Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary crystallographic analysis of thermophilic cellulase from *Fervidobacterium nodosum* Rt17-B1

Baisong Zheng,[a,b] Wen Yang,[a] Yuguo Wang,[b] Yan Feng[b]* and Zhiyong Lou[a]*

[a]Tsinghua–Nankai–IBP Joint Research Group for Structural Biology, Tsinghua University, Beijing 100084, People's Republic of China, and [b]Key Laboratory for Molecular Enzymology and Engineering of the Ministry of Education, Jilin University, Changchun 130023, People's Republic of China

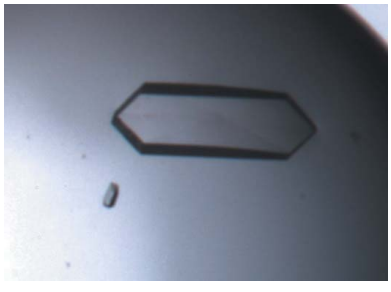
Correspondence e-mail: yfeng@mail.jlu.edu.cn, louzy@xtal.tsinghua.edu.cn

Received 7 October 2008
Accepted 12 January 2009

FnCel5A, a thermostable endoglucanase, is a member of glycohydrolase family 5 which catalyzes the hydrolysis of cellulose to glucose in the thermophilic bacterium *Fervidobacterium nodosum* Rt17-B1. FnCel5A is particularly interesting because of its high thermostability (T_{opt} = 353 K, half-life 48 h) and its high specific activity towards carboxymethylcellulose. These properties make FnCel5A an attractive target for protein engineering to improve cellulase activity. In order to resolve the crystal structure of FnCel5A and to gain a better understanding of its biological function, recombinant FnCel5A was expressed, purified and crystallized at 291 K using NaH_2PO_4/KH_2PO_4 as a precipitant. A 2.4 Å resolution native data set was collected from a single flash-cooled crystal (100 K) using 20%(*v*/*v*) glycerol as a cryoprotectant. These crystals belonged to space group $P2_12_12$, with unit-cell parameters a = 53.5, b = 81.7, c = 85.2 Å, $\alpha = \beta = \gamma = 90^\circ$. One molecule is assumed to be present per asymmetric unit, which gives a Matthews coefficient of 2.5 Å^3 Da^{-1}. A data set was also collected to 1.7 Å resolution from a selenomethionyl derivative; it belonged to space group $P2_12_12_1$ with the same unit-cell parameters as the native crystals.

1. Introduction

Cellulases catalyze the hydrolysis of cellulose, an unbranched β-1,4-linked homopolymer of glucose that is a major structural polysaccharide component of the cell wall and thus of plant biomass (Zhang *et al.*, 2006). They are some of the most important industrial enzymes and are already used in large quantities in the textile industry for the bio-polishing of cotton products, which is essential to remove excess fibres and to provide a soft texture (Ljungdahl, 2008). Moreover, there are many additional potential uses of cellulases, *e.g.* for the treatment of biomass to produce sugars that can be converted into bioethanol (Demain *et al.*, 2005) and the treatment of waste water (Bhat & Bhat, 1997). Various endoglucanases and their genes from eukaryotes and bacteria have been investigated (Kashima *et al.*, 2005; Huang *et al.*, 2005; Ando *et al.*, 2002; Halldorsdottir *et al.*, 1998; Bok *et al.*, 1998). *Fervidobacterium nodosum* Rt17-B1, which is classified as a hyperthermophilic hot-spring bacterium with an optimum temperature of 343 K, was first isolated from a hot spring in New Zealand (Andrews & Patel, 1996). In 2007, genome sequencing of *F. nodosum* Rt17-B1 was completed by the US DOE Joint Genome Institute and several endoglucanases have been identified in the genome sequence (accession No. NC_009718). One of them, FnCel5A, was highly thermally active, with a half-life of 48 h at 353 K, and can be thermally activated by heating. Its specific activity was 803.6 U mg^{-1} towards carboxymethylcellulose (CMC) at pH 5.5 and 353 K. The K_m, V_{max} and k_{cat} towards CMC calculated from kinetic analysis were 12.8 mg ml^{-1}, 0.63 µmol min^{-1} and 408.56 s^{-1}, respectively. The optimal substrate of FnCel5A was β-D-glucan and its activity towards this substrate was four times that towards CMC (Feng *et al.*, unpublished work). These properties make FnCel5A an attractive potential target for protein engineering to improve cellulase activity.



The sequence of FnCel5A (Gene ID 5451775; Fnod_1560) shows that the mature FnCel5A protein has 343 residues and consists of a catalytic domain (residues 25–343) and an N-terminal signal peptide (residues 1–24). Of the 113 known families of glycosyl hydrolases (Carbohydrate-Active Enzymes server; http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html; Henrissat & Bairoch, 1993), sequence comparisons clearly place the catalytic domain of FnCel5A in family 5 (Wang *et al.*, 1993). Family 5 (also known as cellulase family A) is the largest of the β-glycohydrolase families classified to date and includes over 60 bacterial and fungal enzymes which all cleave with retention of configuration. The sequences of family 5 cellulases are rather diverse and have been reported to share only seven conserved residues (Wang *et al.*, 1993). Because of this diversity, the cellulases of family 5 have been further subdivided into five subfamilies within which amino-acid sequence similarities are above 25% and homology-modelling methods can be used (Violot *et al.*, 2005; Huang *et al.*, 2005).

To guide protein engineering of *F. nodosum* FnCel5A, we subcloned the FnCel5A gene encoding the thermostable FnCel5A protein using the polymerase chain reaction (PCR) amplification method. In order to understand the structural basis for the thermostability of FnCel5A and to improve its activity guided by the crystal structure, we report here the crystallization and preliminary crystallographic studies of FnCel5A protein.

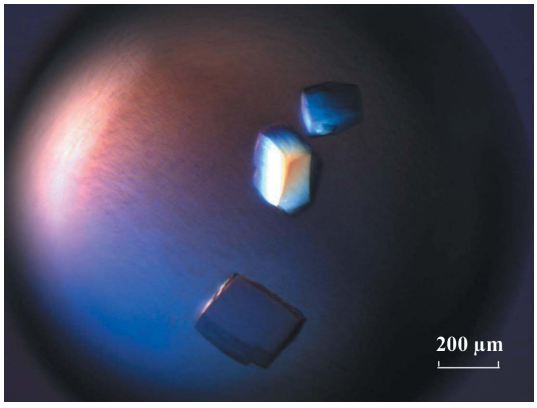


Figure 1
A single crystal of the thermophilic cellulase FnCel5A.

2. Cloning, expression and purification

The primers 5′-CAG CGC **CAT ATG** GAC CAA AGT GTG AGT AA-3′ and 5′-GAC GTC **GGA TCC** TTA TTT TCC AAG TGC AG-3′ were used to amplify the FnCel5A gene from the genome of the thermophilic bacterium *F. nodosum* Rt17-B1. The primers included *Nde*I and *Bam*HI restriction sites (shown in bold). The PCR conditions were 1 min at 367 K, 1 min at 328 K and 2 min at 345 K with 30 cycles. The purified PCR products were digested with *Nde*I and *Bam*HI and then inserted into the expression vector pET-15b with the same digestion sites. After transformation into *Escherichia coli* strain XL1-Blue, the cloned fragments were completely sequenced. The recombinant plasmid was transformed into *E. coli* strain BL21 (DE3). Transformed cells were then cultured at 310 K in LB medium containing 50 µg ml^{-1} ampicillin. When the culture density reached an A_{600} of 0.6–0.8, induction with 1 m*M* IPTG was performed and cell growth continued for 4 h at 310 K. Selenomethionyl (SeMet) labelled FnCel5A was produced in *E. coli* strain B834 (DE3) containing the pET-15b-FnCel5A plasmid with recombinant protein expression induced by 1 m*M* IPTG and incubation at 289 K for 20 h. Cells were harvested by centrifugation, resuspended in 20 m*M* Tris–HCl pH 8.0 buffer containing 10 m*M* NaCl and then homogenized by sonication. Crude bacterial extracts were subjected to heat incubation at 338 K for 30 min and centrifuged at 20 000*g* for 30 min to remove heat-aggregated proteins and cell debris. The supernatant obtained by centrifugation for 30 min at 20 000g was applied onto an Ni^{2+}-chelating affinity column (1.5 ml Ni^{2+}–NTA agarose; Qiagen, USA). Contaminant protein was thoroughly washed off with at least ten bed volumes of wash buffer (20 m*M* Tris–HCl pH 8.0, 10 m*M* NaCl, 20 m*M* imidazole) and the target protein was eluted

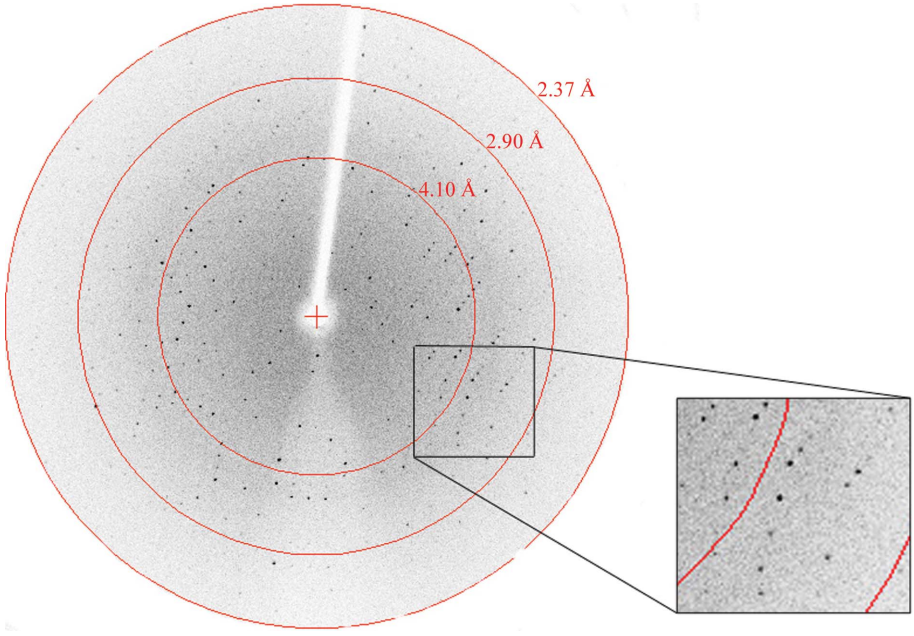


Figure 2
A typical diffraction pattern from a crystal of the native thermophilic cellulase FnCel5A. The exposure time was 120 s, the crystal-to-detector distance was 100 mm and the oscillation range per frame was 1°. The diffraction image was collected on a MAR345dtb image-plate detector.

with 15 ml 20 m*M* Tris–HCl pH 8.0, 10 m*M* NaCl, 200 m*M* imidazole. Resource Q anion-exchange chromatography (GE Healthcare, USA) was subsequently applied using a 0–1 *M* gradient of NaCl in 20 m*M* Tris–HCl pH 8.0 buffer. The target protein was finally eluted with approximately 0.2 *M* NaCl. The purity of FnCel5A was estimated to be greater than 99% by SDS–PAGE analysis.

3. Crystallization

The purified FnCel5A protein was concentrated to ∼20 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0 and 10 m*M* NaCl. Crystallization was performed by the hanging-drop vapour-diffusion method at 291 K in 16-well plates. Each drop contained 1 µl protein solution and 1 µl reservoir solution with 200 µl reservoir solution in the well. Screening was carried out with Hampton Research Crystal Screen kits and positive hits were then optimized. Initial crystals were obtained from condition No. 35 of Crystal Screen I (0.8 *M* sodium phosphate monobasic monohydrate, 0.8 *M* potassium phosphate monobasic, 0.1 *M* Na HEPES pH 6.5) and condition No. 15 of Crystal Screen II (0.5 *M* ammonium sulfate, 1.0 *M* lithium sulfate monohydrate, 0.1 *M* sodium citrate tribasic dihydrate pH 5.6). Using the first condition, a small number of crystals grew in one week from the protein precipitate and were very small with poor X-ray diffraction quality. In contrast, twinned crystals grew rapidly from clear liquor using the second condition. Further crystallization optimization was performed by carefully adjusting the concentration of NaH_2PO_4/KH_2PO_4 and the buffer pH value together with the protein concentration. Fortunately, single crystals were obtained from the optimized reservoir solution [1.2 *M* NaH_2PO_4/KH_2PO_4 (1:1), 0.1 *M* sodium citrate pH 5.6] within 2 d with dimensions of 0.2 × 0.15 × 0.05 mm (Fig. 1) and diffracted to 2.4 Å resolution on the home Rigaku MM-007 X-ray source (Fig. 2). The selenomethionine derivative of FnCel5A (20 mg ml^{-1}) also crystallized using the same condition (Fig. 3) and diffracted to 2.0 Å resolution on the home Rigaku MM-007 X-ray source (Fig. 4).

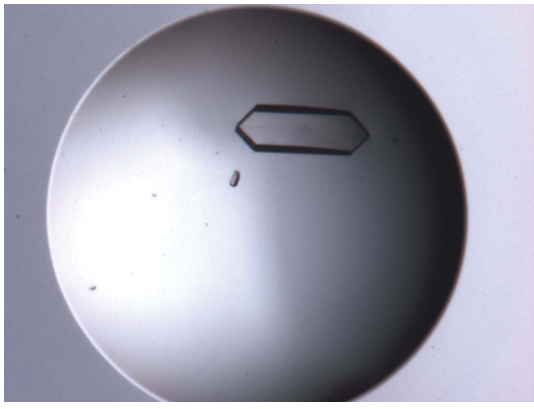


Figure 3
A single crystal of the thermophilic cellulase SeMet-FnCel5A.

4. Data collection and processing

Native diffraction data were collected on a MAR345dtb (MAR Research, Hamburg) image-plate detector at 100 K using a Rigaku MM-007 rotating-anode home X-ray generator operated at 40 kV and 20 mA (λ = 1.5418 Å). The crystal was mounted on a nylon loop and flash-cooled in a nitrogen-gas cryostream at 100 K using an Oxford Cryosystem. Crystals were cryoprotected by the addition of 20%(*v*/*v*) glycerol to the crystallization conditions. A total of 180 frames of data were collected with a 1° oscillation range. All intensity data were indexed, integrated and scaled with the *HKL*-2000 package (Otwinowski & Minor, 1997). The crystal belonged to space group $P2_12_12$, with unit-cell parameters $a = 53.5$, $b = 81.7$, $c = 85.2$ Å, $\alpha = \beta = \gamma = 90°$. We assumed the presence of one molecule per asymmetric unit, which gives a Matthews coefficient of 2.5 Å^3 Da^{-1} with 51% solvent content (Matthews, 1968). Complete data-collection statistics are given in Table 1.

Molecular replacement was performed with the solution structure of *Clostridium thermocellum* endoglucanase CelC (PDB code 1cec;

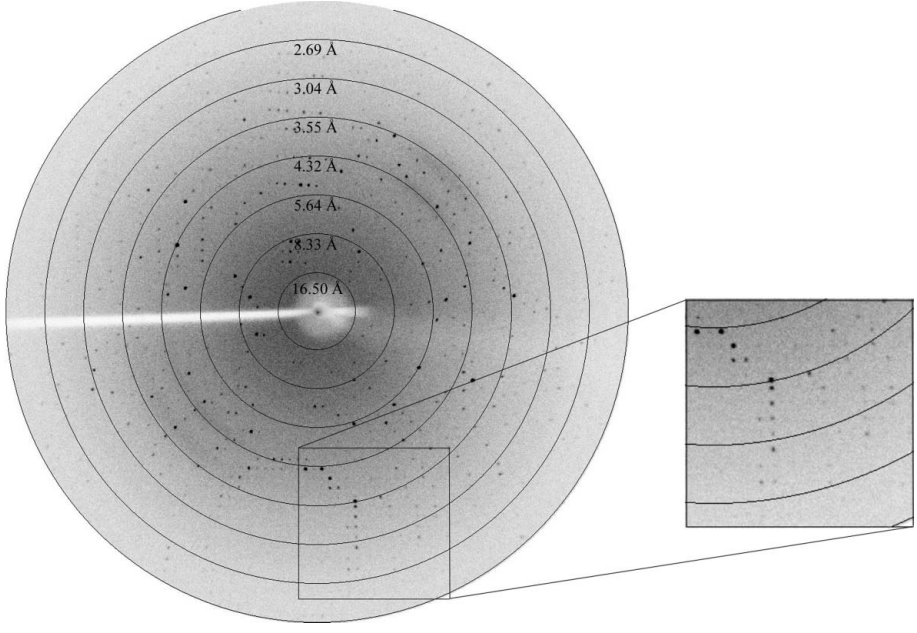


Figure 4
A typical diffraction pattern from a crystal of the thermophilic cellulase SeMet-FnCel5A. The exposure time was 120 s, the crystal-to-detector distance was 100 mm and the oscillation range per frame was 0.5°. The diffraction image was collected on a MAR345 image-plate detector.

Table 1
Data-collection and processing statistics for FnCel5A.

Values in parentheses are for the highest resolution shell.

	Native	SeMet (peak)
Space group	$P2_12_12$	$P2_12_12_1$
Unit-cell parameters (Å)	a = 81.7, b = 85.2, c = 53.5	a = 107.4, b = 82.0, c = 85.8
Wavelength (Å)	1.5418	0.9798
Resolution range (Å)	50.0–2.4 (2.5–2.4)	50.0–1.7 (1.8–1.7)
Total reflections	58334 (5841)	808544 (89834)
Unique reflections	12870 (1298)	122364 (12477)
Redundancy	4.7 (4.5)	6.6 (7.2)
Average $I/\sigma(I)$	9.3 (2.8)	30.7 (4.1)
$R_{\rm merge}$† (%)	13.7 (39.7)	7.2 (40.7)
Data completeness (%)	84.7 (88.4)	96.8 (100.0)
Molecules per ASU	1	2
$V_{\rm M}$ (Å^3 Da^{-1})	2.5	2.7
Sovent content (%)	51	54

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean intensity of the observations $I_i(hkl)$ of reflection hkl.

Dominguez *et al.*, 1995) and the crystal structure of the catalytic domain of endoglucanase A from *C. cellulolyticum* (PDB code 1edg; Ducros *et al.*, 1995) as the initial search models, which showed 24% and 26% sequence similarity to EnCel5, respectively. This procedure was performed using *CNS* v.1.2 (Brünger *et al.*, 1998) and *Phaser* (McCoy *et al.*, 2007), but no obvious correct solution was found according to the rotation and translation functions. The sequences of family 5 members are rather diverse and have been reported to share only seven conserved residues (Wang *et al.*, 1993). This diversity may have caused the failure of molecular replacement for FnCel5A using these search models. Crystals of SeMet-derivative protein were then obtained in order to solve the structure of FnCel5A. Crystals were cryoprotected by the addition of 20%(*v*/*v*) glycerol to the crystallization conditions. The SeMet-derivative crystals belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 107.4$, $b = 82.0$, $c = 85.8$ Å; the space group differed from that of the native crystals and was checked using systematic absences. We assumed the presence of two molecules per asymmetric unit, which gives a Matthews coefficient of 2.7 Å^3 Da^{-1} with 54% solvent content (Matthews, 1968). A 1.7 Å resolution data set was collected from an SeMet-derivative crystal at 100 K using an ADSC Q315 CCD detector on beamline BL5A at the Photon Factory (Japan). Processing of diffraction images and scaling of the integrated intensities were performed using the *HKL*-2000 software package (Otwinowski & Minor, 1997). Six of the eight potential Se atoms in each asymmetric unit were located by *SHELX* and initial phasing and density modification (solvent flipping) were carried out using *SHELXD* and *SHELXE* (Sheldrick, 2008). The final statistics for data collection and processing are summarized in Table 1. Structure determination and refinement is ongoing.

We thank Professor Zihe Rao for generous support and Wen Yang and Zhiyong Lou for their assistance with data collection and technical assistance. This work was supported by the Ministry of Science and Technology of China Project '973' grant 2004CB719606, the Natural Science Foundation of China grants 30670023 and 30870486 and the Excellent Young Teachers Program of the Ministry of Education, China.

References

Ando, S., Ishida, H., Kosugi, Y. & Ishikawa, K. (2002). *Appl. Environ. Microbiol.* **68**, 430–433.
Andrews, K. T. & Patel, B. K. (1996). *Int. J. Syst. Bacteriol.* **46**, 265–269.
Bhat, M. K. & Bhat, S. (1997). *Biotechnol. Adv.* **15**, 583–620.
Bok, J. D., Yernool, D. A. & Eveleigh, D. E. (1998). *Appl. Environ. Microbiol.* **64**, 4774–4781.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Demain, A. L., Newcomb, M. & Wu, J. H. (2005). *Microbiol. Mol. Biol. Rev.* **69**, 124–154.
Dominguez, R., Souchon, H., Spinelli, S., Dauter, Z., Wilson, K. S., Chauvaux, S., Beguin, P. & Alzari, P. M. (1995). *Nature Struct. Biol.* **2**, 569–576.
Ducros, V., Czjzek, M., Belaich, A., Gaudin, C., Fierobe, H. P., Belaich, J. P., Davies, G. J. & Haser, R. (1995). *Structure*, **3**, 939–949.
Halldorsdottir, S., Thorolfsdottir, E. T., Spilliaert, R., Johansson, M., Thorbjarnardottir, S. H., Palsdottir, A., Hreggvidsson, G. O., Kristjansson, J. K., Holst, O. & Eggertsson, G. (1998). *Appl. Microbiol. Biotechnol.* **49**, 277–284.
Henrissat, B. & Bairoch, A. (1993). *Biochem. J.* **293**, 781–788.
Huang, Y., Krauss, G., Cottaz, S., Driguez, H. & Lipps, G. (2005). *Biochem. J.* **385**, 581–588.
Kashima, Y., Mori, K., Fukada, H. & Ishikawa, K. (2005). *Extremophiles*, **9**, 37–43.
Ljungdahl, L. G. (2008). *Ann. NY Acad. Sci.* **1125**, 308–321.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Violot, S., Aghajari, N., Czjzek, M., Feller, G., Sonan, G. K., Gouet, P., Gerday, C., Haser, R. & Receveur-Brechot, V. (2005). *J. Mol. Biol.* **348**, 1211–1224.
Wang, Q., Tull, D., Meinke, A., Gilkes, N. R., Warren, R. A., Aebersold, R. & Withers, S. G. (1993). *J. Biol. Chem.* **268**, 14096–14102.
Zhang, Y.-H. P., Himmel, M. E. & Mielenz, J. R. (2006). *Biotechnol. Adv.* **24**, 452–481.