Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic study of an invertebrate C-type lectin, CEL-I, from the marine invertebrate *Cucumaria echinata*

Tomomitsu Hatakeyama,[a]* Noriaki Matsuo,[a] Haruhiko Aoyagi,[a] Hajime Sugawara,[b,c]† Tatsuya Uchida,[b] Genji Kurisu[b] and Masami Kusunoki[b]

[a]Department of Applied Chemistry, Faculty of Engineering, Nagasaki University, 1-14 Bunkyo-machi, Nagasaki 852-8521, Japan, [b]Research Center for Structural Biology, Institute for Protein Research, Osaka University, 3-2 Yamada-oka, Suita, Osaka 565-0871, Japan, and [c]Computational Science and Technology Division, Department of Advanced Databases, Japan Science and Technology Corporation, Honcho 4-chome, Kawaguchi, Saitama 332-0012, Japan

† Current address: Metabolic Function Research Group, RIKEN Plant Science Center, Hirosawa 2-1, Wako, Saitama 351-0198, Japan.

Correspondence e-mail: thata@net.nagasaki-u.ac.jp

Received 30 May 2001
Accepted 15 October 2001

CEL-I is a GalNAc-specific carbohydrate-binding protein (lectin) isolated from the sea cucumber *Cucumaria echinata*. This protein belongs to the widely distributed C-type lectin family of animal lectins, which require Ca^{2+} for their carbohydrate-binding ability and play important roles in various molecular-recognition processes in organisms. CEL-I was crystallized with 2-methyl-2,4-pentanediol using the hanging-drop vapour-diffusion technique. The CEL-I crystals belong to the monoclinic space group $C2$, with unit-cell parameters $a = 92.38$ (3), $b = 69.94$ (3), $c = 76.69$ (3) Å, $\beta = 136.46$ (2)°. Diffraction data were collected to 2.0 Å resolution using synchrotron radiation. The asymmetric unit contains one CEL-I molecule.

1. Introduction

Lectins are a group of carbohydrate-binding proteins originally isolated from plant seeds. They recognize various types of carbohydrate chains on the cell surface and exhibit cell-agglutinating activity through cross-linking the carbohydrate chains. In addition to agglutination, lectins induce several cellular events, such as proliferation and apoptosis, by stimulating the intracellular signal transduction pathway. Various lectins have also been found in animal tissues and body fluids over the past 10–20 y (Gabius, 1997). Of these, the C-type lectins are so named because of their Ca^{2+}-dependent carbohydrate-binding properties (Drickamer, 1988). C-type lectins constitute one of the major animal lectin families and they have been shown to play important roles in molecular-recognition processes. The C-type lectins in vertebrates fall into seven categories (Day, 1994) and in many cases have multi-domain structures which include C-type carbohydrate-recognition domains (CRDs). They function *via* the binding of CRDs to specific carbohydrate chains on other molecules or the cell surface. In contrast, invertebrate C-type lectins mostly occur as single-domain proteins consisting almost solely of CRDs. Studies have revealed that they play important roles in the host-defence system. Binding to the surface carbohydrates on the invading microorganisms leads to their agglutination and the activation of phagocytes such as macrophages (opsonization). Although C-type CRDs show relatively low homology, tertiary structures indicate that their basic folds are common. In the marine invertebrate *C. echinata* (Holothuroidea), there are four Ca^{2+}-dependent galactose/*N*-acetylgalactosamine (GalNAc) specific lectins (CEL-I, II, III and IV; Hatakeyama *et al.*, 1994). Of these lectins, CEL-IV (Hatakeyama, Ohuchi *et al.*, 1995) and CEL-I (unpublished results) have already been found to belong to the C-type lectin family from their amino-acid sequences, while CEL-III is not a C-type lectin but a novel haemolytic lectin possessing partial homology to the toxic plant lectins ricin and abrin (Nakano *et al.*, 1999). Since a tunicate (*Polyandrocarpa misakiensis*) lectin is the only invertebrate C-type lectin whose tertiary structure has been determined so far, the X-ray crystallographic study of *C. echinata* lectins should deepen our understanding of the carbohydrate-recognition mechanism of invertebrate C-type lectins. CEL-I is especially interesting as this lectin exhibits a much higher specificity for *N*-acetylgalactosamine than the other lectins. CEL-I is composed of two identical subunits with a molecular mass of 16 kDa which are linked by a single disulfide bond. Here, we describe the crystallization and preliminary X-ray diffraction study of CEL-I.

2. Purification of CEL-I

CEL-I was purified from *C. echinata*, which was collected in the sea of Genkai (Fukuoka, Japan). The proteins extracted from the homogenate of *C. echinata* were applied to a lactosyl-Cellulofine column (Teichberg *et al.*, 1988) equilibrated with 0.15 *M* NaCl, 10 m*M* Tris–HCl pH 7.5 (TBS) containing 10 m*M* $CaCl_2$. Adsorbed lectins (CEL-I, CEL-III and

CEL-IV) were eluted with TBS containing 20 m*M* EDTA. The lectins were then separated by a GalNAc-Cellulofine column, utilizing their difference in carbohydrate-binding specificity (Hatakeyama, Nagatomo *et al.*, 1995). After the elution of CEL-III with TBS containing 0.1 *M* lactose, CEL-I and CEL-IV were eluted with TBS containing 20 m*M* EDTA. CEL-I and CEL-IV were finally separated by gel filtration on Sephadex G-75 in TBS.

3. Crystallization of CEL-I

Crystallization of CEL-I was carried out using the hanging-drop vapour-diffusion method. The initial screening for crystallization used screening kits (Crystal Screens 1 and 2, Hampton Research). 2 µl of protein solution (5 mg ml^{-1}) with TBS containing 10 m*M* $CaCl_2$ was mixed with an equal volume of reservoir solution and was allowed to equilibrate against 0.5 ml of reservoir solution at 293 K. Rod-shaped microcrystals were obtained using a reservoir solution with pH 8.5 containing 50%(*v*/*v*) 2-methyl-2,4-pentanediol (MPD) and 0.2 *M* ammonium phosphate. The condition was further optimized with 50–70%(*v*/*v*) MPD, 0.1 *M* Tris–HCl pH 8.0. Clusters of small thin plate-like crystals were first obtained after a week using 65% MPD. Crystal size and quality were further improved with a combination of microseeding and macroseeding techniques (Stura & Wilson, 1992). As a result, single crystals with maximum dimensions of 0.5 × 0.2 × 0.07 mm were obtained in sitting drops in 2–3 weeks using 56%(*v*/*v*) MPD, 10 m*M* $CaCl_2$, 0.1 *M* Tris–HCl pH 8.0 (Fig. 1).

4. Data collection and processing

Crystals of CEL-I were mounted in a nylon CryoLoop (Hampton Research) and placed directly into a nitrogen stream at 100 K. Data collection from the CEL-I crystals was performed on beamline BL18B of the Photon Factory at the High Energy Acceleration Research Organization, Tsukuba, Japan using an ADSC Quantum 4R CCD camera (Watanabe *et al.*, 1995). Image data were processed with the programs *DPS* (Rossmann & van Beek, 1999) and *SCALA* (Collaborative Computational Project, Number 4, 1994). Diffraction data statistics are summarized in Table 1. The space group was determined to be monoclinic *C*2, with unit-cell parameters (standard deviations in parentheses) a = 92.38 (3), b = 69.94 (3), c = 76.69 (3) Å, β = 136.46 (2)°. Assuming one molecule of CEL-I, which contains two identical subunits, in the asymmetric unit, the value of the Matthews constant V_M (Matthews, 1968) is 2.64 Å^3 Da^{-1}, corresponding to a solvent content of 53%, both of which are within the normal values for protein crystals. The reduction of a total of 72 844 reflections from CEL-I crystals yielded 20 936 independent reflections with a completeness of 92.3% in the resolution range 40–2.0 Å and an overall R_{merge} of 6.1%. The completeness and R_{merge} for the highest resolution shell (2.11–2.0 Å) were 82.2 and 20.1%, respectively.

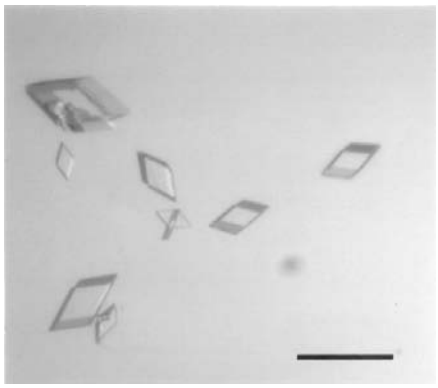

Figure 1
Crystals of CEL-I grown in the presence of $CaCl_2$. Single crystals were obtained by microseeding and macroseeding techniques in sitting drops. The bar represents 0.5 mm.

Molecular-replacement calculations using the CEL-I crystal data were performed with the program *AMoRe* (Navaza, 1994). The search model of the subunit of CEL-I was the human lithostathine structure (PDB code 1lit; Bertrand *et al.*, 1996). Two clear peaks were found with a correlation coefficient (CC) of 0.242 and an *R* factor of 51.7% (10–4.5 Å) after translation-function calculations. From the calculation of the self-rotation function, these two solutions were related by the non-crystallographic twofold axis (Patterson CC of 0.252). The *R* factor was reduced to 48.6% and the CC increased to 0.392 (8–3 Å) with rigid-body refinement. Structure refinement is currently in progress.

We thank Drs N. Sakabe, M. Suzuki and N. Igarashi for kind help with data collection at the Photon Factory. This study was partially supported by a Grant-in-Aid for Scientific Research (B) (#10558109 and #12480181 to MK), a Grant-in-Aid for Scientific Research (C) (#12660082 to TH) and Encouragement of Young Scientists (A) (#12740305 to GK) from Japan Society for the Promotion of Science. This study was also supported in part by a grant from ACT-JST (to MK).

Table 1
Diffraction data statistics.

Values in square brackets refer to the highest resolution shell (2.11–2.0 Å).

X-ray source	Photon Factory BL-18B
Detector	ADSC Quantum 4R CCD
X-ray wavelength (Å)	1.00
Temperature (K)	100
Space group	*C*2
Unit-cell parameters†	
a (Å)	92.38 (3)
b (Å)	69.94 (3)
c (Å)	76.69 (3)
β (°)	136.46 (2)
Resolution limit (Å)	2.0
Total reflections	72844 [5801]
Unique reflections	20936 [2673]
Unique reflections [$I/\sigma(I) > 3$]	19992 [2144]
R_{merge}‡ (%)	6.1 [20.1]
Completeness (%)	92.3 [82.2]
Completeness [$I/\sigma(I) > 3$] (%)	88.1 [65.8]
Multiplicity	3.5 [2.2]
$I/\sigma(I)$	10.5 [2.5]

† Standard deviations in parentheses. ‡ $R_{merge} = \sum_{hkl}\sum_i |I(hkl)_i - \langle I(hkl)\rangle| / \sum_{hkl} I(hkl)$, where $I(hkl)_i$ is the ith measurement of the intensity of reflection *hkl* and $\langle I(hkl)\rangle$ is the mean intensity of reflection *hkl*.

References

Bertrand, J. A., Pignol, D., Bernard, J. P., Verdier, J. M., Dagorn, J. C. & Fontecilla-Camps, J. C. (1996). *EMBO J.* **15**, 2678–2684.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Day, A. J. (1994). *Biochem. Soc. Trans.* **22**, 83–88.
Drickamer, K. (1988). *J. Biol. Chem.* **263**, 9557–9560.
Gabius, H.-J. (1997). *Eur. J. Biochem.* **243**, 543–576.
Hatakeyama, T., Kohzaki, H., Nagatomo, H. & Yamasaki, N. (1994). *J. Biochem.* **116**, 209–214.
Hatakeyama, T., Nagatomo, H. & Yamasaki, N. (1995). *J. Biol. Chem.* **270**, 3560–3564.
Hatakeyama, T., Ohuchi, K., Kuroki, M. & Yamasaki, N. (1995). *Biosci. Biotech. Biochem.* **59**, 1314–1317.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–493.
Nakano, M., Tabata, S., Sugihara, K., Kouzuma, Y., Kimura, M. & Yamasaki, N. (1999). *Biochim. Biophys. Acta*, **1435**, 167–176.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Rossmann, M. G. & van Beek, C. G. (1999). *Acta Cryst.* D**55**, 1631–1640.
Stura, E. A. & Wilson, I. A. (1992). *Crystallization of Nucleic Acids and Proteins: a Practical Approach*, edited by A. Ducruix & R. Giegé, pp. 99–126. Oxford: IRL Press.
Teichberg, V. I., Aberdam, D., Erez, U. & Pinelli, E. (1988). *J. Biol. Chem.* **263**, 14086–14092.
Watanabe, N., Nakagawa, A., Adachi, S. & Sakabe, N. (1995). *Rev. Sci. Instrum.* **66**, 1824–1826.