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

Crystallization and preliminary X-ray analysis of *Acetivibrio cellulolyticus* cellulosomal type II cohesin module: two versions having different linker lengths

Ilit Noach,[a] Orly Alber,[a] Edward A. Bayer,[a]* Raphael Lamed,[b,c] Maly Levy-Assaraf,[b] Linda J. W. Shimon[d] and Felix Frolow[b,c]*

[a]Department of Biological Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel, [b]Department of Molecular Microbiology and Biotechnology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Ramat Aviv 69978, Israel, [c]The Daniella Rich Institute for Structural Biology, Tel Aviv University, Ramat Aviv 69978, Israel, and [d]Department of Chemical Research Support, The Weizmann Institute of Science, Rehovot 76100, Israel

Correspondence e-mail: ed.bayer@weizmann.ac.il, mbfrolow@post.tau.ac.il

Received 15 November 2007
Accepted 13 December 2007

The second type II cohesin module of the cellulosomal scaffoldin polypeptide ScaB from *Acetivibrio cellulolyticus* (CohB2) was cloned into two constructs: one containing a short (five-residue) C-terminal linker (CohB2_S) and the second incorporating the full native 45-residue linker (CohB2_L). Both constructs encode proteins that also include the full native six-residue N-terminal linker. The CohB2_S and CohB2_L proteins were expressed, purified and crystallized in the orthorhombic crystal system, but with different unit cells and symmetries: space group $P2_12_12_1$ with unit-cell parameters $a = 90.36$, $b = 68.65$, $c = 111.29$ Å for CohB2_S and space group $P2_12_12$ with unit-cell parameters $a = 68.76$, $b = 159.22$, $c = 44.21$ Å for CohB2_L. The crystals diffracted to 2.0 and 2.9 Å resolution, respectively. The asymmetric unit of CohB2_S contains three cohesin molecules, while that of CohB2_L contains two molecules.

1. Introduction

The extracellular multi-enzyme cellulosome complex was first described in 1983 (Lamed, Setter & Bayer, 1983) and represents one of nature's elegant examples of 'biomolecular engineering'. In this context, the cellulosome comprises an assorted group of cellulases, hemicellulases and other proteins, all combined into a highly ordered suprastructure that is able to utilize very efficiently the recalcitrant cellulose and hemicellulose polysaccharide substrates of the plant cell wall (Bayer *et al.*, 2000). The first example of a cellulosome complex was discovered in the anaerobic thermophilic bacterium *Clostridium thermocellum* (Lamed, Setter, Kenig *et al.*, 1983; Lamed, Setter & Bayer, 1983). The cellulosomal enzyme subunits were found to be united into the complex by means of a unique class of nonenzymatic multimodular polypeptide subunits termed 'scaffoldin' (Bayer *et al.*, 1994). The scaffoldin contains multiple copies of a distinctive type of module, termed 'cohesin'. The cellulosomal enzyme subunits, on the other hand, contain a complementary type of module, termed 'dockerin'. The high-affinity interaction between the cohesin and dockerin modules provides the definitive molecular mechanism that integrates the enzyme subunits into the cellulosome complex (Lytle *et al.*, 1996; Salamitou *et al.*, 1994).

The cellulosome system of the mesophilic cellulolytic bacterium *Acetivibrio cellulolyticus* is, like *C. thermocellum*, characterized by multi-scaffoldin gene clusters (Bayer *et al.*, 2004) and the cellulosome components are attached to the cell surface *via* an SLH-containing scaffoldin. However, the modular architecture and composition of its components differ distinctly from those of *C. thermocellum*. In the *A. cellulolyticus* system, the products of three successive genes, the primary scaffoldin ScaA, an 'adaptor' scaffoldin ScaB and an anchoring scaffoldin ScaC, presumably self-assemble to generate a supramolecular construction capable of containing 96 individual dockerin-containing enzyme components (Ding *et al.*, 1999; Xu *et al.*, 2003).

Historically, the cohesins and dockerins have been divided into different types according to phylogenetic sequence analysis (Bayer *et*

al., 2004). The type II cohesins differ somewhat from those of type I, having additional segments within the polypeptide and clear diversity in the latter half of the sequence (Leibovitz & Béguin, 1996). The type II cohesin discussed in this communication is the second cohesin (enumerated from the N-terminus) of the ScaB scaffoldin polypeptide chain and shares 50% sequence identity with the first cohesin in this scaffoldin, the crystal structure of which has been solved previously (PDB code 1ohz; Noach *et al.*, 2003). The cohesins on the scaffoldin subunits are separated by distinctive linker sequences, most of which are rich in proline and threonine residues. Such extended flexible linkers physically separate the various modules of the scaffoldins, and the proline-rich regions are believed to promote intermodular and/or intersubunit protein–protein interactions (Woessner & Goodenough, 1992).

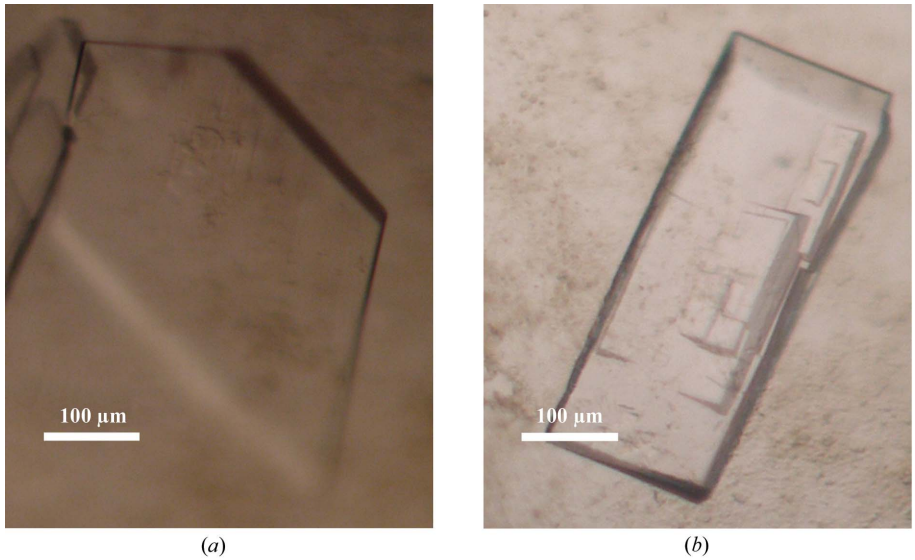
To date, the crystal structures of three different type I cohesin modules have been determined: two of them are from *C. thermocellum* (Shimon *et al.*, 1997; Tavares *et al.*, 1997) and the third is from *C. cellulolyticum* (Spinelli *et al.*, 2000). Following this work, the structure of a type I cohesin–dockerin complex from *C. thermocellum* was determined, which revealed a dual-binding mode of interaction between the two complementary modules (Carvalho *et al.*, 2003). In addition, two type II cohesin structures from *A. cellulolyticus* and *Bacteroides cellulosolvens* (Noach *et al.*, 2003, 2005) have been described, followed by that of another type II cohesin from *C. thermocellum*, first as a single module (Carvalho *et al.*, 2005) and subsequently in a trimodular complex with its counterpart dockerin together with an adjacent X module (Adams *et al.*, 2006).

The cohesin structure forms an elongated nine-stranded β-sandwich. The overall jelly-roll topology of the type II cohesins is very similar to that observed for the type I cohesins, with three additional secondary structures: an α-helix and two 'β-flaps' that disrupt the normal course of a β-strand. In the present communication, we describe the crystallization and preliminary X-ray characterization of two constructs of the second type II cohesin (CohB2) from the *A. cellulolyticus* ScaB scaffoldin. Both constructs contain the native N-terminal linker, but only one of them contains the full-length 45-residue C-terminal linker (CohB2_L); the other contains a short version of this linker (five residues in length; CohB2_S).

2. Materials and methods

2.1. Expression and purification

The DNA encoding the second cohesin module from the *scaB* scaffoldin gene (gi:31540574; Noach *et al.*, 2003; Xu *et al.*, 2003) was cloned, with either the short or full-length C-terminal linker segment, from genomic DNA of *A. cellulolyticus* into the *Nco*I and *Xho*I sites of the pET28a expression vector (Novagen). The constructs, which contained two additional restriction-site residues (Leu-Glu) followed by a hexa-His tag attached to the 3′ end, were expressed in *Escherichia coli* strain BL21 cells grown at 310 K in 1 l Luria–Bertani medium supplemented with 50 µg ml^{-1} kanamycin. Protein expression was induced with 0.1 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) at an OD_{600} of 0.6 and cultivation was continued at 310 K for 3 h. Cells were harvested by centrifugation (3000*g* for 15 min) at 277 K and resuspended in a volume of buffer *A* (25 m*M* Tris pH 7.4, 150 m*M* NaCl, 3 m*M* KCl) equivalent to four times their packed weight. The suspension was kept on ice during sonication (6 × 30 s), after which cell debris was removed by centrifugation (20 000*g* at 277 K for 30 min). The supernatant solution was collected and applied onto a column packed with Ni^{2+}-chelating Sepharose beads (Amersham Pharmacia Biotech) equilibrated with buffer *A*. The column was washed with buffer *A* containing 5 m*M* imidazole and bound protein was eluted with buffer *A* containing 150 m*M* imidazole in 1 ml fractions. Fractions containing protein were analyzed by SDS–PAGE; those containing purified protein were pooled and diluted by a factor of five with double-distilled water to a final solution content of 5 m*M* Tris pH 7.4, 30 m*M* NaCl, 0.6 m*M* KCl and 30 m*M* imidazole. The single-step purification of the recombinant proteins was justified in both cases by the appearance of a single purified band on SDS–PAGE. Typical yields were 130 and 95 mg per litre for CohB2_S and CohB2_L, respectively. The purified proteins (CohB2_S and CohB2_L) were concentrated to about 12 mg ml^{-1} using a Vivaspin centrifugal filter device of 5000 kDa molecular-weight cutoff (VivaScience). The protein concentration was estimated by measuring the absorbance at 280 nm using the calculated extinction coefficient of the proteins (ε_{280} = 20 400 for both CohB2_S and CohB2_L).



(*a*) (*b*)

Figure 1
Crystals of cohesin B2, the second type II cohesin module from scaffoldin B of *A. cellulolyticus*. (*a*) A typical crystal of CohB2_S, containing the short C-terminal linker (five residues). The approximate dimensions of the crystal are 0.6 × 0.3 × 0.06 mm. (*b*) A representative CohB2_L crystal containing the full C-terminal linker length (45 residues). The crystal dimensions are approximately 0.6 × 0.2 × 0.08 mm.

2.2. Crystallization

Crystallization was performed using Hampton Research Crystal Screen and the hanging-drop technique at 293 K. After an initial trial and a round of optimization, the final crystallization conditions were established. Type II CohB2_S crystals were obtained after 10 d in a 6 µl drop containing 4 µl CohB2_S solution (12 mg ml^{-1}) and 2 µl reservoir solution consisting of 0.2 *M* ammonium sulfate, 0.1 *M* sodium acetate trihydrate pH 4.6 and 25%(*w*/*v*) PEG 4000 (condition No. 20; Fig. 1*a*). Type II CohB2_L crystals formed after 5 d in a 4 µl drop containing 2 µl CohB2_L (12 mg ml^{-1}) and 2 µl reservoir solution [0.1 *M* 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid pH 7.5, 10%(*v*/*v*) 2-propanol and 17%(*w*/*v*) PEG 4000; optimized condition No. 41; Fig. 1*b*]. 0.5 ml reservoir volumes were used.

2.3. Data collection and processing

Crystals were harvested from the crystallization drop surrounded by mother liquor using a thin-walled glass capillary. The capillary was sealed from one side using a flame and from the funnel side with high-vacuum grease to facilitate transfer to the synchrotron. X-ray diffraction data sets from both CohB2 crystals were collected on a ADSC CCD detector using synchrotron radiation (ID14-3 beamline at ESRF, Grenoble, France) with an X-ray beam of 0.9393 Å wavelength. The crystals, which were dismounted from the capillary into a solution mimicking the reservoir solution supplemented with 25% ethylene glycol, were incubated for a fraction of a minute, mounted on a MiTeGen stiff micro-mount (http://www.mitegen.com/) made of polyimide and placed in a stream of cold nitrogen at a temperature of 100 K generated by an Oxford Cryosystem (Cosier & Glazer, 1986). All data sets were collected in 0.5° oscillation frames and were indexed, processed and scaled using *DENZO* and *SCALEPACK* as implemented in *HKL*-2000 (Otwinowski & Minor, 1997; Table 1).

CohB2_S crystals diffracted to 2.0 Å resolution and belong to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 90.36$, $b = 68.65$, $c = 111.29$ Å. The presence of several molecules of CohB2_S in the asymmetric unit was detected. The calculated Matthews coefficients (Matthews, 1968) of 2.79 and 2.09 Å^3 Da^{-1} are in accordance with solvent contents of 55.97 and 41.30% and correspond to three and four monomers in the asymmetric unit, respectively, assuming a protein molecular weight of 20.6 kDa. Pseudo-translation was not detected in the Patterson map (data not shown). Two independent peaks were detected on the $\kappa = 180^\circ$ section of the self-rotation function in polar coordinates, calculated by *POLARRFN* as implemented in *CCP*4 (Collaborative Computational Project, Number 4, 1994; Fig. 2*a*). In this case, even the self-rotation function failed to solve the ambiguity in the number of molecules in the asymmetric unit.

CohB2_L crystals diffracted to 2.9 Å and also belong to the primitive orthorhombic space group, with $P2_12_12$ symmetry and unit-cell parameters $a = 68.76$, $b = 159.22$, $c = 44.21$ Å. A calculated Matthews coefficient of 2.46 Å^3 Da^{-1} is in accordance with a solvent content of 50.03% and corresponds to the presence of two molecules in the asymmetric unit, assuming a protein molecular weight of 24.6 kDa (Table 1). No pseudo-translation was detected in the Patterson function map. One independent peak was detected on the $\kappa = 180^\circ$ section of the self-rotation function in polar coordinates (Fig. 2*b*), which corresponds to the presence of two molecules in the asymmetric unit. Inspection of the diffraction data collected from the orthorhombic crystal of CohB2_L relative to the reciprocal-lattice directions revealed anisotropic diffraction, with lower resolution in the c^* direction, indicating that the resolution is orientation- and direction-dependent. The relatively low completeness of the diffraction data of CohB2_L (see Table 1) was caused by overlapping of the diffraction maxima in several regions owing to a combination of experimental factors such as orientation, unit-cell parameters, mosaicity, crystal-to-detector distance, wavelength and the relatively small detector size.

2.4. Initial crystal structure solution

The molecular-replacement program *MOLREP* (Vagin & Teplyakov, 1997) as implemented in *CCP*4 (Collaborative Compu-

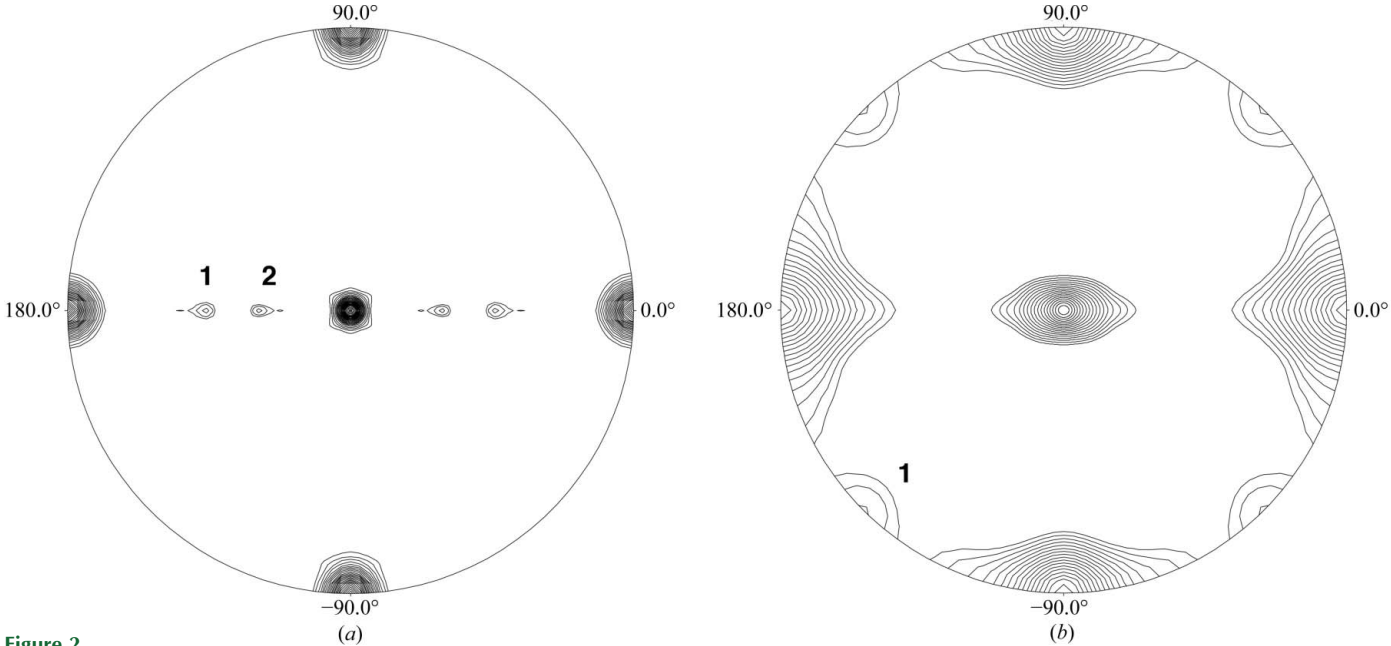


Figure 2
Self-rotation functions, $\kappa = 180^\circ$ section, with an integration radius of 20 Å. (*a*) Self-rotation function for CohB2_S calculated using a 22.7–2.0 Å resolution range and rendered beyond 3σ. Two independent peaks are observed and are numbered 1 and 2. (*b*) Self-rotation function for the CohB2_L molecule calculated using a 38.6–2.5 Å resolution range and rendered beyond 2σ. One peak is observed.

Table 1
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

	CohB2_S	CohB2_L
X-ray source	ESRF (ID14-3)	ESRF (ID14-3)
Wavelength (Å)	0.9393	0.9393
Detector	ADSC CCD	ADSC CCD
Crystal-to-detector distance (mm)	247.5	394.3
Space group	$P2_12_12_1$	$P2_12_12$
Unit-cell parameters		
a (Å)	90.36	68.76
b (Å)	68.65	159.22
c (Å)	111.29	44.21
Resolution (Å)	2.00 (2.07–2.00)	2.90 (2.95–2.90)
V_M (Å^3 Da^{-1})	2.79	2.46
Solvent content (%)	55.97	50.03
Monomers in ASU	3	2
Data processing		
No. of observed reflections	357689	41898
No. of unique reflections	47999	11536
Completeness (%)	95.5 (74.9)	84.7 (79.8)
Mean $I/\sigma(I)$	19.6 (2.1)	22.8 (2.7)
R_{merge}†	0.096 (0.48)	0.064 (0.23)

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \overline{I(hkl)}|/\sum_{hkl}\sum_i I_i(hkl)$, where $\sum_{hkl}$ denotes the sum over all reflections and $\sum_i$ the sum over all equivalent and symmetry-related reflections (Stout & Jensen, 1968).

tational Project, Number 4, 1994) was employed for phase determination of the CohB2_S data set using the previously solved type II cohesin structure (PDB code 1zv9) as a search model. The *MOLREP* statistics yielded an initial solution with an R factor of 0.552 and a correlation coefficient of 0.487. The CohB2_S structure revealed the presence of three independent molecules in the asymmetric unit. In the case of CohB2_L, the structure was solved by *MOLREP* using the coordinates of the newly solved structure of CohB2_S. The initial solution exhibited an R factor of 0.470 and a correlation coefficient of 0.654, with two molecules in the asymmetric unit. Model building and refinement of the structures are currently in progress.

We acknowledge the ESRF for synchrotron beam time and the staff scientists of the ID14 beamline cluster for their assistance. This research was supported by the Israel Science Foundation (grant Nos. 422/05 and 159/07) and by grants from the United States–Israel Binational Science Foundation (BSF), Jerusalem, Israel.

References

Adams, J. J., Pal, G., Jia, Z. & Smith, S. P. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 305–310.
Bayer, E. A., Belaich, J.-P., Shoham, Y. & Lamed, R. (2004). *Annu. Rev. Microbiol.* **58**, 521–554.
Bayer, E. A., Morag, E. & Lamed, R. (1994). *Trends Biotechnol.* **12**, 378–386.
Bayer, E. A., Shoham, Y. & Lamed, R. (2000). *The Prokaryotes: An Evolving Electronic Resource for the Microbiological Community*, 3rd ed., edited by E. Stackebrandt. New York: Springer-Verlag.
Carvalho, A. L., Dias, F. M., Prates, J. A., Nagy, T., Gilbert, H. J., Davies, G. J., Ferreira, L. M., Romão, M. J. & Fontes, C. M. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 13809–13814.
Carvalho, A. L., Pires, V. M., Gloster, T. M., Turkenburg, J. P., Prates, J. A., Ferreira, L. M., Romao, M. J., Davies, G. J., Fontes, C. M. & Gilbert, H. J. (2005). *J. Mol. Biol.* **349**, 909–915.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cosier, J. & Glazer, A. M. (1986). *J. Appl. Cryst.* **19**, 105–107.
Ding, S.-Y., Bayer, E. A., Steiner, D., Shoham, Y. & Lamed, R. (1999). *J. Bacteriol.* **181**, 6720–6729.
Lamed, R., Setter, E. & Bayer, E. A. (1983). *J. Bacteriol.* **156**, 828–836.
Lamed, R., Setter, E., Kenig, R. & Bayer, E. A. (1983). *Biotechnol. Bioeng. Symp.* **13**, 163–181.
Leibovitz, E. & Béguin, P. (1996). *J. Bacteriol.* **178**, 3077–3084.
Lytle, B., Myers, C., Kruus, K. & Wu, J. H. D. (1996). *J. Bacteriol.* **178**, 1200–1203.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Noach, I., Frolow, F., Jakoby, H., Rosenheck, S., Shimon, L. J. W., Lamed, R. & Bayer, E. A. (2005). *J. Mol. Biol.* **348**, 1–12.
Noach, I., Lamed, R., Xu, Q., Rosenheck, S., Shimon, L. J. W., Bayer, E. A. & Frolow, F. (2003). *Acta Cryst.* D**59**, 1670–1673.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Salamitou, S., Raynaud, O., Lemaire, M., Coughlan, M., Béguin, P. & Aubert, J.-P. (1994). *J. Bacteriol.* **176**, 2822–2827.
Shimon, L. J. W., Frolow, F., Yaron, S., Bayer, E. A., Lamed, R., Morag, E. & Shohan, Y. (1997). *Acta Cryst.* D**53**, 114–115.
Stout, G. H. & Jensen, L. H. (1968). *X-ray Structure Determination. A Practical Guide*, p. 402. London: Macmillan.
Spinelli, S., Fierobe, H. P., Belaich, A., Belaich, J. P., Henrissat, B. & Cambillau, C. (2000). *J. Mol. Biol.* **304**, 189–200.
Tavares, G. A., Béguin, P. & Alzari, P. M. (1997). *J. Mol. Biol.* **273**, 701–713.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Woessner, J. P. & Goodenough, U. W. (1992). *Plant Sci.* **83**, 65–76.
Xu, Q., Gao, W., Ding, S.-Y., Kenig, R., Shoham, Y., Bayer, E. A. & Lamed, R. (2003). *J. Bacteriol.* **185**, 4548–4557.