JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Purification, Crystallization and Preliminary X-ray Characterization of Prunin-1, a Major Component of the Almond (*Prunus dulcis*) Allergen Amandin

SILVIA M. ALBILLOS,[†] TENGCHUAN JIN,[‡] ANDREW HOWARD,[‡] YUZHU ZHANG,*[,‡] MAHENDRA H. KOTHARY,[§] AND TONG-JEN FU*[,∥]

Illinois Institute of Technology, National Center for Food Safety and Technology, 6502 South Archer Road, Summit-Argo, Illinois 60501, Department of Biology, Illinois Institute of Technology, 3101 South Dearborn Street, Chicago, Illinois 60616, U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition, 8301 Muirkirk Road, Laurel, Maryland 20708, and U.S. Food and Drug Administration, National Center for Food Safety and Technology, 6502 South Archer Road, Summit-Argo, Illinois 60501

The 11S globulins from plant seeds account for a number of major food allergens. Because of the interest in the structural basis underlying the allergenicity of food allergens, we sought to crystallize the main 11S seed storage protein from almond (*Prunus dulcis*). Prunin-1 (Pru1) was purified from defatted almond flour by water extraction, cryoprecipitation, followed by sequential anion exchange, hydrophobic interaction, and size exclusion chromatography. Single crystals of Pru1 were obtained in a screening with a crystal screen kit, using the hanging-drop vapor diffusion method. Diffraction quality crystals were grown after optimization. The Pru1 crystals diffracted to at least 3.0 Å and belong to the tetragonal space group $P4_122$, with unit cell parameters of $a = b = 150.912$ Å, $c = 165.248$ Å. Self-rotation functions and molecular replacement calculations showed that there are three molecules in the asymmetry unit with water content of 51.41%. The three Pru1 protomers are related by a noncrystallographic 3-fold axis and they form a doughnut-shaped trimer. Two prunin trimers form a homohexamer. Elucidation of prunin structure will allow further characterization of the allergenic features of the 11S protein allergens at the molecular level.

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

INTRODUCTION

As many as 6% of young children and 3% to 4% of adults suffer from food allergies worldwide (*1, 2*). Tree nuts are one of the eight major sources of food allergies (*3–5*), affecting approximately 1% of the general population in the U.S. (*6*). Unlike allergies to other foods, such as egg and milk, which mostly occur during childhood and tend to disappear in adults, nut-induced allergies are often permanent (*7–10*). Almonds are fruit seeds, but are often classified as tree nuts (*11*). Almonds rank first in per capita consumption of tree nuts in the U.S. and are responsible for ~15% of tree nut allergy cases (*12*). The most abundant storage protein in almonds termed amandin (*13*), which accounts for ~65% of the total aqueous extractable almond seed proteins (*14*), has been reported as a major allergen in almonds (*15, 16*).

Amandin is water soluble and has been isolated from almond as early as 100 years ago (*13*). Developing almond seed cDNAs which encode two storage proteins of 61.0 and 55.9 kDa, designated prunin-1 (Pru1) and prunin-2 (Pru2), respectively, have been cloned and sequenced (*17*). Pru1 and Pru2 are each composed of two polypeptides linked by a disulfide bond, a 40 kDa acidic α chain and a 20 kDa basic β chain (*17*). These polypeptides have been shown to correspond to amandin peptide bands of 44, 42, and 28 kDa, as reported by Wolf (*18*), or 41, 39, 22, and 21 kDa, as reported by Sathe (*19*). The sedimentation values of amandin reported in the literature range from 11S to 14S (*14, 20–23*).

Amandin is generally referred to as a member of the 11S seed storage globulin family (*24, 25*). The 11S globulins are widely distributed among plant seeds and are believed to accumulate as protein reserves in developing seeds to act as a reserve for nitrogen, sulfur, and carbon (*26, 27*). Structurally, the 11S proteins are hexameric proteins of MW ~360 kDa, with each subunit comprising an acidic 30−40 kDa polypeptide and

* Address correspondence to these authors. T.-J.F.: phone 708 728 4149, fax 708 728 4177, e-mail tong.fu@fda.hhs.gov. Y.Z.: phone 312 567 3484, fax 312 567 3494, e-mail yuzhu.zhang@iit.edu.

[†] Illinois Institute of Technology, National Center for Food Safety and Technology.

[‡] Department of Biology, Illinois Institute of Technology.

[§] U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition.

[∥] U.S. Food and Drug Administration, National Center for Food Safety and Technology.

10.1021/jf800529k CCC: $40.75 © XXXX American Chemical Society

Published on Web 06/14/2008

a 20 kDa basic polypeptide (*24*). This family of proteins accounts for a number of major food allergens including Ara h 3 in peanut (*28*), Cor a 9 in hazelnut (*29*), Jug r 4 in walnut (*30*), Ana o 2 in cashew nut (*31*), Sin a 2 in mustard seeds (*32*), the globulin in sesame seed (*33*), and glycinin in soybean (*34*).

Most immediate-type food-induced hypersensitivity reactions are believed to be IgE-mediated food allergies (*3, 35*). Structural characterization of allergens is essential for the understanding of the molecular basis of protein allergenicity. Linear IgE-binding epitopes for a number of food allergens have been determined. Using homology-based molecular modeling approaches, the analysis of these epitopes and their comparison to IgE-reactive epitopes in homologous allergens from different foods are being undertaken to characterize structural features important to protein allergenicity (*25, 36, 37*). The homology-based molecular modeling approach has also been applied to characterize the structural basis for IgE-binding cross-reactivity (*38, 39*).

Recently, the structure of Gly m G1, the 11S seed storage protein in soybean, has been determined (*40, 41*). However, there are very few food allergens whose tertiary structures are known and there are no reports of three-dimensional structure of any tree nut allergen at atomic resolution. Additional structural information for the 11S proteins is desired to understand the structural basis and cross-reactivity of this group of proteins. Here, we report our results of the purification, crystallization, and initial structural characterization of Pru1, the major component of amandin.

MATERIALS AND METHODS

Materials. Raw whole Nonpareil Supreme almonds were purchased from VineTree Orchards (Burbank, CA). Sodium chloride, 3-(cyclohexylamino)-1-propane sulfonic acid, sodium azide, β-mercaptoethanol (β-ME), and PBS (phosphate buffered saline pH 7.4) were from Sigma-Aldrich Chemical Co. (St. Louis, MO). Tris (hydroxymethyl)aminomethane was from Bio-Rad Laboratories (Richmond, CA). Ethanol, methanol, hydrochloric acid, sulfuric acid, acetic acid glacial, and Coomassie Brilliant Blue R 250 were reagent or better grade (Fisher Scientific, Pittsburgh, PA). Novex 10−20% Tricine Pre-Cast Mini gels (1.0 mm × 15 well), NuPAGE 3−8% Bis-Tris Pre-Cast Mini gels (1.0 mm × 12 well), and other electrophoresis reagents were from Invitrogen Corporation (Carlsbad, CA). Gradient PhastGels and High Molecular Weight Native Marker Kit were from GE Healthcare (Piscataway, NJ).

Preparation of Defatted Almond Flours. Whole Nonpareil almonds were ground in an analytical mill (Tekmar, Germany). Fat was extracted from the flour in aliquots with hexane in 50 mL centrifugal tubes (flour/solvent ratio was 1:10 w/v). The tubes were kept in constant agitation at 250 rpm in an orbital shaker at room temperature for 3 h. The organic solvent was decanted and the extraction process was repeated twice. After the last extraction, residual hexane in the slurry was evaporated overnight in a fume hood at room temperature. The defatted almond powder was homogenized in a mortar with a pestle and stored at −20 °C in airtight plastic bottles for further use.

Protein Purification. Amandin was first purified by cryoprecipitation. Defatted almond flour was extracted at room temperature with MilliQ water containing 0.02% NaN_3 (flour/H_2O ratio 1:50) with constant agitation in an orbital shaker and vortexing at 10-min intervals. After extracting for 1 h, the mixture was centrifuged (6000*g* for 10 min) and the pellet was re-extracted a second time. The (turbid yellow) supernatants were pooled and filtered through cellulose filters of 0.45 μm pore size (Millipore Corp., Bedford, MA). The filtrate was kept at 4 °C for 12−14 h and the milky, sticky precipitate was collected after centrifugation at 4 °C for 20 min (12000*g*). The precipitate, containing mostly amandin, was then dispersed in 0.01 M phosphate buffered saline (pH 7.4) and dialyzed against distilled water. The dialysis was performed at room temperature with constant stirring for 9 h with reservoir water changed every 3 h. Regenerated cellulose tubular membrane with a 12000−14000 molecular weight cutoff (Fischer Scientific) was used. The cryoprecipitated amandin preparation, with a concentration of ∼40 mg/mL, was stored at −20 °C in airtight plastic bottles.

The amandin preparation was further purified by anion exchange, hydrophobic interaction, and size exclusion chromatography. For anion exchange chromatography, cyroprecipitated amandin was diluted to 1 mg/mL with a 10 mM Tris-HCl buffer, pH 7.9, and loaded onto an 8 mL Source 15Q (GE Healthcare) anion exchange column pre-equilibrated with the Tris-HCl buffer. The column was then washed with the same buffer and eluted with the Tris-HCl buffer plus a linear gradient of NaCl (0−0.3 M).

Fractions containing Pru1 eluted from the anion exchange column were pooled, and ammonium sulfate powder was added to a final concentration of 1.2 M. The protein was then loaded onto two tandemly connected 5 mL phenyl Sepharose columns (GE Healthcare) pre-equilibrated with 10 mM Tris-HCl, pH 7.9, containing 1.2 M ammonium sulfate. The hydrophobic interaction columns were eluted with a 1.2−0 M ammonium sulfate gradient.

Fractions of the eluted Pru1 peak from the hydrophobic interaction chromatography were pooled and further purified with a 300 mL Superdex 200 column (XK 26/70, GE Healthcare) pre-equilibrated with the Tris-HCl buffer (10 mM, pH 7.9) and eluted with the same buffer containing 100 mM NaCl. The Superdex column was calibrated by using thyroglobulin (669 kDa), ferritin (440 kDa), catalase (242 kDa), aldolase (158 kDa), and bovine serum albumin (67 kDa) (GE Healthcare) as molecular weight standards.

Purified Pru1 was pooled and concentrated in a 15 mL Amicon ultracentrifugation filter device (MWCO 30000; Millipore) and the protein concentration was estimated with a Cary 300 UV−visible spectrophotometer by using a theoretical extinction coefficient of $\varepsilon_{280} = 33\,170$ cm^{-1} mol^{-1} calculated according to the Pru1 sequence from the NCBI database. The protein concentration was also determined in triplicate with the BCA (bicinchoninic acid) assay (*42*), using a BCA kit from Pierce Biotechnology, Inc. (Rockford, IL). Bovine serum albumin (BSA) was used as the protein standard. Concentrated Pru1 was dialyzed four times against distilled water over a 24 h period and stored at −20 °C.

Electrophoresis. Electrophoresis with nondenaturing, nondissociating polyacrylamide gel (NDND-PAGE) was performed according to an established protocol (*43*), using NuPAGE 3−8% Tris Acetate Pre-Cast Mini gels. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was carried out according to the method of Fling and Gregerson (*44*), using 10−20% or 4−20% Tricine gels. For reduced samples, 2× Laemmli buffer was added to the samples (1:1 v/v) before they were boiled for 10 min. Running conditions for each type of gel were those suggested by the manufacturer (Invitrogen). The gels were stained with Coomassie Brilliant Blue R 250 0.1% (w/v) in a solution with 50% distilled H_2O, 40% methanol, and 10% acetic acid for 2 h and destained for 12 h. Destaining was carried out in the same solution without the dye. The Precision Plus Protein unstained standards (Bio-Rad) were used for SDS-PAGE and the High Molecular Weight markers (GE Healthcare) were used for NDND-PAGE. Origin Pro 7.0 SR0 software (OriginLab Corp., Northampton, MA) was used to calculate molecular weight values of the polypeptides of interest compared to the standards by exponential curve fitting. UN-SCAN-IT *gel* (Silk Scientific Inc., Orem, UT) was used to quantify percentages of protein in the gels.

N-Terminal Amino Acid Sequencing. Purified Pru1 was subjected to SDS-PAGE with 8% to 25% gradient gels in the PhastSystem (GE Healthcare) and electrophoretically transferred (0.35 A at 25 °C) to a Problott membrane (Applied Biosystems, Foster City, CA) in a transfer buffer [10 mM 3-(cyclohexylamino)-1-propane sulfonic acid containing 10% methanol (pH 11.0)]. The blot was stained with Coomassie Brilliant Blue and the three main peptide bands were excised and subjected to N-terminal amino acid sequencing by Edman degradation, using a Procise model 491 Protein Sequencer (Applied Biosystems). Cysteine residue is chemically labile during Edman degradation and no attempts were made to derivatize it to its PE-Cys form so that it

could be analyzed. A search for homology to the obtained sequences was performed with the Swiss-Prot database at the Swiss Institute of Bioinformatics Website, using BLAST searches.

Crystallization. Purified Pru1 at a concentration of ~40 mg/mL was thawed on ice and subject to centrifugation at 12000*g* for 5 min. The supernatant was concentrated to ~100 mg/mL with an Amicon Ultracentrifugal filter device (Millipore Corp.). The sample was filtered with a 0.2 μm syringe filter (Millipore Corp.) prior to crystallization trial, using the Crystal Screen and the Crystal Screen 2 kits from Hampton Research (Aliso Viejo, CA). The crystallization screen was set up by mixing 1 μL of protein sample with 1 μL of reservoir solution and sealed against 0.5 mL of reservoir solution, using the hanging drop method.

X-ray Diffraction Experiments and Data Processing. Single crystals were briefly soaked in a cryoprotectant solution containing 20% (w/v) sucrose and 2 M ammonium sulfate. The crystals were then picked up with mounted CryoLoops and flash frozen by immersing the crystals in liquid nitrogen. Diffraction data were collected at Ser-CAT beamline 22 BM (Advanced Photon Source, U.S. Argonne National Laboratory) equipped with a Mar-225 CCD detector. A complete data set of 1° per frame with 6 s exposure was collected at 1 Å wavelength. Data processing was carried out with the HKL2000 suite of programs (*45*).

Construction of a Pru1 Structural Model and Phase Solution by Molecular Replacement. A model structure of Pru1 was constructed by homology modeling for solving the phase problem by molecular replacement. Modeling was carried out with the program SCWRL (*46*), using protomer A of the structure Gly m G1 (Protein Data Bank entry 1OD5) from a mutant soybean cultivar containing glycinin composed of only the A3B4 subunit (*40*). Molecular replacement was performed with the program PHASER (*47, 48*).

RESULTS AND DISCUSSION

Prunin-1 Purification and Polypeptide Characterization. Prunin-1 shares the properties of being highly water soluble and readily cold-precipitable with other globular proteins of the 11S family, such as legumin of peas (*49*), arachin of peanuts (*50*), glycinin of soybeans (*51*), and the major globulins of linseed (*52*) and coconut (*53*). In this work, amandin was isolated from defatted almond flour by water extraction and cryoprecipitation and Pru1 was further purified by anion exchange, hydrophobic interaction, and size exclusion chromatography. The purities of the amandin preparations were assessed by nondenaturing, nondissociating polyacrylamide gel electrophoresis (NDND-PAGE). As shown in **Figure 1**, amandin accounted for 74.6% of the total soluble protein in almond. The protein with MW ~500 kDa, previously referred as the 19S component and considered a dimer of the amandin complex (*21*), accounted for another 11.8% of the total protein. Several other proteins can also be detected in the extract, but were not determined in the present study. After cryoprecipitation, only two bands were visible by Coomassie Brilliant Blue staining (**Figure 1**, lanes 2, 4, and 5). Amandin accounted for ~96% of the cryoprecipitated protein preparation and the rest was the 19S protein. These values were comparable to those reported earlier for cryoprecipitated protein and characterized by ultracentifugal analyses (*21*).

On the basis of the distance migrated in NDND-PAGE gels by amandin and the molecular weight markers and with an exponential curve fitting with Origin Pro 7.0, the molecular weight (MW) of amandin was estimated to be 370.0 ± 2.3 kDa ($R^2 = 0.987$). In a recent report, the MW of amandin, obtained by chromatography separations and determined by gel filtration by using the Stokes radius and elution volume, was described to be 427.3 ± 47.6 kDa, but when obtained by means of cryoprecipitation without any further chromatography step, the MW varied broadly and was determined

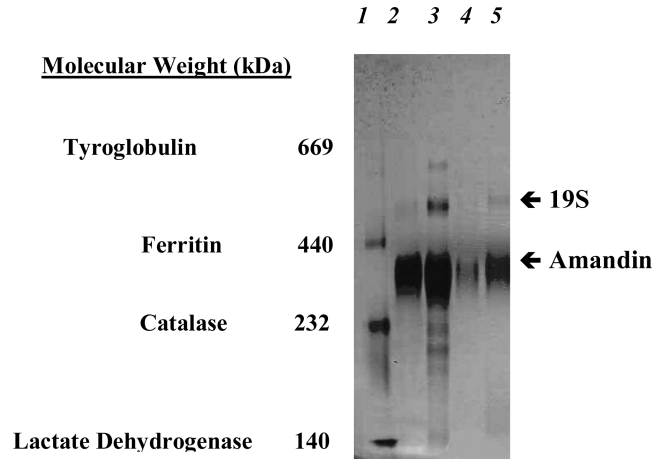

Figure 1. NDND-PAGE of amandin and total soluble protein extract from almond. Ten microliters of each of the samples were loaded in a well of a Novex Tris-Acetate gel and the gel was run at 150 V constant voltage for 2.5 h. Lane 1, Molecular Marker; lane 2, cryoprecipitated amandin (~20 mg/mL); lane 3, total soluble proteins in almond extract (~20 mg/mL); lanes 4 and 5, cryoprecipitated amandin (~0.5 and ~1 mg/mL, respectively).

to be of 368 ± 100 kDa (*22*). It was suggested that the apparent MW determined by gel filtration should be considered with caution because it is based on the hydrodynamic properties of the protein and may be accurate only for globulins of spherical shape (*54*).

To obtain a highly pure protein for crystallization studies, the cryoprecipated amandin was subjected to three different modes of chromatographic separations (**Figure 2**). The purity of the protein fractions collected from each of the purification steps was assessed by SDS-PAGE. As shown in **Figure 3**, anion exchange chromatography effectively removed most of the other protein impurities from the Pru1 fraction. The additional hydrophobic interaction and size exclusion chromatography did not result in further improvement of the purity of the Pru1 sample. The banding pattern of the redissolved crystal was similar to those of the protein preparations resulting from the hydrophobic interaction and size exclusion chromatography suggesting that the protein preparation used for crystallization was highly pure.

SDS-PAGE analysis of the purified protein revealed the presence of three main bands with MW of 22, 40, and 42 kDa (**Figure 3**). The N-terminal amino acid sequences in these three polypeptides were determined by protein sequencing (**Figure 4**). The first 30 amino acids in the 40 and 42 kDa bands are identical. When compared with the published sequences, high similarity was observed to the sequence of aa21−aa50 of the Pru1 with a sequence identity of 96%. In addition, the first 26 amino acids of the 22 kDa band can be matched to aa368−aa393 of Pru1 with a sequence identity of 96%. The identity of both peptides could be 100% if the cysteine residues had been derivatized and analyzed. Thus, in this almond cultivar, amandin could be considered as essentially the result of the *pru1* gene. The 40 kDa and 42 kDa bands correspond to the acidic (theoretical p*I* value of 5.66) subunits of Pru1 and the 22 kDa band corresponds to the basic (theoretical p*I* value of 9.25) subunit of Pru1. Amandin has been identified as a major allergen in almond (*15*). Pru1 is therefore very likely to be an allergen, but future work is needed to test this with serum samples from individuals allergic to almond.

It is known that each isoform of 11S globulin is generally encoded by a single gene (*31*), producing a precursor that is

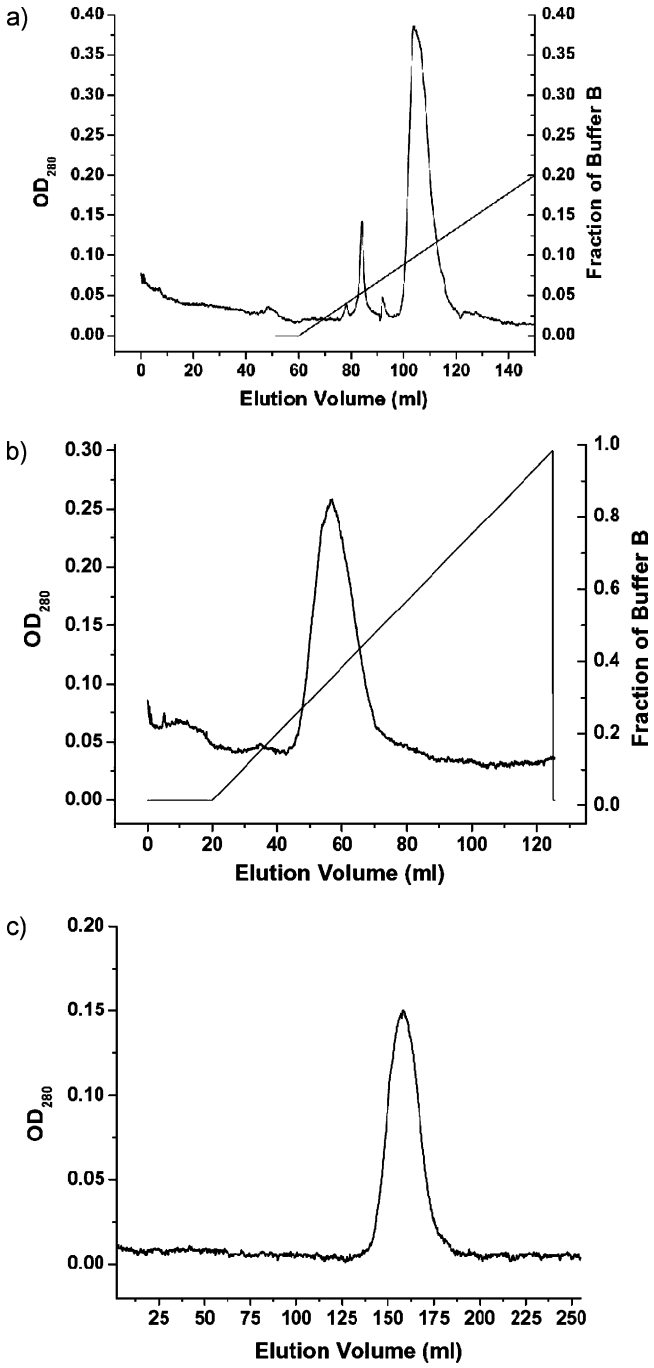

Figure 2. Chromatographic purification of Pru1. **(A)** Anion exchange chromatography. Cryoprecipitated amandin was loaded onto a 8 mL Source 15Q anion exchange column pre-equilibrated with the 10 mM Tris buffer. The column was washed with the Tris buffer and linear NaCl gradient (0−0.3 M over 90 mL) was used to elute the column. The major peak was Pru1. **(B)** Hydrophobic interaction chromatographic purification of Pru1. A Pru1 sample in 1.2 M ammonium sulfate was loaded onto a 10 mL phenyl-sepharose column. After washing with the binding buffer (10 mM Tris-HCl, pH 7.9, 1.2 M ammonium sulfate), the column was eluted with an 105 mL of linear ammonium sulfate gradient of 1.2−0 M. **(C)** Gel filtration purification of Pru1. Twenty milliliters of the Pru1 containing peak from the hydrophobic interaction purification was loaded onto a 300 mL Superdex-200 gel filtration column and eluted by a Tris buffer (10 mM, pH 7.9) containing 100 mM NaCl.

post-translationally cleaved by an asparaginyl endopeptidase after the formation of an interchain disulfide bond between the N-terminal acidic and the C-terminal basic subunits (*55, 56*).

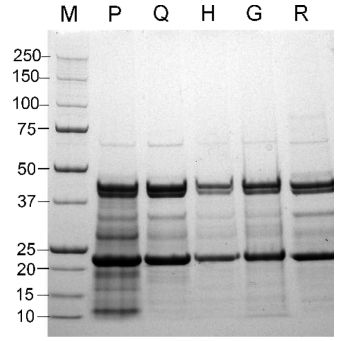

Figure 3. SDS-PAGE analysis of Pru1 samples. Ten microliters of each sample were loaded into a well of a 4−20% SDS gel. The molecular mass of protein bands in the Precision Plus Protein Standards (lane M) is shown to the left of the gel image. Lane P was loaded with the cryoprecipitated amandin; lane Q, Pru1 after anion exchange; lane H, Pru1 after hydrophobic interaction purification; and lane G, Pru1 after size exclusion purification. Lane R was loaded with the Pru1 sample redissolved from Pru1 single crystals. All samples were boiled for 10 min in a SDS sample buffer (50 mM Tris-HCl, pH 6.8, 2% (v/v) SDS, 0.1% (w/v) bromophenol blue, 10% (v/v) glycerol, 100 mM β-mercaptoethanol).

The cleavage site between the Asn-Gly peptide bond is well conserved among a wide variety of plant species (*57*). 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. As shown in **Figure 4**, residues 367−372 of Pru1 are NGLEET and the GLEET sequence is the start of the C-terminal basic subunits, consistent with this motif being an endopeptidase recognition site for post-translational cleavage.

Crystallization and Preliminary X-ray Crystallographic Data. In the crystallization screen, several needle-like crystals appeared at room temperature in the drop hanging over solution No. 4 of the Crystal Screen kit (2 M ammonium sulfate) in 3 days. After optimization of the precipitant concentration, pH, and protein concentration, the starting protein concentration of 60 mg/mL was used with the original no. 4 solution of the Crystal Screen kit as reservoir solution. Under these conditions, long needle-like single crystals of cross-section dimensions of $50 \times 100\ \mu m^2$ were obtained (**Figure 5**). Most of the crystals appeared to have defects and attempts to get good-looking crystals were unsuccessful. However, usable diffraction data could be obtained from these crystals.

The needle-like Pru1 crystals diffracted to 2.65 Å and diffraction decay was obvious during data collection. They typically diffracted to 3.0 Å at the end of a data set of 120 frames with 1-s exposure. A typical diffraction frame is shown in **Figure 6**. The diffraction data were processed with the HKL2000 suite of programs (*45*), revealing a tetragonal crystal system with unit cell parameters $a = b = 150.913$ Å and $c = 165.242$ Å. On the basis of systematic absences in specific reflections in the diffraction, the space group was determined to be $P4_122$ or $P4_322$ (**Table 1**). Assuming three monomers of Pru1 in an asymmetric unit, the Matthews coefficient was 2.53 $Å^3$/Da, corresponding to a solvent content of ~51.41%.

Pru1 has 42% sequence identity with soybean glycinin. Using protomer A of the structure of glycinin (PDB:1OD5) derived from a mutant soybean cultivar containing glycinin composed of only the A3B4 subunit (*40*) as template, we constructed a

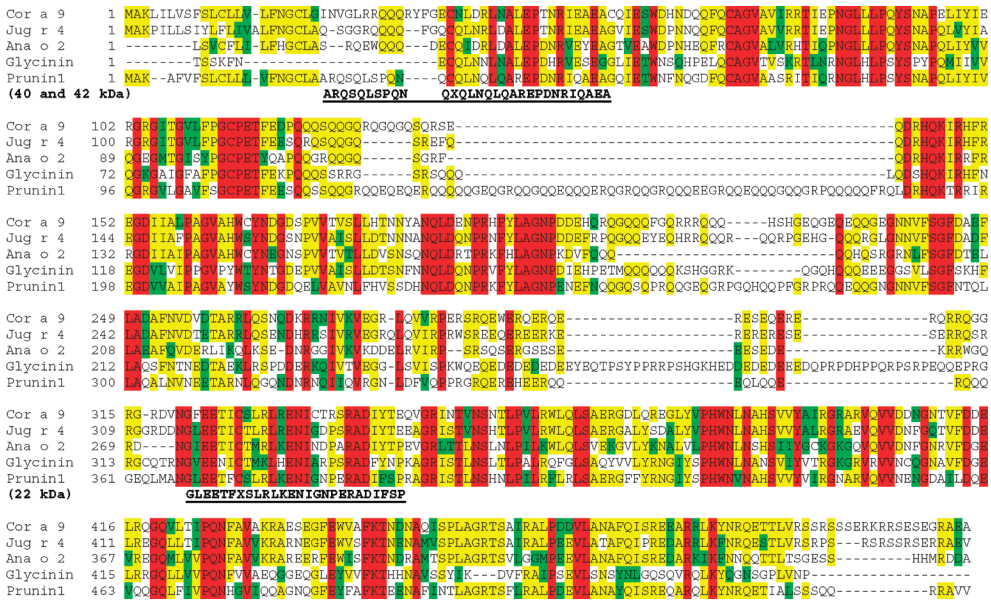

Figure 4. N-Terminal amino acid sequencing identification of Pru1 and sequence alignment of Pru 1 with other known 11S allergens. ClustalW multiple sequence alignment of glycinin, Pru1, and nut allergens Cor a 9, Jug r 4, and Ana o 2 is shown with conserved residues highlighted in red, identical residues in yellow, and similar residues in green. N-Terminal sequencing results for the peptides obtained from denatured Pru1 are shown in underlined boldface below the sequence alignment. Thirty cycles of sequencing indicate that the 40 and 42 kDa bands have identical N-terminus starting from A_{21} of Pru1 (the first underlined boldface sequence), and 26 cycles of sequencing indicate that the 22 kDa band starts from G_{368} of Pru1 (the second underlined bold face sequence).

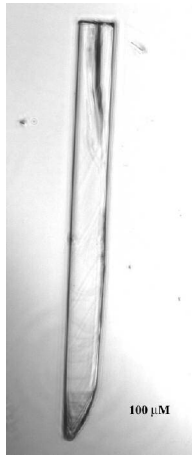

Figure 5. Crystal of Pru1 obtained by vapor diffusion using the hanging drop method.

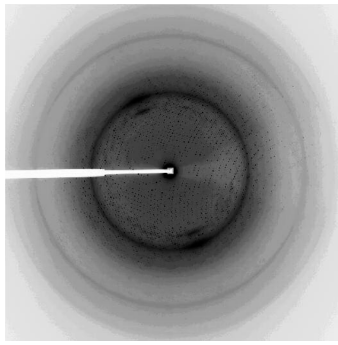

Figure 6. Typical frame of the Pru1 X-ray diffraction data.

Pru1 structural model using the program SCWRL (*46*) (data not shown). The monomeric Pru1 model was used in a molecular replacement calculation, using the program PHASER (*47, 48*). This resulted in a molecular replacement solution with three Pru1 molecules in the asymmetric unit and a log of likelihood of 2453. The *R*-factor of the solution after a round of rigid body and restrained refinement with the program REFMAC (*58*) was 33%.

Table 1. X-ray Data-Collection Statistics[a]

parameter	value
resolution (Å)	50−3.0 (3.11−3.0)
wavelength (Å)	1.0
space group	$P4_122$ or $P4_322$
unit-cell parameters (Å)	
a	150.912
b	150.912
c	165.248
data-collection temp (K)	110
no. of observed reflcns	618080
no. of unique reflcns	38756
completeness (%) (outer shell)	99.9 (100)
$I/\sigma(I)$ (outer shell)	19.62 (4.28)
R_{merge}[b] (%) (outer shell)	14.5 (53.27)

[a] Values in parentheses are for the outer shell. [b] $R_{merge} = \sum_{hkl}\sum_i|I_{hkl,i} - \langle I_{hkl}\rangle|/\sum_{hkl}\sum_i I_{hkl,i}$.

A preliminary inspection of the map calculation and refinement led to an experimental electron-density map at 3.0 Å resolution with a clear protein−solvent boundary. The resulting electron-density map allowed modeling of most of the main-chain atoms for all three monomers (**Figure 7**). The three protomers are related by a noncrystallographic 3-fold axis to form a doughnut-shaped trimer. Nevertheless, implementation of this initial noncrystallographic symmetry relationship in the refinement did not lead to a further improvement of the electron density. Two Pru1 trimers form a homohexamer similar to the glycinin hexamer (*40*). Additional model building and refinement of the structure is currently underway.

Structural determination of Pru1 will provide needed information for comparing the similarities and differences among the 11S protein allergens from different food sources. The 11S seed storage protein from almonds and other plant allergens resulted in multiple bands when analyzed on denaturing SDS

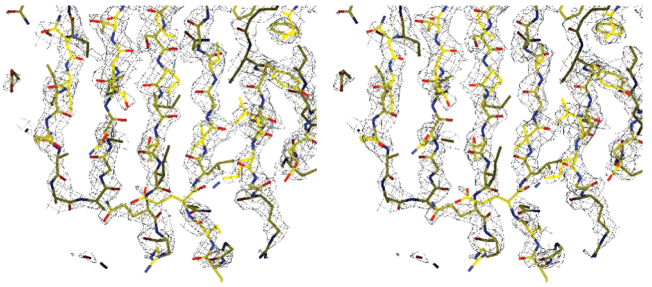

Figure 7. Stereoview of a portion of a $2F_o - F_c$ electron-density map. The map was calculated after one round of rigid body and restrained refinement starting from a PHASER molecular replacement solution. The map was contoured at 1.5σ and the structure model is shown in stick representation.

gels (*28–31*) suggesting the presence of different isoforms. Previous structures of the 11S storage proteins were based on soybean seeds that are either from recombinant source (*41*) or from genetically modified plant (*40*). Structural determination of Pru1 isolated from natural source will provide additional information about the constituents of 11S seed storage proteins which may lead to a better understanding of the allergenicity of this group of proteins.

Elucidation of the three-dimensional structures of the 11S proteins at high resolution can also facilitate the characterization of IgE binding epitopes on the molecular surface of allergenic proteins to characterize the cross-reactivity of this group of allergens, especially those in peanut and tree nuts (*59, 60*). Although the IgE binding epitopes of amandin (or prunin) have not been identified, the IgE binding epitopes for the 11S allergens in peanuts and a number of tree nuts are available (*60*). Comparative study of the structural features of these proteins is difficult currently due to the lack of three-dimensional structures of other 11S proteins of tree nuts and peanuts. Recently, the preliminary crystallographic characterizations of the 11S protein in peanut (Ara h 3) and Brazil nut (Ber e 2) have been attempted (*61, 62*). It is expected that with the availability of an increasing number of three-dimensional structures, progress will be made toward the understanding of the molecular basis of protein allergenicity and allergen cross-reactivity.

LITERATURE CITED

(1) Sicherer, S. H.; Sampson, H. A. 9. Food allergy. *J. Allergy Clin. Immunol.* **2006**, *117*, S470–475.
(2) Moneret-Vautrin, D. A.; Morisset, M. Adult food allergy. *Curr. Allergy Asthma Rep.* **2005**, *5*, 80–85.
(3) Taylor, S. L.; Hefle, S. L. Food allergies and other food sensitivities. *Food Technol.* **2001**, *55*, 68–83.
(4) Lehrer, S. B.; Ayuso, R.; Reese, G. Current understanding of food allergens. *Ann. N.Y. Acad. Sci.* **2002**, *964*, 69–85.
(5) Rosengarten, F. *The Book of Edible Nuts;* Dover Publications: Mineola, NY, 2004; p 384.
(6) 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.
(7) Schreiber, R. A.; Walker, W. A. Food allergy: facts and fiction. *Mayo Clin. Proc.* **1989**, *64*, 1381–1391.
(8) Sicherer, S. H. Clinical update on peanut allergy. *Ann. Allergy Asthma Immunol.* **2002**, *88*, 350–361.
(9) Hourihane, J. O. Prevalence and severity of food allergy-need for control. *Allergy* **1998**, *53* (46 Suppl.), 84–88.
(10) Sampson, H. A. Food allergy. Part 1: immunopathogenesis and clinical disorders. *J. Allergy Clin. Immunol.* **1999**, *103*, 717–728.
(11) Angus, F. Nut allergens. In *Natural toxicants in foods*, Watson, D. H., Ed.; CRC Press LLC: Boca Raton, FL, 1998; pp 84−104.
(12) Teuber, S. S.; Sathe, S. K.; Peterson, W. R.; Roux, K. H. Characterization of the soluble allergenic proteins of cashew nut (*Anacardium occidentale* L.). *J. Agric. Food Chem.* **2002**, *50*, 6543–6549.
(13) Osborne, T. B.; Campbell, G. F. Conglutin and vitellin. *J. Am. Chem. Soc.* **1896**, *18*, 609–623.
(14) Youle, R. J.; Huang, A. H. C. Occurrence of low molecular weight and high cysteine containing albumin storage proteins in oilseeds of diverse species. *Am. J. Bot.* **1981**, *68*, 44–48.
(15) 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.
(16) 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.
(17) 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.
(18) 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.
(19) Sathe, S. K. Solubilization, electrophoretic characterization and in vitro digestibility of almond (*Prunus amygdalus*) proteins. *J. Food Biochem.* **1993**, *16*, 249–264.
(20) Derbyshire, E.; Wright, D. J.; Boulter, D. Legumin and vicilin, storage proteins of legume seeds. *Phytochemistry* **1976**, *15*, 3–24.
(21) Wolf, W. J.; Sathe, S. K. Ultracentrifugal and polyacrylamide gel electrophoretic studies of extractability and stability of almond meal proteins. *J. Sci. Food Agric.* **1998**, *78*, 511–521.
(22) Sathe, S. K.; Wolf, W. J.; Roux, K. H.; Teuber, S. S.; Venkatachalam, M.; Sze-Tao, K. W. Biochemical characterization of amandin, the major storage protein in almond (*Prunus dulcis* L.). *J. Agric. Food Chem.* **2002**, *50*, 4333–4341.
(23) Svedberg, T.; Sjogren, B. The molecular weights of amandin and excelsin. *J. Am. Chem. Soc.* **1930**, *52*, 279–287.
(24) Mills, E. N.; Jenkins, J.; Marigheto, N.; Belton, P. S.; Gunning, A. P.; Morris, V. J. Allergen of the cupin superfamily. *Biochem. Soc. Trans.* **2002**, *30*, 925–929.
(25) Breiteneder, H.; Radauer, C. A classification of plant food allergens. *J. Allergy Clin. Immunol.* **2004**, *113*, 821–830.
(26) Utsumi, S. Plant food protein engineering. *Adv. Food Nutr. Res.* **1992**, *36*, 89–208.
(27) Shewry, P. R. Plant storage proteins. *Biol. Rev. Cambridge Philos. Soc.* **1995**, *70*, 375–426.
(28) 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.
(29) 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.
(30) 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.
(31) Wang, F.; Robotham, J. M.; Teuber, S. S.; Tawde, P.; Sathe, S. K.; Roux, K. H. Ana o 1, a cashew (*Anacardium occidentale*) allergen of the vicilin seed storage protein family. *J. Allergy Clin. Immunol.* **2002**, *110*, 160–166.
(32) Palomares, O.; Cuesta-Herranz, J.; Vereda, A.; Sirvent, S.; Villalba, M.; Rodríguez, R. Isolation and identification of an 11S globulin as a new major allergen in mustard seeds. *Ann. Allergy Asthma Immulog.* **2005**, *94*, 586–592.
(33) Hsiao, E. S. L.; Lin, L.-J.; Li, F.-Y.; Wang, M. M. C.; Liao, M-Y; Tzen, J. T. C. Gene families encoding isoforms of two major sesame seed storage proteins, 11S globulin and 2S albumin. *J. Agri. Food Sci.* **2006**, *54*, 9544–9550.

(34) Helm, R. M.; Cockrell, G.; Connaughton, C.; Sampson, H. A.; Bannon, G. A.; Beilinson, V.; Livingstone, D.; Nielsen, N. C.; Burks, A. W. A soybean G2 glycinin allergen. 1. Identification and characterization. *Int. Arch. Allergy Immunol.* **2000**, *123*, 205–212.

(35) Ebo, D. G.; Stevens, W. J. IgE-mediated food allergy-extensive review of the literature. *Acta Clin. Belg.* **2001**, *56*, 234–247.

(36) Borges, J. P.; Barre, A.; Culerrier, R.; Archimbaud, N.; Didier, A.; Rougé, P. How reliable is the structural prediction of IgE-binding epitopes of allergens? The case study of plant lipid transfer proteins. *Biochimie* **2007**, *89*, 83–91.

(37) Lehmann, K.; Schweimer, K.; Reese, G.; Randow, S.; Suhr, M.; Becker, W. M.; Vieths, S.; Rösch, P. Structure and stability of 2S albumin type peanut allergens: implications for the severity of peanut allergic reactions. *Biochem. J.* **2006**, *395*, 463–472.

(38) de Leon, M. P.; Drew, A. C.; Glaspole, I. N.; Suphioglu, C.; O'Hehir, R. E.; Rolland, J. M. IgE cross-reactivity between the major peanut allergen Ara h 2 and tree nut allergens. *Mol. Immunol.* **2007**, *44*, 463–471.

(39) Barre, A.; Borges, J. P.; Rouge, P. Molecular modeling of the major peanut allergen Ara h 1 and other homotrimeric allergens of the cupin superfamily: a structural basis of their IgE-binding cross-reactivity. *Biochimie* **2005**, *87*, 499–506.

(40) 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.

(41) Adachi, M.; Takenaka, Y.; Gidamis, A. B.; Mikami, B.; Utsumi, S. Crystal structure of soybean proglycinin A1aB1b homotrimer. *J. Mol. Biol.* **2001**, *305*, 291–305.

(42) Smith, P. K.; Krohn, R. I.; Hermanson, G. T.; Mallia, A. K.; Gartner, F. H.; Provenzano, M. D.; Fujimoto, E. K.; Goeke, N. M.; Olson, B. J.; Klenk, D. C. Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **1985**, *150*, 76–85.

(43) Andrews, A. T. *Electrophoresis: Theory, Techniques, and Biochemical and Clinical Applications*, 2nd ed.; Clarendon Press: Oxford, U.K., 1986.

(44) Fling, S. P.; Gregerson, D. S. Peptide and protein molecular weight determination by electrophoresis using a high-molarity tris buffer system without urea. *Anal. Biochem.* **1986**, *155*, 83–88.

(45) Otwinowski, Z.; Minor, W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. In *Methods in Enzymology*; Carter, C. W., Sweet, R. M., Eds.; Academic Press: New York, 1997; Vol. 276, pp 307−326.

(46) Canutescu, A. A.; Shelenkov, A. A.; Dunbrack, R. L., Jr. A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci.* **2003**, *12*, 2001–2014.

(47) McCoy, A. J.; Grosse-Kunstleve, R. W.; Storoni, L. C.; Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2005**, *61*, 458–564.

(48) Storoni, L. C.; McCoy, A. J.; Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 432–438.

(49) Danielson, C. E. Seed globulins of the Gramineae and Leguminosae. 1. Seed globulins of the most common species of the Gramineae and their differentiation in the seed. *Biochem. J.* **1949**, *44*, 387–400.

(50) Neucere, N. J. Isolation of α-arachin, the major peanut globulin. *Anal. Biochem.* **1969**, *27*, 15–24.

(51) Wolf, W. J.; Sly, D. A. Cryoprecipitation of soybean 11S protein. *Cereal Chem.* **1967**, *44*, 653–668.

(52) Dev, D. V.; Sienkiewicz, T. Isolation and subunit composition of 11S globulin in linseed (*Linum usitatissimun* L.). *Nahrung* **1987**, *31*, 767–769.

(53) Rasyid, F.; Manullang, M.; Hansen, P. M. T. Isolation and characterization of coconut protein. *Food Hydrocolloids* **1992**, *6*, 301–314.

(54) Siegel, L. M.; Monty, K. J. Determination of molecular weights and frictional ratios of proteins in impure systems by use of gel filtration and density gradient centrifugation. Application to crude preparations of sulfite and hydroxylamine reductases. *Biochim. Biophys. Acta* **1966**, *112*, 346–362.

(55) Kitamura, K.; Takagi, T.; Shibasaki, K. Subunit structure of soybean 11S globulin. *Agric. Biol. Chem.* **1976**, *40*, 1837–1844.

(56) Badley, R. A.; Atkinson, D.; Hauser, H.; Oldani, D.; Green, J. P.; Stubb, J. M. The structure, physical and chemical properties of soy bean protein glycinin. *Biochim. Biophys. Acta* **1975**, *412*, 214–228.

(57) Dickinson, C. D.; Hussein, E. H.; Nielsen, N. C. Role of posttranslational cleavage in glycinin assembly. *Plant Cell* **1989**, *1*, 459–469.

(58) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol Crystallogr.* **1997**, *53*, 240–55.

(59) Wallowitz, M. L.; Teuber, S.; Beyer, K.; Sampson, H. A.; Roux, K. H.; Sathe, S. K.; Wang, F.; Robotham, J. Cross-reactivity of walnut, cashew, and hazelnut legumin proteins in tree nut allergic patients. *J. Allergy Clin. Immunol.* **2004**, *113*, S156.

(60) Barre, A.; Jacquet, G.; Sordet, C.; Culerrier, R.; Rougé. Homology modeling and conformational analysis of IgE-binding epitopes of Ara h 3 and other legumin allergens with a cupin fold from tree nuts. *Mol. Immunol.* **2007**, *44*, 3243–3255.

(61) 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*, 845–851.

(62) 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.

Received for review February 20, 2008. Accepted April 15, 2008. This work was supported by Cooperative Agreement FD-0004331 between the U.S. Food and Drug Administration and the National Center for Food Safety and Technology and a fund from Illinois Institute of Technology. X-ray diffraction data were collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-BM beamline at the Advanced Photon Source, U.S. Argonne National Laboratory. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract W-31-109-Eng-38. S.M.A. was supported by Fundación Castroverde, Madrid, Spain.

JF800529K