Kate Cameron, Victor D. Alves,
Pedro Bule, Luís M. A. Ferreira,
Carlos M. G. A. Fontes and
Shabir Najmudin*

CIISA – Faculdade de Medicina Veterinária,
Universidade Técnica de Lisboa, Avenida da
Universidade Técnica, 1300-477 Lisboa,
Portugal

Correspondence e-mail: shabir@fmv.ulisboa.pt

Received 19 February 2014

Accepted 3 April 2014

Purification, crystallization and preliminary X-ray characterization of the third ScaB cohesin in complex with an ScaA X-dockerin from *Acetivibrio cellulolyticus*

Interactions between cohesin and dockerin modules are critical for the formation of the cellulosome, which is responsible for the efficient degradation of plant cell-wall carbohydrates by anaerobes. Type I dockerin modules found in modular enzymatic components interact with type I cohesins in primary scaffoldins, enabling the assembly of the multi-enzyme complex. In contrast, type II dockerins located in primary scaffoldins bind to type II cohesins in adaptor scaffoldins or anchoring scaffoldins located at the bacterial envelope, contributing to the cell-surface attachment of the entire complex. *Acetivibrio cellulolyticus* possesses an extremely complex cellulosome arrangement which is organized by a primary enzyme-binding scaffoldin (ScaA), two anchoring scaffoldins (ScaC and ScaD) and an unusual adaptor scaffoldin (ScaB). An ScaA X-dockerin mutated to inactivate one of the two putative cohesin-binding interfaces complexed with the third ScaB cohesin from *A. cellulolyticus* has been purified and crystallized and data were collected to a resolution of 2.41 Å.

1. Introduction

Bacterial multi-enzyme complexes comprising a range of cellulases and hemicellulases, termed cellulosomes, degrade structural polysaccharides located in plant cell walls in a highly efficient and concerted way (Bayer *et al.*, 2004; Fontes & Gilbert, 2010). Cohesin and dockerin recognition pairs are integral to the architecture of the cellulosome. Type I dockerins present in modular cellulosomal catalytic components bind to cohesins located in primary scaffoldins, leading to cellulosome assembly. In contrast, type II dockerins located in primary scaffoldins bind to adaptor scaffoldins, forming polycellulosomes, or to anchoring scaffoldins, contributing to the cell-surface attachment of the entire complex (Adams *et al.*, 2006; Pinheiro *et al.*, 2012). The mesophilic Gram-positive bacterium *Acetivibrio cellulolyticus* possesses an extremely complex cellulosome which can be organized around 16 different scaffoldins that together comprise 41 cohesins (Dassa *et al.*, 2012). The main components in the organization of the *A. cellulolyticus* cellulosome are the primary enzyme-binding scaffoldin (ScaA), two anchoring scaffoldins (ScaC and ScaD) and an unusual adaptor scaffoldin (ScaB) (Xu *et al.*, 2004). The primary scaffoldin ScaA contains a C-terminal type II dockerin domain similar to *Clostridium thermocellum* CipA (Adams *et al.*, 2006). The ScaA dockerin binds to cohesins located in ScaB, which in turn binds to the anchoring scaffoldin ScaC via a third cohesin–dockerin specificity. A fourth scaffoldin, ScaD, is also located at the cell surface but, in contrast to ScaC, contains one cohesin that can bind enzymes directly and two cohesins that are specific for ScaA. ScaC and ScaD interact with the cell surface through S-layer homology (SLH) modules, which enable the entire complex to be tethered to the cell surface by at least two alternative suprastructures (Xu *et al.*, 2004).

Type II cohesin–dockerin interactions usually display a single mode of binding (Fontes & Gilbert, 2010). However, initial crystallization of the wild-type ScaA X-dockerin (XDoc) with the type II cohesin failed, suggesting that this cohesin–dockerin pair might display a dual mode of binding. Closer analysis of the ScaA dockerin sequence revealed the presence of two highly homologous tandem repeats within the ScaA dockerin. Also, the key recognition pair Leu/Glu (LE in helices 1 and 3 at positions 134–135 and 167–168 of the

repeats, respectively) and Phe/Asn (FN at positions 144–145 and 177–178, respectively) are completely conserved. This contradicts what was previously observed in single binding mode type II dockerins, in which a Gly replaces the Asn at one of these two positions (Carvalho *et al.*, 2007). Thus, to promote the crystallization of cohesin–dockerin complexes resulting from an exclusive single mode of binding, we engineered a mutation of Asn to Gly at position 178 which led to crystal formation. We present this result together with a description of the preliminary data collection.

2. Materials and methods

2.1. Protein expression and purification

The genes encoding the *A. cellulolyticus* dockerin and cohesin optimized for expression in *Escherichia coli* were synthesized *in vitro* (NZYTech Ltd, Portugal) and were co-expressed in the same cells as described previously (Cameron *et al.*, 2012). In this construct, the dockerin-encoding gene was introduced at the 5' end and the cohesin-encoding gene at the 3' end, with specifically tailored *NheI* and *NcoI* recognition sites at the 5' end and *XhoI* and *SalI* recognition sites at the 3' end when subcloning into pET028a (Novagen). Thus, the His₆ tag could be introduced either at the N-terminus of the dockerin using *NheI* and *SalI* or at the C-terminus of the cohesin using *NcoI* and *XhoI*, generating protein–protein complexes that could contain a His₆ tag either on the cohesin (AC2_coh) or on the X-dockerin (AC2_doc). Expression screens revealed that the dockerin-tagged complex gave higher expression levels, and it was thus selected for the subsequent work. The AC2_doc dockerin–cohesin complex was expressed in *E. coli* Tuner cells, which were grown at 37°C to an OD₆₀₀ of 0.5. Recombinant protein expression was induced by the addition of 0.2 mM isopropyl β -D-1-thiogalactopyranoside and incubation at 19°C for 16 h. Since the levels of expression of the dockerin were much higher than those of the cohesin, we assumed that most of the cohesin molecules were complexed with dockerin. Thus, the recombinant complex and unbound cohesin were purified by immobilized metal-ion affinity chromatography (IMAC) using Sepharose columns charged with nickel (HisTrap) following conventional protocols (Najmudin *et al.*, 2010).

Figure 1
A Coomassie Brilliant Blue-stained 14% SDS-PAGE gel evaluation of protein purity. Lane 1, Low Molecular Weight Protein Marker (NZYTech Ltd; labelled in kDa); lanes 2 and 3 represent the purified AC2_doc complex separated into the cohesin (top band) and dockerin (bottom band).

Fractions containing the purified cohesin–dockerin complex were buffer-exchanged on a PD-10 Sephadex G-25M gel-filtration column (GE Healthcare) into 20 mM Na HEPES buffer pH 7.5 containing 200 mM NaCl and 2 mM CaCl₂. The protein–protein complexes were loaded at a flow rate of 1 ml min^{−1} onto a HiLoad 16/60 Superdex 75 column (GE Healthcare) previously equilibrated with 20 mM Na HEPES buffer pH 7.5 containing 200 mM NaCl and 2 mM CaCl₂. This step enabled separation of the protein complex from a minor contamination of unbound cohesin. Fractions containing the purified complex were then concentrated with Amicon Ultra-15 centrifugal devices with a 10 kDa cutoff membrane (Millipore) and washed three times with 0.5 mM CaCl₂. The final protein concentration was measured using a NanoDrop 2000C UV–Vis Spectrophotometer (Thermo Scientific) with a molar extinction coefficient (ϵ) of 33 350 M^{−1} cm^{−1} and was adjusted to 10 and 20 mg ml^{−1} AC2_doc in 0.5 mM CaCl₂. A purity evaluation by SDS–PAGE was performed (Fig. 1).

(a)

(b)

Figure 2
(a) Crystals were obtained by sitting-drop vapour-diffusion methods using 20 mg ml^{−1} AC2_doc in 25% PEG 3350, 0.1 M bis-tris pH 5.5, 0.2 M MgCl₂. The black bar represents a length of 100 μ m. (b) Representative diffraction pattern for the monoclinic AC2_doc data (the outer circle corresponds to a resolution of 2.4 $\text{\AA}$).

Table 1

Data-collection statistics.

Values in parentheses are for the lowest/highest resolution shells.

Beamline	ID14-4, ESRF
Data set	AC2_M2
Space group	$P2_1$
Wavelength (Å)	0.97620
Unit-cell parameters	
a (Å)	64.11
b (Å)	38.65
c (Å)	72.32
β (°)	94.07
Resolution limits (Å)	63.95–2.41 (63.95–9.02/2.50–2.41)
Average mosaicity (°)	0.74
No. of observations	52305 (1009/4713)
No. of unique observations	13680 (287/1370)
Multiplicity	3.8 (3.5/3.4)
Completeness (%)	98.1 (96.6/96.3)
$\langle I/\sigma(I) \rangle$	3.0 (4.5/1.3)
$CC_{1/2}^\dagger$	0.951 (0.984/0.487)
$R_{\text{merge}}^\ddagger$ (%)	18.8 (10.7/58.4)
$R_{\text{p.i.m.}}^\S$ (%)	11.6 (5.8/38.0)

† $CC_{1/2}$ is the half-data set correlation coefficient (Diederichs & Karplus, 2013). ‡ $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the i th measurement of reflection hkl and $\langle I(hkl) \rangle$ is the mean value of $I_i(hkl)$ for all i measurements. § $R_{\text{p.i.m.}} = \sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $\langle I(hkl) \rangle$ is the average of symmetry-related observations of a unique reflection.

2.2. Crystallization

The crystallization conditions were screened by the sitting-drop vapour-phase diffusion method using the commercial kits Crystal Screen, Crystal Screen 2, PEG/Ion and PEG/Ion 2 from Hampton Research (California, USA) and JCSG-plus HT96 screens (Molecular Dimensions, UK) with an Oryx8 robotic nanodrop dispensing system (Douglas Instruments). Drops consisting of 0.7 μ l AC2_doc concentrated to 10 or 20 mg ml⁻¹ and 0.7 μ l reservoir solution were prepared at 292 K. Crystals were obtained in 25% PEG 3350, 0.1 M bis-tris pH 5.5, 0.2 M MgCl₂ (Fig. 2a). These crystals were cryocooled in liquid nitrogen after passing them through a cryoprotectant solution consisting of the crystallization buffer containing 30% glycerol.

2.3. Data collection and processing

Data were collected on beamline ID14-4 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France using a Q315r CCD detector (ADSC) from crystals cooled to 100 K using a Cryostream (Oxford Cryosystems). *EDNA* and *MOSFLM* (Leslie, 1992) were used for strategy calculation during data collection. All data were processed with *iMosflm* (Battye *et al.*, 2011) and *AIMLESS* (Evans, 2011) from the *CCP4* suite (Collaborative Computational Project, Number 4, 1994; Winn *et al.*, 2011). The AC2_doc crystals belonged to the monoclinic space group $P2_1$ and the best crystal diffracted to a resolution of 2.41 Å (Fig. 2b). Data-collection statistics are given in Table 1. The Matthews coefficient indicated the presence of a single cohesin–dockerin complex ($V_M = 2.41 \text{ Å}^3 \text{ Da}^{-1}$; solvent content of ~49%; Matthews, 1968). *BALBES* was used to carry out molecular replacement (Long *et al.*, 2008). 41 structures with a sequence identity of greater than 15% were found. The best solution

was found using the cohesin structure with PDB code 2b59 (Adams *et al.*, 2006; 57% sequence identity) and the X-dockerin structure with PDB code 3l8q (Noach *et al.*, 2010; 91% sequence identity). This gave a single Coh–Xdoc complex in space group $P2_1$, with an initial R factor and R_{free} of 34.6 and 44.2%, respectively, using *REFMAC5* (Murshudov *et al.*, 2011) and a Q -factor of 0.688. Structure refinement and analysis are ongoing.

This work was supported in part by Fundação para a Ciência e a Tecnologia (Lisbon, Portugal) through grants PTDC/BIA-PRO/103980/2008 and EXPL/BIA-MIC/1176/2012 and by the European Union Seventh Framework Programme (FP7 2007–2013) under the WallTraC project (grant agreement No. 263916) and BioStruct-X (grant agreement No. 283570) for funding. This paper reflects the views of the authors only. The European Community is not liable for any use that may be made of the information contained herein. We acknowledge ESRF, Grenoble, France (beamline ID14-4) for provision of synchrotron-radiation facilities. The authors would like to thank Drs Ana Luisa Carvalho and Márcia Correia for their help with data collection and Cecília Bonifácio for help with crystallization.

References

- Adams, J. J., Pal, G., Jia, Z. & Smith, S. P. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 305–310.
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst. D* **67**, 271–281.
- Bayer, E. A., Belaich, J.-P., Shoham, Y. & Lamed, R. (2004). *Annu. Rev. Microbiol.* **58**, 521–554.
- Cameron, K., Alves, V. D., Bule, P., Ferreira, L. M. A., Fontes, C. M. G. A. & Najmudin, S. (2012). *Acta Cryst. F* **68**, 1030–1033.
- Carvalho, A. L., Dias, F. M. V., Nagy, T., Prates, J. A. M., Proctor, M. R., Smith, N., Bayer, E. A., Davies, G. J., Ferreira, L. M. A., Romão, M. J., Fontes, C. M. G. A. & Gilbert, H. J. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 3089–3094.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst. D* **50**, 760–763.
- Dassa, B., Borovok, I., Lamed, R., Henrissat, B., Coutinho, P., Hemme, C. L., Huang, Y., Zhou, J. & Bayer, E. A. (2012). *BMC Genomics*, **13**, 210.
- Diederichs, K. & Karplus, P. A. (2013). *Acta Cryst. D* **69**, 1215–1222.
- Evans, P. R. (2011). *Acta Cryst. D* **67**, 282–292.
- Fontes, C. M. G. A. & Gilbert, H. J. (2010). *Annu. Rev. Biochem.* **79**, 655–681.
- Leslie, A. G. W. (1992). *Jnt CCP4/ESF-EACBM News. Protein Crystallogr.* **26**.
- Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). *Acta Cryst. D* **64**, 125–132.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst. D* **67**, 355–367.
- Najmudin, S., Pinheiro, B. A., Prates, J. A. M., Gilbert, H. J., Romão, M. J. & Fontes, C. M. G. A. (2010). *J. Struct. Biol.* **172**, 353–362.
- Noach, I., Levy-Assaraf, M., Lamed, R., Shimon, L. J. W., Frolov, F. & Bayer, E. A. (2010). *J. Mol. Biol.* **399**, 294–305.
- Pinheiro, B. A., Brás, J. L. A., Najmudin, S., Carvalho, A. L., Ferreira, L. M. A., Prates, J. A. M. & Fontes, C. M. G. A. (2012). *Biocatal. Biotransformation*, **30**, 309–315.
- Winn, M. D. *et al.* (2011). *Acta Cryst. D* **67**, 235–242.
- Xu, Q., Barak, Y., Kenig, R., Shoham, Y., Bayer, E. A. & Lamed, R. (2004). *J. Bacteriol.* **186**, 5782–5789.