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

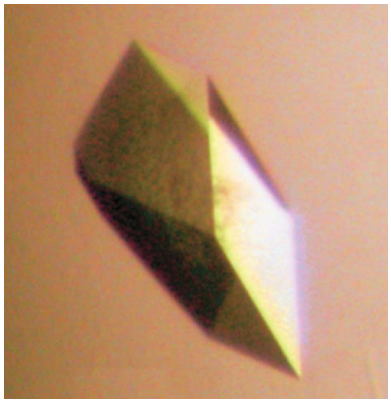
Purification, crystallization and preliminary characterization of a putative LmbE-like deacetylase from *Bacillus cereus*

Vasiliki E. Fadouloglou,[a] Dina Kotsifaki,[b] Anastasia D. Gazi,[a] Georgios Fellas,[a] Chrysi Meramveliotaki,[a] Alexandra Deli,[a] Emmanuel Psylinakis,[a] Vassilis Bouriotis[a,b] and Michael Kokkinidis[a,b]*

[a]University of Crete, Department of Biology, PO Box 2208, GR-71409 Heraklion, Crete, Greece, and [b]Institute of Molecular Biology and Biotechnology, PO Box 1527, Vasilika Vouton, GR-71110 Heraklion, Crete, Greece

Correspondence e-mail: kokkinid@imbb.forth.gr

Received 20 December 2005
Accepted 7 February 2006



The *Bacillus cereus* BC1534 protein, a putative deacetylase from the LmbE family, has been purified to homogeneity and crystallized using the hanging-drop vapour-diffusion method. Crystals of the 26 kDa protein grown from MPD and acetate buffer belong to space group *R*32, with unit-cell parameters $a = b = 76.7$, $c = 410.5$ Å (in the hexagonal setting). A complete native data set was collected to a resolution of 2.5 Å from a single cryoprotected crystal using synchrotron radiation. As BC1534 shows significant sequence homology with an LmbE-like protein of known structure from *Thermus thermophilus*, molecular replacement will be used for crystal structure determination.

1. Introduction

Bacillus cereus is an opportunistic pathogenic bacterium closely related to *B. anthracis*, the causative agent of anthrax in mammals. Recent sequencing and comparative analyses of the *B. anthracis* and *B. cereus* genomes (Ivanova *et al.*, 2003; Read *et al.*, 2003) have revealed that a significant portion (ranging from 66 to 92% depending on the strain) of the *B. cereus* chromosomal genes are common to *B. anthracis*. These homologues include almost all the putative chromosomal virulence and surface proteins of *B. anthracis*, although *B. cereus* is not associated with anthrax. Moreover, homologues of the pXO1 plasmid (one of the two *B. anthracis* plasmids which harbours key virulence genes) have been found in half of the *B. cereus* strains examined. These extensive homologies are strongly suggestive of a common ancestor (Ivanova *et al.*, 2003). Given the striking similarity between the *B. anthracis* and *B. cereus* genomes and the high infectiousness of the former, *B. cereus* offers a good alternative model organism for studying the corresponding proteins of *B. anthracis*.

The BC1534 protein of *B. cereus* is a 26 kDa enzyme which exhibits deacetylase activity with the monosaccharide *N*-acetylglucosamine (GlcNAc) as well as with the N-acetylchitooligosaccharides di-acetylchitobiose ($GlcNAc_2$) and triacetylchitobiose ($GlcNAc_3$); enzymatic activity decreases with increasing chain length, so that GlcNAc is a better substrate than $GlcNAc_2$, which is better than $GlcNAc_3$ (unpublished results). However, the specific substrate and function of the BC1534 deacetylase have not yet been conclusively identified. BC1534 shares 26% amino-acid sequence identity with the product of the *bc3461* gene from *B. cereus* and has three homologues in *B. anthracis* (the products of the *bant_01002171*, *bant_01004539* and *bant_01004184* genes) with identities of 96, 28 and 25%, respectively.

On the basis of its amino-acid sequence, BC1534 can be classified as an LmbE-like protein (Pfam02585). This family comprises mainly deacetylases with acetylchitobiose (GlcNAc) deacetylase activity. The three best characterized members of this family are: (i) the Tk-Dac protein from the archaeon *Thermococcus kodakaraensis* KOD1, a diacetylchitobiose deacetylase which is believed to participate in a novel (probably common in archaea) chitinolytic pathway. This pathway combines diacetylchitobiose ($GlcNAc_2$) degradation with deacetylation (Tanaka *et al.*, 2003, 2004). (ii) The MshB protein from *Mycobacterium tuberculosis*, a mycothiol deacetylase involved in the biosynthetic pathway of mycothiol (Newton *et al.*, 2000). The recently determined crystal structure of MshB has established the location of

a Zn^{2+}-binding site which exhibits considerable similarity to the active sites of metalloproteases (Maynes *et al.*, 2003). (iii) The LmbE-like TT1542 protein from *Thermus thermophilus* HB8, a conserved protein of unknown function whose crystal structure has recently been determined (Handa *et al.*, 2003). On the basis of the structure, a putative catalytic site has been identified and suggestions concerning its function have been made.

BC1534 shares 25% amino-acid sequence identity with the MshB protein; the Zn^{2+}-binding site residues of MshB are all conserved in BC1534, with residues His13, His147 and Asp16 of MshB corresponding to His12, Asp15 and His113 of BC1534. In addition, residue Asp15 of MshB, which has been proposed to play the role of a general base in catalysis, corresponds to Asp16 of BC1534. Interestingly, a general base which activates a water molecule to cleave the peptide bond is a typical feature of metalloproteases and in particular of hydrolases utilizing zinc ions. Furthermore, BC1534 shares an overall sequence identity of 39% with TT1542; the metal-binding site residues and the residue potentially acting as a general base are also all conserved. The crystal structure of TT1542 thus provides an attractive search model for the structure determination of the BC1534 using the molecular-replacement method (Rossmann & Blow, 1962; Rossmann, 1990). Here, we report the purification, crystallization and preliminary crystallographic characterization of the protein BC1534.

2. Protein cloning, expression and purification

The *bc1534* gene was isolated from *B. cereus* ATCC14579 genomic DNA using the polymerase chain reaction (PCR) method. The forward and reverse primers were designed to include *Nde*I and *Xho*I restriction sites, respectively. The product was digested with the restriction endonucleases *Nde*I and *Xho*I and then ligated to the vector pET26b(+). The stop codon TTA had been removed, so a 6×His tag was added. The plasmid was expressed in *Escherichia coli* BL21(DE3) cells as a C-terminally His-tagged protein and purified by affinity chromatography. The resulting amino-acid sequence is MSGLHILAFGAHADDVEIGMAGTIAKYTKQGYEVGICDLTEADLSSNGTIELRKEEAKVAARIMGVKTRLNLAMPDRGLYMKEEYIREIVKVIRTYKPKLVFAPYYEDRHPDHANCAKLVEEAIFSAGIRKYMPELSPHRVESFYNYMINGFHKPNFCIDISEYLSIKVEALEAYESQFSTGSDGVKTPLTEGYVETVIAREKMFGKEVGVLYAEGFMSKKPVLLHADLLGGCK**LGHHHHHH**; the His tag and the two extra C-terminal residues added by the plasmid construction are highlighted.

For a typical preparation, 1 l cell culture was grown at 303 K in LB (Luria–Broth) medium containing 25 µg ml^{-1} kanamycin to an optical density at 600 nm of approximately 0.65. The cells were induced with 1 m*M* IPTG (isopropyl β-D-thiogalactopyranoside) and further incubated at 303 K for 16 h. Harvested cells were resuspended in 50 ml buffer *A* (5 m*M* imidazole, 300 m*M* NaCl, 50 m*M* Na_2HPO_4/NaH_2PO_4 pH 8.0). The cells were disrupted by pulsed sonication for 4 min. The cell debris was removed by centrifugation (13 000*g* for 60 min) and the supernatant was loaded onto a 7 ml Ni–NTA column pre-equilibrated with buffer *A*. The column was washed in steps of 15 column volumes of buffer *B* containing 300 m*M* NaCl, 50 m*M* Na_2HPO_4/NaH_2PO_4 pH 8.0; in each step the concentration of imidazole was successively increased; *i.e.* 10, 20 and 50 m*M* imidazole. The protein was eluted with buffer *B* containing 200 m*M* imidazole. The protein solution was dialysed against 25 m*M* Tris–HCl pH 8.0, 200 m*M* NaCl, concentrated by Amicon Ultra-15 to about 3 ml and applied onto a calibrated S-200 Sephacryl gel-filtration column pre-equilibrated with 25 m*M* Tris–HCl pH 8.0, 200 m*M* NaCl. The column was calibrated using proteins from the Low and High Molecular-Weight Gel-Filtration Calibration Kits from Pharmacia (ovalbumin, 43 kDa; bovine serum albumin, 67 kDa; aldolase, 158 kDa; catalase, 232 kDa). As shown in Fig. 1(*a*), the protein elutes as a single peak at a volume (168 ml) corresponding to an estimated molecular weight of 160 kDa. The protein-containing fractions were pooled, dialysed against 25 m*M* Tris–HCl pH 7.5, 50 m*M* NaCl, concentrated to about 5 mg ml^{-1} and used for crystallization experiments.

Table 1
X-ray data-collection statistics for BC1534 protein crystals.

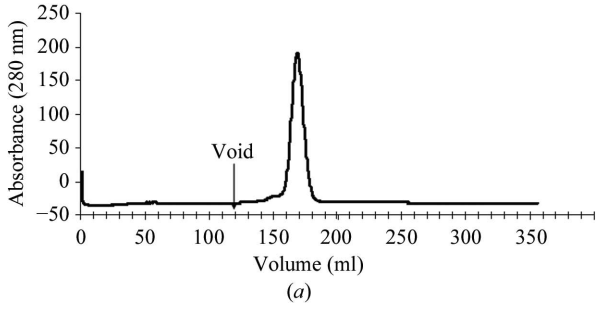
Values in parentheses refer to the outer resolution shell (2.65–2.52 Å).

Synchrotron	Beamline X11, EMBL Outstation, Hamburg
Wavelength (Å)	0.81
Space group	*R*32
Unit-cell parameters, hexagonal setting (Å)	$a = b = 76.7$, $c = 410.5$
No. of monomers in ASU	2
V_M (Å^3 Da^{-1})	2.2
Resolution (Å)	2.52
No. of unique reflections	16341
Multiplicity	10.2
Completeness (%)	99.2 (99.2)
$\langle I/\sigma(I)\rangle$	7.6 (2.1)
R_{sym}† (%)	8.9 (36.2)

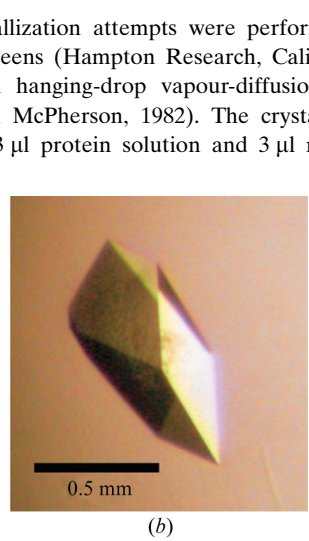
† $R_{sym} = \sum |I(h) - \langle I(h)\rangle| / \sum I(h)$, where $I(h)$ is the observed intensity and $\langle I(h)\rangle$ is the mean intensity of reflection h over all measurements of $I(h)$.

3. Crystallization

Initial crystallization attempts were performed with commercially available screens (Hampton Research, California, USA) using the conventional hanging-drop vapour-diffusion method (Ducruix & Giegé, 1992; McPherson, 1982). The crystallization drops were a mixture of 3 µl protein solution and 3 µl reservoir solution. Very



(*a*)



(*b*)

Figure 1
(*a*) Gel-filtration chromatography curve. (*b*) Crystal of the *B. cereus* BC1534 protein.

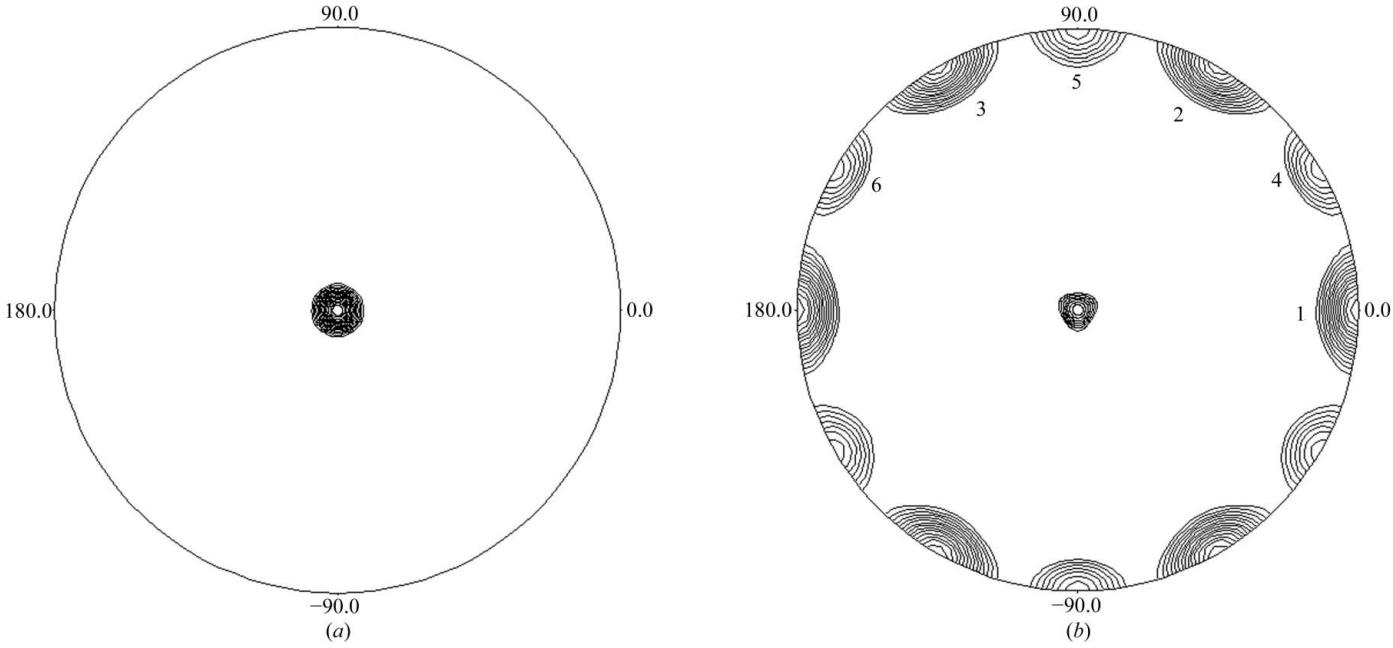


Figure 2
Self-rotation function calculated in the resolution range 12–4 Å using a radius of integration of 23 Å. The crystallographic *a* axis is in the horizontal direction (corresponding to $\varphi = 0$, $\omega = 90^\circ$); the *c* axis is perpendicular to the plane of the paper and directed towards the viewer. (*a*) $\kappa = 120^\circ$ section. (*b*) $\kappa = 180^\circ$ section.

small crystals were obtained with the MPD (2-methyl-2,4-pentanediol) Grid Screen over a wide range of MPD concentrations [8–28%(*v*/*v*)] and acetate buffer (CH_3COOH/CH_3COONa) at pH 5.0. The crystallization conditions were subsequently optimized and the best diamond-shaped crystals (Fig. 1*b*) were obtained with a reservoir solution composed of 14–26%(*v*/*v*) MPD and 100 m*M* acetate buffer pH 5.0. Crystals appear within 2 d at 291 K. They grow to maximum dimensions of approximately 0.5 × 0.5 × 0.2 mm in approximately one month.

4. X-ray data collection

X-ray diffraction data were collected on a MAR 165 mm CCD detector at a wavelength of 0.81 Å using beamline X11 of the EMBL Hamburg Outstation (DESY). Prior to data collection, the crystals were transferred to a cryoprotectant solution containing 35% MPD and 100 m*M* acetate buffer pH 5.0 and then flash-cooled in a cold stream of nitrogen at about 100 K. Using an oscillation range of 0.5°, 360 images were collected to a maximum resolution of 2.52 Å. Data were indexed, integrated and scaled using *MOSFLM* and *SCALA* (Leslie, 1992; Evans, 1997; Collaborative Computational Project, Number 4, 1994).

5. Results and discussion

The protein crystallizes in space group *R*32, with unit-cell parameters $a = b = 76.7$, $c = 410.5$ Å in the hexagonal setting. An acceptable value for the Matthews coefficient V_M (Matthews, 1968) is obtained assuming the presence of two monomers of 26 kDa in the asymmetric unit, corresponding to a solvent content of ∼44%. A summary of crystallographic parameters and data-collection statistics is given in Table 1.

To test for the presence of non-crystallographic symmetry, a self-rotation function was calculated using the program *POLARRFN* from the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994). Fig. 2 shows the $\kappa = 120^\circ$ and $\kappa = 180^\circ$ sections from a self-rotation function calculated with data between 12 and 4 Å and an integration radius of 23 Å. For space group *R*32, four crystallographic rotation axes are anticipated: the threefold axis corresponds to the only peak at $\kappa = 120^\circ$, while three twofold axes perpendicular to the threefold axis and indicated with the numbers 1, 2 and 3 are found at $\kappa = 180^\circ$. Fig. 2(*b*) reveals the presence of at least one additional non-crystallographic twofold axis (corresponding to peak 4 and its symmetry-equivalent peaks 5 and 6). These axes lie on the plane of the crystallographic twofold axes and are spaced every 60° in φ. The peak at $\omega = 0^\circ$ possibly arises from interaction between the crystallographic and non-crystallographic symmetry elements.

The molecular weight of the protein in solution as estimated by gel-filtration experiments (approximately 160 kDa) corresponds to a globular hexamer. This is consistent with the homohexameric form of the TT1542 structure, which consists of a trimer of dimers (Handa *et al.*, 2003).

We are presently pursuing structure determination of the BC1534 protein.

This study was co-funded by European Social Fund and National Resources under the PYTHAGORAS programme. Support in the framework of the European Community Research Infrastructure Action under the FP6 'Structuring the European Research Area Programme' (contract No. RII3/CT/2004/5060008) and provision of beam time at EMBL/DESY, Hamburg are gratefully acknowledged. We thank Dr N. M. Glykos for his useful comments on this manuscript.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Ducruix, A. & Giegé, R. (1992). *Crystallization of Nucleic Acids and Proteins: A Practical Approach*, pp. 73–126. New York: IRL Press.
Evans, P. R. (1997). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **33**, 22–24.

Handa, N., Terada, T., Kamewari, Y., Hamana, H., Tame, J. R. H., Park, S.-Y., Kinoshita, K., Ota, M., Nakamura, H., Kuramitsu, S., Shirouzu, M. & Yokoyama, S. (2003). *Protein Sci.* **12**, 1621–1632.
Ivanova, N. *et al.* (2003). *Nature (London)*, **423**, 87–91.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
McPherson, A. (1982). *Preparation and Analysis of Protein Crystals*. New York: John Wiley.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Maynes, J. T., Garen, C., Cherney, M. M., Newton, G., Arad, D., Av-Gay, Y., Fahey, R. C. & James, M. N. G. (2003). *J. Biol. Chem.* **278**, 47166–47170.
Newton, G. L., Av-Gay, Y. & Fahey, R. C. (2000). *J. Bacteriol.* **182**, 6958–6963.
Read, T. D. *et al.* (2003). *Nature (London)*, **423**, 81–86.
Rossmann, M. G. (1990). *Acta Cryst.* A**46**, 73–82.
Rossmann, M. G. & Blow, D. M. (1962). *Acta Cryst.* **15**, 24–31.
Tanaka, T., Fukui, T., Atomi, H. & Imanaka, T. (2003). *J. Bacteriol.* **185**, 5175–5180.
Tanaka, T., Fukui, T., Fujiwara, S., Atomi, H. & Imanaka, T. (2004). *J. Biol. Chem.* **279**, 30021–30027.