Acta Cryst. (1994). D**50**, 627–631

Crystallization of Seed Globulins from Legumes

BY MICHAEL HENNIG*

Humboldt University, Institute of Biochemistry, Hessische Str. 3–4, 10098 Berlin, Germany

AND BERNHARD SCHLESIER

Institute of Plant Genetics and Crop Plant Research, Correnstrasse 3, 06466 Gatersleben, Germany

(*Received* 10 *September* 1993; *accepted* 22 *November* 1993)

Abstract

Seeds contain large quantities of proteins and are therefore main food sources. In the last century, protein extracts of legume seeds were dialysed against distilled water and in some cases small crystals of pure protein appeared. However, those crystals were generally of poor quality with respect to X-ray diffraction. Recently, the crystallization of some of them was improved and the structures of two 7S globulins, phaseolin from *Phaseolus vulgaris* and canavalin from *Canavalia ensiformis*, have been determined at 3.0 and 2.6 Å resolution, respectively. Efforts to improve the quality of the phaseolin crystals resulted in three new crystal forms which will be discussed in this paper. The only high-resolution X-ray analysis of a seed globulin from legumes is that of narbonin, a 2S protein from *Vicia narbonensis*. The crystal structure at 1.8 Å shows a very compact packing in layers of molecules. The intermolecular contacts include salt bridges and hydrophobic clusters that might facilitate both the aggregation of the molecules and their crystallization. Because the seed globulins appear in large quantities in the protein bodies of the seeds, efficient packing of the molecules similar to the crystal packing can be assumed.

Introduction

The proteins of legume seeds are of interest because of their nutritional value. Even today their solubility properties are used for their classification. According to Osborne (1907) these proteins can be classified as albumins, globulins, gliadins and glutenins depending on their solubility in water, saline solutions, 70% aqueous ethanol and diluted acid or alkali solutions.

Three types of seed globulins are generally distinguished according to their sedimentation coefficients without taking functional properties into account. Most of the legume globulins have sedimentation coefficients of 11–12S (legumin-like globulins) or 7–8S (vicilin-like globulins), but smaller proteins of 2S have also been reported (Derbyshire, Wright & Boulter, 1976; Shotwell & Larkins, 1989; Sun & Larkins, 1992). Independent from this classification, and in addition to their common solubility properties, the globulins of legume seeds share similar biochemical properties. All globulins from legumes are almost insoluble in distilled water, but are soluble at higher ionic strength. In addition, they contain high levels of charged amino acids like glutamic acid, aspartic acid and lysine and are low in sulfur-containing amino acids. All globulins are assumed to lack enzymatic activity and simply provide a source of nitrogen for the developing seedling. For that reason they are commonly called storage proteins.

Seed globulins were amongst the first proteins for which extraction and crystallization succeeded (Ritthausen, 1881). Some of these crystals were utilized for crystallographic characterization, but they were generally not suitable for X-ray structural analysis at atomic resolution. Here we review the conditions for crystallization of the seed globulins from legumes and discuss, using the 2S globulin narbonin as an example, how the solubility properties of these proteins might be explained by their three-dimensional structure.

11S globulins

The most heterogeneous storage proteins are the 11S globulins. These are organized in hexamers of similar subunits and have a molecular weight of about 360 kDa. Every subunit contains two different disulfide-linked polypeptide chains. Until now no crystals of an 11S globulin from legumes have been reported.

However, 11S globulins from other sources have crystallized spontaneously during their purification upon lowering of the ionic strength of the solution. Three proteins, cucurbitin from cucumber, excelsin from brazil nut and edestin from hemp seed were

* Current address: Biozentrum, Department of Structural Biology, University of Basel, Klingelbergstr. 70, 4056 Basel, Switzerland.

ISSN 0907-4449

Table 1. *Crystallization of 7S globulins from legume seeds*

Crystallization conditions	Space group	Cell dimensions (Å, °)	Resolution (Å)	Reference
Phaseolin from *Phaseolus vulgaris*				
70% $(NH_4)_2SO_4$	$P432$	$a = b = c = 67$	2.8	Johnson *et al.* (1982)
Temperature shift	$F4_132$	$a = b = c = 274$	4	This paper
27% ethanol (phytic acid free phaseolin)	$P2_12_12_1$	$a = 113$, $b = 130$, $c = 89$	3	Suzuki *et al.* (1983)
23% ethanol	$P2_12_12$	$a = 128.2$, $b = 135.7$, $c = 161.6$	3	Suzuki *et al.* (1983)
30% ethanol	$P2_12_12$	$a = 124.1$, $b = 133.5$, $c = 162.0$	3	This paper
33% ethanol, $Ba(CH_3COO)_2$	$R3$	$a = b = 194.5$, $c = 207.5$, $\gamma = 120$	3	This paper
Canavalin from *Canavalia ensiformis*				
pH shift from alkaline	$R3$	$a = b = 136$, $c = 75$, $\gamma = 120$	2.8	McPherson & Spencer (1975)
(recombinant canavalin)	$R3$	$a = b = 136$, $c = 75$, $\gamma = 120$	3.2	Ng *et al.* (1993)
pH shift from acidic, EDTA	$P6_3$	$a = b = 126.3$, $c = 51.6$, $\gamma = 120$	2.8	McPherson & Spencer (1975)
			2.6	Ko *et al.* (1993)
pH shift to pH 6.8, 3% NaCl	$C222_1$	$a = 126$, $b = 152$, $c = 131$	2.8	McPherson & Spencer (1975)
pH shift to pH 6.8, 1% NaCl	$P2_13$	$a = 106.0$	2.1	Ko *et al.* (1993)

dissolved in a sodium-chloride-containing buffer. The crystallization was performed by dialysis against a lower concentration of NaCl for cucurbitin (Colman, Suzuki & Van Donkelaar, 1980) or by dialysis against pure water or buffer in the absence of NaCl for excelsin (Drenth & Wiebenga, 1955; Kamiya *et al.*, 1983) and edestin (Drenth, 1957). No attempt has been reported to solve the structure of an 11S globulin with present-day techniques.

7S globulins

The 7S globulins are organized in trimers of about 150 kDa with slightly different subunits. The first crystals of this class were reported for canavalin after tryptic cleavage. However, this protein does not crystallize prior to such cleavage. Crystallization of the cleaved canavalin was performed using the pH dependence of its solubility.

For phaseolin from *Phaseolus vulgaris*, one of the best characterized seed proteins, several crystallization conditions as well as its crystal structure at 3 Å resolution have been published (Lawrence *et al.*, 1990). For the structural analysis, orthorhombic crystals obtained by dialysis against 23% ethanol were used (Table 1). There are two trimers in the asymmetric unit in this crystal form.

Recently, the structure of canavalin has been determined at 2.6 Å for orthorhombic and rhombohedral crystals and at 2.3 Å resolution for the cubic crystal form (Table 1). The crystal packing and intermolecular interactions of the molecules in the crystal forms are different. Crystal contacts are almost exclusively hydrophilic including 3–8 salt bridges between the trimeric molecules; none of the bridges are found in all three crystal forms. Intermolecular interactions involving hydrophobic contacts have not been described, but there are hydrophobic interactions between neighbouring subunits and the two domains (Ko, Ng & McPherson, 1993).

The phaseolin and canavalin structures are very similar. Their monomeric subunits are composed of two almost identical domains consisting of an eight-stranded antiparallel β-barrel with Swiss-roll topology and an extended subdomain with the helix–turn–helix motif. Unfortunately, structural details such as the glycosylation sites in phaseolin have not yet been investigated because of the limited resolution of the analysis.

To improve the quality of the crystals we purified phaseolin by a new method (Schlesier, Manteuffel, Rudolph & Jüttner, 1984). Using ethanol as precipitant we observed similar crystals to those obtained by Suzuki *et al.* (1983) and used for the structural analysis of phaseolin. The addition of divalent metal ions (10 m*M* barium acetate) together with an increase of the ethanol concentration resulted in rather stable rhombohedral crystals (Table 1). This

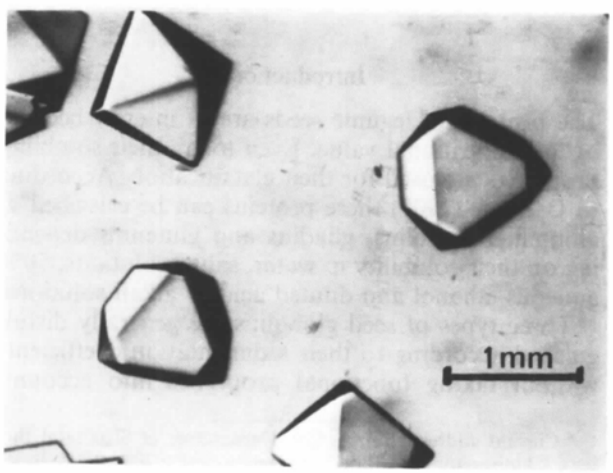

Fig. 1. Crystals of phaseolin with cubic symmetry $F4_132$.

Table 2. *Crystallization of 2S globulins from legume seeds*

Crystallization conditions	Space group	Cell dimensions (Å,°)	Resolution (Å)	Reference
Narbonin from *Vicia narbonensis*				
pH shift from alkaline	$P2_1$	$a = 46.7$, $b = 75.1$, $c = 50.7$, $\beta = 120.5$	1.5	Hennig *et al.* (1990)
Narbonin from *Vicia pannonica*				
pH shift from alkaline	$P2_1$	$a = 46.6$, $b = 73.9$, $c = 48.5$, $\beta = 116.3$	1.9	Schlesier *et al.* (1993)
Concanavalin B from *Canavalia ensiformis*				
32% $(NH_4)_2SO_4$	$P6_1$	$a = b = 80.9$, $c = 102.2$, $\gamma = 120$	5	Morrison *et al.* (1984)

crystallization condition is very reproducible and was applied in the COSIMA protein crystal growth facility on unmanned satellites (Hennig *et al.*, 1994). By slow cooling of a very concentrated (80 mg ml^{-1}) protein solution from 333 to 277 K in 24 h during dialysis against distilled water we obtained cubic crystals (Fig. 1). The size of these crystals was sufficient for X-ray analysis, but the quality of the diffraction could not be significantly improved.

2S globulins

The 2S globulins consist of only one polypeptide chain and have a molecular weight of about 30 kDa. The crystallization of three proteins in this class, the narbonins from *Vicia narbonensis* and *V. pannonica* and concanavalin B from *Canavalia ensiformis*, has been described (Hennig, Schlesier, Pfeffer & Höhne, 1990; Schlesier, Hennig, Kraft & Horstmann, 1993; Morrison, DeLozier, Robinson & McPherson, 1984). They can be purified by dialysis against distilled water. But the methods of crystallization are somewhat different. Concanavalin B was crystallized using $(NH_4)_2SO_4$ as precipitant (Table 2). Its structural analysis is underway (Ko, Ng & McPherson, 1993).

Narbonin is almost insoluble at neutral pH, but can be dissolved at basic pH in a buffer solution containing sodium chloride. Finally, a combination of the salting-in effect and the strong pH dependence of its solubility during dialysis against NaCl free buffer containing acetic acid resulted in crystals of better quality than observed previously for seed globulins. They diffract to 1.5 Å resolution and were used for structure determination (Hennig *et al.*, 1992).

A similar crystallization method was used for canavalin and resulted in crystals of rather different symmetry and crystal packing (*cf.* Tables 1 and 2).

Interrelation between solubility and crystal structure

Because of the very similar solubility properties of all globulins it might be of interest to understand how this property is represented by the three-dimensional structure of the protein. The high-resolution structure of narbonin can be used to discuss features of the packing of the molecules and the crystal contacts as well as hydrophobic properties. Narbonin has an α/β-barrel folding topology, consisting of an eight-fold parallel β-barrel in the centre of the molecule surrounded by α-helices. This folding has been found in at least 21 different proteins, all enzymes, with their active site on the surface at the C-terminal side of the β-strands (Farber, 1993). Narbonin is the first protein with this triose phosphate isomerase-like folding without known enzymatic activity.

Narbonin has an accessible surface of 13 200 Å^2 (Kabsch & Sander, 1983). Most of the surface is covered by hydrophilic residues and many hydrogen bonds are formed between these residues and surrounding solvent molecules. The packing of the narbonin molecules in the crystal is compact (V_m = 2.3 Å^3 Da^{-1}), resulting in many interactions between neighbouring molecules. Within layers of molecules there are many close contacts, but fewer contacts between the layers. In Fig. 2 such a layer is presented, showing the interactions between the loop

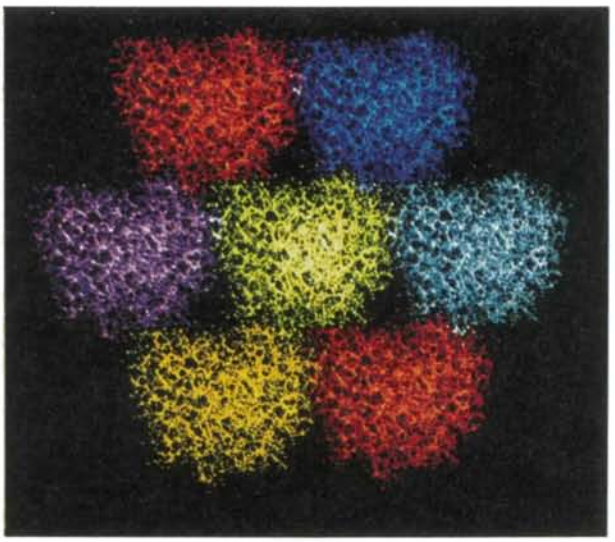

Fig. 2. Packing of molecules in crystals of narbonin.

Table 3. *Intermolecular salt bridges of narbonin*

All salt bridges with distances less than 3.5 Å between the nitrogen and the nearest oxygen are included. The secondary structure is numbered according to β/α-barrel proteins in such a way that the parallel β-barrel is numbered $\beta 1$ to $\beta 8$ and the following α-helices $\alpha 1$ to $\alpha 8$. β' and β'' represent the antiparallel β-sheet following $\beta 2$. α'' is an additional short helix before $\alpha 8$.

Positively charged side chain	Position in the structure	Negatively charged side chain	Position in the structure	Distance (Å)
Lys49 Nζ	β-β''	Glu77 Oε2	$\alpha 2$-$\beta 3$	2.74
	$\alpha 7$	Glu97 Oε2	$\beta 3$-$\alpha 3$	2.76
Arg239 Nε	$\alpha 7$	Glu97 Oε1	$\beta 3$ $\alpha 3$	2.66
Arg271 Nη1	α'' $\alpha 8$	Glu97 Oε2	$\beta 3$ $\alpha 3$	3.02
Lys275 Nζ	α'' $\alpha 8$	Glu107 Oε2	$\alpha 3$	3.48

regions of neighbouring molecules. The C-terminal side of the β-barrel faces a solvent channel.

At least 17 intermolecular hydrogen bonds can be observed, including five salt bridges (Table 3). They stabilize extended loop regions of the molecule. The loop containing Arg271 and Lys275 is fixed by two salt bridges like a claw to α-helix 3 of the neighbouring molecule. Three of the salt bridges are formed by lysin and it is likely that these interactions are weakened by raising the pH.

In addition to hydrophilic residues, we also observed several solvent-accessible hydrophobic residues. Assuming a single molecule in the solution we can identify at least five isoleucines, one phenylalanine, four valines, five leucines and nine prolines which have a solvent accessibility greater than 30 Å^2 per residue (Kabsch & Sander, 1983). In the crystal the number of these residues is reduced, since eight are involved in intermolecular interactions. For example, Phe253 forms a hydrophobic cluster with Ile268 and Val267 of the other molecule (Fig. 3).

Nevertheless, there are several positions in which hydrophobic side chains are exposed to the solvent, rejecting all water molecules in a minimum radius of 3.5 Å. These hydrophobic surface areas cause the very low solubility of narbonin in distilled water. The described crystal contacts can explain quite well why the narbonin crystals are easy to obtain and very stable once they had been grown. The close packing of the molecules seen in the crystal structure is enforced by intermolecular interactions. It may well be that a similar packing of the molecules exists in the seeds.

We gratefully thank Johan N. Jansonius and Betty Shen, Biocentre Basel, for discussions and careful reading of the manuscript. We thank Keith S. Wilson, EMBL Outstation Hamburg, for help during the structural analysis of narbonin and Wolfgang E. Höhne, Humboldt University Berlin, for advice on the crystallization of phaseolin.

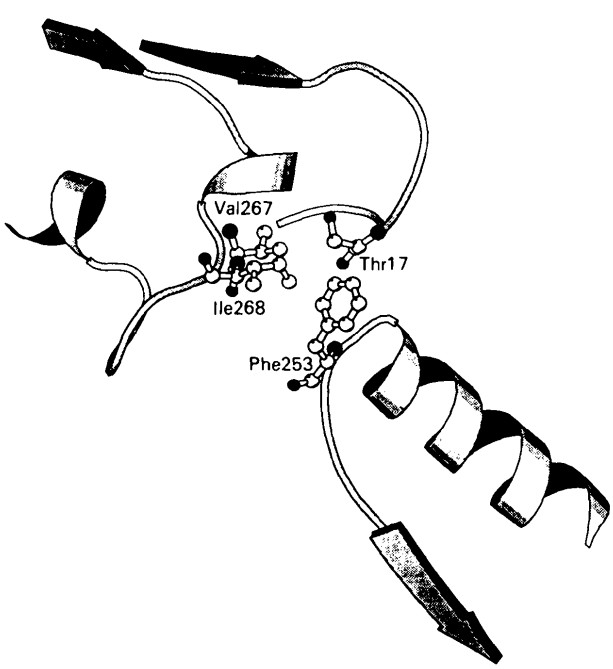

Fig. 3. Hydrophobic interaction between two neighbouring molecules in crystals of narbonin.

References

COLMAN, P. M., SUZUKI, E. & VAN DONKELAAR, A. (1980). *Eur. J. Biochem.* **103**, 585–588.

DERBYSHIRE, E., WRIGHT, D. J. & BOULTER, D. (1976). *Phytochemistry*, **15**, 3–24.

DRENTH, J. (1957). Thesis, Univ. of Groningen, The Netherlands.

DRENTH, J. & WIEBENGA, E. H. (1955). *Recl. Trav. Chim. Pays Bas*, **74**, 813–832.

FARBER, G. K. (1993). *Curr. Opin. Struct. Biol.* **3**, 409–412.

HENNIG, M., SCHLESIER, B., DAUTER, Z., PFEFFER, S., BETZEL, C., HÖHNE, W. E. & WILSON, K. S. (1992). *FEBS Lett.* **306**, 80–84.

HENNIG, M., SCHLESIER, B., PFEFFER, S. & HÖHNE, W. E. (1990). *J. Mol. Biol.* **215**, 339–340.

HENNIG, M., VISANJI, M., WEBER, W., JANCZIKOWSKI, H., PLAAS-LINK, A. & BETZEL, CH. (1994). *J. Cryst. Growth*, **135**, 513–522.

JOHNSON, S., GRAYSON, G., ROBINSON, L., CHAHADE, R. & MCPHERSON, A. (1982). *Biochemistry*, **21**, 4839–4843.

KABSCH, W. & SANDER, CH. (1983). *Biopolymers*, **22**, 2577–2637.

KAMIYA, N., SAKABE, K., SAKABE, N., SASAKI, K., SAKAKIBARA, M. & NOGUCHI, H. (1983). *Agric. Biol. Chem.* **47**, 2091–2098.

KO, T.-P., NG, J. D., DAY, J., GREENWOOD, A. & MCPHERSON, A. (1993). *Acta Cryst.* D**49**, 478–489.

KO, T.-P., NG, J. D. & MCPHERSON, A. (1993). *Plant Physiol.* **101**, 729–744.

LAWRENCE, M. C., SUZUKI, E., VARGHESE, J. N., DAVIS, P. C., VAN DONKELAAR, A., TULLOCH, P. A. & COLMAN, P. M. (1990). *EMBO J.* **9**, 9–15.

MCPHERSON, A. & SPENCER, R. (1975). *Arch. Biochem. Biophys.* **169**, 650–661.

MORRISON, R., DELOZIER, G., ROBINSON, L. & MCPHERSON, A. (1984). *Plant Physiol.* **76**, 175–183.

NG, J. D., KO, T.-P. & MCPHERSON, A. (1993). *Plant Physiol.* **101**, 713–728.

OSBORNE, T. B. (1907). *Carnegie Inst. Washington Publ.* **84**, 1–119.

RITTHAUSEN, H. (1881). *J. Prakt. Chem.* **23**, 481–486.

SCHLESIER, B., HENNIG, M., KRAFT, R. & HORSTMANN, CH. (1993). *Food Proteins Structure and Functionality*, pp. 118–120. Weinheim: VCH Publishers.

SCHLESIER, B., MANTEUFFEL, R., RUDOLPH, A. & JÜTTNER, G. (1984). *Biochem. Physiol. Pflanz.* **179**, 665–678.

SCHLESIER, B. & SCHOLZ, G. (1974). *Biochem. Physiol. Pflanz.* **166**, 367–369.

SHOTWELL, M. A. & LARKINS, B. A. (1989). *Biochemistry of Plants*, Vol. 15, pp. 297–345. New York: Academic Press.

SUN, S. S. M. & LARKINS, B. A. (1992). *Transgenic Plants*, pp. 339–372. New York: Academic Press.

SUZUKI, E., VAN DONKELAAR, A., VARGHESE, J. N., LILLEY, G. G., BLAGROVE, R. J. & COLMAN, P. M. (1983). *J. Biol. Chem.* **258**, 2634–2636.