electronic reprint

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091
Editors: **H. M. Einspahr and J. M. Guss**

Cloning, purification and crystallization of *Bacillus anthracis* class C acid phosphatase

Richard L. Felts, Thomas J. Reilly, Michael J. Calcutt and John J. Tanner

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

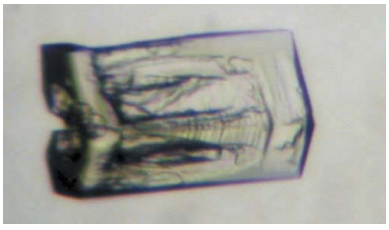
Cloning, purification and crystallization of *Bacillus anthracis* class C acid phosphatase

Richard L. Felts,[a] Thomas J. Reilly,[b,c] Michael J. Calcutt[b] and John J. Tanner[a,d]*

[a]Department of Chemistry, University of Missouri-Columbia, Columbia, MO 65211, USA, [b]Department of Veterinary Pathobiology, University of Missouri-Columbia, Columbia, MO 65211, USA, [c]Veterinary Medical Diagnostic Laboratory, University of Missouri-Columbia, Columbia, MO 65211, USA, and [d]Department of Biochemistry, University of Missouri-Columbia, Columbia, MO 65211, USA

Correspondence e-mail: tannerjj@missouri.edu

Received 2 June 2006
Accepted 22 June 2006



Cloning, expression, purification and crystallization studies of a recombinant class C acid phosphatase from the Category A pathogen *Bacillus anthracis* are reported. Large diffraction-quality crystals were grown in the presence of HEPES and Jeffamine ED-2001 at pH 7.0. The crystals belong to space group $P2_12_12_1$, with unit-cell parameters $a = 53.4$, $b = 90.1$, $c = 104.2$ Å. The asymmetric unit is predicted to contain two protein molecules with a solvent content of 38%. Two native data sets were collected from the same crystal before and after flash-annealing. The first data set had a mosaicity of 1.6° and a high-resolution limit of 1.8 Å. After flash-annealing, the apparent mosaicity decreased to 0.9° and the high-resolution limit of usable data increased to 1.6 Å. This crystal form is currently being used to determine the structure of *B. anthracis* class C acid phosphatase with experimental phasing techniques.

1. Introduction

Phosphatases are ubiquitous enzymes that catalyze phosphoryl transfer from phosphomonoesters to water (Knowles, 1980). The subgroup of phosphatases known as nonspecific acid phosphatases (NSAPs) refers to bacterial polyspecific phosphatases that are secreted to the periplasmic space or cell surface and that exhibit optimum activity at acidic pH. Rossolini and coworkers first recognized NSAPs as a distinct group of phosphatases and further categorized them into three classes denoted A, B and C based on subcellular localization and conserved amino-acid sequence motifs (Rossolini *et al.*, 1998).

Class C NSAPs are ~30 kDa acid phosphatases that are anchored to the outer membrane surface by an N-terminal lipid modification (Reilly & Calcutt, 2004). The signature sequence for class C NSAPs is a bipartite motif consisting of (I/V)-(V/A/L)-D-(I/L)-D-E-T-(V/M)-L-*X*-(N/T)-*XX*-Y near the N-terminus and (I/V)-(L/M)-*XX*-G-D-(N/T)-L-*X*-D-F near the C-terminus, separated by a linker region of 180–220 residues (Thaller *et al.*, 1998). Class C NSAPs characterized to date include enzymes from *Haemophilus influenzae* [*e* (P4); Reilly *et al.*, 1999], *Streptococcus equisimilis* (LlpC; Malke, 1998), *Staphylococcus aureus* (du Plessis *et al.*, 2002), *Helicobacter pylori* (Reilly & Calcutt, 2004) and *Chryseobacterium meningosepticum* (OlpA; Passariello *et al.*, 2003). The crystallization of *e* (P4) has been described (Ou *et al.*, 2006), but no structure of a class C NSAP has yet been reported. Enzymatically inactive mutants of *e* (P4) were recently shown to be potential vaccine candidates for *Ha. influenzae* (Green *et al.*, 2005). Thus, class C enzymes may be useful vaccine reagents for other bacteria.

Class C NSAPs are distantly related to class B NSAPs, which also contain two pairs of absolutely conserved aspartic acid residues (a DDDD motif) separated by a linker region (Thaller *et al.*, 1998). The structure of AphA from *Escherichia coli* shows that class B enzymes have the haloacid dehalogenase (HAD) fold and that the four conserved aspartic acid residues bind a divalent metal ion in the active site (Calderone *et al.*, 2004). The first Asp of the DDDD motif is proposed to be the nucleophile that attacks the P atom of the substrate (Calderone *et al.*, 2006). Although class B and C enzymes have similar bipartite DDDD motifs, they share only 14–22% global sequence identity (Thaller *et al.*, 1998). Thus, it remains to be seen

whether class C NSAPs also have the HAD fold and active-site structure of class B NSAPs.

During our analysis of putative surface proteins in Category A biothreat pathogens, a homolog of the class C NSAP of *H. pylori* was found in the annotated genome sequence of *B. anthracis* strain Sterne. *Bacillus anthracis* is a Gram-positive aerobic spore-forming bacterium that is the causative agent of anthrax. A gene encoding an identical protein is present in *B. anthracis* strains Ames and A2012 (Read *et al.*, 2002, 2003). Closely related proteins (>97% amino-acid sequence identity) are encoded in the genomes of multiple *B. cereus* isolates as well as in strains of *B. thuringiensis*, but no homolog was identified in the complete genome sequences of *B. subtilis*, *B. halodurans* or *B. clausii*. The identified *B. anthracis* protein (RefSeq accession No. NP_846955) shares 32–44% global amino-acid sequence identity with known class C NSAPs from *Ha. influenzae*, *Strep. equisimilis*, *Staph. aureus*, *He. pylori* and *C. meningosepticum*. The defining class C signature motif for this protein is VLDLDET-VLDNSPH in the N-terminal half of the protein and VLFFG-DNLSDF in the C-terminal half.

As part of our ongoing research on the roles of acid phosphatases in bacterial pathogenesis (Felts *et al.*, 2005, 2006; Reilly *et al.*, 2006; Ou *et al.*, 2006), we have initiated structural investigations of *B. anthracis* class C acid phosphatase. We report here the cloning, expression, purification and crystallization of this enzyme as a first step toward high-resolution crystal structure determination.

2. Methods and results

2.1. Cloning, expression and purification

The gene encoding the class C NSAP from *B. anthracis* delta Sterne (strain Sterne lacking plasmid pXO1) was amplified by PCR under standard cycling conditions using primers containing *Nco*I and *Xho*I sites. Following initial cloning into pZero Blunt (Invitrogen), the gene was inserted into the expression vector pET20b (Novagen) using *E. coli* DH10B as cloning host. The resulting recombinant protein consists of 251 residues with an approximate molecular weight of 28 kDa, a hexahistidine tag at the C-terminus and a *pelB* leader sequence for targeting to the *E. coli* periplasm. The *pelB* leader sequence replaces the signal peptide corresponding to residues 1–24 of the native protein. Following confirmation of the insert sequence, the recombinant plasmid, designated pBanxExp, was introduced into the expression host *E. coli* BL21 (DE3) by standard transformation.

A single colony of *E. coli* BL21 (DE3) containing the plasmid pBanxExp was used to inoculate 5 ml LB containing ampicillin (50 µg ml^{-1}) and was incubated overnight at 310 K with constant aeration. A 1:1000 dilution of this overnight culture was then used to inoculate 25 ml fresh medium. The sample was incubated as before until the density of the culture reached an optical density A_{600} of 0.6. The sample was then chilled on ice for 10 min and kept at 277 K for overnight storage. To remove secreted β-lactamases, the *E. coli* starter culture was centrifuged at 3660*g* for 10 min at 277 K and suspended in fresh LB prior to equal distribution into four 1.8 l flasks containing 500 ml LB supplemented with ampicillin (50 µg ml^{-1}) and 0.2%(*w*/*v*) filter-sterilized glucose. The culture was grown with constant aeration at 310 K until A_{600} reached 0.4, at which point IPTG was added to a final concentration of 0.4 m*M*. The cells were harvested 3 h after induction by centrifugation at 5000*g* for 10 min. The resulting pellet was suspended in 25 ml 50 m*M* sodium acetate pH 6.0. All subsequent purification steps were carried out at 277 K unless otherwise noted. The cells were disrupted by two passes through a French pressure cell adjusted to 69 MPa in a 40k rapid fill cell with a flow rate of 20 drops per minute. The resulting extract was then stirred in the presence of 1 *M* NaCl. Unbroken cells and pelletable debris were removed by centrifugation at 31 000*g* for 15 min. Bacterial membranes were pelleted by ultracentrifugation at 192 000*g* for 1 h. The ultracentrifugation supernatant, which contained the majority of the acid phosphatase activity, was dialyzed against 100 volumes of 20 m*M* sodium phosphate pH 7.4 containing 1.0 *M* NaCl (buffer *A*) for 24 h.

The dialyzed sample was applied (5 ml min^{-1}) onto a HiTrap Chelating HP chromatography column (Amersham Biosciences, 5 ml) pre-equilibrated with buffer *A* and charged with 100 m*M* $NiCl_2$. The applied sample was washed with several column volumes of buffer *A*, followed by a second wash with buffer *A* supplemented with 100 m*M* imidazole. The latter wash liberated large amounts of contaminating proteins as shown in Fig. 1 (chromatogram peak *B*). After the washing steps, a linear gradient of 100–500 m*M* imidazole in buffer *A* (over eight column volumes) was performed. The enzyme of interest eluted at approximately 200 m*M* imidazole and was judged by SDS–PAGE to be the correct molecular weight and sufficiently pure for crystallization trials (Fig. 1, chromatogram peak *C*). The purified phosphatase was dialyzed against 50 m*M* sodium acetate pH 6.0 for 24 h and concentrated to 10 mg ml^{-1} using Amicon Ultra 15 centrifugal filter devices (10 kDa molecular-weight cutoff). Protein

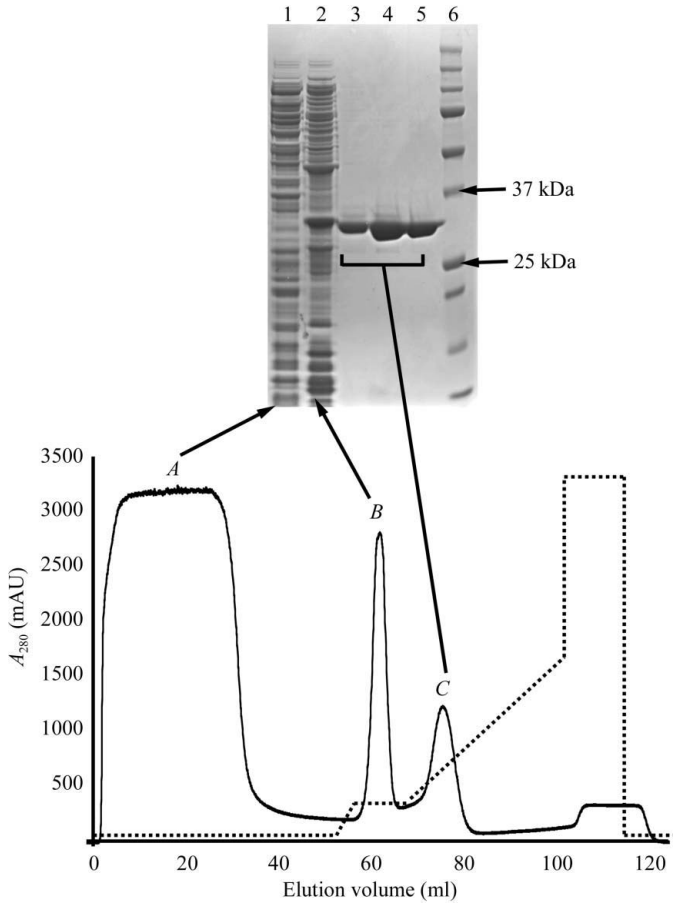


Figure 1
SDS–PAGE analysis of fractions from nickel-affinity chromatography. Lane 1 is the flowthrough indicated as peak *A* on the chromatogram. Lane 2 represents a sample from peak *B*, which resulted from a washing step at 100 m*M* imidazole. Lanes 3, 4 and 5 correspond to the first, middle and last fractions from peak *C*. Peak *C* eluted at 100–200 m*M* imidazole and contained the desired enzyme. Lane 6 shows molecular-weight standards.

Table 1
Data-collection and processing statistics.

Values in parentheses are for the outer resolution shell of data.

	Before annealing	After annealing	After annealing (reprocessed to 1.8 Å)
Wavelength (Å)	1.2037	1.2037	1.2037
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	a = 53.35, b = 90.09, c = 104.15	a = 53.36, b = 90.15, c = 104.23	a = 53.36, b = 90.15, c = 104.23
Diffraction resolution (Å)	45.91–1.80 (1.86–1.80)	45.12–1.60 (1.66–1.60)	45.12–1.80 (1.86–1.80)
Exposure time (s per frame)	7	10	10
Total oscillation range (°)	110	180	95
No. of observations	170571	451942	172003
No. of unique reflections	45878	66821	46860
Redundancy	3.7 (3.7)	6.7 (6.4)	3.8 (3.5)
Refined mosaicity (°)	1.6	0.8	0.8
Completeness (%)	96.9 (95.4)	99.5 (98.5)	98.8 (97.2)
Average $I/\sigma(I)$	8.9 (2.7)	11.1 (3.0)	20 (9.2)
$R_{sym}(I)$	0.066 (0.372)	0.062 (0.443)	0.038 (0.109)

concentration was estimated with the Coomassie Plus kit (Pierce) using bovine serum albumin standards. Purified *B. anthracis* class C NSAP exhibited acid phosphatase activity using the substrates *p*-nitrophenyl phosphate and 4-methylumbelliferyl phosphate.

2.2. Crystallization

Crystals were grown at 293 K using Cryschem 24-well sitting-drop plates (Hampton Research). Initial screening was performed using the Index Screen from Hampton Research. 1.5 µl volumes of the protein at a concentration of 10 mg ml^{-1} and the reservoir were mixed and equilibrated against 1.0 ml reservoir solution. Large well diffracting crystals were obtained directly from Index reagent 39 [0.1 *M* HEPES pH 7.0, 30%(*w*/*v*) Jeffamine ED-2001 pH 7.0]. These crystals typically appeared within two weeks after setup as rectangular blocks with dimensions of 0.4 × 0.2 × 0.2 mm (Fig. 2). Prior to cryogenic data collection, the crystals were soaked in the harvest buffer (0.1 *M* HEPES pH 7.0, 35% Jeffamine ED-2001 pH 7.0) supplemented with 25%(*w*/*v*) PEG 200.

2.3. Data collection, processing and flash-annealing

Two diffraction data sets, corresponding to before and after flash-annealing, were collected from the same crystal at ALS beamline 4.2.2 using a NOIR-1 CCD detector. For both data sets, the detector distance and angle were set to 125 mm and 0°, respectively, and the oscillation width was 0.5° per frame. The two data sets differed in exposure time per frame and total oscillation width. The first data set consisted of 110° of data collected with an exposure time of 7 s per frame, whereas the second data set consisted of 180° of data with exposure time of 10 s per frame.

Processing of the first data set in *d*TREK* (Pflugrath, 1999) suggested space group $P2_12_12_1$, with unit-cell parameters a = 53.3, b = 90.1, c = 104.1 Å. Analysis of crystal packing using the method of Matthews indicated a probable solvent content of 38% with two protein molecules in the asymmetric unit and V_M = 2.8 Å^3 Da^{-1} (Matthews, 1968). This data set exhibited acceptable statistics to 1.8 Å resolution and the refined mosaicity was 1.6° (Table 1).

Although the 1.8 Å data set was quite satisfactory, we nonetheless decided to test the response of the crystal to flash-annealing (Yeh & Hol, 1998). With the crystal remaining in data-collection position, the cold nitrogen stream was blocked for approximately 4 s using a plastic card. Visualization of the crystal through the beamline video system verified that the crystal had indeed thawed during the brief blockage of the stream and that the crystal was refrozen upon removing the card. This procedure was repeated three times.

Test diffraction images collected from the annealed crystal suggested an improvement in resolution and so another complete data set was recorded. This second data set could be processed to 1.6 Å resolution and the mosaicity improved to 0.85° (Table 1). Thus, flash-annealing appeared to cause a useful improvement in diffraction quality.

To more accurately assess the improvement in crystal quality caused by annealing, the data set collected after annealing was truncated at 1.8 Å resolution and only the first 95° of data were used so that the resulting redundancy was comparable to that of pre-annealing data set. Statistics from this calculation are shown in Table 1. Comparison of the two 1.8 Å data sets shows that the data set collected before annealing had an overall average $I/\sigma(I)$ of 8.9, compared with an $I/\sigma(I)$ of 20 after annealing (Table 1). The average $I/\sigma(I)$ in the high-resolution bin improved from 2.7 before annealing to 9.2 after annealing. Likewise, the overall R_{sym} improved from 0.066 (0.372 for the outer shell) before annealing to 0.038 (0.109 for the outer shell) after annealing. We note that the factor of 2–3 increase in $I/\sigma(I)$ exceeds the increase expected from counting statistics arising from the longer exposure time used for the second data set (Dauter & Wilson, 2001). We therefore conclude that the observed improvement in data quality was primarily a consequence of increased lattice order induced by flash-annealing.

The amino-acid sequence of *B. anthracis* class C NSAP was used in a *BLAST* search (Altschul *et al.*, 1990) of the Protein Data Bank (Berman *et al.*, 2000) to determine whether suitable homologs were available for use as search models in molecular-replacement studies. The closest homolog was AphA from *E. coli*, which shares only 17% global sequence identity with *B. anthracis* class C NSAP. Thus, structure determination is being pursued with experimental phasing techniques.

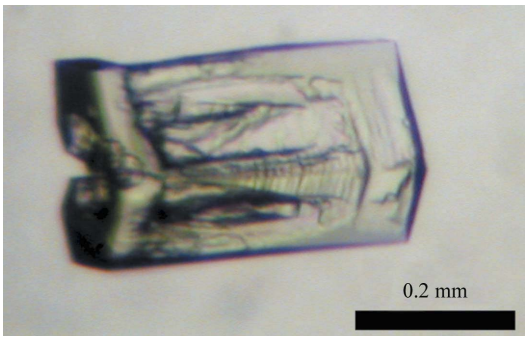


Figure 2
A crystal of *B. anthracis* class C NSAP grown by sitting-drop vapor diffusion over a reservoir containing 0.1 *M* HEPES pH 7.0, 30%(*w*/*v*) Jeffamine ED-2001 pH 7.0.

This research was supported by National Institutes of Health grant U54 AI057160 to the Midwest Regional Center of Excellence for Biodefense and Emerging Infectious Diseases Research (MRCE, JJT and TJR), by the University of Missouri Research Board (JJT) and by a subproject of USDA-ARS Program for Prevention of Animal Infectious Diseases (PPAID) Advanced Technologies for Vaccines and Diagnostics (TJR) under cooperative agreement USDA-ARS 58-1940-5-519. We thank George Stewart and Brian Thompson for providing genomic DNA. We also thank Jay Nix and Darren Sherrell for their assistance at ALS beamline 4.2.2. The Advanced Light

Source is supported by the Director, Office of Science, Office of Basic Energy Sciences and Materials Sciences Division of the US Department of Energy under Contract No. DE-AC03-76SF00098 at Lawrence Berkeley National Laboratory.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.

Calderone, V., Forleo, C., Benvenuti, M., Thaller, M. C., Rossolini, G. M. & Mangani, S. (2004). *J. Mol. Biol.* **335**, 761–773.

Calderone, V., Forleo, C., Benvenuti, M., Thaller, M. C., Rossolini, G. M. & Mangani, S. (2006). *J. Mol. Biol.* **355**, 708–721.

Dauter, Z. & Wilson, K. S. (2001). *International Tables For Crystallography*, Vol. *F*, edited by M. G. Rossmann & E. Arnold, pp. 177–196. Dordrecht: Kluwer Academic Publishers.

Felts, R. L., Reilly, T. J., Calcutt, M. J. & Tanner, J. J. (2006). *Acta Cryst.* F**62**, 32–35.

Felts, R. L., Reilly, T. J. & Tanner, J. J. (2005). *Biochim. Biophys. Acta*, **1752**, 107–110.

Green, B. A., Baranyi, E., Reilly, T. J., Smith, A. L. & Zlotnick, G. W. (2005). *Infect. Immun.* **73**, 4454–4457.

Knowles, J. R. (1980). *Annu. Rev. Biochem.* **49**, 877–919.

Malke, H. (1998). *Appl. Environ. Microbiol.* **64**, 2439–2442.

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

Ou, Z., Felts, R. L., Reilly, T. J., Nix, J. C. & Tanner, J. J. (2006). *Acta Cryst.* F**62**, 464–466.

Passariello, C., Schippa, S., Iori, P., Berlutti, F., Thaller, M. C. & Rossolini, G. M. (2003). *Biochim. Biophys. Acta*, **1648**, 203–209.

Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.

Plessis, E. M. du, Theron, J., Joubert, L., Lotter, T. & Watson, T. G. (2002). *Syst. Appl. Microbiol.* **25**, 21–30.

Read, T. D. *et al.* (2003). *Nature (London)*, **423**, 81–86.

Read, T. D., Salzberg, S. L., Pop, M., Shumway, M., Umayam, L., Jiang, L., Holtzapple, E., Busch, J. D., Smith, K. L., Schupp, J. M., Solomon, D., Keim, P. & Fraser, C. M. (2002). *Science*, **296**, 2028–2033.

Reilly, T. J. & Calcutt, M. J. (2004). *Protein Expr. Purif.* **33**, 48–56.

Reilly, T. J., Chance, D. L. & Smith, A. L. (1999). *J. Bacteriol.* **181**, 6797–6805.

Reilly, T. J., Felts, R. L., Henzl, M. T., Calcutt, M. J. & Tanner, J. J. (2006). *Protein Expr. Purif.* **45**, 132–141.

Rossolini, G. M., Schippa, S., Riccio, M. L., Berlutti, F., Macaskie, L. E. & Thaller, M. C. (1998). *Cell. Mol. Life Sci.* **54**, 833–850.

Thaller, M. C., Schippa, S. & Rossolini, G. M. (1998). *Protein Sci.* **7**, 1647–1652.

Yeh, J. I. & Hol, W. G. (1998). *Acta Cryst.* D**54**, 479–480.