Crystallization and preliminary X-ray diffraction studies of the family 54 carbohydrate-binding module from laminarinase (β -1,3-glucanase) Lic16A of *Clostridium thermocellum*

Yury A. Kislitsyn,^{a*} Valeriya R. Samygina,^a Igor A. Dvortsov,^b Nataliya A. Lunina,^b
Inna P. Kuranova^a and Galina A. Velikodvorskaya^b

Received 12 September 2014

Accepted 11 January 2015

Keywords: carbohydrate-binding module; β -1,3-glucanase; *Clostridium thermocellum*; Lic16A.

Supporting information: this article has supporting information at journals.iucr.org/f

^aA. V. Shubnikov Institute of Crystallography, Russian Academy of Sciences, Leninsky Prospect 59, Moscow 117333, Russian Federation, and ^bInstitute of Molecular Genetics, Russian Academy of Sciences, Kurchatov Square 2, Moscow 123182, Russian Federation. *Correspondence e-mail: kislitsynalex@hotmail.com

The crystallization and preliminary X-ray diffraction analysis of the carbohydrate-binding module (CBM) from laminarinase Lic16A of the hyperthermophilic anaerobic bacterium *Clostridium thermocellum* (ctCBM54) are reported. Recombinant ctCBM54 was prepared using an *Escherichia coli*/pQE30 overexpression system and was crystallized by the hanging-drop vapour-diffusion method. X-ray diffraction data were collected to 2.1 Å resolution using synchrotron radiation. The crystals belonged to space group $P6_322$, with unit-cell parameters $a = b = 130.15$, $c = 131.05$ Å. The three-dimensional structure of ctCBM54 will provide valuable information about the structure–function relation of the laminarinase Lic16A and will allow the exploitation of this binding module in biotechnological applications.

1. Introduction

The specific recognition of polysaccharides by the carbohydrate-binding modules (CBMs) of glycosyl hydrolases plays a key role in important biological processes. The versatile nature of such interactions manifests itself in a wide range of phenomena from the degradation of plant biomass to the treatment of diseases (Sharma & Nakas, 1987; Zverlov *et al.*, 1997; König *et al.*, 2002; John *et al.*, 2011; Pang *et al.*, 2005). It is known that the main functions of CBMs in a multimodular enzyme are the following (Boraston *et al.*, 2004; Hashimoto, 2006). (i) Substrate proximity: the CBMs binds to a polysaccharide, thus placing the catalytic module in close proximity to its substrate. (ii) Substrate targeting: a CBMs having an affinity for the same polysaccharide as the catalytic module can increase its efficiency. (iii) Enzyme stabilization: the presence of auxiliary modules stabilizes the enzyme structure and enhances its temperature resistance. Many CBMs have now been identified experimentally and many other putative CBMs can be further identified. Currently, 69 CBM families have been classified by their amino-acid sequence similarity (Boraston *et al.*, 1999; Coutinho & Henrissat, 1999).

ctCBM54 is a fragment (residues 192–421) of laminarinase (β -1,3-glucanase) Lic16A (EC 3.2.1.39) from *Clostridium thermocellum*. This multi-modular enzyme with a molecular mass of 147 kDa is found *in vivo* bound to the S-layer of the cell wall of clostridia. It shows high activity on soluble polysaccharides with mixed 1,3–1,4- β -linkages (Fuchs *et al.*, 2003). The primary structure of LicA16A is composed of a leader peptide, a threefold repeat of a module homologous to proteins of the cell-surface layer (SLH), the ctCBM54 module,

Table 1

Macromolecule-production information.

Source organism	<i>C. thermocellum</i>
DNA source	Genomic DNA
Forward primer	CAAGTCGGATCCACAATAAAATCTGAAT
Reverse primer	AATGAAGAGTGGCGTTAAAGCTTGAATAA
Cloning vector	pQE30
Expression vector	pQE30
Expression host	<i>E. coli</i> /XL-1
Complete amino-acid sequence of the construct produced†	MGRSHHHHHGSTIKILNKIIPSLYNEKGDYKNE- EVAGNALINTEGVILKDTVINGDLYLAQGIQN- GDVTLDGVNKGTVFVNGGSDSIHFINTKIN- RVVVKNTGVRIVTSGNTSVESVVKSGAKLEE- KELTGDGFKNVTVDSQLSAGNEIFVGDFEQV- DVLADDALLETKEAKMKLRIFGQRIKVNKAI- EKSSKNYIVNGELISTEEEPGSPDAPGAEDDQ- NSGSPGSSTNPAPTKNPNEWR

† The underlined parts are the BamHI and HindIII restrictase sites.

a catalytic GH16 module and a fourfold repeat of a CBM belonging to the CBM4A subfamily (Zverlov *et al.*, 2001; Fig. 1). The binding of the enzyme to the S-layer is performed using the SLH and CBM54 modules. In addition to binding to the polysaccharide structure of the cell wall, ctCBM54 affects the optimum temperature and thermostability of the catalytic module of Lic16A (Dvortsov *et al.*, 2009). ctCBM54 is the first known member of a new family of CBM: family 54. At present, no three-dimensional structures are known of modules belonging to the CBM54 family.

Structure–function studies of carbohydrate-binding modules are crucial for comprehensive understanding of the biological function of laminarinase, which can then be manipulated to develop microbial strains for the production of platform chemicals from plant biomass. In this study, we present the crystallization and the preliminary X-ray data analysis of ctCBM54.

2. Materials and methods

2.1. Macromolecule production

Escherichia coli XL-1 Blue was used for cloning and expression. This strain was transformed by the pQE30 plasmid (Qiagen) carrying the CBM54 DNA fragment and the helper plasmid pACYC-RIL (Stratagene) containing *argU*, *ileY* and *leuW* tRNA genes, as published previously (Dvortsov *et al.*, 2010). Cells were grown in Luria–Bertani (LB) medium with the addition of 200 µg ml^{−1} ampicillin and 40 µg ml^{−1} chloramphenicol at 37°C until the OD₆₀₀ reached 0.6–0.8 and were subsequently induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Expression continued for 3 h at a temperature of 37°C. The cells were harvested by centrifugation (4000g for 30 min at 4°C). The cells were resuspended

Figure 1

Domain structure of LicA from *C. thermocellum*. An explanation of the figure is given in the text.

Table 2

Crystallization.

Method	Hanging-drop vapour diffusion
Plate type	24-well Linbro crystallization plates
Temperature (K)	293
Protein concentration (mg ml ^{−1})	15
Buffer composition of protein solution	10 mM Na HEPES pH 7.0
Composition of reservoir solution	1.8 M sodium malonate pH 7.0
Volume and ratio of drop	3 µl, 1:1
Volume of reservoir (ml)	0.7

in buffer consisting of 50 mM potassium dihydrogen phosphate pH 8.0, 300 mM NaCl, the protease inhibitor PMSF and lysozyme and were sonicated. After removal of the cell debris by centrifugation (18 000g for 15 min), the supernatant was then loaded onto a Ni–NTA agarose column (Qiagen) preloaded with Ni²⁺ and then eluted with a gradient of imidazole (20–250 mM) in buffer consisting of 50 mM potassium dihydrogen phosphate pH 6.0. Samples were run on 10% SDS–PAGE and fractions containing the homogenous protein were pooled, precipitated by ammonium sulfate (final concentration 1.9 M) and stored at 4°C. The macromolecule-production information is summarized in Table 1.

2.2. Crystallization

Samples of the precipitated protein were extensively dialyzed against 10 mM Na HEPES pH 7.0 and concentrated in a centrifugal ultrafiltration device (Millipore). Samples of the protein were diluted to a final concentration of 15 mg ml^{−1} in 10 mM Na HEPES pH 7.0. The protein was crystallized by the hanging-drop vapour-diffusion method at 20°C using 24-well Linbro crystallization plates. The initial screening conditions were set up as described in the manuals for the commercially available kits SaltRx 1 (Hampton Research) and The Protein Complex Suite (Qiagen). In the crystallization droplets, 1.5 µl protein solution (15 mg ml^{−1}) was mixed with an equal volume of reservoir solution on siliconized cover slides (Hampton Research). The cover slides were inverted over the wells of a 24-well plate containing 0.7 ml reservoir solution. The wells were sealed vapour-tight with high-vacuum grease (Dow Corning). The droplets were then equilibrated against the reservoir solution, and after three weeks a mixture of amorphous precipitate and small crystals was observed in some of the drops. To increase the crystal size, we varied the pH value, the protein concentration and the precipitant concentration in subsequent crystallization experiments. Finally, crystals were obtained by equilibrating 3 µl protein drops against a reservoir solution consisting of 1.8 M sodium malonate pH 7.0 (Table 2).

2.3. Data collection and processing

X-ray diffraction experiments were performed using a PILATUS 4M detector on beamline XALOC at ALBA, Cerdanyola del Vallès, Spain (Colldelram *et al.*, 2011). For X-ray data collection, a single crystal was cryoprotected using 3.4 M sodium malonate pH 7.0 and was flash-cooled in a liquid-nitrogen stream. Diffraction data were collected over a

Table 3

Data collection and processing.

Values in parentheses are for the outer shell.

Beamline	XALOC, ALBA
Wavelength (Å)	1.107030
Temperature (K)	100
Detector	PILATUS 4M
Crystal-to-detector distance (mm)	335.6
Rotation range per image (°)	0.3
Total rotation range (°)	60
Exposure time per image (s)	0.25
Space group	<i>P</i> ₆ 22
<i>a</i> , <i>b</i> , <i>c</i> (Å)	130.15, 130.15, 131.05
α , β , γ (°)	90, 90, 120
Mosaicity (°)	0.22
Resolution range (Å)	36.0–2.1 (2.21–2.10)
Total No. of reflections	227687 (32347)
No. of unique reflections	38127 (5501)
Completeness (%)	99.0 (99.6)
Multiplicity	6.0 (5.9)
$\langle I/\sigma(I) \rangle$	13.0 (2.7)
$R_{\text{meas}}^{\dagger}$ (%)	12.4 (77.9)
$R_{\text{merge}}^{\ddagger}$ (%)	11.3 (71.0)
$R_{\text{p.i.m.}}^{\S}$ (%)	5.0 (31.5)
$\text{CC}_{1/2}$ (correlation coefficient)	0.99 (0.78)
CC^*	0.70 (0.66)
Overall <i>B</i> factor from Wilson plot (Å ²)	29.9

$$\dagger R_{\text{meas}} = \frac{\sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}.$$

$$\ddagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}. \quad \S R_{\text{p.i.m.}} = \frac{\sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}.$$

range of 60° with an oscillation angle of 0.3° and a crystal-to-detector distance of 335.6 mm. The obtained diffraction data set was processed using *XDS* (Kabsch, 1993) and *SCALA* (Evans, 2006). The space group was determined using the program *POINTLESS* (Evans, 2006) from the *CCP4* suite (Winn *et al.*, 2011). The data-collection and processing statistics are summarized in Table 3.

3. Results and discussion

The ctCBM54 DNA fragment was cloned into pQE30 expression vector, overexpressed in *E. coli* XL-1 Blue cells and purified to homogeneity using Ni-NTA agarose (Table 1). The purified ctCBM54 was crystallized. The crystals of His6-

Figure 2

Crystals of ctCBM54 obtained using the hanging-drop vapour-diffusion technique (approximate dimensions 0.16 × 0.12 × 0.12 mm).

Table 4

Matthews coefficient based on monomeric ctCBM54 as the molecular unit.

No. of molecules	Probability	V_M (Å ³ Da ⁻¹)	V_s (%)	Molecular weight (Da)
1	0.0052	6.15	79.86	26017.00
2	0.3328	3.08	60.00	52034.00
3	0.6581	2.05	40.01	78051.00
4	0.0038	1.54	20.01	104068.00

ctCBM54 grew to maximum dimensions of 0.16 × 0.12 × 0.12 mm in approximately six weeks (Fig. 2; Table 2). The crystal diffracted to 2.1 Å resolution (Supplementary Fig. S1) and appeared to belong to space group *P*₆22, with unit-cell parameters *a* = *b* = 130.15, *c* = 131.05 Å. The data-collection statistics are summarized in Table 3. In order to define the number of ctCBM54 molecules in the asymmetric unit, the Matthews coefficient (V_M) was calculated (Matthews, 1968). The probabilities were calculated according to the method of Kantardjieff & Rupp (2003); four possible solutions were found with a monomer as the basic unit (Table 4). Based on the total molecular weight of 26 kDa for the monomer, the space group and the unit-cell parameters, V_M calculations indicated that the asymmetric part theoretically contains between one and four monomers, with solvent contents (V_s) ranging from 80 to 20%. The self-rotation function calculated by *MOLREP* (Vagin & Teplyakov, 2010) revealed no non-crystallographic symmetry peaks (χ = 180° section shown in Fig. 3; χ = 120° section, data not shown). ctCBM54 consists of two interconnected parts: an N-terminal submodule of 8 kDa and a C-terminal submodule of 17 kDa (Dvortsov *et al.*, 2010). The N-terminal submodule specializes in the binding of

Figure 3

The self-rotation function map of the ctCBM54 crystal computed by *MOLREP* at χ = 180°.

D-glucose β -linked polymers, for example bacterial crystalline cellulose and Avicel. The C-terminal submodule has a broad substrate affinity: in addition to glucosides, it binds chitin, chitosan and xylan. The binding affinity of the C-terminal submodule can be increased by Ca^{2+} ions (Dvortsov *et al.*, 2010). The CBM2 family and CBM49 family have binding properties relatively similar to that of the N-terminal submodule of ctCBM54. The three-dimensional structures of some modules from the CBM2 family have been solved, for example the chitin-binding module of chitinase Chi18A from *Pyrococcus furiosus* (PDB entry 2cwr; Nakamura *et al.*, 2008). The sequence identity of this module to ctCBM54 does not exceed 5%. The CBM49 family contains no members with a resolved three-dimensional structure. The binding module of Xyn43A from *Paenibacillus polymyxa* (PDB entry 1ux7; Jamal-Talabani *et al.*, 2004) belongs to the CBM36 family and shows calcium-dependent binding of substrates similar to the C-terminal submodule of ctCBM54. However, the modules of Chi18A and Xyn43A have low sequence identity (5 and 6%, respectively) to ctCBM54.

The information above makes it unlikely that molecular replacement could be used to solve the phase problem to obtain the three-dimensional structure of ctCBM54. An attempt to solve the three-dimensional structure by isomorphous replacement using heavy-atom protein derivatives is under way.

Acknowledgements

This work was supported by RFBR (Russian Foundation for Basic Research, grant No. 14-04-00308) and by the Program of the Russian Academy of Sciences 'Molecular and Cell Biology'. We thank Jordi Juanhuix (ALBA beamline BL13/XALOC) for his generous support during data collection.

References

- Boraston, A. B., Bolam, D. N., Gilbert, H. J. & Davies, G. J. (2004). *Biochem. J.* **382**, 769–781.
- Boraston, A. B., McLean, B. W., Kormos, J. M., Alam, M., Gilkes, N. R., Haynes, C. A., Tomme, P., Kilburn, D. G. & Warren, R. A. J. (1999). *Recent Advances in Carbohydrate Bioengineering*, edited by H. J. Gilbert, B. Svensson, G. Davies & B. Henrissat, pp. 202–211. Cambridge: The Royal Society of Chemistry.
- Colldelram, C., Ruget, C. & Nikitina, L. (2011). *Diamond Light Source Proc.* **1**, e44.
- Coutinho, P. M. & Henrissat, B. (1999). *Recent Advances in Carbohydrate Bioengineering*, edited by H. J. Gilbert, G. Davies, B. Henrissat & B. Svensson, pp. 3–12. Cambridge: The Royal Society of Chemistry.
- Dvortsov, I. A., Lunina, N. A., Chekanovskaya, L. A., Schwarz, W. H., Zverlov, V. V. & Velikodvorskaya, G. A. (2009). *Microbiology*, **155**, 2442–2449.
- Dvortsov, I. A., Lunina, N. A., Zverlov, V. V. & Velikodvorskaya, G. A. (2010). *Mol. Biol.* **44**, 591–595.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Fuchs, K. P., Zverlov, V. V., Velikodvorskaya, G. A., Lottspeich, F. & Schwarz, W. H. (2003). *Microbiology*, **149**, 1021–1031.
- Hashimoto, H. (2006). *Cell. Mol. Life Sci.* **63**, 2954–2967.
- Jamal-Talabani, S., Boraston, A. B., Turkenburg, J. P., Tarbouriech, N., Ducros, V. M.-A. & Davies, G. J. (2004). *Structure*, **12**, 1177–1187.
- John, R. P., Anisha, G. S., Nampoothiri, K. M. & Pandey, A. M. (2011). *Bioresour. Technol.* **102**, 186–193.
- Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
- Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
- König, J., Grasser, R., Pikor, H. & Vogel, K. (2002). *Anal. Bioanal. Chem.* **374**, 80–87.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Nakamura, T., Mine, S., Hagihara, Y., Ishikawa, K., Ikegami, T. & Uegaki, K. (2008). *J. Mol. Biol.* **381**, 670–680.
- Pang, Z., Otaka, K., Maoka, T., Hidaka, K., Ishijima, S., Oda, M. & Ohnishi, M. (2005). *Biosci. Biotechnol. Biochem.* **69**, 553–558.
- Sharma, A. & Nakas, J. P. (1987). *Enzyme Microb. Technol.* **9**, 89–93.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Zverlov, V. V., Volkov, I. Y., Velikodvorskaya, T. V. & Schwarz, W. H. (1997). *Microbiology*, **143**, 1701–1708.
- Zverlov, V. V., Volkov, I. Y., Velikodvorskaya, G. A. & Schwarz, W. H. (2001). *Microbiology*, **147**, 621–629.