Molecular cloning, overexpression, purification and crystallographic analysis of a GH43 β -xylosidase from *Bacillus licheniformis*

José Alberto Diogo,^a Leticia Maria Zanphorlin,^a Hélia Harumi Sato,^b Mario Tyago Murakami^{c*} and Roberto Ruller^{a*}

Received 16 January 2015

Accepted 22 May 2015

Edited by B. Hazes, University of Alberta, Canada

Keywords: β -xylosidase; *Bacillus licheniformis*; GH43 family.

^aBrazilian Bioethanol Science and Technology Laboratory (CTBE), National Center for Research in Energy and Materials (CNPem), Giuseppe Maximo Scolfaro 10000, 13083-100 Campinas-SP, Brazil, ^bDepartment of Food Science, University of Campinas, Rua Monteiro Lobato 80, Cidade Universitária, 13081-970 Campinas-SP, Brazil, and ^cBrazilian Biosciences National Laboratory (LNBio), National Center for Research in Energy and Materials (CNPem), Giuseppe Maximo Scolfaro 10000, 13083-100 Campinas-SP, Brazil. *Correspondence e-mail: mario.murakami@lnbio.cnpem.br, roberto.ruller@bioetanol.org.br

β -Xylosidases (EC 3.2.1.37) catalyze the hydrolysis of short xylooligosaccharides into xylose, which is an essential step in the complete depolymerization of xylan, the major hemicellulosic polysaccharide of plant cell walls, and has great biotechnological relevance for the production of lignocellulose-based biofuels and the paper industry. In this study, a GH43 β -xylosidase identified from the bacterium *Bacillus licheniformis* (BlXylA) was cloned into the pET-28a bacterial expression vector, recombinantly overexpressed in *Escherichia coli* BL21(DE3) cells and purified to homogeneity by metal-affinity and size-exclusion chromatography. The protein was crystallized in the presence of the organic solvent 2-methyl-2,4-pentanediol and a single crystal diffracted to 2.49 Å resolution. The X-ray diffraction data were indexed in the monoclinic space group C2, with unit-cell parameters $a = 152.82$, $b = 41.9$, $c = 71.79$ Å, $\beta = 91.7^\circ$. Structural characterization of this enzyme will contribute to a better understanding of the structural requirements for xylooligosaccharide specificity within the GH43 family.

1. Introduction

Xylan is the most abundant hemicellulosic polysaccharide found in plant cell walls and consists of a main chain of β -1,4-linked xylosyl residues, which can be further substituted at the 2-OH and 3-OH positions with a variety of side groups such as arabinose, acetic acid, D-glucuronic acid, *p*-coumaric acid and ferulic acid (Rennie & Scheller, 2014).

The depolymerization of xylan to free xylose has a great impact in the production of biofuels by pentose-fermenting microorganisms (Dodd & Cann, 2009) and in a number of other industrial processes such as the paper, animal-feed and food industries (Zhao *et al.*, 2013; Driss *et al.*, 2014; Valenzuela *et al.*, 2014). The complete enzymatic hydrolysis of xylan requires the synergistic action of several enzymes to cope with the distinct chemical linkages and side groups. For the xylan backbone in particular, endo- β -1,4-xylanases (E.C. 3.2.1.8) are required to cleave internal β -1,4-glycosidic bonds, generating xylooligosaccharides, and β -xylosidases are required to catalyze the breakdown of short xylooligosaccharides into xylose.

These enzymes also complement the activity of cellulases, improving the lignocellulose hydrolysis efficiency (Hu *et al.*, 2011; Gao *et al.*, 2011) and consequently contributing to C6-fermenting technologies. β -Xylosidases, in particular, have an important role in alleviating the inhibition of cellulases by xylooligosaccharides during biomass degradation. On the

other hand, analogously to β -glucosidases, β -xylosidases are also competitively inhibited by high concentrations of free xylose in the course of enzymatic hydrolysis. Therefore, for industrial applications it is of great interest to have β -xylosidases with a high tolerance to xylose (Yan *et al.*, 2008; Shi *et al.*, 2013) and with a stability compatible with the harsh conditions typically found in biotechnological processes involving biomass deconstruction (Bhalla *et al.*, 2013).

These enzymes have been identified, purified and characterized from a number of microorganisms, including actinomycetes, eubacteria and fungi (Holtz *et al.*, 1991; Collins *et al.*, 2005; Polizeli *et al.*, 2005), and they have been found in GH families 1, 3, 30, 31, 39, 43, 51, 52, 54, 116 and 120 (Cantarel *et al.*, 2009). Despite the large expansion of characterized enzymes with this activity, the structural requirements for xylooligosaccharide specificity among the different GH families are still not fully understood and neither is the molecular basis for xylose tolerance. In this study, we report the cloning, overexpression, purification, crystallization and preliminary X-ray analysis of a β -xylosidase belonging to the GH43 family from the bacterium *Bacillus licheniformis* (BIXylA). This enzyme shares approximately 45% sequence similarity with other GH43 β -xylosidases; however, the loops that are considered to be important for substrate binding are not conserved in BIXylA, indicating a distinct functional behaviour, which is being investigated along with this structural characterization.

2. Materials and methods

2.1. Macromolecule production

The open reading frame (ORF) encoding a β -xylosidase belonging to the GH43 family (accession code WP_011197672; Table 1) was amplified from the genomic DNA of *B. licheniformis* by touchdown PCR using Phusion DNA polymerase (Thermo Scientific). The amplified ORF was cloned into the pGEM-T Easy vector (Promega) and further subcloned into the pET-28a bacterial expression vector between the *NheI* and *EcoRI* restriction sites. The resulting construct consisted of the full-length BsXylA sequence with an N-terminal hexahistidine tag.

The recombinant plasmid BIXylA-pET-28a was transformed into *Escherichia coli* strain BL21(DE3) and the cells were grown in selective Luria broth (LB) medium at 310 K and 250 rev min⁻¹ to an A_{600} of ~0.8. Expression was then induced with 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) at 310 K for 4 h. The cells were harvested by centrifugation for 30 min at 8000g and resuspended in lysis buffer [50 mM phosphate pH 7.4, 150 mM NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.5 mg ml⁻¹ lysozyme]. The cell suspension was lysed by sonication and the supernatant was separated by centrifugation for 30 min at 277 K and 12 000g.

The clarified supernatant was subjected to metal-affinity chromatography using a 5 ml HisTrap column (GE Healthcare) pre-equilibrated with 50 mM phosphate pH 7.4, 150 mM

Table 1

Macromolecule-production information.

The restriction sites in the primers and the His₆ tag in the protein sequence are underlined.

Source organism	<i>B. licheniformis</i> (ATCC 14580)
DNA source	<i>B. licheniformis</i> (ATCC 14580)
Forward primer	5'-ATATATGCTAGCATGAGCGGTGAACACACAT- A-3'
Reverse primer	5'-TATATAGAAATCTCTTAATATGCAATATAGCGG- A-3'
Cloning vector	pGEM-T Easy (Promega)
Expression vector	pET-28a (Novagen)
Expression host	<i>E. coli</i> BL21(DE3)
Complete amino-acid sequence of the construct produced	MGSSHHHHHHSSGLVPRGSHMASMSGEHTYTNPV- LTGFHPDPSIIRVGEDYVMNSTFYQYFPAIVI- SHSKDLVHWKIIGHGITENEGLDLSINDSHG- IWAPDISYHNGTFYIFATHRLNGPTVINGRKL- IRRQIMIKSSRPEGYPYKPVFIDEGSGIDPSH- FVDGDKHYMLLSPACTLFLPLNEECTDISGEP- VQIWEGTGRRAPEGPHLLKKDGYYYAILAEGG- TGYSHSITTAARSTHLYGYPYPCYPNPLTQTD- PDAPIQRAGHGKLVETQNGEWAVYLCGRPNQ- GSYTTVGRETALDPVEWTDGWFVINNLKGPS- LVQRANPLPQVKWDEKNFDDFEDTLGLDWQF- VRNPDHSSWSLIERPGYLRLLWTGDWDLHDIRA- KNTVVRREKHHLYSAGVKLDFSPSASGEQAGI- VCYYSTNNYKCLLYEEGLKIKVVENRSGCQ- KTLGKKHAEAGPLFLKAVINKQRDFYYSYEG- KHHWHAGGTEDASFLSDEGSRDAKGHTGTMTVG- IFANNGGSGRKAADFDFWFRYIAY

NaCl, 20 mM imidazole. The bound proteins were eluted using a 0–0.5 M imidazole gradient. Fractions containing the target protein were pooled and concentrated to a final volume of 2 ml for subsequent size-exclusion chromatography on a Superdex 200 column which had been equilibrated with 50 mM phosphate pH 7.4, 150 mM NaCl.

All purification steps were performed using an ÄKTA system (GE Healthcare) and fractions were analyzed by SDS-PAGE (Laemmli, 1970).

The purified BIXylA was concentrated to 15 mg ml⁻¹ using Amicon centrifugal filter units (Millipore) for crystallization experiments. The protein concentration was determined by absorption ($A_{280\text{ nm}}$) using a Nanodrop spectrophotometer (Thermo Scientific) and a protein molar extinction coefficient of 119 220 M⁻¹ cm⁻¹.

Sample monodispersity was checked by dynamic light scattering using a Zetasizer Nano Series (Malvern Instruments Ltd) device. 20 measurements were taken at 291 K and the hydrodynamic parameters were determined using the *Zetasizer Software* v.6.32 (Malvern Instruments Ltd). The hydrodynamic radius (R_h) was extrapolated from the translational diffusion coefficient (D_t) using the Stokes–Einstein equation.

2.2. Crystallization

Initial screening was performed with approximately 600 different crystallization conditions based on the commercial kits Crystal Screen and Crystal Screen 2 (Hampton Research), Wizard I and II (Emerald Bio), JCSG+ (Emerald Bio), Precipitant Synergy (Emerald Bio), PACT (Qiagen) and SaltRx (Hampton Research). The sitting-drop experiments were prepared with a Cartesian Honeybee 963 robot

Table 2
Crystallization.

Method	Sitting-drop vapour diffusion
Plate type	Corning 96-well protein crystallization plates
Temperature (K)	291
Protein concentration (mg ml ⁻¹)	15
Buffer composition of protein solution	20 mM phosphate pH 7.4
Composition of reservoir solution	36% (v/v) 2-methyl-2,4-pentanediol, 5% (w/v) polyethylene glycol 8000, 0.1 M sodium cacodylate pH 6.5
Volume and ratio of drop	1 µl drops, 1:1 ratio
Volume of reservoir (µl)	80

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	W01B-MX2 beamline, LNLS
Wavelength (Å)	1.458
Temperature (K)	100
Detector	Pilatus 2M
Crystal-to-detector distance (mm)	180
Rotation range per image (°)	0.1
Total rotation range (°)	180
Exposure time per image (s)	2
Space group	C2
<i>a</i> , <i>b</i> , <i>c</i> (Å)	152.82, 41.90, 71.79
α , β , γ (°)	90.00, 91.72, 90.00
Mosaicity (°)	0.56
Resolution range (Å)	30.0–2.49
Total No. of reflections	42353
No. of unique reflections	15259
Completeness (%)	94.1 (85.3)
Multiplicity	2.78 (2.16)
$\langle I/\sigma(I) \rangle$	9.6 (1.8)
R_{meas}	0.109 (0.715)
R_{merge}	0.089 (0.554)
Overall <i>B</i> factor from Wilson plot (Å ²)	20.5

(Genomic Solutions) and were systematically analyzed with the assistance of a Rock Imager 1000 automated imaging system (Formulatrix). A single condition [40% (v/v) 2-methyl-2,4-pentanediol, 5% (w/v) polyethylene glycol 8000, 0.1 M sodium cacodylate pH 6.5] yielded small crystals after 10 d. Crystals with adequate dimensions for X-ray diffraction analysis on the MX2 beamline at LNLS, Campinas, Brazil were obtained by reducing the precipitant to 36% in steps of 0.5% (v/v) (Table 2).

2.3. Data collection and processing

For X-ray diffraction data collection, a BIXylA crystal was soaked in a cryoprotectant solution consisting of the reservoir solution supplemented with 20% (v/v) glycerol and was further flash-cooled to 100 K in a nitrogen-gas stream. Shutterless data collection was performed with a single-photon-counting hybrid pixel detector using the fine φ -slicing strategy to optimal data quality. The data were indexed, integrated and scaled using the *XDS* package (Kabsch, 2010*a,b*) taking into account the $CC_{1/2}$ criterion (Karplus & Diederichs, 2012) for high-resolution data cutoff. Additional data-collection information and processing statistics are summarized in Table 3.

3. Results and discussion

The recombinant full-length BIXylA with an N-terminal His₆ tag was purified to homogeneity by two chromatographic steps

Figure 1
The Ni²⁺-chelating affinity (*a*) and size-exclusion (*b*) chromatographic steps employed for purification of BIXylA. An SDS-PAGE analysis of each purification step is shown as an inset.

Figure 2
Photomicrograph of a BIXylA crystal.

Figure 3
X-ray diffraction pattern of a BlXylA crystal collected on the MX2 beamline at LCLS, Campinas, Brazil. The inset is a magnification of the region of the diffraction pattern in the red box.

(Fig. 1), yielding approximately 5 mg target protein per litre of expression medium. The few contaminants that remained from Ni^{2+} -chelating affinity chromatography (Fig. 1a) were successfully removed by a gel-filtration purification step (Fig. 1b), which improved the reproducibility of BlXylA crystallization. The purified protein remained monodisperse in the crystallization concentration (15 mg ml^{-1}) as assessed by dynamic light scattering ($R_h = 4.3 \text{ nm}$ and polydispersity of 25%) and only crystallized in the presence of the organic solvent 2-methyl-2,4-pentanediol. Despite the imperfection of the crystal surface as can be observed in the microphotograph (Fig. 2), the crystal diffracted to $2.49 \text{ \AA}$ resolution and yielded an interpretable diffraction pattern (Fig. 3). The crystal belonged to the base-centred monoclinic space group $C2$ characterized by $\beta = 91.72^\circ$. The calculated Matthews coefficient of $1.90 \text{ \AA}^3 \text{ Da}^{-1}$ corresponds to a solvent content of 35.41% (Matthews, 1968), assuming the molecular weight of the recombinant BlXylA to be 60 387 Da. This analysis statistically confirms the presence of one protomer in the asymmetric unit. BlXylA shares a maximum sequence identity of 45% with homologous β -xylosidases belonging to the GH43 family, and molecular-replacement calculations resulted in an initial model with an R factor of 0.40 and an R_{free} of 0.46,

which were further improved by isotropic restrained refinement to an R factor of 0.34 and an R_{free} of 0.39, indicating it to be a correct solution. The structural characterization of this enzyme combined with site-directed mutagenesis and functional studies will shed light on the molecular requirements for xylooligosaccharide specificity within the enigmatic GH43 family, which have great biotechnological relevance in the conversion of lignocellulosic materials into biofuels and other chemicals.

Acknowledgements

This work was supported by grants from Fundação de Amparo a Pesquisa do Estado de São Paulo (MTM; 2014/07135-1 and 2013/13309-0) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (RR; 481666/2012-5). JAD is the recipient of a FAPESP fellowship (2013/10443-7).

References

- Bhalla, A., Bansal, N., Kumar, S., Bischoff, K. M. & Sani, R. K. (2013). *Bioresour. Technol.* **128**, 751–759.
- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). *Nucleic Acids Res.* **37**, D233–D238.
- Collins, T., Gerday, C. & Feller, G. (2005). *FEMS Microbiol. Rev.* **29**, 3–23.
- Dodd, D. & Cann, I. K. (2009). *Glob. Change Biol. Bioenergy*, **1**, 2–17.
- Driss, D., Driss, Z., Chaari, F. & Chaabouni, S. E. (2014). *Bioengineered*, **5**, 274–279.
- Gao, D., Uppugundla, N., Chundawat, S. P., Yu, X., Hermanson, S., Gowda, K., Brumm, P., Mead, D., Balan, V. & Dale, B. E. (2011). *Biotechnol. Biofuels*, **4**, 5.
- Holtz, C., Kaspary, H. & Klemme, J. H. (1991). *Antonie Van Leeuwenhoek*, **59**, 1–7.
- Hu, J., Arantes, V. & Saddler, J. N. (2011). *Biotechnol. Biofuels*, **4**, 36.
- Kabsch, W. (2010a). *Acta Cryst.* **D66**, 125–132.
- Kabsch, W. (2010b). *Acta Cryst.* **D66**, 133–144.
- Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.
- Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Polizeli, M. L., Rizzatti, A. C., Monti, R., Terenzi, H. F., Jorge, J. A. & Amorim, D. S. (2005). *Appl. Microbiol. Biotechnol.* **67**, 577–591.
- Rennie, E. A. & Scheller, H. V. (2014). *Curr. Opin. Biotechnol.* **26**, 100–107.
- Shi, H., Li, X., Gu, H., Zhang, Y., Huang, Y., Wang, L. & Wang, F. (2013). *Biotechnol. Biofuels*, **6**, 27.
- Valenzuela, S. V., Diaz, P. & Pastor, F. I. (2014). *Appl. Microbiol. Biotechnol.* **98**, 5949–5957.
- Yan, Q. J., Wang, L., Jiang, Z. Q., Yang, S. Q., Zhu, H. F. & Li, L. T. (2008). *Bioresour. Technol.* **99**, 5402–5410.
- Zhao, L., Meng, K., Bai, Y., Shi, P., Huang, H., Luo, H., Wang, Y., Yang, P., Song, W. & Yao, B. (2013). *J. Agric. Food Chem.* **61**, 6880–6889.