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

Crystallization and preliminary X-ray crystallographic analysis of the β-*N*-acetylglucosaminidase CbsA from *Thermotoga neapolitana*

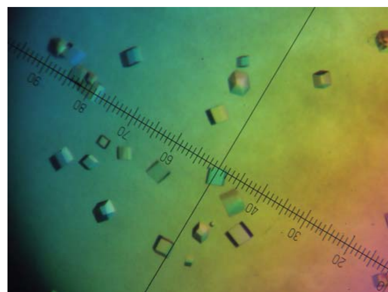
Bo-Young Yoon,[a] Li Jiao,[a] Hyung Ryong Moon,[a] Jaeho Cha[b] and Nam-Chul Ha[a]*

[a]College of Pharmacy and Research Institute for Drug Development, Pusan National University, Busan 609-735, Republic of Korea, and [b]Department of Microbiology, College of Natural Sciences, Pusan National University, Busan 609-735, Republic of Korea

Correspondence e-mail: hnc@pusan.ac.kr

Received 17 October 2011
Accepted 8 November 2011

The β-*N*-acetylglucosaminidase CbsA was cloned from the thermophilic Gram-negative bacterium *Thermotoga neapolitana*. Although CbsA contains a family 3 glycoside hydrolase-type (GH3-type) catalytic domain, it can be distinguished from other GH3-type β-*N*-acetylglucosaminidases by its high activity towards chitobiose. The homodimeric CbsA contains a unique domain at the C-terminus for which the three-dimensional structure is not yet known. In this study, CbsA was overexpressed and the recombinant protein was purified using Ni–NTA affinity and gel-filtration chromatography. The purified CbsA protein was crystallized using the vapour-diffusion method. A diffraction data set was collected to a resolution of 2.0 Å at 100 K. The crystal belonged to space group *R*32. To obtain initial phases, the crystallization of selenomethionyl-substituted protein and the production of heavy-atom derivative crystals are in progress.



© 2012 International Union of Crystallography
All rights reserved

1. Introduction

Thermotoga are thermophilic Gram-negative bacteria that can grow in marine environments at up to 363 K (Belkin *et al.*, 1986). These bacteria utilize various polymeric compounds from the marine environment. In particular, they have many enzymes that hydrolyze diverse carbohydrates, including chitin (a β-1,4-linked polymer of *N*-acetylglucosamine), which is the second most abundant polymer in nature. The genome of *T. maritima* has been sequenced (Nelson *et al.*, 1999) and many genes involved in chitin degradation were found.

In a recent study, *cbsA* was cloned from *T. neapolitana* and the gene product was found to exhibit chitooligomer-degrading activity (β-*N*-acetylglucosaminidase activity; Choi *et al.*, 2009). Sequence analysis suggested that *T. neapolitana* CbsA belongs to the family 3 glycoside hydrolases (GH3), as its N-terminal region (residues 63–280) shows high sequence similarity to GH3 enzymes. The GH3 enzymes have been classified as β-glucosidases, α-L-arabinofuranosidases, β-xylopyranosidases and β-*N*-acetylglucosaminidases with respect to the terminal monosaccharide residue (Harvey *et al.*, 2000). However, CbsA is distinguished from the other members of the GH3-type β-*N*-acetylglucosaminidases by its substrate specificity. CbsA is highly active towards chitobiose, while GH3-type β-*N*-acetylglucosaminidases generally prefer long-chain chitooligosaccharides (Cheng *et al.*, 2000; Chitlaru & Roseman, 1996; Li *et al.*, 2002). CbsA contains an unusual C-terminal region that does not share sequence similarity with other enzymes; this region might be involved in its substrate specificity.

To date, several crystal structures have been determined of GH3 enzymes (Varghese *et al.*, 1999; Litzinger *et al.*, 2010; Yoshida *et al.*, 2010; Pozzo *et al.*, 2010). Whereas the N-terminal catalytic domain is shared among the enzymes, the C-terminal domains are divergent depending on the substrate specificity, and structural data for GH3 representatives are still scarce. In this study, we report the crystallization and preliminary X-ray analysis of full-length CbsA from *T. neapolitana*. The crystal structure of CbsA should allow a deeper structural understanding of GH3-type enzymes and the role of the unusual C-terminal domain.

2. Materials and methods

2.1. DNA construction, protein expression and protein purification

The open reading frame of the *cbsA* gene (residues 1–467; GenBank accession No. AF343913.1) was amplified from the genomic DNA of *T. neapolitana* KCCM41025 using the polymerase chain reaction. The DNA fragment was introduced into the *Nco*I and *Xho*I sites of pPROEX-HTA (Invitrogen, USA), which is an *Escherichia coli* expression vector. The amplified sequence was confirmed using a BigDye terminator cycle sequencing kit and an ABI Prism 3100 genetic analyzer (Perkin Elmer Applied Biosystems, USA). The resulting protein contained three additional amino acids (Gly-His-Met) between the TEV protease cleavage sites as a cloning artifact.

The recombinant CbsA protein was expressed in *E. coli* BL21 (DE3) in 1.5 l LB medium containing ampicillin (50 µg ml^{-1}) at 310 K until the OD_{600} reached 0.6–0.8. Protein expression was induced with 0.5 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) at 303 K. After induction for 4 h, the cells were harvested by centrifugation at 8000*g* for 10 min at 277 K. The harvested cells were disrupted by sonication in lysis buffer consisting of 20 m*M* Tris pH 8.0, 150 m*M* NaCl, 2 m*M* β-mercaptoethanol. The soluble lysate was centrifuged at 25 000*g* for 30 min to remove insoluble materials.

The supernatant was mixed with Ni–NTA affinity resin (Qiagen, Netherlands) that had been pre-incubated with lysis buffer; the mixture was then stirred for 1 h at 277 K. After the slurry has been loaded onto the column, unbound proteins were washed off with lysis buffer supplemented with 20 m*M* imidazole. Recombinant CbsA was eluted with 30 ml lysis buffer supplemented with 250 m*M* imidazole and incubated with recombinant TEV protease at room temperature overnight to remove the hexahistidine tag. The CbsA protein was concentrated using a Centriprep (GE Healthcare, USA) and purified using HiLoad Superdex 200 (GE Healthcare, USA) pre-equilibrated with lysis buffer. The purified protein was concentrated to 10 mg ml^{-1} in 20 m*M* Tris buffer pH 8.0 containing 150 m*M* NaCl and 2 m*M* β-mercaptoethanol and stored at 193 K until use. The protein concentration was determined using a Bio-Rad protein-assay system (Bio-Rad, Hercules, USA) with bovine serum albumin as a standard. The yield was around 10 mg per litre of culture medium.

2.2. Crystallization and data collection

Crystals of CbsA protein were grown at 287 K using the vapour-diffusion method. Crystallization conditions were screened using Crystal Screen HT (Hampton Research, USA) in sitting droplets consisting of 0.5 µl protein solution and 0.5 µl precipitant solution.

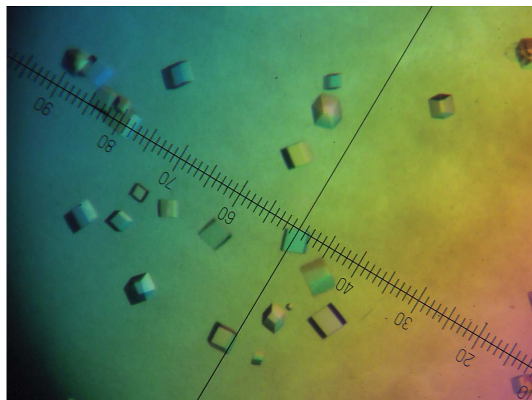


Figure 1
Crystals of *T. neapolitana* CbsA with approximate dimensions 0.1 × 0.1 × 0.1 mm.

Table 1
Diffraction statistics.

Values in parentheses are for the highest resolution shell.

(*a*) Data collection.

X-ray source	BL44, Spring-8
Wavelength (Å)	0.9000
Resolution (Å)	50–2.0 (2.03–2.00)
Space group	*R*32
Unit-cell parameters (Å, °)	$a = b = 158.5$, $c = 517.0$, $\alpha = 90$, $\beta = 90$, $\gamma = 120$
Completeness (%)	93.6 (86.4)
R_{merge}† (%)	10.4 (34.0)
Multiplicity	4.7 (2.6)
Average $I/\sigma(I)$	15.1 (2.2)

(*b*) Cell-content analysis.

No. of molecules in asymmetric unit	4	5	6
V_M (Å^3 Da^{-1})	3.00	2.40	2.00
Solvent content (%)	59.0	48.7	38.5

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ and $\langle I(hkl)\rangle$ are the observed intensities and the mean intensity of related reflections, respectively.

Small cube-shaped crystals grew using a precipitation solution consisting of 0.1 *M* HEPES pH 7.5 buffer containing 1.0 *M* sodium acetate and 0.05 *M* $CdCl_2$. The final crystallization conditions used the same reservoir solution and the hanging-drop vapour-diffusion method in 24-well plates. 1 µl protein solution at a concentration of 10 mg ml^{-1} was mixed with 1 µl reservoir solution and equilibrated against 1 ml reservoir solution. The crystals grew to maximum dimensions (0.1 × 0.1 × 0.1 mm; Fig. 1) within 2 d. X-ray diffraction data were collected from the crystals using a MAR CCD detector on the BL44 beamline at SPring-8, Japan.

3. Results and discussion

CbsA was overexpressed and the recombinant protein was purified using Ni–NTA affinity and gel-filtration chromatography. CbsA eluted as a single peak from the gel-filtration column and appeared to run as a single band on SDS–PAGE with an apparent molecular weight of 52 kDa, which matched the calculated molecular weight (51 370 Da). Crystals suitable for data collection were obtained under the conditions described above (Fig. 1). For data collection under cryogenic conditions, single crystals were soaked in a cryoprotectant solution (reservoir solution supplemented with 25% glycerol) for 1 min prior to flash-cooling in a liquid-nitrogen stream. One set of 360 images (0–360°) was obtained using a 1° oscillation width and 1 s exposure time at 0.9 Å wavelength at 100 K. The collected data were processed and scaled with the *HKL*-2000 package (Otwinowski & Minor, 1997).

Based on the diffraction data, the crystal belonged to space group *R*32, with unit-cell parameters $a = 158.5$, $b = 158.5$, $c = 517.0$ Å in the hexagonal cell. The diffraction data set had a resolution range of 50–2.0 Å with 93.6% completeness and an R_{merge} of 10.4%. Since the self-rotation function from the data set did not indicate the number of molecules per asymmetric unit, we were only able to predict the putative number of molecules in the asymmetric unit from the calculated solvent content. Assuming the presence of five molecules of CbsA in the asymmetric unit, the Matthews coefficient (V_M; Matthews, 1968) was calculated to be 2.40 Å^3 Da^{-1} with a solvent content of 48.7%. Other candidates for the number of molecules per asymmetric unit are listed in Table 1.

In order to solve the structure of CbsA, molecular replacement was attempted using the structures of *Bacillus subtilis* β-*N*-acetylglucosaminidase (PDB entry 3bmx; 80% sequence similarity; Litzinger *et al.*, 2010) and *Vibrio cholerae* GH3 glycoside hydrolase (NagZ; PDB entry 2oxn; 59% sequence similarity; Stubbs *et al.*, 2007) as search models in *MOLREP* (Winn *et al.*, 2011); however, this method was not successful. Given the high methionine frequency in the protein (12 methionine residues in 467 residues), we next intend to obtain experimental phases by the multiple-wavelength anomalous diffraction (MAD) approach using selenomethionine-incorporated crystals. However, crystals of selenomethionine-substituted CbsA have not yet been obtained. In parallel, we are attempting to use the multiple isomorphous replacement or MAD approach to solve the phase problem by producing crystals that are derivatized with an anomalous scatterer or heavy atom.

This work was supported by a Two-Year Research Grant from Pusan National University. This study made use of beamline 44 at SPring-8, Japan.

References

Belkin, S., Wirsen, C. O. & Jannasch, H. W. (1986). *Appl. Environ. Microbiol.* **51**, 1180–1185.

Cheng, Q., Li, H., Merdek, K. & Park, J. T. (2000). *J. Bacteriol.* **182**, 4836–4840.

Chitlaru, E. & Roseman, S. (1996). *J. Biol. Chem.* **271**, 33433–33439.

Choi, K.-H., Seo, J. Y., Park, K.-M., Park, C.-S. & Cha, J. (2009). *J. Biosci. Bioeng.* **108**, 455–459.

Harvey, A. J., Hrmova, M., De Gori, R., Varghese, J. N. & Fincher, G. B. (2000). *Proteins*, **41**, 257–269.

Li, H., Morimoto, K., Katagiri, N., Kimura, T., Sakka, K., Lun, S. & Ohmiya, K. (2002). *Appl. Microbiol. Biotechnol.* **60**, 420–427.

Litzinger, S., Fischer, S., Polzer, P., Diederichs, K., Welte, W. & Mayer, C. (2010). *J. Biol. Chem.* **285**, 35675–35684.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Nelson, K. E. *et al.* (1999). *Nature (London)*, **399**, 323–329.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pozzo, T., Pasten, J. L., Karlsson, E. N. & Logan, D. T. (2010). *J. Mol. Biol.* **397**, 724–739.

Stubbs, K. A., Balcewich, M., Mark, B. L. & Vocadlo, D. J. (2007). *J. Biol. Chem.* **282**, 21382–21391.

Varghese, J. N., Hrmova, M. & Fincher, G. B. (1999). *Structure*, **7**, 179–190.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Yoshida, E., Hidaka, M., Fushinobu, S., Koyanagi, T., Minami, H., Tamaki, H., Kitaoka, M., Katayama, T. & Kumagai, H. (2010). *Biochem. J.* **431**, 39–49.