

Cerrone Cabanos,^a Hiroyuki
Urabe,^a Taro Masuda,^a
Mary Rose Tandang-Silvas,^a
Shigeru Utsumi,^a Bunzo Mikami^b
and Nobuyuki Maruyama^{a*}

^aLaboratory of Food Quality Design and
Development, Graduate School of Agriculture,
Kyoto University, Uji, Kyoto, Japan, and

^bLaboratory of Applied Structural Biology,
Graduate School of Agriculture, Kyoto
University, Uji, Kyoto, Japan

Correspondence e-mail:
marunobu@kais.kyoto-u.ac.jp

Received 17 April 2010
Accepted 20 July 2010

Crystallization and preliminary X-ray analysis of the major peanut allergen Ara h 1 core region

Peanuts contain some of the most potent food allergens known to date. Ara h 1 is one of the three major peanut allergens. As a first step towards three-dimensional structure elucidation, recombinant Ara h 1 core region was cloned, expressed in *Escherichia coli* and purified to homogeneity. Crystals were obtained using 0.1 M sodium citrate pH 5.6, 0.1 M NaCl, 15% PEG 400 as precipitant. The crystals diffracted to 2.25 Å resolution using synchrotron radiation and belonged to the monoclinic space group C2, with unit-cell parameters $a = 156.521$, $b = 88.991$, $c = 158.971$ Å, $\beta = 107.144^\circ$. Data were collected at the BL-38B1 station of SPring-8 (Hyogo, Japan).

1. Introduction

Peanut allergy is a major health concern worldwide (Burks, 2008), with a prevalence of 1% in children. Peanut (*Arachis hypogea*) contains at least 18 different proteins that have been identified as being capable of binding allergen-specific IgE antibodies (Scurlock & Burks, 2004; Dean, 1998). Eleven peanut allergens (Ara h 1–11) which have been officially recognized by the Allergen Nomenclature Sub-Committee of the International Union of Immunological Societies (IUIS) and two others, agglutinin and 18 kDa oleosin, are currently subject to extensive biochemical and immunological studies (Mari & Scala, 2006; <http://www.allergome.org>).

Ara h 1 is one of the three immunodominant peanut allergens. It is a vicilin, a member of the 7S globulin family (Wichers *et al.*, 2004; Kleber-Janke *et al.*, 1999; de Jong *et al.*, 1998), and is also known as conarachin. Solved structures of closely related proteins include those of 7S-1 (46.4% identity; PDB code 2ea7) and 7S-3 (46.2% identity; PDB code 2eaa) from adzuki bean (Fukuda *et al.*, 2008) and the α' (43.1% identity; PDB code 1uik) and β (50.6% identity; PDB code 1ipk) subunits of β -conglycinin from soybean (Maruyama *et al.*, 2001). Ara h 1 is a 65 kDa protein that comprises 12–16% of the total peanut protein (de Jong *et al.*, 1998) and causes sensitization in 35–95% of peanut-allergic patients from different population groups (<http://www.allergome.org>). It has been reported to form a stable trimeric protein (Pomes *et al.*, 2003), but the allergen appeared to exist in an oligomeric structure rather than a trimeric structure upon purification of native Ara h 1 from peanuts using only size-exclusion chromatography (van Boxtel *et al.*, 2006).

To date, there are no available structural data on Ara h 1. The structure of this major allergen should therefore be characterized in order to provide information vital to understanding its allergenicity and possible cross-reactivity with other 7S globulins. Here, we report the crystallization and X-ray data collection of recombinantly produced Ara h 1 core region.

2. Materials and methods

2.1. Reverse transcription–polymerase chain reaction (RT–PCR) and cloning of Ara h 1 core region

Total RNA was isolated from freshly ground peanut seeds using a Trizol kit (Invitrogen). The cDNA was amplified using an RNA LA PCR Kit (AMV) Ver. 1.1 (Takara Bio, Japan). The first-strand cDNA



was synthesized with an oligo(dT) adapter primer containing an M13 primer M4 sequence (a component of the RNA LA PCR Kit). The product was then used for PCR amplification using specifically designed primers based on the core region, amino-acid residues 145–562, of the Ara h 1 sequence (Genbank accession No. L34402). The forward primer 5'-TCTCGGAACAACCCCTTCTACTTCCC-3' starting at residue 145 and the reverse primer 5'-ATAATTAGTCGACTTAACGAGCACTCACAAAGTGAGATTCTCTG-3' ending at residue 562 and containing a stop codon (shown in bold) and a *SalI* restriction site (italicized) were used in the construction of the Ara h 1 core region. The start codon in this construct was part of the *NcoI* site of the vector used. PCR amplification was conducted using KOD-Plus DNA Polymerase (Toyobo, Japan) with initial denaturation at 367 K for 2 min followed by 30 cycles of 371 K for 10 s, 328 K for 30 s and 345 K for 2 min. The amplified fragment with the expected size was then phosphorylated and treated with *SalI*. The fragment was eventually ligated with pET21d (Novagen, California, USA) which had previously been treated with *NcoI*, blunted, treated further with *SalI* and finally dephosphorylated. Sequencing was carried out according to the dideoxy method using an ABI Prism 3100 DNA analyzer (Applied Biosystems).

The pET21d plasmid containing the Ara h 1 core region was transformed into *E. coli* BL21 (DE3) strain (Novagen, California, USA). The recombinant protein was then expressed in Luria–Bertani (LB) medium containing 0.5 M NaCl and 50 µg ml⁻¹ carbenicillin. Cultures were induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM at an OD_{600nm} of 0.60 and incubation was continued for 48 h at 293 K.

2.2. Purification

Cells were harvested by centrifugation at 6700g and the pellet was resuspended and sonicated in buffer A [35 mM potassium phosphate pH 7.4, 0.4 M NaCl, 1 mM EDTA, 0.1 mM (*p*-aminophenyl)-methanesulfonyl fluoride (*p*-APMSF), 1.2 µM leupeptin, 0.2 µM pepstatin A, 0.02% (*w/v*) Na₂S₂O₅]. Insoluble material was removed by centrifugation at 9800g and the supernatant was subjected to ammonium sulfate fractionation. Solid ammonium sulfate was slowly added to the crude extract to 30% saturation with stirring on an ice bath. After the sample had been kept on ice with stirring for 30 min,

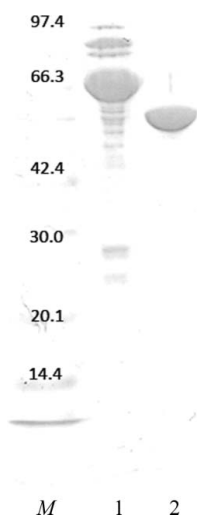


Figure 1
Protein separation by SDS–PAGE (11% gel) stained with Coomassie Brilliant Blue. Lane M, molecular-weight markers (kDa); lane 1, purified full-length mature Ara h 1 (~1 µg); lane 2, purified Ara h 1 core region (~1 µg).

it was centrifuged at 277 K for 30 min (9800g). Further ammonium sulfate was then added to the supernatant to 60% saturation and the supernatant was centrifuged again at 9800g. The final precipitate from 30–60% saturation was resuspended and dialyzed overnight in the same phosphate buffer containing only 0.15 M NaCl (buffer B).

The dialyzed was subsequently applied onto a Q-Sepharose HR 26/10 column (GE Healthcare Life Sciences, New Jersey, USA) equilibrated with buffer B. The protein was eluted with a linear gradient of NaCl from 0.15 to 0.65 M over a period of 160 min at a flow rate of 2 ml min⁻¹. The fraction containing the protein was applied onto a HiLoad 26/60 Superdex 200 prep-grade column (GE Healthcare Life Sciences, New Jersey, USA) equilibrated with buffer A at a flow rate of 2.2 ml min⁻¹. To check for purity, fractions were analyzed by SDS–PAGE using 11% polyacrylamide gels according to the procedure of Laemmli (1979) and those mainly containing bands belonging to the Ara h 1 core (47.4 kDa) were pooled together as one fraction. The protein concentration was determined using a Protein Assay Rapid Kit (Wako, Osaka, Japan) with bovine serum albumin as the standard.

2.3. Crystallization and data collection

Prior to crystallization, the Ara h 1 core was concentrated to 10 mg ml⁻¹ in buffer A using 20 ml Vivaspin concentrators with a 10 000 molecular-weight cutoff membrane (Vivascience, Germany). Screening for crystallization conditions was performed in 24-well Linbro plates using the hanging-drop vapour-diffusion technique against 1 ml well solution. The drops consisted of 1 µl concentrated protein solution mixed with 1 µl well solution. Basic/Extension, Crystal Screen, Crystal Screen 2 and PEG/Ion screen solutions (Hampton Research) were used to identify initial crystallization conditions. The crystallization temperature was set to 293 K.

Crystals were briefly soaked in a cryoprotectant solution obtained by mixing the well solution with 50% PEG 400 in a 1:1 ratio and were flash-cooled at 100 K in a stream of cold N₂ gas. X-ray diffraction images of the crystal were collected on a Jupiter 210 CCD detector (Rigaku, Tokyo, Japan) using synchrotron radiation at a wavelength of 1.0 Å at the BL-38B1 station of SPring-8 (Hyogo, Japan).

3. Results

The cDNA encoding the Ara h 1 core was successfully cloned and expressed in *E. coli* and was purified to homogeneity by ammonium sulfate fractionation followed by ion-exchange and gel-filtration



Figure 2
Ara h 1 core region crystals grown at 293 K in 0.1 M sodium citrate pH 5.6, 0.1 M NaCl, 15% PEG 400 reservoir.

Table 1

Diffraction data-processing statistics for the Ara h 1 core region.

Values in parentheses are for the highest resolution shell.

Space group	C2
Temperature (K)	100
Unit-cell parameters (Å, °)	$a = 156.521, b = 88.991,$ $c = 158.971, \beta = 107.144$
Crystal-to-detector distance (mm)	240.75
Rotation angle per image (°)	0.7
Total rotation range (°)	280
Exposure time per image (s)	10
Wavelength (Å)	1.0
Resolution limits (Å)	49.20–2.25 (2.33–2.25)
No. of observed reflections	483248 (25798)
No. of unique reflections	96996 (8062)
Completeness (%)	97.3 (81.6)
Mean $I/\sigma(I)$	16.7 (4.86)
$R_{\text{merge}}^{\dagger}$ (%)	4.9 (34.5)

$\dagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, where $I_i(hkl)$ is the observed intensity in the i th data set and $\langle I(hkl) \rangle$ is the mean intensity of reflection hkl over all measurements of $I(hkl)$.

chromatography. Fig. 1 shows the SDS–PAGE of the Ara h 1 core region (47.4 kDa) after gel-filtration chromatography. Its purity was estimated to be >95%. Full-length mature Ara h 1 contains three major parts in comparison with previously characterized 7S globulins (Maruyama *et al.*, 2001; Fukuda *et al.*, 2008; Itoh *et al.*, 2006), namely the N-terminal extension region (residues 1–144), the core region (145–562) and the C-terminal extension region (563–601). Both the N- and C-terminal extension regions are variable regions, while the core region is highly conserved in 7S globulins. In this study, we would like to stress that attempts were made to crystallize not only the core region of Ara h 1 but also the full-length mature protein and two extension regions. However, no diffraction-quality crystals were obtained after several months of screening trials and hence here we only report the crystallization of the Ara h 1 core region.

Diffraction-quality crystals were obtained after 2–3 weeks using a reservoir consisting of 0.1 M sodium citrate pH 5.6, 0.1 M NaCl, 15% PEG 400 at 293 K (Fig. 2). A complete data set was collected at 2.25 Å resolution. All images were collected at 100 K using a 0.7° oscillation angle per image. All data were processed and integrated using *DENZO* and scaled using *SCALEPACK* from the *HKL-2000* program suite (Otwinowski & Minor, 1997).

Data-collection and processing statistics are shown in Table 1. The crystals belonged to the monoclinic space group C2, with unit-cell parameters $a = 156.521, b = 88.991, c = 158.971$ Å, $\beta = 107.144^\circ$. The

asymmetric unit is likely to contain two trimers or six monomers as revealed from the initial solution structure. The initial structure was determined by the molecular-replacement technique using *MOLREP* (Vaguine *et al.*, 1999) from the *CCP4* suite (Collaborative Computational Project, Number 4, 1994) with the rigid-body refinement option and using the β subunit of β -conglycinin (PDB code 1lpk; Maruyama *et al.*, 2001) from soybean as a template. In this step, X-ray data used were from 25 to 4 Å resolution. The initial R and R_{free} factors from the resulting model were 45.3% and 45.9%, respectively. The final model structure and further analysis will be described in a subsequent manuscript.

This work was supported in part by grants from the Ministry of Education, Culture, Sports and Science and Technology of Japan (to SU) and from Asahi Breweries Foundation (to NM).

References

- Boxtel, E. L. van, van Beers, M. M. C., Koppelman, S. J., van den Broek, L. A. M. & Gruppen, H. (2006). *J. Agric. Food Chem.* **54**, 7180–7186.
- Burks, A. W. (2008). *Lancet*, **371**, 1538–1546.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Dean, T. P. (1998). *Clin. Exp. Allergy*, **28**, 7–9.
- Fukuda, T., Maruyama, N., Salleh, M. R., Mikami, B. & Utsumi, S. (2008). *J. Agric. Food Chem.* **56**, 4145–4153.
- Itoh, T., Garcia, R. N., Adachi, M., Maruyama, Y., Tecson-Mendoza, E. M., Mikami, B. & Utsumi, S. (2006). *Acta Cryst.* **D62**, 824–832.
- Jong, E. C. de, Van Zijverden, M., Spanhaak, S., Koppelman, S. J., Pellegroni, H. & Penninks, A. H. (1998). *Clin. Exp. Allergy*, **28**, 743–751.
- Kleber-Janke, T., Cramer, R., Appenzeller, U., Schlaak, M. & Becker, W. M. (1999). *Int. Arch. Allergy Immunol.* **119**, 265–274.
- Laemmli, U. K. (1979). *Nature (London)*, **227**, 680–685.
- Mari, A. & Scala, E. (2006). *Arb. Paul Ehrlich Inst. Bundesamt Sera Impfstoffe Frankf. A M.* **95**, 29–39.
- Maruyama, N., Adachi, M., Takahashi, K., Yagasaki, K., Kohno, M., Takenaka, Y., Okuda, E., Nakagawa, S., Mikami, B. & Utsumi, S. (2001). *Eur. J. Biochem.* **268**, 3595–3604.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 301–326.
- Pomes, A., Helm, R. M., Bannon, G. A., Burks, A. W., Tsay, A. & Chapman, M. D. (2003). *J. Allergy Clin. Immunol.* **111**, 640–645.
- Scurlock, A. M. & Burks, A. W. (2004). *Ann. Allergy Asthma Immunol.* **93**, Suppl. 3, S12–S18.
- Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* **D55**, 191–205.
- Wichers, H. J., De Beijer, T., Savelkoul, H. F. & Van Amerongen, A. (2004). *J. Agric. Food Chem.* **52**, 4903–4907.