacother* 61, 50-60. doi:10.1016/j.biopha.2006.10.005
- Müntz K., Belozersky M.A., Dunaevsky Y.E., Schlereth A., Tiedemann J., 2001. Stored proteinases and the initiation of storage protein mobilization in seeds during germination and seedling growth. *J Exp Bot* 52, 1741-1752.
doi:10.1093/jexbot/52.362.1741
- Mutlu A., Chen X., Reddy S.M., Gal S., 1999. The aspartic proteinase is expressed in *Arabidopsis thaliana* seeds and localized in the protein bodies. *Seed Sci Res* 9, 75-84. doi:10.1016/S0031-9422(97)00834-0
- Pan D., Hill A.P., Kashou A., Wilson K.A., Tan-Wilson A., 2011. Electrophoretic transfer protein zymography. *Anal Biochem* 411, 277-283. doi:10.1016/j.ab.2011.01.015
- Park H., Yamanaka N., Mikkonen A., Kusakabe I., Kobayashi H., 2000. Purification and characterization of aspartic proteinase from sunflower seeds. *Biosci, Biotechnol, Biochem* 64, 931-939. doi:10.1271/bbb.64.931
- Pereira C.S., da Costa D.S., Pereira S., de Moura Nogueira F., Albuquerque P.M., Teixeira J., Faro C., Pissarra J., 2008. Cardosins in postembryonic development of cardoon: towards an elucidation of the biological function of plant aspartic proteinases. *Protoplasma* 232, 203-213. doi:10.1007/s00709-008-0288-9
- Piersma S.R., Gaspari M., Hefle S.L., Koppelman S.J., 2005. Proteolytic processing of the peanut allergen Ara h 3. *Mol Nutr Food Res* 49, 744-755.
doi:10.1002/mnfr.200500020

- Qi X., Chen R., Wilson K.A., Tan-Wilson A.L., 1994. Characterization of a soybean beta-conglycinin-degrading protease cleavage site. *Plant Physiol* 104, 127-133.
doi: 10.1104/pp.104.1.127.
- Salvador S.M., Novo C., Domingos A., 2006. Evaluation of the presence of aspartic proteases from *Centaurea calcitrapa* during seed germination. *Enzyme Microb Technol* 38, 893-898. doi:10.1016/j.enzmictec.2005.06.025
- Santos C.N., Alves M.M., Bento I.T., Ferreira R.B., 2012. Missing pieces in protein deposition and mobilization inside legume seed storage vacuoles: calcium and magnesium ions. *Seed Science Research* 22, 249-258.
doi:10.1017/S0960258512000177
- Senyuk V., Rotari V., Becker C., Zakharov A., Horstmann C., Müntz K., Vaintraub I., 1998. Does an asparaginyln-specific cysteine endopeptidase trigger phaseolin degradation in cotyledons of kidney bean seedlings? *Eur J Biochem* 258, 546-558.
doi:10.1016/S0167-4838(00)00277-6
- Seo S., Tan-Wilson A., Wilson K.A., 2001. Protease C2, a cysteine endopeptidase involved in the continuing mobilization of soybean beta-conglycinin seed proteins. *Biochim Biophys Acta* 1545, 192-206. doi:10.1016/S0167-4838(00)00277-6
- Shutov A.D., Baumlein H., Blattner F.R., Müntz K., 2003. Storage and mobilization as antagonistic functional constraints on seed storage globulin evolution. *J Exp Bot* 54, 1645-1654. doi:10.1093/jxb/erg165
- Shutov A.D., Kakhovskaya I.A., Bastrygina A.S., Bulmaga V.P., Horstmann C., Müntz K., 1996. Limited proteolysis of beta-conglycinin and glycinin, the 7S and 11S storage globulins from soybean [*Glycine max* (L.) Merr.]. Structural and evolutionary implications. *Eur J Biochem* 241, 221-228. DOI: 10.1111/j.1432-1033.1996.0221t.x

- Simoes I., Faro C., 2004. Structure and function of plant aspartic proteinases. Eur J Biochem 271, 2067-2075. doi:10.1111/j.1432-1033.2004.04136.xSzecsi P.B., 1992. The aspartic proteases. Scand J Clin Lab Invest Suppl 210, 5-22
- Tan-Wilson A., Bandak B., Prabu-Jeyabalan M., 2012. The PA domain is crucial for determining optimum substrate length for soybean protease C1: structure and kinetics correlate with molecular function. Plant Physiol Biochem 53, 27-32. doi:10.1016/j.plaphy.2012.01.005
- Tan-Wilson A.L., Liu X., Chen R., Qi X., Wilson K.A., 1996. An acidic amino acid-specific protease from germinating soybeans. Phytochemistry 42, 313-319. doi:10.1016/0031-9422(95)00896-9
- Tan-Wilson A.L., Wilson K.A., 2012. Mobilization of seed protein reserves. Physiol Plant 145, 140-153. doi:10.1111/j.1399-3054.2011.01535.x
- Timotijevic G.S., Radovic S.R., Maksimovic V.R., 2003. Characterization of an aspartic proteinase activity in buckwheat (*Fagopyrum esculentum* Moench) seeds. J Agric Food Chem 51, 2100-2104. doi: 10.1021/jf026043b
- Voigt J., Biehl B., Voigt G., Heinrichs H., 1994. Cocoa-specific aroma precursors are generated by proteolytic digestion of the vicilin-like globulin of cocoa seeds. Food Chem 50, 177-184. [http://dx.doi.org/10.1016/0308-8146\(94\)90117-1](http://dx.doi.org/10.1016/0308-8146(94)90117-1)
- Wichers H.J., De Beijer T., Savelkoul H.F., Van Amerongen A., 2004. The major peanut allergen Ara h 1 and its cleaved-off N-terminal peptide; possible implications for peanut allergen detection. J Agric Food Chem 52, 4903-4907. doi:10.1021/jf049697o
- Zakharov A., Carchilan M., Stepurina T., Rotari V., 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. doi:10.1093/jxb/erh247

FIGURE CAPTIONS

Figure 1. Germination and early seedling growth of peanut. A. Changes in the protein complement of cotyledons collected from seedlings at the specified days after imbibition (DAI). Samples of extracts from the equivalent of 2×10^{-4} cotyledons were resolved by SDS-PAGE (4-20% (w/v) total acrylamide monomer, 2.6% cross-linking). Proteins in the dry seed are identified to the left. Newly appearing polypeptides are indicated by arrowheads ($\blacktriangleleft$), with molecular masses indicated to the right of the gel. B. Timing of radicle emergence with respect to DAI.

Figure 2. The pH dependence of the loss of Ara h 1. The percent loss was determined after 20 h of incubation of cotyledon extracts in the presence (filled symbols) and absence (unfilled symbols) of 0.7 mM DTT. The cotyledons were from A: seedlings harvested 1DAI; B: seedlings harvested 3 DAI; and C: immature seeds harvested at mid-maturation (average 2.0 cm long, 860 mg fresh weight compared to 2.3 cm and 1550 mg fresh weight just before desiccation).

Figure 3. The pH dependence of the loss of the arachin acidic and basic chains. Percentages of loss were determined after 20 h of incubation of protein extracts of cotyledons harvested at 1DAI in the presence (filled symbols) and absence (unfilled symbols) of 0.7 mM DTT. Analysis was based on the combined intensities of the non-truncated 35, 39 and 42 kD acidic chains (squares) and separately of the 21 kD basic chains (triangles) of the arachins.

Figure 4. The time course of proteolytic activity directed toward peanut storage proteins. The percentage of Ara h 1 and the non-truncated acidic chains of arachin (42, 39 and 35 kD in Figure 1) lost in 20 h were measured at pH 3 and 37°C. An equivalent number of cotyledons were used for each time point.

Figure 5. *In vitro* proteolysis of peanut seed proteins. Dry seed protein extract from the equivalent of 2×10^{-4} cotyledons (treated with protease inhibitor reagents) was incubated with protein extract from the equivalent of 4×10^{-6} cotyledons (untreated to serve as enzyme source). Cotyledons were from seedlings harvested 1DAI. Incubation was at pH 2.6 and 37°C for 0.5, 1 and 18 h. The resulting proteins were resolved by SDS-PAGE (12.5% (w/v) total acrylamide monomer, 2.6% cross-linking). Newly appearing bands are indicated by leftward arrows; corresponding masses are shown to the right of the gel.

Figure 6. The effect of pH on the *in vitro* cleavage of Ara h 1 and accompanying changes in the intensities of bands found to have Ara h 1 fragments. Reactions were carried out as in Figure 5, and analyzed by SDS-PAGE and densitometry. A. Changes in Ara h 1 and 33/34 kD band intensities; B. Changes in 18, 11 and 10 kD band intensities.

Figure 7. Sequence of Ara h 1 isoforms showing identified tryptic peptides and proposed origins of proteolytic intermediates of Ara h 1. A: Identified tryptic peptides (bold font) mapped to amino acid sequences of Ara h 1 isoforms. Regions are marked (|→ ←|) in accordance with PDB: 3SMH_A (G, glycosylation site). B. Schematic diagram showing proposed origins of the 33/34, 18, 11, and 10 kD intermediates relative to Ara h 1 domains.

Figure 8. Analysis of the protein complement in developing peanut seeds by SDS-PAGE. Proteins from maturing seeds with average lengths of 1.3, 1.8, 2.0, and 2.3 cm (with

average fresh weights of 157, 769, 859, and 1554 mg/seed, respectively) were resolved on SDS-PAGE (4-20% (w/v) total acrylamide monomer, 2.6% cross-linking). Each lane was loaded with protein from the equivalent of 2.5×10^{-3} cotyledons.

Figure 9. Confocal microscopy images of seed and cotyledon slices after incubation with acridine orange. A: Developing seed, B: 3DAI seedling cotyledons, C: 3DAI seedling cotyledons incubated with methylamine in addition to acridine orange. Excitation wavelength 488 nm, emission wavelength 636-688 nm. Acridine orange accumulation is indicated by the bright areas.

CONTRIBUTION

Karl A. Wilson - supervised preliminary work of students, performed many of the assays, the confocal microscopy and the in-gel trypsin digestion for LC-MS/MS analyses, and collaborated on the draft and revisions of the manuscript.

Anna Tan-Wilson - conceived the project, supervised preliminary work of students, contributed to the analysis of gel images, performed the mass spectral data analysis and interpretation, wrote the initial draft and collaborated on revisions of the manuscript.

Table 1. Inhibition of the proteolytic activities directed toward Ara h 1 at pH 3			
Reagent	Percent Inhibition ^a relative to solvent controls ^b		
	Seedling cotyledons		Developing seed
	1 DAI	3 DAI	2 cm length
Pepstatin A, 16 µM	67.9 ± 6.1	65.3 ± 17.7	64.3 ± 11.9
E-64, 16 µM	3.4 ± 4.4	12.6 ± 8.7	0.0 ± 2.5
AEBSF, 4 mM	5.2 ± 7.3	4.6 ± 11.5	4.2 ± 1.9
EDTA, 10 mM	-18.7 ± 3.2 ^c	-12.5 ± 5.8	-6.7 ± 2.9

^a Mean of 3 biological replicates ± standard deviation. ^b Solvent controls: 0.2% (v/v) methanol for pepstatin A, water for AEBSF and EDTA, 0.1% (v/v) ethanol for E-64 ^c Negative value indicates activation of degradation relative to the control.

Table 2. Tryptic peptides identified in bands that appear during Ara h 1 proteolysis

Structural Feature ^a	Ara h 1 Isoform		Tryptic peptide sequences matched to Ara h 1 ^b	Peptides identified by LC-MS/MS						
	P43237	P43238		Band (kD)	Ara h 1	34	18	11	10	25
N-terminal region	✓	✓	94 GRQPGDYDDDR				•			
	✓	✓	124 EREEDWR				•			
	✓		147 EGEQEWGTPGSEVR			•	•			•
		✓	153 EGEQEWGTPGSHVR			•	•			
	✓	✓	172 NNPFYFPSR	•	•	•				•
N-terminal cupin	✓	✓	205 QFQNLQNHR	•	•	•				•
		✓	214 IVQIEAK	•	•	•				
			214 IVQIEAKPNTLVPK	•	•	•				
			229 HADADNILVIQQGQATVTVANGNNR							
	✓	✓	229 HADADNILVIQQGQATVTVANGNNRK	•	•					•
	✓	✓	255 SFNLDEGHALR	•	•				•	
	✓	✓	266 IPSGFISYILNR	•						
	✓	288 ISMPVNTPGQFEDFFPASSR	•	•				•		
Linker	✓	✓	308 DQSSYLQGFSR			•			•	•
	✓	✓	319 NTLEAAAFNAEFNEIR	•	•	•				
	✓	✓	334 RVLLEENAGGEQEER	•	•			•		•
			335 VLLEENAGGEQEER							
		✓	357 SSENNEGVIVK	•				•		
	✓		380 GSEEEEDITNPINLR	•	•			•		•
		✓	387 GSEEEGDITNPINLR							
		386 KGSEEEGDITNPINLR	•	•						
C-terminal cupin	✓		394 DGEPDLSNNFGR							
		✓	402 EGEPDLSNNFGK	•	•					•
	✓	✓	413 LFEVKPDKK							•
	✓	✓	461 GTGNLELVAVR	•	•	•				•
	✓	✓	541 IFLAGDKDNVIDQIEK			•	•			
			548 DNVIDQIEK	•	•	•				•
✓	✓	560 DLAFPGSGEQVEK	•	•	•					

^aBased on assignments in PDB: 3SMH A

^bWhen tryptic peptides match to only one isoform of Ara h 1, sequence numbering is for the specific isoform. When the match is to both isoforms, the number given is that of the P43238 isoform. For the P43237 and P43238 isoforms, the N-termini begin at residues 79 and 86, while the C-termini are at residues 614 and 626, respectively.

Figure 1
Wilson & Tan-Wilson

Wilson and Tan-Wilson
Figure 2

Figure 3
Wilson and Tan-Wilson

Figure 4
Wilson and Tan-Wilson

Figure 5
Wilson & Tan-Wilson

Figure 6
Wilson and Tan-Wilson

A

signal peptide ← → propeptide				
P43237	1	MRGRVSPMLLLGILVLASVSATQAKS-PY-RKTENPCAQRCLQSCQQEPDDLKQKACES	58	
P43238	1	MRGRVSPMLLLGILVLASVSATHAKSSPYQKKTENPCAQRCLQSCQQEPDDLKQKACES	60	
→ N-terminal region				
P43237	59	RCTKLEYDPRCVYD----TGATNQRHPPGERTR GRQPGDYDDDRR QPRREEGGRWGPAEP	114	
P43238	61	RCTKLEYDPRCVYDPRGHTGTTNQSRPPGERTR GRQPGDYDDDRR QPRREEGGRWGPAEP	120	
→cupin				
P43237	115	RER EREEDWR QPREDWRRPSHQQPRKIRPEGREGE QEWGTPGSEVREETS SRNNPFYF PSR	174	
P43238	121	RER EREEDWR QPREDWRRPSHQQPRKIRPEGREGE QEWGTPGSHVREETS SRNNPFYF PSR	180	
P43237	175	RFSTRYGNQNGRIRVLQRFQDQSR QFQNLQNHRI VQIEARPNTLVLPKHADADN ILV IQQ	234	
P43238	181	RFSTRYGNQNGRIRVLQRFQDQSR QFQNLQNHRI VQIEAKPNTLVLPKHADADN ILV IQQ	240	
loop				
P43237	235	GQATVT VANGNNRKS FN LDEGHALRIPSGFISYILNRHDNQNLRVAK ISMPVNT PGQFED	294	
P43238	241	GQATVT VANGNNRKS FN LDEGHALRIPSGFISYILNRHDNQNLRVAK ISMPVNT PGQFED	300	
cupin ← → linker				
P43237	295	FFPASSRDQSSYLQGF SRNTLEAA FN A EFNE IRRVLLEENAGGE QEER GQRRRSTRSS-D	353	
P43238	301	FFPASSRDQSSYLQGF SRNTLEAA FN A EFNE IRRVLLEENAGGE QEER GQRRWSTRSS EN	360	
linker ← → cupin				
P43237	354	NEGVIVKVSKEHVQELTKHAKSVSK KGSEE-EDITNP INLRDGE PDLSNNFGR LFEV KPD	412	
P43238	361	NEGVIVKVSKEHVEELTKHAKSVSKKGSEEEG DTNPINLREGE PDLSNNFG KLFEV KPD	420	
→				
P43237	413	KKNPQLQDLDMMLTCVEIKEGALMLPHFNSKAMVIVVVNKGTGNLEL VAVRKEQQQRGR	472	
P43238	421	KKNPQLQDLDMMLTCVEIKEGALMLPHFNSKAMVIVVVNKGTGNLEL VAVRKEQQQRGR	480	
extended loop ← G				
P43237	473	EQWEEEEEEDEEEEGSNREVVRYTARLKEGDVFIMPA AHPVAINASSELHLLG FGINAEN	532	
P43238	481	EEE---EDEDEEEEGSNREVVRYTARLKEGDVFIMPA AHPVAINASSELHLLG FGINAEN	537	
cupin ← → C-terminal region				
P43237	533	NHR IFLAGDKDNVIDQIEKQAKDLAF PGSG EQVEK L IKNQRESHFVSAR PQS QSP -----	587	
P43238	538	NHR IFLAGDKDNVIDQIEKQAKDLAF PGSG EQVEK L IKNQKESHFVSAR PQS QSP SSP	597	
P43237	588	--SSPEKEDQEEENQGGKGPLLSILKAFN	614	
P43238	598	EKESPEKEDQEEENQGGKGPLLSILKAFN	626	

B

Figure 7
Wilson & Tan-Wilson

Figure 8
Wilson & Tan-Wilson

Figure 9
Wilson and Tan-Wilson

Post-germination proteolysis