STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Received 25 November 2014
Accepted 17 December 2014

Keywords: carbohydrate-binding domain; CBM22; *Paenibacillus*; bacterial xylanase; xylan-binding domain.

Crystallization and preliminary X-ray diffraction analysis of the N-terminal domain of *Paenibacillus barcinonensis* xylanase 10C containing the CBM22-1–CBM22-2 tandem

María Ángela Sainz-Polo,[a] Beatriz González,[a] F. I. Javier Pastor[b] and Julia Sanz-Aparicio[a]*

[a]Department of Crystallography and Structural Biology, Institute of Physical Chemistry Rocasolano, CSIC, Serrano 119, 28006 Madrid, Spain, and [b]Department of Microbiology, Faculty of Biology, University of Barcelona, Avenida Diagonal 643, 08028 Barcelona, Spain. *Correspondence e-mail: xjulia@iqfr.csic.es

A construct containing the CBM22-1–CBM22-2 tandem forming the N-terminal domain of *Paenibacillus barcinonensis* xylanase 10C (Xyn10C) has been purified and crystallized. A xylan-binding function and an affinity for mixed β-1,3/β-1,4 glucans have previously been demonstrated for some members of the CBM22 family. The sequence of the tandem is homologous to the N-terminal domains found in several thermophilic enzymes. Crystals of this tandem were grown by the streak-seeding method after a long optimization strategy. The structure has been determined by molecular replacement to a resolution of 2.43 Å and refinement is under way. This study represents the first structure containing two contiguous CBM22 modules, which will contribute to a better understanding of the role that this multiplicity plays in fine-tuning substrate affinity.

1. Introduction

Microbial hemicellulases, primarily xylanases (EC 3.2.1.8), have essential applications in industry owing to their ability to modify and transform lignocellulose and cell-wall materials of vegetal biomass. These enzymes are used in a wide variety of industrial processes such as the textile, baking and food industries to improve the properties of the products obtained. In the pulp-bleaching process, the application of xylanases reduces the chlorine requirement and thus notably contributes to the sustainability of the pulp and paper industry (Verma & Satyanarayana, 2012). Xylan, the major component of cell walls, is a highly complex polymer and therefore its enzymatic degradation requires a large variety of cooperatively acting enzymes to achieve an efficient process (Biely, 1985; de Vries & Visser, 2001). Among these enzymes, xylanases, which cleave the internal β-1,4 bonds within the main chain of xylan, play a key role.

Xylanases from several microorganisms, including fungi and bacteria, have been characterized and cloned. Sequencing of these genes has revealed that many of the encoded enzymes have a modular structure consisting of a catalytic domain attached to one or more ancillary domains with different functions (Collins *et al.*, 2005). The most common noncatalytic domains in modular xylanases are carbohydrate-binding modules (CBMs; Boraston *et al.*, 2004) that may promote accessibility to xylan, which occurs in spatial proximity to cellulose in natural substrates (Hervé *et al.*, 2010; Gilbert *et al.*, 2013). Analysis of the sequences of xylanases from thermophilic microorganisms identified other conserved sequences that were originally proposed to act as thermostabilizing

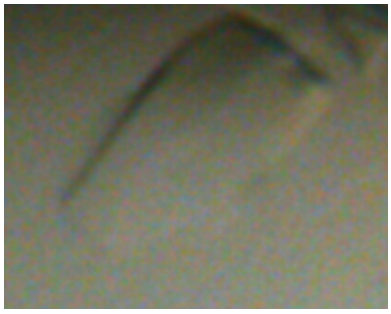


Table 1
Macromolecule-production information.

Source organism	*P. barcinonensis*
DNA source	Genomic DNA/pET-11a
Forward primer†	CGCGGATCCGCGAAGGCGGGGGATATTC
Reverse primer†	GGCCTCGAGTTAAGCAATCGCTTCGGTTGC
Cloning vector	pGEX-4T-2
Expression vector	pGEX-4T-2
Expression host	*E. coli* BL21 (DE3)
Complete amino-acid sequence of the construct produced	MASAAKAGDILLSHSFEEGTTQGWTARGGVKVDV-TAEQAYQGKQSLQTTGRTEAWNGPSLSLTDVV-HKNEVVEISGYVKLVAGSAPADLKFTVERRDG-NGDTQYDQVNAAEQVTDQKWVKLQGQYSYEQG-SSLLLYLESTDAKAAYLLDEFQIRLVKAAPEN-PGEPGEAGQALFKADFEDGNIGNWRARGTEKL-EVVSGIGHNSNRSLKTSSRSETYHGPLVEVLP-YLQKGSTVHISFWAMYDEGPATQVINGSLEKE-FNRDTANLEYAMFASTTLNKGQWKKIEADIIV-PAESTGISGLRMYAETPWKQSSEVTETDTIPF-YVDDVQITATEAIAIE

† BamHI and XhoI restriction sites are underlined.

domains (TSDs; Sunna *et al.*, 2000; Blanco *et al.*, 1999) but were later shown to act as xylan-binding domains which bind xylan and mixed β-1,3/β-1,4-glucan (Meissner *et al.*, 2000) and were classified into family 22 of carbohydrate-binding modules (CBM22).

Xylanase 10C (Xyn10C) from *Paenibacillus barcinonensis* (Sánchez *et al.*, 2005) is a 120 kDa multimodular enzyme that presents the CBM22/GH10/CBM9 modular architecture found in a subset of large GH10 xylanases (Collins *et al.*, 2005). Its N-terminal domain contains two repeats of family 22 carbohydrate-binding modules, both presenting the motifs previously identified as essential determinants for xylan binding (Xie *et al.*, 2001). In addition to its GH10 catalytic domain, two domains homologous to family 9 of cellulose-binding domains (CBM9) are located in the C-terminal region. Xyn10C showed an activity profile similar to those of enzymes from mesophilic microorganisms (Blanco *et al.*, 1999). Here, we report the purification, crystallization and preliminary X-ray crystallographic analysis of the N-terminal domain of *P. barcinonensis* Xyn10C containing the CBM22-1–CBM22-2 tandem (Xyn10C-Nterm). This represents the first structure of two contiguous CBM22 domains to be solved. Elucidation of this structure will contribute to understanding the role of multiplicity in the biological function of different microbial xylanases.

2. Materials and methods

2.1. Macromolecule production

The cDNA of Xyn10C-Nterm was amplified from a XYN10C-Nterm/pET-11a vector (UniProt accession No. O69230) using the forward and reverse oligonucleotides given in Table 1. The Xyn10C-Nterm cDNA generated included restriction sites for BamHI and XhoI enzymes that were used to insert the cDNA into the pGEX-4T-2 vector (GE Healthcare). The sequence of the Xyn10C-Nterm/pGEX-4T-2 plasmid was verified by DNA sequencing.

For protein expression, *Escherichia coli* BL21 (DE3) cells were transformed with the Xyn10C-Nterm/pGEX-4T-2 plasmid and grown in ampicillin-supplemented LB medium (50 µg ml^{-1}) at 310 K until the absorbance at 600 nm reached 0.8. Expression of the GST-fused protein was induced with 0.3 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) for 16 h at 289 K. Cells were collected by centrifugation and kept at 253 K until use.

The cells were suspended in 20 m*M* Tris pH 8.0, 100 m*M* NaCl (buffer *A*) and disrupted by sonication. The cell lysate was centrifuged and the resulting supernatant was filtrated. The clarified lysate was mixed with GST-trap resin (GenScript) equilibrated in buffer *A* and incubated for 1 h at 277 K. The resin was generously washed with buffer *A*. To cleave the GST-fused protein, the resin was incubated with thrombin (10 units per milligram of protein) overnight at 277 K. The cleaved Xyn10C-Nterm was collected, diluted (to reduce the salt to 50 m*M* NaCl) and subjected to ion-exchange chromatography (IEX) on a Mono Q column (ÄKTA; GE Healthcare). The protein was eluted using a gradient buffer of 50 m*M* to 1 *M* NaCl in 20 m*M* Tris pH 8.0. The protein concentration was estimated using a molar extinction coefficient (ε) of 57 870 M^{-1} cm^{-1} at 280 nm with a NanoDrop ND-1000 spectrophotometer and was verified by SDS–PAGE (Fig. 1). Purified Xyn10C-Nterm was concentrated to 47 mg ml^{-1} for protein crystallization assays. The purity of the protein was determined by SDS–PAGE (Laemmli, 1970).

2.2. Crystallization

In order to find a suitable protein concentration, the Pre-Crystallization Test (PCT; Hampton Research) was used, which gave a starting point of 30–50 mg ml^{-1}. Initial crystallization conditions of the purified protein samples were investigated by high-throughput techniques with a NanoDrop

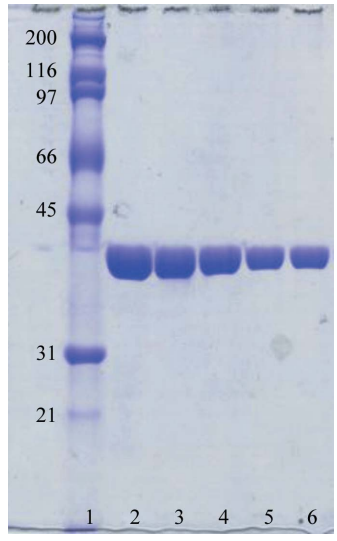


Figure 1
SDS–PAGE analysis of purified Xyn10C-Nterm on 12% polyacrylamide gel. Lane 1, molecular-weight protein markers (labelled in kDa); lanes 2–6, purified Xyn10C-Nterm (Mono Q column fractions). The molecular weight calculated from the sequence is 37.43 kDa.

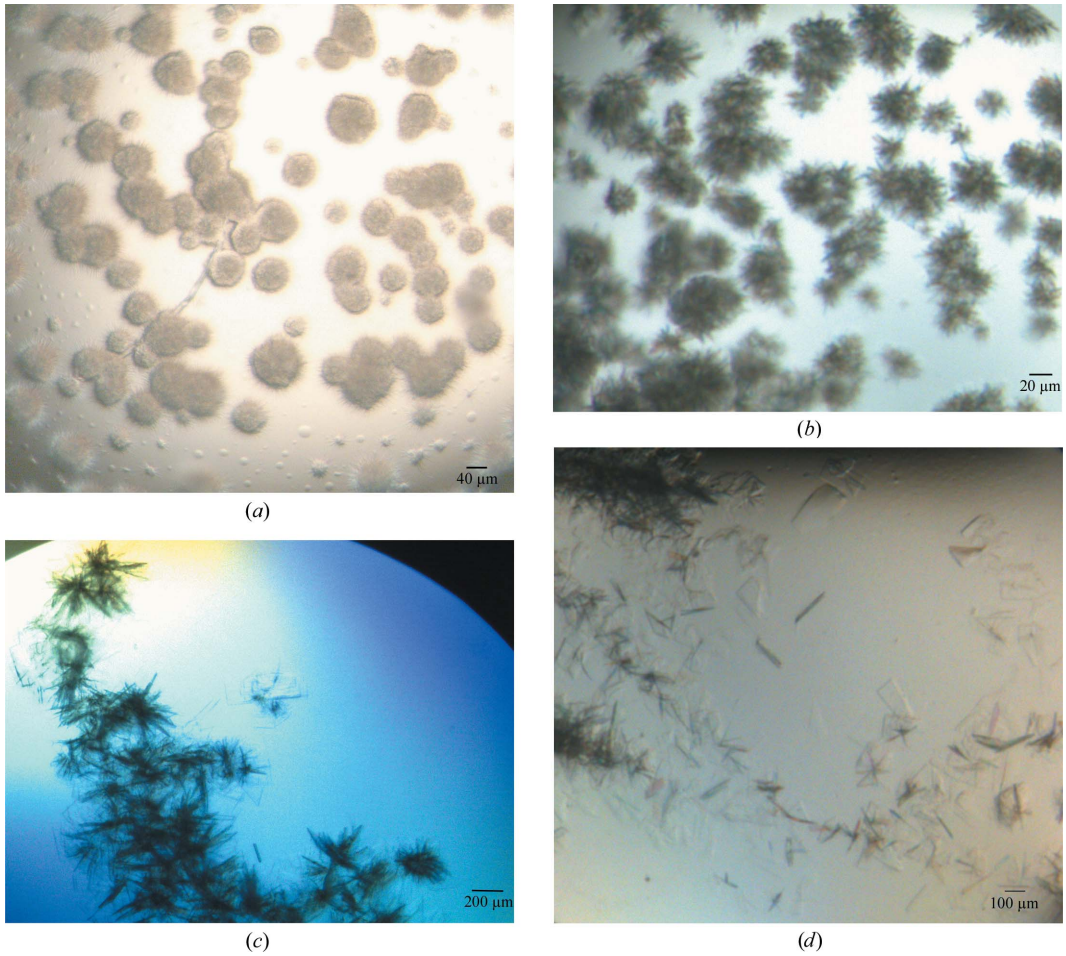
robot (Innovadyne Technologies Inc.) using the commercially available Index, SaltRx and Crystal Screen from Hampton Research and The PACT Suite and The JCSG Suite from Qiagen. The assays were carried out by the sitting-drop vapour-diffusion method at 291 K in Innovaplate SD-2 microplates (Innovadyne Technologies Inc.) by mixing 250 nl protein solution with 250 nl precipitant solution and equilibrating against 60 µl well solution. Optimization experiments were performed as further sitting-drop experiments by mixing 0.5–1 µl protein solution with 0.5–2 µl precipitant solution and equilibrating against 500 µl well solution in Cryschem plates (Hampton Research).

2.3. Data collection and processing

All crystals were soaked in precipitant solution containing an additional 20%(*v*/*v*) glycerol (Garman & Mitchell, 1996) before being flash-cooled to 100 K in liquid nitrogen. Crystals were tested using synchrotron radiation. Diffraction data sets were collected on the ID29 beamline at the European Synchrotron Radiation Facilily (ESRF), Grenoble, France. The data sets were processed using *iMosflm* (Battye *et al.*, 2011) and *AIMLESS* (Evans, 2006) as distributed in the *CCP*4 suite (Winn *et al.*, 2011).

3. Results and discussion

Xyn10C-Nterm was produced from *E. coli* cells in a soluble form and was purified in two steps (GST-tagged protein purification resin and ion-exchange chromatography) with a high degree of purity (Fig. 1, Table 1). A band in the SDS–PAGE analysis showing the expected molecular weight of the affinity-captured protein, which was later cleaved by thrombin, confirmed the identity of the sample. The purified protein was submitted to crystallization trials, with several conditions from commercial screens containing 25–30% PEG 3350 or PEG 8000 and pH 5.5–6.5 resulting in spherulites (Fig. 2*a*). Fine grid screening of these conditions did not improve the results and only the screening of additives led to the formation of small needles, which grew from the spherulites when using 25% PEG 3350 as precipitant and 2%(*v*/*v*) 1,6-hexanediol as an additive (Fig. 2*b*). Additional experiments were performed using different protein:precipitant: additive ratios and sampling the pH, and always led to the



(*a*) (*b*) (*c*) (*d*)

Figure 2
Xyn10C-Nterm crystals obtained from the following conditions. (*a*) 32 mg ml^{-1} protein solution and 25% PEG 3350, 0.2 *M* NaCl, 0.1 *M* bis-tris pH 5.5; 1:1 ratio. (*b*) 47 mg ml^{-1} protein solution and 28% PEG 3350, 0.2 *M* NaCl, 0.1 *M* bis-tris pH 5.5; 1:1 ratio; 2%(*v*/*v*) 1,6-hexanediol was used as an additive. (c) 47 mg ml^{-1} protein solution and 25% PEG 3350, 0.2 *M* NaCl, 0.1 *M* bis-tris pH 5.5, 2%(*v*/*v*) 1,6-hexanediol; 1:1 ratio; streak-seeding from spherulites. (*d*) 47 mg ml^{-1} protein solution and 23% PEG 3350, 0.2 *M* NaCl, 0.1 *M* bis-tris pH 5.5, 2%(*v*/*v*) 1,6-hexanediol; 1:4 ratio.

Table 2
Crystallization.

Method	Vapour diffusion (sitting drop) and streak-seeding
Plate type	Cryschem plates
Temperature (K)	291
Protein concentration (mg ml^{-1})	47
Buffer composition of protein solution	20 m*M* Tris pH 8.0, 80 m*M* NaCl
Composition of reservoir solution	23% PEG 3350, 0.2 *M* NaCl, 0.1 *M* bis-tris pH 5.5
Volume (µl) and ratio of drop	0.5:2:0.3 (protein:reservoir: 2% 1,6-hexanediol)
Volume of reservoir (µl)	500

same crystallization pattern. According to Bergfors (2003), this association trend is ideal for carrying out streak-seeding experiments and, indeed, very thin plates grew from streak-seeding (Fig. 2*c*). The best crystals were obtained by mixing 47 mg ml^{-1} protein solution with 23% PEG 3350, 0.2 *M* NaCl, 0.1 *M* bis-tris pH 5.5, 2%(*v*/*v*) 1,6-hexanediol at 291 K, leaving the mixture to equilibrate overnight and then streak-seeding from a drop containing spherulites with a whisker probe. Isolated thin plates appeared after 2–3 months (Fig. 2*d*). Crystallization information is summarized in Table 2.

The optimized crystals were analysed using a synchrotron-radiation source to obtain high-resolution data (Fig. 3). The best Xyn10C-Nterm crystal diffracted to a resolution of 2.43 Å. The crystals belonged to space group $P2_1$, with unit-cell parameters a = 84.170, b = 110.359, c = 118.522 Å, β = 90.63° (Table 3). As observed by SDS–PAGE and calculated from sequence analysis, the molecular weight of the monomer is 37 kDa (Fig. 1). Assuming a reasonable Matthews coefficient value within the range 3.68–1.84 Å^3 Da^{-1} (Matthews, 1968), corresponding to 66–33% solvent content,

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	Beamline ID29, ESRF
Wavelength (Å)	0.979235
Temperature (K)	120
Detector	PILATUS 6M-F
Crystal-to-detector distance (mm)	499.68
Rotation range per image (°)	0.1
Total rotation range (°)	360
Exposure time per image (s)	0.037
Space group	$P2_1$
a, b, c (Å)	84.170, 110.359, 118.522
β (°)	90.63
Mosaicity (°)	0.44
Resolution range (Å)	84.16–2.43 (2.48–2.43)
Total No. of reflections	510618 (29165)
No. of unique reflections	81116 (4468)
Completeness (%)	99.4 (99.3)
Multiplicity	6.3 (6.5)
$\langle I/\sigma(I)\rangle$	8.6 (2.2)
$R_{\rm p.i.m.}$†	0.060 (0.397)
Overall B factor from Wilson plot (Å^2)	37.24

† $R_{\rm p.i.m.} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the average of symmetry-related observations of a unique reflection.

the presence of 4–8 molecules in the asymmetric unit should be expected. We investigated the local symmetry relating the subunits in the asymmetric unit using *POLARRFN* (Kabsch, 1976) from the *CCP*4 package. Several self-rotation functions were computed in the resolution range 15–3.5 Å with Patterson vectors of 25–35 Å radius of integration. The stereographic projection of the κ = 180° section (Fig. 4) revealed the presence of noncrystallographic twofold symmetry almost parallel to the a and c axes.

Structure determination is in progress using *Phaser* (McCoy *et al.*, 2007) for molecular replacement. The templates used were CBM22-1 and CBM22-2 from *Clostridium thermocellum*

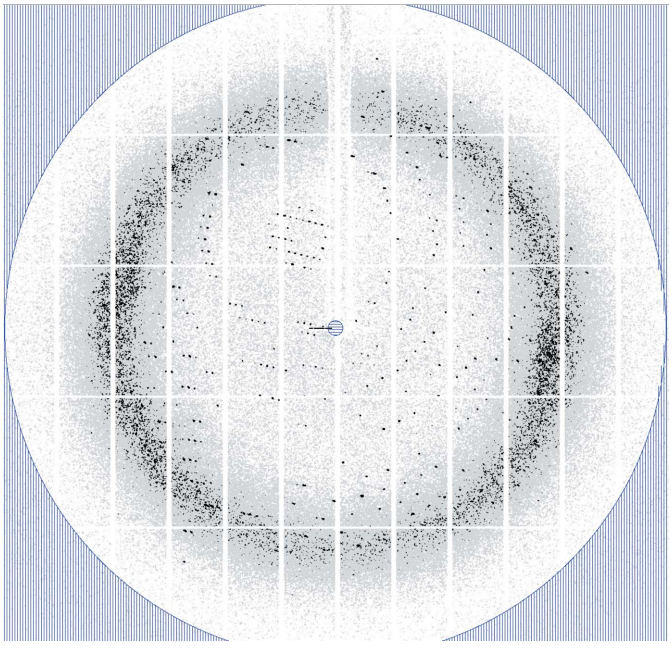


Figure 3
X-ray diffraction pattern of Xyn10C-Nterm obtained using a synchrotron source. The maximum observed resolution is 2.43 Å.

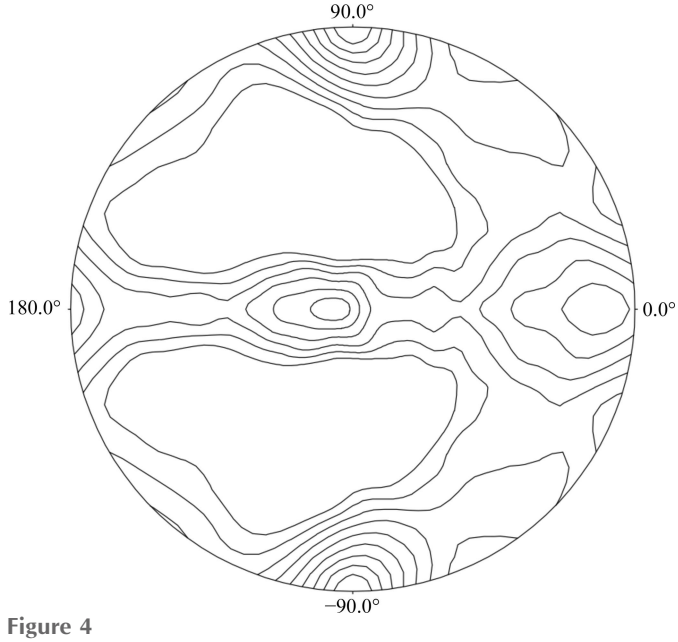


Figure 4
Plot of the self-rotation function of Xyn10C-Nterm crystals in the κ = 180° section.

xylanase Y (PDB entries 2w5f and 1dyo; Najmudin *et al.*, 2010; Charnock *et al.*, 2000), which show about 20% sequence identity with respect to each Xyn10C-Nterm subdomain. A solution with six molecules within the asymmetric unit was obtained and preliminary structural refinement with *REFMAC* (Murshudov *et al.*, 2011) decreased the *R* factor to 0.37 (R_{free} = 0.42). Model building and further refinement are ongoing. All data-collection statistics are shown in Table 3.

Acknowledgements

This work was supported by grants BIO2010-20508-C04-03 and BIO2013-48779-C4-2-R from Dirección General de Investigación, MICINN. This is a product of the Project 'Factoría Española de Cristalización' Ingenio/Consolider 2010. MASP is supported by a JAE-PreDoc fellowship from CSIC. We also thank the ESRF for beamtime and the ID29 staff for providing assistance with data collection.

References

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Bergfors, T. (2003). *J. Struct. Biol.* **142**, 66–76.
Biely, P. (1985). *Trends Biotechnol.* **3**, 286–290.
Blanco, A., Diaz, P., Zueco, J., Parascandola, P. & Pastor, F. I. J. (1999). *Microbiology*, **145**, 2163–2170.
Boraston, A. B., Bolam, D. N., Gilbert, H. J. & Davies, G. J. (2004). *Biochem. J.* **382**, 769–781.
Charnock, S. J., Bolam, D. N., Turkenburg, J. P., Gilbert, H. J., Ferreira, L. M. A., Davies, G. J. & Fontes, C. M. G. A. (2000). *Biochemistry*, **39**, 5013–5021.
Collins, T., Gerday, C. & Feller, G. (2005). *FEMS Microbiol. Rev.* **29**, 3–23.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Garman, E. F. & Mitchell, E. P. (1996). *J. Appl. Cryst.* **29**, 584–587.
Gilbert, H. J., Knox, J. P. & Boraston, A. B. (2013). *Curr. Opin. Struct. Biol.* **23**, 669–677.
Hervé, C., Rogowski, A., Blake, A. W., Marcus, S. E., Gilbert, H. J. & Knox, J. P. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 15293–15298.
Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Meissner, K., Wassenberg, D. & Liebl, W. (2000). *Mol. Microbiol.* **36**, 898–912.
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.
Sánchez, M. M., Fritze, D., Blanco, A., Spröer, C., Tindall, B. J., Schumann, P., Kroppenstedt, R. M., Diaz, P. & Pastor, F. I. J. (2005). *Int. J. Syst. Evol. Microbiol.* **55**, 935–939.
Sunna, A., Gibbs, M. D. & Bergquist, P. L. (2000). *Biochem. J.* **346**, 583–586.
Verma, D. & Satyanarayana, T. (2012). *Bioresour. Technol.* **117**, 360–367.
Vries, R. P. de & Visser, J. (2001). *Microbiol. Mol. Biol. Rev.* **65**, 497–522.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Xie, H., Bolam, D. N., Nagy, T., Szabó, L., Cooper, A., Simpson, P. J., Lakey, J. H., Williamson, M. P. & Gilbert, H. J. (2001). *Biochemistry*, **40**, 5700–5707.