Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of 8Sα globulin, the major seed storage protein of mung bean

Takafumi Itoh,[a] Roberta N. Garcia,[b] Motoyasu Adachi,[a]‡ Yukie Maruyama,[a] Evelyn Mae Tecson-Mendoza,[b] Bunzo Mikami[a] and Shigeru Utsumi[a]*

[a]Food Quality Design and Development, Graduate School of Agriculture, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan, and [b]Institute of Plant Breeding, College of Agriculture, University of the Philippines Los Baños, College, Laguna 4031, Philippines

‡ Present address: Center for Neutron Science, Tokai Research Establishment, Japan Atomic Energy Research Institute, 2-4 Shirakata, Tokai-mura, Ibaraki 319-1195, Japan.

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

Received 10 March 2006
Accepted 16 May 2006

PDB Reference: 8Sα globulin, 2cv6, r2cv6sf.

The 8S globulins of mung bean [*Vigna radiata* (L.) Wilczek] are vicilin-type seed storage globulins which consist of three isoforms: 8Sα, 8Sα' and 8Sβ. The three isoforms have high sequence identities with each other (around 90%). The structure of 8Sα globulin has been determined for the first time by X-ray crystallographic analysis and refined at 2.65 Å resolution with a final R factor of 19.6% for 10–2.65 Å resolution data. The refined 8Sα globulin structure consisted of 366 of the 423 amino-acid residues (one subunit of the biological trimer). With the exception of several disordered regions, the overall 8Sα globulin structure closely resembled those of other seed storage 7S globulins. The 8Sα globulin exhibited the highest degree of sequence identity (68%) and structural similarity (a root-mean-square deviation of 0.6 Å) with soybean β-conglycinin β (7S globulin). Their surface hydrophobicities are also similar to each other, although their solubilities differ under alkaline conditions at low ionic strength. This difference seems to be a consequence of charge–charge interactions and not hydrophobic interactions of the surfaces, based on a comparison of the electrostatic potentials of the molecular surfaces. The thermal stability of 8Sα globulin is lower than that of soybean β-conglycinin β. This correlates with the cavity size derived from the crystal structure, although other structural features also have a small effect on the protein's thermal stability.

1. Introduction

Mung beans [*Vigna radiata* (L.) Wilczek] are a popular food crop in Asia, South America, Australia and the USA and are similar to other legumes such as soybeans, jack beans and kidney beans. Mung bean seeds contain about 20–25% protein and are a major source of protein, especially in developing countries.

The major seed proteins of mung bean are storage globulins of the basic 7S type (~3%), vicilin type (8S; ~90%) and legumin type (11S; ~8%) (Tecson-Mendoza *et al.*, 2001). The 8S globulins, which are the major storage globulins of mung bean, have molecular weights of about 150 kDa and consist of three homologous isoforms, 8Sα, 8Sα' and 8Sβ, which exhibit high homology (about 90% identity) to each other and have a molecular weight of about 49 kDa, indicating that the native 8S globulin consists of heterotrimers (Tecson-Mendoza *et al.*, 2001; Bernardo *et al.*, 2004). Like other 7S globulins, they have no disulfide linkages (Bernardo *et al.*, 2004). Their amino-acid sequences exhibit similarity (about 68%) to soybean β-conglycinin β (a 7S globulin). It has been reported that soybean

Table 1
Data-collection and refinement statistics for an 8Sα globulin crystal.

Values in parentheses are for the highest resolution shell.

Crystal system	Hexagonal observed setting
Space group	*R*3
Unit-cell parameters (Å)	$a = b = 146.6$, $c = 53.3$
Molecules per ASU	1
Data collection	
Resolution limit (Å)	37.0–2.61 (2.71–2.61)
Measured reflections	33649 (1858)
Unique reflections	11640 (1014)
Redundancy	2.9 (1.8)
Completeness (%)	89.6 (71.0)
R_{sym}† (%)	4.9 (17.3)
Refinement	
Final model	366 amino-acid residues, 8 water molecules
Resolution limit (Å)	10.00–2.65 (2.74–2.65)
Reflections used	10853 (848)
Completeness (%)	89.1 (70.6)
Average *B* factor (Å^2)	52.4
R factor‡ (%)	19.6 (36.7)
R_{free}§ (%)	25.9 (46.4)
R.m.s. deviations	
Bonds (Å)	0.008
Angles (°)	1.41

† $R_{sym} = \sum |I_i - \langle I \rangle| / \langle I \rangle \times 100$, where I_i is the intensity of an individual reflection and $\langle I \rangle$ is the mean intensity of all reflections. ‡ *R* factor $= \sum |F_o - F_c| / \sum |F_o| \times 100$, where F_o is the observed structure factor and F_c is the calculated structure factor. § R_{free} was calculated from a randomly chosen 10% of reflections as defined by *CNS* (Brünger *et al.*, 1998).

seed globulins (β-conglycinin and glycinin) have high nutritional value, with a high content of the essential amino acid lysine, and have useful physicochemical properties for food systems, such as heat-induced gel-forming and emulsifying abilities (Utsumi, 1992; Utsumi *et al.*, 1997; Friedman & Brandon, 2001). Therefore, soybean globulins have been used in the production of a variety of processed foods. It has also recently been reported that the soybean globulins have several biological activities, such as a hypocholesterolaemic effect (Carroll & Kurowska, 1995), triglyceride-lowering function (Aoyama *et al.*, 2001) and an appetite-suppressing effect (Nishi *et al.*, 2003). It is expected that further studies on mung bean globulins may result in the enhancement of their characteristics corresponding to those of soybeans and will expand the food utility of mung beans.

In the improvement of the nutritional qualities and functional properties of proteins as food materials, their structural analysis is indispensable. The crystal structures of seed storage globulins such as soybean β-conglycinin α'c (core), consisting of only the core region of α' (Maruyama, Maruyama *et al.*, 2004), and β-conglycinin β (Maruyama *et al.*, 2001), soybean proglycinin A1aB1b (Adachi *et al.*, 2001, 2004; Adachi, Okuda *et al.*, 2003), mature glycinin A3B4 (Adachi, Kanamori *et al.*, 2003), jack bean canavalin (McPherson, 1980; Ko *et al.*, 1993) and kidney bean phaseolin (Lawrence *et al.*, 1994) have been determined. Moreover, the structure–function relationships of soybean proteins have been studied (Maruyama *et al.*, 1998, 1999, 2002*a*,*b*; Maruyama, Prak *et al.*, 2004; Prak *et al.*, 2005, 2006; Tandang *et al.*, 2005). Although these soybean proteins have similar three-dimensional structures, the physicochemical properties of each protein are different and unique. Physicochemical properties are important features for the utilization of these proteins as food materials. The gel-forming ability of the recombinant soybean 11S proglycinin A1aB1b subunit was enhanced by the introduction of cysteine residues and disulfide bonds (Adachi *et al.*, 2004). A large amount of a physiologically active peptide was also introduced into the 11S proglycinin A1aB1b subunit (Prak *et al.*, 2006). The peptides were released by *in vitro* digestion with trypsin. This protein engineering was based on the three-dimensional structure. Thus, the 8Sα globulin structure will provide useful information for enhancing its qualities for use in food production, *e.g.* its gelation characteristics, and for expanding the food utilization, *e.g.* the production of valuable peptides, of mung bean proteins. Therefore, we need to elucidate the structure of the mung bean 8S globulin in order to be able to make precise designs to improve the food quality of mung bean through protein and genetic engineering.

We have expressed the major isoform 8Sα of mung bean 8S globulin in *Escherichia coli* and have successfully crystallized it (Bernardo *et al.*, 2004). Recently, several physicochemical and functional properties of native and recombinant 8Sα globulin have been clarified and will be described. This article describes the crystal structure of 8Sα globulin determined by X-ray crystallography.

2. Materials and methods

2.1. Crystallization and X-ray diffraction

Recombinant 8Sα globulin was overexpressed in *E. coli*, purified and crystallized as described previously (Bernardo *et al.*, 2004). The hanging-drop vapour-diffusion method was used to crystallize the recombinant 8Sα globulin. The hanging drop (6 µl) contained 3 µl protein solution and 3 µl reservoir solution consisting of 12% PEG 1000, 0.2 *M* NaCl and 0.1 *M* MES pH 6.0 and was equilibrated against 1 ml reservoir solution. The protein concentration was 5 mg ml^{-1} and crystallization was allowed to proceed at 293 K. X-ray diffraction images of the 8Sα globulin crystal in a capillary were collected at 293 K with a Bruker Hi-Star multiwire area detector using Cu $K\alpha$ radiation generated by a MacScience M18XHF rotating-anode generator and were processed with *SADIE* and *SAINT* software (Bruker, Karlsruhe, Germany) to a resolution of 2.61 Å (Table 1).

2.2. Structure determination and refinement

The 8Sα globulin crystal structure was determined by the molecular-replacement method as implemented in *CNS* v.1.1 (Brünger *et al.*, 1998). The refined crystal structure of the recombinant soybean β-conglycinin β model was used as the probe structure (PDB code 1ipk). Model building and refinement were performed using *TURBO-FRODO* (AFMB-CNRS, Marseille, France) and *CNS*, respectively, on a Silicon Graphics Octane computer. $F_o - F_c$ and $2F_o - F_c$ maps were used to locate the correct model. Several rounds of positional and *B*-factor refinement followed by manual model building

were performed to improve the model by increasing the data to a resolution of 2.65 Å. Water molecules were incorporated where the difference density exceeded the mean by 3.0σ or more and the $2F_o - F_c$ map showed a density exceeding 1.0σ. The final R factor was 19.6% for 10 853 data points in the resolution range 10.0–2.65 Å (89.1% completeness). The free R value calculated for a randomly separated 10% of the data was 25.9%. The stereo quality of the model was assessed using *PROCHECK* (Laskowski *et al.*, 1993) and *WHAT-CHECK* (Hooft *et al.*, 1996). Structural similarity was searched for in the PDB (Berman *et al.*, 2000) using *DALI* (Holm & Sander, 1993). The coordinates of soybean β-conglycinin β (1ipk), jack bean canavalin (1dgw), kidney bean phaseolin (2phl), soybean 11S proglycinin A1aB1b (1fxz) and oxalate decarboxylase Yvrk from *Bacillus subtilis* sp. 168 (1uw8) were taken from the PDB. These models were superimposed by a fitting program in *TURBO-FRODO*. Ribbon plots were prepared using *MOLSCRIPT* (Kraulis, 1991) and *RASTER3D* (Merritt & Murphy, 1994). The accessible surface areas (ASAs) were calculated with *NACCESS* (Hubbard & Thornton, 1993), which uses the algorithm of Lee & Richards (1971). The probe was taken to be a water molecule of 1.4 Å. Electrostatic surface potential was calculated using *GRASP* (Nicholls *et al.*, 1991). The salt concentration was changed to 0.1 *M* from the default setting. The file `full.crg` was used for charge assignment, where the histidine residue has no charge. The cavity size was estimated by *CASTp* (Liang *et al.*, 1998). The aliphatic index was estimated by the *ProtParam* tool from the *ExPASy* Proteomics Server (Gasteiger *et al.*, 2003).

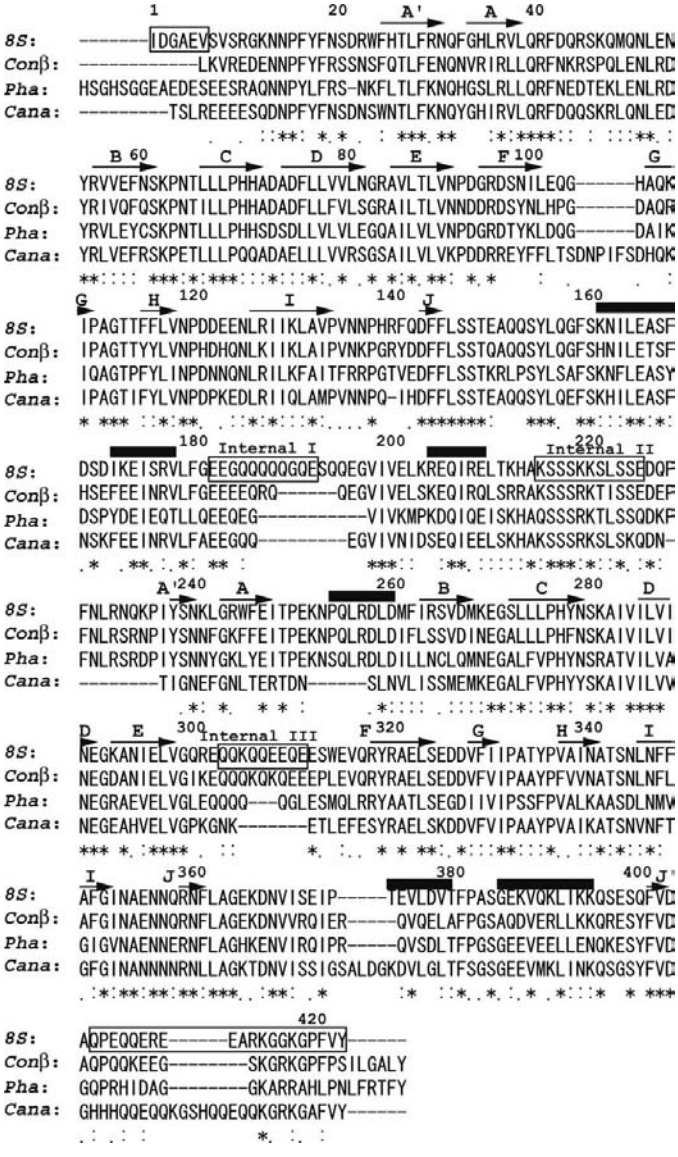

Figure 1
Amino-acid sequence alignment of 8Sα globulin and other seed storage 7S globulins obtained using *ClustalW* (http://align.genome.jp/). *8S*, 8Sα globulin of mung bean (accession No. PRF3021374A); *Con*β, soybean β-conglycinin β (accession No. P25974); *Cana*, jack bean canavalin (accession No. P50477); *Pha*, kidney bean phaseolin (accession No. AAC04316). Identical or similar amino-acid residues among the four proteins are indicated by asterisks or dots, respectively. α-Helices are indicated by black boxes and β-strands by arrows. The surrounding five amino-acid regions are disordered regions.

3. Results and discussion

3.1. Crystallization and structure determination

The recombinant 8Sα globulin of mung bean is a trimer of three identical subunits, each with a molecular weight of about 49 kDa and consisting of 423 amino-acid residues without a signal peptide (Bernardo *et al.*, 2004) (Fig. 1). 8Sα globulin crystals (0.4 × 0.3 × 0.2 mm) were obtained by the hanging-drop vapour-diffusion method as described previously (Bernardo *et al.*, 2004). The space group was determined to be $R3$, with unit-cell parameters $a = b = 146.6$, $c = 53.3$ Å, and the solvent content was 47% for one subunit per asymmetric unit. Results of data collection are summarized in Table 1. The structure of the protein was determined by the molecular-replacement method using the structure of the homologous (68% identity) soybean recombinant β-conglycinin β as the starting model (Fig. 1) and was refined by the simulated-annealing and restrained least-squares methods (Table 1).

3.2. Quality of the refined model

The refined model of 8Sα globulin consisted of 366 amino-acid residues (one subunit of the trimer) and eight water molecules. The electron densities of the main chain and side chain were generally very well defined in the $2F_o - F_c$ map, except for the five disordered regions (described below). Water molecules were also well fitted. The final overall R factor for the refined model was 19.6%, with 10 853 unique reflections in the resolution range 10.0–2.65 Å. The final free R factor was 25.9%. The final r.m.s. deviations from standard geometry were 0.008 Å for bond lengths and 1.41° for bond angles. Based on theoretical curves in the plot calculated according to Luzzati (1952), the absolute positional error was estimated to be close to 0.33 Å at a resolution of 5.0–2.65 Å. Most nonglycine residues (81.0%) lie within the most favoured regions and most other residues (18.1%) lie within the additionally allowed regions of the Ramachandran plot as

defined in *PROCHECK* (Laskowski *et al.*, 1993). However, two residues (0.6%), Arg22 ($\varphi = -175$, $\psi = -44^\circ$) and Asp325 ($\varphi = 85$, $\psi = -16^\circ$), are in generously allowed regions and one residue (0.3%), Gln399 ($\varphi = 51$, $\psi = -65^\circ$), which exhibits well defined density in the $2F_o - F_c$ map, is in a disallowed region. Arg22 and Asp325 were observed in distorted type-I β-turns, while Gln399 was located in a sharp bend of the loop near the C-terminal amino acid. The averaged *B* factor was 52.4 Å^2, which is relatively high (Table 1), probably owing to the loose crystal packing of the trimer in the unit cell.

3.3. Overall 8Sα globulin structure

The overall structure of 8Sα globulin is shown as a ribbon model in Fig. 2. There is one subunit in the asymmetric unit of the crystal. The trimer consisted of three identical subunits related by a threefold axis and has approximate dimensions of 96 × 97 × 48 Å, similar to other seed storage 7S globulins (described below; Figs. 2*a* and 2*b*). The monomer (one subunit) can be divided into two similar modules, the N- and C-terminal modules, related by a pseudo-twofold symmetry axis (Fig. 2*c*). Each module of the subunit consisted of a core β-barrel (jelly-roll) domain and an extended loop domain containing two helices. The secondary-structure elements were named according to the strands in soybean β-conglycinin β (Maruyama *et al.*, 2001). Each core β-barrel domain consisted of two β-sheets, *A′ABIDG* and *J′JCHEF*.

The 8Sα globulin model had five regions that could not be seen in the electron-density maps (the six N-terminal amino acids 1–6; the 20 C-terminal amino acids 404–423; 11 residues in the 'internal I' region, 181–191; 11 residues in the 'internal II' region, 214–224; nine residues in the 'internal III' region, 302–310) (Figs. 1 and 2). Although the total averaged *B* factor

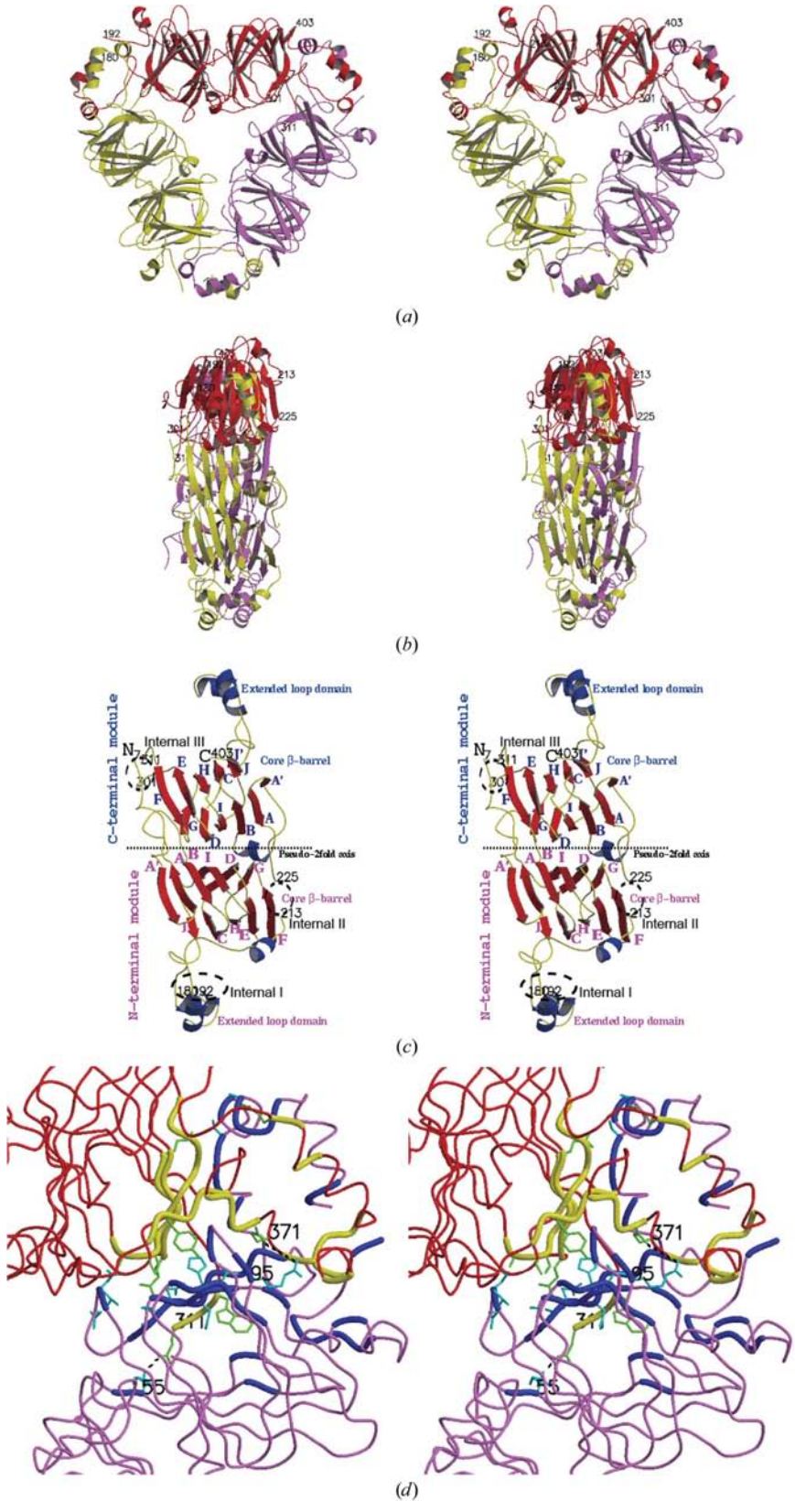

Figure 2
Overall structure of 8Sα globulin. (*a*) The 8Sα globulin trimer is seen along a threefold axis. The three subunits are shown in red, magenta and yellow. Each subunit contains two similar modules (N- and C-terminal modules) consisting of a core β-barrel and extended loop domain. (*b*) View after 90° rotation around the vertical axis of (*a*). (*c*) The 8Sα globulin subunit. Colours denote secondary-structure elements (blue, α-helices; red, β-strands; yellow, loops and coils). The broken line in the figure is the pseudo-twofold axis of the N- and C-terminal modules. (*d*) Intersubunit interface of 8Sα globulin. The structure is represented as a red and magenta ribbon model of 8Sα globulin subunits. In the ribbon model, yellow or blue denotes the residues participating in the hydrophobic interactions. Stick models (green and cyan) designate the residues participating in the hydrogen bond. Black broken lines show the hydrogen bonds of two ion pairs (Arg55–Glu311, Arg95–Glu371).

was 52.4 Å^2 and relatively high (Table 1), the *B* factors at both termini of the invisible regions were higher than the mean value. The values are 79.6 Å^2 for the N-terminal amino acid, 75.7 Å^2 for internal I, 72.3 Å^2 for internal II, 78.0 Å^2 for internal III and 62.5 Å^2 for the C-terminal amino acid. The N-terminal amino-acid sequence and molecular size indicated by SDS–PAGE of the recombinant 8Sα globulin were consistent with the deduced amino-acid sequence from the cDNA as described previously (Bernardo *et al.*, 2004). Therefore, the invisible regions are thought to be disordered.

The ASAs of one subunit and the trimer, using a probe radius of 1.4 Å, were 18 470 and 39 714 Å^2, respectively. 28.3% of the ASA of the subunit is used for formation of the trimer. The association of subunits is dominated by the extended loop domain and the *J′JCHEF* β-sheet (Fig. 2). The nonpolar atoms comprise 67.8% of the intersubunit interface. The number of residues participating in the hydrophobic interactions (C—C contacts <4 Å) were 42 in the N-terminal module and 34 in the C-terminal module, respectively. 20 hydrogen bonds (<3.2 Å) were found. In particular, two ion pairs among these interactions were observed between two modules (Arg55–Glu311, 2.7 Å; Arg95–Glu371, 3.2 Å) (Fig. 2*d*).

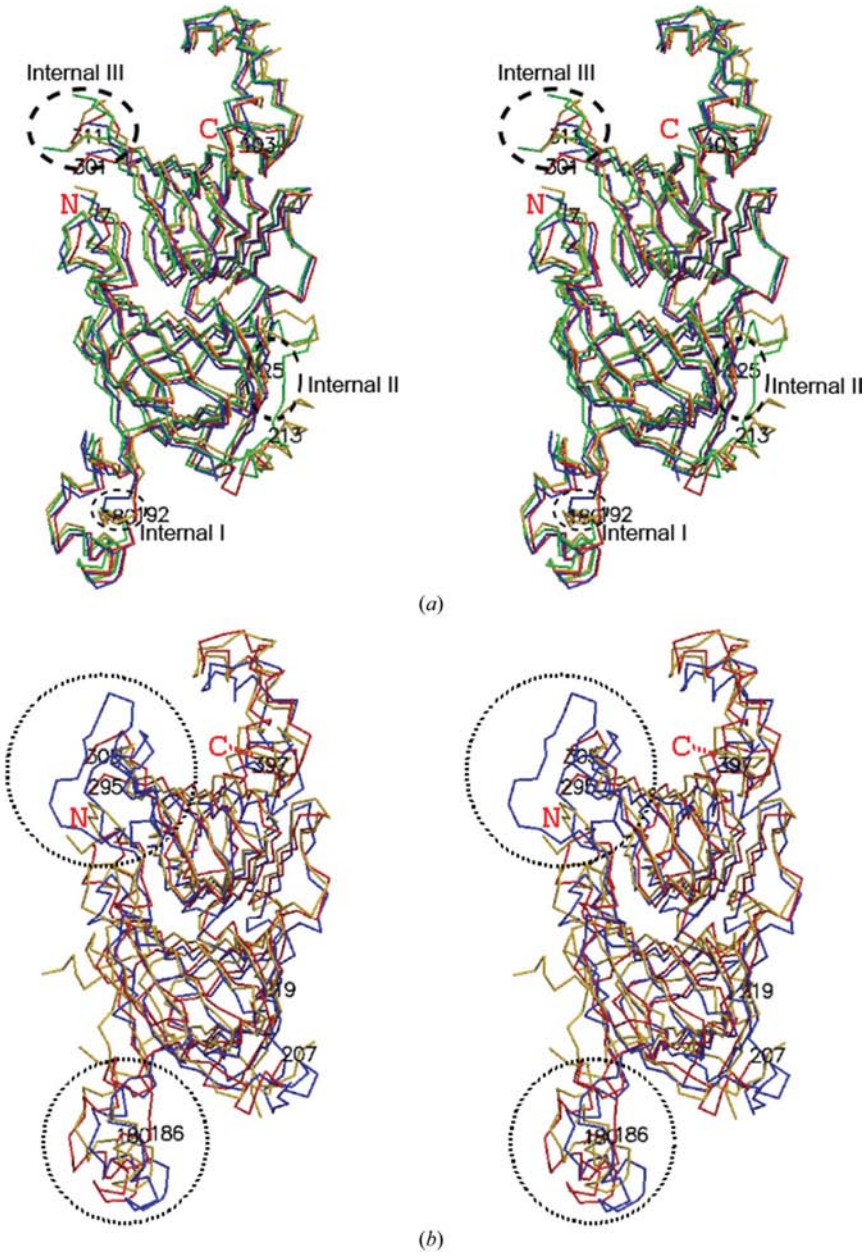

Figure 3
Structural comparison of overall structures. (*a*) 7S globulins (green, soybean β-conglycinin β; blue, jack bean canavalin; yellow, kidney bean phaseolin) superimposed onto 8Sα globulin (red). (*b*) Other structural similar protein and enzyme [yellow, soybean 11S proglycinin (A1aB1b); blue, oxalate decarboxylase Yvrk from *B. subtilis* sp. 168] also superimpose onto 8Sα globulin (red). The black broken line surrounds the site near the N-terminal amino-acid peptide and the N-terminal extended loop region, which forms a different formation in each globulin. The coordinates of β-conglycinin β (1ipk), canavalin (1dgw), phaseolin (2phl), 11S proglycinin (1fxz) and Yvrk (1uw8) were taken from the PDB (Berman *et al.*, 2000).

3.4. Structural comparison

The 8Sα globulin structure consisted of two jelly-roll folds and exhibited similarity to four other reported seed storage 7S globulin structures that have amino-acid sequence similarities to 8Sα globulin, *i.e.* soybean β-conglycinin α'c (Maruyama, Maruyama *et al.*, 2004), β-conglycinin β (Maruyama *et al.*, 2001), jack bean canavalin (McPherson, 1980) and kidney bean phaseolin (Lawrence *et al.*, 1994). The amino-acid sequence similarities between 8Sα globulin and these 7S globulins are about 68% for β-conglycinin α'c and β, about 52% for canavalin and about 58% for phaseolin (Fig. 1). In addition to seed storage 7S globulins, this basic fold is common to 11S globulins [soybean proglycinin A1aB1b (Adachi *et al.*, 2001) and mature glycinin A3B4 (Adachi, Kanamori *et al.*, 2003)], plant germine (an Mn-binding protein with oxalate oxidase and superoxide dismutase activities; Woo *et al.*, 2000), plant auxin-binding protein (Woo *et al.*, 2002) and the bacterial oxalate decarboxylase YvrK (Anand *et al.*, 2002) in the SCOP database (http://scop.mrc-lmb.cam.ac.uk/scop; Murzin *et al.*, 1995).

Fig. 3(*a*) shows the superimposition of 8Sα globulin on other 7S globulins. The r.m.s. deviation was 0.6 Å for the superimposition of 355 common C$^\alpha$ atoms of β-conglycinin β, 0.8 Å for 343 C$^\alpha$ atoms of jack bean canavalin and 0.9 Å for 342 C$^\alpha$ atoms of kidney bean phaseolin calculated by the *RIGID* program implemented in *TURBO-FRODO*. These structures had almost similar topologies, consisting of the

two core β-barrels and extended loop domains. The N- or C-terminal amino acids are disordered in other 7S globulins (Fig. 3*a*). Internal I is in the extended loop region of the N-terminal module, participating in the intersubunit interface (Fig. 2), and was longer than those of any other 7S globulins (Fig. 1). This site is also disordered in β-conglycinin β. However, in phaseolin and canavalin, this site exists as a short loop (Fig. 3*a*). Internal II is in the region connecting the N- and C-terminal modules near the two core β-barrels and does not take part in trimer formation (Fig. 2). This loop is also disordered in the 7S globulins, except for β-conglycinin β. Superimposition of this site was not well fitted in comparison with that of the overall structure (Fig. 3*a*). Therefore, this site is highly flexible within these globulins. Internal III is a neighbour of the loop near the N-terminal amino acid (Fig. 2). Although this site exists as a short loop and a turn in canavalin, this site is also disordered in other globulins (Fig. 3*a*).

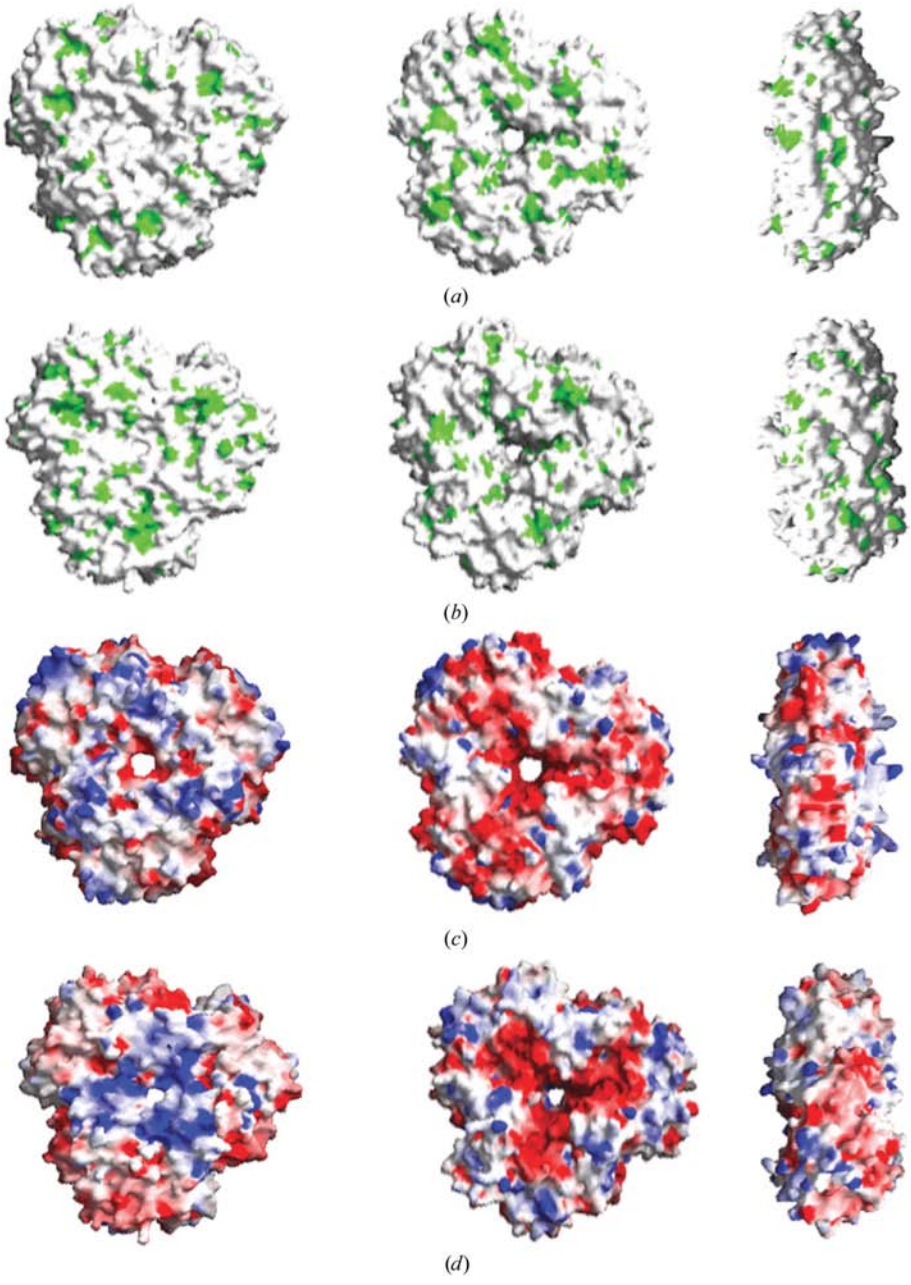

Figure 4
The molecular surface of (*a*) 8Sα globulin and (*b*) soybean β-conglycinin β. Structures are represented as white molecular-surface models. Hydrophobic residues are green. The electrostatic potential surface of (*c*) 8Sα globulin and (*d*) soybean β-conglycinin β. Structures are represented as white molecular-surface models. The electrostatic potential surface are drawn in the range from $-15k_BT$ (red) to $+15k_BT$ (blue), where k_B is Boltzmann's constant and T is the absolute temperature (K).

Fig. 3(*b*) shows the superimposition of 8Sα globulin on soybean proglycinin A1aB1b and oxalate decarboxylase YvrK from *B. subtilis* sp. 168, which also exhibited a high degree of similarity to 8Sα globulin in the PDB (Berman *et al.*, 2000) as observed using *DALI* (Holm & Sander, 1993). The r.m.s. deviation was 1.8 Å for the superimposition of 306 common C^α atoms of A1aB1b and 1.8 Å for 235 C^α atoms of Yvrk. Yvrk is a manganese-dependent enzyme that catalyzes the conversion of oxalate to formate and carbon dioxide and has a hexameric conformation consisting of two trimers on top of one another, differing from mung bean 8S globulins. The overall structures of their subunits resemble each other well. However, some differences exist in the extended loop region located in an intersubunit interface of the N-terminal module and the N-terminal amino-acid residues (Fig. 3*b*). Although the extended loop regions were superimposed well in 7S and 8S globulins (Fig. 3*a*), these regions were not well fitted in 8Sα globulin, 11S glycinin and YvrK. In Yvrk only, the loop located in the N-terminal amino-acid residues is longer and protrudes to the outside (blue C^α model in Fig. 3*b*). The manganese and four manganese-binding residues (three histidines and one glutamate) of Yvrk are not observed in the 8Sα globulin structure. These residues are not conserved in the other seed storage globulins. Furthermore, to the best of our knowledge, there are no reports of any metalloenzymatic activity of these seed storage globulins.

3.5. Molecular surface and physicochemical properties

Owing to their useful physicochemical properties, such as gel-forming and

emulsifying abilities, various seed storage globulins are important for utilization as food materials (Utsumi, 1992; Utsumi *et al.*, 1997). The surface hydrophobicity of a protein is significantly related to its solubility (Nakai & Li-Chan, 1988). Solubility is one of the most important features for the physicochemical properties of these globulins. The ASA ratio of hydrophobic residues (Ala, Val, Leu, Ile, Pro, Met, Phe, Tyr and Trp) on the molecular surface of the trimer was 4957/39714 Å^2/Å^2 (12.5%; Fig. 4*a*). This value was very close to that of soybean β-conglycinin β trimer (5494/40755 Å^2/Å^2; 13.5%; Fig. 4*b*), with which mung bean 8S globulin α has high amino-acid sequence and structural similarities (Figs. 1 and 3*a*). Therefore, the globulins would be expected to exhibit similar hydrophobic properties. Indeed, both exhibited similar retention times on hydrophobic (butyl or phenyl Sepharose) column chromatography (data not shown). The globulin is soluble in salt solution but not in water. 8Sα globulin and soybean β-conglycinin β were soluble in the examined pH range at high ionic strength (μ = 0.5). This is because electrostatic interactions between polypeptides are suppressed by the presence of the salt. However, their solubilities differ under alkaline conditions (pH > 7) at low ionic strength (μ = 0.08) (data not shown). 8Sα globulin is highly soluble at pH > 7 and μ = 0.08, while soybean β-conglycinin β is almost insoluble under these conditions. Considering the close similarity in surface hydrophobicity (Figs. 4*a* and 4*b*), the difference in the solubility under alkaline conditions at low ionic strength may arise from the difference in charge–charge interactions and not the hydrophobic interactions of protein–protein contacts through the molecular surface. Figs. 4(*c*) and 4(*d*) show the electrostatic potential on the molecular surface. In the case of 8Sα globulin, although one surface was covered by mainly negative potentias, the other was not mainly covered by positive potential (Fig. 4*c*). On the other hand, positive or negative potential was closely distributed at the centre of each surface of β-conglycinin β (Fig. 4*d*). It is thought that the predominantly and concentrated electrostatic potential (positive or negative) surfaces lead to charge–charge interactions in β-conglycinin β. Owing to this undesirable interaction through the molecular contacts, β-conglycinin β would precipitate under alkaline conditions (pH > 7) and low ionic strength. Although the effect of the disordered regions discussed above cannot be determined, the internal II disordered region (214–224), which exists on the surface and is visible in β-conglycinin β, might affect the solubility.

Thermal stability is also one of the most important physicochemical properties of these globulins related to food functions such as heat-induced gelation. The thermal stability of 8Sα globulin indicated by DSC analysis is T_m = 350.6 K at ionic strength μ = 0.5 and pH 7.6 (Table 2) (data not shown). For soybean β-conglycinin isoforms, these values under the same conditions are 351.7 K for α, 350.4 K for αc, 355.8 K for α', 356.4 K for α'c and 363.9 K for β (Maruyama *et al.*, 1999). The difference in thermal stability between β-conglycinin α'c, which consists only of the core region without the long N-terminal extension region (141 residues), and β has been accounted for by a combination of several structural features,

Table 2
Thermal stability and structural features.

		β-Conglycinin				
	8Sα globulin	α†	αc†	α'†	α'c	β
Molecular weight (kDa)	49	63	48	67	50	48
T_m‡ (K)	350.6	351.7	350.4	355.8	356.4	363.9
Cavity§ (Å^3)	6440.5	—	—	—	5463.5	4753.7
No. of hydrogen bonds¶	766 (2)	—	—	—	704 (0)	729 (1)
No. of Pro residues	19	38	20	33	19	21
Aliphatic index††	80.6	70.9	88.6	64.7	84.1	85.8

† There are no usable three-dimensional structures. ‡ DSC measurement at ionic strength (μ = 0.5) and pH 7.6 (Maruyama *et al.*, 1999). § Cavity volume was calculated by *CASTp* (Liang *et al.*, 1998). ¶ The values in parentheses represent the number of salt bridges of the intersubunit. †† This value was calculated using the *ProtParam* tool (Gasteiger *et al.*, 2003).

such as cavities, number of hydrogen bonds, intersubunit salt bridges, surface hydrophobicity, number of proline residues and loop region (Maruyama, Maruyama *et al.*, 2004). The T_m values of deletion mutants αc (deletion of the N-terminal extension region; 125 residues) and α'c (deletion of the N-terminal extension region; 141 residues) are very close to those of the α and α' subunits (Table 2; Maruyama *et al.*, 1999). Thus, the thermal stabilities of β-conglycinins are conferred by the core regions. The cavity, which is inside the molecule and which decreases the thermal stability of the 8Sα globulin, was 6441 Å^3 in volume. This value is higher than that for any other β-conglycinin (5463.5 Å^3 for α'c and 4753.7 Å^3 for β; Table 2). Furthermore, the order of these values (8Sα globulin > α'c > β) is correlated with their respective thermal stability values (8Sα globulin < α'c < β). The numbers of hydrogen bonds and salt bridges of 8Sα globulin are greater than those of β (Table 2). Although the hydrogen bonds and salt bridges are thought to be important features for stable packing, this is not consistent with the DSC experimental data. The surface hydrophobicities were very close to each other as described above. The number of proline residues, which decreases the entropy of the denatured structure, is 19 in the 8Sα globulin subunit (Table 2). Although α and α' have a higher number of proline residues than any other globulin, they do not exhibit the highest thermal stability. Also, αc and α'c have a similar number of proline residues to 8Sα globulin and β. Thus, the proline residues in the N-terminal extension region of α or α' have little effect on thermal stability. Shorter loops have been shown to be another stabilizing feature (Maruyama, Maruyama *et al.*, 2004; Chakravarty & Varadarajan, 2002). The effect of disordered regions cannot be discussed completely. The long disordered (or flexible) regions of the seed storage globulins (the N-terminal extension region of β-conglycinin and the hypervariable regions of 11S glycinin) have little influence on their thermal stabilities, as described above (Maruyama *et al.*, 1999; Prak *et al.*, 2005). For soybean β-conglycinins and mung bean 8Sα globulin, there are long disordered regions (internal I, II and III) which are rich in charged residues, *e.g.* glutamic acid and lysine residues, and the amino-acid sequences of the regions are almost conserved between the proteins (Fig. 1). This might affect the thermal

stability a little. The influence of the water molecules on the thermal stability of the storage proteins cannot be discussed here, because the resolutions of the seed storage globulin structures, at around 2.6 Å, are generally too low to observe the water molecules in the model (Maruyama *et al.*, 2001; McPherson, 1980; Ko *et al.*, 1993; Lawrence *et al.*, 1994). To discuss the contribution of water molecules to the difference in the thermal stability among β-conglycinins and 8Sα globulin, higher resolution will be needed. Furthermore, the aliphatic index, which is the relative volume occupied by aliphatic side chains (Ala, Val, Ile and Leu), has been shown to explain the difference in the thermal stability of several 11S globulins (Molina *et al.*, 2004). According to this report, the higher the aliphatic index, the higher the thermal stability. However, this trend is not exhibited among β-conglycinin and 8S globulins (Table 2).

4. Conclusion

The refined 8Sα globulin structure consisted of one subunit of the biological trimer and two jelly-roll folds. There were five invisible regions. The overall structure, with the exception of these disordered regions, very much resembled those of other seed storage 7S globulins, soybean proglycinin A1aB1b and the oxalate decarboxylase Yvrk from *B. subtilis* sp. 168. The surface hydrophobicities of 8Sα globulin and soybean β-conglycinin β were similar to each other. However, their solubilities are different under alkaline conditions at low ionic strength. The difference seems to be a consequence of charge–charge interactions and not of the hydrophobic interactions of the surfaces. The difference in the thermal stability between 8Sα globulin and β-conglycinin β was accounted for by several other structural features. In particular, the cavity correlated with this difference.

Computation time was provided by the Supercomputer Laboratory, Institute for Chemical Research, Kyoto University. This work was supported in part by a grant for the National Project on Protein Structural and Functional Analyses from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Adachi, M., Chunying, H. & Utsumi, S. (2004). *J. Agric. Food Chem.* **52**, 5717–5723.
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., Okuda, E., Kaneda, Y., Hashimoto, A., Shutov, A. D., Becker, C., Muntz, K. & Utsumi, S. (2003). *J. Agric. Food Chem.* **51**, 4633–4639.
Adachi, M., Takenaka, Y., Gidamis, A. B., Mikami, B. & Utsumi, S. (2001). *J. Mol. Biol.* **305**, 291–305.
Anand, R., Dorrestein, P. C., Kinsland, C., Begley, T. P. & Ealick, S. E. (2002). *Biochemistry*, **41**, 7659–7669.
Aoyama, T., Kohno, M., Saito, T., Fukui, K., Takamatsu, K., Hashimoto, T., Yamamoto, Y., Hirotsuka, M. & Kito, M. (2001). *Biosci. Biotechnol. Biochem.* **65**, 1071–1075.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Bernardo, A. E., Garcia, R. N., Adachi, M., Angeles, J. G., Kaga, A., Ishimoto, M., Utsumi, S. & Tecson-Mendoza, E. M. (2004). *J. Agric. Food Chem.* **52**, 2552–2560.
Brünger, 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. (1998). *Acta Cryst.* D**54**, 905–921.
Carroll, K. K. & Kurowska, E. M. (1995). *J. Nutr.* **125**, 594S–597S.
Chakravarty, S. & Varadarajan, R. (2002). *Biochemistry*, **41**, 8152–8161.
Friedman, M. & Brandon, D. L. (2001). *J. Agric. Food Chem.* **49**, 1069–1086.
Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D. & Bairoch, A. (2003). *Nucleic Acids Res.* **31**, 3784–3788.
Holm, L. & Sander, C. (1993). *J. Mol. Biol.* **233**, 123–138.
Hooft, R. W., Vriend, G., Sander, C. & Abola, E. E. (1996). *Nature (London)*, **381**, 272.
Hubbard, S. J. & Thornton, J. M. (1993). *NACCESS Computer Program.* Department of Biochemistry and Molecular Biology, University College, London.
Ko, T.-P., Ng, J. D. & McPherson, A. (1993). *Plant Physiol.* **101**, 729–744.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lawrence, M. C., Izard, T., Beuchat, M., Blagrove, R. J. & Colman, P. M. (1994). *J. Mol. Biol.* **238**, 748–776.
Lee, B. & Richards, F. M. (1971). *J. Mol. Biol.* **55**, 379–400.
Liang, J., Edelsbrunner, H. & Woodward, C. (1998). *Protein Sci.* **7**, 1884–1897.
Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.
McPherson, A. (1980). *J. Biol. Chem.* **255**, 10472–10480.
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.
Maruyama, N., Katsube, T., Wada, Y., Oh, M. H., Barba De La Rosa, A. P., Okuda, E., Nakagawa, S. & Utsumi, S. (1998). *Eur. J. Biochem.* **258**, 854–862.
Maruyama, N., Mohamad Ramlan, M. S., Takahashi, K., Yagasaki, K., Goto, H., Hontani, N., Nakagawa, S. & Utsumi, S. (2002*a*). *J. Am. Oil Chem. Soc.* **79**, 139–144.
Maruyama, N., Mohamed Ramlan, M. S., Takahashi, K., Yagasaki, K., Goto, H., Hontani, N., Nakagawa, S. & Utsumi, S. (2002*b*). *J. Agric. Food Chem.* **50**, 4323–4326.
Maruyama, N., Prak, K., Motoyama, S., Choi, S. K., Yagasaki, K., Ishimoto, M. & Utsumi S. (2004). *J. Agric. Food Chem.* **52**, 8197–8201.
Maruyama, N., Sato, R., Wada, Y., Matsumura, Y., Goto, H., Okuda, E., Nakagawa, S. & Utsumi, S. (1999). *J. Agric. Food Chem.* **47**, 5278–5284.
Maruyama, Y., Maruyama, N., Mikami, B. & Utsumi, S. (2004). *Acta Cryst.* D**60**, 289–297.
Merritt, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.
Molina, M. I., Petruccelli, S. & Añón, C. (2004). *J. Agric. Food. Chem.* **52**, 6023–6029.
Murzin, A. Z., Brenner, S. E., Hubbard, T. J. & Chothia, C. (1995). *J. Mol. Biol.* **247**, 536–540.
Nakai, S. & Li-Chan, E. (1988). *Hydrophobic Interactions in Food Systems*, pp. 43–61. Boca Raton, FL, USA: CRC Press.
Nicholls, A., Sharp, K. & Honig, B. (1991). *Proteins*, **11**, 281–296.
Nishi, T., Hara, H. & Tomita, F. (2003). *J. Nutr.* **133**, 352–357.
Prak, K., Maruyama, Y., Maruyama, N. & Utsumi, S. (2006). *Peptides*, **27**, 1179–1186.

Prak, K., Nakatani, K., Katsube-Tanaka, T., Adachi, M., Maruyama, N. & Utsumi S. (2005). *J. Agric. Food Chem.* **53**, 3650–3657.
Tandang, M. R., Atsuta, N., Maruyama, N., Adachi, M. & Utsumi, S. (2005). *J. Agric. Food. Chem.* **53**, 8736–8744.
Tecson-Mendoza, E. M., Adachi, M., Bernardo, A. E. & Utsumi, S. (2001). *J. Agric. Food Chem.* **49**, 1552–1558.
Utsumi, S. (1992). *Adv. Food Nutr. Res.* **36**, 89–208.
Utsumi, S., Matsumura, Y. & Mori, T. (1997). *Food Proteins and Their Applications*, edited by S. Damodaran & A. Paraf, pp. 257–291. New York: Marcel Dekker.
Woo, E. J., Dunwell, J. M., Goodenough, P. W., Marvier, A. C. & Pickersgill, R. W. (2000). *Nature Struct. Biol.* **7**, 1036–1040.
Woo, E. J., Marshall, J., Bauly, J., Chen, J. G., Venis, M., Napier, R. M. & Pickersgill, R. W. (2002). *EMBO J.* **21**, 2877–2885.