

Cloning and sequencing of a gene encoding a 2S albumin seed storage protein precursor from English walnut (*Juglans regia*), a major food allergen

Suzanne S. Teuber, MD,^{a, b} Abhaya M. Dandekar, PhD,^c W. Richard Peterson, BS,^a and Claudia L. Sellers, MS^d Davis, Pleasant Hill, and Berkeley, Calif.

Background: Walnuts rank third in per capita consumption of tree nuts in the United States and can be associated with systemic IgE-mediated reactions in some individuals.

Objective: The objectives of the study were to clone a gene encoding one of the major food allergens in the walnut kernel and to characterize the recombinant allergen.

Methods: A cDNA expression library in the lambda vector Uni-ZAP, which was prepared from walnut somatic embryos, was screened by using a patient's sera that reacted with multiple protein bands on immunoblotting.

Results: A cDNA clone containing an insert of 663 bp was identified and named Jug r 1. DNA sequence analysis of this clone revealed that it encoded a protein 142 amino acids in length. Comparison of the encoded protein sequences with protein databases revealed that this clone exhibits a 46.1% identity with the Brazil nut (*Bertholletia excelsa*) methionine-rich 2S albumin seed storage protein precursor, Ber e 1. Jug r 1 appears to be an important walnut food allergen; 12 of 16 sera from patients allergic to walnuts demonstrated IgE binding to the 2S albumin seed storage protein precursor fusion protein. An IgE-binding inhibition study suggests that the walnut 2S protein precursor undergoes posttranslational modification into a large and small subunit that is similar to castor seed, cottonseed, mustard seed, and Brazil nut 2S seed storage protein allergens. Interestingly, the gene encoding this allergenic protein in Brazil nuts has recently gained notoriety because of its experimental use as a transgene to enhance the nutritional quality of legumes.

Conclusion: This is now the sixth definitive 2S albumin seed storage protein demonstrated to bind IgE, suggesting that this class of proteins is inherently allergenic. (J Allergy Clin Immunol 1998;101:807-14.)

Key words: 2S albumin seed storage protein, walnut, food allergy, nut allergy

IgE-mediated type I hypersensitivity to foods affects approximately 1% to 3% of the adult population and up to 5% to 8% of the pediatric population.¹ Allergies to

Abbreviations used

GST:	Glutathione-S-transferase
IPTG:	Isopropyl thiogalactose
PBS:	Phosphate-buffered saline
SDS-PAGE:	Sodium dodecylsulfate polyacrylamide gel electrophoresis

milk, egg, wheat, and soy in infants are usually outgrown by age 3 years, but allergy to fish, crustaceans, peanut, and tree nuts are usually permanent and sometimes life-threatening.² The major allergens of cod, shrimp, soy, peanut, milk, and egg have been described, and in some cases the genes encoding these proteins have been cloned.³⁻¹⁰ In contrast, little is known about the allergens involved in tree nut allergy.

California produces 99% of the nation's 213,000 metric tons of English walnuts.¹¹ Walnuts, at 0.47 lbs/person, are a close third to almonds (0.55 lbs) and pecans (0.48 lbs) in per capita consumption in the United States, with much of the total crop being exported.¹² Clinically, patients who are allergic and have anaphylaxis to one tree nut are often allergic to several others. At the molecular level, this relationship is hypothesized to be due to epitope homology between the storage proteins. Previous crossed-radio immunoelectrophoresis studies have shown significant cross-reactivity at the IgE binding level among different tree nut-protein extracts.¹³

Angiosperm seed storage proteins, with the exception of cereals, consist of albumins and globulins. Albumin proteins are soluble in water and low salt concentrations, whereas globulins are soluble in high salt concentrations.¹⁴ Seed proteins have been categorized on the basis of their solubility and sedimentation properties throughout the history of their study.

2S albumin proteins are small globular proteins rich in arginine, glutamine, asparagine, and often cysteine.¹⁵ They are typically heterodimers composed of a large and small subunit joined by one or two disulfide bridges. Posttranslational processing of the precursor into the two subunits occurs at the N-terminus, internally, and at the C-terminus.^{16, 17} We report the cloning and sequencing of a 2S albumin seed storage protein precursor, Jug r 1, from a walnut cDNA library and its sequence

From ^athe Department of Internal Medicine, University of California, Davis, School of Medicine, Davis; ^bVeterans Affairs Northern California Health Care System, Pleasant Hill; ^cthe Department of Pomology, University of California, Davis; and ^dthe Department of Plant Biology, University of California, Berkeley.

Supported in part by National Institutes of Health Clinical Nutrition Research Unit grant no. DK 35747 and an International Life Sciences Institute Allergy and Immunology Institute New Investigator Award. Received for publication Jan. 26, 1998; revised Feb. 6, 1998; accepted for publication Feb. 10, 1998.

No reprints will be available from the author.

Copyright © 1998 by Mosby, Inc.

0091-6749/98 \$5.00 + 0 1/1/89528

TABLE I. Characteristics of patients and patient sera of persons with severe, systemic reactions to walnut kernels

ID no.	Age	Age at onset	Sex	Last reaction to walnut	ImmunoCAP (class)	Percent ImmunoCAP activity absorbed out by rJug r 1*	Reactivity on Western immunoblotting to rJug r 1
1	21	child	F	1993	3†	25	+
2	25	10 yrs	F	1991	3	50	+
3	36	7 yrs	F	1979	2†	63	+
4	55	7 yrs	M	1987	5	84	+
5	50	10 yrs	F	1981	4	0	0
6	21	5 yrs	M	1989	3	0	0
7	19	2 yrs	F	1991	5	47	+
8	50	2 yrs	F	1995	5	64	+
9	34	18 mos	F	1993	4	73	+
10	28	2 yrs	F	1994	3	64	+
11	22	6 yrs	F	1992	3	48	+
12	37	2 yrs	F	1991	5	47	+
13	34	7 yrs	F	1988	2	0	0
14	38	2 yrs	M	1991	3	56	+
15	38	15 yrs	F	1975	3	0	0
16	9	2 yrs	M	1994	2	25	+

*Percent absorption calculated by the following formula: (baseline fluorescent counts – counts after incubation of sera with rJug r1)/(baseline counts). If absorption was below 5%, it was reported as 0 because this is within the reproducibility range of the test.

†Sera drawn and tested in 1993, all others drawn and tested end of 1995 or early 1996. Shown is the Pharmacia ImmunoCAP class level to English walnut kernel from 0 to 5; 5 represents the higher level of specific IgE (international unit scoring method not available from this lab at that time).

homology with other 2S albumin precursor proteins, including Brazil nut proteins.

METHODS

Human sera

The study was approved by the Human Subjects Review Committee at the University of California, Davis. Sera used in this study were collected from 16 patients with a clear, convincing history of life-threatening systemic allergic reactions to walnuts. Such reactions included hypotension or laryngoedema with or without nausea/vomiting, generalized urticaria, and asthma. Patients did not undergo double-blind, placebo-controlled food challenges because of the history of a severe reaction. Patients underwent commercial (Pharmacia ImmunoCAP; Pharmacia, Inc., Diagnostics, Columbus, Ohio) specific IgE testing in our university clinical laboratory against English walnut kernel (Table I). Age of onset was less than 10 years in all but one patient (age 15) and 10 of 16 had had accidental exposure to walnut in the last 5 years. Skin testing on these patients with a history of systemic reactions was not approved by the Committee. Indeed, two of the patients reported past systemic reactions on skin prick testing with walnut extract that required treatment with epinephrine. Control sera were obtained from atopic individuals with allergic rhinoconjunctivitis and high specific IgE (Class 4 or above) to multiple grass pollens, weeds, tree pollens, or a combination thereof, but without any history of walnut food allergy.

Construction and screening of a cDNA expression library

Standard recombinant DNA methodologies were used to construct a cDNA library from walnut somatic embryos. The SU2 somatic embryo line used here was derived from immature cotyledons of open pollinated seed of the walnut cultivar

Sunland, and it has been maintained in culture on hormone-free DKW medium¹⁸ for many years by using in vitro culture procedures previously described.¹⁹ Total RNA was extracted from these somatic embryos by a modification of earlier procedures,²⁰⁻²² and the mRNA was separated.²³ Synthesis and cloning of cDNA was accomplished with a kit from Stratagene. The double-stranded cDNA was cloned directionally into the lambda vector Uni-ZAP, packaged in vitro (Gigapak Gold; Stratagene, La Jolla, Calif.), and transfected into *Escherichia coli* XL1-Blue cells for determination of the percentage of recombinant clones.

The library was amplified and then screened by incubating with *E. coli* (strain Y1090) for 15 minutes at 37° C and then plating in a layer of Luria-Bertani soft agar for approximately 5 hours at 42° C at a density of approximately 20,000 plaques per plate. A nitrocellulose filter soaked in 10 mmol/L isopropyl thiogalactose (IPTG) was dried, placed on the agar plate, and further incubated 10 to 14 hours at 37° C. The filter was placed in blocking solution (4% nonfat milk in phosphate-buffered saline [PBS], pH 7.4) for 1 hour, and primary sera from a patient who reacted to multiple bands on immunoblotting was diluted 1:5 in 4% nonfat milk and allowed to incubate at 4° C overnight with the filter. ¹²⁵I-antihuman IgE (Sanofi Diagnostics/Pasteur, Chaska, Minn.) was added after washing and allowed to incubate overnight. The filter was exposed for 96 hours to Kodak XRT film and developed. Positive plaques were purified to homogeneity. One of the clones, designated Jug r 1 by following the guidelines of the International Union of Immunological Societies' Allergen Nomenclature Subcommittee,²⁴ is further described herein.

Excision and subcloning of cDNA

Recombinant phagemids containing the cDNA were excised from the lambda Uni-ZAP phage by using the ExAssist helper

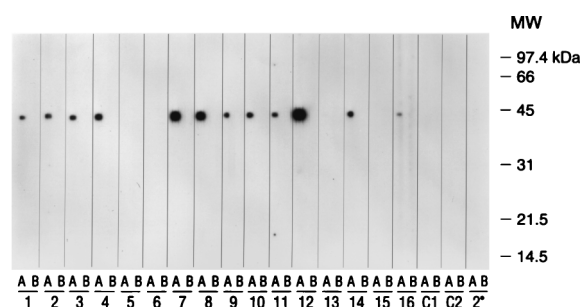


FIG. 1. Western immunoblotting of purified rJug r 1 fusion protein and control protein was performed against sera from 16 patients with systemic reactions to walnut. Numbers 1 to 16 represent incubation with sera from these patients. Lane A in each pair of strips represents purified rJug r 1 fusion protein. Lane B is purified recombinant pyruvate dehydrogenase, prepared in same manner. Shown also are representative strips with sera from two control patients (C1 and C2) and secondary antibody alone (buffer control) at 2°. Strips were exposed to film for 3 days; 12 of 16 patient sera showed IgE binding to fusion protein at approximately 42 kd. An additional patient sera (number 13) showed an extremely light band at 72 hours, which was only faintly enhanced at 17 days; this patient sera is considered negative. Note that there is no IgE binding to recombinant control protein (Lane B), showing that IgE is specifically binding to rJug r 1 and not to GST. After prolonged 17-day exposure, no binding was evident in Lane B, and controls continued to be negative.

phage (Stratagene) in *E. coli* strain SOLR (Stratagene). Plasmid DNA (pBluescript) was isolated by using the alkaline lysis method from individual colonies.²⁵ The cDNA insert of interest was isolated by a double digest with Eco R1 and Xho 1. On agarose gel electrophoresis, the cDNA insert was approximately 660 bp. For later ease in isolation of the rJug r 1 fusion protein, the cDNA insert was then subcloned into the pGEX vector (Pharmacia Biotech, Piscataway, N.J.). The pGEX plasmid expression vector results in the synthesis of foreign proteins as a fusion product with a 26 kd glutathione-S-transferase (GST).

Expression and isolation of fusion protein

Purified fusion protein, rJug r 1, was isolated following the protocol by Frangioni and Neel.²⁶ Briefly, transformed protease-deficient *E. coli* strain BL21 (Pharmacia Biotech) was grown in Luria-Bertani broth with ampicillin at 37° C until the optical density reached 0.6 to 0.9 at 600 nm. IPTG was added to a final concentration of 1.0 mmol/L, and incubation was continued at 37° C. After 3 hours, cells were centrifuged and resuspended in STE (10 mmol/L Tris, pH 8.0, 150 mmol/L NaCl, and 1 mmol/L ethylenediamine tetraacetic acid). Lysozyme was added to 100 µg/ml and incubated on ice for 15 minutes. Then dithiothreitol was added to 5 mmol/L. The bacteria were lysed by the addition of N-laurylsarcosine to 1.5% and then sonicated for 1 minute on ice. The lysate was centrifuged for 5 minutes at 4° C at 10,000 g. The supernatant was removed, Triton X-100 was added to 2% and vortexed, and the fusion protein was isolated by using the Bulk GST Purification Module (Pharmacia Biotech). To check the reactivity and specificity of the purified fusion protein with patient and control sera, approximately 1 µg of fusion protein was loaded per lane for immunoblotting as described below. An irrelevant GST fusion protein, recombinant human pyruvate dehydrogenase, purified in a similar manner was likewise loaded at 1 µg/lane as a control.

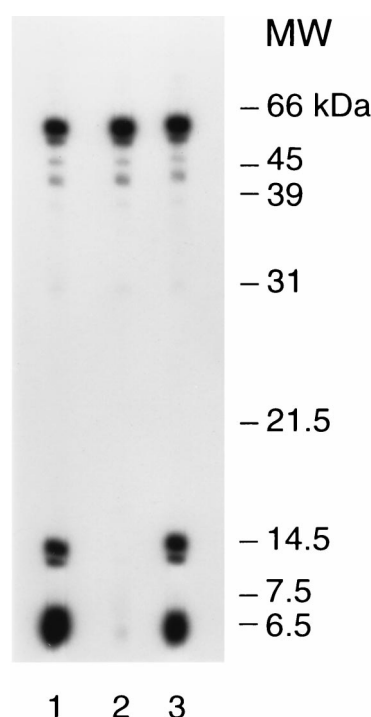


FIG. 2. Western immunoblotting against walnut kernel extract in reducing conditions is shown. Sera from a representative patient who demonstrated IgE binding to fusion protein in Fig. 1 was used. Lane 1, sera diluted 1:5 in PBS/milk alone; lane 2, sera preincubated and diluted 1:5 with PBS/milk and rJug r 1 fusion protein; and lane 3, sera preincubated and diluted 1:5 with PBS/milk and recombinant pyruvate dehydrogenase. Note that bands at about 14, 10, and 5 kd are removed by absorption against rJug r 1, suggesting that native Jug r 1 consists of intact protein at approximately 14 kd and subunits at about 10 and 4 kd, analogous to other 2S albumin seed storage proteins.

Walnut protein extraction

Juglans regia, cultivar Chandler, nuts were obtained from the Pomology Department at the University of California, Davis. Kernels (10 to 12 gm) were initially blended in a Waring Blender (Waring Products Division, New Hartford, Conn.) for 20 to 30 seconds, and 5 to 6 gm of the minced walnut meats were each placed into 50 ml tubes with 15 ml of 0.05 mol/L ammonium bicarbonate, pH 7.8, at 4° C. The tubes were placed on ice, and the kernels were subjected to further pulverization in a polytron (Polytron PT-MR 3000; Brinkmann, Kenemata Ag, Littau, Switzerland) for 10 minutes and then rocked at 4° C for 48 hours. The mixtures were then centrifuged at 45,000 g overnight at 4° C, and the lipid layer was removed. The samples were again centrifuged at 45,000 g for 1 hour, and the lipid layer was completely removed. The aqueous layer was retrieved, placed in aliquots, and frozen at -60° C.

Protein assay

Protein determinations were made with the Bradford Protein Assay kit (Biorad Laboratories, Hercules, Calif.).

Western immunoblotting

Samples were boiled for 5 minutes in sample buffer (50 mmol/L Tris [pH 6.8], 10% glycerol, 2% sodium dodecylsulfate

```

1  TAGCAGCTCTCCTTGTAGCCCTTTTGTTCGTAGCCAACGCCGCTGCCTTCCGCACCACCA  60
   A A L L V A L L F V A N A A A F R T T I -

61  TAACAACCATGGAGATCGACGAGGACATTGACAACCCAAGGCGGCGAGGCGAGGGCTGCC  120
   T T M E I D E D I D N P R R R G E G C R -

121  GGGAGCAGATTCAACGCCAGCAAAACCTCAACCAGTCCAGTACTACCTGAGGCAGCAAA  180
   E Q I Q R Q Q N L N H C Q Y Y L R Q Q S -

181  GCCGGTCAGGGGGTTATGATGAAGACAACCAGCGGCAGCATTTCAGGCAGTGTGCCAGC  240
   R S G G Y D E D N Q R Q H F R Q C C Q Q -

241  AGCTGAGTCAGATGGACGAGCAGTGCAGTGTGAGGGGTTGAGGCAGGTGGTGAGGCGGC  300
   L S Q M D E Q C Q C E G L R Q V V R R Q -

301  AGCAGCAGCAGCAGGGTCTCCGGGGTGAGGAAATGGAGGAGATGGTGCACTCTGCTAGGG  360
   Q Q Q Q G L R G E E M E E M V Q S A R D -

361  ACTTGCCAAATGAATGCGGTATCAGCAGCCAGCGCTGTGAAATTCGAGAAGCTGGTCTCT  420
   L P N E C G I S S Q R C E I R R S W F

421  AGATCAGGACATCATGCACCATAATAAGGGTCCCTGTTAGCTCTAGCATTGTGCTGCTGA  480

481  AGATAACTGGTGACTAGCTTTAGCAACTAGCAAGCACATCCTCGTGTGTAATGCAATATG  540

541  TATCCTGGCCTTTAAGTAGTTTAGGGCCTTTGTATGTAGGCTTGCTTAAGTAAAAATAAAA  600

601  ACTTGTTAAATAAAATGCCTCGTCTTGCTAATTAAAAAATAAAAAA  649

```

FIG. 3. Nucleotide sequence and deduced amino acid sequence for Jug r 1. Open reading frame encodes polypeptide 139 amino acids in length. The nucleotide sequence has been submitted to GenBank under accession number U66866.

[SDS], 100 mmol/L dithiothreitol, and 0.0025% bromophenol blue). Samples were run on an SE 600 Vertical Slab Gel Unit (Hoefel Scientific Instruments, San Francisco, Calif.) at a constant current of 30 mA. A 15% SDS-polyacrylamide Laemmli gel was used. Proteins were transferred to 0.22 μ m nitrocellulose membranes (Schleicher & Schuell, Keen, N.H.) overnight at 35 V by using a TE 42 Transphor Electro-Transfer Unit (Hoefel Scientific Instruments). The blots were blocked in a mixture of PBS and 3% nonfat dry milk for 30 minutes and then incubated overnight at room temperature in sera diluted 1:5 in PBS/milk. The following day, strips were washed for 20 minutes three times in PBS/0.05% Tween-20 at room temperature and incubated overnight with ^{125} I-antihuman IgE (Sanofi Diagnostics/Pasteur). The strips were again washed, dried, and exposed to x-ray film for 18 to 72 hours at -70° C.

Absorption of sera with rJug r 1 polypeptide and ImmunoCAP against English walnut

The purified fusion protein containing rJug r 1 obtained above was used to preabsorb sera before standard commercial ImmunoCAP (Pharmacia) against English walnut kernel. Sera were prepared as follows. Forty microliters of sera was incubated overnight at room temperature with either 20 μ l of PBS (diluent control to bring samples up to the same volume of 60 μ l for determination of baseline fluorescent counts), 18 μ l of PBS plus 2 μ l walnut extract (10 μ g protein, the amount predetermined to maximally absorb out the IgE activity of this volume of sera), or 10 μ l of PBS plus 10

μ l of the rJug r 1 purified fusion protein (the amount predetermined in test runs for this volume of sera to reach maximum absorbance of activity on the ImmunoCAP). Sera were then submitted for standard ImmunoCAP assay against English walnut kernel, and the results were obtained in counts of fluorescence. The percent absorbance was calculated by the following formula: (baseline counts – counts after preabsorbance with rJug r 1 fusion protein)/(baseline counts).

Absorption of human sera with rJug r 1 polypeptide and immunoblotting against native walnut kernel extract

Bacterial lysates of rJug r 1 were prepared following the protocol in the GST Gene Fusion System Manual (Pharmacia Biotech). Briefly, IPTG-induced 100 ml cultures were pelleted, resuspended in 5 ml PBS, and sonicated. Triton X-100 was added to 1% and mixed gently for 30 minutes. After centrifuging at 12,000 g for 10 minutes at 4° C, 250 ml of the supernatant containing the fusion protein was added to human sera diluted 1:5 in PBS/3% nonfat milk for a total volume of 500 μ l and incubated overnight at room temperature. The negative control protein, recombinant pyruvate dehydrogenase fusion protein (an irrelevant protein to screen for nonspecific binding to GST or *E. coli* contaminants), was obtained from the pGEX vector and prepared and incubated with sera by using identical procedures. Immunoblotting against *J. regia* kernel cultivar Chandler was carried out as described above.

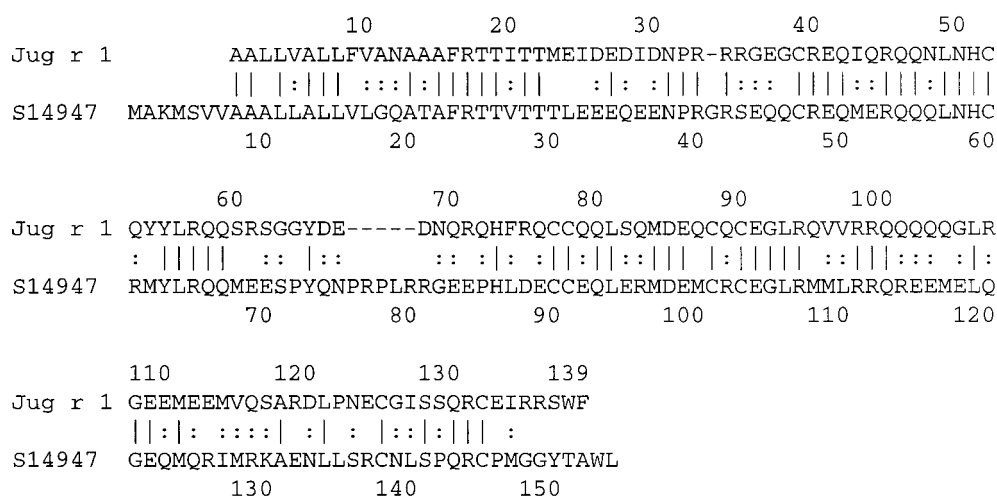


FIG. 4. Alignment of deduced amino acid sequence of Jug r 1 precursor with Brazil nut Ber e 1 precursor²⁴ (Protein Identification Resource accession number S14947); 46.1% identity exists between them. Note that all eight cysteine residues are preserved but that walnut 2S albumin does not have significant methionine content; only two of 13 methionine residues seen in Brazil nut precursor are seen in the walnut polypeptide.

Nucleotide sequencing

The cDNA was sequenced by using the Sequenase Quick-Denature Plasmid Sequencing Kit (Amersham, Arlington Heights, Ill.) on both strands. Nucleotide and protein databases were searched by using GCG programming (GCG programming, Sequence Analysis Software Package; Genetics Computer Group, University of Wisconsin Biotechnology Center, Madison, Wis.).

RESULTS

Identification of an IgE-reactive clone from *J. regia* cDNA library

The primary library from *J. regia* embryo had approximately 1.5×10^6 pfu/ml and was amplified to produce high-titer phage lysates (9.5×10^9 pfu/ml). Analysis of the library revealed that 98% were recombinant, indicating that they contained inserts, presumably cDNA. Three clones were identified by screening one plate of 20,000 plaques. One clone has been designated and submitted as Jug r 1 to the International Union of Immunological Societies' Allergen Nomenclature Subcommittee, and the others are yet to be characterized. Digestion of the plasmid DNA by Eco R1 and Xho I produced an approximately 660 bp cDNA fragment on agarose gel (data not shown).

Expression of rJug r 1

The Jug r 1 cDNA was subcloned into the frame-shifted plasmid expression vector pGEX 2. Fusion protein was induced and purified by glutathione agarose, and the purified fusion protein was run on sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and detected as a protein band at approximately 42 kd. Because GST has a molecular weight of 26 kd, the deduced molecular weight of rJug r 1 was estimated at 15 to 16 kd. Immunoblotting of

the purified fusion protein rJug r 1 versus a control fusion protein (recombinant human pyruvate dehydrogenase) with sera from 16 patients with walnut food allergy showed that the fusion protein bound IgE in 12 of 16 patients (75%), indicating that Jug r 1 is a major allergen (Fig. 1). One of the patient sera considered negative (number 13) did have trace IgE binding that showed up only weakly at 72 hours and was thus considered negative compared with the others. Confirming that we had cloned a potential food allergen, none of the control atopic sera or the secondary antibody control alone showed detectable IgE binding to rJug r 1. No binding was evident to the recombinant control fusion protein, suggesting that patient IgE was binding specifically with epitopes present on the rJug r 1 part of the fusion protein and not binding to GST. When the strips were exposed to film for 17 days, the atopic control sera still did not show binding to the antigen at 42 kd (data not shown).

Absorption of sera with rJug r 1 and ImmunoCAP against English walnut

This assay was performed to assess the relative importance of the allergen to the total IgE response against walnut kernel in each patient. The results are shown in Table I and ranged from 0% ImmunoCAP reactivity absorbed out to 84% absorbance. The four patient sera that had negative results on Western immunoblotting to rJug r 1 were all negative on the absorption screen as expected. Of the other 12 patients, the percent activity absorbed out by rJug r 1 was 50% or greater in seven of them, indicating that this protein most likely represents an important allergen in terms of the quantitative IgE response.

Jug r 1	 AAL	L.VALLFVAN	AA.AFRTTIT	TMEIDEDIDN	PRRRGE.GCR
Ber e 1	MAKMSVVAAA	L.LALLVLGQ	AT.AFRTTIT	TLEEEQEEN	PRGRSEQQCR
Castor	MAKLIPTIAL	VSVLLFIIAN	ASFAYRTTIT	TIEIDESKGE	REGSSSQQCR
Cotton	MAKLAVYLAT	LALILF.LAN	ASI....TSV	TVESEENRDS	 CE
Sin a 1				P	AGPFRIPKCR
Jug r 1	EQIQROQNLN	HCQYYLRQOS	RSGGYDED..	...NQR....	..QHFRQCCQ
Ber e 1	EQMERQQQLN	HCRMYLRQQM	EESPYQNPRP	LRRGEE....	..PHLDECCQ
Castor	QEVQR.KDLS	SCERYLRQSS	SRSPGEEVL	RMPGDENQQQ	ESQQLQCCCN
Cotton	QQIRKQAHLK	HCQKYMEEL	GGEKSDN...	I	AGGYIDS CCQ
Sin a 1	KEFQQAQHLR	ACQOWLHKQA	MQSGSG....	..PSPQGPQQ	RPPL LQCCCN
Jug r 1	QLSQMDEOCQ	CEGLRQVRR	QQQ..QQGL.	.RGEEMEMV	QS....ARDL
Ber e 1	QLERMDEMCR	CEGLRMMLRR	QRE..EMEL.	.QGEQMORIM	RK....AENL
Castor	QVKQVRDECQ	CEAIKYIAED	QIQ..QGQL.	.HGEESERVA	QR....AGEI
Cotton	QLEKMDTQCR	CQGLRHATMQ	QMOMQGM.	.GSKQMRIM	QKV...TKKI
Sin a 1	ELHQEEPLCV	CPTLKGASKA	VKQQVRQGLE	QQGQQGPHVI	SRIYTATHL
Jug r 1	PNECGISSQR	 CEIR.	.RSWF.....		
Ber e 1	LSRCNLSPQR	CPMGG	YTAWL.....		
Castor	VSSCGVRCMR	QTRTNPSQOG	CRGQIQEQQN		
Cotton	MSECEMEPGR	CDTPS	.RSLI.....		
Sin a 1	PKVCNIPQVS	V....CPFKK	TMPGPS....		

FIG. 5. Alignment of deduced amino acid sequence of Jug r 1 precursor with Brazil nut Ber e 1 precursor,²⁹ castor seed 2S albumin precursor,³⁰ cottonseed 2S albumin precursor,³¹ and small and large subunits (not precursor peptide) of mustard seed allergen Sin a 1.³² Identity with Jug r 1 is indicated by bold type. Identity between all five sequences is noted by underlining. Note preservation of cysteine residues. Although only 46.1% identity exists between the Jug r 1 and Ber e 1 precursors, the pairwise similarity between two peptides is 82% (GCG Sequence Analysis Software Package). Pairwise similarity between Jug r 1 precursor and castor seed precursor is 69%, with cottonseed precursor at 59% and mustard seed 2S albumin at 45%.

Absorption of sera with rJug r 1 and immunoblotting against native walnut kernel extract

Absorption experiments were also carried out to show that we had cloned a gene encoding an allergen present in the mature walnut kernel that could be ingested by allergic patients, rather than an irrelevant IgE binding protein from walnut somatic embryo. Preabsorption of patient sera with rJug r 1 abolished IgE binding on Western immunoblotting (reducing conditions) to proteins from mature walnut kernel at approximately 14, 10-12, and 5 kd, even at the low 1:5 sera dilution used (Fig. 2). On the basis of sequence similarities to Brazil nut and other 2S albumin seed storage proteins, we infer that the native Jug r 1 allergen consists of a large subunit of about 10 kd and a small subunit of about 4 to 5 kd. To reconcile this with the data from the immunoblot, we are possibly seeing an intact 14 kd heterodimeric protein incompletely reduced because of partial inaccessibility of the disulfide bonds or reassociation after reduction, as well as the large and small subunit components. It is, however, often difficult to assign definite molecular weights in the lower ranges in this gel system. Therefore it is also possible that we are seeing the large subunit, an isoform of the large subunit, and the small subunit. On a nonreducing immunoblot, we see an indistinct blur of IgE binding at about 14 kd (data not shown). Isolation of the separate bands on higher resolution gels and N-terminal sequencing will be needed to exactly determine the molecular weights and identification of each band. Absorption of sera with an irrelevant GST fusion protein did not reduce IgE binding to any walnut kernel anti-

gens. Note that the three-band Jug r 1 complex is not the only allergen recognized by this serum from a patient allergic to walnuts, which exhibits a typical pattern of banding.

Nucleotide and deduced amino acid sequence of Jug r 1

The nucleotide and deduced amino acid sequences are shown in Fig. 3. A Protein Identification Resource data bank search revealed that Jug r 1 encodes a 2S albumin seed storage protein precursor 139 amino acids in length with significant homology to the allergenic methionine-rich 2S albumin seed storage protein precursor from Brazil nuts, Ber e 1, which has recently gained attention because of transgenic expression in soybean.²⁷ Alignment of the cloned walnut Jug r 1 precursor with the Ber e 1 precursor is shown in Fig. 4, revealing 46% identity.

It must be noted that these are precursor proteins. Described 2S albumin seed storage proteins undergo significant posttranslational modification, with most undergoing cleavage into a large and small subunit held together by interchain disulfide bonds at the cysteine residues. The deduced molecular weight of the precursor, based on an estimate from the sequence itself and SDS-PAGE of the fusion protein, is 15 to 16 kd. Precursor proteins are not detectable on routine extraction of mature nuts and seeds because processing occurs in the cell just after translation before the 2S albumin is compartmentalized into seed storage protein bodies. As an example, in Brazil nuts the precursor protein is 15 kd, but the final intact 2S albumin obtained from nut extract

is 13 kd, with large and small subunits extremely similar to what we believe is occurring in walnuts.²⁸ Fig. 5 shows the amino acid sequence comparison between the walnut 2S albumin precursor and the 2S albumin allergens from Brazil nuts, castor seed, cottonseed, and mustard seed. All are shown as protein precursors except mustard seed, in which the sequence was obtained by N-terminal sequencing of each native protein subunit. A key characteristic of the sequences is the preservation of the cysteine residues. Further studies making use of N-terminal sequencing of subunits will need to be performed to determine the exact location of the signal peptide, discard peptides, and actual subunits within the precursor in *J. regia*.

DISCUSSION

Two seed storage proteins have been known to be allergenic for over 50 years,^{33, 34} and both were shown to be 2S albumins in the 1970s by Youle et al.^{35, 36} First, the 2S albumin seed storage protein obtained from castor seed was shown to be a major aeroallergen sensitizer around castor-oil mills worldwide.^{35, 37, 38} Second, the cottonseed 2S albumin seed storage protein was found to be allergenic on ingestion or on inhalation in some individuals.^{34, 36, 39, 40} The yellow mustard seed 2S albumin precursor has also been cloned and sequenced, and interestingly, it is a major allergen in mustard seed.⁴¹ The same group has also found that the oriental mustard (*Brassica juncea*) food allergen, Bra j 1, is also a 2S albumin.⁴² Recently, Nordlee et al.²⁷ showed that the 2S albumin seed storage protein from Brazil nuts could bind IgE from patients with Brazil nut allergy and appeared to be the predominant protein responsible for IgE binding in their series of nine patients. However, not all of their patients had ever eaten Brazil nuts but had avoided all tree nuts after an initial tree nut allergy was discovered (offending nuts not identified in the report). The peanut allergen, Ara h 2, a 17 kd and 18 kd protein that does not appear to be made up of subunits like Jug r 1 or Ber e 1, was also recently reported to have homology (30%) to 2S albumins, but exhibited the closest homology with conglutin- δ from lupin.⁴³

Thus for many years, 2S albumin seed storage proteins have been associated with IgE responses in the literature. It is not surprising that one of the allergens from walnut kernel is also a 2S albumin protein. Additionally, a recent report of occupational asthma caused by sesame seed exposure revealed the allergen to be a probable 2S albumin on the basis of its SDS-PAGE appearance of an intact protein at approximately 14 kd and a large and small subunit.⁴⁴

The 2S albumins are unique among seed proteins in terms of the higher sulfur content in some species compared with other storage proteins. This is due to higher levels of cysteine and of the essential amino acid methionine in some seeds, although not in legumes.¹⁵ Genetic selection for seeds enriched in this storage protein has been suggested as a favorable goal that could enhance human nutrition under marginal circumstances,

but traditional breeding methodologies have only slightly increased methionine content in legumes.¹⁵ The methionine content of some 2S albumins, however, is tremendously high (e.g., Brazil nuts), which has made the Brazil nut 2S albumin a target for genetic engineering. Indeed, the nutritional quality of several different legumes, including soybean, has been experimentally enhanced through transgenic expression of the Brazil nut 2S albumin protein.²⁷ Apparently it did not come as a surprise to those involved in 2S albumin research that the Brazil nut 2S albumin bound IgE and was a probable major food allergen. Indeed, as prophetically cautioned by Youle and Huang in 1981,¹⁵ "the allergenic properties of the 2S proteins deserve consideration in the nutritional evaluation of this class of seed protein."

It is now known that two tree nut allergens are 2S albumin seed storage proteins with significant homology. Brazil nuts and walnuts are not closely related among the tree nuts; more significant homology is expected to be seen with pecan, which is the only commercially important nut in the same order as *J. regia*: Juglandales (black walnuts and butternuts are in this order also). Although few studies exist, Gillespie et al.⁴⁵ showed in 1976 that black walnuts significantly inhibited a pecan RAST by 80%, followed by Brazil nuts at 49%, peanuts at 48%, and almonds with no inhibition at 0% in one patient who reported allergy to many nuts. Only two of our 16 subjects had ever eaten Brazil nuts, and anaphylactic reactions were reported by both. One of those patients was patient number 6, who did not react to the recombinant 2S albumin protein precursor and whose serum did not bind to the 14, 10, or 5 kd antigens on Western immunoblotting of walnut kernel extract (the serum binds a protein at approximately 34 kd, data not shown). As more of the seed storage proteins are elucidated and the patterns of reactivity to tree nuts among various patients are mapped, it is expected that logic will prevail as to why some individuals are sensitive to many tree nuts and others to only one or two.

It is becoming apparent that the 2S albumin storage proteins can be viewed as universal allergens among seeds in which they occur, although they are not necessarily cross-reactive, and that they are associated with IgE responses through airborne and oral exposure. 2S albumins were shown to be abundant in 12 genetically diverse seeds based on sedimentation sucrose gradient centrifugation, ranging from a low value of 20% of total seed protein in peanut to 62% in oilseed rape and sunflower.¹⁵ Of interest, peanut 2S albumins were noted to be uniquely low in cysteine and methionine.¹⁵ Given the broad allergenicity of this protein class, IgE-epitope mapping and T-cell epitope research may provide further insight into common motifs involved in IgE recognition and induction in general.

We thank Dr. Bob Buchanan for helpful discussions and Tom Kenny for assistance with sequence analysis.

REFERENCES

- Zeiger RS. Prevention of food allergy in infancy. *Ann Allergy* 1990;65:430-41.
- Schreiber RA, Walker WA. Food allergy: facts and fiction. *Mayo Clin Proc* 1989;64:1381-91.
- Elsayed S, Bennich H. The primary structure of allergen M from cod. *Scand J Immunol* 1975;4:203-8.
- Leung PSC, Chu KH, Chow WK, Ansari A, Bandea CI, Kwan HS, et al. Cloning, expression, and primary structure of *Metapenaeus ensis* tropomyosin, the major heat-stable shrimp allergen. *J Allergy Clin Immunol* 1994;94:882-90.
- Burks AW, Brooks JR, Sampson HA. Allergenicity of major component proteins of soybean determined by ELISA and immunoblotting in children with atopic dermatitis and positive soy challenges. *J Allergy Clin Immunol* 1988;81:1135-42.
- Burks AW, Williams LW, Helm RM, Connaughton CA, Cockrell G, O'Brien TJ. Identification of a major peanut allergen, *Ara h* 1, in patients with atopic dermatitis and positive peanut challenges. *J Allergy Clin Immunol* 1991;88:172-9.
- Burks AW, Williams LW, Connaughton C, Cockrell G, O'Brien TJ, Helm RM. Identification and characterization of a second major peanut allergen, *Ara h* II, with use of the sera of patients with atopic dermatitis and positive peanut challenges. *J Allergy Clin Immunol* 1992;90:962-9.
- Bleumink E, Young E. Identification of the atopic allergen in cow's milk. *Int Arch Allergy* 1968;34:521-43.
- Alexander LJ, Hayes G, Pearse MJ, Beattie CW, Stewart AF, Willis IM, et al. Complete sequence of the bovine β -lactoglobulin cDNA. *Nucleic Acids Res* 1989;17:6739.
- Holen E, Elsayed S. Characterization of four major allergens of hen egg white by IEF/SDS-PAGE combined with electrophoretic transfer and IgE-immunautoradiography. *Int Arch Allergy Appl Immunol* 1990;91:136-41.
- Walnut Marketing Board Statistics and Economic Data, 1996. Sacramento (CA): Walnut Marketing Board; 1996.
- Commodity Economics Division of the Economic Research Service. Washington (DC): United States Department of Agriculture; 1995.
- Goetz DW, Whisman BA, Goetz AD. Crossed immunoelectrophoresis identification of nut cross-reactivities [abstract]. *J Allergy Clin Immunol* 1993;91:360.
- Osborne TB. The vegetable proteins. 2nd Ed. London: Longman-Green; 1924.
- Youle R, Huang AHC. Occurrence of low molecular weight and high cysteine containing albumin storage proteins in oilseeds of diverse species. *Am J Botany* 1981;68:44-8.
- Krebbes E, Herdies L, De Clercq A, Seurinck J, Leemans J, Van Damme J, et al. Determination of the processing sites of an *Arabidopsis* 2S albumin and characterization of the complete gene family. *Plant Physiol* 1988;87:859-66.
- D'Hondt K, Van Damme J, Van Den Bossche C, Leejeerajumenean S, De Rycke R, Derksen J, et al. Studies of the role of the propeptides of the *Arabidopsis thaliana* 2S albumin. *Plant Physiol* 1993;102:425-33.
- Driver JA, Kuniyuki AH. In vitro propagation of Paradox walnut rootstock. *Hortscience* 1984;19:507-9.
- McGranahan GH, Driver JA, Tulecke W. Tissue culture of *Juglans*. In: Bonga JM, Durzan DJ, editors. Cell and tissue culture in forestry. Boston: Martinus Nijhoff; 1987. p. 261-71.
- Cathala G, Savouret J-F, Mendez B, West BL, Karin M, Martial JA, et al. Laboratory methods: a method for isolating intact translationally active ribonucleic acid. *DNA* 1983;2:329-35.
- Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WJ. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 1979;18:5294-9.
- Verwoerd TC, Dekker BMM, Hoekema A. A small-scale procedure for the rapid isolation of plant RNAs. *Nucleic Acids Res* 1989;17:2362.
- Aviv H, Leder P. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc Natl Acad Sci* 1972;69:1408-13.
- King TP, Hoffman D, Löwenstein H, Marsh DG, Platts-Mills TAE, Thomas W. Allergen nomenclature. *J Allergy Clin Immunol* 1995;96:5-14.
- Ausubel FM, Brent R, Kingston RE, Moore DD, Smith JA, Seidman JG, et al. Current protocols in molecular biology. Vol 1. New York: John Wiley & Sons, Inc.; 1994. p. 261-71.
- Frangioni JV, Neel BG. Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. *Anal Biochem* 1993;210:179-87.
- Nordlee JA, Taylor SL, Townsend JA, Thomas LA, Bush RK. Identification of a Brazil-nut allergen in transgenic soybeans. *N Engl J Med* 1996;334:688-92.
- de Castro LAB, Lacerda Z, Aramayo RA, Sampaio MJAM, Gander ES. Evidence for a precursor molecule of Brazil nut 2S seed proteins from biosynthesis and cDNA analysis. *Mol Gen Genet* 1987;206:338-43.
- Gander ES, Holmstroem KO, de Paiva GR, de Castro LAB, Carneiro M, Grossi de Sa MF. Isolation, characterization and expression of a gene coding for a 2S albumin from *Bertholletia excelsa* (Brazil nut). *Plant Mol Biol* 1991;16:437-48.
- Irwin SD, Lord JM. Nucleotide sequence of a *Ricinus communis* 2S albumin precursor gene. *Nucleic Acids Res* 1990;18:5890.
- Galau GA, Wang HYC, Hughes DW. Cotton *Mat5-A* (C164) gene and *Mat5-D* cDNAs encoding methionine-rich 2S albumin storage proteins. *Plant Physiol* 1992;99:779-82.
- Menendez-Arias L, Moneo I, Dominguez J, Rodriguez R. Primary structure of the major allergen of yellow mustard (*Sinapis alba* L.) seed, *Sin a* I. *Eur J Biochem* 1988;177:159-66.
- Spies JR, Coulson EJ. The chemistry of allergens: isolation and purification of active protein-polysaccharide fraction CB-1A from castor beans. *J Am Chem Soc* 1943;65:1720-35.
- Spies JR, Bernton HS, Stevens H. The chemistry of allergens. I. Isolation of an active fraction from cottonseed. *J Allergy* 1939;10:113-30.
- Youle RJ, Huang AHC. Evidence that the castor bean allergens are the albumin storage proteins in the protein bodies of castor bean. *Plant Physiol* 1978;61:1040-2.
- Youle RJ, Huang AHC. Albumin storage proteins and allergens in cottonseed. *J Agric Food Chem* 1979;27:500-3.
- Ordman D. An outbreak of bronchial asthma in South Africa, affecting more than 200 persons, caused by castor bean dust from an oil processing factory. *Int Arch Allergy Appl Immunol* 1955;7:10-24.
- Panzani RC. Respiratory castor bean dust allergy in the south of France with special reference to Marseilles. *Int Arch Allergy Appl Immunol* 1957;11:224-36.
- Bernton HS, Spies JR, Stevens H. Significance of cottonseed sensitivity. *J Allergy* 1940;11:138-46.
- Atkins FM, Wilson M, Bock SA. Cottonseed hypersensitivity: new concerns over an old problem. *J Allergy Clin Immunol* 1988;82:242-50.
- Gonzalez de la Pena MA, Villalba M, Garcia-Lopez JL, Rodriguez R. Cloning and expression of the major allergen from yellow mustard seeds, *Sin a* I. *Biochem Biophys Res Commun* 1993;190:648-53.
- Monsalve RI, Gonzalez de la Pena MA, Menendez-Arias L, Lopez-Otin C, Villalba M, Rodriguez R. Characterization of a new oriental-mustard (*Brassica juncea*) allergen, *Bra j* I: detection of an allergenic epitope. *Biochem J* 1993;293:625-32.
- Stanley JS, King N, Burks AW, Huang SK, Sampson H, Cockrell G, et al. Identification and mutational analysis of the immunodominant IgE binding epitopes of the major peanut allergen *Ara h* 2. *Arch Biochem Biophys* 1997;342:244-53.
- Alday E, Curel G, Lopez-Gil MJ, Carreño D, Moneo I. Occupational hypersensitivity to sesame seeds. *Allergy* 1996;51:69-70.
- Gillespie DN, Nakajima S, Gleich GJ. Detection of allergy to nuts by the radioallergosorbent test. *J Allergy Clin Immunol* 1976;57:302-9.