crystallization communications

Purification, identification and preliminary crystallographic studies of Pru du amandin, an allergenic protein from *Prunus dulcis*

Vineet Gaur, Dhruv K. Sethi and Dinakar M. Salunke*

National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110 067, India

Correspondence e-mail: dinakar@nii.res.in

Food allergies appear to be one of the foremost causes of hypersensitivity reactions. Nut allergies account for most food allergies and are often permanent. The 360 kDa hexameric protein Pru du amandin, a known allergen, was purified from almonds (*Prunus dulcis*) by ammonium sulfate fractionation and ion-exchange chromatography. The protein was identified by a *BLAST* homology search against the nonredundant sequence database. Pru du amandin belongs to the 11S legumin family of seed storage proteins characterized by the presence of a cupin motif. Crystals were obtained by the hanging-drop vapour-diffusion method. The crystals belong to space group $P4_1$ (or $P4_3$), with unit-cell parameters $a = b = 150.7$, $c = 164.9$ Å.
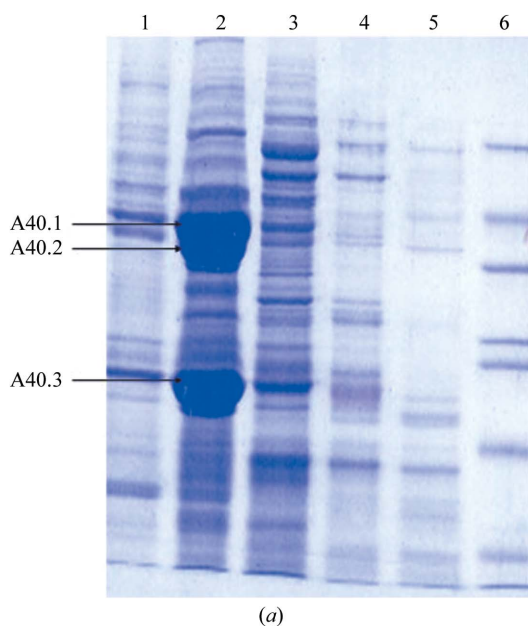
1. Introduction

Allergy is a hypersensitivity reaction initiated by immunological mechanisms, perhaps arising from an interaction between innocuous xenobiotics with a previously primed immune system in susceptible individuals. Allergies resulting from various pharmacological agents have been very well addressed (Demoly, 2006). However, food allergies remain inadequately addressed despite their ubiquitous presence. With increases in hygiene standards, a marked increase in cases of food allergy have been observed, especially in Western countries (Wills-Karp *et al.*, 2001), the prevalence being of the order of 1.4–2.4%, with a higher prevalence in children than in adults (Schafer *et al.*, 2001).

Food allergy is an adverse immune reaction to food that depends on various factors such as diet and culture (Dalal *et al.*, 2002), route of exposure, the genetic makeup of susceptible individuals (Bischoff & Crowe, 2005) and the inherent properties of the allergen itself (Breiteneder & Mills, 2005). Various biochemical and biophysical characteristics, such as thermal stability, pH stability, resistance to proteolysis, ligand interaction *etc.* have been associated with the allergenicity of proteins (Breiteneder & Mills, 2005). However, none of these satisfactorily explain the allergenicity of certain proteins over others. As far as food allergens are concerned, it has been observed that most food allergens belong to only a few protein families. Thus, establishing a structure–function relationship for allergens would not only provide insight into the mechanism of allergies but would also help in the risk assessment for genetically modified proteins. Determination of the three-dimensional structures of homologous food proteins with and without allergenic potential would facilitate the understanding of the structural basis of allergy vis-à-vis their normal physiological role.

Most food allergens of plant origin belong to one of three protein families: the cupin superfamily, prolamins and pathogen-related proteins (Shewry *et al.*, 2002). Legumes and nuts are the leading causes of allergies (Breiteneder & Radauer, 2004). Although almonds belong to the family of Rosaceae fruits, almonds are most often classified as tree nuts. Allergy to almonds has frequently been reported, making this tree nut one of the most common food allergens, with symptoms ranging from mild oral symptoms to life-threatening systemic symptoms (Poltronieri *et al.*, 2002). Three allergens from almond have been identified: 2S albumin, Pru du amandin (almond major protein, AMP or prunin) and conglutin γ.

Pru du amandin belongs to the 11S legumin family, which has a cupin fold. 11S legumin proteins constitute a major group of seed storage proteins that are present in dicots, monocots (including cereals and palms) and fern spores. Like all other legumins, Pru du amandin contains two polypeptide chains linked by a disulfide bond (Sathe *et al.*, 2002) that are produced as a result of post-translational proteolysis. Pru du amandin shows 42% sequence identity to soybean glycinin A3B4 subunit, the X-ray structure of which is known in a homohexameric form (PDB code 1od5).
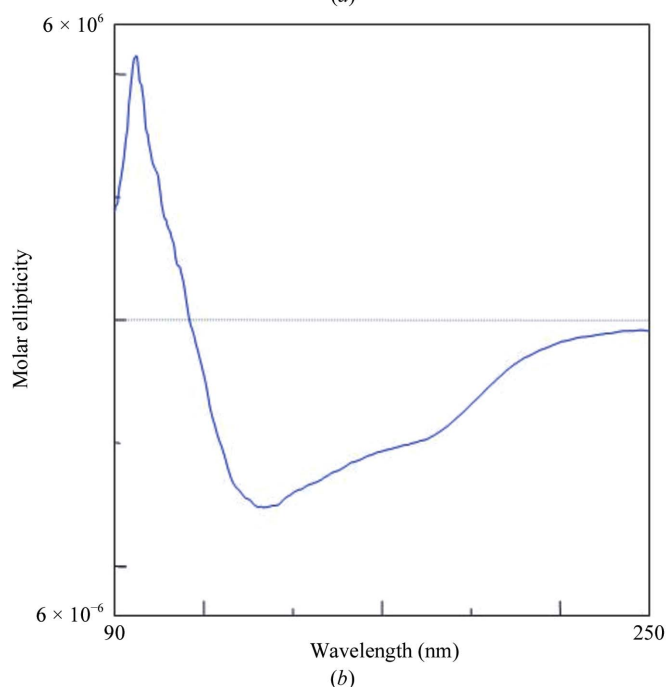
As part of our continuing efforts to identify and characterize proteins with allergenic potential from seeds, our group has previously reported the purification and crystallization of novel proteins from *Vigna unguiculata* (Chanana *et al.*, 2004) and *Lathyrus sativus* (Qureshi *et al.*, 2006). Here, we report the identification, purification and crystallization of Pru du amandin.

2. Experimental methods

2.1. Protein purification

Prunus dulcis nuts were ground and then defatted with petroleum ether. The defatted powder was homogenized in 50 mM Tris–HCl pH 7.5 containing 150 mM NaCl and protease-inhibitor cocktail (Sigma, P-9599) by continuous stirring for 4 h at 277 K. The crude extract obtained was then centrifuged at 16 900g for 60 min at 277 K and subjected to salt fractionation over a concentration range of 20–90%(w/v) ammonium sulfate, separating the proteins on the basis of their differential solubilities. The pellets were resuspended in
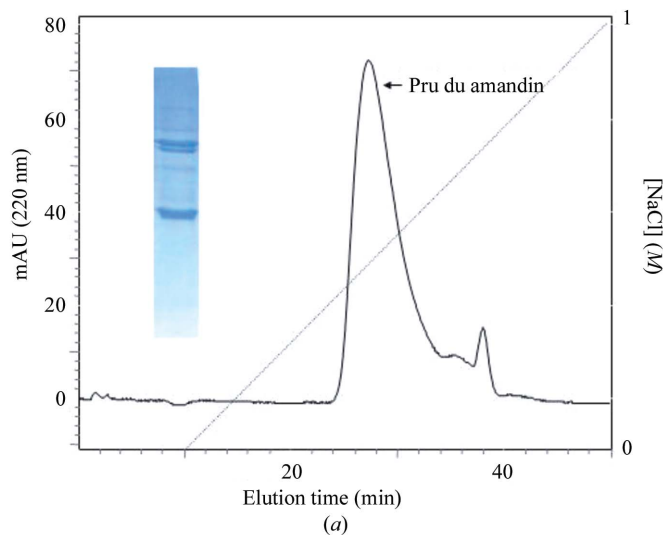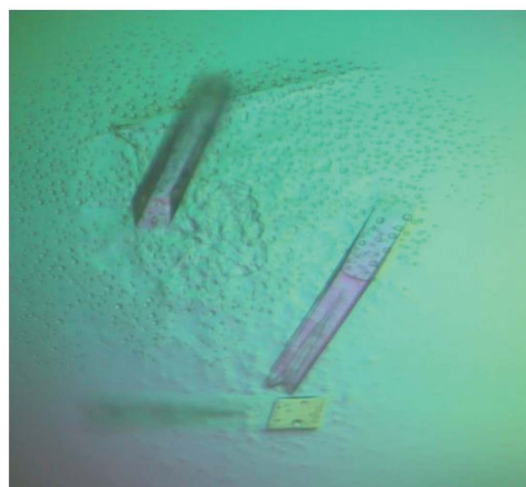

(*a*)

(*b*)

Figure 1
(*a*) Fractionation profile of almond-seed proteins. Lanes 1–6 correspond to 20, 40, 60, 80 and 90% salt fractions and marker proteins, respectively. The N-terminal sequences of the abundantly present proteins in the 40% ammonium sulfate fraction were A40.1, ARQSQLSPQNQXQLN; A40.2, ARQSQLSPQNQXQLN; A40.3, GLEETFXSLRLKENIGNPXR (where an X in the sequence denotes an unidentified residue). (*b*) CD spectrum of Pru du amandin recorded at 283 K and pH 7.5 in the far-UV region.

(*a*)

(*b*)

Figure 2
(*a*) Weak anion-exchange (PI/M Pharmacia) chromatogram; the peak corresponds to Pru du amandin from *P. dulcis*. The dotted line indicates the NaCl gradient. SDS–PAGE of the purified protein is shown as an inset. (*b*) The crystals of Pru du amandin used for diffraction studies. The dimensions of a typical crystal are about $0.3 \times 0.1 \times 0.1$ mm.

Table 1
Data-collection and processing statistics.

Values in parentheses are for the outer shell.

Space group	$P4_1$ or $P4_3$
Unit-cell parameters (Å)	$a = 150.7$, $c = 164.9$
Resolution limits (Å)	30–3.0 (3.1–3.0)
Total reflections	174223 (15413)
Unique reflections	69655 (7020)
Completeness (%)	95.5 (96.9)
Multiplicity	2.49 (2.18)
Mean $I/\sigma(I)$	3.7 (1.2)
R_{merge} (%)	13.5 (39.3)

50 mM Tris–HCl buffer pH 7.5 (Fig. 1a). The protein sample from the 40% salt fraction was subjected to a Sephacryl S-200 gel-filtration column (Pharmacia) which had been pre-equilibrated with 50 mM Tris–HCl pH 7.5 containing 140 mM NaCl. The purified protein was concentrated to 10 mg ml^{-1} using ultrafiltration (Amicon, 100 kDa cutoff, Millipore) and dialyzed against 50 mM Tris–HCl pH 7.5. The protein was further purified using ion-exchange (PI/M weak anion exchanger, Pharmacia) chromatography (Fig. 2a). 50 mM Tris–HCl pH 7.5 was used as running buffer and elution was carried out using a gradient of 0–1 M NaCl in 50 mM Tris–HCl pH 7.5 over 40 min with a flow rate of 5 ml min^{-1}, corresponding to 25.3 column bed volumes. The concentration of purified protein was determined by the BCA protein assay (Pierce Biotechnology) using bovine serum albumin as the standard. The protein was concentrated using Centriplus YM-100 (Millipore).

2.2. Protein-sequence analysis

The protein bands A40.1, A40.2 and A40.3 from the 40% ammonium sulfate cut were transferred onto polyvinylidene fluoride (PVDF) membrane (Applied Biosystems) and individually subjected to N-terminal sequencing by Edman degradation on a Procise protein sequencer (Applied Biosystems). The N-terminal sequence was used to identify the purified protein by a *BLAST* homology search optimized for the detection of short nearly exact matches in the non-redundant database (http://www.nlm.nih.gov/BLAST).

2.3. Secondary-structure analysis

A CD spectrum in the far-ultraviolet region was recorded on a Jasco J-710 spectropolarimeter connected to a Peltier temperature controller set to 283 K. The spectrum was recorded in 10 mM phosphate buffer pH 7.5 using a sample concentration of 3 μM and a cuvette of 2 cm path length. An inert nitrogen atmosphere was maintained during the measurements and each spectrum was recorded as an average of 30 scans with data collected at 0.1 nm intervals in the 190–250 nm range (Fig. 1b).

2.4. Crystallization

The protein concentration was adjusted to 15 mg ml^{-1} for crystallization. Crystallization was performed at room temperature using the hanging-drop vapour-diffusion method (McPherson, 1982). Initial crystallization experiments were set up using Crystal Screen (HR2-110, Hampton Research). The drops were composed of equal volumes (5 μl) of protein solution and precipitant solution and were equilibrated against 1 ml reservoir volume. The final crystallization conditions were optimized to 0.1 M Tris–HCl pH 8.5 containing 2.0 M ammonium sulfate as precipitant (Fig. 2b).

2.5. X-ray diffraction analysis

X-ray diffraction data were collected at 120 K using a 1:1 mixture of Paratone-N oil and paraffin oil as a cryoprotectant (HR2-132, Hampton Research). X-ray diffraction intensities were recorded from a single crystal on a MAR345dtb detector (MAR Research, Germany) with an incident X-ray wavelength of 1.5418 Å generated by a Rigaku RU-H3R copper rotating-anode generator equipped with Osmic focusing mirrors. The crystal-to-detector distance, oscillation angle and exposure time for each of the 95 images were set to 120 mm, 1° and 4 min, respectively. The data were processed using *AUTOMAR* and scaled using *SCALEPACK*.

3. Results and discussion

In the 40% ammonium sulfate fraction, two bands at 40–42 kDa, A40.1 and A40.2, share the same N-terminal sequence, ARQSQL-SPQNXQLN, whereas the N-terminal sequence of a 20–22 kDa band, A40.3, is GLEETFXSLRLKENIGNPXR (Fig. 1a). These sequences are a perfect match to the N-terminal and C-terminal subunits of Pru du amandin, respectively, with the ambiguous residues (identified by the symbol X) corresponding to two cysteine and one glutamate residue in the known sequence of Pru du amandin (PIR CAA55009). Secondary-structure analysis carried out by circular dichroism corresponded to an $\alpha+\beta$ protein with 16.2% α-helix, 49.8% β-sheet and the remainder irregular structure (Fig. 1b). This agrees with the known secondary-structural features of 11S seed legumins from other sources (Adachi *et al.*, 2001).

The protein crystals were obtained using 2.0 M ammonium sulfate in 0.1 M Tris–HCl pH 8.5 (Fig. 2b). Initially, diffraction data were collected at room temperature using capillary mounting and the crystals diffracted to a maximum resolution of 3.2 Å. Owing to the high concentration of salt in the mother liquor, none of the conventional cryoprotectants could be used for collecting diffraction data at low temperature owing to either phase separation or osmotic shock. Sucrose, xylitol, glucose, lithium sulfate and sodium malonate were all tried and all caused a deterioration in the crystal quality. Diffraction data with a completeness of 95.5% at 3.0 Å resolution were finally collected at 120 K using Paratone-N oil and paraffin oil in a 1:1 ratio as a cryoprotectant (Hampton Research). Systematic absences suggested that the crystals belong to the primitive tetragonal space group $P4_1$ (or $P4_3$), with unit-cell parameters $a = 150.7$, $c = 164.9$ Å. Details of the data collection and the processing statistics are given in Table 1. Pru du amandin shows about 42% sequence identity to soybean glycinin, which has been crystallized in a homohexameric form (Adachi *et al.*, 2003). Ara h 3, major peanut allergen and pea legumin exhibit a similar homohexameric state in solution (Breiteneder & Radauer, 2004). Assuming the presence of six protomers in the asymmetric unit, the Matthews coefficient (V_M) was calculated to be 2.6 Å^3 Da^{-1} and the solvent content was calculated to be 52.6%, which are consistent with the values for typical protein crystals.

The crystal structure of an 11S legumin from *Glycine max* (PDB code 1od5) shows the presence of two cupin folds in each protomer (Adachi *et al.*, 2003). The double-stranded β-helix that comprises the cupin fold appears to be a remarkably stable structural motif, resisting both proteolysis and thermal denaturation and thus contributing towards the allergenicity of these proteins. The core structure of Pru du amandin is expected to be very similar, as suggested by the 42% sequence identity, thereby facilitating the molecular-replacement structure-solution approach. However, there are major regional differences among the primary sequences of the two proteins, struc-

tural comparison of which may be of interest in identifying the causes of allergy. Pru du amandin appears to be exceptionally rich in polyglutamine stretches, a characteristic also found in glutelin from wheat, a completely unrelated allergenic protein. Thus, structure determination of Pru du amandin may shed light on the structural basis of the allergenicity of almonds.

References

Adachi, M., Kanamori, J., Masuda, T., Yagasaki, K., Kitamura, K., Mikami, B. & Utsumi, S. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 7395–7400.
Adachi, M., Takenaka, Y., Gidamis, A. B., Mikami, B. & Utsumi, S. (2001). *J. Mol. Biol.* **305**, 291–305.
Bischoff, S. & Crowe, S. (2005). *Gastroenterology*, **128**, 1089–1113.
Breiteneder, H. & Mills, E. N. (2005). *J. Allergy Clin. Immunol.* **115**, 14–23.
Breiteneder, H. & Radauer, C. (2004). *J. Allergy Clin. Immunol.* **113**, 821–830.
Chanana, V., Kaur, K. J. & Salunke, D. M. (2004). *Acta Cryst.* D**60**, 2100–2103.
Dalal, I., Binson, I., Reifen, R., Amitai, Z., Shohat, T., Rahmani, S., Levine, A., Ballin, A. & Somekh, E. (2002). *Allergy*, **57**, 362–365.
Demoly, P. (2006). *Allergy*, **61**, 907–909.
McPherson, A. (1982). *Preparation and Analysis of Protein Crystals*, pp. 96–97. New York: John Wiley & Sons.
Poltronieri, P., Cappello, M. S., Dohmae, N., Conti, A., Fortunato, D., Pastorello, E. A., Ortolani, C. & Zacheo, G. (2002). *Int. Arch. Allergy Immunol.* **128**, 97–104.
Qureshi, I. A., Sethi, D. K. & Salunke, D. M. (2006). *Acta Cryst.* F**62**, 869–872.
Sathe, S. K., Wolf, W. J., Roux, K. H., Teuber, S. S., Venkatachalam, M. & Sze-Tao, K. W. (2002). *J. Agric. Food Chem.* **50**, 4333–4341.
Schafer, T., Bohler, E., Ruhdorfer, S., Weigl, L., Wessner, D., Heinrich, J., Filipiak, B., Wichmann, H. E. & Ring, J. (2001). *Allergy*, **56**, 1172–1179.
Shewry, P. R., Beaudoin, F., Jenkins, J., Griffiths-Jones, S. & Mills, E. N. (2002). *Biochem. Soc. Trans.* **30**, 906–910.
Wills-Karp, M., Santeliz, J. & Karp, C. L. (2001). *Nature Rev. Immunol.* **1**, 69–75.