Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of lectin C from the roots of pokeweed (*Phytolacca americana*)

Minoru Hayashida,[a] Tomomi Fujii,[a] Mika Hamasu,[b] Masatsune Ishiguro[c]† and Yasuo Hata[a]*

[a]Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan, [b]Laboratory of Protein Chemistry and Engineering, Department of Genetic Resources Technology, Faculty of Agriculture, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan, and [c]Laboratory of Protein Chemistry and Engineering, Graduate School of Bioresources and Bioenvironmental Sciences, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

† Present address: Department of Food and Nutrition, Faculty of Food and Nutrition, Beppu University, Beppu, Ooita 874-8501, Japan.

Correspondence e-mail: hata@scl.kyoto-u.ac.jp

Received 21 December 2002
Accepted 17 April 2003

Lectin C from the roots of pokeweed (*Phytolacca americana*) (PL-C; 13 747 Da, 126 amino-acid residues), which consists of three chitin-binding domains, was initially crystallized in two crystal forms. One form, obtained in the presence of 30%(*w*/*v*) PEG 4000, belongs to the tetragonal system. The other, obtained in the presence of 2.0 *M* ammonium sulfate, belongs to the rhombohedral system. Statistical analyses of the X-ray diffraction intensities showed that both crystals were twinned. Single crystals suitable for a diffraction experiment were obtained by the addition of 0.5%(*v*/*v*) dioxane to the latter precipitant solution. They belong to space group *R*3, with unit-cell parameters $a = b = 104.1$, $c = 69.7$ Å, and diffract X-rays to 1.8 Å resolution. A heavy-atom derivative crystal has been obtained and structure determination is presently ongoing using the SIRAS method.

1. Introduction

Lectins are carbohydrate-binding proteins that cause cell agglutination and are widely distributed in nature. Binding of lectins to glycoconjugates on cell surfaces has been shown to be involved in a variety of biological processes (Brandley & Schnaar, 1986; Hadden, 1988; Sairam, 1989; Kooijman *et al.*, 1989; Perillo *et al.*, 1995). The most remarkable functional feature of lectins is to induce the proliferation and maturation of resting lymphocytes, which is termed mitogenic stimulation and is key to the immune response to the antigen.

Pokeweed lectin is known to be a mitogenic lectin that exceptionally activates both T and B cells, in contrast to almost all other mitogenic lectins which only stimulate T cells (Borrebaeck & Carlsson, 1989; Di Sabato *et al.*, 1987). The roots of pokeweed contain five lectins (PLs), designated PL-A, PL-B, PL-C, PL-D1 and PL-D2. Amino-acid sequence analyses have shown that PLs are composed of several chitin-binding domains comprising about 40 residues, which are homologous to one another in sequence and have conserved positions of four disulfide bridges and three putative aromatic residues for carbohydrate binding (Yamaguchi *et al.*, 1995, 1996, 1997). PL-A (22 kDa) is an N-terminal half-fragment of PL-B. PL-B (35 kDa) and PL-C (13 kDa) are novel lectins consisting of seven and three chitin-binding domains, respectively, in contrast to almost all other lectins which contain even numbers of chitin-binding domains (usually two or four). PL-D1 (9 kDa) comprises two chitin-binding domains. PL-D2 (9 kDa) is exactly the same as PL-D1 in amino-acid sequence except for the lack of two C-terminal residues, Leu and Thr. Thus, PLs are characterized by two, three or seven chitin-binding domains. This characteristic suggests that PLs have evolved by a different mechanism from 'the subsequent duplication mechanism' proposed for *Urtica dioica* agglutinin and wheat germ agglutinin (Wright *et al.*, 1991; Yamaguchi *et al.*, 1997). In the subsequent duplication mechanism, a single-domain protein, the ancestor of a multicomponent protein, first evolves into a two-domain intermediate, which subsequently evolves into a four-domain protein. Of the five pokeweed lectins, PL-B is the glycoprotein with the most potent haemagglutination and mitogenic activities. PL-C has mitogenic activity but does not have haemagglutination activity. PL-D2 has both haemagglutination and mitogenic activities, while PL-D1 has haemagglutination activity but does not have mitogenic activity. In this way, the PLs have acquired various physiological properties, such as haemagglutination, mitogenicity and carbohydrate-binding specificity, by changing the number of domains in the polypeptide chain in the evolutionary process.

In order to understand the relationships between the domain-repeat structures and their physiological properties, we have carried out a series of crystal structure determinations

of PLs. Here, we report the characterization of the crystals of PL-C, which is a novel chitin-binding protein consisting of three repeat domains.

2. Materials and methods

2.1. Crystallization

PL-C was purified to homogeneity and lyophilized as described previously (Kino *et al.*, 1995). The lyophilized protein was dissolved in 100 m*M* Tris–HCl buffer pH 7.5 to a final concentration of 10 mg ml^{-1}. Crystallization was performed using the hanging-drop vapour-diffusion method at 298 K. Droplets were prepared by mixing 1 µl of protein solution and 1 µl of precipitant solution. Two crystal forms were obtained by equilibrating drops of the protein solution against 500 µl reservoir solutions containing different precipitants. Crystals belonging to the tetragonal system (Fig. 1*a*), with dimensions of 0.2 × 0.2 × 0.2 mm, were grown in 5 d in 100 m*M* Tris–HCl buffer pH 8.5 containing 30%(*w*/*v*) PEG 4000 and 200 m*M* magnesium chloride. Crystals belonging to the rhombohedral system, with dimensions of 0.2 × 0.2 × 0.2 mm, were grown in one week in 100 m*M* sodium acetate pH 4.6 containing 2.0 *M* ammonium sulfate. These two crystal forms were twinned, as described later.

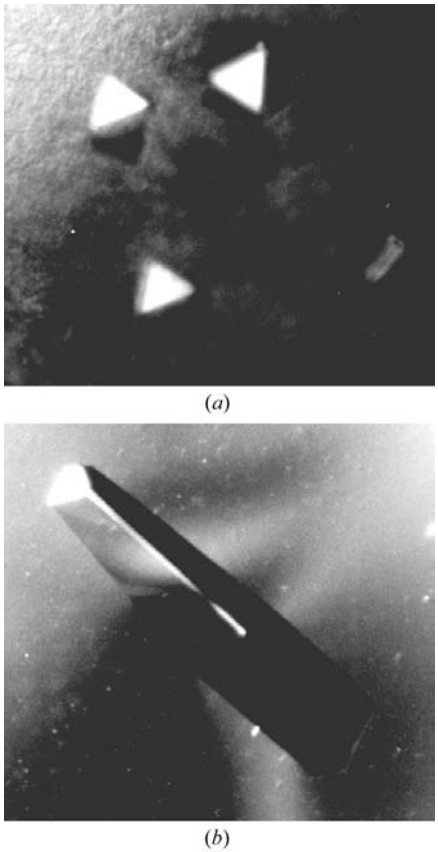

Figure 1
Crystals of PL-C. (*a*) Tetragonal twinned crystal with approximate dimensions of 0.2 × 0.2 × 0.2 mm and (*b*) rhombohedral single crystal with approximate dimensions of 0.3 × 0.3 × 1.0 mm.

Attempts to further search for crystallization conditions were made in order to circumvent the twinning problem entirely. Single crystals suitable for X-ray diffraction were obtained by the addition of 0.5%(*v*/*v*) dioxane to the crystallization solution for the rhombohedral crystal (Fig. 1*b*). Large single crystals, with dimensions of 0.3 × 0.3 × 1.0 mm, were obtained in two weeks by the repeated microseeding method, in which a slight amount of the precipitant solution containing only a few microseeds was added to 1 µl of the protein solution just prior to preparing the hanging drop. Typical crystals diffracted X-rays to 1.8 Å resolution.

2.2. X-ray data collection

Diffraction data from the tetragonal and rhombohedral twinned crystals were collected on an R-AXIS IIc imaging-plate system using Cu $K\alpha$ radiation from an RU-300 X-ray generator. For the tetragonal crystal, the crystal-to-detector distance was set to 95 mm and each oscillation frame was taken with a 2.0° oscillation range. For the rhombohedral crystal, each frame was taken with a 1.5° oscillation range at a crystal-to-detector distance of 95 mm. The collected intensities were processed with *DENZO* and scaled with *SCALEPACK* (Otwinowski & Minor, 1997). Diffraction data from the rhombohedral single crystal were collected using a Quantum-4 CCD detector (ADSC) at the BL-18B station (λ = 1.000 Å) of the Photon Factory (PF), National Laboratory of High Energy Physics, Tsukuba, Japan. The crystal-to-detector distance was set to 115 mm and each oscillation frame was taken with a 1.5° oscillation range for 50 s. The intensities were integrated with *MOSFLM* (Leslie, 1992) and scaled with *SCALA* (Evans, 1993) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). The structure-factor amplitudes for all data sets were produced with *TRUNCATE* (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

3.1. Detection of twinning

Details of the data-collection statistics for the twinned crystals are summarized in Table 1. The tetragonal crystal belongs to space group $P4_1$ or $P4_3$, with unit-cell parameters $a = b = 47.9$, $c = 72.2$ Å. The 2.8 Å resolution data set of 3707 independent reflections was collected with an $R_{\rm merge}$ of 5.4% (the completeness was 90.6%). The presence of one molecule in the asymmetric unit corresponds to a reasonable Matthews coefficient (Matthews, 1968) of 3.01 Å^3 Da^{-1} and a solvent content of approximately 60%. However, the self-rotation function in space group $P4_1$ or $P4_3$ for this crystal showed the presence of a twofold axis along [110], which runs along the diagonal between the *a* and *b* axes. Assuming that the asymmetric unit of the crystal contains two molecules, the Matthews coefficient is calculated to be 1.50 Å^3 Da^{-1}, which is extremely low. Such a phenomenon may be observed as a possible symptom of perfect twinning (Yeates, 1997). For protein crystals, hemihedral twinning is the most common. In this twinning, the crystal consists of two distinct components which are oriented so that their diffraction patterns exactly overlap. Consequently, two twin-related reflections independently contribute to each measured diffraction intensity and to peaks in the Patterson map. Thus, the twofold symmetry observed in the self-rotation function seemed to correspond to the twin operator. In the case of the tetragonal symmetry, the reflections of indices (hkl) and $(kh\bar{l})$ are related by the twin operator.

Table 1
Data-collection statistics for twinned crystals of PL-C.

Values in parentheses refer to the highest resolution shell.

Data set	Tetragonal	Rhombohedral
X-ray source	RU-300	RU-300
Wavelength (Å)	1.542	1.542
Resolution limit (Å)	2.80 (2.90–2.80)	2.00 (2.07–2.00)
No. of observations	32067	146035
No. of unique reflections	3707	18134
Redundancy	8.7	8.1
Completeness (%)	90.6 (69.3)	94.6 (85.0)
$R_{\rm merge}$† (%)	5.4 (27.6)	12.8 (29.4)
$I/\sigma(I)$	11.3 (1.7)	7.9 (2.0)

† $R_{\rm merge} = \sum_i |I_i - \langle I_i\rangle| / \sum_i I_i$, where $\langle I_i\rangle$ is the average of I_i over all symmetry equivalents.

It is possible to detect and characterize twinning by analyzing the intensity statistics. A cumulative distribution function $N(Z)$ gives the fraction of reflections having an intensity less than or equal to Z, where Z represents the intensity relative to the mean intensity ($Z = I/\langle I\rangle$) (Howells *et al.*, 1950). For a twinned crystal, fewer reflections are observed with extremely high or low intensity and the distribution curve for acentric reflections has a sigmoidal shape. The $N(Z)$ plot for the tetragonal crystal revealed that

it was twinned (Fig. 2*a*). Therefore, the twin fraction (α) was estimated according to the Britton plot (Britton, 1972; Fisher & Sweet, 1980). The twin fraction is the fractional volume of the minor twin component with respect to the whole crystal and hence $0.0 < \alpha < 0.5$. The Britton plot for the tetragonal crystal showed that the twin fraction was approximately 0.40 (Fig. 2*b*).

The rhombohedral crystal belongs to space group *R*3, with unit-cell parameters $a = b = 104.1$, $c = 69.7$ Å (in the hexagonal setting). A 2.0 Å resolution data set of 18 134 independent reflections was collected with an $R_{\rm merge}$ of 12.8% (the completeness was 94.6%) (Table 2). The high value of $R_{\rm merge}$ may arise from inherent characteristics of the present rhombohedral twinned crystal, because the value is slightly high even at low resolution, although the signal-to-noise ratio is reasonable. Assuming the presence of two molecules in the asymmetric unit, the $V_{\rm M}$ value is calculated to be 2.64 Å^3 Da^{-1}, with a solvent content of 53%. The self-rotation function for this crystal showed the presence of a twofold axis ($\theta = 90.00$, $\varphi = -64.07$, $\chi = 180.00^\circ$). Therefore, it seemed reasonable that the asymmetric unit contained two molecules related by a twofold axis. However, the $N(Z)$ plot for the rhombohedral crystal indicated that twinning was likely (Fig. 3*a*), as was the case with the tetragonal crystal. The α value was estimated to be 0.20 using the Britton plot (Fig. 3*b*). From these results, we concluded that the tetragonal and rhombohedral crystals were heavily twinned and that application of detwinning was not suitable for the present twinned data because of the high twin fraction.

The crystals obtained by the addition of dioxane, on the other hand, were shown to be single by the $N(Z)$ plot and Britton plot.

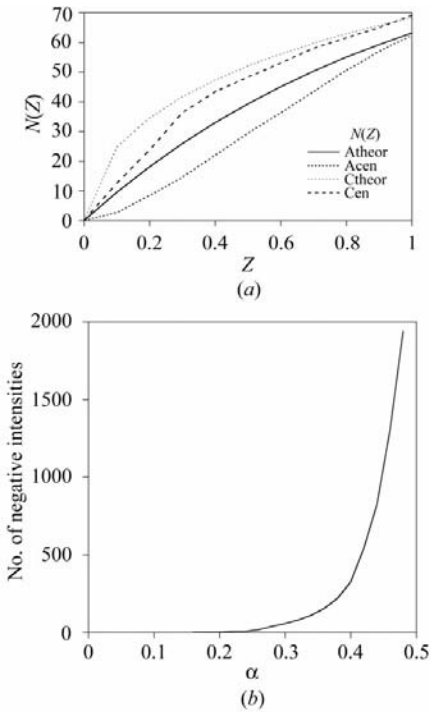

Figure 2
Analysis of the intensity statistics for the tetragonal twinned crystal. (*a*) A graphical representation of $N(Z)$ curves against Z for acentric data. The cumulative distribution function $N(Z)$ gives the fraction of reflections having an intensity less than Z ($Z = I/\langle I\rangle$, where I is the observed intensity and $\langle I\rangle$ is the mean intensity). Atheor, theoretical distribution for acentric reflections; Acen, observed distribution for acentric reflections; Ctheor, theoretical distribution for centric reflections; Cen, observed distribution for centric reflections. The figure was produced with *TRUNCATE* (Collaborative Computational Project, Number 4, 1994). (*b*) Britton plot (Britton, 1972; Fisher & Sweet, 1980) showing the number of negative intensities as a function of the assumed twin fraction. The figure was produced with *DETWIN* (Collaborative Computational Project, Number 4, 1994).

3.2. Quality of the single crystal and preparation of a heavy-atom derivative

The rhombohedral single crystal has the same crystallographic parameters as those of the rhombohedral twinned crystal. The crystal contains two molecules in the asymmetric unit. The diffraction data were collected at 1.8 Å resolution. The data set had an $R_{\rm merge}$ of 6.70% for 26 091 independent reflections (the completeness was 99.9%) derived from 134 896 observations. In the highest resolution shell (1.89–1.80 Å), the data set was 99.9% complete, with a signal-to-noise ratio [$I/\sigma(I)$] of 2.4. Details of the data-collection statistics are summarized in Table 2.

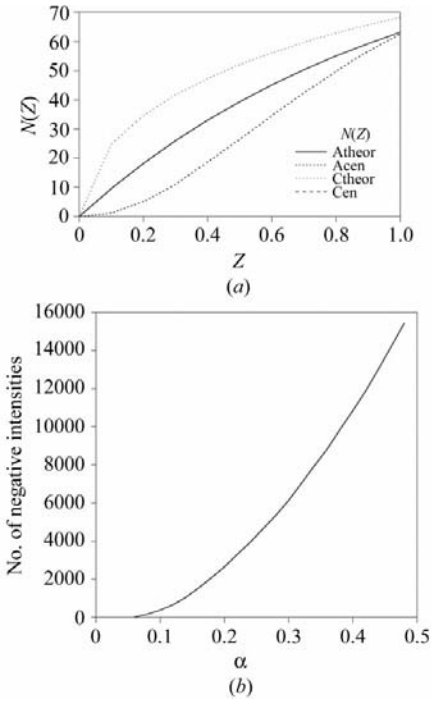

Figure 3
Analysis of the intensity statistics for the rhombohedral twinned crystal. (*a*) Cumulative intensity distribution curves. The crystal belongs to space group *R*3; hence there are no observed centric reflections. (*b*) Britton plot.

Table 2
Crystal preparation and data-collection statistics of PL-C single crystal.

Values in parentheses refer to the highest resolution shell.

Data set	Native	Hg
Soaking conditions		
Reagent	—	$HgCl_2$
Concentration (m*M*)	—	1
Time (h)	—	5
Data collection		
X-ray source	PF (BL-18B)	PF (BL-18B)
Wavelength (Å)	1.000	1.000
Resolution limit (Å)	1.80 (1.89–1.80)	3.00 (3.16–3.00)
No. of observations	134896	26694
No. of unique reflections	26091	5680
Redundancy	5.2	4.7
Completeness (%)	99.9 (99.9)	99.7 (97.9)
$R_{\rm merge}$† (%)	6.7 (27.9)	6.5 (11.8)
$R_{\rm iso}$‡ (%)	—	14.0
$I/\sigma(I)$	7.1 (2.4)	9.4 (5.3)

† $R_{\rm merge} = \sum_i |I_i - \langle I_i\rangle| / \sum_i I_i$, where $\langle I_i\rangle$ is the average of I_i over all symmetry equivalents. ‡ $R_{\rm iso} = \sum \big||F_{PH}| - |F_P|\big| / \sum |F_P|$, where $|F_P|$ and $|F_{PH}|$ are the native and derivative structure-factor amplitudes, respectively.

For the structure solution of PL-C using the rhombohedral single crystal, an isomorphous heavy-atom derivative crystal was prepared by soaking the native crystal in a reservoir solution containing 1 m*M* mercury chloride for 5 h at room temperature. Intensity data from the derivative crystal were collected to 3.0 Å resolution on the BL-18B station of PF, Tsukuba, Japan, in the same manner as for the data collection from the native crystal (Table 1). An isomorphous difference Patterson map was calculated with the program *FFT* from the *CCP4* suite (Collaborative Computational Project, Number 4, 1994). This map revealed the

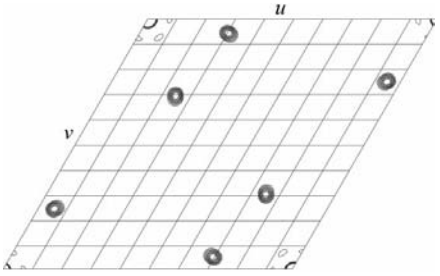

Figure 4
Harker section ($w = 0$) of the difference Patterson map computed using 10–5 Å resolution data from the native and mercury-derivative crystals. The contouring starts from 2σ at intervals of 1σ.

presence of a single Hg-atom position in the asymmetric unit of the derivative crystal (Fig. 4). The structure determination of PL-C is under way using the SIRAS method. Concurrently, preparation of another isomorphous derivative crystal has been tried using a soaking method.

We would like to thank Drs N. Sakabe, M. Suzuki and N. Igarashi of the Photon Factory for their kind help with the synchrotron experiments (Proposal No. 2000G313). This work was supported in part by the National Project on Protein Structural and Functional Analyses and Grants-in-Aid for Scientific Research (13033020 and 14560065 to YH) from the Ministry of Education, Culture, Sports, Science and Technology of Japan and by the Rice Genome Project (PR-2202 to YH) of the Ministry of Agriculture, Forestry and Fisheries of Japan. YH is a member of the Structural Biology Sakabe Project.

References

Borrebaeck, C. A. K. & Carlsson, R. (1989). *Adv. Lectin Res.* **2**, 10–27.
Brandley, B. K. & Schnaar, R. L. (1986). *J. Leukoc. Biol.* **40**, 97–111.
Britton, D. (1972). *Acta Cryst.* A**28**, 296–297.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Di Sabato, G., Hall, J. M. & Thompson, L. (1987). *Methods Enzymol.* **150**, 3–17.
Evans, P. R. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 114–122. Warrington: Daresbury Laboratory.
Fisher, R. G. & Sweet, R. M. (1980). *Acta Cryst.* A**36**, 755–760.
Hadden, J. W. (1988). *Mol. Immunol.* **25**, 1105–1112.
Howells, E. R., Phillips, D. C. & Rogers, D. (1950). *Acta Cryst.* **3**, 210–214.
Kino, M., Yamaguchi, K., Umekawa, H. & Funatsu, G. (1995). *Biosci. Biotech. Biochem.* **59**, 683–688.
Kooijman, R., de Wildt, P., Beumer, S., van der Vliet, G., Homan, W., Kalshoven, H., Musgrave, A. & van den Ende, H. (1989). *J. Cell Biol.* **109**, 1677–1687.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACMB Newsl. Protein Crystallogr.* **26**.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods. Enzymol.* **267**, 307–326.
Perillo, N. L., Pace, K. E., Seilhamer, J. J. & Baum, L. G. (1995). *Nature (London)*, **378**, 736–739.
Sairam, M. R. (1989). *FASEB J.* **3**, 1915–1926.
Wright, H. T., Sandrasegaram, G. & Wright, C. S. (1991). *J. Mol. Evol.* **33**, 283–294.
Yamaguchi, K., Mori, A. & Funatsu, G. (1995). *Biosci. Biotech. Biochem.* **59**, 1384–1385.
Yamaguchi, K., Mori, A. & Funatsu, G. (1996). *Biosci. Biotech. Biochem.* **60**, 1380–1382.
Yamaguchi, K., Yurino, N., Kino, M., Ishiguro, M. & Funatsu, G. (1997). *Biosci. Biotech. Biochem.* **61**, 690–698.
Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.