Hiroshi Yokoyama,^{a*} Takahiro Yamashita,^a Naoki Horikoshi,^b Hitoshi Kurumizaka^b and Wataru Kagawa^{c*}

^aAnimal Waste Management and Environment Division, National Institute of Livestock and Grassland Science, 2 Ikenodai, Tsukuba, Ibaraki 305-0901, Japan, ^bLaboratory of Structural Biology, Graduate School of Advanced Science and Engineering, Waseda University, 2-2 Wakamatsu-cho, Shinjuku-ku, Tokyo 162-8480, Japan, and ^cProgram in Chemistry and Life Science, Department of Interdisciplinary Science and Engineering, School of Science and Engineering, Meisei University, 2-1-1 Hodokubo, Hino-shi, Tokyo 191-8506, Japan

Correspondence e-mail: hiroshiy@affrc.go.jp, wataru.kagawa@meisei-u.ac.jp

Received 4 February 2013

Accepted 26 February 2013

Crystallization and preliminary X-ray diffraction analysis of the secreted protein Athe_0614 from *Caldicellulosiruptor bescii*

The Athe_0614 protein is a component of the extracellular proteins secreted by the anaerobic, extremely thermophilic and cellulolytic bacterium *Caldicellulosiruptor bescii*. The recombinant protein was expressed in *Escherichia coli*, purified to near-homogeneity and crystallized using polyethylene glycol 2000 monomethyl ether as a precipitant. The crystals belonged to the monoclinic space group $P2_1$, with unit-cell parameters $a = 48.4$, $b = 42.2$, $c = 97.8$ Å, $\beta = 96.1^\circ$, and diffracted to 2.7 Å resolution using synchrotron radiation.

1. Introduction

The enzymatic conversion of cellulosic biomass to fermentable sugars has attracted particular attention with regard to biofuel production. *Caldicellulosiruptor bescii*, formerly known as *Anaerocellum thermophilum* (Svetlichnyi *et al.*, 1990), is an extremely thermophilic bacterium with a growth-temperature range of 315–363 K (optimum range 351–353 K; Yang *et al.*, 2010). *C. bescii* degrades microcrystalline cellulose and native grasses (Yang *et al.*, 2009) and shows higher degrading activities than the cellulases from the fungus *Trichoderma reesei* (Kanafusa-Shinkai *et al.*, 2013), which form the most common enzyme system used for plant-biomass saccharification. Therefore, the cellulase system of *C. bescii* is a potential enzyme for biofuel production from plant biomass.

C. bescii secretes a number of extracellular proteins, including multidomain cellulases for the degradation of plant biomass. The Athe_0614 protein (49 kDa) is a component of the secreted extracellular proteins (Dam *et al.*, 2011), but its biological function remains unclear. The protein shows 35% amino-acid sequence identity to the xylooligosaccharide-binding protein (XBP1) from *Caldanaerobius polysaccharolyticus* (Han *et al.*, 2012), implying that the protein might be involved in xylan utilization. A computer search using the Conserved Domain Database (Marchler-Bauer *et al.*, 2011) suggested that the protein consists of a single domain (amino-acid residues 30–418) that is structurally quite similar to the PBP2 LTTR substrate superfamily (cl11398). This superfamily consists of various protein families, including the LysR substrate-binding domain family (Pfam 03466) and bacterial extracellular solute-binding protein family 1 (Pfam 01547). LysR-type transcriptional regulators and maltose-binding proteins are included in these families. In the present study, we report the crystallization and preliminary X-ray diffraction analysis of the Athe_0614 protein.

2. Materials and methods

2.1. Protein expression and purification

The DNA fragment encoding amino-acid residues 27–435 of the Athe_0614 protein (accession No. ACM59738) was amplified from the genomic DNA of *C. bescii* DSM 6725 by PCR using primers F27 (5'-CATGC TAGCT CTGAA AAGGC TTTCA GCTCA GCTGG TGGAA) and R435 (5'-CTGCG GCCGC TTACT TCTTG AAATA TTCCT TAGCA GCTTT). The truncated N-terminal region (amino-acid residues 1–26) contains a putative secretion signal (Emanuelsson

et al., 2007). The resulting DNA fragment was digested with *NheI* and *NotI*, and was cloned into pET28b (Novagen) vector such that a hexahistidine-tag sequence was fused to the N-terminus of the protein product. Recombinant hexahistidine-tagged Athe_0614_{27–435} protein was overexpressed in *Escherichia coli* strain Rosetta 2 (Novagen) by induction with 0.2 mM isopropyl β -D-1-thiogalactopyranoside. The cells producing recombinant protein were resuspended in a binding buffer consisting of 20 mM HEPES pH 7.2, 20 mM imidazole, 0.3 M NaCl, 1 mM DTT, 0.025% Tween-20. After cell disruption by sonication, the cell debris was removed by centrifugation. The soluble fraction was heat-treated at 348 K for 20 min and centrifuged again to remove insoluble materials. The soluble fraction was applied onto HisTrap HP nickel resin (GE Healthcare Biosciences). After washing with binding buffer containing 50 mM imidazole, the recombinant protein was eluted with binding buffer containing 0.5 M imidazole. The eluted protein was dialyzed against 20 mM HEPES pH 7.5 buffer containing 2 mM β -mercaptoethanol. The hexahistidine tag was removed by thrombin protease treatment (approximately 2 U per milligram of protein) during dialysis. The protein was further purified using HiTrap SP HP ion-exchange resin (GE Healthcare Biosciences) with a linear gradient elution from 0 to 1 M NaCl in the buffer used in the dialysis step. The purified protein was dialyzed against 20 mM MES pH 6.0 buffer, 2 mM β -mercaptoethanol and was concentrated to 80–100 mg ml⁻¹ using a Vivaspin 20 centrifugal filter (Sartorius Stedim Biotech) for crystallization.

2.2. Crystallization and diffraction data collection

The initial screening of crystallization conditions was conducted in 96-well plates by the sitting-drop vapour-diffusion technique using a Hydra II crystallization robot (Thermo Fisher Scientific) with commercially available screening kits (Crystal Screen and Crystal Screen 2 from Hampton Research and JCSG Core Suites I–IV from Qiagen). Droplets consisting of 0.2 μ l each of protein solution (40–50 mg ml⁻¹) and reservoir solution were equilibrated against 70 μ l reservoir solution at 293 K and crystal formation was observed after 7 d. Crystallization conditions were optimized in 24-well plates using the hanging-drop vapour-diffusion technique. Droplets consisting of 1 μ l each of protein solution (40–50 mg ml⁻¹) and reservoir solution

Figure 1
SDS-PAGE analysis. The Athe_0614_{27–435} protein was separated by 11.5% SDS-PAGE and stained with Coomassie Brilliant Blue. The positions of molecular-mass markers are labelled in kDa.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

Beamline	BL17A, Photon Factory
Wavelength (Å)	0.98
Space group	<i>P</i> 2 ₁
Unit-cell parameters (Å, °)	<i>a</i> = 48.4, <i>b</i> = 42.2, <i>c</i> = 97.8, β = 96.1
Resolution range (Å)	50–2.7 (2.76–2.70)
Observed reflections	78291
Unique reflections	11036
Oscillation width (°)	0.5
Total rotation range (°)	360
Multiplicity	7.1 (7.2)
Mosaicity (°)	0.6
Completeness (%)	99.9 (100)
<i>R</i> _{merge} [†] (%)	8.6 (18.7)
$\langle I/\sigma(I) \rangle$	14.4 (8.4)

[†] $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation and $\langle I(hkl) \rangle$ is the mean value for that reflection.

were equilibrated against 500 μ l reservoir solution at 293 K. The crystals were cryoprotected in reservoir solution supplemented with 10% (v/v) glycerol and were flash-cooled in a liquid-nitrogen stream at 100 K. X-ray diffraction data were collected using an ADSC Quantum 315 CCD detector on beamline 17A of the Photon Factory, Tsukuba, Japan. The data sets were processed with *MOSFLM* (Leslie, 2006) and scaled with *AIMLESS/SCALA* (Winn *et al.*, 2011) using the *iMosflm* interface (Battye *et al.*, 2011).

3. Results and discussion

The Athe_0614_{27–435} protein expressed in *E. coli* as a soluble fraction was purified to near-homogeneity and the protein yield was 8 mg per litre of bacterial culture. The purity of the protein was estimated to be greater than 95% by SDS-PAGE with Coomassie Brilliant Blue staining (Fig. 1). Crystals were obtained from five crystallization conditions (conditions Nos. 67, 76 and 80 of JCSG Core Suite III and Nos. 14 and 59 of JCSG Core Suite IV) during the initial screening. Condition No. 80 of JCSG Core Suite III [0.1 M citric acid pH 4.0, 30% (w/v) PEG 6000] was further optimized to 0.1 M citric acid pH 4.0, 25% (w/v) PEG 2000 monomethyl ether. Thin rod-like polycrystals of the Athe_0614_{27–435} protein were obtained after 2–3 d under the optimized condition (Fig. 2). The crystals were cracked with a cryoloop and single crystals were picked up. A diffraction data set was collected to 2.7 Å resolution from a single crystal (Table 1;

Figure 2
Crystals of the Athe_0614_{27–435} protein. The scale bar represents 100 μ m.

Figure 3
Diffraction image of a *C. bescii* Athe_0614 crystal.

Fig. 3). The data analysis showed that the crystal belonged to the monoclinic space group $P2_1$, with unit-cell parameters $a = 48.4$, $b = 42.2$, $c = 97.8$ Å, $\beta = 96.1^\circ$. The asymmetric unit was expected to contain one molecule of the Athe_0614 protein, with a crystal volume per unit molecular weight of $2.22 \text{ Å}^3 \text{ Da}^{-1}$, which corresponds to a solvent content of 45%. As mentioned above, the Athe_0614 protein is suggested to have structural similarities to XBP1, LysR-type transcriptional regulators and maltose-binding proteins. Data analysis by the molecular-replacement method using these proteins as search models is currently in progress.

We thank Dr K. Saito and Mr K. Onodera for their assistance in the initial screening for crystallization. This work was supported in part by Grants-in-Aid for Scientific Research (C) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Dam, P. *et al.* (2011). *Nucleic Acids Res.* **39**, 3240–3254.
- Emanuelsson, O., Brunak, S., von Heijne, G. & Nielsen, H. (2007). *Nature Protoc.* **2**, 953–971.
- Han, Y., Agarwal, V., Dodd, D., Kim, J., Bae, B., Mackie, R. I., Nair, S. K. & Cann, I. K. (2012). *J. Biol. Chem.* **287**, 34946–34960.
- Kanafusa-Shinkai, S., Wakayama, J., Tsukamoto, K., Hayashi, N., Miyazaki, Y., Ohmori, H., Tajima, K. & Yokoyama, H. (2013). *J. Biosci. Bioeng.* **115**, 64–70.
- Leslie, A. G. W. (2006). *Acta Cryst.* **D62**, 48–57.
- Marchler-Bauer, A. *et al.* (2011). *Nucleic Acids Res.* **39**, D225–D229.
- Svetlichnyi, V. A., Svetlichnaya, T. P., Chernykh, N. A. & Zavarzin, G. A. (1990). *Microbiology*, **59**, 598–604.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yang, S.-J., Kataeva, I., Hamilton-Brehm, S. D., Engle, N. L., Tschaplinski, T. J., Doeppke, C., Davis, M., Westpheling, J. & Adams, M. W. W. (2009). *Appl. Environ. Microbiol.* **75**, 4762–4769.
- Yang, S.-J., Kataeva, I., Wiegel, J., Yin, Y., Dam, P., Xu, Y., Westpheling, J. & Adams, M. W. W. (2010). *Int. J. Syst. Evol. Microbiol.* **60**, 2011–2015.