

Molecular cloning and epitope analysis of the peanut allergen Ara h 3

Pat Rabjohn,¹ Erica M. Helm,¹ J. Steven Stanley,² C. Michael West,² Hugh A. Sampson,³ A. Wesley Burks,² and Gary A. Bannon¹

¹Department of Biochemistry and Molecular Biology, and

²Department of Pediatrics, University of Arkansas for Medical Sciences, Arkansas Children's Hospital Research Institute, Little Rock, Arkansas 72205, USA

³Department of Pediatrics, Mount Sinai Medical School, New York, New York 10029, USA

Address correspondence to: Gary A. Bannon, University of Arkansas for Medical Sciences, Slot 516, 4301 W. Markham, Little Rock, Arkansas 72205, USA. Phone: (501) 686-5787; Fax: (501) 686-8169; E-mail: Bannongarya@exchange.uams.edu

Received for publication September 28, 1998, and accepted in revised form January 1, 1999.

Peanut allergy is a significant IgE-mediated health problem because of the increased prevalence, potential severity, and chronicity of the reaction. Following our characterization of the two peanut allergens Ara h 1 and Ara h 2, we have isolated a cDNA clone encoding a third peanut allergen, Ara h 3. The deduced amino acid sequence of Ara h 3 shows homology to 11S seed-storage proteins. The recombinant form of this protein was expressed in a bacterial system and was recognized by serum IgE from ~45% of our peanut-allergic patient population. Serum IgE from these patients and overlapping, synthetic peptides were used to map the linear, IgE-binding epitopes of Ara h 3. Four epitopes, between 10 and 15 amino acids in length, were found within the primary sequence, with no obvious sequence motif shared by the peptides. One epitope is recognized by all Ara h 3-allergic patients. Mutational analysis of the epitopes revealed that single amino acid changes within these peptides could lead to a reduction or loss of IgE binding. By determining which amino acids are critical for IgE binding, it might be possible to alter the Ara h 3 cDNA to encode a protein with a reduced IgE-binding capacity. These results will enable the design of improved diagnostic and therapeutic approaches for food-hypersensitivity reactions.

J. Clin. Invest. 103:535-542 (1999).

Introduction

Peanut allergy is a major health concern because of the severity of the allergic reaction, the persistence of the allergic response throughout the lifetime of the individual, and the ubiquitous use of peanut as a protein supplement in processed foods. Food-hypersensitivity reactions affect ~8% of children and 1%–2% of adults (1, 2). Peanuts, tree nuts, and shellfish are responsible for the majority of food-hypersensitivity reactions in adults, whereas peanuts, milk, and eggs account for the majority of reactions in children (3). The reaction to peanut is generally more severe than the reaction to other foods, often resulting in fatal anaphylaxis. Although most children outgrow allergies to milk and eggs, peanut allergies persist into adulthood, lasting the entire lifetime of the individual (4). Currently, avoidance is the only treatment for patients with peanut allergies. Unfortunately, the inclusion of peanut as a protein supplement in processed foods makes accidental consumption almost inevitable. Despite the prevalence of peanut hypersensitivity in children and an increasing number of deaths each year from peanut-induced anaphylaxis, the identification and characterization of unique, clinically relevant allergens from peanut are incomplete, limiting our understanding of their role in the immunobiology of hypersensitivity reactions.

Exposure to an allergen induces the production of allergen-specific IgE antibodies and the subsequent development of an allergic response in atopic individu-

als (5). In recent years, a number of allergens have been identified that stimulate IgE production and cause IgE-mediated disease (6–9). Although considerable information exists on the identification, purification, and cloning of inhaled allergens (pollens, dust mites, animal danders, insects, and fungi), few food allergens have been identified and studied.

The peanut (*Arachis hypogaea*) is a member of the legume (Leguminosae) family, which includes the soybean, jack bean, lima bean, and pea (10). Peanut shares many cross-reacting proteins with these other members; however, ingestion of these other legumes rarely elicits a clinical reaction in peanut-allergic patients (11). To identify allergenic proteins unique to the peanut, cross-reacting antibodies to soybean were removed from the sera of peanut-allergic patients. This soy-adsorbed sera, containing IgE antibodies unique to peanut, identified several allergenic fractions on immunoblots to whole peanut protein (12). Two of these fractions, Ara h 1 (63.5 kDa) and Ara h 2 (17 kDa), have been extensively characterized in our laboratory. These proteins have been isolated, and their corresponding cDNA has been cloned from a mature peanut cDNA library (13, 14). Both proteins are recognized by serum IgE from >90% of peanut-allergic patients, establishing them as major peanut allergens. Serum IgE recognition of these allergens appears to be due primarily to amino acids comprising linear epitopes in the absence of carbohydrates (14, 15). Mutational analysis of these immunodominant epitopes

Table 1

Ara h 3 IgE-binding epitopes

Epitope	Amino acid sequence	Ara h 3 position	Recognition (%)
1	IETWNPNNQEFECAG	33–47	25
2	GNIFSGFTPEFLEQA	240–254	38
3	VTVRGGLRILSPDRK	279–293	100
4	DEDEYEYDEEDRG	303–317	38

The Ara h 3 IgE-binding epitopes are indicated as the single-letter amino acid code. The position of each peptide with respect to the Ara h 3 protein coding sequence is indicated. The percent recognition is the percentage of patients previously shown to recognize recombinant Ara h 3 whose serum IgE recognized that particular synthetic epitope.

has revealed that IgE binding can be abolished by single amino acid mutations within each epitope (14, 15). These results indicate that recombinant protein can be considered for use in both diagnostic and immunotherapeutic approaches to peanut hypersensitivity.

In this article, we report the cloning of the cDNA encoding a third peanut allergen, Ara h 3. Sequencing of the Ara h 3 cDNA identified this allergen as a legumin-like seed-storage protein. The recombinant form of this protein has been expressed in a bacterial system and is recognized by serum IgE from 44% (8/18) of our peanut-hypersensitive patient population. The derived amino acid sequence has been used to construct synthetic peptides for use in performing a detailed analysis of the linear IgE-binding epitopes of this protein. These results will enable the design of improved diagnostic and therapeutic approaches to treat individuals with peanut hypersensitivity.

Methods

Patients. Serum from patients with documented peanut hypersensitivity was used to probe recombinant protein and identify the Ara h 3 IgE-binding epitopes. Each patient had a positive immediate skin prick test to peanut and either a positive double-blind, placebo-controlled food challenge or a convincing history of peanut anaphylaxis (laryngeal edema, severe wheezing, and/or hypotension). One individual with elevated serum IgE levels (who did not have peanut-specific IgE or exhibit peanut hypersensitivity) served as a control in these studies. In some instances a serum pool, consisting of equal aliquots of serum IgE from each of the patients, was used in immunoblot analysis experiments to determine the IgE-binding characteristics of the population. Details outlining the challenge procedure and collection of IgE serum have been discussed previously (16). All studies were approved by the Human Use Advisory Committee at the University of Arkansas for Medical Sciences.

Isolation and amino acid sequence analysis of peanut allergen Ara h 3. Gel slices containing Ara h 3 were sent to the W.M. Keck Foundation (Biotechnology Resource Laboratory, Yale University, New Haven, Connecticut, USA) for amino acid sequencing. The NH₂-terminal amino acid sequence of Ara h 3 was determined by performing Edman degradation on an Applied Biosystems Inc. (Foster City, California, USA) gas-phase sequencer with an online HPLC column that was eluted with increasing concentrations of acetonitrile.

Identification of Ara h 3 cDNA clones. A mature peanut cDNA library was screened using a γ -ATP, 5' end-labeled, degenerate 23-bp oligonucleotide derived from the NH₂-terminal amino acid sequence (ISFRQPEENA). Positive plaques were subjected to *in vivo* excision to remove phagemid from the vector using the R408 Helper Phage (Stratagene, La Jolla, California, USA) according to a protocol supplied by the manufacturer. Super-

natants containing the excised phagemid pBluescript packaged as filamentous phage particles were decanted into sterile tubes. For DNA preparation, rescued phagemids were plated on LB-ampicillin plates using XL-1 Blue cells and incubated overnight at 37°C. Colonies appearing on the plate contain the pBluescript double-stranded plasmid with the cloned insert. DNA was prepared using the Plasmid Spin Miniprep kit (QIAGEN Inc., Valencia, California, USA) and sequenced as described later here. Several clones were identified in this manner, all of which were lacking ~300 bp of the 5' end.

Amplification of glycinin cDNA ends. The CapFinder PCR cDNA Library Construction Kit (CLONTECH Laboratories Inc., Palo Alto, California, USA) was used to selectively amplify the 5' portion of the cDNA encoding Ara h 3. Poly (A)⁺ RNA was isolated as described previously (13). For first-strand cDNA synthesis, 0.5 μ g poly (A)⁺ RNA, 1 μ l CapSwitch oligonucleotide, 5 pmol of 3F.A.S. (GCACCTTCTGGTACTATC), an antisense primer derived from a conserved nucleotide sequence present in glycinins, were incubated at 72°C for 2 min, then placed on ice for 2 min before being added to a mixture of 5 \times first-strand buffer, 2 mM DTT, 1 mM dNTP, and 100 U of Moloney murine leukemia virus reverse transcriptase. This reaction proceeded at 42°C for 1 h and was placed on ice. To amplify double-stranded cDNA, 2 μ l of first-strand cDNA and 5 pmol each of 3F.A.S. and H1 (to serve as primers) were added to a reaction mixture containing 1 \times KlenTaq PCR buffer, 0.2 mM dNTP, 1 \times KlenTaq Polymerase Mix, and dH₂O. The PCR reaction commenced with a 1-min denaturation at 95°C, followed by 22 cycles of denaturation at 95°C and annealing/elongation for 5 min at 68°C. The amplified 5' portion of the Ara h 3 cDNA was cloned into a pGEM-T vector by standard protocols supplied by Promega Corp. (Madison, Wisconsin, USA).

PCR amplification of the Ara h 3 mRNA sequence. Two oligonucleotides — *Rab-1* (CGNCAGCAACCGGAGGAGAACGC), derived from nucleotide sequence obtained from selective amplification of the 5' end of peanut glycinins, and T7+ (CGACTCACTATAGGGCGAATTGG), an oligonucleotide derived from the pBluescript vector sequence — served as primers for a PCR reaction to selectively amplify the nucleotide sequence encoding the Ara h 3 protein. Our mature peanut cDNA library served as template and was concentrated by standard phenol/chloroform extraction followed by ethanol precipitation. Each PCR reaction consisted of 1 μ l of concentrated cDNA library, 5 pmol of each primer, 0.2 mM dNTP, and 1.25 U of *Taq* DNA polymerase. These reactions were carried out in a buffer containing 3 mM MgCl₂, 500 mM KCl, and 100 mM Tris-HCl (pH 9.0). After an initial denaturation cycle at 94°C for 2 min, 30 cycles of PCR 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 were carried out in a thermocycler (Perkin-Elmer Corp., Norwalk, Connecticut, USA). After separating by electrophoresis on a 1% agarose gel and purification, products of the appropriate size were inserted into a pGEM-T vector.

DNA sequencing and analysis. Sequencing was performed according to the methods of Sanger *et al.* (17) using oligonucleotide primers directed to different regions of the clone and the femtomole DNA Cycle Sequencing System (Promega Corp.). Sequence analysis was performed on the University of Arkansas for Medical Science's Vax computer using the Wisconsin DNA analysis software package (18).

Bacterial expression and purification of recombinant Ara h 3. A cDNA corresponding to the Ara h 3 sequence was amplified by PCR and cloned into a pET vector. This plasmid allowed expression of a recombinant protein that included the addition of Ala¹, Ser², and Phe³ at the NH₂-terminus, three amino acids not encoded by our clone. Ser² and Phe³ coincide with amino acids in the native protein; however, Ile¹ in the native protein was altered to Ala¹ in the recombinant for ease of expression (19).

Ara h 3

1	<u>I S F R Q Q P E E N A C Q F Q R L N A Q R P D</u>	N
	CGG CAG CAA CCG GAG GAG AAC GGG TGC CAG TTC CAG CGC CTC AAT GCG CAG AGA CCT GAC AAT	72
25	R I E S E G G Y I E T W N P N N Q E F E C A G V	
	CGC ATT GAA TCA GAG GGC GGT TAC ATT GAG ACT TGG AAC CCC AAC CAG GAG TTC GAA TGC GCC GGC GTC	144
49	A L S R L V L R R N A L R R P F Y S N A P Q E I	
	GCC CTC TCT CGC TTA GTC CTC CGC CGC AAC GCC CTT CGT AGG CCT TTC TAC TCC AAT GCT CCC CAG GAG ATC	216
73	F I Q Q G R G Y F G L I F P G C P R H Y E E P H	
	TTC ATC CAG CAA GGA AGG GGA TAC TTT GGG TTG ATA TTC CCT GGT TGT CCT AGA CAC TAT GAA GAG CCT CAC	288
97	T Q G R R S Q S Q R P P R R L Q G E D Q S Q Q Q	
	ACA CAA GGT CGT CGA TCT CAG TCC CAA AGA CCA CCA AGA CGT CTC CAA GGA GAA GAC CAA AGC CAA CAG CAA	360
121	R D S H Q K V H R F D E G D L I A V P T G V A F	
	CGA GAT AGT CAC CAG AAG GTG CAC GGT TTC GAT GAT CTC ATT GCA GTT CCC ACC GGT GTT GCT TTC	432
145	W L Y N D H D T D V V A V S L T D T N N N D N Q	
	TGG CTC TAC AAC GAC CAC GAC ACT GAT GTT GTT GCT GTT TCT CTT ACT GAC ACC AAC AAC AAC GAC AAC CAG	504
169	L D Q F P R R R F N L A G N T E Q E F L R Y Q Q Q	
	CTT GAT CAG TTC CCC AGG AGA TTC AAT TTG GCT GGG AAC ACG GAG CAA GAG TTC TTA AGG TAC CAG CAA CAA	576
193	S R Q S R R R S L P Y S P Y S P Q S Q P R Q E E	
	AGC AGA CAA AGC AGA CGA AGA AGC TTA CCA TAT AGC CCA TAC AGC CCG CAA AGT CAG CCT AGA CAA GAA GAG	648
217	R E F S P R G Q H S R R E R A G Q E E N E G G	
	CGT GAA TTT AGC CCT CGA GGA CAG CAC AGC CGC AGA GAA CGA GCA GGA CAA GAA GAA AAC GAA GGT GGA	720
241	N I F S G G F T P E F L E Q A F C Q V D D R Q I V Q	
	AAC ATC TTC AGC TTC ACG CCG GAG TTC CTG GAA CAA GCC TTC CAG GTT GAC GAC AGA CAG ATA GTG CAA	792
265	N L R G E T E S E E G A I V T V R G G L R I L	
	AAC CTA AGA GSC GAG ACC GAG AGT GAA GAA GAG GGA GCC ATT GTG ACA GTG AGG GGA GSC CTC AGA ATC TTG	864
289	S P D R K R R A D E E E E Y D E D E Y E Y D E E	
	AGC CCA GAT AGA AAG AGA CGT GCC GAA GAA GAG GAA TAC GAT GAA GAA TAT GAA TAC GAT GAA GAG	936
313	D R R R R G R G S R G R G N G I E E T I C T A S A	
	GAT AGA AGG CGT GGC AGG GGA AGC AGA GGC AGG GGG AAT GGT ATT GAA GAG ACG ATC TGC ACC GCA AGT GCT	1008
337	K K N I G R N R S P D I Y N P Q A G S L K T A N	
	AAA AAG AAC ATT GGT AGA AAC AGA TCC CCT GAC ATC TAC AAC CCT CAA GCT GGT TCA CTC AAA ACT GCC AAC	1080
361	D L N L L I L R W L G P S A E Y G N L Y R N A L	
	GAT CTC AAC CTT CTA ATA CTT AGG TGG CTT GGA CCT AGT GCT GAA TAT GGA AAT CTC TAC AGG AAT GCA TTG	1152
385	F V A H Y N T N A H S I I Y R L R G R A H V Q V	
	TTT GTC GCT CAC TAC AAC ACC AAC GCA CAC AGC ATC ATA TAT CGA TTG AGG GGA CGG GCT CAC GTG CAA GTC	1224
409	V D S N G N R V Y D E E L Q E G H V L V V P Q N	
	GTG GAC AGC AAC GGC AAC AGA GTG TAC GAC GAG GAG CTT CAA GAG GGT CAC GTG CTT GTG GTG CCA CAG AAC	1296
433	F A V A G K S Q S E N F E Y V A F K T D S R P S	
	TTC GCC GTC GCT GGA AAG TCC CAG AGC GAG AAC TTC GAA TAC GTG GCA TTC AAG ACA GAC TCA AGG CCC AGC	1368
457	I A N L A G E N S V I D N L P E E V V A N S Y G	
	ATA GCC AAC CTC GCC GGT GAA AAC TCC GTC ATA GAT AAC CTG CCG GAG GAG GTG GTT GCA AAT TCA TAT GGC	1440
481	L Q R R E Q A R Q L K N N N P F K F V P P S Q Q	
	CTC CAA AGG GAG CAG GCA AGG CAG CTT AAG AAC AAC CCC TTC AAG TTC TTC GTT CCA CCG TCT CAG CAG	1512
505	S P R A V A *	
	TCT CCG AGG GCT GTG GCT TAA	1533

Figure 1

Nucleotide sequence of an Ara h 3 cDNA clone. The nucleotide sequence is on the second line. The first line is the derived amino acid sequence. Boxed amino acids correspond to the amino acid sequence determined from the NH₂-terminus of the Ara h 3 protein. The numbers on the right side of the figure indicate the position of the nucleotide sequence relative to the first nucleotide in the insert. The numbers on the left side of the figure indicate the position of the amino acid sequence relative to the first determined amino acid from the protein. The underlined amino acids are found at the NH₂-terminal end of the Ara h 3 protein, but are not encoded by the cDNA clone. The stop codon TAA is indicated by an asterisk. Ara h 3 GenBank accession no. AF093541.

Primers for PCR were designed to include an *NheI* site at the 5' end of the cDNA and a *Sall* site at the 3' end of the cDNA. The primers used were 5'-TATGGCTAGCTTCCGGCAGCAACCGGAGGAG-3' (5' primer) and 5'-CCGTGACAGCCACAGCCCTCGGAGA-3' (3' primer). PCR products were cloned into the *NheI/Sall* restriction sites of the plasmid pET 24(B)⁺ under the control of the T7 lac promoter. This expression vector contains the gene encoding kanamycin resistance and coding sequence for a His₆ tag produced at the COOH-terminus of the recombinant protein. Protein expression in the *Escherichia coli* strain BL21 (DE3) was induced by the addition of isopropyl-β-D-thiogalactopyranoside to a final concentration of 1 mM once the culture reached A₆₀₀ = 0.6. The cells were harvested at 1-h intervals, resuspended in SDS sample buffer containing DTT, and boiled at 100 °C for 5 min. Samples were either used immediately for immunoblot analysis, or samples were pelleted, washed with 50 mM Tris-HCl, and stored for later use as a frozen pellet at -70 °C.

Recombinant Ara h 3 was purified from bacterial lysates under denaturing conditions using the His-Bind Purification Kit (Novagen Inc., Madison, Wisconsin, USA). Cell extracts were resuspended in 4 ml of cold Binding Buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, and 6 M urea; supplied with Novagen kit), sonicated to shear DNA, and incubated on ice for 1 h. Next, the lysate was centrifuged at 12,000 g for 45 min to remove cellular debris. The postcentrifugation supernatant was

prepared for loading onto the column by passing it through a 0.45-μm membrane using a syringe-end filter. A His-Bind Quick Column (Novagen) was packed with His-Bind metal chelation resin, washed with deionized H₂O, and charged until saturation with Charge Buffer (50 mM NiSO₄; Novagen). After equilibration of the column with Binding Buffer, 2 volumes of supernatant were loaded onto the column. The column was washed with 10 volumes of Binding Buffer and 6 volumes of Wash Buffer (20 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, and 6 M urea). Elution was achieved with 5 volumes of Elution Buffer (1 M imidazole, 0.5 M NaCl, 20 mM Tris-HCl, and 6 M urea; Novagen). Fractions collected over the course of the experiment containing recombinant Ara h 3 were lyophilized and stored in 1× PBS.

SDS-PAGE, Western blots, and IgE-binding assay. Purified recombinant Ara h 3 was analyzed by SDS-PAGE using precast 12% Tris-glycine gels (Novex, San Diego, California, USA). Samples were electrophoresed for 90 min at 125 V. Proteins were visualized by either Coomassie blue staining or by using Gelcode Blue Stain Reagent (Pierce Chemical Co., Rockford, Illinois, USA) according to the manufacturer's protocol. For immunoblot analysis, proteins were electroblotted onto nitrocellulose at 30 V for 90 min. After transfer, blots were blocked using a solution containing Tris-NaCl and 3% BSA. Alternatively, cellulose membranes containing synthetic peptides were blocked in a solution provided by Genosys Biotechnologies, Inc. (The Woodlands,

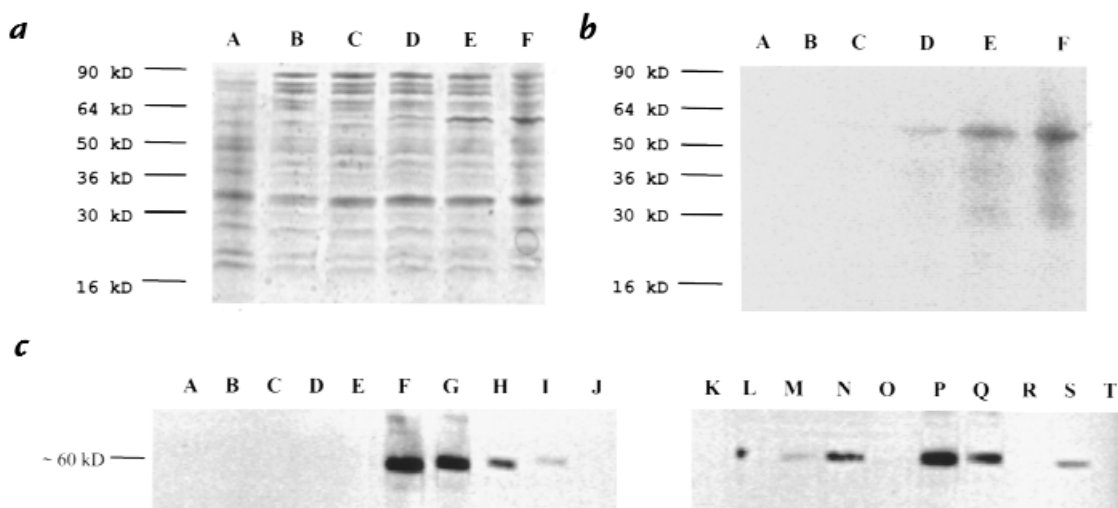


Figure 2

Bacterial expression and immunoblot analysis of recombinant Ara h 3. (a) 10- μ l samples of bacterial extract were electrophoresed in a 12% Tris-glycine gel. Proteins were visualized by Coomassie blue staining. Lane A, 4-h induction of vector containing no insert; lane B, uninduced Ara h 3 (vector with *inset*); lane C, 1-h induction; lane D, 2-h induction; lane E, 3-h induction; lane F, 4-h induction. (b) Immunoblot of a with a pool of patient serum as described in Methods. (c) Immunoblot of purified recombinant Ara h 3 with serum IgE from individual patients (lanes A–R were patients with documented peanut hypersensitivity) and a pool of serum IgE from peanut-hypersensitive patients (lane S). A nonallergic patient with elevated serum IgE served as a negative control (lane T).

Texas, USA). All blots were incubated with a serum pool from patients with documented peanut hypersensitivity or individual sera diluted (1:5) in a solution containing Tris-NaCl and 1% BSA for 16 h at 4°C. Primary antibody was detected with ¹²⁵I-labeled anti-IgE antibody (Sanofi Diagnostics Pasteur Inc., Paris, France).

Peptide synthesis. Individual peptides were synthesized with Fluorenylmethoxycarbonyl (Fmoc) amino acids on a derivatized cellulose membrane containing free hydroxyl groups according to manufacturer's instructions (Genosys Biotechnologies). Briefly, synthesis of each peptide began by esterifying an Fmoc amino acid to the cellulose membrane. Coupling reactions are followed by acetylation with acetic anhydride in *N,N*-dimethylformamide to render peptides unreactive during the subsequent steps. After acetylation, Fmoc protective groups are removed by the addition of piperidine to render nascent peptides reactive. The remaining amino acids are added by this same process of coupling, blocking, and deprotection, until the desired peptide is generated. Upon addition of the last amino acid, the side chains of the peptide are deprotected with a 1:1:0.05 mixture of dichloromethane/trifluoroacetic acid/triisobutylsilane and washed with methanol. Membranes containing synthetic peptides were either probed immediately with serum IgE or stored at -20°C until needed.

Results

Molecular cloning and sequence of the Ara h 3 cDNA. The NH₂-terminus of a purified 14-kDa protein identified by soy-adsorbed IgE serum from peanut-hypersensitive patients was used for amino acid sequencing (12). Degenerate oligonucleotides derived from the amino acid sequence were used to screen a mature peanut cDNA library. Details concerning the procedures used to clone the Ara h 3 cDNA are outlined in Methods. The sequence of the Ara h 3 cDNA and the deduced amino acid sequence are shown in Fig. 1. The Ara h 3 cDNA is an open-reading frame of 1,530 nucleotides, coding for

510 amino acids. This reading frame starts with a CGG codon and ends with a TAA stop codon at nucleotide position 1,533. The calculated size of the protein encoded by this open reading frame is ~57 kDa. The 24 amino acids obtained from NH₂-terminal sequencing of the 14-kDa protein correspond to the amino acids encoded by the nucleotides located at the 5' end of the cDNA clone. The 14-kDa protein appears to be an NH₂-terminal breakdown product of a larger allergen. This cDNA clone appears to be lacking the extreme 5' end that would encode a signal peptide and the initiator methionine. Database searches for sequence similarity revealed that the Ara h 3 cDNA encoded an 11S seed-storage protein. Ara h 3 showed 62%–72% sequence identity to glycinins from *Glycine max* (soybean) and legumins from *Pisum sativum* (pea). In particular, 24 of 26 residues thought to be important for the tertiary structure of these storage proteins (20) are present in the Ara h 3 primary sequence, including a conserved cleavage site at Asn-325 and Gly-326. There was no homology noted between this allergen and the other major peanut allergens already identified (Ara h 1 or Ara h 2).

Expression, antigenicity, and purification of recombinant Ara h 3. The Ara h 3 cDNA was cloned into a pET 24 plasmid and expressed in a bacterial system. Optimal expression was obtained following a four-hour induction by isopropyl *B*-D-thiogalactopyranoside (Fig. 2a, lane F). The immunoblot in Fig. 2b was performed using serum IgE from a pool of patients with peanut hypersensitivity to determine the molecular weight and specificity of IgE binding. From the blot, the estimated size of the recombinant protein produced by bacterial cells is ~57 kDa, which corresponds to the predicted molecular mass encoded by the clone. Figure 2c shows 20 immunoblot strips of purified recombinant

Ara h 3 incubated with different patient sera. Forty-four percent (8/18) of the patients tested had IgE that recognized the recombinant protein (Fig. 2c, lanes A–R). The difference in binding intensities between Ara h 3–allergic patients could be due to the amount of peanut-specific IgE in each individual or differences in affinity of patient-specific IgE to this allergen.

Multiple IgE-binding regions located throughout the Ara h 3 protein. Sixty-three overlapping peptides were synthesized to determine which regions of the Ara h 3 protein were recognized by serum IgE. Each peptide synthesized was 15 amino acids long and offset from the previous peptide by eight amino acids. This approach allowed the analysis of the entire Ara h 3 primary sequence in large, overlapping fragments. These peptides were probed with a serum pool of IgE from peanut-hypersensitive patients who had previously been shown to recognize recombinant Ara h 3. Figure 3 shows the four IgE-binding regions and their corresponding location within the Ara h 3 primary amino acid sequence. These IgE-binding regions were represented by amino acid residues 21–55, 134–154, 231–269, and 271–328.

To determine the exact amino acid sequence of the IgE-binding regions, synthetic peptides (15 amino acids offset by two amino acids) representing the larger IgE-binding regions were generated and probed with a serum

pool of IgE from patients who recognize recombinant Ara h 3. This process made it possible to distinguish individual IgE-binding epitopes within the larger IgE-binding regions of the Ara h 3 protein. Figure 4a is an immunoblot of six synthetic peptides, whereas Fig. 4b shows the amino acid sequence representing this region and the amino acid sequences represented by each individual peptide. The shaded area in Fig. 4b represents the core epitope. The four IgE-binding epitopes identified in this manner are shown in Table 1, with the epitopes ranging from nine to 13 amino acids in length.

Immunodominance and characterization of the Ara h 3 epitopes. To determine whether any of the four epitopes were immunodominant (within the Ara h 3–allergic population), each set of four peptides was probed individually with serum IgE from the eight patients previously shown to recognize recombinant Ara h 3 (results summarized in Table 1 as percentage recognition). Epitope 1 was recognized by serum IgE from 25% (2/8) of the patients tested, whereas epitopes 2 and 4 were recognized by serum IgE from 38% (3/8) of the eight patients tested. Interestingly, epitopes 2 and 4 were recognized by the same three patients. Epitope 3 was recognized by serum IgE from 100% (8/8) of the Ara h 3–allergic patients, classifying it as an immunodominant epitope within the Ara h 3–allergic population. Sixty-eight percent of the amino acids constituting the epitopes were

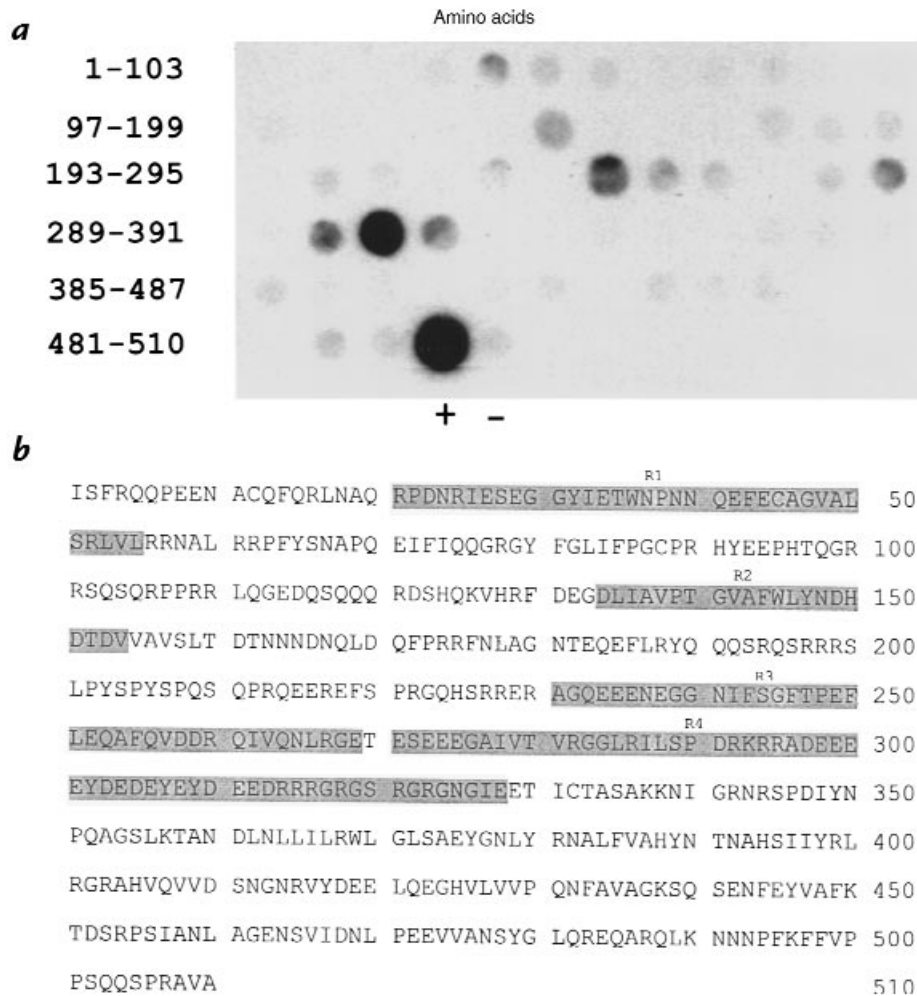
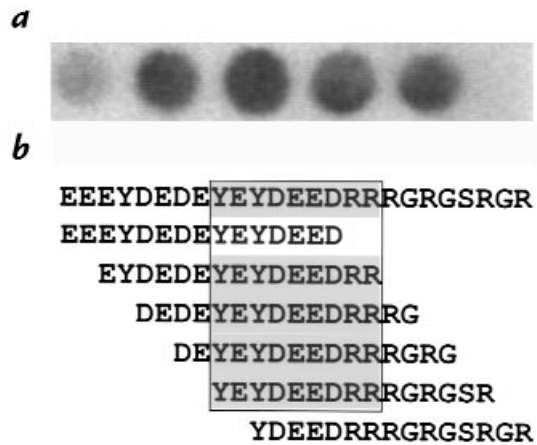


Figure 3

Multiple IgE-binding regions identified on the Ara h 3 allergen. (a) The Ara h 3 primary sequence was synthesized as 15 amino acid–long peptides offset from each other by eight residues. These peptides were probed with a pool of serum IgE from peanut-hypersensitive patients. The position of the peptides within the Ara h 3 protein are indicated on the left. (+) indicates an immunodominant peptide from Ara h 2 that served as a positive control, and (–) indicates a peptide synthesized to serve as a negative control. (b) The amino acid sequence of the Ara h 3 protein is shown. The shaded areas (R1–R4) correspond to the IgE-binding regions shown in a.

Figure 4

Identification of a core IgE-binding epitope on the Ara h 3 allergen. (a) Detailed epitope mapping was performed on IgE-binding regions by synthesizing overlapping peptides 15 amino acids in length offset from the previous peptide by two residues. These peptides were probed with a pool of serum IgE from eight patients previously shown to recognize recombinant Ara h 3. The data shown represents amino acids 299–321. (b) The amino acid sequence (residues 299–321) of Ara h 3 that was tested in a is shown. Residues in boldface correspond to common IgE-binding amino acids of the spots shown in a.



either polar uncharged or apolar residues. However, there was no obvious sequence motif with respect to position or polarity shared by the individual epitopes.

Mutations at specific residues eliminate IgE binding. The amino acids essential for IgE binding to the Ara h 3 epitopes were determined by synthesizing multiple peptides with single amino acid changes at each position. These peptides were probed with a pool of serum IgE from patients who had previously recognized the wild-type peptide, to determine whether amino acid changes affected peanut-specific IgE binding. Fig. 5 shows an immunoblot strip containing the wild-type and mutant peptides for epitope 4. The pool of serum IgE did not recognize the peptide, or a decrease in binding was observed when alanine was substituted for the wild-type amino acid at positions 308, 309, 310, 311, 312, and 314. Interestingly, it appears as if an alanine substitution increases IgE binding at positions 304 and 305. The remaining Ara h 3 epitopes were analyzed in the same manner. In general, each epitope could be altered to a non-IgE-binding peptide by the replacement of the wild-type amino acid residue with alanine. The critical residues for IgE binding within each Ara h 3 epitope are shown in Table 2. It appears that the central amino acids within each epitope are favored for mutation. All mutations that led to a significant decrease in IgE binding were located at residues found within each core epitope (as identified in Fig. 4). There was no obvious consensus in the type of amino acid that, when mutated to alanine, leads to complete loss or a decrease in IgE binding.

Table 2
Amino acids critical to IgE binding

Epitope	Amino acid sequence	Ara h 3 position
1	IETWNP NN QEFECAG	33–47
2	GN I FSG F TPE F LEQA	240–254
3	VTVRG GLR IL S PDRK	279–293
4	DEDE YEYDEED RG	303–317

The Ara h 3 IgE-binding epitopes are indicated as the single-letter amino acid code. The position of each peptide with respect to the Ara h 3 primary sequence is indicated in the right column. The amino acids that, when altered, led to a decrease in IgE binding are shown as the bold, underlined residues.

Discussion

We have reported the cDNA cloning, expression, and epitope analysis of Ara h 3, an allergenic, 11S storage protein from the peanut, *A. hypogaea*. Although these are predominant proteins in legumes, this is the first time that the cDNA from an 11S storage protein has been cloned and shown to encode an allergenic protein in the peanut. 11S storage proteins are initially synthesized as 60-kDa preproglobulins consisting of covalently linked acidic and basic polypeptides. The precursors are deposited in storage bodies where they aggregate into trimers, before being cleaved by an asparagine-dependent endopeptidase (21, 22). This results in an NH₂-terminal acidic chain of ~35 kDa and a COOH-terminal basic chain of ~20 kDa, which later become linked by a disulfide bridge (23). 11S storage proteins are then assembled into their mature form as hexameric oligomers consisting of six similar subunits (24). The Ara h 3 cDNA represents the coding region for the 60-kDa preproglobulin.

We have demonstrated high-level expression of recombinant Ara h 3 in a bacterial system. Serum IgE from 44% (8/18) of our peanut-allergic patient population recognized recombinant Ara h 3, designating it as a minor allergen. This is in contrast to the other peanut allergens, Ara h 1 and Ara h 2, both of which are major allergens (25), recognized by >90% of the patient population. All three of these allergens share similar functional properties; they are all seed-storage proteins with no enzymatic activity. However, no direct evidence exists as to why only a portion of the patient population recognizes Ara h 3. The ability of 11S storage proteins to oligomerize into hexamers and the position of the epitopes at the tertiary level of protein structure may provide insight into this issue. Another possibility is the level of sequence similarity retained between these proteins from different legumes. Ara h 3 exhibits higher sequence identity with legume storage proteins from soybean and pea (62%–72%) than Ara h 1 exhibits with vicilins (40%) or Ara h 2 exhibits with conglutinins (39%) (14, 15). The percentage of patients with allergen-specific IgE may depend on unique sequences not conserved between protein families of different legume species. This would account for the lower percentage of peanut-allergic patients with IgE to Ara h 3 and the smaller number of linear, IgE-binding regions found within this allergen. Several characteristics

have been proposed to justify why particular proteins are allergenic. These reasons include their increased abundance in the food supply, their durability during food processing (blanching), and their resistance to digestion in the gastrointestinal tract (26, 27). Whether these traits play a role in why this particular protein is a minor allergen still needs to be investigated.

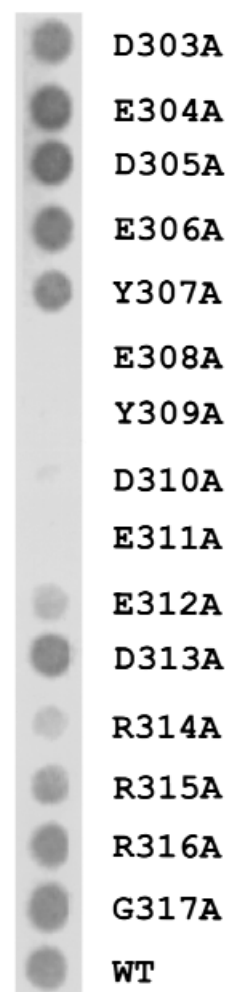
Given that allergen-specific IgE plays such a critical role in the etiology of allergic disease, determination of allergen-specific, IgE-binding epitopes is an important first step toward understanding the complexity of hypersensitivity reactions. By generating synthetic, overlapping peptides representing the entire primary sequence of the protein, we were able to determine that there are four distinct IgE-recognition sites distributed throughout the primary sequence of the protein. One of these sites (epitope 3) was recognized by serum IgE from every Ara h 3-allergic patient, designating it as an immunodominant epitope. Interestingly, epitopes 3 and 4 are located within the hypervariable region of the acidic chain, a stretch of amino acids that is highly variable in length among 11S storage proteins. This region contains a high proportion of glutamate, aspartate, and arginine residues and will tolerate large, naturally occurring insertions or deletions. Computer predictions from other studies suggest that this region is exposed on the surface of the protein (28), making it accessible to IgE binding. Multiple IgE-binding epitopes have been identified for allergens from cow milk (29), codfish (30), hazel (31), soybean (32), shrimp (33), and others. The observation of multiple IgE-binding sites probably reflects the polyclonal nature of the immune response and may be a necessary step in establishing a protein as an allergen. Also, cross-linking of IgE on the surface of the allergen is required for mast-cell degranulation, leading to the release of histamine and other mediators of the immune response. For cross-linking to occur, multiple IgE-binding sites would have to be present.

Soybean glycinins have long been the target of genetic engineering studies aimed at improving their nutritional value (34–37). Initial experiments revealed that mutations within the 35-kDa acidic subunit had little effect on the ability of glycinins to oligomerize into higher-order structures. We chose alanine for our mutational analysis based on its ability to be tolerated in both α -helix and β -sheet secondary structure. All four Ara h 3 epitopes reside within the 35-kDa acidic subunit. Because there is no overlap between the amino acids considered to be important for glycinin structure (20) and the critical residues required for IgE binding, we feel confident that the Ara h 3 cDNA can be altered to encode a nonallergenic protein that retains its native structure and function. This altered gene could be used to modify peanut plants, and the expressed protein could be used for recombinant immunotherapy (38).

Currently, the only therapeutic option available for the prevention of food allergic reactions is food avoidance. Because peanut is used as a protein supplement in a wide variety of processed foods, accidental consumption is almost inevitable (39). Recombinant allergens represent promising tools for diagnosing and treating food-hypersensitivity patients. Characteristics of recombinant aller-

Figure 5

Ara h 3 epitopes can be mutated to non-IgE-binding peptides. Epitope 4 was synthesized with an alanine residue substituted for one of the amino acids at each position in the peptide. The synthesized peptides were probed with a pool of serum IgE from peanut-hypersensitive patients whose IgE has previously been shown to recognize this peptide. The letters to the right of each lane indicate the one-letter amino acid code for the residue normally at that position and the amino acid substituted for that residue. The numbers indicate the position of each residue in the Ara h 3 protein. *WT* indicates the wild-type peptide with no amino acid substitutions.



gens that would make them useful in a clinical setting include a reduced capacity to bind serum IgE, ensuring a lower risk of IgE-mediated anaphylactic side effects, and retention of their T-cell epitopes, allowing modulation of the immune response. The elucidation of the major IgE-binding epitopes on Ara h 3 and the determination of which amino acids within these epitopes are critical for IgE binding provides the information necessary to alter the *Ara h 3* gene by site-directed mutagenesis to encode a protein that escapes IgE recognition. Because of the extensive use of peanut protein in consumer foods and the potential severity of a peanut-induced allergic reaction, our future studies are aimed at producing a cDNA that encodes a nonallergenic 11S storage protein for both immunotherapeutic purposes and stable transformation into plants to produce a hypoallergenic peanut.

Acknowledgments

This work was supported by grant RO1-AI33596 from the National Institutes of Health and by the Clarissa Sosin Allergy Research Foundation.

1. Jansen, J.J., *et al.* 1994. Prevalence of food allergy and intolerance in the adult Dutch population. *J. Allergy Clin. Immunol.* **93**:446–456.
2. Burks, A.W., and Sampson, H.A. 1993. Food allergies in children. *Curr. Prob. Pediatr.* **23**:230–252.
3. Sampson, H.A. 1988. The role of food allergy and mediator release in atopic dermatitis. *J. Allergy Clin. Immunol.* **81**:635–645.

4. James, J.M., and Sampson, H.A. 1992. An overview of food hypersensitivity. *Pediatr. Allergy Immunol.* **3**:67-78.
5. Sutton B.J., and Gould, H.J. 1993. The human IgE network. *Nature.* **366**:421-428.
6. Arruda, L.K., Vailes, L.D., Platts-Mills, T.A.E., Hayden, M.L., and Chapman, M.D. 1997. Induction of IgE antibody responses by glutathione S-transferase from the German cockroach. *J. Biol. Chem.* **272**:20907-20912.
7. Akasawa, A., Hsieh, L.S., Martin, B.M., Liu, T., and Lin, Y. 1996. A novel acidic allergen, Hev b 5, in latex. *J. Biol. Chem.* **271**: 25389-25393.
8. Gregoire, C., et al. 1996. cDNA cloning and sequencing reveal the major horse allergen Equ c1 to be a glycoprotein member of the lipocalin superfamily. *J. Biol. Chem.* **271**:32951-32959.
9. Engel, E., et al. 1997. Immunological and biological properties of Bet v 4, a novel birch pollen allergen with two EF-hand calcium-binding domains. *J. Biol. Chem.* **272**:28630-28637.
10. Gibbs, P.E.M., Strongin K.B., and McPherson, A. 1989. Evolution of legume seed storage proteins—a domain common to legumins and vicilins is duplicated in vicilins. *Mol. Biol. Evol.* **6**:614-623.
11. Broadbent, J.B., and Sampson, H.A. 1989. Cross-allergenicity in the legume botanical family in children with food hypersensitivity. *J. Allergy Clin. Immunol.* **84**:701-709.
12. Eigenmann, P.A., Burks, A.W., Bannon, G.A., and Sampson, H.A. 1996. Identification of unique peanut and soy allergens in sera adsorbed with cross-reacting antibodies. *J. Allergy Clin. Immunol.* **98**:969-978.
13. Burks, A.W., Cockrell, G., Stanley, J.S., Helm, R.M., and Bannon, G.A. 1995. Recombinant peanut allergen Ara h 1 expression and IgE-binding in patients with peanut hypersensitivity. *J. Clin. Invest.* **96**: 1715-1721.
14. Stanley, J.S., et al. 1997. Identification and mutational analysis of the immunodominant IgE-binding epitopes of the major peanut allergen Ara h 2. *Arch. Biochem. Biophys.* **342**:244-253.
15. Burks, A.W., et al. 1997. Mapping and mutational analysis of the IgE-binding epitopes on Ara h 1, a legume vicilin protein and a major allergen in peanut hypersensitivity. *Eur. J. Biochem.* **245**:334-339.
16. Burks, A.W., et al. 1991. Identification of a major peanut allergen, Ara h 1, in patients with atopic dermatitis and positive peanut challenge. *J. Allergy Clin. Immunol.* **88**:172-179.
17. Sanger, F., Nicklen, S., and Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74**:5463-5467.
18. Devereux, J., Haerberli, P., and Smithies, O. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**:387-395.
19. Tobias, J.W., Shrader, T.E., Rocap, G., and Varshavsky, A. 1991. The N-end rule in bacteria. *Science.* **254**:1374-1377.
20. Bairoch, A., and Bucher, P. 1994. PROSITE: recent developments. *Nucleic Acids Res.* **22**:3584-3589.
21. Tumer, N.E., Richter, J.D., and Nielsen, N.C. 1982. Structural characterization of the glycinin precursors. *J. Biol. Chem.* **257**:4016-4018.
22. Barton, K.A., et al. 1982. The biosynthesis and processing of high molecular weight precursors of soybean glycinin subunits. *J. Biol. Chem.* **257**:6089-6095.
23. Staswick, P.E., Hermodson, M.A., and Nielsen, N.C. 1981. Identification of the acidic and basic subunit complexes of glycinin. *J. Biol. Chem.* **256**:8752-8755.
24. Staswick, P.E., Hermodson, M.A., and Nielsen, N.C. 1984. The amino acid sequence of the A2B1a subunit of glycinin. *J. Biol. Chem.* **259**:13431-13435.
25. Shin, D.S., et al. 1998. Biochemical and structural analysis of the IgE-binding sites on Ara h 1, and abundant and highly allergic peanut protein. *J. Biol. Chem.* **273**:13753-13759.
26. Astwood, J.D., Leach, J.N., and Fuchs, R.L. 1996. Stability of food allergens to digestion in vitro. *Nat. Biotechnol.* **14**:1269-1273.
27. Veiths, S., Aulepp, J., Becker, W.M., and Buschmann, L. 1996. Characterization of labile and stable allergens in foods of plant origin. In *Food allergies and intolerances*. G. Eisenband, H. Aulepp, A.D. Dayan, P.S. Elias, and W. Grunow, editors. VCH Verlagsgesellschaft. Weinheim, Germany. 130-149.
28. Nielsen, N.C., Scott, M.P., and Lago, W.J.P. 1990. Assembly properties of modified subunits in the glycinin subunit family. In *NATO Advanced Study Institute on Plant Molecular Biology*. R. Hermann and B. Larkins, editors. Plenum Press. New York, NY. 635-640.
29. Ball, G., et al. 1994. A major continuous epitope of bovine beta-lactoglobulin recognized by human IgE binding. *Clin. Exp. Allergy.* **24**:758-764.
30. Aas, K., and Elsayed, S. 1975. Physico-chemical properties and specific activity of a purified allergen (codfish). *Dev. Biol. Stand.* **29**:90-98.
31. Elsayed, S., Holen, E., and Dybendal, T. 1989. Synthetic allergenic epitopes from the amino-terminal regions of the major allergens of hazel and birch pollen. *Int. Arch. Allergy Appl. Immunol.* **89**:410-415.
32. Herian, A.M., Taylor, S.L., and Bush, R.K. 1990. Identification of soybean allergens by immunoblotting with sera from soy-allergic adults. *Int. Arch. Allergy Appl. Immunol.* **92**:193-198.
33. Shanti, K.N., Martin, B.M., Nagpal, S., Metcalfe, D.D., and Rao, P.V. 1993. Identification of tropomyosin as the major shrimp allergen and characterization of its IgE-binding epitopes. *J. Immunol.* **151**:5354-5363.
34. Utsumi, S., et al. 1993. Synthesis, processing and accumulation of modified glycinins of soybean in the seeds, leaves and stems of transgenic tobacco. *Plant Sci.* **92**:191-202.
35. Utsumi, S., Katsube, T., Ishige, T., and Takaiwa, F. 1997. Molecular design of soybean glycinins with enhanced food qualities and development of crops producing such glycinins. *Adv. Exp. Med. Biol.* **415**:1-15.
36. Dickinson, C.D., Scott, M.P., Hussein, E.H., Argos, P., and Nielsen, N.C. 1990. Effect of structural modifications on the assembly of a glycinin subunit. *Plant Cell.* **2**:403-413.
37. Lawrence, M.C., Izard, T., Beuchat, M., Blagrove, R.J., and Colman, P.M. 1994. Structure of phaseolin at 2.2 Å resolution. Implications for a common vicilin/legumin structure and the genetic engineering of seed storage proteins. *J. Mol. Biol.* **238**:748-776.
38. Ferreira, F., et al. 1998. Modulation of IgE reactivity of allergens by site-directed mutagenesis: potential use of hypoallergenic variants for immunotherapy. *FASEB J.* **12**:231-242.
39. Sicherer, S.H., Burks, A.W., and Sampson, H.A. 1998. Clinical features of acute allergic reactions to peanut and tree nuts in children. *Pediatrics.* **102**:6-16.