Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of family 39 β-D-xylosidase from *Geobacillus stearothermophilus* T-6

Mirjam Czjzek,[a]* Tsafrir Bravman,[b] Bernard Henrissat[a] and Yuval Shoham[b]

[a]Architecture et Fonction des Macromolécules Biologiques, CNRS–Université de Provence, Université d'Aix-Marseille II, IBSM, 31 Chemin Joseph-Aiguier, 13402 Marseille CEDEX 20, France, and [b]Department of Biotechnology and Food Engineering and Institute of Catalysis Science and Technology, Technion–Israel Institute of Technology, Haifa 32000, Israel

Correspondence e-mail: czjzek@afmb.cnrs-mrs.fr

Received 12 November 2003
Accepted 14 January 2004

β-D-Xylosidases (EC 3.2.1.37) are hemicellulases that hydrolyze short xylooligosaccharides into single xylose units. In this study, the crystallization and preliminary X-ray analysis of the β-D-xylosidase (XynB1) from *Geobacillus stearothermophilus* T-6, a family 39 glycoside hydrolase, are described. XynB1 is a tetrameric protein consisting of four identical subunits of 503 amino acids and with a calculated molecular weight of 58 001 Da. Both the native and the selenomethionine-containing XynB1 were crystallized by the hanging-drop vapour-diffusion method and the crystals were found to belong to space group $P2_12_12_1$, with unit-cell parameters a = 92.7, b = 165.7, c = 311.0 Å. The native crystals diffracted X-rays to a resolution of 2.1 Å.

1. Introduction

Xylans, the major hemicellulosic polysaccharides in the plant cell wall, are composed of a β-1,4-linked D-xylopyranosyl backbone substituted with glucuronosyl, arabinosyl and acetyl side chains. Because of the structural complexity of xylan, its complete breakdown requires the synergistic action of several hemicellulolytic enzymes with diverse specificity (Beg *et al.*, 2001; Shallom & Shoham, 2003). Among others, these include endo-1,4-β-xylanases (EC 3.2.1.8), which hydrolyze the xylan backbone, and β-D-xylosidases (EC 3.2.1.37), which cleave the resulting xylooligomers to free xylose. Hemicellulases are vital for maintaining the carbon cycle in nature, since they are responsible for the complete degradation of the plant biomass to soluble saccharides, which in turn can be utilized as carbon and energy sources for microorganisms and higher animals. Interest in xylan-degrading enzymes stems from their potential applications in the paper and pulp industry for biobleaching and for the bioconversion of lignocellulose material to fermentative products (Beg *et al.*, 2001; Galbe & Zacchi, 2002; Mielenz, 2001; Shallom & Shoham, 2003). Recently, these enzymes as well as other glycosidases were demonstrated to be potential glycosynthases for oligosaccharide and thioglycoside synthesis (Jahn *et al.*, 2003; Mackenzie *et al.*, 1998).

Hemicellulases, in common with all glycoside hydrolases, hydrolyze the glycosidic bond either by retention or inversion of the anomeric configuration of the substrate. Inverting glycosidases use a single-displacement mechanism with the assistance of general acid and general base residues. Retaining glycosidases follow a two-step double-displacement mechanism involving two key active-site residues, one functioning as the nucleophile and the other as the acid/base (Davies *et al.*, 1998).

Based on sequence similarities, β-D-xylosidases are currently assigned to glycoside hydrolase families 3, 39, 43, 52 and 54 (Henrissat & Bairoch, 1996; Henrissat & Davies, 1997). These families (together with all other glycoside hydrolase families) are constantly updated at the Carbohydrate-Active Enzymes server (http://afmb.cnrs-mrs.fr/CAZY). The β-D-xylosidases from families 39 and 43 have been shown to cleave the glycosidic bond with retention and inversion of the anomeric configuration, respectively (Armand *et al.*, 1996; Braun *et al.*, 1993). The enzymes of family 43 are probably the most studied β-D-xylosidases; the three-dimensional structure of a member of this family was recently determined and revealed a unique five-bladed β-propeller fold (Nurizzo *et al.*, 2002). The stereochemical hydrolysis course of family 52 enzymes proceeds *via* retention of the anomeric configuration (Bravman, Zolotnitsky *et al.*, 2001) and a detailed kinetic analysis together with the identification of the catalytic pair of this family has been described very recently (Bravman, Belakhov *et al.*, 2003; Bravman, Zolotnitsky *et al.*, 2003). Family 39 enzymes belong to the largest glycoside hydrolase clan (GH-A), in which all its members are believed to share a TIM-barrel $(\beta/\alpha)_8$ structure. Enzymes of this clan are sometimes described as the 4/7 superfamily, because the proton donor and nucleophile are found on strands 4 and 7 of the $(\beta/\alpha)_8$ barrel, respectively (Henrissat *et al.*, 1995). In addition to xylosidases, family 39 also contains

α-L-iduronidases, which are involved in the lysosomal degradation of glycosaminoglycans in higher organisms. Deficiencies in these enzymes can lead to several severe ailments, since accumulation of partially degraded glycosaminoglycans results in mucopolysaccharidosis type I, known as Hurler or Scheie syndrome (Unger *et al.*, 1994). The nucleophile of family 39 was identified in *Thermoanaerobacterium saccharolyticum* β-D-xylosidase and in human α-L-iduronidase using different mechanism-based inactivators (Nieman *et al.*, 2003; Vocadlo *et al.*, 1998). The acid–base catalyst of this family was identified through detailed kinetic analysis of catalytic mutants (Bravman, Mechaly *et al.*, 2001).

As part of our ongoing efforts to study the hemicellulolytic system in *Geobacillus stearothermophilus* T-6, we have recently cloned and sequenced a 23.5 kbp chromosomal segment that contains a cluster of genes involved in the degradation of xylan. This region includes a β-D-xylosidase gene (*xynB1*) showing homology to family 39 glycoside hydrolases. The *xynB1* gene is part of an operon responsible for the intake and catabolism of aldotetraouronic acid [2-*O*-α-(4-*O*-methyl-α-D-glucuronosyl)-xylotriose] formed by the primary degradation of xylan (Shulami *et al.*, 1999). The *xynB1* gene encodes a 503-residue polypeptide chain with a calculated molecular weight of 58 001 Da and, based on gel filtration, the protein consists of four identical subunits.

Two β-xylosidases from glycoside hydrolase families 3 and 39 have previously been crystallized (Yang *et al.*, 2002; Golubev *et al.*, 2000). Here, we describe the crystallization and preliminary X-ray analysis of the family 39 β-D-xylosidase from *G. stearothermophilus* T-6. The three-dimensional structure of this enzyme will provide crucial and complementary information about the catalytic mechanism and binding properties of family 39 glycoside hydrolases. Two β-xylosidases from glycoside hydrolase families 3 and 39 have previously been crystallized (Yang *et al.*, 2002; Golubev *et al.*, 2000) and more recently the crystal structure of one member of familly 39 has been reported (Yang *et al.*, 2004).

2. Experimental

2.1. Overexpression and purification of XynB1

The *xynB1* gene (GenBank accession No. AF098273) from *G. stearothermophilus* T-6 was cloned into the pET-9d vector and overexpressed in *Escherichia coli* BL21 (DE3). The purification procedure included two steps, heat treatment at 333 K and gel filtration, as previously described (Bravman, Mechaly *et al.*, 2001) and resulted in near-gram quantities of >99% purified enzyme. Production of the selenomethionine derivative of XynB1 was carried out in the methionine-auxotrophic *E. coli* strain B834 (DE3) (Novagen), essentially as described previously (Mechaly *et al.*, 2000), and protein purification proceeded as for the wild-type enzyme.

2.2. Protein crystallization

The protein was kept at a concentration of 3.5 mg ml^{-1} in 50 m*M* Tris buffer pH 7, 100 m*M* NaCl and 0.02% NaN_3. Initial crystallization trials were performed with MDL (Molecular Dimensions Limited) and Decode Genetics (Wizard 1 and 2, Emerald Biostructures) screens; that is, a total of 192 trials in two 96-well crystallization plates from Greiner. These trials were set up using a Cartesian crystallization robot, mixing 300, 200 and 100 nl of protein solution with 100 nl of reservoir solution. The most promising initial crystallization conditions were refined through variation of the precipitant, additive and protein concentration, pH and drop volume.

From these optimized conditions, crystals were grown by the hanging-drop vapour-diffusion method using 24-well Linbro plates. Hanging drops were prepared on siliconized cover slips by mixing 2 µl of the protein solution with 1 µl of reservoir solution consisting of 100 m*M* HEPES pH 7.5, 0–100 m*M* NaCl and 17–22%(*w*/*v*) PEG 8000. They were then placed over 0.5 ml of the reservoir solution. Crystals of the SeMet-substituted protein were subsequently grown by hanging-drop vapour-diffusion starting from similar conditions to the native protein. The optimized conditions for the SeMet-substituted protein are 100 m*M* HEPES pH 7.5 and 16–19%(*w*/*v*) PEG 8000 for hanging drops containing 2 µl protein solution and 1 µl reservoir solution.

2.3. Data collection and processing

X-ray diffraction data were collected from a native XynB1 crystal at 100 K on beamline ID14-EH2 at the ESRF (Grenoble, France) using an ADSC Quantum 4R CCD detector. All crystals were flash-cooled in a liquid-nitrogen stream with 15%(*v*/*v*) glycerol as a cryoprotectant. The wavelength of the synchrotron X-rays was 0.933 Å. The crystal was rotated through 110° with a 0.8° oscillation range per frame. Attempts to collect X-ray diffraction data from the SeMet-substituted XynB1 cryocooled crystals were made at beamlines ID29 and ID14-EH4. However, the crystals suffered from radiation damage and data could only be collected at one wavelength. The wavelength of the synchrotron X-rays at ID14-EH4 was 0.9793 Å. The crystal was rotated through 110° with a 0.5° oscillation range per frame. A second data set at beamline ID29 was collected at a wavelength of 0.9796 Å; the oscillation angle was again 0.5° and a total of 110° was covered. All raw data were processed using the program *MOSFLM* (Leslie, 1990). The resultant data were merged and scaled using the program *SCALA* (Collaborative Computational Project, Number 4, 1994).

3. Results

XynB1 as well as SeMet-substituted XynB1 have been overexpressed in soluble form in sufficient quantities for crystallization. The optimized reservoir conditions are 100 m*M* HEPES pH 7.5, 50 m*M* NaCl, 20%(*w*/*v*) PEG 8000 for XynB1 and 100 m*M* HEPES pH 7.5 with 18%(*w*/*v*) PEG 8000 for SeMet-XynB1. Crystals of native XynB1 grew to maximum dimensions of 0.08 × 0.2 ×

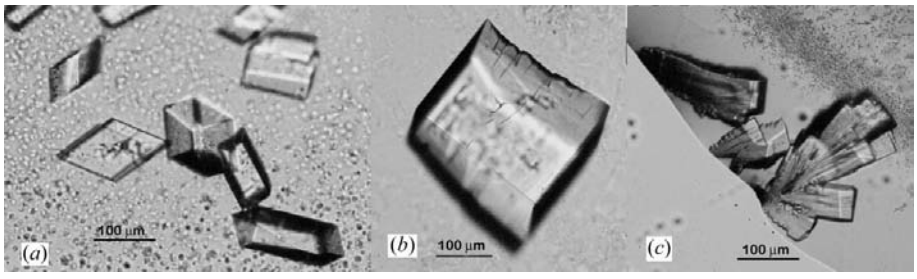

Figure 1
(*a*) Regular-shaped crystals of native XynB1 grown using 100 m*M* HEPES pH 7.5, 100 m*M* NaCl, 22%(*w*/*v*) PEG 8000. (*b*) A single crystal of native XynB1 grown using 100 m*M* HEPES pH 7.5, 50 m*M* NaCl and 20%(*w*/*v*) PEG 8000. Its approximate dimensions are 0.08 × 0.2 × 0.15 mm. (*c*) Cluster of crystals of SeMet-XynB1 grown from 100 m*M* HEPES pH 7.5 and 18%(*w*/*v*) PEG 8000.

Table 1
Data-collection statistics.

Values in parentheses correspond to the highest resolution shell.

	XynB1	SeMet-XynB1	SeMet-XynB1
ESRF beamline	ID14-EH2	ID14-EH4	ID29
Wavelength (Å)	0.933	0.9793	0.9796
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)			
a	92.7	92.6	92.6
b	165.7	165.6	165.6
c	311.0	310.6	310.6
Resolution (Å)	24.9–2.1 (2.2–2.1)	29.9–3.0 (3.2–3.0)	29.8–3.0 (3.2–3.0)
No. observations	1009913	337917	339457
No. unique reflections	273112	88124	87790
R_{sym}† (%)	15.5 (44)	7.9 (12.0)	9.8 (15.4)
R_{ano}‡ (%)	—	9.2 (9.8)	10.1 (15.2)
$\langle I/\sigma(I)\rangle$	3.5 (1.9)	7.0 (5.0)	5.7 (4.4)
Redundancy	3.7 (3.7)	3.8 (3.8)	3.9 (2.6)
Completeness (%)	98.2 (98.2)	91.8 (72.5)	91.1 (66.7)

† $R_{sym} = \sum |I - I_{av}| / \sum I$, where the summation is over all symmetry-equivalent reflections. ‡ $R_{ano} = \sum |I(+) - I(-)| / \sum |I_{av}$, where I_{av} is the average of Friedel amplitudes at a single wavelength.

0.15 mm within two weeks, while the SeMet-XynB1 crystals had maximum dimensions of $0.06 \times 0.06 \times 0.1$ mm and took at least three weeks to grow (Fig. 1).

The native crystals diffracted to 1.9 Å resolution; however, owing to one particularly large unit-cell parameter (c = 311 Å), the geometric set-up for data collection was optimized to have data complete to 2.1 Å resolution. The MAD data set was only complete to 3.0 Å resolution. The data-collection statistics for a native data set and a two-wavelength data set of SeMet-XynB1 are summarized in Table 1. Both the native and the SeMet-substituted crystals belong to space group $P2_12_12_1$. Biochemical characterization revealed that the protein forms tetramers. Accordingly, two tetrameric molecules of XynB1 are present in the asymmetric unit, giving a crystal volume per protein mass (V_M) of 2.57 Å^3 Da^{-1} and a solvent content of 52% by volume (Matthews, 1968). The number of SeMet substitutions has been verified by electrospray mass spectroscopy. All eight methionines present in the primary structure have been substituted. The unit cell therefore contains a total of 64 Se atoms. Crystal structure determination using the SAD (SeMet) method is currently under way.

This research was supported by grants from GIF, the German–Israeli Foundation for Scientific Research and Development (to YS), and the French–Israeli Association for Scientific and Technological Research (AFIRST, to YS and BH), Jerusalem, Israel. Additional support was provided by the Fund for the Promotion of Research at the Technion and by the Otto Meyerhof Center for Biotechnology, established by the Minerva Foundation (Munich, Germany).

References

Armand, S., Vieille, C., Gey, C., Heyraud, A., Zeikus, J. G. & Henrissat, B. (1996). *Eur. J. Biochem.* **236**, 706–713.
Beg, Q. K., Kapoor, M., Mahajan, L. & Hoondal, G. S. (2001). *Appl. Microbiol. Biotechnol.* **56**, 326–338.
Braun, C., Meinke, A., Ziser, L. & Withers, S. G. (1993). *Anal. Biochem.* **212**, 259–262.
Bravman, T., Belakhov, V., Solomon, D., Shoham, G., Henrissat, B., Baasov, T. & Shoham, Y. (2003). *J. Biol. Chem.* **278**, 26742–26749.
Bravman, T., Mechaly, A., Shulami, S., Belakhov, V., Baasov, T., Shoham, G. & Shoham, Y. (2001). *FEBS Lett.* **495**, 115–119.
Bravman, T., Zolotnitsky, G., Belakhov, V., Shoham, G., Henrissat, B., Baasov, T. & Shoham, Y. (2003). *Biochemistry*, **42**, 10528–10536.
Bravman, T., Zolotnitsky, G., Shulami, S., Belakhov, V., Solomon, D., Baasov, T., Shoham, G. & Shoham, Y. (2001). *FEBS Lett.* **495**, 39–43.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Davies, G., Sinnott, M. L. & Withers, S. G. (1998). *Comprehensive Biological Catalysis*, edited by M. L. Sinnott, pp. 119–209. London: Academic Press.
Galbe, M. & Zacchi, G. (2002). *Appl. Microbiol. Biotechnol.* **59**, 618–628.
Golubev, A. M., Brandao Neto, J. R., Eneyskaya, E. V., Kulminskaya, A. V., Kerzhner, M. A., Neustroev, K. N. & Polikarpov, I. (2000). *Acta Cryst.* D**56**, 1058–1060.
Henrissat, B. & Bairoch, A. (1996). *Biochem. J.* **316**, 695–696.
Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J. P. & Davies, G. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 7090–7094.
Henrissat, B. & Davies, G. (1997). *Curr. Opin. Struct. Biol.* **7**, 637–644.
Jahn, M., Marles, J., Warren, R. A. & Withers, S. G. (2003). *Angew. Chem. Int. Ed. Engl.* **42**, 352–354.
Leslie, A. G. W. (1990). *Crystallographic Computing*, edited by D. Moras, A. D. Podjarny & J. C. Thierry, pp. 50–61. Oxford University Press.
Mackenzie, L. F., Wang, Q., Warren, R. A. & Withers, S. G. (1998). *J. Am. Chem. Soc.* **120**, 5583–5584.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mechaly, A., Teplitsky, A., Belakhov, V., Baasov, T., Shoham, G. & Shoham, Y. (2000). *J. Biotechnol.* **78**, 83–86.
Mielenz, J. R. (2001). *Curr. Opin. Microbiol.* **4**, 324–329.
Nieman, C. E., Wong, A. W., He, S., Clarke, L., Hopwood, J. J. & Withers, S. G. (2003). *Biochemistry*, **42**, 8054–8065.
Nurizzo, D., Turkenburg, J. P., Charnock, S. J., Roberts, S. M., Dodson, E. J., McKie, V. A., Taylor, E. J., Gilbert, H. J. & Davies, G. J. (2002). *Nature Struct. Biol.* **9**, 665–668.
Shallom, D. & Shoham, Y. (2003). *Curr. Opin. Microbiol.* **6**, 219–228.
Shulami, S., Gat, O., Sonenshein, A. L. & Shoham, Y. (1999). *J. Bacteriol.* **181**, 3695–3704.
Unger, E. G., Durrant, J., Anson, D. S. & Hopwood, J. J. (1994). *Biochem. J.* **304**, 43–49.
Vocadlo, D. J., MacKenzie, L. F., He, S., Zeikus, G. J. & Withers, S. G. (1998). *Biochem. J.* **335**, 449–455.
Yang, J. K., Yoon, H.-J., Ahn, H. J., Lee, B. I., Kim, H.-W., Laivenieks, M., Vieille, C., Zeikus, J. G. & Suh, S. W. (2002). *Acta Cryst.* D**58**, 531–532.
Yang, J. K., Yoon, H.-J., Ahn, H. J., Lee, B. I., Pedelacq, J. D., Liong, E. C., Berendzen, J., Laivenieks, M., Vieille, C., Zeikus, G. J., Vocadlo, D. J., Withers, S. G. & Suh, S.-W. (2004). *J. Mol. Biol.* **335**, 155–165.