Evandro Ares de Araújo,^a Atílio Tomazini Jr,^a Marco Antonio Seiki Kadowaki,^a Mário Tyago Murakami^b and Igor Polikarpov^{a*}

^aInstituto de Física de São Carlos, Universidade de São Paulo, 400 Trabalhador Saocarlense, 13566-590 São Carlos-SP, Brazil, and

^bLaboratório Nacional de Biociências, CNPEM, 10000 Giuseppe Máximo Scolfaro, 13083-100, Campinas-SP, Brazil

Correspondence e-mail: ipolikarpov@ifsc.usp.br

Received 2 April 2013

Accepted 9 May 2013

Crystallization and preliminary X-ray diffraction analysis of a new xyloglucanase from *Xanthomonas campestris* pv. *campestris*

Xyloglucanases (Xghs) are important enzymes involved in xyloglucan modification and degradation. *Xanthomonas campestris* pv. *campestris* (Xcc) is a phytopathogenic bacterium which produces a large number of glycosyl hydrolases (GH), but has only one family 74 GH (Xcc-Xgh). This enzyme was overexpressed in *Escherichia coli*, purified and crystallized. Diffraction data sets were collected for the native enzyme and its complex with glucose to maximum resolutions of 2.0 and 2.1 Å, respectively. The data were indexed in a hexagonal crystal system with unit-cell parameters $a = b = 153.4$, $c = 84.9$ Å. As indicated by molecular-replacement solution, the crystals belonged to space group $P6_1$.

1. Introduction

Exploration of new energy sources is a scientific and technologic challenge to modern societies. Renewable energy is an important alternative to fossil fuels (Goldemberg, 2008; Balat & Balat, 2009). Vegetal biomass is a renewable, sustainable and inexpensive material suitable for second-generation bioethanol production. The plant cell wall is constituted of cellulose fibres embedded in a polymer matrix and is composed of cellulose (~15–40%), hemicellulose-pectin (~30–40%) and lignin (~20%) (Lynd *et al.*, 2002; Himmel *et al.*, 2007). Many microorganisms, such as bacteria and fungi, produce extracellular enzymes that are capable of degrading plant cell wall material by converting cellulose and hemicelluloses to fermentable sugars (Chundawat *et al.*, 2011).

As part of the hemicellulose fraction of biomass, xyloglucan forms a complex network which functions as a bridge between cellulose microfibrils, giving rigidity and extension capacity to the cell walls of dicotyledons and some monocotyledons (Martinez-Fleites *et al.*, 2006). This branched polymer is composed of a cellulose-like backbone of β -(1,4)-glucose with side chains of α -D-xylose, xylose-galactose or xylose-galactose-fucose and their diversity is plant dependent (Hayashi, 1989; Peng *et al.*, 2012; Park & Cosgrove, 2012). Xyloglucan is a structural element that hinders access of enzymes to cellulose and thus biomass conversion by cellulases (Ding *et al.*, 2012; Dixon, 2013). Thus, the efficient removal of xyloglucan is required for complete lignocellulosic biomass decomposition by the concerted action of cellulases and hemicellulases (Hayashi, 1989; Scheller & Ulvskov, 2010).

Xyloglucan-specific endo- β -1,4-glucanases or xyloglucanases (EC 3.2.1.151) are grouped into glycoside hydrolase families 5, 12, 16, 44 and 74 in the Carbohydrate-Active enZymes Database (CAZy; <http://www.cazy.org/>; Cantarel *et al.*, 2009; Gilbert *et al.*, 2008; Damásio *et al.*, 2012). Glycoside hydrolase family 74 (GH74) enzymes exhibit high specificity towards xyloglucan, except for Cel74 from *Thermotoga maritima*, which has been characterized as a cellobiohydrolase (Chhabra & Kelly, 2002).

It has been reported that endo- β -1,4-glucanases display hydrolytic activity towards xyloglucan and oligoxyloglucan (Yaoi & Mitsuishi, 2002; Bauer *et al.*, 2005; Desmet *et al.*, 2007; Irwin *et al.*, 2003; Yaoi *et al.*, 2004, 2009; Martinez-Fleites *et al.*, 2006; Hasper *et al.*, 2002; Ichinose *et al.*, 2012; Enkhbaatar *et al.*, 2012). From a structural point of view, this is a poorly studied group of enzymes, with only two crystal structures available in the Protein Data Bank (PDB): one from *Clostridium thermocellum* F7/YS (PDB entries 2cn2 and 2cn3;

Martinez-Fleites *et al.*, 2006) and one from *Geotrichum* sp. M128 (PDB entries 1sqj, 2ebs and 3aof; Yaoi *et al.*, 2004, 2007, 2009).

Xanthomonas campestris pv. *campestris* (*Xcc*) is a Gram-negative bacterium that is able to synthesize a wide range of extracellular enzymes which hydrolyse different kinds of substrates. Analysing the genomic sequence data of *Xcc* strain ATCC33913 in CAZY (Cantarel *et al.*, 2009), the uncharacterized gene *xcc1752* (GenBank NP_637119.1) was identified. This gene was annotated as a putative cellulase (81.14 kDa) belonging to GH74.

We set out to determine the three-dimensional structure of this xyloglucanase from *Xcc* (*Xcc*-Xgh) in the hope that its structural characterization will contribute to the catalogue of GH74 structures and might provide insights into the molecular basis of the mode of action of bacterial xyloglucanases on xyloglucan polysaccharides.

2. Materials and methods

2.1. Protein preparation

The coding region of the *Xcc*-Xgh protein was amplified by polymerase chain reaction (PCR) from *Xcc* genomic DNA using the primer pair *Xcc*_1752F (5'-ACATATGGCCACGTCCGGG-CCCTACCAGTG-3') and *Xcc*_1752R (5'-AGAGCTCATGGC-ATGGCCACGTCCGGGCCCTACCAG-3') containing *Nde*I and

*Sac*I restriction sites (shown in bold), respectively. The 2136 bp amplified fragment was introduced into pTZ57R/T vector (Fermentas, USA) and then subcloned into pSMT3 expression vector.

The recombinant plasmid pSMT3-*Xcc*-Xgh was transformed into *Escherichia coli* BL21 (DE3) cells and induced using 1 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) at 293 K for 24 h. Cells were harvested by centrifugation at 6000g and resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl). Cells were lysed by sonication and the supernatant was centrifuged at 1800g to remove cell debris. The supernatant was subjected to affinity purification using Ni-NTA resin (Qiagen) equilibrated with 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 20 mM imidazole buffer. The *Xcc*-Xgh protein eluted at approximately 300 mM imidazole in a 50–300 mM imidazole gradient. The protein was dialysed against buffer consisting of 20 mM Tris-HCl pH 8.0, 150 mM NaCl and then subjected to ubiquitin-like protein-specific protease 1 (ULP1) digestion at 277 K for 24 h to remove the His tag. The resulting mixture was again loaded onto Ni-NTA resin and untagged *Xcc*-Xgh was collected in the flowthrough. Fractions containing the target protein were loaded onto a HiLoad Superdex 75 16/60 size-exclusion chromatography column (GE Healthcare) pre-equilibrated with 20 mM Tris-HCl pH 8.0, 150 mM NaCl using an ÄKTA purifier 10 system (GE Healthcare Life Sciences) to remove aggregates and impurities from the protein of interest.

The final protein sample has >95% purity as assessed by 15% SDS-PAGE stained with Coomassie Brilliant Blue (Laemmli, 1970). The purified *Xcc*-Xgh was concentrated to 42 mg ml⁻¹ in 20 mM Tris-HCl pH 8.0, 50 mM NaCl for crystallization assays. The protein concentration was determined by measuring absorption at 280 nm using a NanoDrop spectrophotometer (Thermo Scientific).

2.2. Crystallization and data collection

Initial crystallization screens were performed using a Honeybee automated robotic system at the Laboratório Robotizado de Cristalização de Proteínas (Robolab-LNLS, Campinas, Brazil) using 480 different reservoir conditions from kits from Hampton Research (Laguna Niguel, California, USA) and Emerald BioSystems (Bainbridge Island, Washington, USA), including Crystal Screen, Crystal Screen 2, Wizard I and II, and SaltRx. The sitting-drop vapour-diffusion method was used, in which 1 μ l *Xcc*-Xgh solution (in 20 mM Tris-HCl pH 8.0, 50 mM NaCl) was mixed with 1 μ l reservoir solution in 96-well Cryschem Plates (Hampton Research). For data collection, a single apoprotein crystal was transferred to a cryoprotectant solution containing 15% (v/v) ethylene glycol and was flash-cooled at 100 K. A diffraction data set for *Xcc*-Xgh complexed with glucose was collected from a crystal soaked in cryoprotectant solution containing 1 M glucose.

Diffraction data were collected on the MX2 beamline of the Brazilian Synchrotron Light Laboratory, Campinas, Brazil, using a wavelength of 1.48 Å (Guimarães *et al.*, 2009). A total of 360 images covering an oscillation range of 180° were collected using a MAR Mosaic 225 mm CCD detector (MAR Research). The diffraction images were processed using *XDS* (Kabsch, 2010) and reduced with *SCALA* from *CCP4* (Evans, 2006, 2011; Winn *et al.*, 2011). *POINTLESS* (Evans, 2006) was used to examine the symmetry of the diffraction pattern and score the possible crystallographic symmetry and the *phenix.xtriage* module of *PHENIX* (Adams *et al.*, 2010) was used to analyse the collected data for twinning.

Figure 1

Crystals of *Xcc*-Xgh obtained during initial screening (a) and after optimization of crystallization conditions (b).

Table 1

Crystallographic data and refinement statistics.

Values in parentheses are for the outermost resolution shell.

	Apoenzyme	Complex with glucose
Source	MX2, LNLS	MX2, LNLS
Wavelength (Å)	1.4586	1.4586
Space group	$P6_1$	$P6_1$
Unit-cell parameters (Å)		
$a = b$	153.43	150.53
c	84.92	84.20
Resolution (Å)	19.7–2.00 (2.30–2.00)	42.5–2.10 (2.40–2.10)
No. of reflections	837436	149921
No. of unique reflections	76752	55997
$R_{\text{merge}}^{\dagger}$ (%)	6.5 (28.6)	7.1 (72.4)
$\langle I/\sigma(I) \rangle$	26.86 (9.08)	29.1 (1.5)
Completeness (%)	99.8 (100)	99.3 (100)
Multiplicity	10.9 (10.7)	2.7 (2.6)

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

3. Results and discussion

The Xcc-Xgh protein was submitted to several crystallization trials and small crystals were obtained by a sitting-drop method in a condition consisting of 100 mM sodium cacodylate buffer pH 6.5, 200 mM calcium acetate, 18% (w/v) polyethylene glycol 8000; however, the macroscopic form of the crystals resembled a shower of needles (Fig. 1a). In order to obtain diffraction-quality crystals, the condition of initial crystallization was exhaustively optimized and the best crystals with a typical size of 0.2 mm in the largest dimension were obtained in 100 mM sodium cacodylate buffer pH 6.5, 18% PEG 8000, 200 mM calcium acetate at 291 K and grew in 1 d (Fig. 1b).

A diffraction data set was collected to 2.0 Å resolution at 100 K from a cryoprotected apoprotein crystal (Table 1). A second diffraction data set for the enzyme in complex with glucose was collected to 2.1 Å resolution. The glucose-soaked crystal has slightly different unit-cell parameters (Table 1) compared with the uncomplexed crystal, suggesting an impact of glucose on the crystal lattice. Assuming the presence of one molecule in the asymmetric unit, the solvent content of the crystals was estimated to be ~50.0% based on the Matthews coefficient (Matthews, 1968). The data scaling suggested the point group $P6$ and systematic absences indicated the enantiomorphic space groups $P6_1$ or $P6_5$. This ambiguity was solved by the molecular-replacement process, which was conducted using the crystallographic structure of the homologous enzyme from *C. thermocellum* F7/YS (PDB entry 2cn2; Martinez-Fleites *et al.*, 2006) and clearly revealed $P6_1$ as the correct space group. Refinement of the structures is in progress.

We are grateful to the Brazilian National Synchrotron Light Source (LNLS) and Laboratório Nacional de Biociências (LNBio, CNPEM). This work was supported by Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) via research grant Nos. 2008/56255-9, 2007/08706-9, 2010/52362-5 and 2009/05349-6, by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and by Conselho Nacional de Desenvolvimento Científico

e Tecnológico (CNPq) via INCT do Bioetanol and grant Nos. 471834/2009-2, 301981/2011-6 and 550931/2011-2. We are grateful to Livia Regina Manzi for suggestions about the writing of this paper and for discussions about molecular biology.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Balat, M. & Balat, H. (2009). *Appl. Energy*, **86**, 2273–2282.
- Bauer, S., Vasu, P., Mort, A. J. & Somerville, C. R. (2005). *Carbohydr. Res.* **340**, 2590–2597.
- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). *Nucleic Acids Res.* **37**, D233–D238.
- Chhabra, S. R. & Kelly, R. M. (2002). *FEBS Lett.* **531**, 375–380.
- Chundawat, S. P., Beckham, G. T., Himmel, M. E. & Dale, B. E. (2011). *Annu. Rev. Chem. Biomol. Eng.* **2**, 121–145.
- Damásio, A. R., Ribeiro, L. F., Furtado, G. P., Segato, F., Almeida, F. B., Crivellari, A. C., Buckridge, M. S., Souza, T. A., Murakami, M. T., Ward, R. J., Prade, R. A. & Polizeli, M. L. (2012). *Biochim. Biophys. Acta*, **1824**, 461–467.
- Desmet, T., Cantarel, T., Gualfetti, P., Nerinckx, W., Gross, L., Mitchinson, C. & Piens, K. (2007). *FEBS J.* **274**, 356–363.
- Ding, S.-Y., Liu, Y.-S., Zeng, Y., Himmel, M. E., Baker, J. O. & Bayer, E. A. (2012). *Science*, **338**, 1055–1060.
- Dixon, R. A. (2013). *Nature (London)*, **493**, 36–37.
- Enkhbaatar, B., Temuujin, U., Lim, J.-H., Chi, W.-J., Chang, Y.-K. & Hong, S.-K. (2012). *Appl. Environ. Microbiol.* **78**, 607–611.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Evans, P. R. (2011). *Acta Cryst.* **D67**, 282–292.
- Gilbert, H. J., Ståhlbrand, H. & Brumer, H. (2008). *Curr. Opin. Plant Biol.* **11**, 338–348.
- Goldemberg, J. (2008). *Conference on the Ecological Dimensions of Biofuels*, March 10, 2008, Washington DC. http://www.esa.org/biofuels/presentations/Goldemberg_BiofuelsPresentation.pdf.
- Guimarães, B. G., Sanfelici, L., Neuenschwander, R. T., Rodrigues, F., Grizolli, W. C., Raulik, M. A., Piton, J. R., Meyer, B. C., Nascimento, A. S. & Polikarpov, I. (2009). *J. Synchrotron Rad.* **16**, 69–75.
- Hasper, A. A., Dekkers, E., van Mil, M., van de Vondervoort, P. J. & de Graaff, L. H. (2002). *Appl. Environ. Microbiol.* **68**, 1556–1560.
- Hayashi, T. (1989). *Plant Physiol. Plant Mol. Biol.* **40**, 139–168.
- Himmel, M. E., Ding, S.-Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W. & Foust, T. D. (2007). *Science*, **315**, 804–807.
- Ichinose, H., Araki, Y., Michikawa, M., Harazono, K., Yaoi, K., Karita, S. & Kaneko, S. (2012). *Appl. Environ. Microbiol.* **78**, 7939–7945.
- Irwin, D. C., Cheng, M., Xiang, B., Rose, J. K. & Wilson, D. B. (2003). *Eur. J. Biochem.* **270**, 3083–3091.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
- Lynd, L. R., Weimer, P. J., van Zyl, W. H. & Pretorius, I. S. (2002). *Microbiol. Mol. Biol. Rev.* **66**, 506–577.
- Martinez-Fleites, C., Guerreiro, C. I., Baumann, M. J., Taylor, E. J., Prates, J. A., Ferreira, L. M., Fontes, C. M., Brumer, H. & Davies, G. J. (2006). *J. Biol. Chem.* **281**, 24922–24933.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Park, Y. B. & Cosgrove, D. J. (2012). *Plant Physiol.* **158**, 1933–1943.
- Peng, F., Peng, P., Xu, F. & Sun, R.-C. (2012). *Biotechnol. Adv.* **30**, 879–903.
- Scheller, H. V. & Ulvskov, P. (2010). *Annu. Rev. Plant Biol.* **61**, 263–289.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yaoi, K., Kondo, H., Hiyoshi, A., Noro, N., Sugimoto, H., Tsuda, S., Mitsuishi, Y. & Miyazaki, K. (2007). *J. Mol. Biol.* **370**, 53–62.
- Yaoi, K., Kondo, H., Hiyoshi, A., Noro, N., Sugimoto, H., Tsuda, S. & Miyazaki, K. (2009). *FEBS J.* **276**, 5094–5100.
- Yaoi, K., Kondo, H., Noro, N., Suzuki, M., Tsuda, S. & Mitsuishi, Y. (2004). *Structure*, **12**, 1209–1217.
- Yaoi, K. & Mitsuishi, Y. (2002). *J. Biol. Chem.* **277**, 48276–48281.