Acta Cryst. (1997). D53, 114–115

Crystallization and preliminary X-ray analysis of a cohesin domain of the cellulosome from *Clostridium thermocellum*

LINDA J. W. SHIMON,[a]* FELIX FROLOW,[a] SIMA YARON,[b] EDWARD A. BAYER,[c] RAPHAEL LAMED,[d] ELY MORAG[c] AND YUVAL SHOHAM[b] at [a]*Departments of Chemical Services and Structural Biology, The Weizmann Institute of Science, Rehovot,* [b]*Department of Food Engineering and Biotechnology, The Technion, Haifa,* [c]*Department of Membrane Research and Biophysics, The Weizmann Institute of Science, Rehovot, Israel, and* [d]*Department of Molecular Microbiology and Biotechnology, Tel Aviv University, Ramat Aviv. E-mail: cvlinda@weizmann.weizmann.ac.il*

(*Received 7 June 1996; accepted 19 September 1996*)

Abstract

Recombinant cohesin-2, a unique type of protein-recognition domain from the cellulosome of *Clostridium thermocellum*, has been crystallized by the hanging-drop vapor-diffusion method. The crystals are monoclinic, space group $C2$ with unit-cell dimensions $a = 79.91$, $b = 47.86$, $c = 51.13$ Å, $\beta = 126.77°$. There is most likely to be one molecule per asymmetric unit, corresponding to a packing density of 2.16 Å^3 Da^{-1}. The crystals diffract to beyond 2.3 Å on a conventional laboratory rotating-anode source.

1. Introduction

In recent years interest has focused on the efficient degradation of cellulose (a major pollutant worldwide in the form of waste paper) by a highly specialized class of cellulolytic micro-organisms (Béguin & Aubert, 1994). *Clostridium thermocellum* is a particularly efficient cellulolytic bacterium which occurs in anaerobic thermophilic ecosystems. Its cellulase activity is carried out predominantly by a multi-functional, multi-enzyme complex, termed the cellulosome (Lamed, Setter, Kenig & Bayer, 1983; Lamed & Bayer, 1988). In order to form this very large complex, the bacterium produces a special multiple-domain polypeptide, called scaffoldin, which organizes the catalytic (enzymatic) subunits into the complex (Bayer, Setter & Lamed, 1985; Bayer, Morag & Lamed, 1994; Wu, Orme-Johnson & Demain, 1988; Gerngross, Romaniec, Kobayashi, Huskisson & Demain, 1993). The 210 kDa scaffoldin subunit comprises a single cellulose-binding domain and nine distinct but closely related copies of a strong protein-recognizing domain, the cohesin domain. The cohesin domains serve to bind the different cellulose-degrading enzymes into the cellulosome complex by interacting selectively with complementary domains (called dockerin domains) located on the individual enzymes. The cohesin–dockerin interaction thus appears to be the key to cellulosome integrity (Tokatlidis, Salamitou, Béguin, Dhurjati & Aubert, 1991; Yaron, Morag, Bayer, Lamed & Shoham, 1995; Kruus, Lua, Demain & Wu, 1995), and determination of the structures of the complementary domains is crucial to the understanding of the assembly of the cellulosome. In this communication we report the crystallization and preliminary X-ray analysis of one of the cohesin domains (cohesin-2) from the cellulosome of *C. thermocellum*.

2. Materials and methods

2.1. *Purification and crystallization*

Cohesin-2 of the *C. thermocellum* cellulosome was over-expressed with a 6xHis tag in *Escherichia coli* XL-1 Blue strain and was purified to electrophoretic homogeneity by isolation using Ni-NTA affinity chromatography followed by gel filtration using fast protein liquid chromatography. Details of gene cloning and protein purification are described elsewhere (Yaron *et al.*, 1997). Protein samples were dialyzed into 50 m*M* 2[*N*-morpholino]ethane-sulfonic acid buffer pH 7.0, 200 m*M* NaCl and stored at 2 mg ml^{-1} at 277 K until further use.

Cohesin-2 was crystallized at room temperature using the hanging-drop vapor-diffusion method (McPherson, 1982). Preliminary crystallization conditions were determined by the sparse-matrix sampling method (Jancarik & Kim, 1991). Tetragonal, hexagonal and monoclinic forms were obtained, however, monoclinic crystals were chosen for further characterization as they were found to give the highest resolution. In the final conditions for crystallization the reservoirs contained 28%(*w*/*v*) polyethylene glycol 6000, 750 m*M* LiCl and 100 m*M* Tris buffer pH 8.6. Protein samples (2 mg ml^{-1}) were mixed in equal amounts with the reservoir solution and allowed to equilibrate. Crystals grew within 3–4 weeks but the larger ones were often twinned. To obtain single crystals, small crystals were used as seeds and transferred to other drops, set up as above, but where no crystals had yet appeared. These seeds most often remained single and grew to a final size of 0.3 × 0.3 × 0.2 mm (Fig. 1).

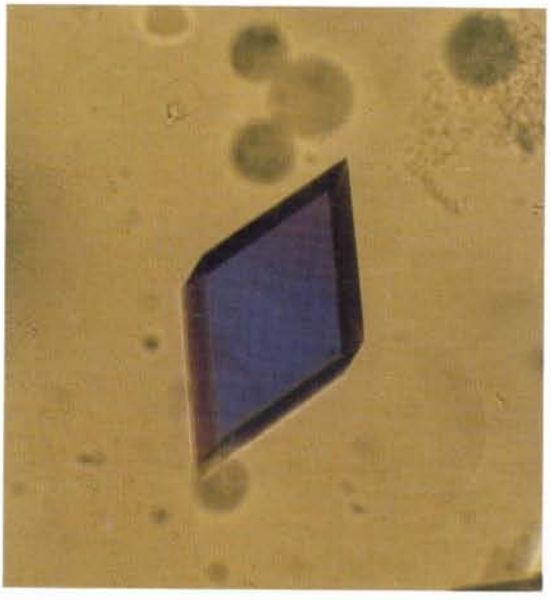

Fig. 1. Crystal of cohesin-2 from *Clostridium thermocellum*, monoclinic form. The largest dimension is 0.3 mm.

Table 1. *Statistics for the X-ray data as a function of the resolution*

Resolution (Å)	R_{merge}	No. of unique reflections	Completeness (%)	Multiplicity	$\langle I\rangle/\sigma(I)$	$I > 3\sigma(I)$ (%)
50.00–4.90	0.044	738	98.7	3.20	37.40	99.6
4.90–3.89	0.053	720	99.4	3.27	34.24	98.3
3.89–3.40	0.067	717	99.6	3.31	28.21	95.6
3.40–3.09	0.081	714	99.7	3.33	22.53	95.1
3.09–2.87	0.094	717	99.6	3.33	17.16	91.5
2.87–2.70	0.120	704	99.7	3.30	11.97	79.7
2.70–2.56	0.139	711	99.9	3.34	9.96	74.4
2.56–2.45	0.176	718	100.0	3.29	7.84	68.9
Overall	0.063	5739	99.6	3.30	27.24	88.0

2.2. *X-ray analysis*

Crystals were mounted in thin-walled glass capillaries. A full native data set to 2.45 Å resolution was collected as 1.5° oscillation frames on a Rigaku R-AXIS IIC detector (equipped with blue image plates) mounted on a Rigaku FRC Cu rotating anode (focal cup 0.1 × 0.1 mm, operating at 50 kV, 55 mA) equipped with focusing mirrors. Data frames were processed with the *DENZO* and *SCALEPACK* packages (Otwinowski, 1993).

3. Results and discussion

Single crystals of cohesin-2 were obtained by seeding. The crystals were stable in the X-ray beam and diffract to at least 2.3 Å on a conventional X-ray source. A full native data set to 2.45 Å resolution was collected. Indexing by the autoindexing routine in *DENZO* gave a monoclinic space group $C2$ with unit-cell dimensions of $a = 79.51$, $b = 47.86$, $c = 51.13$ Å, $\beta = 126.77°$. Data processing gave an R_{merge} of 6.6% (based on intensities). The data set is 99.6% complete to 2.45 Å (100% complete in the range 2.6–2.45 Å) (Table 1). Assuming a calculated molecular weight of 17 993 Da and one molecule per asymmetric unit, the V_m value is 2.16 Å^3 Da^{-1}, which lies in the normal range for globular proteins (Matthews, 1968). Screening for heavy-atom derivatives is currently under way.

References

Bayer, E. A., Morag, E. & Lamed, R. (1994). *Trends Biotechnol.* **12**, 378–386.

Bayer, E. A., Setter, E. & Lamed, R. (1985). *J. Bacteriol.* **163**, 552–559.

Béguin, P. & Aubert, J.-P. (1994). *FEMS Microbiol. Lett.* **13**, 25–58.

Gerngross, U. T., Romaniec, M. P. M., Kobayashi, T., Huskisson, N. S. & Demain, A. L. (1993). *Mol. Microbiol.* **8**, 325–334.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

Kruus, K., Lua, A. C., Demain, A. L. & Wu, J. H. D. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 9254–9258.

Lamed, R. & Bayer, E. A. (1988). *Adv. Appl. Microbiol.* **33**, 1–46.

Lamed, R., Setter, E., Kenig, R. & Bayer, E. A. (1983). *Biotechnol. Bioeng. Symp.* **13**, 163–181.

McPherson, A. (1982). *Preparation and Analysis of Protein Crystals.* New York: John Wiley.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Otwinowski, Z. (1993). *Data Collection and Processing, Proceedings of the CCP4 Study Weekend*, pp. 56–62. Warrington: Daresbury Laboratory.

Tokatlidis, K., Salamitou, S., Béguin, P., Dhurjati, P. & Aubert, J.-P. (1991). *FEBS Lett.* **291**, 185–188.

Wu, J. H. D., Orme-Johnson, W. H. & Demain, A. L. (1988). *Biochemistry*, **27**, 1703–1709.

Yaron, S., Morag, W., Bayer, E. A., Lamed, R. & Shoham, Y. (1995). *FEBS Lett.* **360**, 121–124.

Yaron, S., Shimon, L. J. W., Frolow, F., Lamed, R., Morag, W., Shoham, Y. & Bayer, E. A. (1997). *J. Biotechnol.* In the press.