Expression, purification, crystallization and preliminary X-ray diffraction analysis of *Aspergillus terreus* endo- β -1,4-glucanase from glycoside hydrolase family 12

Fernando Segato,[‡] Gabriela L. Berto,[‡] Evandro Ares de Araújo, João Renato Muniz and Igor Polikarpov*

Instituto de Física de São Carlos, Universidade de São Paulo, Avenida Trabalhador Saocarlense 400, 13566-590 São Carlos-SP, Brazil

[‡] These authors contributed equally.

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

Received 13 October 2013

Accepted 30 December 2013

Endoglucanases are important enzymes that are involved in the modification and degradation of cellulose. Filamentous fungi such as *Aspergillus terreus* are effective biomass degraders in nature owing to their capacity to produce an enzymatic arsenal of glycoside hydrolases, including endoglucanase from glycoside hydrolase family 12 (GH12). The *A. terreus* GH12 endoglucanase was cloned and overexpressed in *A. nidulans*, purified and crystallized. A single crystal was obtained from a solution consisting of 2 M ammonium sulfate, 5% (v/v) 2-propanol. X-ray diffraction data were collected to a resolution of 1.85 Å using synchrotron radiation and a preliminary molecular-replacement solution was obtained in the trigonal space group $P3_221$. The unit-cell parameters were $a = b = 103.24$, $c = 48.96$ Å.

1. Introduction

Concerns over global climate change, increasing demand for energy and the reduction in oil supplies have fostered the need for the development of alternative and renewable energy resources which can replace fossil fuels. Biomass is an abundant feedstock, which can be converted into different types of energy sources and chemicals. It consists of three main components, cellulose, hemicellulose and lignin, the first two of which are composed of carbohydrates that can be depolymerized and fermented to ethanol and/or converted to a number of chemical products (Himmel *et al.*, 2007). The main technological barrier to biomass utilization is the absence of low-cost technologies to decrease the recalcitrance of cellulosic biomass. The most promising strategy to overcome this barrier involves the use of mixtures of microbial cellulolytic enzymes to break down biomass (Lynd *et al.*, 2002; Wang *et al.*, 2011).

Cellulases are glycoside hydrolases found in at least 12 enzyme families (<http://www.cazy.org>) which can be broadly separated into endoglucanases (EC 3.2.1.4), which randomly hydrolyze internal glycoside bonds, exoglucanases (EC 3.2.1.91/176), which split off cellobiose from cellulose termini, and β -glucosidases (EC 3.2.1.21), which hydrolyze cellobiose and oligomers to produce glucose (Zhang & Lynd, 2004). The endoglucanases from GH (glycoside hydrolase) family 12 (GH12) hydrolyze the β -1,4-glycosidic bond in cellulose (Henrissat, 1991) and are widely used in industrial processes and in animal feed to improve the digestibility of glucans (Picart *et al.*, 2012).

Aspergillus terreus is an important filamentous fungus with significant commercial relevance in industrial biotechnology. This fungus is frequently isolated from agricultural wastes and cellulosic materials, and *A. terreus* enzymes have been extensively studied and utilized in various bioprocesses (Han *et al.*, 2010; Chidi *et al.*, 2008). Furthermore, *A. terreus* has been shown to produce secondary metabolites of significant industrial interest using biomass as a substrate (Jahromi *et al.*, 2012). In this work, we report the cloning, heterologous overexpression, purification and crystallization of an endoglucanase member of GH family 12 from *A. terreus* strain NIH2624.

2. Materials and methods

2.1. Protein expression and purification

The *A. terreus* GH12 (AtGH12) gene sequence (NCBI entry XP_001218516) was analyzed using the *SignalP* 4.1 server (Center for Biological Sequence Analyses – CBS; Petersen *et al.*, 2011). The program identified 45 nucleotides that encode a 15-amino-acid leader-peptide sequence; this sequence was omitted from the final construct by an appropriate choice of primers. The gene was amplified from the *A. terreus* NIH2624 genomic DNA by the polymerase chain reaction (PCR) method using the oligonucleotide primers AT09894F (5'-TATAGCGGCCGCCAGGAGCTCTGCGAGCAATATGG-3') and AT09894R (5'-TATATCTAGATGCCACACTAGCAGACCACT-3') containing *NotI* and *XbaI* restriction sites (shown in bold), respectively. To generate an expression plasmid, the amplified 660 bp fragment was inserted into the pEXPYR vector, which expresses the GH12 gene under the control of the *A. niger* glucoamylase promoter and includes the glucoamylase leader-peptide sequence (*glaAp*). The vector also contains an auxotrophic marker (*pyrG*), which can be recycled, thus allowing multiple sequential transformations of the strain (Segato, Damásio, Gonçalves, de Lucas *et al.*, 2012). The recombinant plasmid was transformed in *Escherichia coli* One Shot TOP10 competent cells (Invitrogen) to propagate copies, and was then extracted and used to transform the *A. nidulans* A773 host strain (Segato, Damásio, Gonçalves, de Lucas *et al.*, 2012).

For heterologous protein expression, saline solution [0.9%(w/v) NaCl] containing 10^7 – 10^8 spores per millilitre was inoculated in 300 ml of liquid minimal medium supplemented with 2% maltose (Sigma–Aldrich, USA). The medium was distributed onto dishes (100 ml in 150 mm Petri dishes) and incubated at 310 K for 2 d (Segato, Damásio, Gonçalves, Murakami *et al.*, 2012). The mycelial mat was lifted with a spatula and discarded. The medium was collected by filtration, centrifuged at 10 000g for 10 min prior to concentration by ultrafiltration (10 000 Da cutoff Amicon concentrator) and the buffer was exchanged for 20 mM bis-tris pH 6.0, 100 mM NaCl. The enzyme was then purified by chromatography using a Q Sepharose column (GE Healthcare Life Science) equili-

brated with 20 mM bis-tris, 150 mM NaCl and eluted with a linear gradient of 0–1 M NaCl in the same buffer. The AtGH12-containing fractions were concentrated with an Amicon concentrator and the buffer was exchanged to 20 mM bis-tris pH 6.0 before further purification at room temperature by size-exclusion chromatography on a Superdex G-75 column (GE Healthcare Life Science) equilibrated with the same buffer using an ÄKTApurifier 10 system (GE Healthcare Life Science).

The protein fractions were collected and monitored for endoglucanase activity using 100 µl β -glucan (Megazyme) solution at 0.5%(w/v) concentration in phosphate buffer pH 6.5 and 1 µg purified GH12. The mixture was incubated for 15 min at the optimal temperature of the enzyme (328 K; results not shown). The reaction was stopped by adding 100 µl 3,5-dinitrosalicylic acid (DNS) and boiling for 5 min, and the solution was analyzed in a spectrophotometer at a wavelength of 575 nm for quantification of reducing sugars (Miller, 1959). The fractions containing endoglucanase activity were pooled and concentrated again in an Amicon (10 000 Da cutoff). Enzymatic activity against xyloglucan and konjac mannan was also determined by the DNS method (Miller, 1959) using the same procedure.

The purified *A. terreus* GH12 was concentrated to 9 mg ml⁻¹ in 20 mM bis-tris buffer pH 6.0 for crystallization assays. The protein concentration was determined by measuring the absorption at 280 nm using a NanoDrop spectrophotometer (Thermo Scientific). The extinction coefficient was deduced from the protein sequence using the *ProtParam* software (<http://web.expasy.org/protparam/>; Gasteiger *et al.*, 2005).

2.2. Crystallization

The purified AtGH12 (molecular mass 23 979 Da) at a concentration of 9 mg ml⁻¹ was submitted to crystallization screening using the sitting-drop vapour-diffusion technique. Drops of 2 µl final volume (1:1 ratio of protein and screening solutions) were set up automatically with a Honeybee 931 crystallization robot (Genomic Solutions Inc.) using commercial screening reagent kits and maintained at a temperature of 291 K. Once initial hits for crystallization had been determined, the crystals were further optimized and the best condition was obtained when a 1 µl droplet was mixed with a reservoir solution consisting of 2.0 M ammonium sulfate, 5%(v/v) 2-propanol.

2.3. Data collection and processing

For data collection, a single crystal was transferred to a cryo-protectant solution consisting of 15%(v/v) ethylene glycol added to the reservoir solution and was flash-cooled in a liquid-nitrogen vapour stream at 100 K (Oxford Cryosystems). Diffraction data were collected on the MX2 beamline at the Brazilian Synchrotron Light Laboratory (LNLS), Campinas, Brazil (Guimarães *et al.*, 2009). The wavelength of the radiation source was set to 1.459 Å to increase the diffraction efficiency (Polikarpov *et al.*, 1997) and to optimize the synchrotron-radiation flux. A total of 360 images with an oscillation angle of 0.5°, corresponding to a range of 180°, were collected using a MAR Mosaic 225 mm CCD detector (MAR Research). The data were indexed and scaled using *XDS* (Kabsch, 2010) and reduced with *SCALA* from *CCP4* (Evans, 2006; Winn *et al.*, 2011). *POINTLESS* (Evans, 2006) and the *phenix.xtriage* module of *PHENIX* (Adams *et al.*, 2010) were used to examine the symmetry of the diffraction pattern.

Figure 1

Secretion in *A. nidulans*: pEXPYR plasmid carrying *A. terreus* GH12 was transformed into *A. nidulans* A773 (*pyrG*). Transformants were grown on medium containing 5% maltose for 2 d and secreted proteins were analyzed by SDS–PAGE. Lane M, molecular-mass standards (labelled in kDa); lane 1, A773 carrying empty vector; lane 2, A773 carrying vector containing the *A. terreus* GH12 gene; lane 3, purified *A. terreus* GH12.

3. Results and discussion

The final protein sample was more than 95% pure as assessed by SDS-PAGE stained with Coomassie Brilliant Blue (Shapiro *et al.*, 1967; Fig. 1). We presume that the enzyme is partially glycosylated, which promotes differences in protein migration as revealed by a doublet in lane 3 showing the purified protein (Fig. 1). However, the heterogeneity of the protein did not affect its crystallization. The final yield of pure protein was approximately 30 mg per litre of culture. Amino-acid sequence analysis of the GH12 translated gene from *A. terreus* using *BLASTp* (<http://blast.ncbi.nlm.nih.gov>) showed significant similarity to the endoglucanases from *A. niger* (PDB entry 1ks4; Khademi *et al.*, 2002), *Trichoderma harzianum* (PDB entry 4h7m; Prates *et al.*, 2013) and *T. reesei* (PDB entry 1h8v; Sandgren *et al.*, 2001), as visualized in the alignment formatted with *T_EXshade* (Beitz, 2000; Fig. 2). The purified recombinant GH12 from *A. terreus* was active towards β -glucan, konjac mannan and xyloglucan, with particularly high activity towards the latter substrate. These results support the notion that the enzyme encoded by the GH12 gene from *A. terreus* is an endoglucanase with high xyloglucanase activity (results not shown).

The AtGH12 was submitted to crystallization trials and crystals were obtained by the sitting-drop method after 2 d using a reservoir solution consisting of 2 M ammonium sulfate, 5% (v/v) 2-propanol. A single crystal (Fig. 3) was used for X-ray diffraction experiments and data were collected to 1.85 Å resolution under cryogenic conditions (100 K). Scaling of the diffraction data indicated the space group was either $P3_121$ or $P3_221$, with unit-cell parameters $a = b = 103.24$, $c = 48.96$ Å. Taking into consideration the molecular mass of 23 979 Da, the calculated Matthews coefficient was $3.11 \text{ Å}^3 \text{ Da}^{-1}$ for one enzyme molecule in the asymmetric unit, with a solvent content of 60.5% (Matthews, 1968). The data-processing statistics are presented in Table 1. The atomic coordinates of an endo- β -1,4-glucanase found in the Protein Data Bank (PDB entry 1ks4; Khademi *et al.*, 2002), which displays a sequence identity of 58%, were used as a search model for molecular replacement using the program *Phaser* (McCoy *et al.*, 2007). A final translation-function Z-score (TFZ) of 17.0 was a good indication that the correct molecular-replacement solution had been found, and further refinement steps, which led to a clear decrease in R_{work} and R_{free} to 29.90 and 32.40%, respectively, revealed that the correct space group was $P3_221$. Refinement of the structure is in progress. In parallel with the structural studies,

Figure 2 Multiple sequence alignment of endo- β -1,4-endoglucanases from *T. reesei* (PDB entry 1h8v; Sandgren *et al.*, 2001), *T. harzianum* (PDB entry 4h7m; Prates *et al.*, 2013), *A. terreus* and *A. niger* (PDB entry 1ks4; Khademi *et al.*, 2002). The conserved amino-acid residues are highlighted in greyscale and the consensus is shown as a logo representation. The sequence identities of *A. terreus* endoglucanase to *A. niger*, *T. reesei* and *T. harzianum* endoglucanases are 58, 56 and 55%, respectively.

Table 1

Data-collection statistics.

Values in parentheses are for the outermost resolution shell.

Source	MX2, LNLS
Wavelength (Å)	1.459
Resolution (Å)	25.81–1.85 (1.89–1.85)
No. of reflections	252490 (15098)
No. of unique reflections	25703 (1551)
Multiplicity	9.8 (9.7)
$R_{\text{merge}}^{\dagger}$ (%)	10.2 (96.7)
Completeness (%)	99.3 (97.8)
Mean $I/\sigma(I)$	11.9 (2.2)
Mosaicity (°)	0.66

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the i th replicate of reflection hkl and $\langle I(hkl) \rangle$ is the mean value, and the summations are over all i replicates and then over all hkl .

Figure 3

Photomicrograph of *A. terreus* endoglucanase GH12. The minimum scale division corresponds to 1.0 μm .

comprehensive biochemical and enzymatic analyses including substrate-specificity and hydrolysis-product pattern evaluation are being carried out.

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 e Pesquisa do Estado de São Paulo (FAPESP) via research grants 2008/05637-9, 2007/0870-9, 2010/16947-9 and INCT Bioetanol (FAPESP/CNPq), by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) via grants 301981/2011-6, 403090/2012-1, 373143/

2012-5 and 150600/2013-5 and the MCT/CNPq/FAPESP EU–Brazil Collaboration program in Second Generation Biofuels (CeProBio Project; FAPESP 2009/52840-7 and CNPq 490022/2009-0), and by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Beitz, E. (2000). *Bioinformatics*, **16**, 135–139.
- Chidi, S. B., Godana, B., Ncube, I., Van Rensburg, E. J., Cronshaw, A. & Abotsi, E. K. (2008). *Afr. J. Biotechnol.* **7**, 3939–3948.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. M. Walker, pp. 571–607. Totowa: Humana Press.
- 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.
- Han, M.-J., Kim, N.-J., Lee, S. Y. & Chang, H. N. (2010). *Curr. Genet.* **56**, 369–382.
- Henrissat, B. (1991). *J. Biochem.* **280**, 309–316.
- 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.
- Jahromi, M. F., Liang, J. B., Ho, Y. W., Mohamad, R., Goh, Y. M. & Shokryazdan, P. (2012). *J. Biomed. Biotechnol.* **2012**, 196264.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Khademi, S., Zhang, D., Swanson, S. M., Wartenberg, A., Witte, K. & Meyer, E. F. (2002). *Acta Cryst.* **D58**, 660–667.
- Lynd, L. R., Weimer, P. J., van Zyl, W. H. & Pretorius, I. S. (2002). *Microbiol. Mol. Biol. Rev.* **66**, 506–577.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Miller, G. L. (1959). *Anal. Chem.* **31**, 426–428.
- Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. (2011). *Nature Methods*, **8**, 785–786.
- Picart, P., Goedegebuur, F., Díaz, P. & Pastor, F. I. J. (2012). *Fungal Biol.* **116**, 443–451.
- Polikarpov, I., Teplyakov, A. & Oliva, G. (1997). *Acta Cryst.* **D53**, 734–737.
- Prates, É. T., Stankovic, I., Silveira, R. L., Liberato, M. V., Henrique-Silva, F., Pereira, N. Jr, Polikarpov, I. & Skaf, M. S. (2013). *PLoS One*, **8**, e59069.
- Sandgren, M., Shaw, A., Ropp, T. H., Wu, S., Bott, R., Cameron, A. D., Ståhlberg, J., Mitchinson, C. & Jones, A. (2001). *J. Mol. Biol.* **308**, 295–310.
- Segato, F., Damásio, A. R. L., Gonçalves, T. A., de Lucas, R. C., Squina, F. M., Decker, S. R. & Prade, R. A. (2012). *Enzyme Microb. Technol.* **51**, 100–106.
- Segato, F., Damasio, A. R. L., Gonçalves, T. A., Murakami, M. T., Squina, F. M., Polizeli, M. T. M., Mort, A. J. & Prade, R. A. (2012). *Biotechnol. Biofuels*, **5**, 21.
- Shapiro, A. L., Viñuela, E. & Maizel, J. V. Jr (1967). *Biochem. Biophys. Res. Commun.* **28**, 815–820.
- Wang, H., Squina, F., Segato, F., Mort, A., Lee, D., Pappan, K. & Prade, R. (2011). *Appl. Environ. Microbiol.* **77**, 5199–5206.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Zhang, Y.-H. P. & Lynd, L. R. (2004). *Biotechnol. Bioeng.* **88**, 797–824.