Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic analysis of a family 45 endoglucanase from the thermophilic fungus *Melanocarpus albomyces*

Mika Hirvonen and Anastassios C. Papageorgiou*

Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, PO Box 123, Turku 20521, Finland

Correspondence e-mail: apapageo@btk.utu.fi

Received 1 November 2001
Accepted 23 November 2001

20K-cellulase, an endoglucanase produced by the thermophilic fungus *Melanocarpus albomyces*, has been crystallized. It belongs to the family 45 of glycoside hydrolases and has sequence homology with *Humicola insolens* endoglucanase V (EGV). However, in contrast to EGV, it does not harbour a cellulose-binding module. Optimization of the crystallization conditions using a PEG/ion combination and microseeding techniques was employed to improve the quality and the size of the crystals. A complete data set to 2.2 Å resolution was collected using synchrotron radiation. Preliminary crystallographic analysis showed that the crystals belong to the tetragonal space group $P4_12_12/P4_32_12$, with unit-cell parameters $a = 47.3$, $b = 47.3$, $c = 177.3$ Å and one molecule per asymmetric unit.

1. Introduction

Cellulose, a polymer of β-1,4-glucosyl units, is the major polysaccharide component of plant cell walls and is therefore the most abundant carbon resource on Earth, accounting for almost half of the organic material in the biosphere. It is widely used in the paper and textile industry and as a raw material for processed food. Biological processing of cellulose is of great ecological and commercial interest as an alternative for the production of ethanol or other commodity products (Himmel *et al.*, 1999).

Cellulose can form a unique crystalline 'insoluble' substrate, termed cellulose I, that is particularly resistant to microbial degradation. A number of bacterial and fungal organisms use specific enzymes (cellulases) that act synergistically to convert insoluble cellulosic substrates into soluble sugars (Béguin & Aubert, 1994; Bayer *et al.*, 1998). Cellulases catalyse the hydrolysis of β-1,4-glucosidic bonds in cellulose and belong to the glycoside hydrolase families of enzymes that hydrolyse oligosaccharides and/or polysaccharides. Based on sequence similarities, more than 60 families of glycoside hydrolases have been identified thus far, with cellulases being found in at least 13 of them (Henrissat & Bairoch, 1996). Furthermore, cellulases are classified into two large families: cellobiohydrolases (EC 3.2.1.91), which hydrolyse crystalline cellulose by cleavage from the chain ends, and endoglucanases (EC 3.2.1.4), which primarily hydrolyse amorphous and disordered ('soluble') regions of cellulose, cutting at internal glycosidic bonds. Many cellulases are modular enzymes that consist of a catalytic core connected by a flexible and highly glycosylated linker region to a cellulose-binding domain (Gilkes *et al.*, 1991).

Cellulases are used in a wide variety of manufacturing activities. However, the unique characteristics of each cellulase make some of them more desirable for certain purposes than others. Effective use of cellulases in applied processes requires a full understanding of their structure and mechanism. An endoglucanase of molecular weight 20 kDa, namely 20K-cellulase, from the thermophilic fungus *M. albomyces* has been found to be particularly useful in the textile industry for the biostoning of denims (Vehmaanperä *et al.*, 1997). In fabric experiments, 20K-cellulase displayed excellent performance at neutral pH and temperatures up to 343 K. Unlike most cellulases, 20K-cellulase possesses neither a cellulose-binding domain nor a linker region. Sequence alignment has shown that it belongs to the family 45 of glycoside hydrolases and exhibits ~75% sequence identity with the *H. insolens* endoglucanase V catalytic core. Structure determination will allow a detailed structural comparison with *H. insolens* endoglucanase V, the only member of family 45 of glycoside hydrolases whose catalytic core structure is presently known at high resolution (Davies *et al.*, 1995). A methodology for growing single crystals of 20 K-cellulase suitable for X-ray analysis is described.

2. Crystallization

The protein was supplied by Macrocrystal Oy (Espoo, Finland) in a solution containing 2 *M* NaCl. SDS–PAGE showed a single band in the

© 2002 International Union of Crystallography
Printed in Denmark – all rights reserved

expected molecular weight range (~20 kDa) even with overloaded samples; hence, no further purification was pursued. In native PAGE, the protein was found to migrate as a monomer. N-terminal sequence analysis with Edman degradation (Applied Biosystems 477 A Protein Sequencer) confirmed the identity of the protein according to its earlier reported full sequence (Vehmaanperä *et al.*, 1997). The protein solution (~44 mg ml^{-1}) was initially passed through a gel-filtration PD-10 mini-column (Pharmacia) equilibrated with 10 m*M* HEPES buffer pH 7.0 to remove NaCl. Fractions of 0.5 ml were collected and protein concentration was assessed by the Bradford Assay (Biorad). Only fractions that contained more that 10 mg ml^{-1} protein were used for crystallization. Crystals were grown at 289 K by the hanging-drop vapour-diffusion method. Initial crystallization trials produced plate-like crystals by mixing 2 µl of endoglucanase (~15 mg ml^{-1}) with an equal volume of a reservoir solution containing 15%(*w*/*v*) PEG 4000, 0.1 *M* sodium acetate buffer pH 4.6. However, these crystals were extremely thin and diffracted to only 4.0 Å resolution using an R-AXIS II detector mounted on a Rigaku Rotaflex rotating-anode generator (Cu $K\alpha$, λ = 1.5418 Å) operating at 50 kV and 100 mA. Moreover, they always grew in layers from a single nucleation site and displayed profound twinning problems, as became evident from the appearance of multiple diffraction patterns in single images. Although they could be separated carefully by the use of microneedles to produce smaller pieces for X-ray analysis, the overall quality of the resultant diffraction was problematic and not suitable for further structural studies. Optimization of the crystallization conditions was therefore necessary. As the crystals grew in the presence of PEG 4000, a PEG/ion combination based on the PEG/ion screen (Hampton Research, USA) was used as an optimization procedure. Crystals of identical morphology but of considerably better quality were grown using 15%(*w*/*v*) PEG 4000, 0.15 *M* potassium acetate, 0.1 *M* sodium acetate buffer pH 4.6. Although the new crystals displayed the same tendency to grow from a single nucleation site, they were thicker than the previous ones and could be separated more easily. Sodium iodide was also found to have a beneficial effect on the size of the crystals, but the diffraction of these crystals was not good. Further improvement of crystal quality and size was achieved by microseeding. In a typical microseeding experiment, two or three small crystal pieces were crushed in 10 µl of reservoir solution (15% PEG 4000, 0.15 *M* potassium acetate, 0.1 *M* sodium acetate buffer pH 4.6). 1 µl of the seed-stock solution was further diluted to 50 µl with reservoir solution. From this solution, 0.1 µl aliquots were used to seed 2.5 µl drops that had been equilibrated for at least 1 d against the reservoir solution at 289 K. Single crystals suitable for data collection grew within 1–5 d of seeding (Fig. 1).

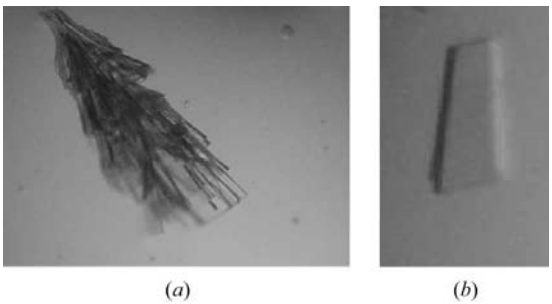

Figure 1
(*a*) A typical cluster of endoglucanase crystals before optimization. (*b*) A single crystal of endoglucanase grown after optimization of the crystallization conditions and microseeding. The dimensions of this crystal are approximately 0.2 × 0.1 × 0.05 mm.

3. Data collection

A complete data set to 2.2 Å resolution was collected from a single crystal at station X11, EMBL Hamburg c/o DESY. The crystal was initially transferred to a reservoir solution containing 25% glycerol as cryoprotectant for ~30 s and was then flash-cooled to 100 K in a nitrogen-gas cold stream (Oxford Cryosystems Cryostream Cooler). Diffraction data were recorded on a MARCCD detector with a diameter of 165 mm. The exposure time was 10–15 s per image (dose mode) and the oscillation range was 0.5°. 120 images were collected in total at a wavelength of 0.8499 Å. Processing, scaling and merging were accomplished with the *HKL* package (Otwinowski & Minor, 1997). Complete data-collection statistics are shown in Table 1. The crystals belong to the tetragonal space group $P4_12_12/P4_32_12$, with unit-cell parameters a = 47.3, b = 47.3, c = 177.3 Å. Assuming one molecule in the asymmetric unit, the Matthews (Matthews, 1968) coefficient is 2.5 Å^3 Da^{-1}, which corresponds to a solvent content of ~50%. The presence of a monomer in the crystal is in agreement with the results of the native gel electrophoresis.

Table 1
Data-collection statistics.

Values in parentheses are for the last resolution shell (2.29–2.20 Å).

Space group	$P4_12_12/P4_32_12$
Unit-cell parameters (Å)	a = 47.3, b = 47.3, c = 177.3
No. of crystals	1
Temperature (K)	100
No. of measurements	228394
No. of unique reflections	10219
Resolution range (Å)	50.0–2.2
Mosaicity (°)	0.45
$R_{\rm merge}$† (%)	6.1 (25.0)
Completeness (%)	92.5 (94.4)
$I/\sigma(I)$	8.0 (3.0)
B factor‡ (Å^2)	24.2

† $R_{\rm merge} = \sum(|I_j - \langle I\rangle|)/\sum\langle I\rangle$, where I_j is the observed intensity of reflection j and $\langle I\rangle$ is the average intensity of multiple observations. ‡ Wilson plot.

4. Discussion

20K-cellulase from the thermophilic fungus *M. albomyces* was crystallized and preliminary crystallographic analysis was carried out. Initial crystals were not suitable for X-ray analysis owing to their thinness and severe twinning problems. Crystals of significantly better quality and larger size were produced by adding potassium acetate to the crystallization medium and using microseeding techniques. The approach described may be of general interest for improvement of crystal quality, particularly when PEG is used as the main precipitant. A complete data set to 2.2 Å resolution was collected. The structure will be determined by molecular replacement using *H. insolens* EGV as a search model. Preliminary results with *AMoRe* (Navaza, 1994) have shown better statistics in the translation function when space group $P4_32_12$ was used. Successful determination of the crystal structure will resolve the space group.

We thank the Academy of Finland for support to initiate this project, Jukka Hellman for the N-terminal sequence analysis, Macrocrystal Oy for providing us with protein and the staff at EMBL Hamburg outstation (c/o DESY, Germany) for help and advice during data collection. Access to EMBL Hamburg (EU contract No. HPRI-CT-1999-00017) is greatly acknowledged.

References

Bayer, E. A., Chanzy, H., Lamed, R. & Sholam, Y. (1998). *Curr. Opin. Struct. Biol.* **8**, 548–557.
Béguin, P. & Aubert, J.-P. (1994). *FEMS Microbiol. Rev.* **13**, 25–58.
Davies, G. J., Tolley, S. P., Henrissat, B., Hjort, C. & Schülein, M. (1995). *Biochemistry*, **34**, 16210–

16220.
Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C. Jr & Warren, R. A. J. (1991). *Microbiol. Rev.* **55**, 303–315.
Henrissat, B. & Bairoch, A. (1996). *Biochem. J.* **316**, 695–696.
Himmel, M. E., Ruth, M. F. & Wyman, C. E. (1999). *Curr. Opin. Biotechnol.* **10**, 358–364.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Vehmaanperä, J., Elovainio, M., Haakana, H., Joutsjoki, V., Lantto, R., Mäntylä, A., Paloheimo, M., Suominen, P., Londesborough, J. & Miettinen-Oinonen, A. (1997). Patent No. WO9714804.