Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of the *Clostridium thermocellum* cellulosome xylanase Z feruloyl esterase domain

David L. Blum,† Florian D. Schubot, Lars G. Ljungdahl, John P. Rose and Bi-Cheng Wang*

Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA

† Present address: Diversa Corporation, 10665 Sorrento Valley Road, San Diego, CA 92121-1609, USA.

Correspondence e-mail: wang@bcl1.bmb.uga.edu

Received 27 December 1999
Accepted 5 May 2000

Feruloyl esterases cleave ferulic acid from arabinoxylan and pectin. Feruloyl groups are believed to crosslink the polysaccharide chain within the polymer and to link hemicellulose to lignin, which may play a role in controlling the growth of plants. The *Clostridium thermocellum* cellulosome xylanase Z feruloyl esterase was expressed in *Escherichia coli*, purified and crystallized. The crystals diffract to 2.4 Å resolution and belong to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 43.14, b = 63.77, c = 79.57 Å. Assuming one molecule per asymmetric unit, the Matthews coefficient is calculated to be 1.87 Å^3 Da^{-1}, which corresponds to a solvent content of 34%.

1. Introduction

Plant cell walls constitute the most abundant renewable natural resource on earth. Their main constituents are cellulose, hemicellulose, pectin and lignin. Hemicellulose in the form of arabinoxylan is predominant in grasses of the Graminae family. Ferulic acid is a constituent of arabinoxylan that is composed of a chain of β1–4 linked xylose units with acetyl, methyl-glucuronyl and arabinosyl substituents (Puls & Schuseil, 1993). The arabinose residues have an ester-linked ferulic acid on carbon 5 which may form cross-links with other xylan chains (Borneman *et al.*, 1993). Hemicellulose and lignin have been shown to be crosslinked by dimers of ferulic acid (Ralph *et al.*, 1995). Feruloyl esterases (FAE) hydrolyze the ferulic acid from arabinoxylan (Fig. 1) and are important for complete cell-wall degradation along with xylanases, which cleave the backbone chain of arabinoxylan (Borneman *et al.*, 1993). A substrate for FAEs is the trisaccharide FAXX (*O*-{5-*O*-[(*E*)-feruloyl]-α-L-arabinofuranosyl}-(1→3)-*O*-β-D-xylopyranosyl-(1→4)-D-xylopyranose) derived from wheat bran (Borneman *et al.*, 1990). A number of bacterial and fungal species that degrade plant cell materials have been found to contain FAEs (Christov & Prior, 1993; Williamson *et al.*, 1998). Sequences of genes encoding FAEs from the bacteria *Butyribivrio fibriosolvens* and *Pseudomonas fluorescens* (Dalrymple & Swadling, 1997; Ferreira *et al.*, 1993) have been reported, as have sequences from the aerobic fungi *Aspergillus niger* and *A. tubingensis* (de Vries *et al.*, 1997) and the anaerobic fungi *Piromyces equi* (Fillingham *et al.*, 1999) and *Orpinomyces* sp. strain PC-2 (Blum *et al.*, 2000). The FAE (EstA) of *P. equi* was shown to be part of a cellulase–hemicellulase complex similar to that of the cellusome of *C. thermocellum*, whereas the *Orpinomyces* FaeA was not part of this fungus. The sequence of the latter showed homology to domains of unknown function in xylanase Y and xylanase Z from *C. thermocellum* (Fontes *et al.*, 1995; Grépinet *et al.*, 1988). No crystal structure of FAE or of other phenolic esterases have been reported.

C. thermocellum is an anaerobic bacterium that produces an enzyme complex termed the cellulosome (Bayer *et al.*, 1998). The cellulosome is composed of a scaffolding protein with nine repeated protein sequences called cohesin domains that bind catalytic subunits through their dockerin domain. One of the catalytic subunits, xylanase Z (XynZ), has a modular structure (Fig. 2) composed of a xylanase domain, a dockerin domain for binding to the cohesin domain of the scaffolding protein, a family VI cellulose-binding domain and the unknown domain recently identified as FAE (Blum *et al.*, 2000). While the xylanase domain of XynZ has been crystallized and its crystal structure determined at 1.4 Å resolution (Souchon *et al.*, 1994; Dominguez *et al.*, 1995), little structural information is available for the other domains of XynZ. Here, we report the crystallization and preliminary X-ray diffraction analysis of the XynZ FAE domain (residues 20–286 of XynZ) from *C. thermocellum*. The three-dimensional structure of this enzyme will contribute to our understanding of the enzymatic mechanism and substrate specificity of feruloyl esterases and of the interaction between two domains of a protein catalyzing different reactions in the hydrolysis of a common substrate.

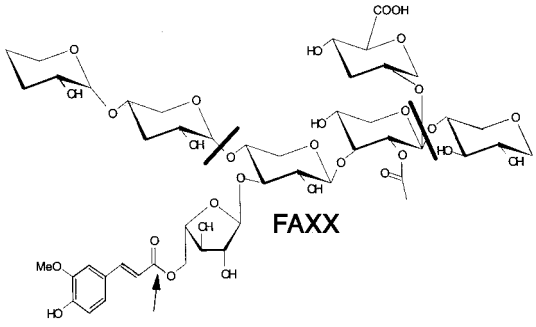

Figure 1
A schematic drawing of FAXX (inside the black bars) within a simplified arabinoxylan structure. An arrow signifies the bond broken by FAE.

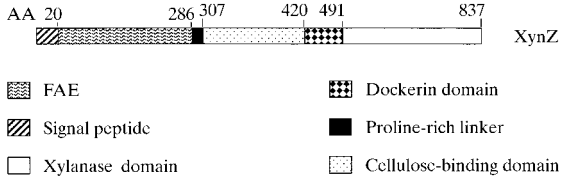

Figure 2
A schematic diagram of the domain organization of the xylanase Z gene from *C. thermocellum*.

Figure 3
Crystal of the *C. thermocellum* cellulosome xylanase Z feruloyl esterase domain. The crystal shown is approximately 0.01 × 0.05 × 0.3 mm. The crystals belong to space group $P2_12_12_1$, with unit-cell parameters a = 43.2, b = 63.9, c = 79.6 Å and diffract to 2.4 Å resolution.

2. Materials and methods

2.1. Crystallization

The XynZ FAE domain from *C. thermocellum* was expressed in *E. coli* and purified as reported previously (Blum *et al.*, 2000) by heat treatment at 343 K for 30 min followed by gel filtration with a TSK 3000SW column (Tosohaas, Montgomeryville, PA, USA). The expressed protein released ferulic acid from FAXX and other substrates, demonstrating that it is enzymatically active.

The enzyme was crystallized by the hanging-drop vapor-diffusion technique at 291 K using 2 µl drops. Each drop contained equal volumes of protein solution (~20 mg ml^{-1}) and reservoir solution containing 0.1 *M* sodium acetate trihydrate pH 4.6, 0.2 *M* ammonium sulfate, 30%(*w*/*v*) polyethylene glycol monomethyl ether 2000 (Hampton Research) and 15%(*v*/*v*) glycerol.

2.2. Data collection

For data collection, the crystal was mounted in a 0.2 mm diameter rayon loop (Teng, 1990) containing a minimal amount of mother liquor and was flash-cooled (Hope, 1988) to 93 K in a nitrogen-gas cold stream (Oxford Instruments). No cryoprotectant was required to stabilize the crystal. Data to 2.4 Å resolution (λ = 1.54 Å) were recorded on a MAR Research 165 mm CCD detector at ID17 (IMCA-CAT), Advanced Photon Source, Argonne National Laboratory. Details of the data collection are given in Table 1. The data were integrated and scaled using *HKL*2000 (Otwinowski & Minor, 1997). The data set was 99.7% complete to 2.4 Å resolution and gave an internal agreement (R_{sym}) of 8.2%.

Table 1
Data-collection parameters.

Values in parentheses refer to the highest resolution shell.

Crystal dimensions (mm)	0.01 × 0.05 × 0.30
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 43.2, b = 63.9, c = 79.6
Molecules per asymmetric unit	1
Wavelength (Å)	1.54
Crystal-to-detector distance (mm)	100
Detector 2θ (°)	0.0
Maximum resolution of data (Å)	2.4
Rotation for each exposure (°)	1
Temperature (K)	93
Time for each exposure (s)	4
Total rotation range (°)	540
Average $I/\sigma(I)$	28 (7)
Average redundancy	>19 (13)
R_{merge} (I)	0.082 (0.27)
Completeness (%)	99.9 (99.7)

3. Results and discussion

Crystallization trials were originally attempted with a larger protein that contained both the FAE and cellulose-binding domains (Fig. 2) and had a molecular mass of 45 kDa (Blum *et al.*, 2000). The solubility of this construct was poor and dynamic light scattering showed severe aggregation; all attempts to crystallize this construct failed to produce crystals.

Crystallization trials on the FAE domain alone, however, produced crystals suitable for diffraction analysis. The crystals appeared in 7–14 d and grew to dimensions of 0.01 × 0.05 × 0.3 mm (Fig. 3) in four to five weeks. The crystal indexed in an orthorhombic lattice with unit-cell parameters a = 43.14, b = 63.77, c = 79.57 Å. Analysis of the three-dimensional diffraction pattern (*XPREP*; Sheldrick, 1991) indicated that the space group was $P2_12_12_1$. Based on one molecule per asymmetric unit, the Matthews coefficient (Matthews, 1968) is calculated to be 1.87 Å^3 Da^{-1}, corresponding to a solvent content of 34%, which is well within the range for most protein crystals. The structure determination is in progress. Since the search for a molecular-replacement solution was not successful, the structure will be solved *via* heavy-atom derivatization of the crystals.

Use of the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science under Contract No. W-31-109-Eng-38. Data were collected at beamline 17-ID in the facilities of the Industrial Macromolecular Crystallography Association Collaborative Access Team (IMCA-CAT). These facilities are supported by the companies of the Industrial Macromolecular Crystallography Association through a contract with Illinois Institute of Technology (IIT) executed through the IIT's Center for Synchrotron Radiation Research and Instrumentation. Bayer Corporation and Arthur Robbins are thanked for the use of Bayer's share in access to the IMCA-CAT beamline. Work supported in part by grant DE-FG02-93ER20127 from the department of Energy (LGL).

References

Bayer, E. A., Shimon, L. J. W., Shoham, Y. & Lamed, R. (1998). *J. Struct. Biol.* **124**, 221–234.
Blum, D. L., Kataeva, I., Li, X.-L. & Ljungdahl, L. G. (2000). Submitted.
Borneman, W. S., Hartley, R. D., Himmelsbach, D. S. & Ljungdahl, L. G. (1990). *Anal. Biochem.* **190**, 129–133.

Borneman, W. S., Ljungdahl, L. G., Hartley, R. D. & Akin, D. E. (1993). *Hemicellulose and Hemicellulases*, edited by M. P. Coughlan & G. P. Hazlewood, pp. 85–102. London: Portland Press.
Christov, L. P. & Prior, B. A. (1993). *Enzyme Microb. Technol.* **15**, 460–475.
Dalrymple, B. D. & Swadling, Y. (1997). *Microbiology*, **143**, 1203–1210.
Dominguez, R., Souchon, H., Spinelli, S., Dauter, Z., Wilson, K. S., Chavaux, S., Beguin, P. & Alzari, P. M. (1995). *Nature Struct. Biol.* **2**, 569–576.
Ferreira, L. M. A., Wood, T. M., Williamson, G., Faulds, C., Hazlewood, G. P., Black, G. W. & Gilbert, H. J. (1993). *Biochem. J.* **294**, 349–355.
Fillingham, I. J., Kroon, P. A., Williamson, G., Gilbert, H. J. & Hazlewood, G. P. (1999). *Biochem. J.* **343**, 215–224.
Fontes, C. M. G. A., Hazlewood, G. P., Morag, E., Hall, J., Hirst, B. H. & Gilbert, H. J. (1995). *Biochem. J.* **307**, 151–158.
Grépinet, O., Chebrou, M.-C. & Béguin, P. (1988). *J. Bacteriol.* **170**, 4576–4581.
Hope, H. (1988). *Acta Cryst.* B**44**, 22–26.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Puls, J. & Schuseil, J. (1993). *Hemicellulose and Hemicellulase*, edited by M. P. Coughlan & G. P. Hazlewood, pp. 1–28. London: Portland Press.
Ralph, J., Grabber, J. H. & Hatfield, R. D. (1995). *Carbohydr. Res.* **275**, 167–178.
Sheldrick, G. (1991). *XPREP – Space Group Determination and Reciprocal Space Plots.* Siemens Analytical X-ray Instruments, Madison, WI, USA.
Souchon, H., Spinelli, S., Beguin, P. & Alzari, P. M. (1994). *J. Mol. Biol.* **235**, 1348–1350.
Teng, T. Y. (1990). *J. Appl. Cryst.* **23**, 387–391.
Vries, R. P. de, Michaelsen, B., Poulsen, C. H., Kroon, P. A., van den Heuvel, R. H. H., Fauds, C. B., Williamson, G., van den Hombergh, J. P. T. & Visser, J. (1997). *Appl. Environ. Microbiol.* **63**, 4638–4644.
Williamson, G., Kroon, P. A. & Faulds, C. B. (1998). *Microbiology*, **144**, 2011–2023.