JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY
ARTICLE

J. Agric. Food Chem. XXXX, XXX, 000–000 A
DOI:10.1021/jf9017355

# Crystal Structure of Prunin-1, a Major Component of the Almond (*Prunus dulcis*) Allergen Amandin

TENGCHUAN JIN,[†,⊥] SILVIA M. ALBILLOS,[‡,#] FENG GUO,[†] ANDREW HOWARD,[†] TONG-JEN FU,*[,§] MAHENDRA H. KOTHARY,[∥] AND YU-ZHU ZHANG*[,†,‡]

[†]Department of Biological, Chemical, and Physical Sciences, Illinois Institute of Technology, 3101 South Dearborn Street, Chicago, Illinois 60616, [‡]Illinois Institute of Technology, National Center for Food Safety and Technology, 6502 South Archer Road, Summit-Argo, Illinois 60501, [§]National Center for Food Safety and Technology, U.S. Food and Drug Administration, 6502 South Archer Road, Summit-Argo, Illinois 60501, and [∥]Center for Food Safety and Applied Nutrition, U.S. Food and Drug Administration, 8301 Muirkirk Road, Laurel, Maryland 20708. [⊥]Present address: National Institute of Allergy and Infectious Diseases, National Institutes of Health, 4 Memorial Drive, Bethesda, Maryland 20892. [#]Present address: Institute of Biotechnology, INBIOTEC, Av del Real 1, 24006 León, Spain

Downloaded by UNIV DI ROMA TOR VERGATA on September 1, 2009 | http://pubs.acs.org
Publication Date (Web): August 20, 2009 | doi: 10.1021/jf9017355

Seed storage proteins are accumulated during seed development and act as a reserve of nutrition for seed germination and young sprout growth. Plant seeds play an important role in human nutrition by providing a relatively inexpensive source of protein. However, many plant foods contain allergenic proteins, and the number of people suffering from food allergies has increased rapidly in recent years. The 11S globulins are the most widespread seed storage proteins, present in monocotyledonous and dicotyledonous seeds as well as in gymnosperms (conifers) and other spermatophytes. This family of proteins accounts for a number of known major food allergens. They are of interest to both the public and industry due to food safety concerns. Because of the interests in the structural basis of the allergenicity of food allergens, we sought to determine the crystal structure of Pru1, the major component of the 11 S storage protein from almonds. The structure was refined to 2.4 Å, and the R/Rfree for the final refined structure is 17.2/22.9. Pru1 is a hexamer made of two trimers. Most of the back-to-back trimer−trimer association was contributed by monomer−monomer interactions. An α helix (helix 6) at the C-terminal end of the acidic domain of one of the interacting monomers lies at the cleft of the two protomers. The residues in this helix correspond to a flexible region in the peanut allergen Ara h 3 that encompasses a previously defined linear IgE epitope.

**KEYWORDS: Almond (*Prunus dulcis*); prunin; Pru1; amandin; food allergy; X-ray crystallography**

## INTRODUCTION

Seed storage proteins are accumulated in developing seeds. They are believed to act as a reserve for nitrogen (*1*, *2*). Plant seeds play an important role in human nutrition by providing a relatively inexpensive protein source. However, many plant foods contain proteins that elicit allergic responses in a subset of the human population. The number of people suffering from food allergies has increased in recent years (*3*−*5*). Inadvertent exposure to allergens has become an important food safety issue faced by both the consumer and the food industry.

There are thousands of proteins in mature seeds of a plant (*6*, *7*), but only a few of them are known to elicit allergic reactions in susceptible patients. These include the seed storage proteins, with the 11S globulins being the most widespread. The 11S proteins are present in both monocotyledonous and dicotyledonous seeds as well as in gymnosperms (conifers) and other spermatophytes. This family of proteins includes many known major food allergens, including Ara h 3 and Ara h 4 in peanut (*Arachis hypogaea*) (*8*), Cor a 9 in hazelnut (*Corylus avellana*) (*9*), Jug r 4 in English walnut (*Juglans regia*) (*10*), Ana o 2 in cashew nut (*Anacardium occidentale*) (*11*), Sin a 2 in yellow mustard (*Sinapis alba*) (*12*), Ses i 6 and Ses i 7 in sesame (*Sesamum indicum*) (*13*), and amandin in almond (*Prunus dulcis*) (*14*).

Almonds are fruit seeds but are often classified as tree nuts (*15*). Tree nuts are one of the eight major sources of food allergies (*16*−*18*), affecting approximately 1% of the general population in the United States (*19*). Unlike allergies to other foods, such as egg and milk, which mostly affect children and tend to disappear in adults, nut-induced allergies are often permanent (*20*−*23*). In a study of 54 patients with tree nut allergies (*24*), almond was the second highest cause of allergic reactions.

The most abundant storage protein in almond is amandin, which has been reported as a major allergen (*14*, *25*). Amandin belongs to the 11S globulin family (*26*). Sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS-PAGE) analysis of amandin showed the presence of peptide bands of molecular masses of 44, 42, and 28 kDa (*27*) or 41, 39, 22, and 21 kDa (*28*). cDNAs encoding two almond seed storage proteins designated

*To whom correspondence should addressed. (T.-J.F.) Tel: 708-728-4149. Fax: 708-728-4177. E-mail: tong.fu@fda.hhs.gov. (Y.-Z.Z.) Tel: 312-567-3484. Fax: 312-567-3494. E-mail: yuzhu.zhang@iit.edu.

© XXXX American Chemical Society

prunin-1 (Pru1) and prunin-2 (Pru2) have been cloned and sequenced (*29*). Pru1 and Pru2 are each composed of two polypeptides, a 40 kDa N-terminal subunit and a 20 kDa C-terminal subunit, which are linked by a disulfide bond (*29*). Recently, Pru1 was identified as the major component of the almond allergen amandin (*30*). The characterization of the acidic and basic polypeptides of this almond allergen (*31*) has been reported.

Food allergies are mostly IgE-mediated hypersensitivity diseases and are potentially life-threatening (*32*). In recent years, animal models and clinical studies have helped understand the pathophysiology of food allergies (*33*). However, to date, no effective treatment for food allergies is available. The recommended care for individuals with food allergies focuses on avoidance of the offending food, recognition of early symptoms of an allergic reaction, and prompt administration of epinephrine injection (*34*) in cases of accidental ingestion.

The concept of utilizing blocking IgG antibodies to compete with IgE for binding with allergens and to prevent IgE-mediated responses has existed for more than 70 years (*35*). The clinical benefits of specific-allergen immunotherapy are also well-documented (*36*). However, immunotherapy with food allergen extracts is associated with high anaphylactic risk and is not advised (*37*, *38*). Nevertheless, treatments with short peptides based on dominant T-cell epitopes, allergen fragments, and mutant hypoallergenic recombinant allergens in combination with Th1-inducing adjuvants appear to be the most promising approaches for treating allergies (*37*, *39*).

Knowledge of the molecular nature of allergen−antibody interactions is crucial for understanding the mechanism of conventional immunotherapy as well as for designing alternative immunotherapeutic strategies. Molecular characterization of food allergens along with increased understanding of the pathophysiology and immunologic mechanisms of food allergy should lead to more effective and safer ways of combating allergy-related diseases. As information concerning the IgE reactive sites is crucial in this regard, major research efforts have been made during recent years to characterize allergen epitopes (*40*, *41*). Although most of these studies have focused on linear epitopes, the importance of conformational epitopes has also been recognized (*42*). Determination of the three-dimensional structures of allergens is required for the understanding of conformational epitopes of food allergens. In this paper, we report the X-ray crystal structure of Pru1.

## MATERIALS AND METHODS

**Structure Refinement.** The purification, crystallization, and characterization of the crystals of Pru1 have been reported (*30*). Since then, new data have been collected at the SER-CAT 22BM beamline at the Advanced Photon Source (APS), Argonne National Laboratory. Molec-

**Table 1.** X-ray Crystallographic Statistics and Refinement

| | |
|---|---|
| data collection | |
| wavelength (Å) | 1.0 |
| temperature (K) | 110 |
| space group | $P4_1$ |
| $a = b$, $c$ (Å) | 151.02, 165.37 |
| resolution limits (Å)[a] | 36.08−2.40 (2.49−2.40) |
| no. of observed reflections | 826259 |
| no. of unique reflections[a] | 142516 (13994) |
| completeness (%)[a] | 99.2 (97.7) |
| mean$I(\sigma)$[a] | 7.78 (1.63) |
| $R_{sym}$(%) [a,b] | 12.9 (60.8) |
| refinement | |
| resolution range (Å) | 36.08−2.40 |
| no. of reflections | |
| working set | 122566 |
| test set[c] | 3818 |
| no. of atoms | |
| protein | 19199 |
| water | 743 |
| ligands and ion | 15 |
| $R/R_{free}$ (%) | 17.2/22.9 |
| rmsd from ideal geometry | |
| bond lengths (Å) | 0.015 |
| bond angles (deg) | 2.001 |

[a] The numbers in parentheses are in the outer shell. [b] $R_{sym} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i (hkl)$, where $I_i(hkl)$ is the $i$th observation of reflection $hkl$ and $\langle I(hkl)\rangle$ is the weighted average intensity for all observations $i$ of reflection $hkl$. [c] Three percent of the total reflections from thin shells in the center of each reflection bin.

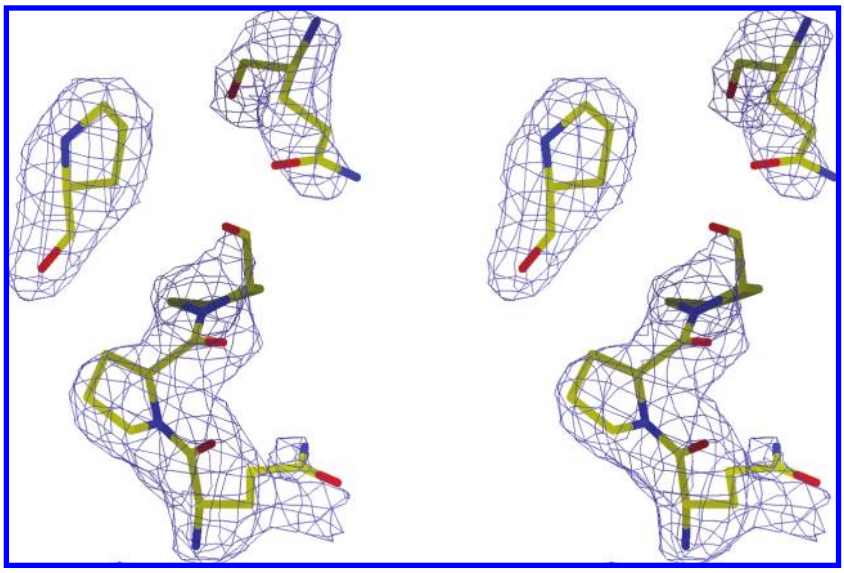

**Figure 1.** Stereoview of a shake-and-omit Fo-Fc electron density map contoured at 3.0 $\sigma$. The map was calculated with Q335, P336, and P337 from protomer E and Q352 and P385 from protomer B (which are the closest residues to P337/E, the last residue can be located for the N-terminal domain of protomer E) omitted. The omitted residues in the final structure are shown in a stick representation with the CPK coloring scheme.

Downloaded by UNIV DI ROMA TOR VERGATA on September 1, 2009 | http://pubs.acs.org
Publication Date (Web): August 20, 2009 | doi: 10.1021/jf9017355

Downloaded by UNIV DI ROMA TOR VERGATA on September 1, 2009 | http://pubs.acs.org
Publication Date (Web): August 20, 2009 | doi: 10.1021/jf9017355

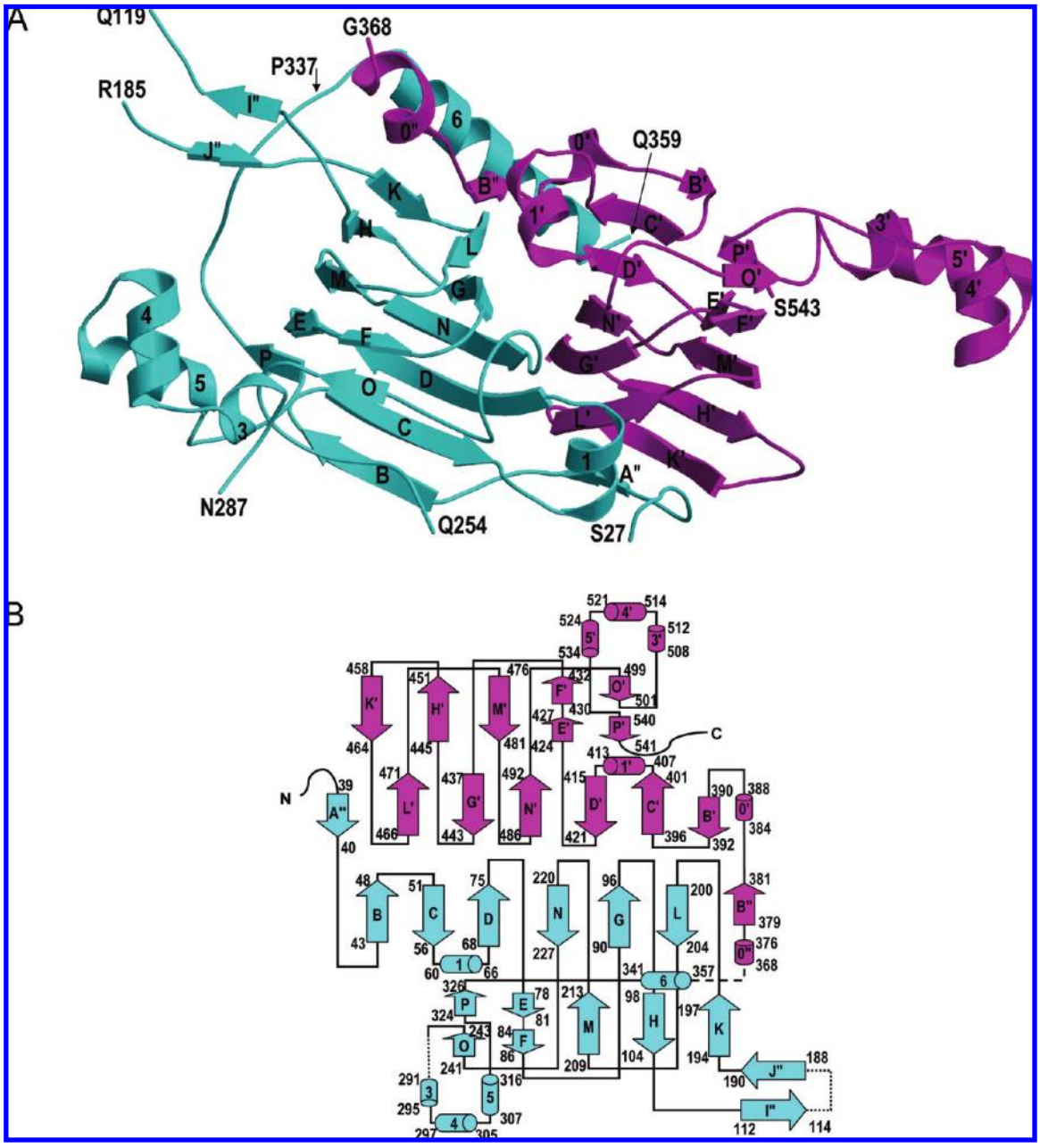


**Figure 2.** Structure of Pru1. (**A**) A ribbon diagram of protomer F. The N-terminal domain is shown in cyan, and the C-terminal domain is shown in magenta. Strands are labeled alphabetically, and helixes are numbered. Element labels in the C-terminal domain are primed. The first strand in the N-terminal peptide is part of a $\beta$-sheet of the C-terminal domain and is double-primed. Similarly, the first strand in the C-terminal peptide is double-primed. Strands I and J in the N-terminal domain are also double primed because they would become part of a $\beta$-sheet of the C-terminal of another protomer in the hexameric native allergen. (**B**) A schematic diagram shows the N- and C-terminal domain cupin fold. The coloring and labeling schemes are the same as in part A. Dotted lines indicate unlocated flexible loops, and the broken line indicates that the peptide connection between the N- and C-terminal domains was cleaved. The residue positions of the starts and ends of the helixes and strands are indicated.

ular replacement calculations were carried out using the program Phaser (*43*, *44*), implemented in CCP4i (*45*). The model was first refined using CNS (*46*) with simulated annealing and then using the program REFMAC5 (*47*) implemented in CCP4i (*45*). Structure refinement was alternated with model building and model improvement using Coot (*48*).

The final structure was refined with all data to 2.40 Å resolution, and the final model was checked by PROCHECK (*49*) and Molprobity Validation (*50*). The quality of the structure model was also checked with a shake-and-omit protocol by first introducing random errors up to 0.3 Å to the coordinates of the final refined structure using the program *pdbset* distributed with CCP4 (*45*). Each region to be tested was manually omitted from the shaken structure. Then, 20 cycles of restrained refinements were carried out using REFMAC5 (*47*). This was followed by inspecting the Fo-Fc map together with the final refined structure. The atomic coordinates and structure factors for Pru1 were deposited in the Protein Data Bank (3FZ3). Molecular graphic figures were prepared using the programs Rasmol (*51*), MolScript (*52*), Raster3D (*53*), and Pymol (*54*). Structure alignment was carried out with the program MUSTANG (*55*).

**Preparation of Crude Protein Extract from Tree Nuts for SDS-PAGE Analysis.** Crude protein extracts from hazelnut, pecan, walnut, and almond were prepared separately using a common protocol. The nuts were first ground and then extracted in 1 M NaCl at 60 °C for 20 min. The extract was centrifuged at 20199*g* for 5 min, and the aqueous layer was collected and centrifuged again. The crude protein extracts were mixed with two volumes of 8 M urea to denature the proteins. The samples were

Downloaded by UNIV DI ROMA TOR VERGATA on September 1, 2009 | http://pubs.acs.org
Publication Date (Web): August 20, 2009 | doi: 10.1021/jf9017355

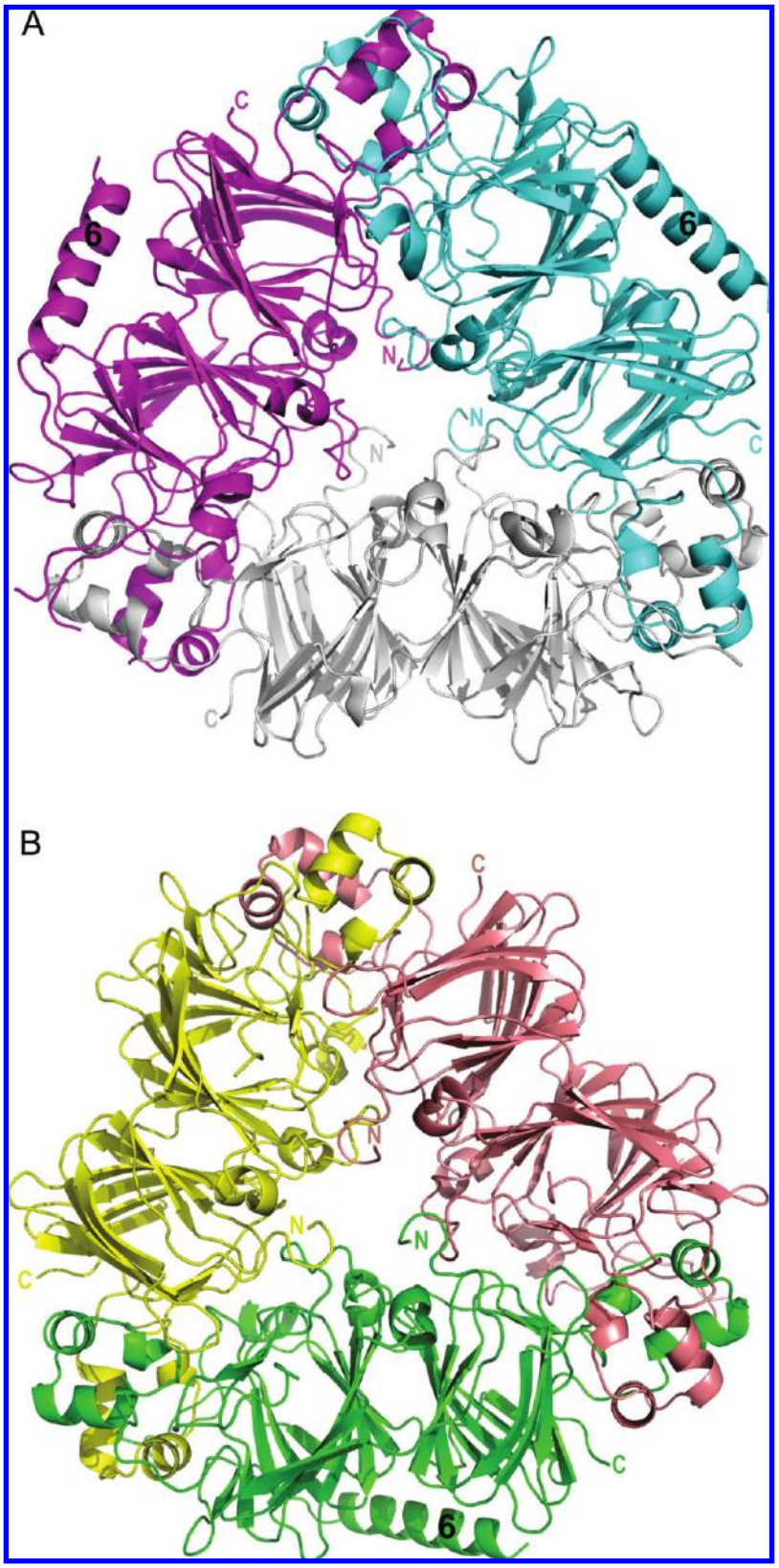


**Figure 3.** Ribbon diagram presentation of Pru1 trimers. (**A**) Protomers A (gray), B (cyan), and C (magenta) form a trimer with only protomer A lacking helix 6. The trimer is viewed from the trimer−trimer interface. (**B**) Protomers D (yellow), E (pink), and F (green) form a trimer with only protomer F containing helix 6. This trimer is viewed from the side opposite to the trimer−trimer interface. The N- and C-terminal of the Pru1 structure as well as helix 6 are labeled.

then mixed with an equal volume of 2× SDS sample buffer [50 mM Tris−HCl, pH 6.8, 2% (v/v) SDS, 0.1% (w/v) bromophenol blue, 10% (v/v) glycerol, and 100 mM dithiothreitol] and boiled for 5 min. The samples were centrifuged at 20199*g* for 1 min and were analyzed by SDS-PAGE using homecast 4−20% gradient gels.

## RESULTS

**Structure of Pru1.** Although the data could be indexed in the space group $P4_122/P4_322$ (*30*), processing the data in the lower symmetry space group $P4_1$ resulted in a slightly lower $R_{sym}$. Using

Downloaded by UNIV DI ROMA TOR VERGATA on September 1, 2009 | http://pubs.acs.org
Publication Date (Web): August 20, 2009 | doi: 10.1021/jf9017355

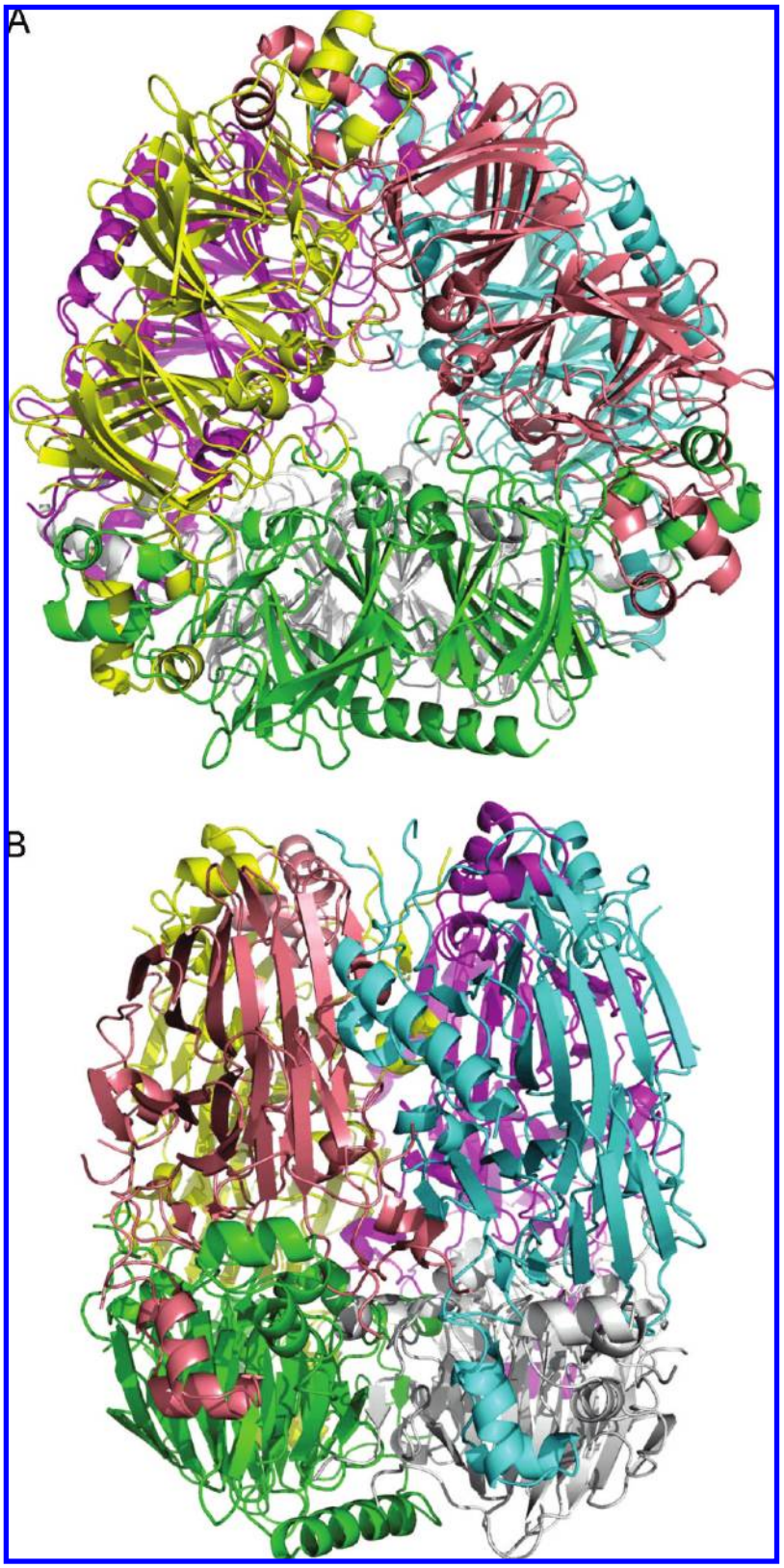


**Figure 4.** Ribbon diagram of a Pru1 hexamer. (**A**) The hexamer is formed by the "back-to-back" association of two trimers. (**B**) The hexamer viewed after part **A** is rotated 90° about a vertical axis pointing down.

the initial structure model of the old data set (*30*) as a template, a Pru1 hexameric structure in the $P4_1$ asymmetric unit was obtained by molecular replacement calculation. The final model of the Pru1 structure determined at 2.40 Å gave $R/R_{free}$ values of 17.2%/22.9% (**Table 1** and **Figure 1**). It included a total of 2394 amino acid residues, 743 water molecules, 6 $Ca^{2+}$, and 9 $Na^{2+}$. The rmsd for bonds was 0.015 Å from ideal small molecule values, and all covalent bonds lie within a 6.0 $\sigma$ range from the standard values. The rmsd for angles was 2.00° with two angle deviations slightly larger than 6.0 $\sigma$ from the standard values. On the Ramachandran plot for nonglycine or nonproline residues, 85.6% of the residues were in the most favored regions, 13.5%

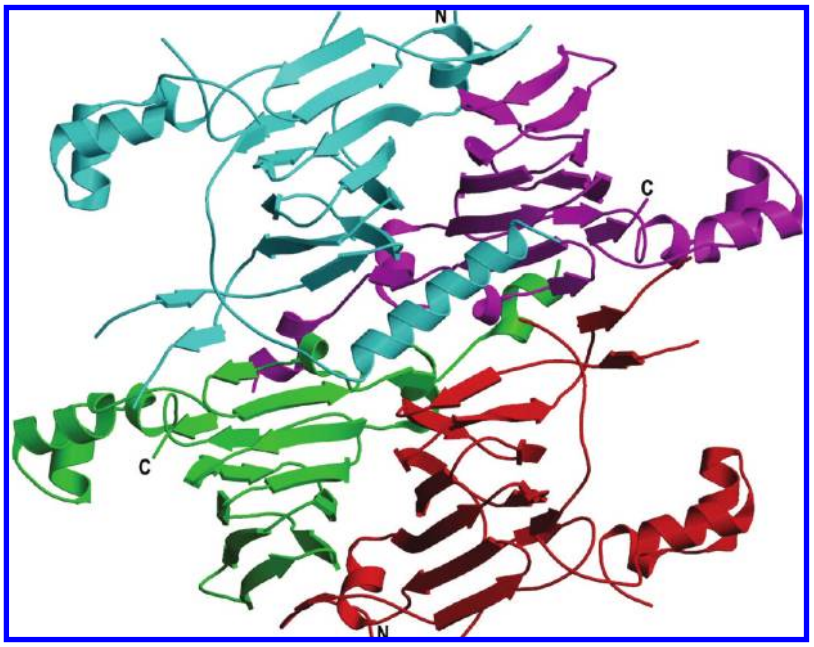


**Figure 5.** Ribbon diagram showing helix 6 situated at the cleft of two interacting protomers from different trimers in the hexamer. The N- and C-terminal domains of protomer A are shown in red and green, respectively, and the N- and C-terminal domains of protomer F are shown in cyan and magenta, respectively.

Downloaded by UNIV DI ROMA TOR VERGATA on September 1, 2009 | http://pubs.acs.org
Publication Date (Web): August 20, 2009 | doi: 10.1021/jf9017355

were in the additional allowed regions, and 0.9% were in the generously allowed regions, with one outlier falling just outside of the allowed regions.

The final refined structure of Pru1 contained six protomers in the crystallographic asymmetric unit. Each protomer consisted of an N-terminal subunit and a C-terminal subunit that are known to be separated into two peptide chains by post-translational peptidase cleavage. Both the N- and the C-terminal domains contained a cupin fold, and each cupin structure was connected with a helical region containing three short helixes. The N- and C-terminal domains were related by a pseudodyad axis (**Figure 2A**, **B**). For the N-terminal domain, two flexible loops (S118−F184 and Q255−G286) could not be located in the electron density map and were not included in the model. In addition, the N-terminal residues A21−L26 and the C-terminal residues R338−N367 in protomers A, D, and E and G360−N367 in protomers B, C, and F were also not located in the electron density map (**Figure 2A**). For the C-terminal domain, eight residues at its C-terminal (S544−V551) were not located in the electron density map. For the side chains of residues R394/A, N287/B, R340/B, N287/C, Q341/C, N287/F, and R340/F, only the $\beta$-carbon could be located and included in the final structure, and for residue R394/E, only the $\beta$- and $\gamma$-carbons of its side chain were located and included.

The six protomers formed a tightly packed hexamer that could be considered as two trimers associated back-to-back. One trimer consisted of two protomers with helix 6 and one protomer without helix 6 (**Figure 3A**), while the other trimer consisted of two protomers without helix 6 and one protomer with helix 6 (**Figure 3B**). Within a trimer, the helical regions of the adjacent protomers were packed against each other, and they contribute significantly to the interactions between adjacent protomers (**Figure 3**). Head-to-tail associations between monomers result in a doughnut-shaped trimer with a pseudo noncrystallographic 3-fold axis.

The back-to-back association of two trimers formed the native Pru1 hexamer (**Figure 4**). Most of the trimer−trimer interaction was contributed by monomer−monomer interactions. As shown in **Figure 5**, the two associating monomers were coupled by $\beta$-strands paring between the N-terminal domain of one monomer and the C-terminal domain of the other. They were also related by a pseudo noncrystallographic 2-fold axis, but only one of them had helix 6 occupying the groove at their interface (**Figures 4** and **5**). This monomer−monomer interaction buried $2932 \pm 66$ Å$^2$ solvent accessible areas. In comparison, the solvent accessible areas involved in intermolecular surfaces for each monomer within the doughnut-shaped trimer was $5343 \pm 44$ Å$^2$. Implementation of noncrystallographic symmetry relationships in the refinement did not lead to an improvement of the structural factor.

**Comparison with the 11S Seed Storage Proteins from Other Organisms.** The 11S seed storage protein is known to be a hexamer in its native state (*56*). The structures of one isoform of Ara h 3, an 11S storage protein and a major allergen from peanut (*57*), and of the soybean glycinin A3B4 produced in a genetically modified soybean (*58*) have been reported. Structure alignment using the program MUSTANG (*55*) showed that without helix 6, the overall structure of Pru1 was very similar to that of Ara h 3 and glycinin A3B4 (data not shown). However, neither of these two structures contained a helix corresponding to helix 6 in Pru1.

As shown in **Figure 6A**, multiple sequence alignment of the 11S seed storage proteins from different species indicated that the C terminus of the N-terminal domain was not conserved. In pistachio and Brazil nuts, this region was very short, while in the 11S proteins of some other species, including soybean and peanut, it was much longer. SDS-PAGE analyses (**Figure 6B**) showed that the N-terminal domain of Pru1, as well as the 11S proteins in hazelnut, pecan, and walnut migrated as two bands, whereas previous reports had shown that the N-terminal domains of peanut Ara h 3 and Brazil nut Ber e 2 appeared to be a single band (*59*, *60*).

## DISCUSSION

A number of plants are known to contain different isoforms of the 11S globulins, and it is known that each of the isoforms

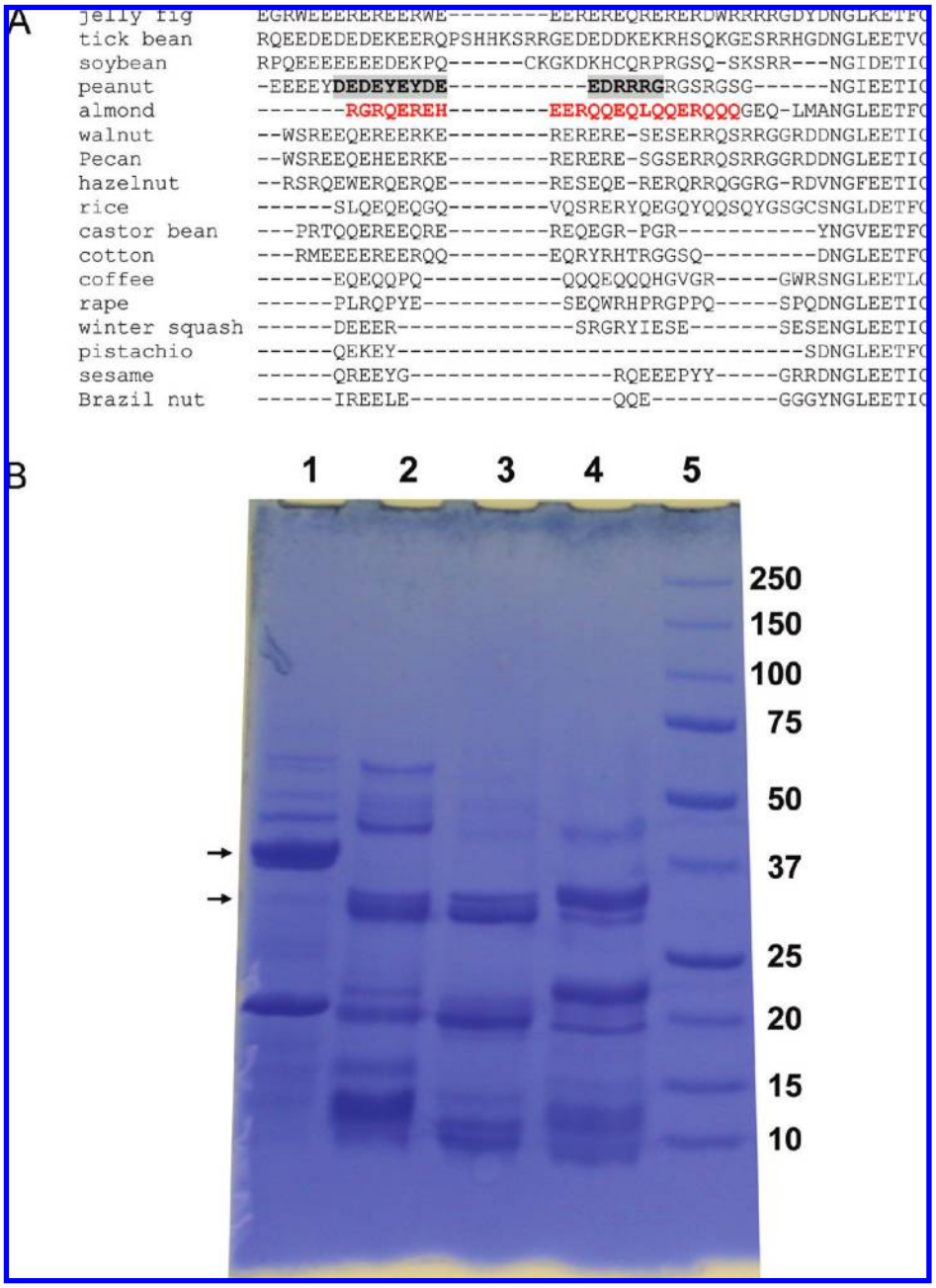


**Figure 6.** Comparison of the 11S seed storage proteins from different species. (**A**) A portion of a ClustalW multiple sequence alignment of the 11S proteins covering helix 6 in Pru1 (red bold face letters) and the corresponding regions in the orthologues: The database access IDs are as follows: jelly fig (*Ficus pumila*), ABK80751; tick bean (*Vicia faba*), CAA38758; soybean (*Glycine max*), P04776; peanut (*A. hypogaea*), ACH91862; almond (*P. dulcis*), CAA55009; English walnut (*J. regia*), AAW29810; pecan (*Carya illinoinensis*), ABW86979; hazelnut (*C. avellana*), AAL73404; rice (*Oryza sativa*), 1210248A; castor bean (*Ricinus communis*), AAF73007; cotton (*Gossypium hirsutum*), P09802; coffee (*Coffea arabica*), AAC61983; rape (*Brassica napus*), P33525; winter squash (*Cucurbita maxima*), P13744; pistachio (*Pistacia vera*), ABU42022; sesame (*S. indicum*), ABB60055; and Brazil nut (*Bertholletia excelsa*), AAO38859. Linear IgE epitope 4 in peanut Ara h 3 as defined by Rabjohn et al. (*8*) is shown with a shaded background. (**B**) SDS-PAGE analysis suggests the presence of helix 6 in a subset of the 11S seed storage proteins. Crude protein extracts from almond, pecan, walnut, and hazelnut (lanes 1−4, respectively) were analyzed by SDS-PAGE using 4−20% gradient gels. On the basis of the property and the abundance of the 11S seed storage proteins, the double bands at ∼35−43 kDa indicate the presence of different post-translationally processed N-terminal domains. The molecular masses (in kDa) of the protein standards (lane 5) are given at the right side of the gel image. Arrows point to the doublets in the proteins.

of the 11S globulin is genetically encoded by a single gene (*13*, *29*). Its protein precursor is post-translationally cleaved by an asparaginyl endopeptidase, leaving a disulfide bond between the N-terminal and the C-terminal subunits as the only interchain link (*61*). The cleavage site between Asn-Gly peptide bond is well-conserved among a wide variety of plant species (*62*). Pru1 and several other 11S allergens from tree nuts and peanut possess the NGXEET motif: NGLEET in Pru1 from almond and Jug r 4 from walnut, NGFEET in Cor a 9 from hazelnut, and NGIEET in both Ana o 2 from cashew and Ara h 3 from peanut.

With N-terminal amino acid sequencing using Edman degradation, we have shown that purified Pru1 is the result of a single gene *pru1* and the mature Pru1 starts from residue A21 of the proprotein translated from the open reading frame of the *prunin*−1 gene (*30*). We have also shown that the NGLEET sequence is a bona fide recognition site for post-translational cleavage by an

Downloaded by UNIV DI ROMA TOR VERGATA on September 1, 2009 | http://pubs.acs.org
Publication Date (Web): August 20, 2009 | doi: 10.1021/jf9017355

endopeptidase that cleaves it between N367 and G368 (*30*). In the Pru1 structure, three of the protomers did not contain helix 6. The C terminus (P337) of the N-terminal domain of these protomers was not buried inside the native hexamer. Although it is possible that residues 337−367 are flexible in these protomers, the structural data would support the possibility that 30 residues at the C terminus may have been cleaved from the N-terminal domains of these protomers. As a result, the N-terminal domains of three protomers are 30 residues shorter than those of the other three protomers in the native hexamer. This is consistent with the N-terminal domain of Pru1 migrating at two distinctive rates in SDS-PAGE gels (*30*). Data on the C terminus of the N-terminal domains are not available as a large amount of the two forms of the N-terminal domain is needed to determine their C-terminal amino acid sequences by C-terminal peptide sequencing. As the sequences flanking the peptide bond between P337 and R338 are not conserved, it seems unlikely that there is a conserved peptidase that is responsible for the cleavage of the peptide bond. Nonetheless, the N-terminal domains of the 11S seed storage proteins from a number of different species also migrated with two distinct rates in SDS-PAGE gels. Assuming that Pru1 and the 11S proteins in a subset of plant species are processed to have different N-terminal domains, it remains to be determined whether the C-terminal peptide of the N-terminal domain exists in the seed as a separate unit. In addition, it remains to be investigated whether the truncation of the C terminus of this domain is caused by an endopeptidase-catalyzed cleavage or by stepwise C-terminal degradation.

The crystal structures of the soybean glycinin A3B4 and the peanut major allergen Ara h 3 have been reported (*57*,*58*). In both proteins, the residues corresponding to those in helix 6 of Pru1 (**Figure 6A**) were not located in the electron density map. Whether these residues were cleaved in the seed and not present in the crystals remains to be determined. In the case of peanut Ara h 3, four dominant linear IgE epitopes were previously defined (*8*), and one of these epitopes (the last one in the peptide sequence) resides in the region corresponding to helix 6 in Pru1 in the sequence alignment (**Figure 6A**). This suggests that the C terminus of the N-terminal domain of Ara h 3 may be present in the seed regardless of whether it is an integral part of the Ara h 3 molecule or as a free peptide. The property of this region may need to be considered when the allergenicity of the 11S proteins in different seeds is assessed (*63*).

In summary, the crystal structure of Pru1 was determined at 2.40 Å. Analysis of the structure provided a possible explanation for the phenomenon whereby the N-terminal peptides of many 11S seed storage proteins appear as doublets in SDS-PAGE gels. The IgE epitopes of Pru1 or amandin have not been identified, and it remains to be determined whether any IgE binding epitope(s) resides in helix 6. The information regarding Pru1 obtained in this study along with the crystal structures of other 11S allergens in peanut and tree nuts will facilitate future characterization of IgE binding epitopes of the 11S allergens and may also help in characterizing the cross-reactivity between these proteins.

## LITERATURE CITED


(1) Utsumi, S. Plant food protein engineering. *Adv. Food Nutr. Res.* **1992**, *36*, 89–208.

(2) Shewry, P. R. Plant storage proteins. *Biol. Rev. Camb. Philos. Soc.* **1995**, *70*, 375–426.

(3) Branum, A. M.; Lukacs, S. L. *Food Allergy among U.S. Children: Trends in Prevalence and Hospitalizations*. NCHS data brief No. 10. Hyattsville, MD: National Center for Health Statistics, 2008.

(4) Sampson, H. A. Update on food allergy. *J. Allergy Clin. Immunol.* **2004**, *113*, 805−819; quiz 820.

(5) Wang, J.; Sampson, H. A. Food anaphylaxis. *Clin. Exp. Allergy* **2007**, *37*, 651–660.

(6) Gallardo, K.; Job, C.; Groot, S. P.; Puype, M.; Demol, H.; Vandekerckhove, J.; Job, D. Proteomic analysis of arabidopsis seed germination and priming. *Plant Physiol.* **2001**, *126*, 835–848.

(7) Mooney, B. P.; Krishnan, H. B.; Thelen, J. J. High-throughput peptide mass fingerprinting of soybean seed proteins: Automated workflow and utility of UniGene expressed sequence tag databases for protein identification. *Phytochemistry* **2004**, *65*, 1733–1744.

(8) Rabjohn, P.; Helm, E. M.; Stanley, J. S.; West, C. M.; Sampson, H. A.; Burks, A. W.; Bannon, G. A. Molecular cloning and epitope analysis of the peanut allergen Ara h 3. *J. Clin. Invest.* **1999**, *103*, 535–542.

(9) Beyer, K.; Grishina, G.; Bardina, L.; Grishin, A.; Sampson, H. A. Identification of an 11S globulin as a major hazelnut food allergen in hazelnut-induced systemic reactions. *J. Allergy Clin. Immunol.* **2002**, *110*, 517–523.

(10) Wallowitz, M.; Peterson, W. R.; Uratsu, S.; Comstock, S. S.; Dandekar, A. M.; Teuber, S. S. Jug r 4, a legumin group food allergen from walnut (*Juglans regia* Cv. Chandler). *J. Agric. Food Chem.* **2006**, *54*, 8369–8375.

(11) Wang, F.; Robotham, J. M.; Teuber, S. S.; Sathe, S. K.; Roux, K. H. Ana o 2, a major cashew (*Anacardium occidentale* L.) nut allergen of the legumin family. *Int. Arch. Allergy Immunol.* **2003**, *132*, 27–39.

(12) Palomares, O.; Cuesta-Herranz, J.; Vereda, A.; Sirvent, S.; Villalba, M.; Rodriguez, R. Isolation and identification of an 11S globulin as a new major allergen in mustard seeds. *Ann. Allergy Asthma Immunol.* **2005**, *94*, 586–592.

(13) Hsiao, E. S.; Lin, L. J.; Li, F. Y.; Wang, M. M.; Liao, M. Y.; Tzen, J. T. Gene families encoding isoforms of two major sesame seed storage proteins, 11S globulin and 2S albumin. *J. Agric. Food Chem.* **2006**, *54*, 9544–9450.

(14) Roux, K. H.; Sathe, S. K.; Peterson, W. R.; Teuber, S. S. The major seed storage protein of almond (almond major protein) is an allergen. *J. Allergy Clin. Immunol.* **1999**, *103*, S66.

(15) Angus, F. Nut allergens. In *Natural Toxicants in Foods*; Watson, D. H., Ed.; CRC Press LLC: Boca Raton, FL, 1998; pp 84−104.

(16) Food and Agriculture Organization of the United Nations, *Report of the FAO Technical Consultation on Food Allergies*; November 13−14, 1995.

(17) U.S. Food and Drug Administration, Section 555−250 Statement of policy for labeling and preventing cross-contact of common food allergens. Revised **2005**. Webpage: http://www.fda.gov/ora/compliance_ref/cpg/cpgfod/cpg555-250.htm (Accessed: May 26, **2007**).

(18) Lehrer, S. B.; Ayuso, R.; Reese, G. Current understanding of food allergens. *Ann. N.Y. Acad. Sci.* **2002**, *964*, 69–85.

(19) Crespo, J. F.; James, J. M.; Fernandez-Rodriguez, C.; Rodriguez, J. Food allergy: Nuts and tree nuts. *Br. J. Nutr.* **2006**, *96* (Suppl. 2), S95–S102.

(20) Schreiber, R. A.; Walker, W. A. Food allergy: Facts and fiction. *Mayo Clin. Proc.* **1989**, *64*, 1381–1391.

(21) Sicherer, S. H. Clinical update on peanut allergy. *Ann. Allergy Asthma Immunol.* **2002**, *88*, 350−361; quiz 361−362, 394.

(22) Hourihane, J. O. Prevalence and severity of food allergy—Need for control. *Allergy* **1998**, *53*, 84–88.

(23) Sampson, H. A. Food allergy. Part 1: Immunopathogenesis and clinical disorders. *J. Allergy Clin. Immunol.* **1999**, *103*, 717–728.

(24) Sicherer, S. H.; Burks, A. W.; Sampson, H. A. Clinical features of acute allergic reactions to peanut and tree nuts in children. *Pediatrics* **1998**, *102*, e6.

(25) Roux, K. H.; Teuber, S. S.; Robotham, J. M.; Sathe, S. K. Detection and stability of the major almond allergen in foods. *J. Agric. Food Chem.* **2001**, *49*, 2131–2136.

(26) Svedberg, T.; Sjogren, B. The molecular weights of amandin and excelsin. *J. Am. Chem. Soc.* **1930**, *52*, 279–287.

(27) Wolf, W. J. Gel electrophoresis and amino acid analysis of the nonprotein nitrogen fractions of defatted soybean and almond meals. *Cereal Chem.* **1995**, *72*, 115–121.

(28) Sathe, S. K. Solubilization, electrophoretic characterization and *in vitro* digestibility of almond (*Prunus amygdalus*) proteins. *J. Food Biochem.* **1993**, *16*, 249–264.


Downloaded by UNIV DI ROMA TOR VERGATA on September 1, 2009 | http://pubs.acs.org
Publication Date (Web): August 20, 2009 | doi: 10.1021/jf9017355

Downloaded by UNIV DI ROMA TOR VERGATA on September 1, 2009 | http://pubs.acs.org
Publication Date (Web): August 20, 2009 | doi: 10.1021/jf9017355

(29) Garcia-Mas, J.; Messeguer, R.; Arus, P.; Puigdomenech, P. Molecular characterization of cDNAs corresponding to genes expressed during almond (*Prunus amygdalus* Batsch) seed development. *Plant Mol. Biol.* **1995**, *27*, 205–210.
(30) Albillos, S. M.; Jin, T.; Howard, A.; Zhang, Y.; Kothary, M. H.; Fu, T. J. Purification, crystallization and preliminary X-ray characterization of prunin-1, a major component of the almond (*Prunus dulcis*) allergen amandin. *J. Agric. Food Chem.* **2008**, *56*, 5352–5358.
(31) Albillos, S. M.; Menhart, N.; Fu, T.-J. Structural stability of amendin, a major allergen from almond (*Prunus dulcis*), and its acidic and basic polypepides. *J. Agric. Food Chem.* **2009**, *57*, 4698−4705.
(32) Bock, S. A.; Munoz-Furlong, A.; Sampson, H. A. Fatalities due to anaphylactic reactions to foods. *J. Allergy Clin. Immunol.* **2001**, *107*, 191–193.
(33) Untersmayr, E.; Jensen-Jarolim, E. Mechanisms of type I food allergy. *Pharmacol. Ther.* **2006**, *112*, 787–798.
(34) Grimshaw, K. E. Dietary management of food allergy in children. *Proc. Nutr. Soc.* **2006**, *65*, 412–417.
(35) Cooke, R. A.; Barnard, J. H.; Hebald, S.; et al. Serological evidence of immunity with coexisting sensitization in a type of human allergy (hay fever). *J. Exp. Med.* **1935**, *62*, 733–750.
(36) Bousquet, J.; Lockey, R.; Malling, H. J.; Alvarez-Cuesta, E.; Canonica, G. W.; Chapman, M. D.; Creticos, P. J.; Dayer, J. M.; Durham, S. R.; Demoly, P.; Goldstein, R. J.; Ishikawa, T.; Ito, K.; Kraft, D.; Lambert, P. H.; Lowenstein, H.; Muller, U.; Norman, P. S.; Reisman, R. E.; Valenta, R.; Valovirta, E.; Yssel, H. Allergen immunotherapy: Therapeutic vaccines for allergic diseases. World Health Organization. American Academy of Allergy, Asthma and Immunology. *Ann. Allergy Asthma Immunol.* **1998**, *81*, 401–405.
(37) Nieuwenhuizen, N. E.; Lopata, A. L. Fighting food allergy: Current approaches. *Ann. N.Y. Acad. Sci.* **2005**, *1056*, 30–45.
(38) de Leon, M. P.; Rolland, J. M.; O'Hehir, R, E. The peanut allergy epidemic: Allergen molecular characterisation and prospects for specific therapy. *Expert Rev. Mol. Med.* **2007**, *9*, 1–18.
(39) Rolland, J. M.; Douglass, J.; O'Hehir, R. E. Allergen immunotherapy: Current and new therapeutic strategies. *Expert Opin. Invest. Drugs* **2000**, *9*, 515–527.
(40) Bohle, B. T-cell epitopes of food allergens. *Clin. Rev. Allergy Immunol.* **2006**, *30*, 97–108.
(41) Tanabe, S. Epitope peptides and immunotherapy. *Curr. Protein Pept. Sci.* **2007**, *8*, 109–118.
(42) Nowak-Wegrzyn, A.; Bloom, K. A.; Sicherer, S. H.; Shreffler, W. G.; Noone, S.; Wanich, N.; Sampson, H. A. Tolerance to extensively heated milk in children with cow's milk allergy. *J. Allergy Clin. Immunol.* **2008**, *122*, 342−347, 347 e1-2.
(43) McCoy, A. J.; Grosse-Kunstleve, R. W.; Storoni, L. C.; Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr.* **2005**, *D61*, 458–464.
(44) Storoni, L. C.; McCoy, A. J.; Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr.* **2004**, *D60*, 432–438.
(45) Potterton, E.; Briggs, P.; Turkenburg, M.; Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003**, *59*, 1131–1137.
(46) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr.* **1998**, *D54*, 905–921.
(47) Vagin, A. A.; Steiner, R. A.; Lebedev, A. A.; Potterton, L.; McNicholas, S.; Long, F.; Murshudov, G. N. REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2184–2195.
(48) Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126–2132.
(49) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.
(50) Lovell, S. C.; Davis, I. W.; Arendall, W. B.3rd; de Bakker, P. I.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **2003**, *50*, 437–450.
(51) Sayle, R. A.; Milner-White, E. J. RASMOL: Biomolecular graphics for all. *Trends Biochem. Sci.* **1995**, *20*, 374.
(52) Kraulis, P. J. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **1991**, *24*, 946–950.
(53) Merritt, E. A.; Murphy, M. E. Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr.* **1994**, *D50*, 869–873.
(54) PyMol homepage. Webpage: http://pymol.org/ (Accessed: May 26, **2007**).
(55) Konagurthu, A. S.; Whisstock, J. C.; Stuckey, P. J.; Lesk, A. M. MUSTANG: A multiple structural alignment algorithm. *Proteins* **2006**, *64*, 559–574.
(56) Kao, O. H.; Berns, D. S.; Town, W. R. The characterization of C-phycocyanin from an extremely halo-tolerant blue-green alga, *Coccochloris elabens*. *Biochem. J.* **1973**, *131*, 39–50.
(57) Jin, T.; Guo, F.; Chen, Y.-w.; Howard, A.; Zhang, Y.-Z. Crystal structure of Ara h 3, a major allergen in peanut. *Mol. Immunol.* **2009**, *46*, 1796–1804.
(58) Adachi, M.; Kanamori, J.; Masuda, T.; Yagasaki, K.; Kitamura, K.; Mikami, B.; Utsumi, S. Crystal structure of soybean 11S globulin: glycinin A3B4 homohexamer. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 7395–7400.
(59) Guo, F.; Jin, T.; Howard, A.; Zhang, Y. Z. Purification, crystallization and initial crystallographic characterization of brazil-nut allergen Ber e 2. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2007**, *63*, 976–979.
(60) Jin, T.; Howard, A.; Zhang, Y. Z. Purification, crystallization and initial crystallographic characterization of peanut major allergen Ara h 3. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2007**, *63*, 848–851.
(61) Badley, R. A.; Atkinson, D.; Hauser, H.; Oldani, D.; Green, J. P.; Stubb, J. M. The structure, physical and chemical properties of the soy bean protein glycinin. *Biochim. Biophys. Acta* **1975**, *412*, 214–228.
(62) Dickinson, C. D.; Hussein, E. H.; Nielsen, N. C. Role of posttranslational cleavage in glycinin assembly. *Plant Cell* **1989**, *1*, 459–469.
(63) Hansen, K. S.; Ballmer-Weber, B. K.; Luttkopf, D.; Skov, P. S.; Wuthrich, B.; Bindslev-Jensen, C.; Vieths, S.; Poulsen, L. K. Roasted hazelnuts—Allergenic activity evaluated by double-blind, placebo-controlled food challenge. *Allergy* **2003**, *58*, 132–138.

**Received May 22, 2009. Revised manuscript received July 16, 2009. Accepted July 23, 2009. This work was supported by a fund from the Illinois Institute of Technology and the Cooperative Agreement FD-000431 between the U.S. Food and Drug Administration and the National Center for Food Safety and Technology. X-ray diffraction data were collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-BM beamline at the APS, Argonne National Laboratory. Use of the APS was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract W-31-109-Eng-38.**