Min-Kyu Kim,^a Young Jun An,^a
Chang-Sook Jeong,^a Jung Min
Song,^b Mee Hye Kang,^b Yoon-Ho
Lee^b and Sun-Shin Cha^{a,c,d,*}

^aMarine Biotechnology Research Division,
Korea Institute of Ocean Science and
Technology, Ansan 426-744, Republic of Korea,
^bMarine Ecosystem Research Division, Korea
Institute of Ocean Science and Technology,
Ansan 426-744, Republic of Korea, ^cOcean
Science and Technology School, Korea Maritime
University, Pusan 606-791, Republic of Korea,
and ^dDepartment of Marine Biotechnology,
University of Science and Technology,
Daejeon 305-333, Republic of Korea

Correspondence e-mail: chajung@kiost.ac

Received 10 June 2013

Accepted 24 July 2013

Expression at 279 K, purification, crystallization and preliminary X-ray crystallographic analysis of a novel cold-active β -1,4-D-mannanase from the Antarctic springtail *Cryptopygus antarcticus*

The *CaMan* gene product from *Cryptopygus antarcticus*, which belongs to the glycoside hydrolase family 5 type β -1,4-D-mannanases, has been crystallized using a precipitant solution consisting of 0.1 M Tris-HCl pH 8.5, 25% (w/v) polyethylene glycol 3350 by the microbatch crystallization method at 295 K. The *CaMan* protein crystal belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 73.40$, $b = 83.81$, $c = 163.55$ Å. Assuming the presence of two molecules in the asymmetric unit, the solvent content was estimated to be about 61.29%. *CaMan*-mannopentaose (M5) complex crystals that were isomorphous to the *CaMan* crystals were obtained using the same mother liquor containing 1 mM M5.

1. Introduction

Mannan is a major component of hemicelluloses in plants, plant seeds and some algae. There are four kinds of mannan: two linear mannans (pure mannan and glucomannan) and two branched mannans (galactomannan and galactoglucomannan) (Petkiewicz *et al.*, 2001; Moreira & Filho, 2008; Tailford *et al.*, 2009). Each of these polysaccharides presents a β -1,4-linked backbone containing mannoses (pure mannan) or a combination of glucoses and mannoses (glucomannan) (Liepman *et al.*, 2007). In addition, the mannan backbone can be branched with side chains of α -1,6-linked galactose residues (galactomannan or galactoglucomannan). In some cases, such as acemannans, the mannose and glucose residues in the backbone are acetylated at C2 or C3 (Jansson *et al.*, 1975). The degradation of mannans requires the concerted action of a variety of hydrolytic enzymes owing to their structural complexity. β -Mannanases (β -1,4-D-mannan mannohydrolases; EC 3.2.1.78), β -mannosidases (β -1,4-D-mannopyranoside hydrolases; EC 3.2.1.25) and β -glucosidases (β -1,4-D-glucoside glucosylhydrolases; EC 3.2.1.21) degrade the β -1,4-linkage in the mannan backbone. In addition, acetyl mannan esterases (EC 3.1.1.6) and α -galactosidases (α -1,6-D-galactoside galactohydrolases; EC 3.2.1.22) are required to remove the acetyl group and side-chain substituents, respectively, which are attached at various points on the mannan structures (Moreira & Filho, 2008). These enzymes are important for microorganisms and animals that utilize various plant β -mannans as food sources. They have also gained industrial interest owing to their ability to modify and degrade mannan-containing polysaccharides in paper production and the food and feed industries (Marga *et al.*, 1996).

Based on their primary sequences, β -mannanases are grouped mainly into glycoside hydrolase (GH) families 5 and 26, both of which belong to the largest clan GH-A featuring the TIM (triosephosphate isomerase) (β/α)₈-barrel fold (Zhao *et al.*, 2009; Gilbert, 2010). The crystal structures of β -mannanases belonging to both GH families from a wide range of bacteria, fungi, plant and animals have been studied (Hogg *et al.*, 2001; Le Nours *et al.*, 2005; Cartmell *et al.*, 2008; Tailford *et al.*, 2009; Songsiriritthigul *et al.*, 2011; Zhao *et al.*, 2011; Hilge *et al.*, 1998; Sabini *et al.*, 2000; Bourgault *et al.*, 2005; Larsson *et al.*, 2006; Yan *et al.*, 2008; Zhang *et al.*, 2008; Mizutani *et al.*, 2012; Gonçalves *et al.*, 2012; Couturier *et al.*, 2013), revealing an open active-site cleft with at least four subsites and two catalytic gluta-

mates (nucleophiles and acid/base). Ligand-complex structures indicate that GH5 and GH26 β -mannanases have aromatic platforms distributed in the active-site cleft that interact with the hydrophobic α -faces of substrate sugar rings (Gilbert *et al.*, 2008).

The freezing-intolerant Antarctic springtail *Cryptopygus antarcticus* Willem (Collembola, Isotomidae) is the most abundant and widespread terrestrial micro-arthropod in the maritime Antarctic region. The adult reaches approximately 1.2 mm in body length with a live weight of 55–78 μ g and contains 69–85% water. This organism feeds on fungi, unicellular algae and detritus. Therefore, it is thought that *C. antarcticus* contains various kinds of cold-active enzymes that hydrolyse carbohydrates. From an expressed sequence tags (ESTs) library of *C. antarcticus*, we identified the β -1,4-D-mannanase gene (*CaMan*) that belongs to the GH5 family. *CaMan* is a 382-residue protein with a putative signal peptide and exhibits high specific activity towards locust bean gum at an optimal temperature of 303 K and an optimal pH of 3.5 (Song *et al.*, 2008). Its optimal temperature is the lowest among those of the known mannanases and the optimal pH is also the lowest except for those of the *Sclerotium rolfii* and *Aspergillus sulfuris* enzymes (pH 2.9 and pH 2.4, respectively; Gübitz *et al.*, 1996; Chen *et al.*, 2007). Even at 273–278 K, *CaMan* retains 20–40% of its maximum activity. It also shows typical features of a cold-active enzyme, with a high frequency of polar residues such as Asn, Gln and Ser and a low frequency of hydrophobic residues as well as a low ratio of Arg/(Arg+Lys) compared with the mesophilic β -mannanases (Russell, 2000; Margesin & Schinner, 1999). Here, we report the overexpression, purification, crystallization and preliminary X-ray crystallographic analyses of *CaMan*, which is the first β -1,4-D-mannanase from an arthropod and from a multi-cellular animal other than molluscs, and its complex with mannopentaose (M5).

2. Materials and methods

2.1. Expression at 279 K and purification of the *CaMan* protein

The *CaMan* gene (GenBank accession No. ABV68808.1) without the signal peptide-coding region was amplified by the polymerase chain reaction using the EST library of *C. antarcticus* as a template (Song *et al.*, 2008). The gene was inserted downstream of the T7 promoter of the expression plasmid pET-24a (Novagen, Madison, Wisconsin, USA) and the resulting construct expressed residues 22–382 of the *CaMan* protein with a C-terminal noncleavable His₆ tag (LEHHHHHH). After verifying the DNA sequence, plasmid DNA was transformed into *Escherichia coli* strain Rosetta-gami2 (DE3) (Stratagene, La Jolla, California, USA). The cells were grown to an OD₆₀₀ of approximately 0.5 in Luria–Bertani medium containing 50 μ g ml^{−1} kanamycin (Duchefa) at 310 K and expression was induced by 1 mM isopropyl β -D-1-thiogalactopyranoside (IPTG; Duchefa). To solubilize the active *CaMan* protein, we introduced the extremely low temperature induction system (Song *et al.*, 2012). After 7 d induction at 279 K, the cells were harvested and resuspended in 20 mM Tris–HCl pH 7.4 containing 5 mM imidazole. The cells were disrupted by sonication and the cell debris was discarded by centrifugation at 20 000g for 30 min at 277 K. The resulting supernatant was loaded onto TALON metal-affinity resin (Clontech, USA). The column was washed with buffer consisting of 20 mM Tris–HCl pH 7.4, 20 mM imidazole. The *CaMan* protein was eluted with the same buffer containing 300 mM imidazole. The protein fraction partially purified by the TALON resin was dialysed into 20 mM Tris–HCl pH 7.4, 1 mM dithiothreitol (DTT) and loaded onto a Superdex 75 HR 16/60 column (GE Healthcare) pre-equilibrated with buffer

Table 1

Data-collection and processing statistics.

Values in parentheses are for the outermost resolution shell.

	<i>CaMan</i>	<i>CaMan</i> –M5
Beamline	BL-17A, PF	BL-4A, PAL
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 73.40, b = 83.81,$ $c = 163.55$	$a = 74.01, b = 82.31,$ $c = 164.26$
Wavelength (Å)	0.96416	1.0000
Resolution (Å)	50.00–2.35 (2.39–2.35)	50.00–2.40 (2.44–2.40)
Completeness (>0 σ) (%)	96.3 (88.9)	97.7 (96.3)
$R_{\text{merge}}^{\dagger}$ (%)	5.0 (18.4)	9.6 (30.7)
Average $I/\sigma(I)$	24.5 (3.1)	11.7 (1.9)
No. of observations	232051	214319
No. of unique reflections	41345 (1891)	39246 (1907)
Multiplicity	5.6 (4.1)	5.5 (3.5)

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of observed reflection hkl and $\langle I(hkl) \rangle$ is the mean intensity of replicate measurements.

consisting of 20 mM Tris–HCl pH 7.4, 300 mM NaCl, 1 mM DTT (buffer A). The *CaMan* protein eluted at $\sim$ 70.1 min at a flow rate of 1.5 ml min^{−1}. The purified *CaMan* in buffer A was concentrated to approximately 24 mg ml^{−1} for crystallization.

2.2. Crystallization and X-ray data collection

The *CaMan* protein was crystallized and optimized by the microbatch crystallization method at 295 K. Small drops composed of 1 μ l protein solution and an equal volume of crystallization reagent were pipetted under a layer of a 1:1 mixture of silicon oil and paraffin oil in 96-well IMPACT plates (Greiner Bio-One) or 72-well HLA plates (Nunc). Screening for crystallization conditions was performed with all available screening kits from Hampton Research, Axygen Biosciences and Emerald BioSystems. Initial crystals (Fig. 1) were grown in a precipitant solution consisting of 0.1 M Tris–HCl pH 8.5, 25% (w/v) polyethylene glycol (PEG) 3350 (condition No. 45 of Index from Hampton Research). The rod-shaped crystals were produced in 3 d. To obtain the *CaMan*–M5 complex crystal, *CaMan* protein was incubated with 1 mM M5 and crystallized under the same conditions as the native *CaMan* crystal.

The *CaMan* and *CaMan*–M5 crystals were mounted using a nylon loop (50 μ m Mounted CryoLoop from Hampton Research) for data collection and were cooled to 100 K using a Cryostream cooler (Oxford Cryosystems) after slowly concentrating a drop of the mother liquor by leaving the drop open to air and allowing the drop to slowly evaporate for 8 min. A 2.35 Å resolution data set for

Figure 1
Crystal of the *CaMan* protein.

CaMan was collected using an ADSC Quantum 315r CCD on beamline BL-17A of the Photon Factory, Tsukuba, Japan. The exposure time to the synchrotron radiation was 5 s. A total of 191 frames of 1° oscillation were measured with the crystal-to-detector distance set to 285 mm. A 2.40 Å resolution data set for the *CaMan*–M5 complex was collected using an ADSC Quantum 210 CCD on beamline BL-4A at Pohang Light Source, Pohang, Republic of Korea. The exposure time to the synchrotron radiation was 4 s. A total of 240 frames of 1° oscillation were measured with the crystal-to-detector distance set to 350 mm. Diffraction data were processed and scaled using *DENZO* and *SCALEPACK* from the *HKL-2000* program suite (Otwinowski & Minor, 1997). The data-collection statistics are shown in Table 1.

3. Results and discussion

The *CaMan* protein formed inclusion bodies when overexpressed at 288–310 K. This insolubility problem was overcome by simply lowering the temperature to 279 K as previously described (Song *et al.*, 2012). Recombinant *CaMan* protein expressed at this extremely low temperature exhibited mannanase activity (Song *et al.*, 2012) and was successfully crystallized. *CaMan* shows the highest specific

activity towards locust bean gum and is also active towards konjac and β -1,4-mannan (Song *et al.*, 2008). It also digests mannotetraose (M4) and mannopentaose (M5) to give mannobiose (M2) and mannotriose (M3) with a little production of mannose (M1). However, *CaMan* hardly degraded M2 and M3 (Song *et al.*, 2008). To elucidate the structural basis for its catalytic features, we tried to obtain native *CaMan* and *CaMan*–M5 complex crystals. The *CaMan* crystal belonged to the primitive orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 73.40$, $b = 83.81$, $c = 163.55$ Å. The crystal volume per unit molecular weight (V_M) was about $3.20 \text{ Å}^3 \text{ Da}^{-1}$ with a solvent content of 61.29% by volume (Matthews, 1968) when the asymmetric unit was assumed to contain two molecules. The *CaMan*–M5 complex crystal belonged to the primitive orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 74.01$, $b = 82.31$, $c = 164.26$ Å. The crystal volume per unit molecular weight (V_M) was about $3.18 \text{ Å}^3 \text{ Da}^{-1}$ with a solvent content of 61.29% by volume when the asymmetric unit was assumed to contain two molecules (Table 1).

CaMan shows 41, 40, 41 and 40% sequence identity to other invertebrate β -mannanases from the abalone *Haliotis discus hannai* (*HdMan*), the common sea hare *Aplysia kurodai* (*AkMan*), the freshwater snail *Biomphalaria glabrata* (*BgMan*) and the blue mussel *Mytilus edulis* (*MeMan5A*), respectively. The primary sequence of *CaMan* aligns well with these β -mannanases, all of which belong

Figure 2 Sequence alignment of *C. antarcticus* β -mannanase (*CaMan*) with those of *H. discus hannai* (*HdMan*), *A. kurodai* (*AkMan*), *B. glabrata* (*BgMan*) and *M. edulis* (*MeMan5A*). Amino acids highlighted with triangles indicate the catalytic glutamates (black), active-site residues (blue), conserved glycines (green), other highly conserved residues (orange) and GH5 subfamily 10-specific tryptophans (purple). *CaMan* has an extended loop region (349–352; black box).

to GH5 subfamily 10. However, this enzyme shows low sequence identities ranging from 13.8 to 8% to other β -mannanases from plants, fungi and bacteria. The sequence alignment reveals that CaMan has an extended loop region (349–352; Fig. 2). Hitherto, two crystal structures of invertebrate mollusc β -mannanases [AkMan (PDB entry 3vup; Mizutani *et al.*, 2012) and MeMan5A (PDB entry 2c0h; Larsson *et al.*, 2006)] have been available. According to these structures, the extended loop of CaMan is highly likely to be located at the entrance of the active-site cleft. Structural or sequential modifications of loops in the active site have been reported to affect the discrimination of mannan substrates (Gonçalves *et al.*, 2012; Cartmell *et al.*, 2008). From this perspective, the extended loop of CaMan in the active-site cleft is likely to be critical to its catalytic features. Molecular replacement was performed with *MOLREP* (Vagin & Teplyakov, 2010) using the structure of MeMan5A (Larsson *et al.*, 2006) as a search model. Using the apo CaMan data set, the rotation search, which was performed within the resolution range 45.46–2.73 Å, yielded two top solutions with R_f/σ of 9.22 and 7.90 (in contrast to 5.85 for the third solution), indicating the presence of two monomers per asymmetric unit. The best solution after the translation function had an R factor of 65.60%. Restrained refinement of this solution with *REFMAC5* (Murshudov *et al.*, 2011) significantly lowered the R factors to $R = 37.93\%$ and $R_{\text{free}} = 44.12\%$ with a figure of merit of 0.57 in the 45.80–2.35 Å resolution range. Using the CaMan–M5 data set, the rotation search, which was performed within the resolution range 35.97–2.56 Å, yielded two top solutions with R_f/σ of 8.08 and 6.95 (in contrast to 4.27 for the third solution), indicating the presence of two monomers per asymmetric unit. Restrained refinement of this solution with *REFMAC5* (Murshudov *et al.*, 2011) lowered the R factors to $R = 38.76\%$ and $R_{\text{free}} = 44.37\%$ with a figure of merit of 0.57 in the 35.99–2.40 Å resolution range. From difference Fourier maps of the CaMan–M5 complex, we could identify M5 in the active site. To our knowledge, this is the first structure of a GH5 β -mannanase in complex with M5 and thus will provide detailed insights into the substrate-recognition mechanism and enzymatic features of CaMan as well as the functional implications of the extended loop. Manual model improvement and refinement are now in progress.

We thank the staff of beamlines BL-4A at Pohang Light Source, Republic of Korea and BL-17A at the Photon Factory, Japan for data-collection support. This study was supported by the National Research Foundation of Korea grant No. 2012005978, the CAP through KRCF, KIST and KIOST, and the KIOST in-house program (PE98933).

References

- Bourgault, R., Oakley, A. J., Bewley, J. D. & Wilce, M. C. (2005). *Protein Sci.* **14**, 1233–1241.
- Cartmell, A., Topakas, E., Ducros, V. M., Suits, M. D., Davies, G. J. & Gilbert, H. J. (2008). *J. Biol. Chem.* **283**, 34403–34413.
- Chen, X., Cao, Y., Ding, Y., Lu, W. & Li, D. (2007). *J. Biotechnol.* **128**, 452–461.
- Couturier, M., Roussel, A., Rosengren, A., Leone, P., Stålbrand, H. & Berrin, J. G. (2013). *J. Biol. Chem.* **288**, 14624–14635.
- Gilbert, H. J. (2010). *Plant Physiol.* **153**, 444–455.
- Gilbert, H. J., Stålbrand, H. & Brumer, H. (2008). *Curr. Opin. Plant Biol.* **11**, 338–348.
- Gonçalves, A. M. D., Silva, C. S., Madeira, T. I., Coelho, R., de Sanctis, D., San Romão, M. V. & Bento, I. (2012). *Acta Cryst.* **D68**, 1468–1478.
- Gübitz, G., Hayn, M., Sommerauer, M. & Steiner, W. (1996). *Bioresour. Technol.* **58**, 127–135.
- Hilge, M., Gloor, S. M., Rypniewski, W., Sauer, O., Heightman, T. D., Zimmermann, W., Winterhalter, K. & Piontek, K. (1998). *Structure*, **6**, 1433–1444.
- Hogg, D., Woo, E. J., Bolam, D. N., McKie, V. A., Gilbert, H. J. & Pickersgill, R. W. (2001). *J. Biol. Chem.* **276**, 31186–31192.
- Jansson, P. E., Kenne, L. & Lindberg, B. (1975). *Carbohydr. Res.* **45**, 275–282.
- Larsson, A. M., Anderson, L., Xu, B., Muñoz, I. G., Usón, I., Janson, J. C., Stålbrand, H. & Ståhlberg, J. (2006). *J. Mol. Biol.* **357**, 1500–1510.
- Le Nours, J., Anderson, L., Stoll, D., Stålbrand, H. & Lo Leggio, L. (2005). *Biochemistry*, **44**, 12700–12708.
- Liepmann, A. H., Nairn, C. J., Willats, W. G., Sørensen, I., Roberts, A. W. & Keegstra, K. (2007). *Plant Physiol.* **143**, 1881–1893.
- Marga, F., Ghakis, C., Dupont, C., Morosoli, R. & Kluepfel, D. (1996). *Appl. Environ. Microbiol.* **62**, 4656–4658.
- Margesin, R. & Schinner, F. (1999). *Cold-Adapted Organisms: Ecology, Physiology, Enzymology and Molecular Biology*. Berlin, New York: Springer.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Mizutani, K., Tsuchiya, S., Toyoda, M., Nanbu, Y., Tominaga, K., Yuasa, K., Takahashi, N., Tsuji, A. & Mikami, B. (2012). *Acta Cryst.* **F68**, 1164–1168.
- Moreira, L. R. & Filho, E. X. (2008). *Appl. Microbiol. Biotechnol.* **79**, 165–178.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Petkowicz, C. L. de O., Reicher, F., Chanzy, H., Taravel, F. R. & Vuong, R. (2001). *Carbohydr. Polym.* **44**, 107–112.
- Russell, N. J. (2000). *Extremophiles*, **4**, 83–90.
- Sabini, E., Schubert, H., Murshudov, G., Wilson, K. S., Siika-Aho, M. & Penttilä, M. (2000). *Acta Cryst.* **D56**, 3–13.
- Song, J. M., An, Y. J., Kang, M. H., Lee, Y.-H. & Cha, S.-S. (2012). *Protein Expr. Purif.* **82**, 297–301.
- Song, J. M., Nam, K.-W., Kang, S. G., Kim, C.-G., Kwon, S.-T. & Lee, Y.-H. (2008). *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **151**, 32–40.
- Songsiririthigul, C., Lapboonrueng, S., Roytrakul, S., Haltrich, D. & Yamabhai, M. (2011). *Acta Cryst.* **F67**, 217–220.
- Tailford, L. E., Ducros, V. M., Flint, J. E., Roberts, S. M., Morland, C., Zechel, D. L., Smith, N., Bjørnvad, M. E., Borchert, T. V., Wilson, K. S., Davies, G. J. & Gilbert, H. J. (2009). *Biochemistry*, **48**, 7009–7018.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Yan, X.-X., An, X.-M., Gui, L.-L. & Liang, D.-C. (2008). *J. Mol. Biol.* **379**, 535–544.