Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

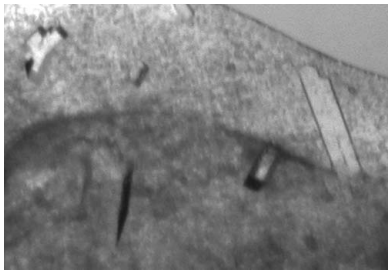
Purification, crystallization and initial crystallographic characterization of the *Ginkgo biloba* 11S seed globulin ginnacin

Tengchuan Jin, Yu-Wei Chen, Andrew Howard and Yu-Zhu Zhang*

Department of Biological, Chemical and Physical Sciences, Illinois Institute of Technology, Chicago, IL 60616, USA

Correspondence e-mail: zhangy@iit.edu

Received 22 February 2008
Accepted 29 May 2008



Ginkgo biloba, a well known 'living fossil' native to China, is grown worldwide as an ornamental shade plant. Medicinal and nutritional uses of *G. biloba* in Asia have a long history. However, ginkgo seed proteins have not been well studied at the biochemical and molecular level. In this study, the *G. biloba* 11S seed storage protein ginnacin was purified by sequential anion-exchange and gel-filtration chromatography. A crystallization screen was performed and well diffracting single crystals were obtained by the vapor-diffusion method. A molecular-replacement structural solution has been obtained. There are six protomers in an asymmetric unit. Structure refinement is currently in progress.

1. Introduction

Phylogenetic studies place ginkgo (*Ginkgo biloba* L.) in its own division (Gold *et al.*, 2002), making it the world's oldest living species of tree. As the only genetic link between lower and higher plants, the ginkgo's evolutionary history shows divergence from ferns and conifers. A deciduous tree, ginkgo has a life span more than 1000 y. Originating from China, ginkgo was introduced to Japan and Korea centuries ago and subsequently to Europe in the 1700s.

Ginkgo leaves and seeds have been used in traditional Chinese medicine and as a foodstuff for centuries throughout Asia. Literature searches found numerous published studies on the medical use of *G. biloba.* Today, *G. biloba* leaf extract is recognized as the most widely used herbal treatment for improving cognitive functions such as memory, learning, alertness and mood (Gold *et al.*, 2002, 2003).

The legumin- and vicilin-like seed storage globulins of spermatophytes are accumulated as a protein body during embryogenesis and seed development. They function as a nitrogen and carbon reserve during seedling growth. It is generally accepted that legumins and vicilins evolved from a common ancestral gene encoding the prototype of the present common (cupin) domain in the subunits of these globulins (Baumlein *et al.*, 1995; Shutov *et al.*, 1995). The corresponding vicilin-type seed storage protein has not been identified in ginkgo. The 11S legumin-type seed storage protein has recently been purified, characterized and cloned (Arahira & Fukazawa, 1994; Hager *et al.*, 1995). However, its three-dimensional structure has not been solved.

Because of the special status of *G. biloba* in the evolutionary tree, not only one-dimensional nucleotide information but also three-dimensional structure information is needed to study its molecular evolution. The evolution of the cupin-protein family and structural comparison of cupin-family proteins from different species are especially sought because numerous members of this family have been identified as food allergens. Here, we report the crystallization, X-ray data collection and initial crystallographic characterization of ginnacin.

2. Methods

2.1. Purification

Fresh ginkgo seeds were collected from Morton Arboretum (Illinois, USA). Seed proteins were extracted from 30 g of ground shelled

seeds in 170 ml extraction buffer (1 *M* NaCl, 20 m*M* Tris–HCl pH 7.9). The extract was subjected to centrifugation at 20 000*g* for 10 min and the supernatant was collected as crude extract for further use. Using a dialysis tube with a molecular-weight cutoff of 5 kDa, the crude extract was dialyzed twice for 12 h against distilled water at 277 K. The dialyzed sample was again subjected to centrifugation at 20 000*g* for 10 min and the supernatant was filtered with 0.45 µm pore-size syringe filters. All the above procedures were carried out in the presence of protease inhibitors (100 n*M* aprotinin, 50 m*M* antipain, 50 m*M* leupeptin and 0.5 mg ml^{-1} pepstatin) and antibiotics (50 mg ml^{-1} ampicillin and 50 mg ml^{-1} kanamycin). The filtrate was then loaded onto an 8 ml Source Q15 column (GE Healthcare, Piscataway, New Jersey, USA) pre-equilibrated with 10 m*M* Tris buffer (10 m*M* Tris–HCl pH 7.9) and eluted with a 0–0.6 *M* linear NaCl gradient in 100 ml Tris buffer.

The 11S globulin-containing fractions were pooled and loaded onto a 300 ml Superdex 200 column (GE Healthcare) pre-equilibrated and eluted with buffer *A* (10 m*M* Tris–HCl pH 7.9, 100 m*M* NaCl). The gel-filtration column was calibrated with thyroglobulin (669 kDa), ferritin (440 kDa), catalase (242 kDa), aldolase (158 kDa) and bovine serum albumin (67 kDa) obtained from a protein standard kit (GE Healthcare). The flow rate was 1 ml min^{-1} for all the purification and calibration runs.

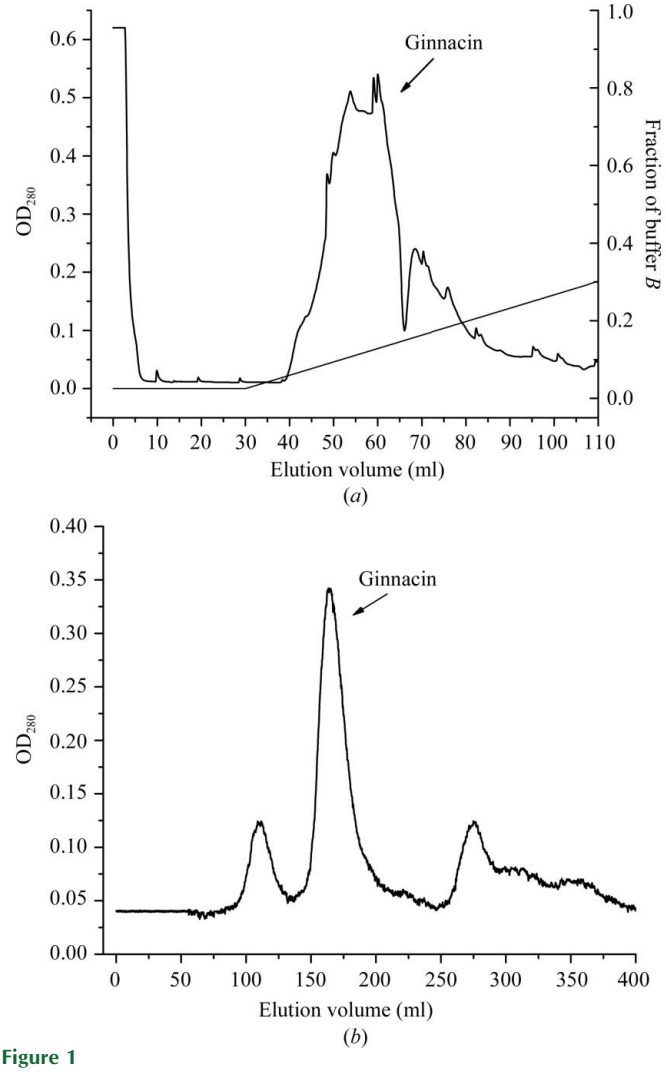


Figure 1
(*a*) Purification of ginnacin using an anion-exchange column. Dialyzed ginkgo-seed high-salt extract was loaded onto an 8 ml Source 15Q column pre-equilibrated with 10 m*M* Tris buffer. The bound protein was eluted with a linear 0–0.6 *M* NaCl gradient in 10 m*M* Tris buffer. Ginnacin eluted at ∼0.16 *M* NaCl. (*b*) Gel-filtration purification of ginnacin. Proteins in the ginnacin-containing fractions from anion-exchange purification were separated using a 300 ml Superdex-200 gel-filtration column. Ginnacin migrates at 171 ml, corresponding to a protein with a molecular weight of ∼280 kDa.

2.2. Crystallization

Purified ginnacin was concentrated to ∼15 mg ml^{-1} with Ultracel-30k filter devices (Millipore, Bedford, Massachusetts, USA) and the buffer was changed to H_2O by dilution and concentration using the same filter devices. A crystallization screen was set up at room temperature using Crystal Screen and Crystal Screen 2 (Hampton Research, Aliso Viejo, California, USA) with the hanging-drop vapor-diffusion method. For each condition, 1 µl protein solution was mixed with 1 µl crystallization solution and sealed against 1 ml crystallization solution in a well of a 24-well Linbro plate. Crystallization optimization was performed by varying the protein and precipitant concentrations. Cryoprotectants (sucrose, glycerol, PEG 400 and ethylene) were screened for conditions under which the crystals did not undergo visible morphological changes after immersion in a cryoprotectant solution for at least 5 min. Cryoprotectant solutions at final concentrations of 10–25%(*w*/*v*) or (*v*/*v*) were prepared by mixing the cryoprotectants in the reservoir solution.

2.3. X-ray diffraction experiments and crystal characterization

Single crystals were picked up in nylon loops, briefly immersed in a cryoprotectant solution and then flash-cooled in liquid nitrogen and subsequently stored in a liquid-nitrogen tank. X-ray data collection was performed using a MAR300 CCD detector on the SER-CAT 22ID beamline at the Advanced Photon Source (APS), Argonne National Laboratory. A typical data set consists of 150 1° frames collected at 1 Å wavelength with 1 s exposure. The diffraction data were processed using the *HKL*-2000 suite of programs (Otwinowski & Minor, 1997) and *XGEN* (Howard, 2000). The protein sequence of ginnacin is available in the protein database at NCBI (CAA53177). Molecular-replacement search models were created using *CHAIN-*

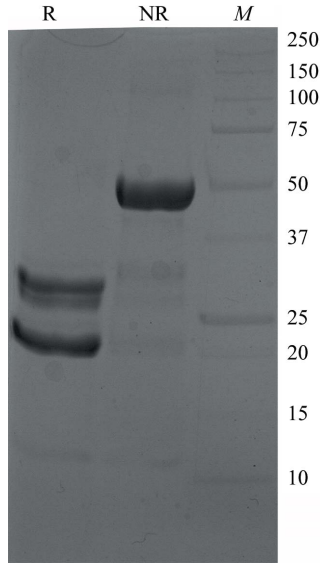


Figure 2
SDS–PAGE analysis of ginnacin. Reduced (lane R) and nonreducing (lane NR) samples were analyzed using a 4–25% SDS gel. The molecular weights of the protein bands in Precision Plus Protein Standards (lane *M*; kDa) are shown to the right of the gel.

SAW (Schwarzenbacher *et al.*, 2004) and the *CCP*4 interface (Potterton *et al.*, 2003) starting from the structure of a homologous protein and a sequence alignment compiled by *ClustalW* (Larkin *et al.*, 2007). Molecular-replacement calculations were performed with the program *Phaser* (McCoy *et al.*, 2005; Storoni *et al.*, 2004).

3. Results and discussion

3.1. Protein purification

As shown in Fig. 1, ginnacin can bind to the anion-exchange column. In the background of Tris–HCl buffer pH 7.9, it eluted at ∼160 m*M* NaCl (Fig. 1*a*). Ginnacin-containing fractions from the ion-exchange column were further purified by size exclusion (Fig. 1*b*). Ginnacin eluted at ∼171 ml from the Superdex 200 column, corresponding to a molecular weight of ∼280 kDa (Fig. 1*b*). As shown in Fig. 2, reduced SDS–PAGE indicated that there are two major peptides in the purified ginnacin. The apparent molecular weights of the two bands are ∼20 and ∼28 kDa, respectively. SDS–PAGE under nonreducing conditions mainly showed a single band with an apparent molecular weight of ∼49 kDa. This is consistent with ginnacin being a homohexamer with each monomer post-translationally processed into an N-terminal acidic chain of ∼28 kDa that is covalently linked by interchain disulfide bond(s) to a C-terminal basic chain of ∼20.1 kDa (Hager *et al.*, 1995). Further purification with a hydrophobic interaction column did not result in the removal of a noticeable amount of impurities.

3.2. Crystallization

Ginnacin purified from fresh ginkgo seeds readily crystallized under various conditions in the crystal screen within a week. The best diffracting crystals were obtained in 30% polyethylene glycol 400, 0.2 *M* magnesium chloride, 0.1 *M* HEPES pH 7.5. Typical long crystals with cross-section of 0.1 × 0.1 mm could be obtained in two weeks at room temperature (Fig. 3). Polyethylene glycol 400 at 30% concentration can also protect the crystal from freezing and no additional cryoprotectant was used. It is interesting to note that proteins purified using the same strategy from dried ginkgo seeds did not yield any crystals in screens with the same crystallization screening kits.

3.3. Data processing and molecular-replacement solution

The ginnacin crystals typically diffracted to 3.0 Å resolution, but crystal damage was obvious during data collection. The best data set obtained during our recent trip to APS was collected with 1 s exposure. The data were processed using the *HKL*-2000 suite of programs (Otwinowski & Minor, 1997) and *XGEN* (Howard, 2000), revealing a primitive orthorhombic crystal system with unit-cell parameters a = 148.20, b = 173.56, c = 99.64 Å. From systematic absences of specific reflections in the diffraction, the space group was determined to be $P2_12_12$ (Table 1). As the 11S seed storage proteins are known to form dimers of trimers (Mills *et al.*, 2002), based on the theoretical molecular weight of 51 kDa for ginnacin the asymmetric unit of the crystal is most likely to consist of six protomers, with a Matthews coefficient of 2.09 Å^3 Da^{-1} and a corresponding solvent content of 41.2%.

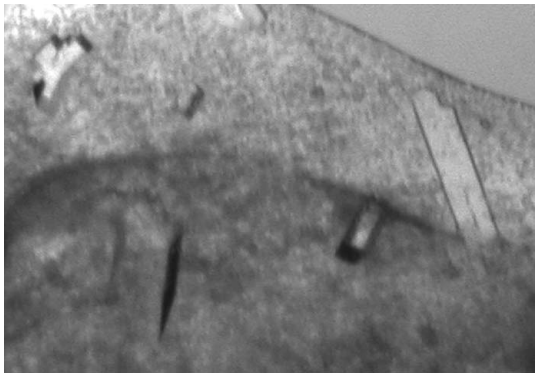


Figure 3
Single crystals of ginnacin obtained by vapor diffusion against 30% PEG 400, 0.2 *M* $MgCl_2$, 0.1 *M* HEPES pH 7.5.

Table 1
X-ray data-collection statistics.

Values in parentheses are for the outer shell.

Resolution (Å)	50–3.00 (3.11–3.00)
Wavelength (Å)	1.0
Data-collection temperature (K)	110
Space group	$P2_12_12$
Unit-cell parameters (Å)	
a	148.20
b	173.56
c	99.64
No. of observed reflections	417426
No. of unique reflections	51879
Redundancy	8.0 (7.9)
Completeness (%)	99.4 (99.0)
Mean $I/\sigma(I)$	12.8 (4.32)
$R_{\rm merge}$† (%)	11.4 (42.0)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

In order to solve the phase problem, molecular-replacement calculations were carried out using *Phaser* as a *CCP*4*i* module. Using the ginnacin sequence as the query, a *BLAST* search against the PDB database reported several hit structures, including a pumpkin seed globulin (PDB code 2evx) and soybean glycinins (PDB codes 1fxz and 1od5). Search models were generated with *CHAINSAW* as a *CCP*4*i* module with sequence-alignment files created using *ClustalW* and the corresponding hit PDB files as input files. Using these search models, however, molecular-replacement calculations did not result in a solution. A *Phaser* calculation starting with a *CHAINSAW* model of ginnacin using our nearly finished structure of Korean pine 7S vicilin (which had been refined to R and $R_{\rm free}$ values of 20.4% and 27.9%, respectively, to 2.22 Å resolution) as a template was successful. The log-likelihood gain is 3362 and the R factor and $R_{\rm free}$ after a few cycles of restrained refinement with *REFMAC* were 29.5% and 36.3%, respectively; the r.m.s.d. is 0.014 Å for bond lengths and 2.2° for bond angles. As expected, there are six ginnacin protomers in the asymmetric unit. The sequence identity between Loblolly pine vicilin and ginnacin is 21% over 437 amino acids. This is lower than the sequence identity between ginnacin and glycinin (37% over 476 amino acids) or between ginnacin and the pumpkin 11S seed globulin (39% over 441 amino acids). Currently, we are in the process of determining the sequence of Korean pine vicilin. It would be interesting to see whether the 11S seed storage protein of the 'living fossil' *G. biloba* is structurally more closely related to the 7S vicilin from Korean pine than it is to the 11S proteins from pumpkins and beans. Structural information on ginnacin may also shed light on the evolution of the cupin-protein families.

This work was supported by a fund from Illinois Institute of Technology. X-ray diffraction data were collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at

the APS, Argonne National Laboratory. Use of the APS was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract W-31-109-Eng-38.

References

Arahira, M. & Fukazawa, C. (1994). *Plant Mol. Biol.* **25**, 597–605.
Baumlein, H., Braun, H., Kakhovskaya, I. A. & Shutov, A. D. (1995). *J. Mol. Evol.* **41**, 1070–1075.
Gold, P. E., Cahill, L. & Wenk, G. L. (2002). *Psychol. Sci. Public Interest*, **3**, 2–11.
Gold, P. E., Cahill, L. & Wenk, G. L. (2003). *Sci. Am.* **288**, 86–91.
Hager, K. P., Braun, H., Czihal, A., Muller, B. & Baumlein, H. (1995). *J. Mol. Evol.* **41**, 457–466.
Howard, A. J. (2000). In *Crystallographic Computing*, vol. 7, edited by P. E. Bourne & K. D. Watenpaugh. Oxford University Press.
Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). *Bioinformatics*, **23**, 2947–2948.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Mills, E. N., Jenkins, J., Marigheto, N., Belton, P. S., Gunning, A. P. & Morris, V. J. (2002). *Biochem. Soc. Trans.* **30**, 925–929.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. (2003). *Acta Cryst.* D**59**, 1131–1137.
Schwarzenbacher, R., Godzik, A., Grzechnik, S. K. & Jaroszewski, L. (2004). *Acta Cryst.* D**60**, 1229–1236.
Shutov, A. D., Kakhovskaya, I. A., Braun, H., Baumlein, H. & Muntz, K. (1995). *J. Mol. Evol.* **41**, 1057–1069.
Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* D**60**, 432–438.