



ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

PLANT
SCIENCE

Plant Science xxx (2006) xxx–xxx

www.elsevier.com/locate/plantsci

Cloning and characterization of a novel peanut allergen Ara h 3 isoform displaying potentially decreased allergenicity

I.-H. Kang¹, M. Gallo^{*}

Agronomy Department, Plant Molecular and Cellular Biology Program, The Genetics Institute,
University of Florida, Gainesville, FL 32611, United States

Received 25 July 2006; received in revised form 21 September 2006; accepted 22 September 2006

Abstract

Peanut allergy is one of the most serious food allergies. The major peanut allergens are designated Ara h 1, Ara h 2 and Ara h 3. A previously unidentified *ara h3*-related cDNA, *ara h3-im*, was isolated and its corresponding protein product confirmed in peanut seed by N-terminal sequencing. Ara h3-im is distinct from other related allergens in having a novel N-terminal sequence, lacking two potential N-glycosylation sites, missing most of the fourth IgE-binding epitope found in Ara h 3, and having a different predicted three-dimensional structure. Immunoblotting showed that the 36 kDa Ara h 3-im polypeptide was not recognized by IgE from the sera of peanut sensitive patients that recognized other major peanut allergen polypeptides. *ara h 3-im* is a new peanut arachin gene that encodes an expressed protein in peanut seed, Ara h 3-im, that shares features with other 11S globulin genes. Initial results indicate that Ara h 3-im has potentially lower allergenic properties than previously characterized peanut allergens which may aid in the production of a hypoallergenic peanut.

© 2006 Published by Elsevier Ireland Ltd.

Keywords: Ara h 1; Ara h 2; Ara h 3; Epitopes; Peanut allergens; Seed storage proteins

1. Introduction

Peanut (*Arachis hypogaea* L.) allergy is one of the most serious food allergies, eliciting a reaction through immunoglobulin E (IgE). Approximately, 1.5 million people are allergic to peanuts and peanut allergy causes as many as 100 deaths each year in the U.S. Allergies to peanut can persist throughout the lifetime of an individual [1,2]. Additionally, the number of children having an allergic reaction to peanut has nearly doubled during the last decade in the U.S. [3,4].

Most peanut allergens are seed storage proteins. Although about 20 peanut allergens have been reported, Ara h 1, Ara h 2 and Ara h 3 are classified as important major allergens which are recognized by more than 50% of peanut allergic patients [5]. Ara h 1 and Ara h 2 are recognized by 70–90% of patients with peanut allergy [6,7], and Ara h 3 is recognized by serum

IgE from approximately 44 to 54% of different patient populations with a history of peanut sensitivity [8,9].

Ara h 3 was originally identified as a 14 kDa glycinin whose N-terminal sequence was determined by Edmann degradation after separation by SDS-PAGE [7,10]. An *ara h 3* gene was cloned using degenerate cDNA probes whose sequences were determined based on N-terminal sequence information. This gene was expressed as a 60 kDa IgE-binding protein containing both acidic and basic chains [8]. Independently, another peanut allergen, Ara h 4, was described [9]. However, sequence comparisons between *ara h 3* and *ara h 4* cDNAs revealed high sequence homology, and now *ara h 3* and *ara h 4* are considered to be essentially the same gene.

The nucleotide sequence of *ara h 3* encodes 507 amino acid residues and the calculated size of the deduced protein is ~60 kDa. It is a legumin-like seed storage protein that has high sequence homology to glycinin, which is the major 11S globulin seed storage protein in soybean [8]. The 11S globulins are common storage proteins in legumes that function as a source of nitrogen for the developing plant and are initially synthesized as 60 kDa precursor globulins. Database searches for sequence similarity revealed that there are several *ara h*

* Corresponding author. Tel.: +1 352 392 1823; fax: +1 352 392 7248.

E-mail address: mamea@ifas.ufl.edu (M. Gallo).

¹ Present address: Department of Biology, University of Utah, Salt Lake City, UT 84112, United States.

3-related genes including *ara h 4*, *ara h 3/ara h 4*, and *gly1* in peanut. However, a 14 kDa protein was originally identified as Ara h 3 [7,10]. Therefore, this suggests that the 14 kDa protein is a processed form of Ara h 3 following translation of the expected whole protein. In fact, it was reported that there are several post-translational processing sites such as Asn-Gln, Asn-Gly and cysteine residues which can form disulfide bonds in Ara h 3 [8]. Recently, protein extracts from peanut seed run on SDS-PAGE revealed a series of Ara h 3 polypeptides ranging in size from 14 to 45 kDa [11]. Based on their N-terminal sequences, the observed set of polypeptides is the result of post-translational proteolysis [7,11,12,13].

Linear IgE-binding epitope analysis of a recombinant Ara h 3 revealed that four IgE-binding epitopes are present in the protein and mutational analysis of these epitopes revealed that single changes at particular amino acid residues resulted in a significant decrease in IgE-binding [8]. An interesting finding was that all four critical IgE-binding epitopes reside within the 40 kDa acidic subunit [8]. Previous studies also demonstrated that the main IgE reactivity is directed toward acidic subunits, whereas basic subunits contain only a minority of IgE epitopes [11]. However, a recent study showed that a basic subunit of Ara h 3 contained dominant immunoreactivity [14].

In this study, we report the cloning of a novel *ara h3*-related cDNA, *ara h 3-im* (GenBank accession no. AY618460), along with identification and characterization of its corresponding protein in peanut seed. Studies of this new peanut allergen, Ara h 3-im, indicate that it may have lower allergenicity compared to other known Ara h 3 proteins.

2. Materials and methods

2.1. Southern blot hybridization

Genomic DNA was extracted from peanut leaf tissue (*Arachis hypogaea* L. cv. Georgia Green) by a modification of the CTAB method [15]. Twenty micrograms of genomic DNA was digested with *EcoRI*, *HindIII*, *DraI* and *BamHI* (Promega, Madison, WI). After staining with ethidium bromide (0.5 µg/ml) and destaining, gels were transferred onto Hybond-N membranes (Amersham-Pharmacia Biotech, Inc., Piscataway, NJ) by capillary transfer with 6 X SSC after depurination (0.2N HCl), denaturation (1.5 M NaCl, 0.5N NaOH), and neutralization (1 M Tris (pH 7.4), 1.5 M NaCl) for 15, 30, and 45 min, respectively. After overnight transfer, the membranes were UV cross-linked at 120 J/cm² for 5 min (SpectroLinker XL-10000, Spectronics Co., Westbury, NY).

A partial cDNA fragment of *ara h3* (576 nucleotides) was generated from a λgt11 cDNA library constructed with RNA from developing peanut seeds of F78-1339 (Dr. A. Abbot, Clemson University) by PCR amplification with gene-specific primers (forward: 5'GTGCAAACCTAAGAGGCGAG3', reverse: 5'CCTTGAGTCTGTGTTGAATGC3') for use as a probe. PCR mixture consisted of 2 µl of concentrated cDNA library as template, 1 µM of each primer, 0.2 mM dNTP, 1× ExTaq Buffer and 5 U of ExTaq DNA polymerase (Takara, Japan) in a total volume of 30 µl. After an initial denaturation

cycle at 94 °C for 2 min, there were 30 cycles consisting of a 30 s denaturation step at 94 °C followed by annealing at 60 °C for 30 s and elongation at 72 °C for 1 min and this was performed using a GeneAmp PCR System 2400 (Perkin-Elmer Corp., Norwalk, Connecticut, USA). The ³²P-labeled probe for *ara h 3* was synthesized using the Prime-a Gene[®] Labeling System (Promega, Madison, WI) as suggested by the manufacturer.

Blots were prehybridized for 1 h at 65 °C in sodium phosphate buffer (Na₂HPO₄ and NaH₂PO₄, pH 7.2) containing 7% SDS and denatured salmon sperm DNA (100 µg/ml). Hybridization was in the same buffer for 16–18 h at 65 °C. Blots were washed once for 15 min at 65 °C in 20 mM sodium phosphate buffer (pH 7.2) containing 5% SDS, and then twice for 15 min at 65 °C with 20 mM sodium phosphate buffer (pH 7.2) plus 1% SDS. Autoradiographic images were obtained after exposure to X-ray film, X-OmatTM XFS-1 (Eastman Kodak Co., Rochester, NY) using intensifying screens for 3–10 d at –70 °C.

2.2. cDNA library screening and cloning of *ara h 3-im*

The same ³²P-labeled 576 bp *ara h 3* PCR product described above was used to screen the seed-derived cDNA library by plaque hybridization. Positive signals from duplicated membranes were compared to the original plates. Once the signals were confirmed, positive plaques were removed from the plate and diffused in 1 ml SM buffer (NaCl, MgSO₄, 1 M Tris–HCl; pH 7.5) with one drop of chloroform and stored for 1–2 h at 4 °C. The selected positive plaques were screened two additional times. After the third screening, 15 positive plaques were selected for further analysis. Instead of in vivo excision procedures, the target DNA from the 15 positive plaques were amplified by PCR reactions with λgt11 forward and reverse primers (New England BioLabs, Beverly, MA). The PCR products were cloned into pCR2.1 (Invitrogen, Carlsbad, CA) and sequenced as described below.

2.3. Sequence and structural analyses

The putative *ara h 3* cDNAs were sequenced in both directions (Interdisciplinary Center for Biotechnology Research, DNA Sequencing Core Lab, University of Florida, Gainesville, FL). DNA sequences were analyzed to find homologous sequences using the standard nucleotide–nucleotide BLAST (blastn) (Basic Local Alignment Search Tool on NCBI web site, <http://www.ncbi.nlm.nih.gov/>) search. Translate (<http://au.expasy.org/tools/dna.html>) on an ExPASy Molecular Biology server (<http://au.expasy.org/tools/>) at the Swiss Institute of Bioinformatics was used to translate the nucleotide sequences to protein sequences. Nucleotide and amino acid sequence alignments were carried out using T-COFFEE (<http://www.ch.embnet.org/software/TCoffee.html>) [16] and Pairwise Sequence Alignment Programs (<http://genome.cs.mtu.edu/sas.html>) with the gap program [17]. The prediction of putative post-translational modification sites was performed by predic-

tion tools such as NetNGlyc (<http://www.cbs.dtu.dk/services/NetNGlyc/>) found on the ExPASy Molecular Biology server. The amino acid sequences for Ara h 3-im and Ara h 3 were submitted for modeling of monomeric structures to SWISS-MODEL (<http://www.expasy.org/swissmod/SWISS-MODEL.html>) [18] using the soybean glycinin G1 structure (PDB ID:1FXZ) [19] as a template.

2.4. N-terminal sequencing and analysis

Peanut proteins were extracted from seeds by a modification of the method by Koppelman et al. [5]. After separation of peanut protein extracts by SDS-PAGE, the separated proteins were electrophoretically transferred to PVDF membranes (BioRad, Hercules, CA) generally as described by Towbin et al. [20]. After Coomassie blue staining, the protein bands were cut from a PVDF-blot membrane and their N-termini were sequenced (Interdisciplinary Center for Biotechnology Research, Protein Chemistry Core Facility, University of Florida, Gainesville, FL).

2.5. Peanut-positive human patient sera and Western blot analysis

In order to study the immunoreactivity of Ara h 3-im, a modified method of Xiang et al. [21] was used for immunoblotting. Sera (#18500-DB and #9735-RE) from two patients with a history of peanut allergy were purchased from Plasma-Lab International (<http://allergicdonorinfo.com/>; Everett, WA). The IgE level of these sera were 192 IU/ml (#18500-DB) and more than 5000 IU/ml (#9735-RE). One patient serum (#18140-KT) with a high IgE level (approximately 1000 IU/ml), but no reported soybean- or peanut-allergic reactions was used as a negative control. After application of the peanut patient serum, blots were incubated with horseradish peroxidase conjugated to goat anti-human IgE (Bethyl-Labs, Montgomery, TX) and bound IgE was detected by staining for peroxidase activity using the ECL plus reagent (Amersham-Pharmacia Biotech, Inc., Piscataway, NJ) and imaged using Biomax ML film (Eastman Kodak Co., Rochester, NY).

3. Results

3.1. Genomic Southern analysis

Southern blot analysis was used to determine *ara h 3* copy number. As shown in Fig. 1, multiple bands hybridized with the probe in all digests. Four hybridization bands ranging from 2 to 6.5 kb were observed in the *Eco*RI digested genomic DNA and the hybridization intensity was strongest for the 5 kb fragment. When hybridization was done with genomic DNA digested with *Hind*III and *Dra*I, five (1.7–6 kb) and seven (1.2–6 kb) fragments were detected, respectively. Three strong bands of ~6, 7 and 8 kb were detected in the *Bam*HI digest. Collectively, these results indicate that there are more than two copies of *ara h 3* in the tetraploid peanut genome.

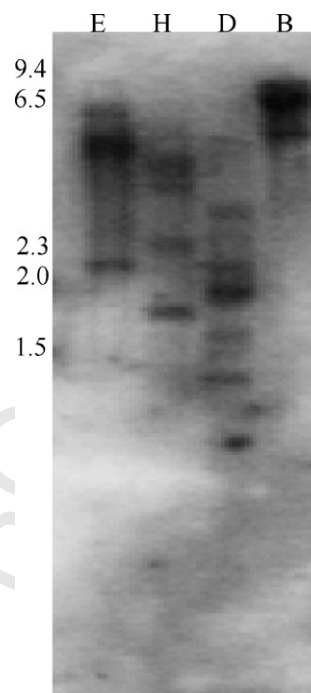


Fig. 1. Copy number determination for *ara h 3*. E: *Eco*RI, H: *Hind*III, D: *Dra*I, B: *Bam*HI. Numbers on the left of the panel refer in kb to the 1 kb ladder size marker (Promega, WI, USA).

3.2. Cloning and characterization of a novel peanut allergen gene

Because the genomic Southern blot analysis indicated the existence of multiple *ara h 3* genes in peanut, a cDNA library was screened to identify previously uncharacterized *ara h 3*-related genes. From this screening, *ara h 3-im* (GenBank accession no. AY618460) was cloned. The nucleotide and the deduced amino acid sequences of *ara h 3-im* are shown in Fig. 2. The cDNA is 1,533 nucleotides in length. Sequence homology at the nucleotide level among *gly1*, *ara h 3*, and *ara h 4* genes is 85% or greater, but the percent identity between *ara h 3-im* and these genes are 75, 73, and 71% for *gly1*, *ara h 3* and *ara h 4*, respectively. Because *ara h 3/ara h 4* is a genomic clone, it was not used for nucleotide sequence comparison.

This *ara h 3-im* cDNA has an open reading frame coding for 510 amino acids (Fig. 2). However, it is lacking the initiator methionine at the N-terminus, similar to the previously cloned *ara h 3* [8]. Thus, this is a partial cDNA clone missing a portion of the 5' end. The reading frame begins with an AAG (Lys; K) codon and ends with an UAA stop codon. The calculated size of the protein encoded by this open reading frame is ~58 kDa. The percent identity among the deduced amino acid sequences for Ara h 3/Ara h 4, Ara h 4, Gly1, and Ara h 3 is more than 87%. However, sequence homology between Ara h 3-im and these four proteins is 73% for Ara h 3/Ara h 4, 71% for Ara h 4, 70% for Gly1, and 67% for Ara h 3.

Amino acid sequences of these five proteins were aligned (Fig. 3). They are composed of two polypeptide domains, an acidic and a basic chain, which can be separated through the cleavage of an Asn-Gly (N-G) peptide bond. An Asn-Gln (N-Q)

1	AAG CTT CTT GCG CTT TCG CTT TGC TTT TGC GTA CTG GTT CTG GGC GCC AGT AGC GTC ACC TTC AGG CAG GGG	72
1	K L L A L S L C F C V L V L G A S S V T F R Q G	24
73	GGG GAG GAG AAT GAG TGC CAG TTC CAG CGC CTC AAT GCC CAG AGA CCC GAC AAC CGC ATT GAG TCT GAA GGC	144
25	G E E N E C Q F Q R L N A Q R P D N R I E S E G	48
145	GGT TAC ATT GAG ACT TGG AAC CCC AAC AAC CAG GAG TTC CAG TGC GCC GGC GTC GCC CTC TCT CGC ACC GTC	216
49	G Y I E T W N P N Q E F Q C A G V A L S R T V	72
217	CTC CGC CGC AAT GCC CTT CGC AGG CCT TTC TAC TCC AAT GCT CCC CTC GAG ATC TAC GTT CAG CAA GGA AGC	288
73	L R R N A L R R P F Y S N A P L E I Y V Q Q G S	96
289	GGA TAC TTT GGG TTG ATA TTC CCT GGT TGT CCT AGT ACA TAT GAA GAG CCT GCA CAA GAA GGA CGC CGA TAT	360
97	G Y F G L I F P G C P S T Y E E P A Q E G R R Y	120
361	CAG TCC CAA AAG CCA TCA AGA CGT TTT CAA GTA GGG CAG GAC GAC CCA AGC CAA CAG CAA CAG GAT AGT CAC	432
121	Q S Q K P S R R F Q V G Q D D P S Q Q Q Q D S H	144
433	CAG AAG GTG CAC CGC TTC GAT GAG GGT GAT CTC ATT GCA GTT CCC ACC GGT GTT GCT TTC TGG ATG TAC AAC	504
145	Q K V H R F D E G D L I A V P T G V A F W M Y N	168
505	GAC GAA GAC ACT GAT GTT GTT ACC GTT ACT CTT AGT GAC ACT AGC AGT ATC CAC AAC CAG CTC GAC CAG TTT	576
169	D E D T D V V T V T L S D T S S I H N Q L D Q F	192
577	CCC AGG AGA TTC TAT CTG GCC GGA AAC CAA GAG CAA GAG TTC TTA AGG TAC CAG CAG CAG CAA GGT AGT CGT	648
193	P R R F Y L A G N Q E Q E F L R Y Q Q Q Q G S R	216
649	CCA CAT TAT CGA CAA ATT AGC CCA AGG GTA CGA GGT GAC GAA CAA GAA AAC GAA GGT AGC AAT ATC TTC AGC	720
217	P H Y R Q I S P R V R G D E Q E N E G S N I F S	240
721	GGC TTT GCA CAG GAG TTC CTT CAA CAC GCC TTC CAG GTT GAT AGA CAA ACC GTA GAA AAT CTA AGA GGA GAG	792
241	G F A Q E F L Q H A F Q V D R Q T V E N L R G E	264
793	AAC GAG CGC GAG GAA CAA GGA GCC ATT GTT ACA GTC AAG GGA GGC CTC AGA ATC TTA AGC CCA GAT GAA GAA	864
265	N E R E E Q G A I V T V K G G L R L S P D E E	288
865	GAT GAG AGC TCG AGG TCG CCT CCC AGC AGA CGA GAG GAA TTC GAT GAG GAT CGA TCT AGG CCA CAA CAA CGT	936
289	D E S S R S P P S R R E E F D E D R S R P Q Q R	312
937	GGT AAA TAC GAT GAG AAC CGC AGA GGC TAC AAA AAT GGT ATT GAG GAA ACC ATC TGC TCT GCC AGT GTT AAG	1008
313	G K Y D E N R R G Y K N G I E E T I C S A S V K	336
1009	AAG AAC CTT GGC AGA TCC TCA AAC CCT GAC ATC TAC AAC CCT CAA GCT GGT AGC CTC AGA AGC GTC AAC GAA	1080
337	K N L G R S S N P D I Y N P Q A G S L R S V N E	360
1081	CTC GAC CTC CCA ATT CTC GGC TGG CTC GGT CTC AGT GCT CAA CAT GGA ACA ATC TAT AGG AAT GCA ATG TTT	1152
361	L D L P I L G W L G L S A Q H G T I Y R N A M F	384
1153	GTG CCT CAC TAC ACC CTC AAC GCA CAC ACC ATC GTT GTG GCA TTG AAC GGA CGG GCA CAT GTG CAA GTG GTG	1224
385	V P H Y T L N A H T I V V A L N G R A H V Q V V	408
1225	GAC AGC AAC GGT AAC AGA GTG TAC GAC GAG GAG CTT CAA GAG GGT CAC GTG CTT GTG GTG CCA CAG AAC TTC	1296
409	D S N G N R V Y D E E L Q E G H V L V V P Q N F	432
1297	GCC GTC GCA GCA AAG GCC CAG AGC GAG AAC TAC GAA TAC TTG GCG TTC AAG ACA GAC TCA AGG CCC AGC ATA	1368
433	A V A A K A Q S E N Y E Y L A F K T D S R P S I	456
1369	GCC AAC CTG GCC GGC GAA AAC TCC ATC ATA GAT AAC TTG CCG GAG GAG GTG GTT GCA AAT TCT TAC CGC CTC	1440
457	A N L A G E N S I I D N L P E E V V A N S Y R L	480
1441	CCA AGG GAG CAG GCA AGG CAA CTT AAG AAC AAC AAC CCC TTC AAG TTC TTC GTT CCA CCT TTC GAT CAT CAG	1512
481	P R E Q A R Q L K N N N P F K F F V P P F D H Q	504
1513	TCT ATG AGG GAG GTG GCT TAA	1533
505	S M R E V A *	510

Fig. 2. Nucleotide and deduced amino acid sequences of *ara h 3-im* (GenBank accession no. AY618460). Top line is the nucleotide sequence, and the lower line is the corresponding amino acid sequence. Sequences (E1–E4) highlighted in yellow show predicted IgE-binding epitopes, and the predicted critical amino acids for IgE-binding are highlighted in red as determined from a previous Ara h 3 epitope study [8]. Four cysteine residues are highlighted in green. Two peptide bonds, Asn (N₁₈₇)–Gln (Q₁₈₈) highlighted in pink, and Asn (N₃₂₄)–Gly (G₃₂₅) highlighted in light blue are shown. An asterisk (*) indicates the stop codon.

peptide bond and four conserved cysteine residues, which can form inter- and intra-disulfide bonds, are also present. Specifically, Ara h 3-im has an acidic (36 kDa) and a basic (20 kDa) chain based on the presence of an Asn₃₂₄–Gly₃₂₅ peptide bond that is critical for acidic and basic chain separation. There are four cysteine residues (Cys₃₀, Cys₆₃, Cys₁₀₆, and Cys₃₃₁) and one dominant post-translational cleavage site, Asn₁₈₇–Gln₁₈₈.

However, Ara h 3-im has several divergent regions when compared to the other Ara h 3-related proteins that make it distinct. Its N-terminal sequence (VTFRQGGEENE) shares only 7 of the first 12 amino acid residues in common with Ara h 3 and the other proteins (Fig. 3) which was confirmed by N-terminal sequencing (Fig. 4). Additionally, the potential N-glycosylation site found in the acidic (NES) chain of Ara h 4, Gly1 and Ara h 3/Ara h 4, and the potential N-glycosylation site

			E1	
Ara h 3	-----ISFRQQPEENACQFQRLNAQRPDNRIESEG	YIETWNPNNQEFECAGVA	49	
Ara h 4	---MAKLELSFCFCFLVLGASSISFRQQPEENACQFQRLNAQRPDNRIESEG	YIETWNPNNQEFECAGVA	69	
Gly1	MIRGRALSVCFCLVLGASSISFRQQPEENACQFQRLNAQRPDNRLESEG	YIETWNPNNQEFECAGVA	70	
Ara h 3/4	---MGKLLALSVCFCFLVLGASSISFRQQPEENACQFQRLNAQRPDNRIESEG	YIETWNPNNQEFECAGVA	69	
Ara h 3-im	---KLLALSVCFCFLVLGASSVTFRQGGEEENECQFQRLNAQRPDNRIESEG	YIETWNPNNQEFECAGVA	67	
Ara h 3	LSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFPGCPHYEPTQGRRSQSQRPPRLQ-GEDQ-S		117	
Ara h 4	LSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFPGCPSTYEPAQQGRYQSQRPPRLQ-EEDQ-S		137	
Gly1	LSRLVLN---LRRPFYSNAPQEIFIQQGRGYFGLIFPGCPSTYEPAQQGRRHQSQRAPRRFE-GEDQ-S		135	
Ara h 3/4	LSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFPGCPSTYEPAQQGRRHQSQRPPRRFQ-GQDQ-S		137	
Ara h 3-im	LSRTVLRNALRRPFYSNAPLEIYVQQSGGYFGLIFPGCPSTYEPAQQGRYQSQKPSRRFQVQGDQDS		137	
Ara h 3	QQQR-DSHQKVHRFDEGDLIAVPTGVAFWLYNDHDTVAVSLTDNNNDNQLDQFPRRFLNAGNTEQEF		186	
Ara h 4	QQQQ-DSHQKVHRFNEGDLIAVPTGVAFWLYNDHDTVAVSLTDNNNDNQLDQFPRRFLNAGNHEQEF		206	
Gly1	QQQQQDSHQKVHRFDEGDLIAVPTGVAFWLYNDHDTVAVSLTDNNNDNQLDQFPRRFLNAGNHEQEF		205	
Ara h 3/4	QQQQ-DSHQKVHRFDEGDLIAVPTGVAFWLYNDHDTVAVSLTDNNNDNQLDQFPRRFLNAGNHEQEF		206	
Ara h 3-im	QQQQ-DSHQKVHRFDEGDLIAVPTGVAFWLYNDHDTVAVSLTDNNNDNQLDQFPRRFLNAGNHEQEF		206	
Ara h 3	LRYYQQSRQSRRRSLPSPSPSQPRQEEREFSPRGQHSRRERAGQEEENEGGNIFSGFTPEFLQAQFQ		256	
Ara h 4	LRYYQQSRQSRRRSLPSPSPSPSPSRPREEREFSPRGQHSRRERAGQEEENEGGNIFSGFTPEFLQAQFQ		276	
Gly1	LRYYQQS---RRRSLPSPSPSQPRQEEREFSPRGQHSRRERAGQEEENEGGNIFSGFTPEFLQAQFQ		272	
Ara h 3/4	LRYYQQS---RRRSLPSPSPSPQTPKQEDREFSPRGQHSRRERAGQEEENEGGNIFSGFTPEFLQAQFQ		273	
Ara h 3-im	LRYYQQS---GSRP-HY-----RQISPRVR-----GDEQENEGSNIFSGFAQEFLOHAQFQ		252	
Ara h 3	VDDRQIVQNLRGETESEEEGAI	VTVRGGLRLILSPDRKRA-----DEEEYDEDEYEDDE-----DRRRG	317	
Ara h 4	VDDRQIVQNLRGENESEEEGAI	VTVRGGLRLILSPDGTRGA-----DEEEYDEDEYEDDE-----DGRRG	337	
Gly1	VDDRQIVQNLRGENESEEEGAI	VTVRGGLRLILSPDRKRA-----DEEEYDEDEYEDDE-----DRRRG	333	
Ara h 3/4	VDDRQIVQNLRGENESEDEGAI	VTVRGGLRLILSPDRKRRQYERPEDEEEYDEDEYEDDEERQDQRRRG	343	
Ara h 3-im	VD-RQTVENLRGENEREDEGAI	VTVRGGLRLILSPDEEDSSRSPPSRREEFDEDR-SRPQ---QRGKYD	316	
Ara h 3	RGSRRGNIGIEETICTASAKKNIGNRSPDIYNPQAGSLKTANDL--NLLILRWLGPSAEYGNLYRNALF		385	
Ara h 4	RGSRRGNIGIEETICTACVKKNIGNRSPHIYDPQWFTQCHDL--NLLILRWLGLSAEYGNLYRNALF		405	
Gly1	RGSRRGNIGIEETICTASVKKNIGNRSPDIYNPQAGSLKTANDL--NLLILRWLGLSAEYGNLYRNALF		401	
Ara h 3/4	RGSRRGNIGIEETICTASFKNIGNRSPDIYNPQAGSLKTANELQLNLLILRWLGLSAEYGNLYRNALF		413	
Ara h 3-im	ENRRGYKNIGIEETICTASVKKNLGRSSNPDIYNPQAGSLRSVNEL--DLPILRWLGLSAQHGITYRNAMF		384	
Ara h 3	VAHYNTNAHSIIYALRGRAHVQVVDSSNGNRVYDEELQEGHVLVVPQNFVAVAGKSQSENFEYVAFKTDSP		455	
Ara h 4	VPHYNTNAHSIIYALRGRAHVQVVDSSNGNRVYDEELQEGHVLVVPQNFVAVAGKSQSENFEYVAFKTDSP		475	
Gly1	VPHYNTNAHSIIYALRGRAHVQVVDSSNGNRVYDEELQEGHVLVVPQNFVAVAGKSQSDNFEYVAFKTDSP		471	
Ara h 3/4	VPHYNTNAHSIIYALRGRAHVQVVDSSNGDRVDEELQEGHVLVVPQNFVAVAGKSQSENFEYVAFKTDSP		483	
Ara h 3-im	VPHYTLNAHTIVVALNGRAHVQVVDSSNGNRVYDEELQEGHVLVVPQNFVAVAKAQSENFEYVAFKTDSP		454	
Ara h 3	SIANLAGENSVIDNLPEEVVANSYGLPREQARQLKNNNPFFKFFVPP-SQQSPRAVA		510	
Ara h 4	SIANFAGENSVIDNLPEEVVANSYGLPREQARQLKNNNPFFKFFVPP-FQQSPRAVA		530	
Gly1	NIANFAGENSVIDNLPEEVVANSYGLPREQARQLKNNNPFFKFFVPP-SQQSLRAVA		526	
Ara h 3/4	SIANLAGENSVIDNLPEEVVANSYGLPREQARQLKNNNPFFKFFVPP-SEQSLRAVA		538	
Ara h 3-im	SIANLAGENSVIDNLPEEVVANSYRLPREQARQLKNNNPFFKFFVPPFDHQMREVA		510	

Fig. 3. Amino acid sequence alignment of Ara h 3-im and related proteins. The position of amino acids is indicated in the right column. Four epitopes (E1–E4) are indicated with the critical amino acids highlighted in red within each epitope region. The NG di-peptide, highlighted in yellow, is the predicted cleavage site for the acidic (the sequences inside of the black box) and basic chains. Four conserved Cys residues, highlighted in green, are shown. Ara h 4 has an additional Cys residue instead of Ser residue at position 355. In addition to the conserved NQ peptide bond, highlighted in blue, two regions (NES and NRS) are indicated in purple as potential *N*-glycosylation sites.

found in the basic (NRS) chain in Ara h 3, Ara h 4, Gly1 and Ara h 3/Ara h 4, are absent in Ara h 3-im (Fig. 3).

A previous study using a recombinant Ara h 3 discovered four IgE epitopes containing critical amino acid residues [8]. Ara h 3-im contains IgE epitopes designated E1–E3, represented by amino acid residues 51–65, 236–249, and 274–288, and is missing most of E4 (Fig. 3). According to the identified E4 [8], it is composed of 15 amino acid residues (DEDEYEDDEDRRG), recognized by serum IgE from 38% of peanut patients tested and the underlined amino acid residues, particularly the sixth residue, glutamic acid (E), are the critical residues for IgE-binding of E4. However, Ara h 3-im has only the first three amino acid residues (DED) (303–305),

therefore missing the critical residues for IgE-binding of E4 found in all of the other Ara h 3-related proteins (Fig. 3).

3.3. Identification of peanut seed proteins by N-terminal sequencing

Fig. 4A shows a peanut seed protein profile displaying 12 polypeptide bands. The molecular weights of these polypeptides are approximately 63, 44, 40, 36, 30, 26, 25, 22, 20, 17, 16, and 14 kDa, respectively. N-terminal sequencing was performed on eight of the bands to determine their identity. For the 30, 26, 25, and 16 kDa polypeptides, the amount of protein recovered was inadequate for sequencing. The 63 kDa band

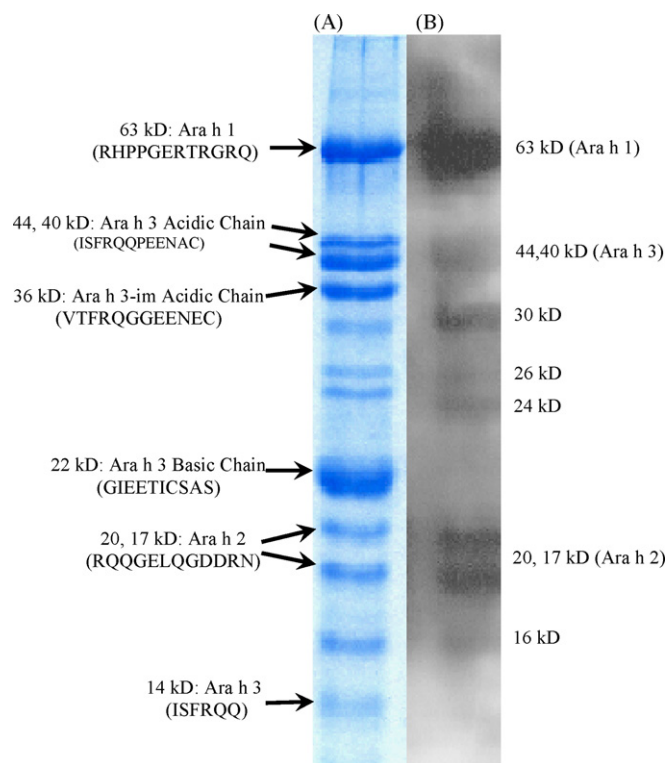


Fig. 4. N-terminal sequence analysis of peanut seed polypeptides (A) and immunoblot with a peanut-allergic patient serum (B). Protein was extracted from Georgia Green seed (developmental stage 3). The IgE source was 18500-DB, however, 9735-RE gave similar results (Plasma Lab International, Everett, WA). Protein molecular weights were estimated based on SDS-PAGE using a protein size marker (Novagen, Madison, WI).

h 3 [8]. The 36 kDa band contained sequences identical to the amino acid sequences of the N-terminus of Ara h 3-im, and this was the calculated size of the Ara h 3-im acidic chain based on the deduced amino acid sequence.

3.4. IgE-binding tests for peanut seed proteins

In order to begin characterizing the IgE-binding properties of the Ara h 3-im polypeptide, immunoblotting was performed with two different peanut-specific patient sera. The IgE reactivity of these two human sera was predominantly directed toward Ara h 1, Ara h 2, and the 44 and 40 kDa polypeptides corresponding to Ara h 3, with Ara h 1 being the most reactive (Fig. 4B). However, the 36 kDa band of Ara h 3-im was not recognized by either of the tested human sera (Fig. 4B).

3.5. Three-dimensional molecular modeling

Computer-assisted, three-dimensional models of Ara h 3-im and Ara h 3 were generated using the crystal structure for soybean glycinin G1 (PDB ID:1FXZ) [19] as a template (Fig. 5). The IgE epitopes found in Ara h 3-im (Fig. 5A) and Ara h 3 (Fig. 5B) are displayed based on their corresponding positions in the primary structures. It is evident that Ara h 3-im has a different predicted secondary structure compared to Ara h 3, particularly in the acidic chain region. Ara h 3 has an alpha helical region following the E3-containing region and the E4-containing region clearly extends out from the body of the protein; these secondary motives are missing in Ara h 3-im.

4. Discussion

Based on the peanut genomic Southern blot analysis, it is clear that there are several genes with high homology to *ara h 3* in the peanut tetraploid genome. This is not a surprise based on the fact that Ara h 3 probably functions as a hetero-hexamer and several *ara h 3*-related genes such as *ara h 4*, *gly 1*, *ara h 3/ara h 4* having more than 85% sequence homology have been cloned.

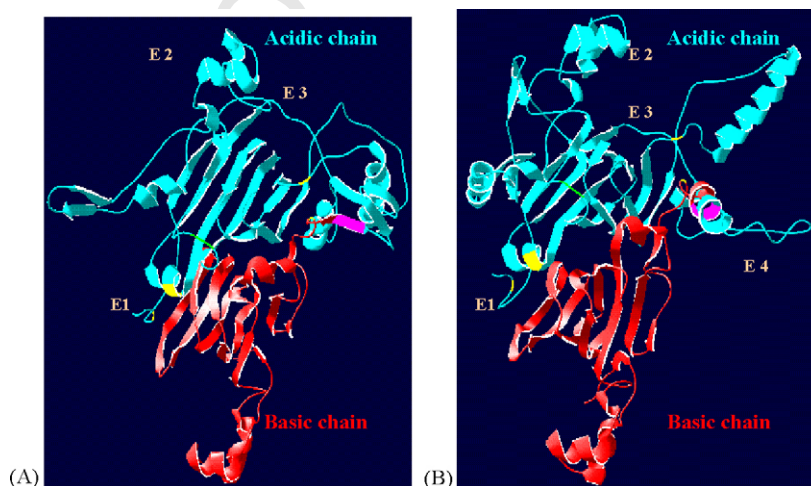


Fig. 5. Molecular models of Ara h 3-im (A) and Ara h 3 (B). The structures were modeled using soybean glycinin G1 as a template by SWISS-MODEL. Four epitopes (E1–E4) are shown. The acidic chain is displayed in blue, and the basic chain is displayed in red.

Recently, two additional cDNA clones encoding Ara h 412 [24] and a peanut putative BBTI [25] were reported as Ara h 3 isoforms with more than 90% homology at the amino acid level to Ara h 3. However, *ara h 3-im* shows much less sequence similarity (less than 75%) with these other *ara h 3* clones at both the nucleotide and deduced amino acid levels.

Ara h 3 and related proteins belong to the 11S globulin storage protein family that is characterized by three common features. The first one is that they contain an acidic and basic chain separated by a conserved Asn-Gly (N-G) peptide bond. Secondly, the formation of intra- and inter-disulfide bonds is observed due to four conserved cysteine residues. Thirdly, an Asn-Gln (N-Q) peptide bond is present that serves as a potential proteolytic cleavage site. It has been demonstrated that the vacuolar processing enzymes (VPE) mediate Asn-specific proteolytic processing in legumin-type globulin proteins and that seed-type proteolytic enzymes, including VPEs, can process storage proteins into chains capable of stable accumulation in mature seeds [26–29]. It has been shown that the assembly of soybean glycinin is regulated in part by the activity of an asparaginyl endopeptidase that cleaves asparagine (Asn; N) and glycine (Gly; G) peptide bonds [30,31]. Also, it has been demonstrated that post-translational cleavage of soybean glycinin precursors into acidic and basic polypeptides can regulate the aggregation of subunits in trimers into hexamers [27,28]. Therefore, these potential proteolytic processing sites should be essential for legume-type 11S globulin proteins [29], and accordingly are found in Ara h 3-related proteins, including the newly identified Ara h 3-im.

It has been shown that legumin-like proteins show higher sequence homology than other seed storage proteins; glycinin

subunits of soybean have been well characterized in this regard [32,33]. There are several glycinin and glycinin-like genes in the soybean genome [32], and it has been reported that other plants such as pea [34], field bean [35], and cotton [36] have several legumin-like globulin proteins. Ara h 3 and Ara h 3-im show higher sequence homology with soybean G1 and other legume LegA proteins than with soybean G4 and other 11S globulin proteins (Table 1). In fact, a study reported that Ara h 3 exhibits 62–72% sequence identity with soybean and pea homologs [8]. The soybean glycinin subunits can be separated into at least two distinct subfamilies. One subfamily includes the G1, G2 and G3 subunits, and the other subfamily contains G4 and G5 [32]. Homologies within the same family range from 80 to 90%, but sequence identities between the two families are less than 50%. Two subunits of legumin-like proteins are also present in pea [34]. Interestingly, the homology between the LegA subunit of pea and the soybean G1 subunit is higher than the similarity between two soybean glycinin subunit families [32]. Our data show that Ara h 3-im and Ara h 3 are more related to the soybean G1 subunits than to the soybean G4 subunits. Based on the fact that the existence of two distinct subfamilies is common in legumin-like proteins, it can be said that all tested Ara h 3-related proteins belong to one subfamily of peanut legumin-like proteins. Therefore, it can be suggested that another protein group homologous with soybean G4 subunits may be present in peanut.

Previous studies showed that the major IgE-binding epitopes in Ara h 3 are located in the acidic chain. The basic subunits of Ara h 3 contain incomplete IgE-binding sites; only two amino acid residues in E4 reside in the basic chain [8,11]. We also observed a lack of binding of the 22 kDa Ara h 3 basic chain by IgE. Therefore, the acidic subunits are dominantly recognized by IgE and are important in formation of the multimeric structures at least in peanut and soybean. Another study reported the same results for soybean glycinin G1, which shares similar IgE-binding epitopes with Ara h 3 [37]. However, a recent paper reported that a basic domain showed major allergenic properties in a study using a group of children allergic to peanut [14].

Although the overall predicted three-dimensional structures are similar between Ara h 3-im and Ara h 3, Ara h 3-im lacks most of E4, including the critical amino acid residues for IgE-binding. This results in a clearly different predicted secondary structure for Ara h 3-im in that region. The effect that this deletion may have on allergenicity could be important. It has been reported that mutations of any critical amino acid in an epitope leads to a significant decrease in IgE-binding [8]. Although the serum of only two patients was used, our immunoblotting recognized the major allergen polypeptides, but failed to recognize the 36 kDa protein band of Ara h 3-im. Additionally, based on the structures of soybean glycinin G1, G2 and Ara h 3, it was shown that an IgE epitope is accessible in the trimeric and hexameric state, because trimers stack upon one another with their large surface areas at the interface [21]. Therefore, we can hypothesize that there may be multiple Ara h 3 hexamers having different allergenic properties due to the variable allergenicities of the Ara h 3 monomers comprising it.

Table 1

Amino acid sequence identity among 11S globulin subunits

Protein	AI	G1	G4	LA	L2	LV	Ha	Ca	Se	Ri	Oa
A3	67	54	37	49	50	53	42	39	32	31	30
AI	—	56	36	52	53	55	42	40	33	34	30
G1	—	—	44	60	62	61	44	44	36	34	30
G4	—	—	—	45	43	44	40	35	32	31	31
LA	—	—	—	—	86	86	44	41	35	32	31
L2	—	—	—	—	—	93	47	42	37	34	30
LV	—	—	—	—	—	—	48	42	36	34	30
Ha	—	—	—	—	—	—	—	49	47	45	40
Ca	—	—	—	—	—	—	—	—	43	41	33
Se	—	—	—	—	—	—	—	—	—	41	36
Ri	—	—	—	—	—	—	—	—	—	—	58

A3: Ara h 3 from peanut (*A. hypogaea* L.) [8]. AI: Ara h 3-im protein from peanut (*A. hypogaea* L.). G1: Glycinin subunit G1 from soybean (*Glycine max* L.) [38]. G4: Glycinin subunit G4 from soybean (*Glycine max* L.) [32]. LA: Legumin A subunit from pea (*Pisum sativum* L.) [39]. L2: Legumin A2 subunit from field bean (*Vicia faba* L.) [40]. LV: Legumin A subunit from vetch (*Vicia sativa* L.) (Accession no. S44294). Ha: 11S globulin-like protein from Hazelnut (*Corylus avellana* L.) [41]. Ca: Allergen Ana 02 from cashew (*Anacardium occidentale* L.) (Accession #AAN76862). Se: 11S globulin from sesame (*Sesamum indicum* L.) [42]. Ri: Glutelin from rice (*Oryza sativa* L.) [43]. Oa: 11S globulin from oat (*Avena sativa* L.) [44]. Homologies >80% are shaded purple, 50–70% in green, and <50% in yellow.

Consequently, we conclude that Ara h 3-im may have altered allergenicity which could be lower than other Ara h 3 and related proteins. However, further studies using a larger sampling of peanut patient serum IgE along with purified natural or recombinant Ara h 3-im will be conducted to test this hypothesis. If it is true that Ara h 3-im has lower allergenic properties than other Ara h 3 proteins, this study may provide the information necessary to produce a hypoallergenic peanut through silencing of the major allergens and selecting for the reduced allergenic polypeptides via mutational breeding and/or genetic engineering.

Acknowledgements

This research was supported by a USDA-CSREES-administered special grant (00-34420-9178), and the Florida Agricultural Experiment Station.

References

- [1] M.W. Yocum, D.A. Khan, Assessment of patients who have experienced anaphylaxis: a 3-year survey, *Mayo Clin. Proc.* 69 (1994) 16–23.
- [2] S.H. Sicherer, A. Munoz-Furlong, A.W. Burks, H.A. Sampson, Prevalence of peanut and tree nut allergy in the U.S. determined by a random digit dial telephone survey, *J. Allergy Clin. Immunol.* 103 (1999) 559–562.
- [3] H.A. Sampson, Managing peanut allergy, *Br. Med. J.* 312 (1996) 1050–1051.
- [4] A.W. Burks, Peanut allergy: a growing phenomenon, *J. Clin. Invest.* 111 (2003) 950–952.
- [5] S.J. Koppelman, R.A. Vlooswijk, L.M. Knippels, M. Hessing, E.F. Knol, Van F.C. Reijssen, C.A. Bruijnzeel-Koomen, Quantification of major peanut allergens Ara h 1 and Ara h 2 in the peanut varieties Runner, Spanish, Virginia, and Valencia, bred in different parts of the world, *Allergy* 56 (2001) 132–137.
- [6] M.C. Clarke, S.A. Kilburn, J.O. Hourihane, K.R. Dean, J.O. Warner, T.P. Dean, Serological characteristics of peanut allergy, *Clin. Exp. Allergy* 28 (1998) 1251–1257.
- [7] A.W. Burks, H.A. Sampson, G.A. Bannon, Review Article Series II. Peanut allergens, *Allergy* 53 (1998) 725–730.
- [8] P. Rabjohn, E.M. Helm, J.S. Stanley, C.M. West, H.A. Sampson, A.W. Burks, G.A. Bannon, Molecular cloning and epitope analysis of the peanut allergen Ara h 3, *J. Clin. Invest.* 103 (1999) 535–542.
- [9] T. Kleber-Janke, R. Cramer, U. Appenzeller, M. Schlaak, W.M. Becker, Selective cloning of peanut allergens, including profilin and 2S albumins, by phage display technology, *Int. Arch. Allergy Immunol.* 119 (1999) 265–274.
- [10] P.A. Eigenmann, A.W. Burks, G.A. Bannon, H.A. Sampson, Identification of unique peanut and soy allergens in sera adsorbed with cross-reacting antibodies, *J. Allergy Clin. Immunol.* 98 (1996) 969–978.
- [11] S.J. Koppelman, E.F. Knol, R.A.A. Vlooswijk, M. Wensing, A.C. Knulst, S.L. Hefle, H. Gruppen, S. Piersma, Peanut allergen Ara h 3: isolation from peanuts and biochemical characterization, *Allergy* 58 (2003) 1144–1151.
- [12] S.J. Koppelman, C.A. Bruijnzeel-Koomen, M. Hessing, H.H. De Jong, Heat-induced conformational changes of Ara h 1, a major peanut allergen, do not affect its allergenic properties, *J. Biol. Chem.* 274 (1999) 4770–4777.
- [13] A. Boldt, D. Fortunato, A. Conti, A. Petersen, B. Ballmer-Weber, U. Lepp, G. Reese, W.M. Becker, Analysis of the composition of an immunoglobulin E reactive high molecular weight protein complex of peanut extract containing Ara h 1 and Ara h 3/4, *Proteomics* 5 (2005) 675–686.
- [14] P. Restani, C. Ballabio, E. Corsini, A. Flocchi, P. Isoardi, C. Magni, C. Poiesi, L. Terracciano, M. Duranti, Identification of the basic subunit of Ara h 3 as the major allergen in a group of children allergic to peanuts, *Ann. Allergy Asthma Immunol.* 94 (2005) 262–266.

- [15] S.O. Rogers, A.J. Bendich, Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissue, *Plant Mol. Biol.* 5 (1985) 69–76.
- [16] C. Notredame, D. Higgins, J. Heringa, T-Coffee: a novel method for multiple sequence alignments, *J. Mol. Biol.* 302 (2000) 205–217.
- [17] X. Huang, On global sequence alignment, *Comput. Appl. Biosci.* 10 (1994) 227–235.
- [18] N. Guex, M.C. Peitsch, SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling, *Electrophoresis* 18 (1997) 2714–2723.
- [19] M. Adachi, Y. Takenaka, A.B. Gidamis, B. Mikami, S. Utsumi, Crystal structure of soybean proglycinin A1aB1b homotrimer, *J. Mol. Biol.* 305 (2001) 291–305.
- [20] H.K. Towbin, T.H. Staehelin, J. Gordon, Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets, *Proc. Natl. Acad. Sci. U.S.A.* 76 (1979) 4350–4354.
- [21] P. Xiang, T.A. Beardslee, M.G. Zeece, J. Markwell, G. Sarath G, Identification and analysis of a conserved immunoglobulin E-binding epitope in soybean G1a and G2a and peanut Ara h 3 glycinins, *Arch. Biochem. Biophys.* 408 (2002) 51–57.
- [22] A.W. Burks, G. Cockrell, J.S. Stanley, R.M. Helm, G.A. Bannon, Recombinant peanut allergen Ara h I expression and IgE binding in patients with peanut hypersensitivity, *J. Clin. Invest.* 96 (1995) 1715–1721.
- [23] J.S. Stanley, N. King, A.W. Burks, S.K. Huang, H. Sampson, G. Cockrell, R.M. Helm, C.M. West, G.A. Bannon, Identification and mutational analysis of the immunodominant IgE binding epitopes of the major peanut allergen Ara h 2, *Arch. Biochem. Biophys.* 342 (1997) 244–253.
- [24] J.M. Chatel, L. Song, B. Bhogal, F.M. Orson, Various factors (allergen nature, mouse strain, CpG/recombinant protein expressed) influence the immune response elicited by genetic immunization, *Allergy* 58 (2003) 641–647.
- [25] H.W. Dodo, O.M. Viquez, S.J. Maleki, K.N. Konan, cDNA clone of a putative peanut (*Arachis hypogaea* L.) trypsin inhibitor has homology with peanut allergens Ara h 3 and Ara h 4, *J. Agric. Food Chem.* 52 (2004) 1404–1409.
- [26] K. Yamada, T. Shimada, M. Kondo, M. Nishimura, I. Hhara-Nishimura, Multiple functional proteins are produced by cleaving Asn-Gln bonds of a single precursor by vacuolar processing enzyme, *J. Biol. Chem.* 274 (1999) 2563–2570.
- [27] D.F. Gruis, D.A. Selinger, J.M. Curran, R. Jung, Redundant proteolytic mechanisms process seed storage proteins in the absence of seed-type members of the vacuolar processing enzyme family of cysteine proteases, *Plant Cell* 14 (2002) 2863–2882.
- [28] D. Gruis, J. Schulze, R. Jung, Storage protein accumulation in the absence of the vacuolar processing enzyme family of cysteine proteases, *Plant Cell* 16 (2004) 270–290.
- [29] S.R. Piersma, M. Gaspari, S.L. Hefle, S.J. Koppelman, Proteolytic processing of the peanut allergen Ara h 3, *Mol. Nutr. Food Res.* 49 (2005) 744–755.
- [30] C.D. Dickinson, E.H.A. Hussein, N.C. Nielsen, Role of posttranslational cleavage in glycinin assembly, *Plant Cell* 1 (1989) 459–469.
- [31] M.P. Scott, R. Jung, K. Muntz, N.C. Nielsen, A protease responsible for post-translational cleavage of a conserved Asn-Gly linkage in glycinin, the major seed storage protein of soybean, *Proc. Natl. Acad. Sci. U.S.A.* 89 (1992) 658–662.
- [32] N.C. Nielsen, C.D. Dickinson, T.J. Cho, V.H. Thanh, B.J. Scallan, R.L. Fischer, T.L. Sims, G.N. Drews, R.B. Goldberg, Characterization of the glycinin gene family in soybean, *Plant Cell* 1 (1989) 313–328.
- [33] R. Jung, M.P. Scott, Y.W. Nam, T.W. Beaman, R. Bassuner, I. Sallbach, K. Muntz, N.C. Nielsen, The role of proteolysis in the processing and assembly of 11S seed globulins, *Plant Cell* 10 (1998) 343–357.
- [34] C. Domoney, R. Casey, Measurement of gene number for seed storage proteins in Pisum, *Nucleic Acids Res.* 13 (1985) 687–699.
- [35] H. Baumlein, U. Wobus, J. Pustell, F.C. Kafatos, The legumin gene family: structure of a B type gene of Vicia faba and a possible legumin gene specific regulatory element, *Nucleic Acids Res.* 14 (1986) 2707–2720.
- [36] C.A. Chlan, J. Pyle, A. Legocki, L. Dure III, Developmental biochemistry of cottonseed embryogenesis and germination. XVIII. cDNA and amino

- acid sequences of the members of the storage protein families, Plant Mol. Biol. 7 (1986) 475–489.
- [37] T.A. Beardslee, M.G. Zeece, G. Sarath, J.P. Markwell, Soybean glycinin G1 acidic chain shares IgE epitopes with peanut allergen Ara h 3, Int. Arch. Allergy Immunol. 123 (2000) 299–307.
- [38] T.L. Sims, R.B. Goldberg, The glycinin Gy1 gene from soybean, Nucleic Acids Res. 17 (1989) 4386.
- [39] G.W. Lycett, R.R. Croy, A.H. Shirsat, D. Boulter, The complete nucleotide sequence of a legumin gene from pea (*Pisum sativum* L.), Nucleic Acids Res. 12 (1984) 4493–4506.
- [40] B. Schlesier, R. Bassuner, N. Van Hai, K. Muntz, The cDNA derived primary structure of two distinct legumin A subunit precursors from field bean (*Vicia faba* L.), Nucleic Acids Res. 18 (1990) 7146.
- [41] K. Beyer, G. Grishina, L. Bardina, A. Grishin, H.A. Sampson, Identification of an 11S globulin as a major hazelnut food allergen in hazelnut-induced systemic reactions, J. Allergy Clin. Immunol. 110 (2002) 517–523.
- [42] S.S. Tai, L.S. Wu, E.C. Chen, J.T. Tzen, Molecular cloning of 11S globulin and 2S albumin, the two major seed storage proteins in sesame, J. Agric. Food Chem. 47 (1999) 4932–4938.
- [43] T.W. Okita, Y.S. Hwang, J. Hnilo, W.T. Kim, A.P. Aryan, R. Larson, H.B. Krishnan, Structure and expression of the rice glutelin multigene family, J. Biol. Chem. 264 (1989) 12573–12581.
- [44] M.A. Tanchak, M. Giband, B. Potier, J.P. Scherthaner, S. Dukiandjiev, I. Altosaar, Genomic clones encoding 11S globulins in oats (*Avena sativa* L.), Genome 38 (1995) 627–634.