Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

K. Manikandan,[a] S. Jagtap,[b] M. Rao[b] and S. Ramakumar[a,c]*

[a]Department of Physics, Indian Institute of Science, Bangalore 560 012, India, [b]Division of Biochemical Sciences, National Chemical Laboratory, Pune 411 008, India, and [c]Bioinformatics Centre, Indian Institute of Science, Bangalore 560 012, India

Correspondence e-mail: ramak@physics.iisc.ernet.in

Received 7 February 2006
Accepted 4 March 2006

Crystallization and preliminary X-ray characterization of a thermostable low-molecular-weight 1,4-β-D-glucan glucohydrolase from an alkalothermophilic *Thermomonospora* sp.

Cellulases catalyze the hydrolysis of β-1,4-glycosidic linkages within cellulose, the most abundant organic polymer on earth. The cellulase (TSC; EC 3.2.1.4) from an alkalothermophilic *Thermomonospora* sp. has a low molecular weight of 14.2 kDa. It is optimally active at 323 K and stable over the wide pH range of 5–9. Moreover, it has bifunctional activity against cellulose and xylan polymers. In this study, TSC was purified from the native source and crystallized by the hanging-drop vapour-diffusion method. The crystals belong to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 49.9, b = 79.5, c = 99.7 Å, and diffract to better than 2.3 Å resolution.

1. Introduction

Cellulose is the predominant component of plant biomass and as such is the most abundant organic polymer on earth (Sandgren *et al.*, 2005). It is a major constituent of the complex matrix of the plant cell wall. The total amount of cellulose on earth has been estimated at 7×10^{11} tons (Coughlan, 1985). It is a linear homo-polysaccharide consisting of anhydrous glucose units that are linked by β-1,4-glycosidic bonds. Cellulases are enzymes that hydrolyse the β-1,4-glycosidic bonds of celluloses and have been classified broadly into two types: (i) endoglucanases (EC 3.2.1.4), which hydrolyze bonds internally in cellulose chains, and (ii) cellobiohydrolases (EC 3.2.1.91), which act preferentially on chain ends and progressively cleave off cellobiose as the main product. Cellulases play a significant role in plant-biomass degradation and in maintaining the recycling of carbon on earth and have tremendous biotechnological applications in industries such as textiles, animal feed, ethanol production and the clarification of juices and beverages.

Cellulases have been assigned to 12 of the glycosyl hydrolase (GH) families: 5–9, 12, 26, 44, 45, 48, 61 and 74 (http://afmb.cnrs-mrs.fr/CAZY; Henrissat & Bairoch, 1996; Henrissat & Davies, 1997). In general, cellulases are modular enzymes with molecular weights varying from approximately 20 to 90 kDa consisting of a catalytic domain connected to a cellulose-binding domain by a flexible linker region.

The cellulase (TSC) from an alkalothermophilic *Thermomonospora* sp. is an extracellular enzyme. It has a low molecular weight of 14.2 kDa. The purification and properties of TSC have been described recently by Jagtap & Rao (2005). Interestingly, it has been shown that TSC has only one active site for the substrates carboxymethyl cellulose (CMC) and xylan with an endo mode of action (Jagtap & Rao, 2005). The primary structure of TSC is not known. The three-dimensional structure of TSC would therefore be a representative of such low-molecular-weight cellulases and should provide information on the active-site residues that are responsible for the bifunctional activity of the enzyme.

2. Experimental procedures

2.1. Crystallization

TSC was purified from the native source as described by Jagtap & Rao (2005). Briefly, the enzyme was purified by fractional ammonium sulfate (30–55%) precipitation followed by sequential chromato-

Table 1
Crystallographic data-collection statistics.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	1.5418
Temperature (K)	100
Space group	$P2_12_12_1$
Unit-cell parameters (Å, °)	a = 49.9, b = 79.5, c = 99.7, $\alpha = \beta = \gamma$ = 90
Unit-cell volume (Å^3)	391552
Matthews coefficient (Å^3 Da^{-1})	2.7
Solvent content (%)	55
No. of molecules in AU	3
Resolution (Å)	20.00–2.30 (2.38–2.30)
Observed reflections	146987
Unique reflections	18357 (1806)
Redundancy	8.0
Completeness (%)	100.0 (100.0)
$R_{\rm merge}$ (%)	10.6 (43.9)
Average $I/\sigma(I)$	18.1 (4.9)

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

graphy on a DEAE-Sephadex A-50 (Sigma) ion-exchange column and a Biogel P100 (Bio-rad) gel-filtration column. The purity of the enzyme was analyzed by SDS–PAGE and was electrophoretically homogeneous, revealing a single 14.2 kDa band. About 10 mg pure protein was obtained from 4 l culture filtrate. The native enzyme was used for all crystallization setups. Initial crystallization conditions were screened by the hanging-drop vapour-diffusion method using commercial crystallization screening kits (Crystal Screen 1 and Crystal Screen 2) from Hampton Research, USA. A protein concentration of 20 mg ml^{-1} (in triple-distilled water), a reservoir volume of 300 µl and a drop size of 4 µl (1:1 ratio of protein and reservoir solution) were used for initial screening. Long irregular rod-shaped crystals (Fig. 1) appeared in a week using the following crystallization conditions from Crystal Screen 1: condition No. 6 (0.2 *M* $MgCl_2$, 0.1 *M* Tris–HCl pH 8.5, 30% PEG 4000), condition No. 18 (0.2 *M* magnesium acetate, 0.1 *M* sodium cacodylate pH 6.5, 20% PEG 8000), condition No. 28 (0.2 *M* sodium acetate, 0.1 *M* sodium cacodylate pH 6.5, 30% PEG 8000) and condition No. 46 (0.2 *M* calcium acetate, 0.1 *M* sodium cacodylate pH 6.5, 18% PEG 8000). The crystal obtained in condition No. 28 was found to be of good quality and size compared with the crystals grown under the other conditions, which were not suitable for diffraction experiments.

2.2. Data collection

A single crystal grown in condition No. 28 (Fig. 1*e*) of approximate dimensions 1.0 × 0.5 × 0.2 mm was used for full diffraction data collection. The diffraction intensity data were measured at 100 K using a MAR Research imaging-plate detector (345 mm diameter) mounted on a Rigaku RU-200 rotating-anode X-ray generator (Cu $K\alpha$ radiation; λ = 1.5418 Å). Crystal freezing included a 3 min soaking of the crystal in a cryoprotectant solution containing the original crystallization solution and 20% glycerol. The pre-soaked crystal was then submitted to immediate flash-freezing directly within a cold nitrogen-gas stream (Oxford Cryosystems). The TSC crystal diffracted X-rays to beyond 2.3 Å. A total of 216 frames were collected with 10 min exposure per frame, 1° oscillation and a crystal-to-detector distance of 120 mm and data were processed using the programs *DENZO* and *SCALEPACK* from the *HKL*2000 package (Otwinowski & Minor, 1997).

3. Results and discussion

Native TSC crystals were grown by the hanging-drop method and diffraction-quality crystals were obtained using crystallization condition No. 28 of Crystal Screen I (Hampton Research, USA). A

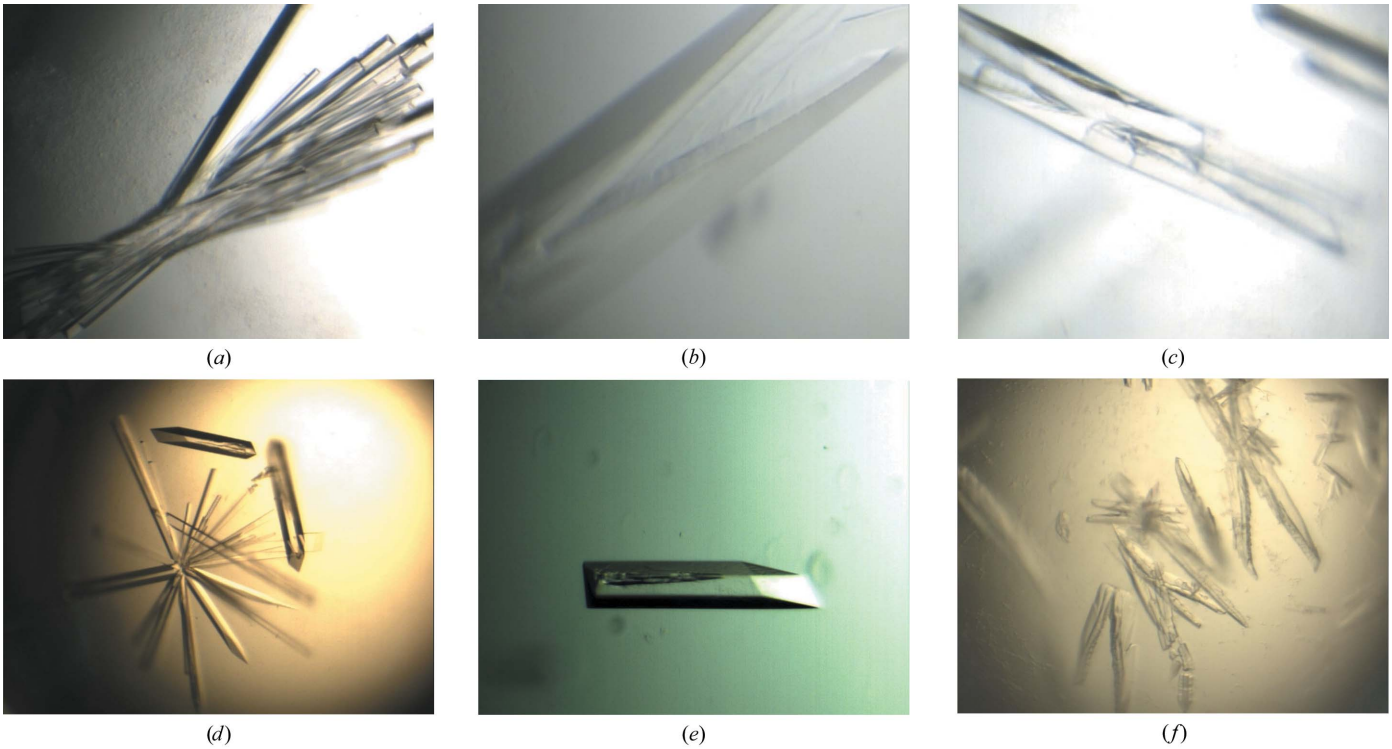

Figure 1
Crystals of native TSC grown in different crystallization conditions from Hampton Crystal Screen 1. The corresponding crystallization condition numbers were (*a*) 6, (*b*) and (*c*) 18, (*d*) and (*e*) 28, (*f*) 46. Except for those in (*d*) and (*e*), the crystals were small and were not used for diffraction experiments. The crystal used for data collection (Fig. 1*e*), with a maximum dimension of 1 mm, was about ten times the linear dimensions of the crystals in (*a*), (*b*), (*c*) and (*f*).

crystal grown in a reservoir solution (300 µl volume) consisting of 0.2 *M* sodium acetate, 0.1 *M* sodium cacodylate pH 6.5 and 20% PEG 8000 was used for diffraction data collection. The crystals belonged to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 49.9, b = 79.5, c = 99.7 Å. Statistics for data collection and processing to 2.3 Å resolution are summarized in Table 1. The Matthews coefficient V_M (Matthews, 1968) was calculated to be 2.7 Å^3 Da^{-1}, suggesting the presence of three molecules in the asymmetric unit. This V_M value corresponds to a solvent content of approximately 55%.

Structure solution is not feasible by the molecular-replacement method as the primary structure of TSC is not known, although the sequences of several trypsin-digested fragments are available, which we hope signals that a full sequence will be forthcoming. Nevertheless, we are in the process of growing better quality crystals in order to obtain an atomic resolution data set which would help in solving the structure by *ab initio* methods. We are also exploring the possibility of obtaining heavy-atom derivatives in order to determine the structure of the TSC enzyme using the method of multiple isomorphous replacement (MIR).

The intensity data were collected at the X-ray Facility for Structural Biology at the Indian Institute of Science, supported by the Department of Science and Technology (DST) and the Department of Biotechnology (DBT). KM and SJ thank CSIR (India) and UGC (India), respectively, for fellowships.

References

Coughlan, M. (1985). *Biochem. Soc. Trans.* **13**, 405–406.
Henrissat, B. & Bairoch, A. (1996). *Biochem. J.* **316**, 695–696.
Henrissat, B. & Davies, G. (1997). *Curr. Opin. Struct. Biol.* **7**, 637–644.
Jagtap, S. & Rao, M. (2005). *Biochem. Biophys. Res. Commun.* **329**, 111–116.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sandgren, M., Stahlberg, J. & Mitchinson, C. (2005). *Prog. Biophys. Mol. Biol.* **89**, 246–291.