Solution Structure and Stability against Digestion of rproBnIb, a Recombinant 2S Albumin from Rapeseed: Relationship to Its Allergenic Properties[†,‡]

David Pantoja-Uceda,[§] Oscar Palomares,[||] Marta Bruix,[§] Mayte Villalba,[||] Rosalía Rodríguez,[||] Manuel Rico,[§] and Jorge Santoro*[,§]

Instituto de Química Física "Rocasolano", CSIC, Serrano 119, Madrid 28006, Spain, and Departamento de Bioquímica y Biología Molecular, Facultad de Química, Universidad Complutense, Madrid 28040, Spain

Received September 8, 2004; Revised Manuscript Received October 11, 2004

ABSTRACT: NMR spectroscopy has been used to determine the solution structure of the precursor form of the recombinant napin BnIb, rproBnIb, a 2S albumin, 109-residue protein from the seeds of *Brassica napus*. More than 90% of the side-chain proton resonances were unambiguously assigned from the analysis of two-dimensional correlation (COSY), total correlation (TOCSY), and nuclear Overhauser effect (NOESY) spectra. The final structures were computed by using restrained molecular dynamics on the basis of 1316 upper-limit distance constraints derived from NOE cross-correlation intensities. The computed structures exhibited a root-mean-square deviation (RMSD) radius of 0.66 Å for the backbone and 1.16 Å for the side-chain heavy atoms of the structural core. The resulting structure consists of five amphipathic helices arranged in a right-handed super helix, a folding motif found in other proteins of the prolamin superfamily. As in the case of the mature protein, the recombinant precursor behaves as a plant food allergen. To trace out the origin and characteristics of its allergenic properties, rproBnIb was assayed against simulated gastric fluid and found to be very resistant to proteolysis. Also, heat treatment of the protein followed up to 85 °C by circular dichroism showed a very limited unfolding, which was recovered after cooling to 20 °C, indicating a high thermal stability. These results suggest that rproBnIb, as other 2S albumins, may be able to reach the gut immune system intact. A comparison of the putative epitopes against IgE antibodies of the three members of the prolamine family [2S albumins, nonspecific lipid transfer proteins (nsLTPs), and α-amylase/trypsin inhibitors] indicates that there are not common surfaces of interaction with IgE. Though the epitopes appear to be located in different regions of the proteins, they do comply with the requirements of being solvent-exposed and flexible.

The 2S albumin class of proteins constitutes the major seed storage protein group in *Brassica napus*, representing about 20% of the total protein content in mature rape seeds. These proteins represent an excellent model for studying both the expression of plant genes and the protein maturation processes in plant cells (*1*). The interest in Brassicaceae seeds has increased because they represent one of the most important resources for animal nutrition and industrial oil and are susceptible to genetic manipulation. Several functions or activities have been assigned to this family of proteins: nitrogen and sulfur storage, antifungal activity, and calmodulin antagonist capability, among others. 2S Albumins from several species such as mustard, castor bean, Brazil nut, English walnut, sunflower, and peanut have been shown to be type I allergy inducers of remarkable incidence, suggesting that this family of storage proteins is intrinsically allergenic (*2*, *3*). The importance of 2S albumins as allergens is increasing because many of these seeds are frequently hidden as condiment in manufactured foods (*4*, *5*).

2S Albumins consist of two different polypeptide chains linked by disulfide bridges. The two chains derive by proteolytic cleavage from a single polypeptide precursor. BnIb (12.7 kDa) is a representative member of a distinct group of rapeseed 2S albumins, referred to as "low molecular weight napins" (LMW-napins)[1] to distinguish them from the more common and abundant group of "high molecular mass napins" (14.0−14.7 kDa) (*6*).

Napins are polymorphic proteins due to their origin from multigene families. As a result, their isolation from the seeds renders a microheterogeneous material not suitable for three-dimensional structure determination, by either X-ray diffraction or NMR. In fact, we used a sample of BnIb from natural sources to determine its solution structure and were

[†] This work was supported by Projects BIO2002-00720 and SAF2002-02711 of the Spanish Ministry of Science and Technology granted to J.S. and R.R., respectively. D.P.-U. was supported by a predoctoral grant from the Spanish Ministry of Science and Technology.

[‡] Coordinates for the 20 minimized structures of BnIb have been deposited in the Protein Data Bank (entry 1SM7).

* To whom correspondence should be addressed. E-mail: jsantoro@iqfr.csic.es.

[§] CSIC.

[||] Universidad Complutense.

[1] Abbreviations: BSA, bovine serum albumin; CD, circular dichroism; COSY, correlated spectroscopy; ECL, enhanced chemiluminescence; LMW-napin, low molecular mass napin; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; nsLTP, nonspecific lipid transfer protein; PCR, polymerase chain reaction; RMSD, root-mean-square deviation; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; SGF, simulated gastric fluid; TOCSY, total correlated spectroscopy; TSP, trimethylsilyl propionate.

10.1021/bi048069x CCC: $27.50
Published on Web 11/25/2004

able just to define its global fold (*7*). On the other hand, the heterodimeric subunit organization of 2S albumins as well as the abundant disulfide bond content makes their recombinant expression difficult to perform in prokaryote systems. Therefore, we chose the yeast *Pichia pastoris* to produce the recombinant form of the precursor of BnIb (rproBnIb), which rendered a single polypeptide and highly soluble protein (*8*). In fact, the expression of rproBnIb in *P. pastoris* has been shown to be a successful method for high-yield production of homogeneous and properly folded protein since it was immunologically equivalent to the natural and processed form (*8*).

In this paper, we describe the high-resolution solution structure of rproBnIb as obtained by NMR methods. Having available as many as possible structures of related allergenic proteins is crucial for a rational mapping of their dominant epitopes, what in turn is important for clinical purposes. We also have examined the stability of rproBnIb against low-pH proteolysis and temperature, so as to get insights into the physicochemical and structural attributes that make this protein an allergen.

MATERIALS AND METHODS

Protein Sample. rproBnIb was expressed and purified as described previously (*8*). The protein's sequence shows two differences with respect to that given for napin in ref *7*: a Trp instead of Ser36 and a Ser substituting Trp43. These changes must not be considered as artifacts of the PCR reaction but as a result of the natural polymorphism of the protein, as previously reported (*8*).

SDS−PAGE and Immunoblotting. SDS−PAGE was carried out as described (*9*) in 15% polyacrylamide gels. Proteins were stained with Coomassie Blue or electrophoretically transferred onto nitrocellulose membranes. For immunoblotting, membranes were incubated with rproBnIb-specific rabbit antiserum (diluted 1:80 000) (*8*). Afterward, blots were incubated with horseradish peroxidase-labeled goat anti-rabbit IgG (Tago, Inc., Burlingame, CA) diluted 1:3000. The peroxidase reaction was developed by the ECL Western blotting reagent (Amersham Pharmacia Biotech, Uppsala, Sweden) as described (*10*).

Digestion in Simulated Gastric Fluid. The stability to proteolytic digestion of rproBnIb in SGF [pepsin 0.32% (w/v) and 30 mM NaCl, pH 1.2] was examined as described (*11*) with the following modifications: 40 μg of protein were dissolved in 200 μL of SGF. Incubations were maintained at 37 °C in a shaking water bath, and an aliquot of the digested sample was periodically withdrawn (at 10 and 60 s and 4, 16, and 60 min). The digestion was stopped with 0.2 M Na_2CO_3, and samples were mixed with a sample loading buffer for SDS−PAGE analysis. As a control, BSA was used as a sensitive protein to digestion by SGF.

Spectroscopic Analyses. Circular dichroism (CD) spectra of rproBnIb were obtained at different temperatures in the far-UV range (200−250 nm) on a Jasco J-715 spectropolarimeter (Japan Spectroscopic Co., Tokio, Japan) as described (*12*). The protein concentration was 0.20 mg/mL in 20 mM sodium phosphate buffer, pH 7.0. Mean residue mass ellipticity was calculated on the basis of 114.95 as the average molecular mass/residue, obtained from the amino acid composition, and expressed in terms of $[\theta]_{MWR}$ (degree·centimeter^2·decimole^{-1}).

Sequence Alignments. The routine application of alignment programs such as BLAST (*13*) or CLUSTAL W (*14*) did not align properly all cysteine residues. Therefore, manual reelaborations of the CLUSTAL W results, forcing the alignment of all eight cysteines, were performed.

NMR Spectroscopy. All NMR spectra of rproBnIb were acquired on a Bruker AV 600 NMR spectrometer at a protein concentration of 1.2 mM in 95% H_2O/5% D_2O or in a D_2O solution containing sodium phosphate buffer and TSP at pH 3.0 and 308 K. The pH of the samples was checked with a glass microelectrode and was not corrected for isotope effects. The temperature of the NMR probe was calibrated by use of a methanol sample (*15*). For the resonance assignments and collection of conformational constraints the following experiments were recorded: two-dimensional (2D) COSY, TOCSY with a mixing time of 60 ms, and NOESY with a mixing time of 80 ms. Proton chemical shifts were referenced to internal trimethylsilyl propionate (TSP). Water suppression was achieved either by selective presaturation or by including a WATERGATE module (*16*) in the original pulse sequences prior to acquisition. All experiments were acquired in the phase-sensitive mode (*17*−*19*), and conventional pulse sequences and phase-cycling procedures were used (*20*).

The size of the acquisition data matrix was 2048 × 512 data points in the direct and indirect dimension, respectively. Data matrixes were zero-filled up to duplicate the number of data points in both dimensions prior to Fourier transformation. Resolution enhancement methods included sine-bell or square-sine-bell windows with optimized shifts for every spectrum.

The processing of the spectra was carried out by using XWINNMR (Bruker Biospin, Karlsruhe, Germany). The spectral analysis, assignment, and cross-peak volume determinations were performed with NMRView (*21*).

Distance Constraints. The volumes obtained from integration of the NOESY cross-peaks were converted into upper distance limits by using CALIBA (*22*). The limits were restricted within the range of 2.7−6.0 Å. The strongest sequential HN_i/HN_{i+1} connectivity in helical regions was taken as the reference for the 2.8 Å distance. Finally, after DYANA filtering, 1316 upper-limit distance constraints, conformationally relevant, were obtained, distributed in 346 intraresidual, 409 sequential, 327 short-range, and 234 long-range constraints. Additionally, as the rproBnIb contains eight cysteine residues, four sets of standard upper and lower limits for each identified disulfide bond, 2.1/2.0 Å for $S\gamma(i)-S\gamma(j)$ and 3.1/3.0 Å for $C\beta(i)-S\gamma(j)$ and $C\beta(j)-S\gamma(i)$, were introduced during the rounds of calculations. No stereospecific assignments of prochiral protons were introduced in the initial calculations, and the usual correction for pseudoatoms (*23*) was added in all cases. In the final steps, the stereospecific assignments of $H\beta-H\beta'$ resonances of residues Glu8, Leu26, Asn50, Glu51, and Cys59, the $H\delta-H\delta'$ resonances of Pro79 and the methyl γ resonances of Val58 were obtained with GLOMSA (*22*) on the basis of a set of 20 protein structures and the interproton distance upper limits.

NOE Assignment and Structure Calculation. The three-dimensional solution structure of rproBnIb was determined by simulated annealing combined with molecular dynamics in torsion angle space followed by a final minimization in Cartesian space.

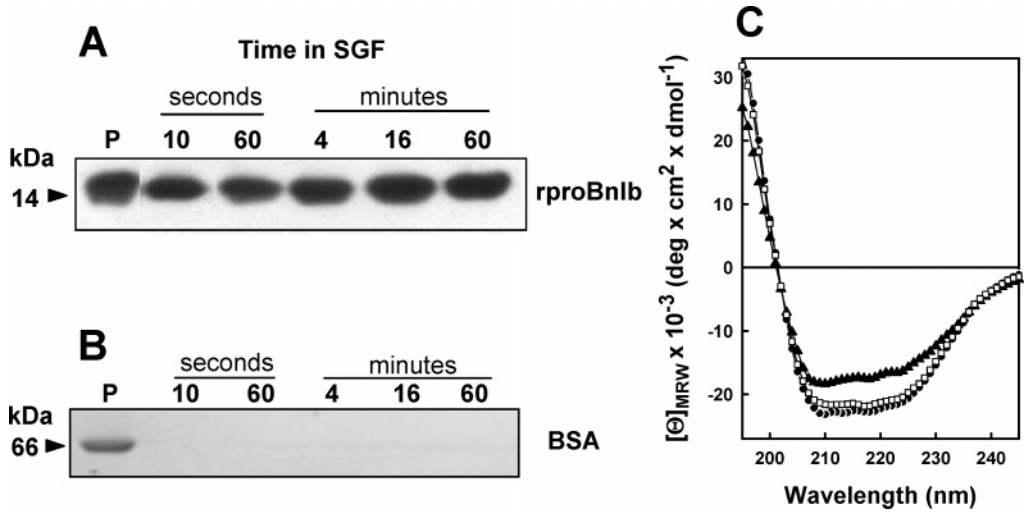

FIGURE 1: (A) Stability of rproBnIb to digestion by SGF. SDS−PAGE in 15% polyacrylamide gels and immunodetection of the samples transferred to membranes with rabbit anti-rproBnIb serum at different times of treatment with SGF are shown. rproBnIb sample was also run without treatment for control (lane P). (B) SDS−PAGE and Coomassie Blue staining of a nonallergenic protein (BSA) subjected to digestion by SGF and the corresponding control without treatment (lane P). (C) CD spectra of rproBnIb in the far-UV (195−245 nm) at different temperatures: (●) 20 °C, (▲) after heating at 80 °C, and (□) recooling at 20 °C. $[\Theta]_{MRW}$, mean residue weight ellipticity.

An initial molecular model was built based on the homology modeling technique. The previous structure of BnIb (*7*), where only the global fold could be determined because of the residue heterogeneity present in the natural sample, and rRicC3 (*24*), a suitable representative structure for the broad family of seed 2S albumin proteins with 20% sequence identity to BnIb, were used to build a model with the program MODELLER 6 version 1 (*25*). The resulting three-dimensional model together with a set of unambiguous distance constraints were used to run an automatic peak assignment in the NOESY spectra by use of the algorithms NOAH (*26*) and CANDID (*27*), as implemented in the programs DYANA 1.5 (*28*) and CYANA 1.0 (*29*), respectively. On the basis of the consistency with these initial structures, an additional set of NOEs could then be assigned and subsequently used to calculate a new set of structures. Several such cycles of structure calculation and NOE assignment were applied to assign the maximum number of NOE peaks possible. From these distance constraints, a set of 50 structures was calculated with DYANA version 1.5. All His, Lys, and Arg side chains were treated as positively charged, and all Glu and Asp side chains were considered to be negatively charged. The 20 structures with the lowest DYANA target function were refined by energy minimization in vacuo (500 steps) in the presence of NMR constraints by use of AMBER7 (*30*). PROCHECK-NMR (*31*) and MOLMOL (*32*) were used respectively to analyze the quality of the refined structures and to prepare the images of the molecules.

RESULTS

Stability to Digestion by SGF and Heat Treatment of rproBnIb. The resistance of rproBnIb to proteolytic digestion was assayed by use of SGF. The time course of the digestion was followed by immunostaining with rproBnIb-specific rabbit antiserum after SDS−PAGE and transference to membranes. As shown in Figure 1A, rproBnIb was resistant to digestion by SGF. After 1 h of treatment, rproBnIb remained undigested and conserved the antigenic epitopes. BSA was included in the experiment as a nonplant protein and was found to be very sensitive to SGF as it was completely digested within 10 s (Figure 1B).

Figure 1C shows the thermal unfolding of rproBnIb. CD analysis of the protein at different temperatures indicated that rproBnIb was very resistant to heat treatment. Application of the convex-constrain-analysis method of Perczel et al. (*33*) to the CD spectrum of rproBnIb at 20 °C gave the following composition: 83% α-helix and 17% nonregular conformation. At 80 °C rproBnIb was slightly perturbed, losing 27% of the α-helix content, but upon cooling at 20 °C the protein totally recovered the initial structure, indicating that the changes induced by heat treatment were completely reversible.

Resonance Assignment. Proton resonances were assigned by combining information extracted from 2D experiments, by a standard strategy (*23*). In Figure 2, the amide−amide region of the NOESY spectrum in H_2O is shown, together with the assignment of some helical segments. Most of the proton resonance assignments are coincident with the previous ones reported for BnIb (*7*). They have been deposited in the BioMagResBank database (*34*) under Accession Number BMRB-6137.

Disulfide Bonds. Four disulfide bonds stabilize the 3D structure of rproBnIb. The individual covalent links between the cysteine residues could be determined by a careful analysis of the NOE spectra (*35*). Several intercysteine NOE connectivities allowed us to identify clearly three disulfide bonds: Cys5−Cys59, Cys18−Cys48, and Cys61−Cys104. The only possibility for the remaining bond was then Cys49−Cys96. Long-range NOEs between residues in the vicinity of the cysteines confirm the disulfide pattern, which coincides with the one determined for other members of the 2S albumin family (*7*, *24*, *36*, *37*).

Quality of the Final Structure. Statistics on geometrical constraints and details of quality and precision for the 20 final structures are summarized in Table 1, and a superposition of the structures is shown in Figure 3. The structures show good geometry, low DYANA function target, low AMBER energy, and few NOE violations. The 20 final minimized structures are consistent with the resonance and

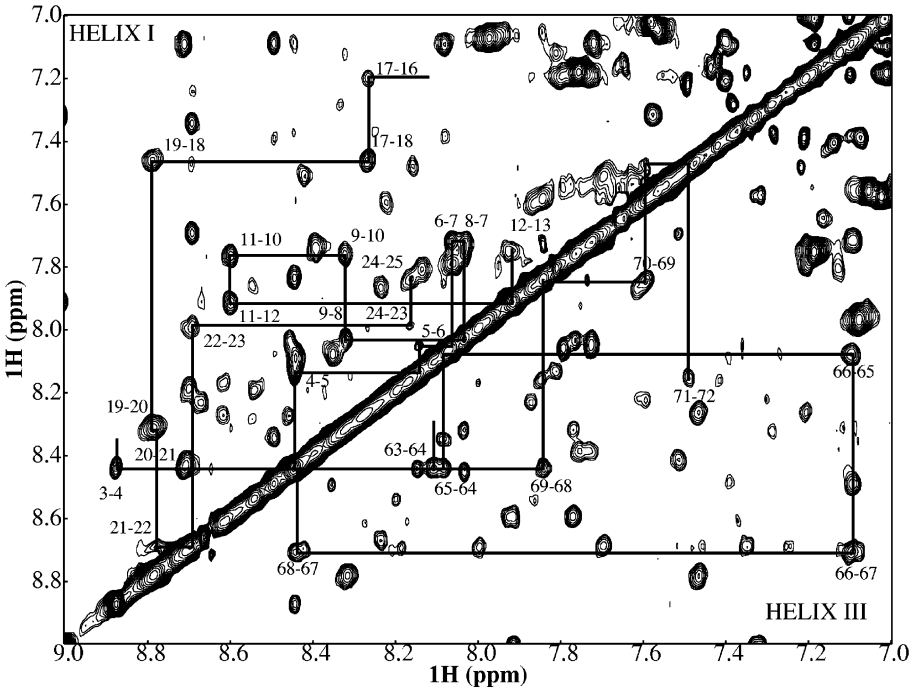

FIGURE 2: Amide−amide region of the 2D-NOESY 600 MHz spectrum, with links between sequential H^N−H^N cross-correlations of helix Ia and Ib above the diagonal and helix III below the diagonal.

Table 1: Structural Statistics of the 20 Best Structures of RproBnIb

Constraints			
intraresidual distances			346
sequential distances			409
medium-range distances: $(i - j) < 5$			327
long-range distances: $(i - j) \geq 5$			234
DYANA			
	average	minimal	maximal
target function	2.61	1.84	3.10
maximal violation (Å)	0.35	0.24	0.44
sum of violations (Å)	10.9	9.2	13.5
AMBER			
mean energy (kcal/mol)			−1207
RMS deviations from ideal geometry			
bond lengths (Å)			0.0103
bond angles (deg)			2.243
distance constraints violations			
number > 0.4 Å			1
0.4 Å ≥ number > 0.3 Å			7
0.3 Å ≥ number > 0.2 Å			35
maximal violation (Å)			0.46
RMSD			
		backbone	heavy atoms
global (3−25, 44−71, 81−104)		0.66	1.16
helix Ia		0.45	1.15
helix Ib		0.33	1.22
helix II		0.24	0.65
helix III		0.39	0.76
helix IV		0.35	0.87
Ramachandran			
favorable regions (%)			92.6
additional allowed (%)			7.2
disallowed (%)			0.1

NOE cross-correlation assignments, where the maximum violation is 0.46 Å. These structures show small deviations from the ideal geometry, 0.01 Å and 2.24° for bond lengths and angles, respectively. The average RMSD from the mean for backbone atoms of the most structured regions is 0.66 and 1.16 Å for all heavy atoms, excluding unstructured regions at the N- and C-termini (residues 1−2 and 105−109, respectively) and the loop regions of residues Leu26−Ser43 and Arg72−Phe80. The good convergence of the structures is founded on the high number (12.1 as an average) of distance constraints per residue and the fine definition of the polypeptide chain derives from the large number of NOEs in the core of the protein.

The Ramachandran plot for the 20 final minimized structures indicates that 99.9% of the backbone dihedral angles are in allowed regions and 0.1% in the disallowed regions according to the program PROCHECK-NMR (*31*). Two poorly defined regions are the segment Leu26−Ser43 and the hypervariable region (Arg72−Phe80), both showing a large scatter in their backbone dihedral angles.

Global Fold of the Main Chain. The 3D structure of rproBnIb was calculated on the basis of 1316 interproton upper distance constraints following the procedure described under Materials and Methods. The overall topology of the rproBnIb is the one already observed in the all-α proteins of the prolamin superfamily covering the nonspecific lipid transfer proteins (nsLTPs) (*38*−*41*), proteinase/α-amylase inhibitors (*42*−*44*), and 2S albumins (*7*, *24*, *45*). It consists of a globular four-helical motif with a simple "up and down" topology, whose helices are arranged in a right-handed superhelix. In rproBnIb, as occurs in rRicC3 (*24*) and SFA-8 (*45*), the first helix is split in two, so that the five helical regions extend from residues 3−11 (helix Ia), 16−25 (helix Ib), 44−54 (helix II), 57−71 (helix III), and 81−95 (helix IV).

Helices Ia and Ib run approximately coaxial and are connected by a nonhelical turn, L1 (12−15). Between helix Ib and helix II there is a long unstructured loop, L2, encompassing the residues Leu26−Ser43. Helix II runs antiparallel to helices Ia and Ib. Both helices I (Ia + Ib) and

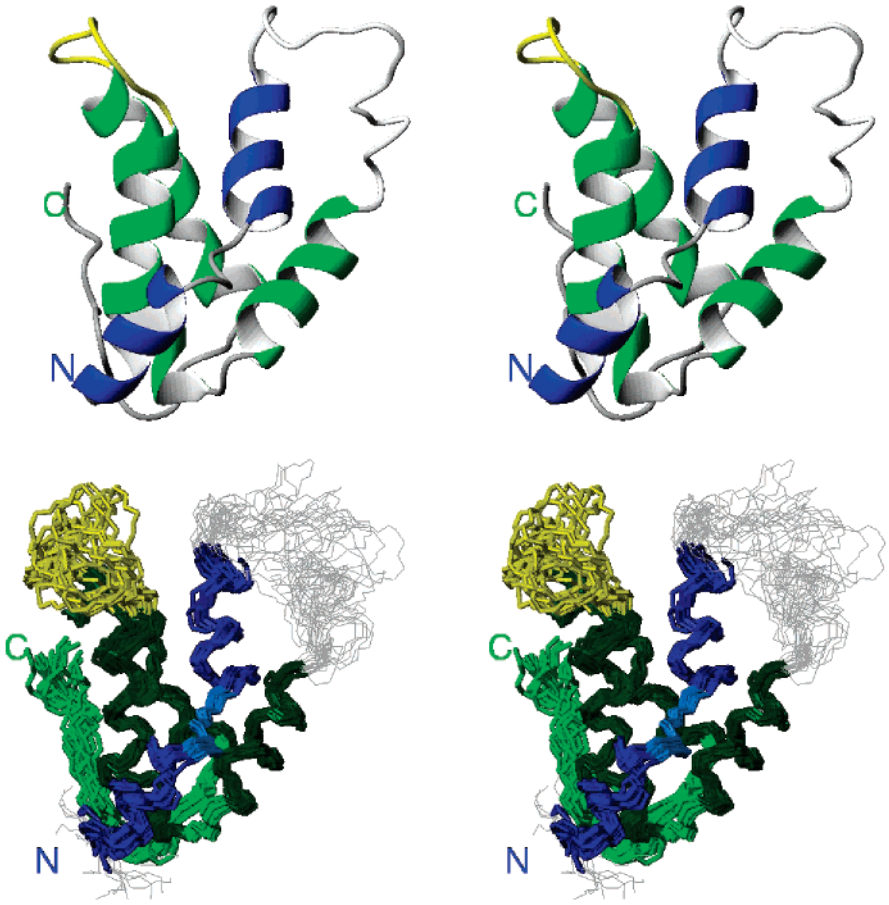

FIGURE 3: Top: Stereoview of a ribbon representation of the solution structure of rproBnIb. Bottom: Stereoview of the superposition of the 20 best DYANA structures of rproBnIb after restrained energy minimization.

II are coupled to the bundle formed by helices III and IV, by means of two disulfide bridges, Cys5 (helix Ia)−Cys59 (helix III) and Cys49 (helix II)−Cys96 (helix IV). A third disulfide bond is formed between Cys18 (helix Ib) and Cys48 (helix II), thus situating the N-terminus of helix Ib close to the middle of helix II. Two residues, Glu55 and Asp56, constitute the interhelical loop between helix II and helix III, L3. The 57−71 helix III is kinked, with a kink angle around 20°. This perturbation is related to the presence of Pro62 in the helix (*46*), which in addition breaks the hydrogen-bond network. Helices III and IV run almost antiparallel to each other and are connected by a short segment constituted by residues Arg72−Phe80, known as the "hypervariable region" in 2S albumins (*6*, *47*, *48*), because of the high variability in length and composition between the different members of the family. In helix IV, an intrahelical proline residue is found, Pro93, that breaks the regular i, $i + 4$ hydrogen-bonding pattern of the ideal helical structure but does not produces a detectable change in the direction of the helix axis. Finally, helix IV is followed by a long loop devoid of any regular secondary structure, which packs against helix III. The disulfide bridge between Cys61 (helix III) and Cys104 anchors this segment to the protein core.

Side Chains. Most of the side chains of the hydrophobic residues are significantly protected from the solvent and point toward the core of the protein. However, Leu26, Ala27, Phe31, Val58, Val73, Phe80, Val95, Ile101, Ile107, Ala108, and Ile109 present a fractional accessible surface area above 0.4. These residues are located in interhelical loops and in the C-terminal segment. The presence of hydrophobic residues exposed to the solvent is important, since they may define the regions of interaction with IgE antibodies in allergenic proteins. The side chains of most hydrophobic residues are well packed, resulting in small χ_1 circular variances. Nevertheless, Cys5 and Cys61, located in helices Ia and III; Phe31 and Val73, in the interhelical loops L2 and L3, respectively; and Cys104, Phe106, Ile107, and Ile109, in the C-terminal tail, display significant variability with circular variances greater than 0.15 ($\chi_1 = 0$, fixed conformation; $\chi_1 = 1$, random value). The core of the protein is formed by hydrophobic interactions between side chains of helix Ia, the interhelical loop L1, the N-terminus of helix Ib, the C-terminus of helix II, the N-terminus of helix III, and the C-terminal tail. It includes residues Phe9, Leu15, Cys18, Cys48, Leu52, Val60, Leu64, Ala67, Leu92, and Met98 and is surrounded by the remaining disulfide bonds Cys5−Cys59, Cys49−Cys96, and Cys61−Cys104. Another aggregate of hydrophobic side chains is found in the confluence of the C-termini of helix Ib and II and the middle of helix IV, close to the protein surface, and involves Ile22, Val71, Ile85, and Ile88. Furthermore, three hydrophobic patches with solvent-accessible surface areas ranging from 200 to 300 Å^2 were detected by the QUILT program (*49*). The first one is found in the confluence of the C-terminus of helix Ib, the N-terminus of helix IV, and the hypervariable loop (Trp21, Ile22, Leu26, and Ile88). The second one is placed in the channel between helix III and IV (Pro62, Tyr86,

Table 2: Summary of Structural Alignment Statistics

	Z-score DALI	RMSD (Å)	aligned Cα/ total Cα	sequence identity (%)
α-amylase inhibitor from maize	7.7	3.1	93/116	19
HPS protein	4.5	3.1	67/75	12
nsLTP from maize	2.8	2.6	61/93	23

Pro105, Phe106, and Ile109), and a third one is located around the disulfide bond Cys5−Cys59, the interhelical loop L3 (Val58), and the C-terminal tail (Ile101).

rproBnIb has seven negative (Asp and Glu) and 13 positive (Arg, His, and Lys) charged residues. Most of these charged residues are located on the surface of the protein and show high fractional accessibility values.

DISCUSSION

Comparison with Other 2S Albumins. Up to now, the structures of three 2S albumins have been reported: those from the wild-type protein, BnIb (*7*) (1PNB), 2S albumin from *Ricinus communis*, RicC3 (*24*) (1PSY), and 2S albumin from *Helianthus annuus* L., SFA-8 (*45*) (1S6D).

In contrast with BnIb, rproBnIb consists of a single polypeptide chain, containing the entire segment between helices Ib and II of the precursor protein, which includes the fragment Ser32-Glu33-Asn34 eliminated in the mature protein by proteolytic cleavage during its biological processing. The 3D structures of rproBnIb, rRicC3, and SFA-8 show that the loop between helices Ib and II is unstructured and does not interact with any other part of the protein, which evidences that the linker region is not needed for a proper folding of the members of this family of proteins. Consequently, it was to be expected that the 3D core structure of the wild-type protein did not change much with respect to that of rproBnIb.

Although the residue heterogeneity present in the sample of wild-type BnIb precluded the determination of a high-resolution 3D structure, and only its global fold could be advanced (*7*), the agreement with the structure of rproBnIb here determined is most satisfactory, in general terms. This is right in regard to the extension and length of helices I−III but not for helix IV, which is six residues longer in the latter structure. The two proteins adopt similar architectures, but an analysis of the interhelical angles shows a shift of about 35° in the orientation of helix Ib. Nevertheless, the most divergent feature in the structures of BnIb and rproBnIb is the arrangement of the C-terminal tail. In rproBnIb, that segment is packed against helices III and IV, whereas in BnIb it is folded over itself and penetrates into the core of the protein. rproBnIb adopts a molecular architecture very similar to those of rRicC3 (*24*) and SFA-8 (*45*). Therefore, all the credit must be given to the arrangement seen in rproBnIb, the observed differences with respect to BnIb being most probably due to artifacts in the 3D experimental model of the later (*7*).

Comparison with Other Members of the Prolamin Superfamily. Table 2 summarizes some parameters derived from the structural alignment of some representative members of the prolamin superfamily: the α-amylase from maize (1BFA, *42*), the hydrophobic soybean protein, HSP (1HYP, *50*), and the nsLTP from maize (1MZM, *51*). Two families of nsLTPs are to be distinguished, one with a molecular mass of ∼9 kDa referred to as nsLTP1, and other with molecular mass of 7 kDa as nsLTP2s (*52*). Data in Table 2 indicate that rproBnIb is more similar to the α-amylase inhibitors than to the nsLTP protein family.

The pattern of disulfide bonds of rproBnIb is identical to those of the 2S albumins SFA-8, RicC3, and α-amylase inhibitors but differs from the nsLTP1s in that the links in the CXC segment are exchanged. As in rRicC3 and SFA-8, the different orientation shown by helix III in rproBnIb with respect to the one in nsLTP1s, about 20°, impedes the formation of the cavity running through the whole molecule found in nsLTP1s (*41*), which serves to host one lipid molecule. No evidence for the existence of such a hydrophobic tunnellike cavity was found in rproBnIb, either by close examination of the structures or by the application of the VOIDOO program (*53*).

Analysis of Multiple Sequence Alignments. Analysis of multiple sequence alignments of 2S albumins reveals that, in addition to the eight conserved cysteines, there are six conserved hydrophobic positions (see Figure 4). If we extend the analysis to include sequences from other members of the prolamin superfamily, we find that these hydrophobic positions are still conserved except position 5 in nsLTP1s, where an Asn is typically found. In 2S albumins, residues at three of these positions, 1, 4, and 5, belong to the protein core, while these at positions 3 and 6 form a second hydrophobic cluster surrounded by helices Ib, III, and IV. All these nonpolar contacts contribute to the molecule compactness and stability.

The subfamily of nsLTP1s is unique in having a disulfide mismatch at the Cys-X-Cys motif. Some authors (*54*) have related the hydrophobicity of the residue X with the pairing of the disulfide bridges: nsLTP2s have a hydrophobic residue at X, buried within the core of the protein, whereas nsLTP1s possess a hydrophilic residue at X whose exposure to solvent requires a rotation of the Cys-X-Cys motif and thereby induces a mismatch in the cysteine pairing. Whereas this reasoning applies to nsLTP1s and nsLTP2s, it no longer holds for other members of the prolamin superfamily, such as the 2S albumins RicC1 (*56*) and Ber e 1 (*36*) or the α-amylase inhibitor from ragi seeds (*44*), which have a polar residue at X but exhibit the same S−S pairing as in nsLTP2s. A possible explanation is that the hydrophobicity of the residue at position 5 is the key determinant of the disulfide pairing.

In a previous report (*24*), we proposed that the pattern of cysteine pairing determines whether a hydrophobic tunnel, able to host a lipid molecule, could be formed or not in the interior of a protein with the prolamin family scaffold. The fact that nsLTP2s, having the opposite pairing to nsLTP1s, also appear to form an internal cavity (*54*, *57*), does not invalidate our proposal, since its location and geometry are quite different from the ones in LTP1s.

High Stability and Compactness of the Structure of rproBnIb. The high resistance to digestion by SGF and heat treatment of rproBnIb derives from its highly stable and compact 3D structure. Disulfide bridges play an important role in that regard by maintaining intact the global fold of the molecule during those treatments (*58*). A high resistance to digestion was previously documented for Sin a 1 (*11*), the major allergen and 2S albumin from yellow mustard seeds (*59*). Sin a 1 must have a very similar 3D structure to

```
                                  1     23
rproBnIb  ------QPQKCQREFQQEQ-HLRACQQWIRQQLAGSPFSEN
rRicC3    GEREGSSSQQCRQEVQRKD--LSSCERYLRQSSS-linker
SFA-8     PYGRGRTESGCYQQMEEAE-MLNHCGMYLMKNLGERSQV--
1AFH      -------AISCGQVASA----IAPCISYARG----------
1N89      ---------ACQASQ------LAVCASAILS----------
1B1U      -----SVGTSCIPGMAIPHNPLDSCRWYVSTRTCGVGP---
1HYP      ---ALITRPSCP--------DLSICLNILGG----------

                             4                    5
rproBnIb  QWGPQQGPSLREQCCNELYQ----------EDQVCVCPTLKQAAKSVRVQ--
rRicC3    ---QQQESQQLQQCCNQVKQVR--------DE--CQCEAIKYIAEDQIQQGQ
SFA-8     -SPRMREEDHKQLCCMQLKNL---------DE-KCMCPAIMMMLNEPMWIRM
1AFH      -----QGSGPSAGCCSGVRSLNNAARTTADRRAACNCLKNAAAGV-------
1N89      ------GAKPSGECCGNLRAQ-----------QGCFCQYAKDPTY-------
1B1U      ---RLATQEMKARCCRQLEAIP----------AYCRCEAVRILMDGVVTPSG
1HYP      ------SLGTVDDCCALIGGL------GDIEAIVCLCIQLRALG--------

                                    6
rproBnIb  ---------GQHGPFQSTRIYQIAKNLPNVCNMK------QIGTCPFIAI---
rRicC3    ----------LHG-EESERVAQRAGEIVSSCGV----------RCM-------
SFA-8     ----------------RDQVMSMAHNLPIECNLMSQ-------PCQM------
1AFH      ----------------SGLNAGNAASIPSKCGVSIPYTISTSTDCSRVN----
1N89      ---------------GQYIRSPHARDTLTSCGLAVP-------HC--------
1B1U      QHEGRLLQDLPGCPRQVQRAFAPKLVTEVECNLATIHGGP---FCLSLLGAGE
1HYP      ----------------ILNLNRNLQLILNSCGRSYPSNA----TCPRT-----
```

FIGURE 4: Structure alignment of representative proteins of the prolamin superfamily. Abbreviations: rproBnIb, this study; RicC3, 2S albumin storage protein from *Ricinus communis* (*24*); SFA-8, methionine-rich 2S albumin from sunflower seeds (*45*); 1AFH, nsLTP from maize seeds (*39*); 1N89, wheat nsLTP2 (*54*); 1B1U, bifunctional inhibitor of trypsin and α-amylase from ragi (*55*); and 1HYP, the hydrophobic soybean protein (*50*). Residues involved in helices are underlined. Positions marked with numbers 1−6 host conserved hydrophobic residues.

that of rproBnIb because of its high sequence homology, so we may conclude that in general 2S albumins are highly stable against low pH, proteolysis, and temperature. This exceptional stability can be explained by taking into account the high compactness of their structure. The volume of rproBnIb as calculated with the program VADAR (*60*) is 11 635 Å^3, a value 20% lower than that expected from its molecular weight (14 731 Å^3). The high stability of 2S albumins is crucial in relation to the allergenicity presented by most members of this family of proteins. As previously recognized (*61*), by being highly stable at low pH, 2S albumins are able to reach intact the gut immune system so as to induce sensitization and elicitation of allergic reactions.

The Hypervariable Region and Epitope Mapping in 2S Albumins. The segment connecting helices III and IV in 2S albumins, known as the hypervariable region, has been described as a segment exhibiting a high variability in the number and composition of residues among the different members of the family (*5*, *47*, *48*). Immunological studies performed on the 2S albumins of mustard seeds, Sin a 1 (*59*) and Bra j 1E (*62*), and the one from English walnut, Jug r 1, (*63*), have shown that the hypervariable segment constitutes an important antigenic region. The segment forms an unstructured loop, and the residues therein present a high value of the fractional accessible surface area, what indicates that the loop is highly exposed to solvent. Moreover, an empirical measure of local protein flexibility, based on the estimation of the NMR S^2 order parameters of the N−H vectors of the protein backbone (*64*), indicates that the hypervariable loop in rproBnIb exhibits great mobility, as occurs in the analogue protein from ricin, RicC3 (*24*). The hypervariable loop has the sequence VRVQGQHGPFQS in rproBnIb, containing two charged amino acids, Arg72 and His 77, which are strong candidates to be involved in potential IgE-binding sites.

Ara h 2, a major peanut allergen, has a high degree of identity with 2S albumin proteins (30−35%) (*65*) and exhibits an identical disulfide bridge signature. The main divergence with 2S albumins consists of the great length (34 residues) of the loop equivalent to L1. A common structure with those of the 2S albumins is then to be expected. The major linear immunoglobulin- (IgE-) binding epitopes of this allergen were mapped (*65*) by use of overlapping peptides covering the entire sequence of Ara h 2 (see Figure 5). Three main epitopes were recognized by sera of 15 peanut-sensitive patients: Asp27−Arg36, Ser57−Gln66, and Ser65−Tyr74, corresponding in rproBnIb to the segments 2−11 (helix Ia) and two segments of the loop L1 with repeated sequences. Interestingly, the peptide DRLQGRQQEQ located at the hypervariable region was also recognized as an epitope, albeit with less generality than the dominant ones.

Epitope Mapping in Other Members of the Prolamin Family. Pru p 3 is the major allergen from *Prunnus persica*. Its amino acid sequence presents high identity (53−62%) to those of nsLTPs of cereals, for which 3D experimental structures are available. In a recent study (*68*), a 3D model was built for Pru p 3 by using the structure of the nsLTPs from rice (1RZL, *67*) as a template. The main IgE-binding regions were identified by testing some specially designed protein variants, as well as synthetic peptides covering the full Pru p 3 sequence. IgE immunodetection of synthetic peptides led to the identification of three regions as major IgE epitopes (see Figure 5): residues around the second Cys until the CC motif, residues between motifs CC and CXC, and residues in the region of the penultimate Cys. However, only the triple mutant R39A/T40A/R44A showed a substantial decrease of IgE binding, thus pointing toward the connection between helices II and III and the first turn of helix III as the dominant allergen of Pru p 3 (residues 41−47).

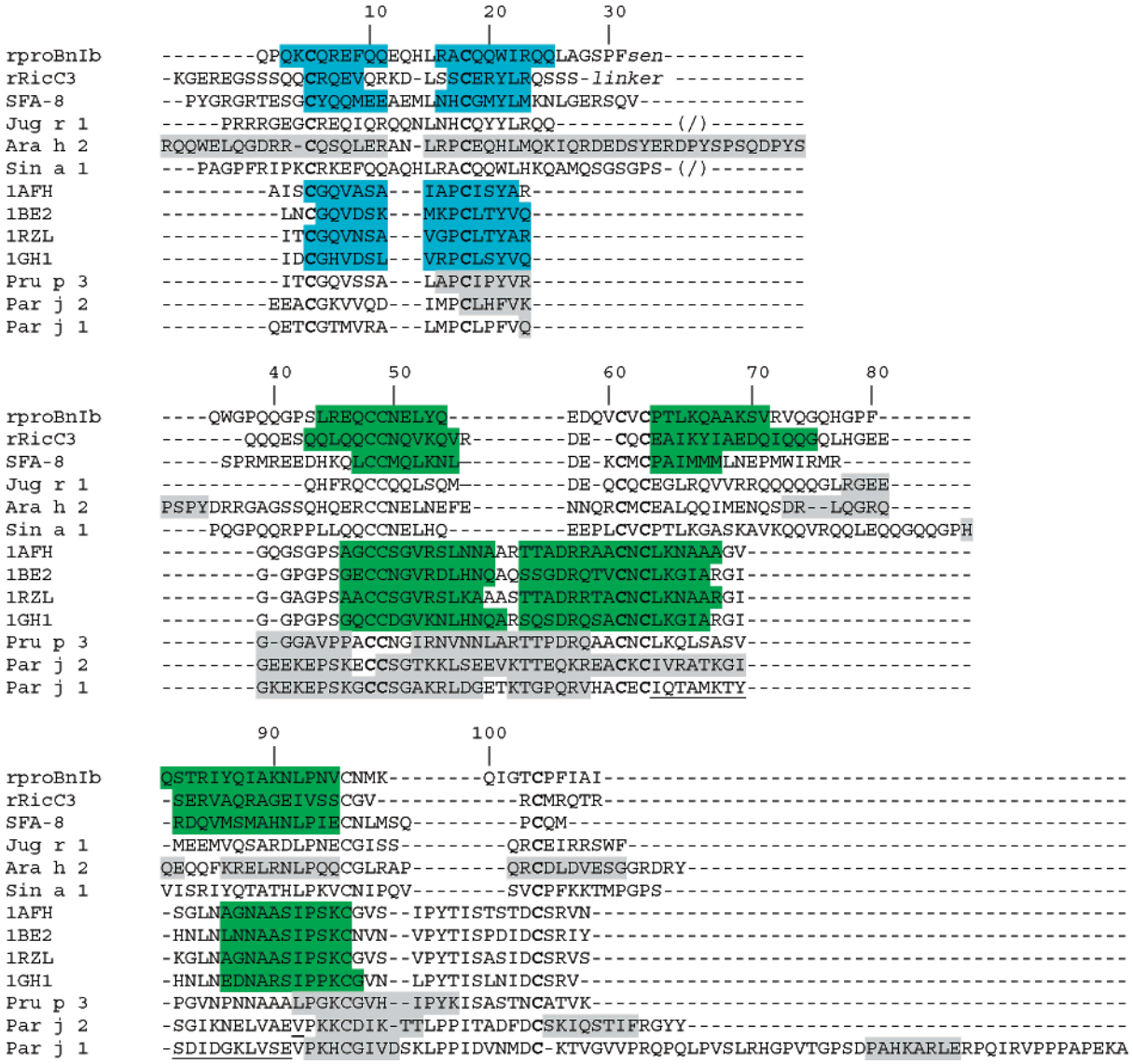

FIGURE 5: Sequence alignment of 2S albumins and nsLTPs with epitope mapping information. Cysteine residues are marked in bold. The residues involved in helices are shown in blue (short-chain) and green (long-chain) boxes. Linear IgE epitopes are boxed in gray, and T cell epitopes are underlined. Abbreviations: rproBnIb, this study; rRicC3, 2S albumin storage protein from *Ricinus communis* (*24*); SFA-8, methionine-rich 2S albumin from sunflower seeds (*45*); Jug r 1, 2S albumins from English walnut (*63*); Ara h 2, a major peanut allergen (*65*); Sin a 1, 2S albumin of yellow mustard seeds (*59*); 1AFH, nonspecific lipid transfer protein extracted from maize seeds (*39*); 1BE2, nonspecific lipid transfer protein extracted from barley (*65*); 1RZL, nonspecific lipid transfer protein extracted from rice (*67*); 1GH1, nonspecific lipid transfer protein extracted from wheat (*38*); Pru p 3, the major peach allergen (*68*); Par j 2 and Par j 1, the major allergens from *Parietaria judaica* pollen (*69*). The residue numbering refers to rproBnIb.

Also recently, a similar study has been carried out with the objective of identifying IgE-binding epitopes on Par j 1 and Par j 2, the major allergens from *Parietaria judaica* pollen (*69*). As in the previous study, overlapping peptides representing the complete length of the allergens were tested against the sera of allergic patients. No mutant allergen was tested in this case. Three common allergenic epitopes with high homology and identical conformation were detected (see Figure 5): the segment 18−27, a loop connecting helices Ib and helix II; the segment 30−38 in the C-terminal of helix II; and the segment 73−79 at the end of helix IV and the beginning of the unstructured C-terminus. When the data from Pru p 3 and those from Par j 1 and Par j 2 are considered together, four consensus regions are delineated: 18−27 (*P. judaica* numbering), 33−48, 41−47, and 72−79. The first, third, and fourth segments belong to flexible and solvent-exposed regions, whereas the second one has lower solvent accessibility. The segment 41−47, covering the loop between helices II and III and the N-terminal region of helix III, is the best candidate to be the dominant IgE epitope in nsLTPs, on the basis of the results obtained for Pru p 3 with the triple mutant R39A/T40A/R44A (R41A/T42A/R46A according to the numbering of *P. judaica*).

As a case study for a proposed method for epitope mapping (*70*), an 18 amino acid residue peptide from the sequence of the α-amylase inhibitor, a major allergen from wheat, was identified, encompassing residues 9−26. That segment corresponds to the N-terminal unstructured residues, together with helix Ia, that in the α-amylase family is missing, a region which obviously is highly flexible and solvent exposed.

In summary, it appears that the main role of the highly stable structure common to the members of the prolamin family is that of reinforcing their ability to reach intact the

gut immune system. It appears also that the structure per se does not prefigure common regions of IgE binding among the different members of the superfamily. However, it is true that the different regions detected with IgE-binding capability in the different families do comply with the conditions of being flexible and solvent-exposed.

REFERENCES

1. Byczynska, A., and Barciszewski, J. (1999) The biosynthesis, structure and properties of napin − the storage protein from rape seeds, *J. Plant Physiol. 154*, 417−425.
2. Monsalve, R. I., Villalba, M., Rico, M., Shewry, P. R., and Rodríguez, R. (2004) The 2S albumin proteins, in *Plant Food Allergens* (Mills, E. N. C., and Shewry, P. R., Eds.), pp 42−56, Blackwell Science, Oxford, U.K.
3. Teuber, S. S., Dandekar, A. M., Peterson, W. R., and Sellers, C. L. (1998) Cloning and sequencing of a gene encoding a 2S albumin seed storage protein precursor from English walnut (*Juglans regia*), a major food allergen, *J. Allergy Clin. Immunol. 101*, 807−814.
4. Crevel, R. (2002) Industrial dimensions of food allergy, *Biochem. Soc. Trans. 30*, 941−944.
5. Monsalve, R. I., Villalba, M., and Rodríguez, R. (2001) Allergy to mustard seeds: The importance of 2S albumins as food allergens, *Internet Symp. Food Allergens 3*, 57−69.
6. Monsalve, R. I., Lopez-Otín, C., Villalba, M., and Rodríguez, R. (1991) A new distinct group of 2S albumins from rapeseed. Amino acid sequence of two low molecular weight napins, *FEBS Lett. 295*, 207−210.
7. Rico, M., Bruix, M., González, C., Monsalve, R. I., and Rodríguez, R. (1996) ^{1}H NMR assignment and global fold of napin BnIb, a representative 2S albumin seed protein, *Biochemistry 35*, 15672−15682.
8. Palomares, O., Monsalve, R. I., Rodríguez, R., and Villalba, M. (2002) Recombinant pronapin precursor produced in *Pichia pastoris* displays structural and immunologic equivalent properties to its mature product isolated from rapeseed, *Eur. J. Biochem. 269*, 2538−2545.
9. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227*, 680−685.
10. Villalba, M., Batanero, E., Monsalve, R. I., González de la Peña, M. A., Lahoz, C., and Rodríguez, R. (1994) Cloning and expression of Ole e 1, the major allergen from olive tree pollen. Polymorphism analysis and tissue specificity, *J. Biol. Chem. 269*, 15217−15222.
11. Astwood, J. D., Leach, J. N., and Fuchs, R. L. (1996) Stability of food allergens to digestion in vitro, *Nat. Biotechnol. 14*, 1269−1273.
12. Huecas, S., Villalba, M., González, E., Martínez-Ruiz, A., and Rodríguez, R. (1999) Production and detailed characterization of biologically active olive pollen allergen Ole e 1 secreted by the yeast *Pichia pastoris*, *Eur. J. Biochem. 261*, 539−546.
13. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res. 25*, 3389−3402.
14. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.
15. Duerst, R., and Merbach, A. (1965) Accurate NMR temperature measurements, *Rev. Sci. Instr. 36*, 1896−1897.
16. Piotto, M., Saudek, V., and Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions, *J. Biomol. NMR 2*, 661−665.
17. States, D., Haberkorn, R. A., and Ruben, D. J. (1982) A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants, *J. Magn. Reson. 48*, 286−292.
18. Redfield, A. G., and Kuntz, S. D. (1975) Quadrature Fourier NMR detection: Simple multiplex for dual detection and discussion, *J. Magn. Reson. 19*, 250−254.
19. Marion, D., and Wüthrich, K. (1983) Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^{1}H−^{1}H spin−spin coupling constants in proteins, *Biochem. Biophys. Res. Commun. 113*, 967−974.
20. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., and Skelton, N. J. (1996) *Protein NMR Spectroscopy*, Academic Press Inc., San Diego, CA.
21. Johnson, B. A., and Blevins, R. A. (1994) NMRView: A computer program for the visualization and analysis of NMR data, *J. Biomol. NMR 4*, 603−614.
22. Güntert, P., Braun, W., and Wüthrich, K. (1991) Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA, *J. Mol. Biol. 217*, 517−530.
23. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
24. Pantoja-Uceda, D., Bruix, M., Giménez-Gallego, G., Rico, M., and Santoro, J. (2003) Solution structure of RicC3, a 2S albumin storage protein from *Ricinus communis*, *Biochemistry 42*, 13839−13847.
25. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol. 234*, 779−815.
26. Mumenthaler, C., Güntert, P., Braun, W., and Wüthrich, K. (1997) Automated combined assignment of NOESY spectra and three-dimensional protein structure determination, *J. Biomol. NMR 10*, 351−362.
27. Herrmann, T., Güntert, P., and Wüthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA, *J. Mol. Biol. 319*, 209−227.
28. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA, *J. Mol. Biol. 273*, 283−298.
29. Güntert, P. (2003) Automated NMR protein structure calculation, *Prog. NMR Spectrosc. 43*, 105−125.
30. Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E., III, Wang, J., Ross, W. S., Simmerling, C., Darden, T., Merz, K. M., Stanton, R. V., Cheng, A., Vincent, J. J., Crowley, M., Tsui, V., Gohlke, H., Radmer, R., Duan, Y., Pitera, J., Massova, I., Seibel, G. L., Singh, U. C., Weiner, P., and Kollman, P. A. (2002) AMBER 7, University of California, San Francisco.
31. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR 8*, 477−486.
32. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures, *J. Mol. Graphics 14*, 51−55.
33. Perczel, A., Hollosi, M., Tusnady, G., and Fasman, G. D. (1991) Convex constraint analysis: a natural deconvolution of circular dichroism curves of proteins, *Protein Eng. 4*, 669−679.
34. Seavey, B. R., Farr, E. A., Westler, W. M., and Markley, J. L. (1991) A relational database for sequence-specific protein NMR data, *J. Biomol. NMR 1*, 217−236.
35. Williamson, M. P., Havel, T. F., and Wüthrich, K. (1985) Solution conformation of proteinase inhibitor IIA from bull seminal plasma by ^{1}H nuclear magnetic resonance and distance geometry, *J. Mol. Biol. 182*, 295−315.
36. Alcocer, M. J., Murtagh, G. J., Bailey, K., Dumoulin, M., Meseguer, A. S., Parker, M. J., and Archer, D. B. (2002) The disulphide mapping, folding and characterisation of recombinant Ber e 1, an allergenic protein, and SFA-8, two sulphur-rich 2S plant albumins, *J. Mol. Biol. 324*, 165−175.
37. Egorov, T. A., Odintsova, T. I., Musolyamov, A., Fido, R., Tatham, A. S., and Shewry, P. R. (1996) Disulphide structure of a sunflower seed albumin: conserved and variant disulphide bonds in the cereal prolamin superfamily, *FEBS Lett. 396*, 285−288.
38. Gincel, E., Simorre, J. P., Caille, A., Marion, D., Ptak, M., and Vovelle, F. (1994) Three-dimensional structure in solution of a wheat lipid-transfer protein from multidimensional ^{1}H NMR data. A new folding for lipid carriers, *Eur. J. Biochem. 226*, 413−422.
39. Gomar, J., Petit, M. C., Sodano, P., Sy, D., Marion, D., Kader, J. C., Vovelle, F., and Ptak, M. (1996) Solution structure and lipid binding of a nonspecific lipid transfer protein extracted from maize seeds, *Protein Sci. 5*, 565−577.
40. Heinemann, B., Andersen, K. V., Nielsen, P. R., Bech, L. M., and Poulsen, F. M. (1996) Structure in solution of a four-helix lipid binding protein, *Protein Sci. 5*, 13−23.

41. Poznanski, J., Sodano, P., Suh, S. W., Lee, J. Y., Ptak, M., and Vovelle, F. (1999) Solution structure of a lipid transfer protein extracted from rice seeds. Comparison with homologous proteins, *Eur. J. Biochem. 259*, 692−708.
42. Behnke, C. A., Yee, V. C., Trong, I. L., Pedersen, L. C., Stenkamp, R. E., Kim, S. S., Reeck, G. R., and Teller, D. C. (1998) Structural determinants of the bifunctional corn Hageman factor inhibitor: X-ray crystal structure at 1.95 Å resolution, *Biochemistry 37*, 15277−15288.
43. Oda, Y., Matsunaga, T., Fukuyama, K., Miyazaki, T., and Morimoto, T. (1997) Tertiary and quaternary structures of 0.19 alpha-amylase inhibitor from wheat kernel determined by X-ray analysis at 2.06 Å resolution, *Biochemistry 36*, 13503−13511.
44. Strobl, S., Mühlhahn, P., Bernstein, R., Wiltscheck, R., Maskos, K., Wunderlich, M., Huber, R., Glockshuber, R., and Holak, T. A. (1995) Determination of the three-dimensional structure of the bifunctional alpha-amylase/trypsin inhibitor from ragi seeds by NMR spectroscopy, *Biochemistry 34*, 8281−8293.
45. Pantoja-Uceda, D., Shewry, P. R., Bruix, M., Tatham, A. S., Santoro, J., and Rico, M. (2004) Solution structure of a methionine-rich 2S albumin from sunflower seeds: Relationship to its allergenic and emulsifying properties, *Biochemistry 43*, 6976−6986.
46. Barlow, D. J., and Thornton, J. M. (1988) Helix geometry in proteins, *J. Mol. Biol. 201*, 601−619.
47. Raynal, M., Depigny, D., Grellet, F., and Delseny, M. (1991) Characterization and evolution of napin-encoding genes in radish and related crucifers, *Gene 99*, 77−86.
48. Krebbers, E., Herdies, L., De Clercq, A., Seurinck, J., Leemans, J., Van Damme, J., Segura, M., Gheysen, G., Van Montagu, M., and Vandekerckhove, J. (1988) Determination of the processing sites of an *Arabidopsis* 2S albumin and characterization of the complete gene family, *Plant Physiol. 87*, 859−866.
49. Lijnzaad, P., Berendsen, H. J., and Argos, P. (1996) Hydrophobic patches on the surfaces of protein structures, *Proteins: Struct., Funct., Genet. 25*, 389−397.
50. Baud, F., Pebay-Peyroula, E., Cohen-Addad, C., Odani, S., and Lehmann, M. S. (1993) Crystal structure of hydrophobic protein from soybean; a member of a new cysteine-rich family, *J. Mol. Biol. 231*, 877−887.
51. Shin, D. H., Lee, J. Y., Hwang, K. Y., Kim, K. K., and Suh, S. W. (1995) High-resolution crystal structure of the nonspecific lipid-transfer protein from maize seedlings, *Structure 3*, 189−199.
52. Kader, J. C. (1996) Lipid-transfer proteins in plants, *Annu. Rev. Plant Physiol. Plant Mol. Biol. 47*, 627−654.
53. Kleywegt, G. J., and Jones, T. A. (1994) Detection, delineation, measurement and display of cavities in macromolecular structures, *Acta Crystallogr. D50*, 178−185.
54. Pons, J. L., de Lamotte, F., Gautier, M. F., and Delsuc, M. A. (2003) Refined solution structure of a liganded type 2 wheat nonspecific lipid transfer protein, *J. Biol. Chem. 278*, 14249−14256.
55. Gourinath, S., Alam, N., Srinivasan, A., Betzel, C., and Singh, T. P. (2000) Structure of the bifunctional inhibitor of trypsin and alpha-amylase from ragi seeds at 2.2 Å resolution, *Acta Crystallogr. D: Biol. Crystallogr. 56*, 287−293.
56. Bashir, M. E., Hubatsch, I., Leinenbach, H. P., Zeppezauer, M., Panzani, R., and Hussein, I. H. (1998) RicC1 and RicC3, the allergenic 2S albumin Storage Proteins of *Ricinus communis*: Complete Primary Structures and Phylogenetic Relationships, *Int. Arch. Allergy Immunol. 115*, 73−82.
57. Samuel, D., Liu, Y. J., Cheng, C. S., and Lyu, P. C. (2002) Solution structure of plant nonspecific lipid transfer protein-2 from rice (*Oryza sativa*), *J. Biol. Chem. 277*, 35267−35273.
58. Gekko, K., Kimoto, A., and Kamiyama, T. (2003) Effects of disulfide bonds on compactness of protein molecules revealed by volume, compressibility, and expansibility changes during reduction, *Biochemistry 42*, 13746−13753.
59. Menéndez-Arias, L., Domínguez, J., Moneo, I., and Rodríguez, R. (1990) Epitope mapping of the major allergen from yellow mustard (*Sinapis alba* L.), *Mol. Immunol. 27*, 143−150.
60. Wishart, D., Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R., and Sykes, B. (2003) Program Vadar, Faculty of Pharmacy, University of Alberta, Edmonton, Canada.
61. Mills, E. N. C., Jenkins, J. A., and Shewry, P. R. (2004) The role of common properties in determining plant food protein allergenicity, in *Plant Food Allergens* (Mills, E. N. C., and Shewry, P. R., Eds.), pp 158−170, Blackwell Science, Oxford, U.K.
62. Monsalve, R. I., González de la Peña, M. A., Menéndez-Arias, L., López-Otín, C., Villalba, M., and Rodríguez, R. (1993) Characterization of a new oriental mustard (*Brassica juncea*) allergen, Bra j 1E: detection of an allergen epitope, *Biochem. J. 293*, 625−632.
63. Robotham, J. M., Suzanne, S., Teuber, S., Sathe, S. K., and Roux, K. H. (2002) Linear IgE epitope mapping of the English walnut (*Juglans regia*) major food allergen, Jug r 1, *J. Allergy Clin. Immunol. 109*, 143−149.
64. Zhang, F., and Brüschweiler, R. (2002) Contact Model for the Prediction of NMR N−H Order Parameters in Globular Proteins, *J. Am. Chem. Soc. 124*, 12654−12655.
65. Stanley, J. S., King, N., Burks, A. W., Huang, S. K., Sampson, H., Cockrell, G., Helm, R. M., West, C. M., and Bannon, G. A. (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.
66. Lerche, M. H., and Poulsen, F. M. (1998) Solution structure of barley lipid transfer protein complexed with palmitate. Two different binding modes of palmitate in the homologous maize and barley nonspecific lipid transfer proteins, *Protein Sci. 7*, 2490−2498.
67. Lee, J. Y., Min, K., Cha, H., Shin, D. H., Hwang, K. Y., and Suh, S. W. (1998) Rice nonspecific lipid transfer protein: the 1.6 Å crystal structure in the unliganded state reveals a small hydrophobic cavity, *J. Mol. Biol. 276*, 437−448.
68. García-Casado, G., Pacios, L. F., Díaz-Perales, A., Sánchez-Monge, R., Lombardero, M., García-Selles, F. J., Polo, F., Barber, D., and Salcedo, G. (2003) Identification of IgE-binding epitopes of the major peach allergen Pru p 3, *J. Allergy Clin. Immunol. 112*, 599−605.
69. Asturias, J. A., Gomez-Bayon, N., Eseverri, J. L., and Martinez, A. (2003) Par j 1 and Par j 2, the major allergens from *Parietaria judaica* pollen, have similar immunoglobulin E epitopes, *Clin. Exp. Allergy 33*, 518−524.
70. Walsh, B. J., and Howden, M. E. (1989) A method for the detection of IgE binding sequences of allergens based on a modification of epitope mapping, *J. Immunol. Methods 121*, 275−280.

BI048069X