Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic studies of a novel antifungal protein with five disulfide bridges from *Eucommia ulmoides* Oliver

Ye Xiang,[a] Ren-Huai Huang,[a] Wei Liu,[a] Gen-Pei Li,[a,b] Xiao-Zhu Liu[c] and Da-Cheng Wang[a,b]*

[a]Center for Molecular Biology, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, People's Republic of China, [b]National Laboratory of Microgravity, Chinese Academy of Sciences, Beijing 100101, People's Republic of China, and [c]Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650204, People's Republic of China

Correspondence e-mail: dcwang@sun5.ibp.ac.cn

Two antifungal proteins, named *Eucommia* antifungal peptides 1 and 2 (EAFP1 and EAFP2), have been purified from the bark of the tree *E. ulmoides* Oliver and show a relatively broad spectrum of antifungal activity against several agriculturally important plant pathogens. One of these small proteins (EAFP2) has been crystallized. The crystal belongs to space group $P2_1$, with unit-cell parameters a = 19.01, b = 23.16, c = 30.69 Å, β = 98.54°. 1.0 Å resolution data were collected from an EAFP2 crystal and have been used to obtain phase information directly by an *ab initio* method.

1. Introduction

Plants have evolved highly effective defence mechanisms, in either a constitutive or an inducible manner, to protect them against pathogenic invasion. It has been found that a wide array of peptides and proteins are involved in these mechanisms (Garcia-Olmedo *et al.*, 1998; Broekaert *et al.*, 1997). The mode of action of these proteins is still much debated; therefore, it is essential to obtain three-dimensional structures of the respective antimicrobial proteins in order to understand their molecular mechanisms. The cysteine-rich peptides form a large group of antifungal proteins in which certain disulfide motifs stabilize the molecular scaffold. Accordingly, these plant antimicrobial proteins are usually classified by their distinct disulfide motif and their sequence homology (Broekaert *et al.*, 1997). Thionins (Orru *et al.*, 1997), plant defensins (Segura *et al.*, 1998) and hevein-like (Nielsen *et al.*, 1997) and knotin-like proteins (Gao *et al.*, 2001) belong to these classes. To date, a maximum of eight cysteines pairing into four disulfide bridges have been found in these small proteins.

Recently, two novel 41-residue antifungal peptides, named *Eucommia* antifungal peptides 1 and 2 (EAFP1 and EAFP2), have been purified from the bark of the tree *E. ulmoides* Oliver. This tree was originally found in China and has the common name hardy rubber tree or guttapercha tree. It is usually slender, up to 15 m in height, of open habit with furrowed grey bark. A bioassay showed that EAFPs have potent antifungal activity against a broad spectrum of chitin-containing and chitin-free fungi, including eight pathogenic fungi from cotton, wheat, potato, tomato and tobacco; their antifungal activity can be blocked by calcium ions. The amino-acid sequences of EAFP1 and EAFP2 are identical except for one residue at site 27. The most striking feature of these small proteins is the ten cysteines cross-linked to form five disulfide bridges with the pairing pattern C1–C5, C2–C9, C3–C6, C4–C7, C8–C10 (Fig. 1; Huang *et al.*, 2002). This is the first plant antifungal protein to be found with such a five-disulfide motif. Therefore, the three-dimensional structures of EAFPs become of great interest. As a first step, we report here

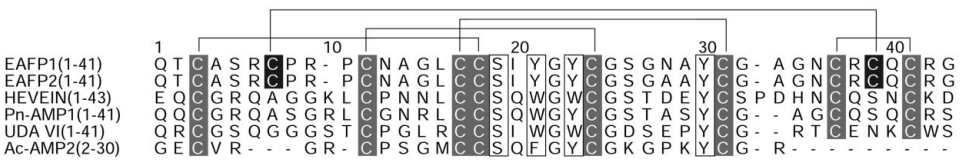

Figure 1
Sequences of EAFPs compared with those of hevein-like proteins. The pairing pattern of the distinct five-disulfide motif is indicated at the top. The conserved cysteines are highlighted in white on a grey background and the two additional cysteines of the EAFPs and the residues related to the activity are highlighted in white on a black background and in boxes, respectively. Hevein, rubber latex hevein (Broekaert *et al.*, 1990); Pn-AMP1, antimicrobial peptide from *Pharbitis nil* (Koo *et al.*, 1998); Ac-AMP2, antimicrobial peptide from *Amaranthus caudatus* (Broekaert *et al.*, 1992); UDA VI, *Urtica dioica* agglutinin (Harata & Muraki, 2000). The figure was prepared with *ALSCRIPT* (Barton, 1993).

the crystallization and preliminary crystallographic studies of this protein, including the initial solution by *ab initio* structure determination.

2. Experimental and results

2.1. Purification and characterization

Crude EAFPs were isolated from bark pieces from the tree *E. ulmoides* Oliver as described previously (Liu *et al.*, 1994). The crude proteins were further purified using HPLC with a reverse-phase column (YWG-PAK C18, Waters, Milfold, MA, USA). Two main fractions named EAFP1 and EAFP2, both with antifungal activity, were obtained.

MALDI–TOF mass-spectrometry analysis showed that both EAFP1 and EAFP2 have a high purity; their molecular weights are 4201.4 and 4158.9 Da, respectively. The pIs (isoelectric points) of EAFP1 and EAFP2 determined by isoelectrofocusing electrophoresis exceed 11. The full sequences of the EAFPs were elucidated by automated Edman degradation in combination with tandem mass spectra and C-terminal ladder sequencing analysis; the results showed a distinct five-disulfide motif (Fig. 1). Details of the sequencing and bioassay have been published elsewhere (Huang *et al.*, 2002).

2.2. Crystallization

EAFP1 and EAFP2 differ by only one residue. Since one of them, EAFP2, had a rather large yield from purification, it was chosen for crystallization experiments and further structure determination. EAFP2 was dissolved in water at a concentration of 20 mg ml^{-1} for crystallization. Initial screening was performed using the sparse-matrix method (Jancarik & Kim, 1991) with Hampton Research Crystal Screen I at 293 K. The hanging-drop vapour-diffusion method was used. Each drop contained equal amounts (1 µl) of protein (20 mg ml^{-1}) and reservoir solution and was equilibrated against 500 µl of reservoir solution in the well. In the initial screen, crystallites and small twinned crystals appeared from several conditions. Through optimization, diffraction-quality crystals were grown under the following conditions: drops formed by mixing equal volumes (3 µl) of 15 mg ml^{-1} protein in water and 1.0 *M* sodium acetate–acetic acid buffer pH 5.5 containing 5%(*v*/*v*) 2-propanol were equilibrated with the same buffer solution. The crystals appeared in one week and grew to their full size in three weeks. They are typically large thin plates with maximum dimensions of 0.6 × 0.5 × 0.06 mm (Fig. 2) and diffract to very high resolution. In the optimization procedure, we found that the growth of EAFP2 crystals was very sensitive to the concentration of sodium acetate. When this concentration was lowered even by 0.1 *M*, no crystals could be obtained. Other essential factors are the protein concentration and the addition of 2-propanol. At a protein concentration of 20 mg ml^{-1} or without 2-propanol, the crystals were badly twinned. When the protein concentration was changed to 15 mg ml^{-1} and 2-propanol was added, single crystals appeared in the drops. The pH seems to have little effect on the crystal growth, as crystals could be obtained over a wide range of pH values (4.5–7.5).

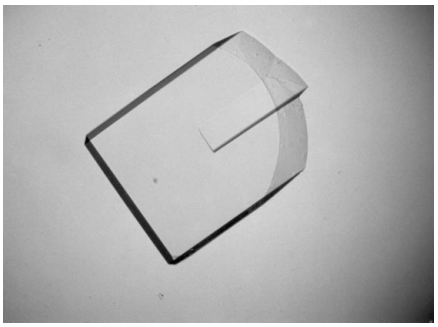

Figure 2
Large plate-like crystal of EAFP grown from sodium acetate as precipitant at 293 K with dimensions of 0.6 × 0.5 × 0.06 mm.

2.3. Data collection and crystallographic analysis

X-ray diffraction data from EAFP2 crystals were first collected at 1.66 Å resolution on a MAR 345 image-plate detector using Cu $K\alpha$ radiation (λ = 1.5418 Å) from an in-house generator operating at 40 kV and 50 mA. In order to use direct methods for the structural analysis, data collection to high resolution (1.0 Å) was then carried out using synchrotron radiation at beamline BL-18B of the Photon Factory, KEK, Tsukuba, Japan. The wavelength used was 1.0 Å, with the incident beam collimated to 0.1 mm in diameter in order to obtain high-resolution data. Data were collected on an ADSC Quantum 4R CCD detector at 283 K with a crystal-to-detector distance of 50 mm, $\Delta\varphi = 2^\circ$, λ = 1.00 Å and 60 s exposure time. A total of 105 frames were collected. All data were processed and scaled with the *HKL* program suite (Otwinowski & Minor, 1997). Details of the collected data are listed in Table 1.

Table 1
Data-collection statistics of EAFP2.

Values in parentheses refer to the highest resolution shells: 1.72–1.66 Å and 1.04–1.00 Å for data sets 1 and 2, respectively.

	Data set 1	Data set 2
Unit-cell parameters		
a (Å)	19.00	19.01
b (Å)	23.17	23.16
c (Å)	30.56	30.69
β (°)	98.00	98.54
Source	Cu $K\alpha$	BL18B (PF)
Space group	$P2_1$	$P2_1$
Wavelength (Å)	1.54	1.00
Resolution (Å)	30.00–1.66	30.00–1.00
No. of observations	10062	66812
No. of unique reflections	2804	13659
Completeness (%)	87.4 (66.5)	94.4 (90.3)
R_{sym} (%)	4.7 (10.9)	5.3 (19.8)
$I/\sigma(I)$	24.8 (10.7)	25.2 (6.0)

Preliminary analysis on the 1.0 Å data showed that the crystal belongs to space group $P2_1$. The unit-cell parameters are a = 19.01, b = 23.16, c = 30.69 Å, $\alpha = \gamma = 90^\circ$, β = 98.54°. The merged 1.66 Å data set consists of 2804 unique reflections with a completeness of 87.4% and an R_{sym} of 4.7%, while the 1.0 Å data set consists of 13 659 unique reflections with a completeness of 94.4% and an R_{sym} of 5.3% (19.8% for the last shell, 1.04–1.00 Å). Assuming one molecule to be present in the asymmetric unit, the Matthews coefficient V_M (Matthews, 1968) is estimated as 1.64 Å^3 Da^{-1}, with a corresponding solvent content of 21.4%, which is well within the range found in protein crystals.

2.4. Preliminary structure analysis by direct methods

As the protein is small (41 residues) with high-resolution data (1.0 Å), we tried to use an *ab initio* method to determine its struc-

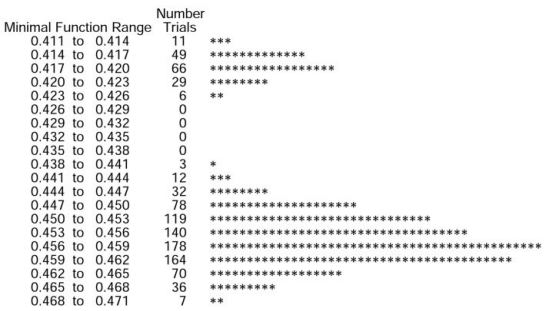

Figure 3
Histogram of the *SnB* results showing the number of trials *versus* the minimal function value (R_{min}). The distribution of R_{min} is clearly bimodal and separates about 150 possible solutions from non-solutions.

ture. Preliminary results showed this to be successful. We used the *Shake-and-Bake* dual-space direct-method procedure (Weeks *et al.*, 1994) programmed in *SnB* (Weeks & Miller, 1999). The normalized structure magnitudes (E values) of the 1.0 Å resolution data were calculated using the program *DREAR* (Blessing & Smith, 1999) built into the *SnB* program. 1000 trial structures were generated and subjected to 500 cycles of reciprocal-space phase refinement and real-space density modification. It took about 36 h for *SnB* to complete the computation on an Intel P4 1.7 Linux PC. The final minimal function value ($R_{\rm min}$) distribution of these trials was clearly bimodal (Fig. 3) and the minimal $R_{\rm min}$ of these trials was 0.424, with a correlation coefficient (CC) value of 0.68. These results suggested that a correct solution had been obtained. This was further confirmed by checking the results with the *SnB* graphic display, in which five disulfide bonds could be defined. Some of the side chains and part of the backbone also appeared clearly in the graphics window. Further model building and structure refinement are now in process.

This work was supported by NNSFC (30070162), NKBRSP (G1999075604) and the State Basic Research Program of China. We would like to thank KEK for providing beam time and Dr N. Sakabe and Dr M. Suzuki for their kind help during data collection at BL18B of the Photon Factory, KEK, Japan.

References

Barton, G. J. (1993). *Protein Eng.* **6**, 37–40.
Blessing, R. H. & Smith, G. D. (1999). *J. Appl. Cryst.* **32**, 664–670.
Broekaert, W. F., Cammue, B. P. A., De Bolle, M. F. C., Thevissen, K., De Samblanx, G. W. & Osborn, R. W. (1997). *Crit. Rev. Plant Sci.* **16**, 297–323.
Broekaert, W., Lee, H., Kush, A., Chua, N. & Raikhel, N. (1990). *Proc. Natl Acad. Sci. USA*, **87**, 7633–7637.
Broekaert, W. F., Marien, W., Terreas, F. R. G., De Bolle, M. F. C., Proost, P., Van Damme, J., Dillen, L., Claeys, M., Rees, S. B., Vanderleyden, J. & Cammue, B. P. A. (1992). *Biochemistry*, **31**, 4308–4314.
Gao, G. H., Liu, W., Dai, J. X., Wang, J. F., Hu, Z., Zhang, Y. & Wang, D. C. (2001). *Biochemistry*, **40**, 10973–10978.
Garcia-Olmedo, F., Molina, A., Alamillo, J. M. & Rodriguez-Palenzuela, P. (1998). *Biopolymers*, **47**, 479–491.
Harata, K. & Muraki, M. (2000). *J. Mol. Biol.* **297**, 673–681.
Huang, R. H., Xiang, Y., Liu, X. Z., Zhang, Y., Hu, Z. & Wang, D. C. (2002). *FEBS Lett.* **521**, 87–90.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Koo, J. C., Lee, S. Y., Chun, H. J., Cheong, Y. H., Choi, J. S., Kawabata, S., Miyagi, M., Tsunasawa, S., Ha, K. S., Bae, D., Han, C., Lee, B. L. & Cho, M. J. (1998). *Biochim. Biophys. Acta*, **1382**, 80–90.
Liu, X. Z., Hu, Z., Li, Y., Yang, J. B. & Li, B. J. (1994). *Acta Bot. Yunnanica*, **16**, 385–391.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Nielsen, K. K., Nielsen, J. E., Madrid, S. M. & Mikkelsen, J. D. (1997). *Plant Physiol.* **113**, 83–89.
Orru, S., Scaloni, A., Gianneattasio, M., Urech, K., Pucci, P. & Schaller, G. (1997). *Biol. Chem.* **378**, 989–996.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Segura, A., Moreno, M., Molina, A. & Garcia-Olmedo, F. (1998). *FEBS Lett.* **435**, 159–162.
Weeks, C. M., DeTitta, G. T., Hauptman, H. A., Thuman, P. & Miller, R. (1994). *Acta Cryst.* A**50**, 210–220.
Weeks, C. M. & Miller, R. (1999). *J. Appl. Cryst.* **32**, 120–124.