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

Heterologous expression, crystallization and preliminary X-ray characterization of CcCel6C, a glycoside hydrolase family 6 enzyme from the basidiomycete *Coprinopsis cinerea*

Yuma Kurakata,[a] Takashi Tonozuka,[a] Yuan Liu,[b] Satoshi Kaneko,[c] Atsushi Nishikawa,[a] Kiyoharu Fukuda[b] and Makoto Yoshida[b]*

[a]Department of Applied Biological Science, Tokyo University of Agriculture and Technology, Tokyo 183-8509, Japan, [b]Department of Environmental and Natural Resource Science, Tokyo University of Agriculture and Technology, Tokyo 183-8509, Japan, and [c]Food Biotechnology Division, National Food Research Institute, Ibaraki 305-8643, Japan

Correspondence e-mail: ymakoto@cc.tuat.ac.jp

Received 24 September 2008
Accepted 16 December 2008

CcCel6C is a gene that encodes a glycoside hydrolase family 6 (GH6) enzyme in the *Coprinopsis cinerea* genome. In the evolutionary tree of GH6 enzymes, the encoded enzyme was closely related to Cel6B from *Humicola insolens*, previously called endoglucanase VI, while its amino-acid sequence revealed a region corresponding to the C-terminal active-site-enclosing loop typical of cellobiohydrolase II. Here, the crystallization of CcCel6C produced in *Escherichia coli* is reported. The square prismatic crystal belonged to the triclinic space group $P1$, with unit-cell parameters $a = 44.04$, $b = 45.11$, $c = 48.90$ Å, $\alpha = 77.81$, $\beta = 87.34$, $\gamma = 68.79^\circ$. Diffraction data were collected to 1.6 Å resolution.

© 2009 International Union of Crystallography
All rights reserved

1. Introduction

Cellulose, which consists of linear polymers made up of glucose units linked by β-1,4-glucosidic bonds, is the predominant structural component of plant cell walls and is the most abundant biomass resource on earth. Cellulases hydrolyze the β-1,4-glucosidic bonds of cellulose chains and are traditionally classified as endoglucanases (EGs; EC 3.2.1.4) or cellobiohydrolases (CBHs; EC 3.2.1.91). EGs randomly cleave the internal β-1,4-glucosidic bond of cellulose, whereas CBHs preferentially act on the end of the chain and progressively cleave off cellobiose as the main product. Cellulases have been classified into 12 glycoside hydrolase (GH) families in the CAZy database (http://www.cazy.org/; Henrissat & Bairoch, 1996; Henrissat & Davies, 1997). The members of glycoside hydrolase family 6 (GH6) are broadly divided into CBH-type and EG-type enzymes based on their activity profiles. To date, five three-dimensional structures of GH6 enzymes (two ascomycete CBHIIs, one ascomycete EG and two bacterial EGs) have been reported and their overall structures display a distorted β/α-barrel. The three-dimensional structures of CBHIIs demonstrate that they have active sites enclosed by N-terminal and C-terminal loops that form a tunnel (Rouvinen *et al.*, 1990; Varrot, Hastrup *et al.*, 1999; Varrot, Schülein *et al.*, 1999). The enclosed active sites permit the cleavage of several sequential bonds of the substrate by trapping the cellulose chain in the tunnel, which delays the dissociation of the enzyme and substrate. In contrast, the structures of EGs show active sites in a cleft formed by a C-terminal loop deletion coupled with the peeling open of an N-terminal loop (Davies *et al.*, 2000; Spezio *et al.*, 1993; Varrot *et al.*, 2005).

Recently, we cloned five genes encoding GH6 enzymes from the basidiomycete *Coprinopsis cinerea* (*Coprinus cinereus*) and designated them *CcCel6A*, *CcCel6B*, *CcCel6C*, *CcCel6D* and *CcCel6E* (Yoshida *et al.*, 2009). The amino-acid sequence of CcCel6A is quite similar to those of known CBHIIs, suggesting that the enzyme corresponds to CBHII. The deduced amino-acid sequences of CcCel6B, CcCel6C, CcCel6D and CcCel6E also reveal regions corresponding to the active-site-enclosing loops of CBHIIs. In contrast to CcCel6A, however, CcCel6B, CcCel6C, CcCel6D and CcCel6E formed a clearly distinct group from known CBHIIs in

evolutionary tree analysis; they were located near EGVI from *Humicola insolens*. The features of the amino-acid sequences encouraged us to investigate the enzymatic properties of CcCel6B, CcCel6C, CcCel6D and CcCel6E. Therefore, we attempted heterologous expression of the enzymes using an *Escherichia coli* expression system and recombinant CcCel6C was successfully produced. Here, we report the crystallization and preliminary X-ray characterization of the recombinant enzyme. To our knowledge, this is the first report of the crystallization and X-ray characterization of a basidiomycete GH6 enzyme.

2. Experimental procedures

2.1. Construction of the expression vector

Total RNA was extracted from *C. cinerea* strain 5338 (Yanagi *et al.*, 1988) grown in Kremer and Wood medium (Kremer & Wood, 1992) containing 2% Avicel for 12 d using the RNeasy Plant Mini Kit (Qiagen GmbH, Hilden, Germany). The first-strand cDNA was synthesized with SuperScript III Reverse Transcriptase (Invitrogen, Carlsbad, California, USA) using the oligo(dT) primer (TaKaRa Bio, Shiga, Japan). As presented in Table 1, the oligonucleotide primers were designed based on the cDNA sequence encoding the CcCel6C mature protein (accession No. AB433539). PCR was carried out with the primer pair using KOD-plus-DNA polymerase (Toyobo, Osaka, Japan) as the DNA polymerase and the synthesized first-strand cDNA as a template. The PCR product was subcloned into pGEM-T Easy Vector (Promega, Madison, Wisconsin, USA) followed by sequencing with a 3130 Genetic Analyzer (Applied Biosystems, Foster City, California, USA). The target gene was digested with *Nde*I and *Not*I and ligated into the corresponding restriction site of the pET21a vector (Novagen, Madison, Wisconsin, USA). The nucleotide sequence of the inserted cDNA was confirmed again by sequence analysis.

Table 1
Nucleotide sequences of the primers used in the present study.

The primers contained two thymine residues to prevent the removal of the terminal nucleotide by the exonuclease activity of high-fidelity DNA polymerase (shown in italics), the recognition site of *Nde*I or *Not*I (bold) and a stretch of nucleotide residues encoding the N-terminal or C-terminal amino acids of the mature CcCel6C.

Primer	Sequence
CcCel6C-F	5′-*TT***CATATG**GCTCCCTCTGCCTCCTTC-3′
CcCel6C-R	5′-*TT***GCGGCCGC**CGCGTAAGTAGGCTCAAGCG-3′

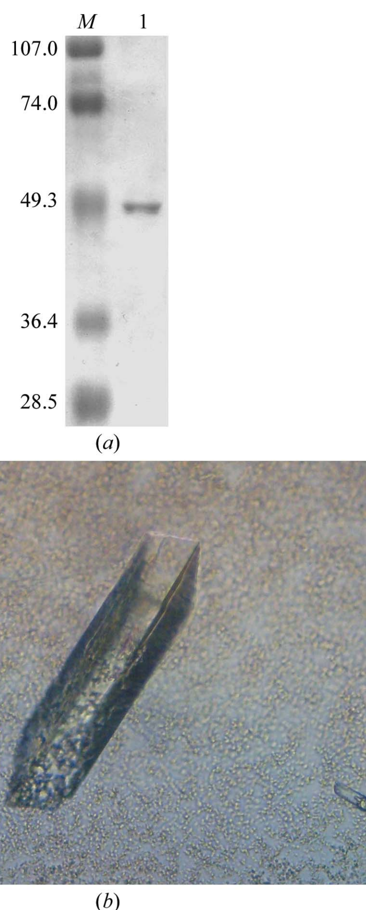

Figure 1
(*a*) SDS–PAGE of purified CcCel6C. The samples were separated on a 10%(*w*/*v*) polyacrylamide gel. Lane *M*, molecular-weight standards; lane 1, recombinant CcCel6C. (*b*) A crystal of CcCel6C grown by hanging-drop vapour diffusion at 293 K. The crystallization drop consisted of 1.0 µl CcCel6C solution at a concentration of 21.5 mg ml^{-1} in 20 m*M* phosphate buffer pH 7.4 and an equal volume of crystallization reservoir solution containing 100 m*M* HEPES–KOH buffer pH 7.0 with 30%(*w*/*v*) PEG 8000 and 150 m*M* magnesium acetate. The dimensions of this crystal were around 0.3 × 0.05 × 0.05 mm.

2.2. Heterologous expression and purification of the recombinant enzyme

E. coli strain BL21 (DE3) (Stratagene, La Jolla, California, USA) was transformed by the obtained plasmid. The transformants were grown in 100 ml LB medium at 310 K until they reached an optical density of 0.2 at 600 nm and expression was induced with 0.2 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) for 24 h at 291 K. The cells were harvested and then resuspended in 5 ml 20 m*M* phosphate buffer pH 7.4 followed by sonication for 10 min on ice. After centrifugation to remove insoluble material, the supernatant was applied onto a 5 × 50 mm Ni–NTA agarose (GE Healthcare, Buckinghamshire, England) column equilibrated with 20 m*M* phosphate buffer pH 7.4 containing 500 m*M* sodium chloride and 20 m*M* imidazole. The column was washed with the same buffer and the recombinant protein was eluted with 20 m*M* phosphate buffer pH 7.4 containing 500 m*M* sodium chloride and 500 m*M* imidazole. The fraction was dialyzed against 20 m*M* phosphate buffer pH 7.4 and the purity of the eluted enzyme was analyzed by SDS–PAGE. The final preparation was used as the purified enzyme.

2.3. Crystallization and data collection

For crystallization trials, CcCel6C solution was prepared at a concentration of 9.96 or 21.5 mg ml^{-1} in 20 m*M* phosphate buffer pH 7.4. All crystallization trials were carried out using the hanging-drop vapour-diffusion method (McPherson, 1982) at 293 K, in which 1.0 µl enzyme solution was mixed with an equal volume of crystallization reservoir solution and each hanging drop was placed over 500 µl reservoir solution. Some crystallization conditions similar to those used for *H. insolens* CBHII (Varrot, Hastrup *et al.*, 1999), which shares 44% amino-acid sequence identity with CcCel6C, were examined. The reservoir solution used for crystallization of *H. insolens* CBHII contained 100 m*M* triethanolamine buffer pH 7.0, 22% PEG 8000 and 200 m*M* magnesium acetate. In the present study, the initial screening, which varied the concentration of the precipitant, the concentration of the salt and the pH and composition of the buffer, was performed with 9.96 mg ml^{-1} CcCel6C solution. In the screening, crystals were obtained under a wide range of conditions and these were optimized manually with 21.5 mg ml^{-1} CcCel6C solution. The best crystal was obtained in a crystallization drop

containing 100 m*M* HEPES–KOH pH 7.0, 30% PEG 8000, 150 m*M* magnesium acetate after one week.

To perform data collection at cryogenic temperatures, the crystals were immersed in a cryoprotectant solution consisting of the well solution and PEG 8000 to a final concentration of 40%(*w*/*v*) and flash-frozen in a stream of nitrogen gas. Diffraction data were collected on the PF-AR NW12 beamline (Photon Factory, Tsukuba, Japan) and the data set was processed and scaled using *HKL*-2000 (Otwinowski & Minor, 1997). Preliminary analysis of the data set was performed using programs from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

We attempted heterologous expression of CcCel6C with a C-terminal His_6 tag using an *E. coli* expression system. To optimize production, *CcCel6C* expression was induced for 4 h at 310 K, for 18 h at 298 K or for 24 h at 291 K using 0.2 m*M* IPTG as an inducer. Recombinant CcCel6C was produced as an insoluble form when it was cultivated for 4 h at 310 K or for 18 h at 298 K. In contrast, soluble CcCel6C was successfully produced on cultivation at 291 K for 24 h. The recombinant enzyme was purified using an Ni–NTA column and the purified CcCel6C produced a single band with molecular mass 48 951 Da, which is higher than that calculated from the amino-acid sequence (43 717 Da), on SDS–PAGE analysis (Fig. 1*a*).

Crystallization of the recombinant enzyme was attempted using crystallization conditions based on those previously reported for *H. insolens* CBHII (Varrot, Hastrup *et al.*, 1999) and square prismatic crystals were observed. The most suitable crystal for diffraction analysis was obtained in 100 m*M* HEPES–KOH buffer pH 7.0 with 30%(*w*/*v*) PEG 8000 and 150 m*M* magnesium acetate. The crystal reached dimensions of 0.3 × 0.05 × 0.05 mm within one week (Fig. 1*b*).

Using synchrotron radiation, the crystal of CcCel6C diffracted to 1.6 Å resolution (Fig. 2) and a complete data set was recorded. Data statistics are given in Table 2. Diffraction analysis showed that the crystal belonged to the triclinic space group *P*1, with unit-cell parameters $a = 44.04$, $b = 45.11$, $c = 48.90$ Å, $\alpha = 77.81$, $\beta = 87.34$, $\gamma = 68.79°$. The Matthews coefficient (V_M; Matthews, 1968) was calculated to be 2.02 $Å^3\ Da^{-1}$, suggesting the presence of one molecule in the asymmetric unit. This V_M value corresponded to a solvent content of approximately 39.2%.

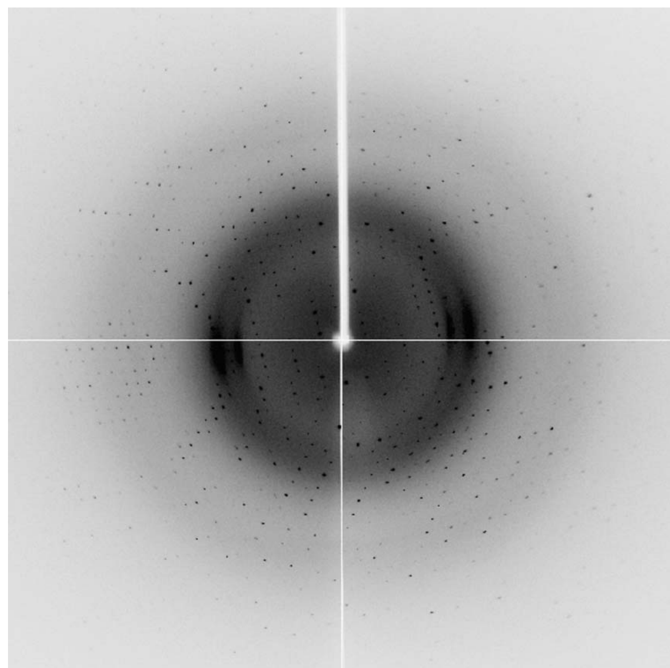

Figure 2
X-ray diffraction pattern to a resolution of 1.6 Å.

Table 2
Diffraction data-collection statistics.

Values in parentheses are for the highest resolution shell.

Beamline	PF-AR NW12
Detector	ADSC Quantum 210
Temperature (K)	100
Wavelength (Å)	1.000
Space group	*P*1
Unit-cell parameters	
a (Å)	44.04
b (Å)	45.11
c (Å)	48.90
α (°)	77.81
β (°)	87.34
γ (°)	68.79
Resolution (Å)	50.0–1.60 (1.66–1.60)
No. of measured reflections	85577
No. of unique reflections	43370
Completeness (%)	95.8 (94.0)
R_{merge} (%)†	2.5 (9.0)
$I/\sigma(I)$	33.2 (9.21)

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

Phase determination was carried out by the molecular-replacement method using the *MOLREP* program in the *CCP*4 suite (Vagin & Teplyakov, 1997), with the *H. insolens* CBHII structure (PDB code 1bvw) and a predicted structural model generated by the SWISS-MODEL automated homology-modelling server of the Swiss Institute of Bioinformatics (http://www.expasy.ch/; Guex & Peitsch, 1997) as the search models. One clear solution was obtained using the predicted structure. Further model building and refinement are in progress.

This research was financially supported in part by the Green Biomass Research for Improvement of Local Energy Self-sufficiency program of the Ministry of Education, Culture, Sports, Science and Technology of Japan and in part by the Development of Biomass Utilization Technologies for Revitalizing Rural Areas program of the Ministry of Agriculture, Forestry and Fisheries of Japan. This research was performed with the approval of the Photon Factory Advisory Committee (2007G010 and 2008G013) of the National Laboratory for High Energy Physics (Tsukuba, Japan).

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Davies, G. J., Brozozowski, A. M., Dauter, M., Varrot, A. & Schülein, M. (2000). *Biochem. J.* **348**, 201–207.
Guex, N. & Peitsch, M. C. (1997). *Electrophoresis*, **18**, 2714–2723.
Henrissat, B. & Bairoch, A. (1996). *Biochem. J.* **316**, 695–696.
Henrissat, B. & Davies, G. J. (1997). *Curr. Opin. Struct. Biol.* **7**, 637–644.
Kremer, S. M. & Wood, P. M. (1992). *Eur. J. Biochem.* **205**, 133–138.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McPherson, A. (1982). *Preparation and Analysis of Protein Crystals*, pp. 82–159. New York: John Wiley.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J. K. & Jones, T. A. (1990). *Science*, **249**, 380–386.
Spezio, M., Wilson, D. B. & Karplus, P. A. (1993). *Biochemistry*, **32**, 9906–9916.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Varrot, A., Hastrup, S., Schülein, M. & Davies, G. J. (1999). *Biochem. J.* **337**, 297–304.
Varrot, A., Leydier, S., Pell, G., Macdonald, J. M., Stick, R. V., Henrissat, B., Gilbert, H. J. & Davies, G. J. (2005). *J. Biol. Chem* **280**, 20181–20184.
Varrot, A., Schülein, M. & Davies, G. J. (1999). *Biochemistry*, **38**, 8884–8891.
Yanagi, S. O., Kawasumi, T., Takebe, I. & Takemaru, T. (1988). *Agric. Biol. Chem.* **52**, 281–284.
Yoshida, M., Sato, K., Kaneko, S. & Fukuda, K. (2009). In the press.