Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray crystallographic studies of the C-terminal domain of galactose-binding lectin EW29 from the earthworm *Lumbricus terrestris*

Ryuichiro Suzuki,[a,b] Zui Fujimoto,[b]* Atsushi Kuno,[c] Jun Hirabayashi,[c] Ken-ichi Kasai[d] and Tsunemi Hasegawa[a]

[a]Department of Material and Biological Chemistry, Faculty of Science, Yamagata University, Yamagata 990-8560, Japan, [b]Department of Biochemistry, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8602, Japan, [c]Research Center for Glycoscience, National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki 305-8566, Japan, and [d]Department of Biological Chemistry, Faculty of Pharmaceutical Sciences, Teikyo University, Kanagawa 199-0195, Japan

Correspondence e-mail: zui@affrc.go.jp

Received 15 June 2004
Accepted 2 August 2004

The galactose-binding lectin EW29 from the earthworm *Lumbricus terrestris* is composed of two homologous domains, both of which are members of the R-type lectin family. The truncated mutant rC-half comprising the C-terminal domain was crystallized by the hanging-drop vapour-diffusion method. The crystal belonged to space group $P4_32_12$, with unit-cell parameters $a = b = 61.2$, $c = 175.6$ Å, and diffracted to beyond 1.9 Å resolution. Matthews coefficient calculations suggested that this crystal contained two molecules per asymmetric unit.

1. Introduction

The carbohydrate-binding proteins known as lectins exist ubiquitously in both animals and plants, but lectins from the annelid phylum have rarely been reported. Recently, genes for the galactose-binding lectin EW29 from the earthworm *Lumbricus terrestris* were cloned and its sugar-binding specificity was characterized (Hirabayashi *et al.*, 1998). According to sequence analysis, EW29 was found to be composed of two homologous domains (14.5 kDa), *i.e.* N-terminal and C-terminal domains, forming a tandem-repeat type structure; both domains contained triple-repeated Q*X*W motifs, as described previously (Rutenber & Robertus, 1991; Hazes, 1996). The lectin domain containing this motif was defined as an R-type lectin (Villafranca & Robertus, 1981; Rutenber *et al.*, 1987). In general, every domain of R-type lectins, such as *Ricinus communis* toxin B chain (RTB) and *Sambucus sieboldina* agglutinin (SSA), potentially binds to sugar chains; in some cases, their multimerization causes them to possess haemagglutination activity (Villafranca & Robertus, 1981; Kaku *et al.*, 1996). The R-type lectin family is classified into two groups: the lectin group and the enzyme group, as described by Hirabayashi *et al.* (1998). Several three-dimensional structures of R-type lectins have been determined for ricin (Rutenber & Robertus, 1991), abrin (Tahirov *et al.*, 1995) and ebulin (Pascal *et al.*, 2001) in the lectin group and two xylan-binding domains of *So*CBM13 (Fujimoto *et al.*, 2002) and *Sl*CBM13 (Notenboom *et al.*, 2002) in the enzyme group. All proteins possessed the common β-trefoil fold structure, but their sugar binding differs according to their ligand specificities.

Recent biochemical analyses of EW29 have indicated that galactose-binding ability of the N-terminal domain was much weaker than that of the C-terminal domain (Hirabayashi *et al.*, 1998). Interestingly, the truncated mutant comprising only the C-terminal domain (C-half) retained haemagglutination activity (unpublished data). Furthermore, the sugar-binding specificity of the C-half was revealed to be different from that of RTB using frontal affinity chromatography (FAC) (unpublished data). Therefore, three-dimensional structural studies of C-half are expected to elucidate the molecular basis for these biological and biochemical observations.

This paper describes the crystallization and preliminary X-ray crystallographic studies of the C-half of EW29. There is no previous information on the structure of a single-domain type of haemagglutinin from the R-type lectin family. This is also the first report of a crystallization of an annelid phylum lectin.

2. Experimental procedures and results

2.1. Protein preparation and crystallization

The overexpression vector for the recombinant C-half (rC-half) was constructed previously (Hirabayashi *et al.*, 1998). The plasmid was transformed into *Escherichia coli* BL21 (DE3) strain and the transformant was cultivated in 1 l of LB medium containing kanamycin (0.025 mg ml^{-1}) at 310 K with shaking to an optical density at 600 nm of 0.4. After addition of isopropyl-1-thio-β-D-galactoside (IPTG) to a final concentration of 1 m*M*, the culture was incubated at 298 K with agitation for 12 h. The cells were harvested and lysed using a sonifier and the cell lysate was subjected to lactose-Sepharose affinity chromatography. Lactose-Sepharose resin was prepared as described previously (Ito *et al.*, 2004). The lactose-Sepharose column (5 ml bed volume) was equilibrated with 20 m*M* phosphate buffer pH 7.2 at room temperature

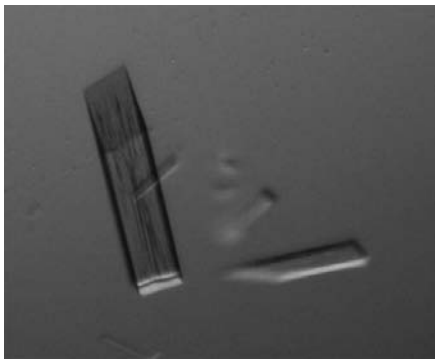

Figure 1
Typical crystals of rC-half grown by the hanging-drop vapour-diffusion method.

and 50 ml of cell lysate solution filtrated with a 0.45 µm Millex filter (Millipore, Bedford, MA, USA) was loaded onto the column. The column was thoroughly washed with 20 m*M* phosphate buffer pH 7.2 and bound proteins were then eluted with 200 m*M* lactose in the same buffer. The purified protein solution was desalted with a PD-10 desalting column (Amersham Biosciences, Uppsala, Sweden) and then concentrated with Centriprep (Millipore, Bedford, MA, USA) to approximately 5 ml (20 mg ml^{-1}). The purity of the protein was checked by SDS–PAGE.

Crystallization conditions were screened by the sitting-drop vapour-diffusion technique using Wizard Screens I and II (Emerald Biostructures, Bainbridge Island, USA). A protein solution of 15–20 mg ml^{-1} was used for initial screening. A droplet composed of 0.5 µl protein solution containing 15 mg ml^{-1} D-lactose and 0.5 µl reservoir solution was equilibrated against 100 µl reservoir solution at 293 K. Small or thin needle-shaped crystals of rC-half were obtained in several buffer conditions containing imidazole. A second screening of crystallization conditions was performed by the hanging-drop vapour-diffusion method using dipotassium hydrogen phosphate and sodium dihydrogen phosphate as precipitants. 5 µl protein solution (containing 15 mg ml^{-1} D-lactose and 0.01 *M* cadmium chloride as additives) and 5 µl reservoir solution were mixed and equilibrated against 1 ml of the same reservoir solution. After refinement of the crystallization conditions, crystals of rC-half (0.4 × 0.1 × 0.05 mm; Fig. 1) were obtained at 293 K within a week using 20 mg ml^{-1} protein solution and the following reservoir solution: 0.2 *M* sodium chloride, 1.6 *M* dipotassium hydrogen phosphate, 0.4 *M* sodium dihydrogen phosphate, 0.1 *M* imidazole buffer pH 8.0. As the addition of imidazole is necessary for crystallization of rC-half, it may be that it produces better crystal packing.

2.2. Data collection

Diffraction experiments were conducted at beamline NW12, Photon Factory, Advanced Ring (PF-AR), Tsukuba, Japan (λ = 0.978 Å). Crystals of rC-half were mounted in nylon loops (Hampton Research, CA, USA) after soaking in a cryoprotectant solution (20% glycerol in the precipitant solution). The crystals were then flash-frozen in a nitrogen stream at 95 K. Diffraction data were collected using a Quantum 210 CCD X-ray detector (Area Detector Systems Corporation, Poway, CA, USA) in 1° oscillation steps over a range of 180°. The data set was processed and scaled using the programs *DENZO* and *SCALEPACK* from the *HKL*2000 package (Otwinowski & Minor, 1997). The crystals of rC-half belonged to the tetragonal space group $P4_32_12$, with unit-cell parameters $a = b = 61.2$, $c = 175.6$ Å, and diffracted to beyond 1.9 Å resolution. The data-collection statistics are shown in Table 1. Matthews coefficient calculations suggested that the V_M value of the crystal was 2.9 Å^3 Da^{-1}, assuming the presence of two rC-half molecules in the asymmetric unit (Matthews, 1968). This V_M value corresponds to a solvent content of 58%.

The molecular-replacement method was used for structural analysis of rC-half using the crystal structure of *So*CBM13 as a search model. *So*CBM13 is a xylan-binding domain of the family 10 xylanase from *Streptomyces olivaceoviridis* E-86 (PDB code 1xyf) and folds in the same manner as the RTB (Rutenber & Robertus, 1991; Fujimoto *et al.*, 2000). Phasing and model building using the programs *MOLREP* (Vagin & Teplyakov, 1997) from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994) and *ARP*/*wARP* (Perrakis *et al.*, 1999) are now under way.

Table 1
Data-collection statistics of rC-half.

Values in parentheses refer to the highest resolution shell.

X-ray source	PF-AR NW12
Space group	$P4_32_12$
Unit-cell parameters (Å)	$a = b = 61.2$, $c = 175.6$
Wavelength (Å)	0.978
Resolution (Å)	50.0–1.90 (1.97–1.90)
R_{merge} (%)	6.5 (29.4)
Completeness (%)	99.9 (100.0)
Average $I/\sigma(I)$	22.2 (3.3)
Average redundancy	7.0 (7.2)
Unique reflections	27297 (2660)
Observed reflections	191587

We thank the beamline scientists at PF-AR for their assistance with X-ray data collection.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Fujimoto, Z., Kuno, A., Kaneko, S., Kobayashi, H., Kusakabe, I. & Mizuno, H. (2002). *J. Mol. Biol.* **316**, 65–78.
Fujimoto, Z., Kuno, A., Kaneko, S., Yoshida, S., Kobayashi, H., Kusakabe, I. & Mizuno, H. (2000). *J. Mol. Biol.* **300**, 575–585.
Hazes, B. (1996). *Protein Sci.* **5**, 1490–1501.
Hirabayashi, J., Dutta, S. K. & Kasai, K. (1998). *J. Biol. Chem.* **273**, 14450–14460.
Ito, S., Kuno, A., Suzuki, R., Kaneko, S., Kawabata, Y., Kusakabe, I. & Hasegawa, T. (2004). *J. Biotechnol.* **110**, 137–142.
Kaku, H., Tanaka, Y., Tazaki, K., Minami, E., Mizuno, H. & Shibuya, N. (1996). *J. Biol. Chem.* **271**, 1480–1485.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Notenboom, V., Boraston, A. B., Williams, S. J., Kilburn, D. G. & Rose, D. R. (2002). *Biochemistry*, **41**, 4246–4254.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pascal, J. M., Day, P. J., Monzingo, A. F., Ernst, S. R., Robertus, J. D., Iglesias, R., Perez, Y., Ferreras, J. M., Citores, L. & Girbes, T. (2001). *Proteins Struct. Funct. Genet.* **43**, 319–326.
Perrakis, A., Morris, R. J. H. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Rutenber, E., Ready, M. & Robertus, J. D. (1987). *Nature (London)*, **326**, 624-626.
Rutenber, E. & Robertus, J. D. (1991). *Proteins Struct. Funct. Genet.* **10**, 260–269.
Tahirov, T. H., Lu, T. H., Liaw, Y. C., Chen, Y. L. & Lin, J. Y. (1995). *J. Mol. Biol.* **250**, 354–367.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Villafranca, J. E. & Robertus, J. D. (1981). *J. Biol. Chem.* **256**, 554–556.