Anil Kumar Verma,^a Arun Goyal,^{a*} Filipe Freire,^b Pedro Bule,^b Immacolata Venditto,^b Joana L. A. Brás,^b Helena Santos,^b Vânia Cardoso,^b Cecília Bonifácio,^c Andrew Thompson,^d Maria João Romão,^c José A. M. Prates,^b Luís M. A. Ferreira,^b Carlos M. G. A. Fontes^b and Shabir Najmudin^{b*}

^aDepartment of Biotechnology, Indian Institute of Technology Guwahati, Guwahati 781 039, India, ^bCIISA—Faculdade de Medicina Veterinária, Universidade Técnica de Lisboa, Avenida da Universidade Técnica, 1300-477 Lisboa, Portugal, ^cREQUIMTE—CQFB, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal, and ^dSynchrotron SOLEIL, L'Orme des Merisiers, Saint-Aubin, 91192 Gif-sur-Yvette, France

Correspondence e-mail: arungoyl@iitg.ernet.in, shabir@fmv.utl.pt

Received 6 August 2013
Accepted 9 September 2013

Overexpression, crystallization and preliminary X-ray crystallographic analysis of glucuronoxylan xylanohydrolase (Xyn30A) from *Clostridium thermocellum*

The modular carbohydrate-active enzyme belonging to glycoside hydrolase family 30 (GH30) from *Clostridium thermocellum* (CtXynGH30) is a cellulosomal protein which plays an important role in plant cell-wall degradation. The full-length CtXynGH30 contains an N-terminal catalytic module (Xyn30A) followed by a family 6 carbohydrate-binding module (CBM6) and a dockerin at the C-terminus. The recombinant protein has a molecular mass of 45 kDa. Preliminary structural characterization was carried out on Xyn30A crystallized in different conditions. All tested crystals belonged to space group *P*1 with one molecule in the asymmetric unit. Molecular replacement has been used to solve the Xyn30A structure.

1. Introduction

The modular carbohydrate-active enzyme belonging to glycoside hydrolase family 30 (GH30) from *Clostridium thermocellum* (CtXynGH30) is a cellulosomal protein which plays an important role in plant cell-wall degradation. It contains an N-terminal signal peptide followed by a catalytic module (Xyn30A), a C-terminal carbohydrate-binding module belonging to family 6 (CBM6) and a dockerin module. CtXynGH30 was originally classified as a glycoside hydrolase (GH) family 5 member in the CAZy database. It has now been reclassified as a member of the GH family 30 glucuronoxylan xylanohydrolases, subfamily H (in the CAZy database it is in subfamily 8; Cantarel *et al.*, 2009; St John *et al.*, 2010; Aspeborg *et al.*, 2012). Structurally characterized proteins of this subfamily include XynA from *Erwinia chrysanthemi* (Hurlbert & Preston, 2001; Vrsanská *et al.*, 2007) and XynC from *Bacillus subtilis* (St John *et al.*, 2009). These display preferential hydrolysis of glucuronoxylan to branched xylo-oligosaccharides and negligible activity towards linear β -1–4-linked xylo-oligosaccharides (St John *et al.*, 2006, 2011). Sequence alignment, protein-modelling and molecular-docking studies of Xyn30A showed that the modelled structure is close to the XynC family GH30 from *B. subtilis* 168, presenting a common $(\beta\alpha)_8$ -barrel fold with distinguishing structural features of family GH30 xylanases such as the obligate 'side β -structure' similar to CBM (St John *et al.*, 2006). Determining the three-dimensional structure of *C. thermocellum* Xyn30A is an essential step for the further biochemical understanding of the protein activity towards its different substrates. In the present study, recombinant protein Xyn30A was cloned, overexpressed, purified and crystallized. X-ray diffraction data were collected for preliminary structure analysis.

2. Materials and methods

2.1. Bacterial strains and plasmids

Escherichia coli DH5 α cells were used for cloning and *E. coli* BL21 (DE3) cells were used as the expression host. The plasmid used for cloning and expression was pET-28a(+). All of the above-mentioned items were procured from Novagen (Madison, Wisconsin, USA). pGEM-T Easy vector was purchased from Promega Corporation (Madison, Wisconsin, USA).

2.2. Gene amplification and cloning

Oligonucleotide primers containing *Nhe*I and *Xho*I restriction sites were designed. The primers used for amplifying Xyn30A were the forward primer 5'-CTCTGCTAGCGCAACAATCAACTTGTCG-3' and the reverse primer 5'-CACACTCGAGTTATACAAAAGTTGTAACGC-3'. PCR was carried out using denaturation at 367 K for 5 min and 30 cycles of (i) denaturation at 367 K for 30 s, (ii) annealing at 328 K for 60 s and (iii) extension at 345 K for 2 min followed by extension at 345 K for 5 min. The PCR product was run on 0.8% agarose gel, purified using a QIAquick gel-extraction kit, ligated into pGEM-T Easy vector and positive clones were screened (Suzuki *et al.*, 1997). The Xyn30A gene fragment was extracted and cloned into pET-28a(+) vector using *Nhe*I and *Xho*I restriction enzymes. *E. coli* DH5 α cells were transformed by the heat-shock method using cloned DNA. The transformed cells were grown on LB-agar plates supplemented with kanamycin (50 μ g ml⁻¹) at 310 K. Positive clones were confirmed by restriction digestion of the recombinant plasmid (pXyn30A) purified using a QIAprep Spin Miniprep Kit.

2.3. Expression and purification of recombinant Xyn30A

E. coli BL21(DE3) cells were transformed with the recombinant plasmid pXyn30A. Cells harbouring pXyn30A were grown in LB containing kanamycin (50 μ g ml⁻¹) at 310 K at 180 rev min⁻¹ until the mid-exponential phase ($A_{550} \approx 0.6$), recombinant protein expression was induced with 1.0 mM isopropyl β -D-1-thiogalactopyranoside and the cells were left for an additional 16 h at 292 K and 180 rev min⁻¹ (Ichinose *et al.*, 2008). The cells were harvested by centrifugation at 8000g at 277 K for 15 min and the resulting cell pellet was resuspended in 50 mM HEPES buffer containing 10 mM imidazole, 1 M NaCl, 3 mM CaCl₂·2H₂O pH 7.5. The cells were then sonicated on ice for 15 min (8 s on/16 s off pulse, 33% amplitude) and centrifuged again at 17 000g at 277 K for 40 min to obtain the cell-free extract. The recombinant Xyn30A containing a His₆ tag was purified from the cell-free extract by immobilized metal-ion affinity chromatography (IMAC) using a Sepharose column (HiTrap Chelating, GE Healthcare) and was eluted using 50 mM HEPES buffer containing 300 mM imidazole, 1 M NaCl, 3 mM CaCl₂·2H₂O pH 7.5. Purified Xyn30A was buffer-exchanged into 50 mM HEPES buffer pH 7.5 containing 200 mM NaCl, 5 mM CaCl₂ using a PD-10 column and was then subjected to gel filtration using a HiLoad 16/60 Superdex 75 column (GE Healthcare) at a flow rate of 1 ml min⁻¹. Purified Xyn30A was concentrated using a 10 kDa cutoff centrifugal concentrator (Millipore, USA) and washed three times with 0.5 M CaCl₂. The purity and molecular mass of the recombinant protein were analysed by 13% (w/v) SDS-PAGE (Laemmli, 1970; Fig. 1a). The purified recombinant Xyn30A contains an N-terminal His₆ tag (MGSSHHHHHHSSGLVPRGSHMAS), giving a total of 409 amino-acid residues with a molecular mass of 45.18 kDa. The protein concentration was estimated using a molar extinction coefficient (ϵ) of 91 853 M⁻¹ cm⁻¹ with a NanoDrop 2000c spectrophotometer (Thermo Scientific).

2.4. Crystallization

The crystallization conditions were screened by the sitting-drop vapour-diffusion method using the commercial screens Crystal Screen, Crystal Screen 2 and PEG/Ion from Hampton Research (California, USA), JCSG from Molecular Dimensions and an in-house-prepared sparse-matrix screen of 80 conditions using a NanoDrop robotic dispensing system Oryx8 (Douglas Instruments). Drops consisting of 0.9 μ l Xyn30A solution at either 16 or 32 mg ml⁻¹

and 0.9 μ l reservoir solution were prepared at 292 K. Crystals appeared within 1 d, reaching their maximum size within a few days in many different conditions: (i) 2 M ammonium sulfate; (ii) 0.1 M HEPES pH 7.5, 10% polyethylene glycol (PEG) 6000, 5% (v/v) ($\pm$)-2-methyl-2,4-pentanediol; (iii) 0.2 M sodium fluoride, 20% (w/v) PEG 3350; (iv) 0.2 M potassium sulfate, 20% (w/v) PEG 3350 (Fig. 1b); (v) 0.2 M sodium acetate trihydrate pH 7.0, 20% PEG 3350; and (vi) 2% (v/v) Tacsimate pH 4.0, 0.1 M sodium acetate trihydrate 4.6, 16% PEG 3350. The crystals were cryocooled in liquid nitrogen after soaking in the appropriate cryoprotectant [30% (v/v) glycerol added to the crystallization buffer] for a few seconds, except for the crystal grown in 0.2 M potassium sulfate, 20% (w/v) PEG 3350, which was cryocooled in Paratone.

2.5. Data collection and processing

Data from all the different crystals obtained were collected on beamline ID29 at the European Synchrotron Radiation Facility (ESRF), Grenoble (de Sanctis *et al.*, 2012) and at PROXIMA-1 at SOLEIL (Orsay, France) using a PILATUS 6M detector with the crystals cooled to 100 K using a Cryostream (Oxford Cryosystems). Data from the ESRF and SOLEIL were processed using the

Figure 1
(a) SDS-PAGE [13% (w/v)] showing overexpression and purification of Xyn30A. Lane 1, Page Ruler protein marker (labelled in kDa); lane 2, uninduced Xyn30A cells; lane 3, cell-free extract; lane 4, last wash from column; lane 5, purified Xyn30A (~45 kDa). (b) Crystals of Xyn30A obtained by sitting-drop vapour diffusion. The largest crystals are approximately 200 $\times$ 120 $\times$ 60 μ m in size.

Table 1

Data-collection statistics.

Values in parentheses are for the lowest/highest resolution shells.

Data set	Xyn30A
Beamline	PROXIMA-1, SOLEIL
Wavelength (Å)	0.826
Space group	<i>P</i> 1
Unit-cell parameters (Å, °)	$a = 46.09$, $b = 47.46$, $c = 53.62$, $\alpha = 83.15$, $\beta = 73.41$, $\gamma = 65.87$
Resolution limits (Å)	42.93–1.40 (42.93–4.43/1.48–1.40)
Mosaicity (°)	0.39
No. of observations	347015 (11412/50059)
No. of unique observations	76088 (2464/10874)
Multiplicity	4.6 (4.6/4.6)
Completeness (%)	97.2 (99.3/95.6)
$\langle I/\sigma(I) \rangle$	3.5 (7.3/1.3)
$CC_{1/2}^\dagger$	0.971 (0.974/0.536)
$R_{\text{merge}}^\ddagger$	0.196 (0.100/0.860)
$R_{\text{p.i.m.}}^\S$	0.101 (0.052/0.449)

† $CC_{1/2}$ is the half-data-set correlation coefficient (Diederichs & Karplus, 2013). ‡ $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the i th intensity measurement of reflection hkl , including symmetry-related reflections and $\langle I(hkl) \rangle$ is its average. § $R_{\text{p.i.m.}} = \sum_{hkl} \{ [1/N(hkl) - 1]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl) \}$, where $\langle I(hkl) \rangle$ is the average of symmetry-related observations of a unique reflection.

Figure 2

Representative diffraction pattern of a Xyn30A crystal with resolution rings shown as dashed circles. The inset shows an enlargement of the contents of the red box, showing spots at ~ 1.2 Å resolution.

programs *iMOSFLM* (Battye *et al.*, 2011) and *AIMLESS* (Evans, 2011) from the *CCP4* suite (Winn *et al.*, 2011) or *XDS* (Kabsch, 2010). All of the crystals tested belonged to space group *P*1 with one molecule in the asymmetric unit and a solvent content ranging from 45 to 47% and had similar data-collection statistics, despite being grown in different conditions and having different crystal morphologies.

The data collected at SOLEIL were from the best diffracting crystal (Fig. 1*b*) and were used to solve the Xyn30A structure. 470° of data were collected with a $\Delta\phi$ of 0.2° . The crystal diffracted to a resolution of 1.1 Å (Fig. 2). The crystal was broken/multiple and a second low-resolution data set was needed to index and to find the

correct lattice before integrating the high-resolution data with *XDS*. The resolution was cut at a conservative 1.4 Å (the $CC_{1/2}$ value is below 0.5 for higher resolution data). The slightly high R_{merge} value at this resolution is due to radiation damage from a longer exposure for a *P*1 data collection. The Matthews coefficient of $2.33 \text{ Å}^3 \text{ Da}^{-1}$ (Matthews, 1968) indicates the presence of one molecule in the asymmetric unit with a solvent content of 47.2%. Data-collection statistics are given in Table 1.

The three-dimensional structure was solved by molecular replacement using the program *Phaser* (McCoy *et al.*, 2007) as implemented in the *PHENIX* software package (Adams *et al.*, 2010) and using the crystal structure of XynC from *B. subtilis* 168 (PDB code 3gtg; St John *et al.*, 2009) as a search model, with the final solution giving an LLG score of 181 and a TFZ of 24.5. Structure refinement and analysis are ongoing.

This research work was supported in part by a Cutting-edge Research Enhancement and Scientific Training (CREST) Fellowship from the Department of Biotechnology, Ministry of Science and Technology to AG. A scholarship from IIT Guwahati to AKV is gratefully acknowledged. The authors also acknowledge financial support from Fundação para a Ciência e a Tecnologia, Portugal through projects PTDC/BIA-PRO/103980/2008 and PTDC/BIA-PRO/100359/2008. We acknowledge SOLEIL for provision of synchrotron-radiation facilities (proposal ID 20120372) and Dr Victor D. Alves for help with data collection. We would also like to thank the ESRF, Grenoble (beamline ID29) for access and technical support during data collection.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
 Aspeborg, H., Coutinho, P. M., Wang, Y., Brumer, H. III & Henrissat, B. (2012). *BMC Evol. Biol.* **12**, 186.
 Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
 Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). *Nucleic Acids Res.* **37**, D233–D238.
 Diederichs, K. & Karplus, P. A. (2013). *Acta Cryst.* **D69**, 1215–1222.
 Evans, P. R. (2011). *Acta Cryst.* **D67**, 282–292.
 Hurlbert, J. C. & Preston, J. F. III (2001). *J. Bacteriol.* **183**, 2093–2100.
 Ichinose, H., Yoshida, M., Fujimoto, Z. & Kaneko, S. (2008). *Appl. Microbiol. Biotechnol.* **80**, 399–408.
 Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
 Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
 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.
 Sanctis, D. de *et al.* (2012). *J. Synchrotron Rad.* **19**, 455–461.
 St John, F. J., Godwin, D. K., Preston, J. F., Pozharski, E. & Hurlbert, J. C. (2009). *Acta Cryst.* **F65**, 499–503.
 St John, F. J., González, J. M. & Pozharski, E. (2010). *FEBS Lett.* **584**, 4435–4441.
 St John, F. J., Hurlbert, J. C., Rice, J. D., Preston, J. F. & Pozharski, E. (2011). *J. Mol. Biol.* **407**, 92–109.
 St John, F. J., Rice, J. D. & Preston, J. F. (2006). *J. Bacteriol.* **188**, 8617–8626.
 Suzuki, T., Ibata, K., Hatsu, M., Takamizawa, K. & Kawai, K. (1997). *J. Ferment. Bioeng.* **84**, 86–89.
 Vrsanská, M., Kolenová, K., Puchart, V. & Biely, P. (2007). *FEBS J.* **274**, 1666–1677.
 Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.