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

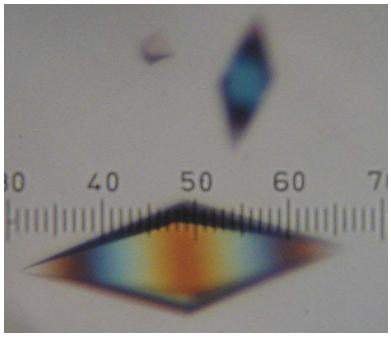
Crystallization of a newly discovered histidine acid phosphatase from *Francisella tularensis*

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, Missouri 65211, USA, [b]Department of Veterinary Pathobiology, College of Veterinary Medicine, University of Missouri-Columbia, Columbia, Missouri 65212, USA, [c]Veterinary Medical Diagnostic Laboratory, College of Veterinary Medicine, University of Missouri-Columbia, Columbia, Missouri 65212, USA, and [d]Department of Biochemistry, University of Missouri-Columbia, Columbia, Missouri 65211, USA

Correspondence e-mail: tannerjj@missouri.edu

Received 4 October 2005
Accepted 29 November 2005
Online 16 December 2005



Francisella tularensis is a highly infectious bacterial pathogen that is considered by the Centers for Disease Control and Prevention to be a potential bioterrorism weapon. Here, the crystallization of a 37.2 kDa phosphatase encoded by the genome of *F. tularensis* subsp. *holarctica* live vaccine strain is reported. This enzyme shares 41% amino-acid sequence identity with *Legionella pneumophila* major acid phosphatase and contains the RHGXRXP motif that is characteristic of the histidine acid phosphatase family. Large diffraction-quality crystals were grown in the presence of Tacsimate, HEPES and PEG 3350. The crystals belong to space group $P4_12_12$, with unit-cell parameters a = 61.96, c = 210.78 Å. The asymmetric unit is predicted to contain one protein molecule, with a solvent content of 53%. A 1.75 Å resolution native data set was recorded at beamline 4.2.2 of the Lawrence Berkeley National Laboratory Advanced Light Source. Molecular-replacement trials using the human prostatic acid phosphatase structure as the search model (28% amino-acid sequence identity) did not produce a satisfactory solution. Therefore, the structure of *F. tularensis* histidine acid phosphatase will be determined by multiwavelength anomalous dispersion phasing using a selenomethionyl derivative.

1. Introduction

Francisella tularensis is a Gram-negative facultative intracellular coccobacillus that causes the zoonotic disease tularemia (Ellis *et al.*, 2002). The organism grows readily in broth culture, it can be isolated from numerous rodent and arthropod vectors and it is highly infectious (Dennis *et al.*, 2001; Oyston *et al.*, 2004). Ulceroglandular tularemia results from infections acquired through the skin or mucous membranes. Of chief concern in the context of biodefense is the possibility of infection by inhalation of the aerosolized pathogen, which results in the pneumonic form of the disease (Dennis *et al.*, 2001). Inhalation tularemia is by far the most dangerous form of tularemia and has a case fatality rate of up to 30% if untreated (Oyston *et al.*, 2004). Accordingly, the Centers for Disease Control and Prevention have designated *F. tularensis* to be a Category A pathogen, a class that includes the causative pathogens of anthrax, botulism, plague and smallpox. Thus, there is renewed interest in understanding the basic biochemistry and biology of *F. tularensis* in order to facilitate development of improved vaccines and antimicrobial agents that will provide protection and treatment in the event of a *Francisella*-based attack.

As part of our ongoing research on the roles of phosphatases in *F. tularensis* intracellular survival and virulence (Baron *et al.*, 1999; Reilly *et al.*, 1996, 2006; Felts *et al.*, 2005), we analyzed the complete but unannotated genome sequence of *F. tularensis* subsp. *holarctica* live vaccine strain (LVS) in search of possible phosphatase genes. Comparison of the open reading frames of the LVS genome with current sequence databases using *BLAST* (Altschul *et al.*, 1990) identified a 37.2 kDa ortholog of the major acid phosphatase from *Legionella pneumophila* (MAP; Aragon *et al.*, 2001). The deduced amino-acid sequence of the *F. tularensis* protein is 41% identical (134 out of 330 residues) to that of MAP (Fig. 1) and exhibited lower but significant homology to several phosphatases from eukaryotic sources. These enzymes share the conserved RHGXRXP motif (Fig. 1) characteristic of the histidine acid phosphatase (HAP) family (Van Etten *et al.*, 1991).

MAP is thought to be secreted *via* the type II secretion system and may be involved in intracellular infection (Aragon *et al.*, 2001). It is not known whether *F. tularensis* HAP (FtHAP) is secreted by an analogous system or whether FtHAP plays a role in virulence or intracellular survival. In parallel with our studies of substrate specificity, kinetics and biological function of this newly discovered enzyme, we have crystallized FtHAP as a first step toward structure determination.

2. Methods and results

2.1. Cloning, expression and protein purification

The *hap* gene was amplified by PCR from genomic DNA obtained from the *F. tularensis* subsp. *holarctica* live vaccine strain and cloned into pET20b (Novagen) using *Nco*I and *Xho*I sites. The recombinant protein was expressed in *Escherichia coli* using procedures similar to those described previously for the AcpA phosphatase (Reilly *et al.*, 2006). Harvested cells were resuspended in 50 m*M* sodium acetate pH 6.0 (buffer *A*) and lysed in a French pressure cell. The resulting mixture exhibited strong phosphatase activity as measured by a discontinuous colorimetric activity assay using *p*-nitrophenylphosphate as the substrate (Reilly *et al.*, 1999). This assay was also performed after each step in the following purification procedure to identify the enzyme of interest and its relative activity in the collected fractions. Sodium chloride was added to the broken cells to a final concentration of 1.0 *M* and two centrifugation cycles of (i) 27 200*g* for 20 min and (ii) 184 000*g* for 1.5 h were performed. The supernatant from the latter centrifugation step was dialyzed against buffer *A* for 24 h. The sample was applied onto a HiTrap SP HP cation-exchange column (Amersham Biosciences) that had been equilibrated with buffer *A*. A linear NaCl gradient of 0–1.0 *M* over eight column volumes was used to elute the protein from the column. FtHAP was eluted at 500 m*M* NaCl. The protein was then dialyzed against 10 m*M* Na_3PO_4 pH 7.0 (buffer *B*) for 24 h. The dialyzed sample was applied onto a HiTrap Chelating HP column (Amersham Biosciences) that had been charged with 0.1 *M* $NiSO_4$ and equilibrated with buffer *B*. A linear imidazole gradient of 0–1.0 *M* was used to elute the protein from the column. FtHAP eluted at 350 m*M* imidazole. The protein was then dialyzed against buffer *A* for 24 h and concentrated to 10 mg ml^{-1}. Protein concentration was determined by absorption spectroscopy using an extinction coefficient (λ = 280 nm) of 48 360 M^{-1} cm^{-1} predicted by the ExPASy server (Gasteiger *et al.*, 2005). Protein purity was evaluated by SDS–PAGE.

2.2. Crystallization

Crystallization trials were performed using the sitting-drop method of vapor diffusion at 293 K. Initial screening for crystallization conditions was performed using the Index Screen from Hampton Research. Equal volumes of the protein (1.5 µl) and the reservoir (1.5 µl) were mixed and allowed to equilibrate with 1.0 ml of reservoir for 24–48 h. Index Screen reagents 3–6, 21, 39, 42, 57, 63, 65, 67–69, 71, 72, 75–80, 83–85, 87, 92 and 94 produced crystals of various size and quality. Condition 63 produced the largest and most well defined crystals. This condition contained 5%(*v*/*v*) Tacsimate, 0.1 *M* HEPES pH 7.0 and 25%(*w*/*v*) PEG 3350. Tacsimate (Hampton Research, HR2-755) is a mixture of organic acids that includes sodium malonate, sodium acetate, triammonium citrate, succinic acid, DL-malic acid and sodium formate (Bob Cudney, personal communication). The initial crystals were diamond-shaped with jagged edges and in some cases multiple crystals were fused together (Fig. 2*a*). One round of optimization resulted in large single crystals having sharp edges and dimensions of 0.4 × 0.4 × 0.7 mm (Fig. 2*b*). The final optimized

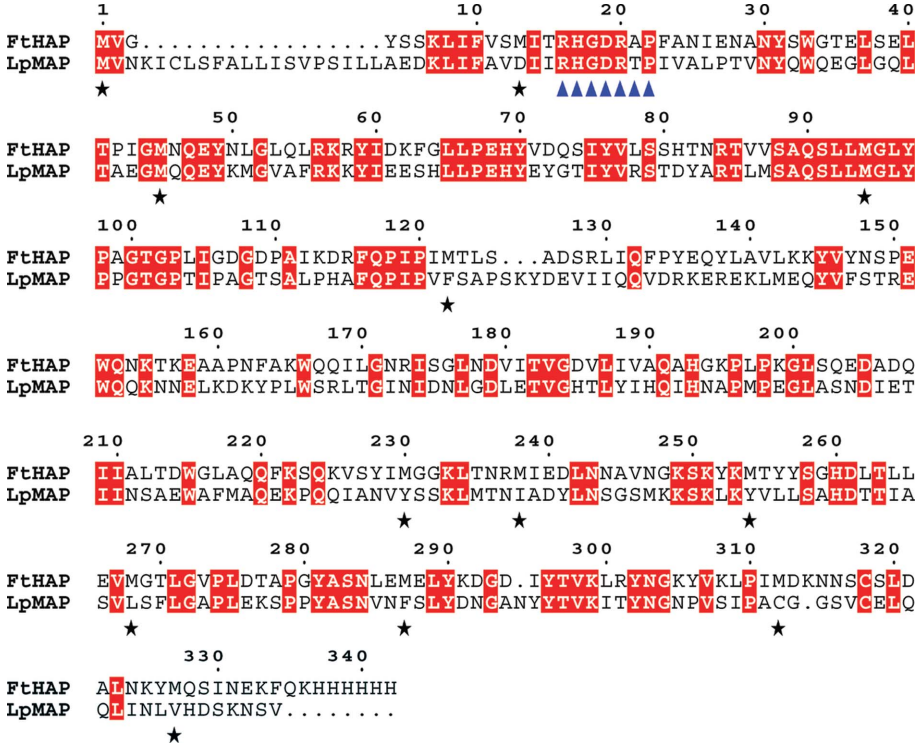


Figure 1
Amino-acid sequence alignment of the *F. tularensis* histidine acid phosphatase (FtHAP) construct used for crystallization studies and the major acid phosphatase of *L. pneumophila* (LpMAP). Identical residues are highlighted in red, the RHGXRXP signature motif of histidine acid phosphatases is denoted by blue triangles and black stars denote the Met residues of FtHAP.

Table 1
Data-collection and processing statistics.

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

Space group	$P4_12_12$
Unit-cell parameters (Å)	a = 61.96, c = 210.78
Wavelength (Å)	1.12718
Resolution range (Å)	46.47–1.75 (1.81–1.75)
Total reflections	243916
Unique reflections	41088
Redundancy	5.94 (4.72)
Mosaicity (°)	0.466
R_{merge}	0.041 (0.299)
Completeness (%)	96.2 (91.7)
Average $I/\sigma(I)$	25.4 (5.2)

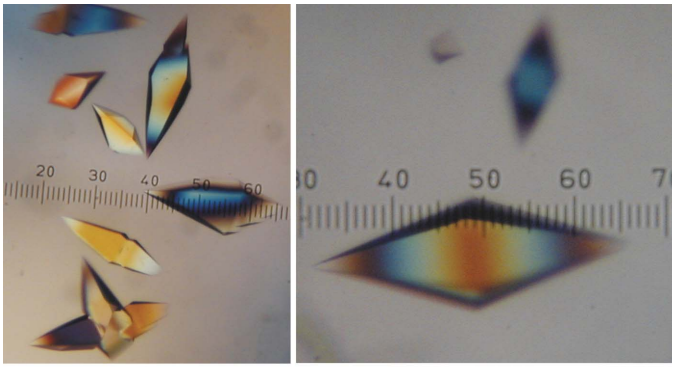
condition contained 10%(*v*/*v*) Tascimate, 0.1 *M* HEPES pH 7.0 and 19%(*w*/*v*) PEG 3350. These crystals typically appeared within 48 h of setup.

The crystals were cryoprotected in the harvest buffer [10%(*v*/*v*) Tacsimate, 0.1 *M* HEPES pH 7.0, 24% PEG 3350] supplemented with 25% PEG 200 as follows. Firstly, 10–30 µl of harvest buffer was added to a sitting drop that contained crystals. Next, the liquid in the drop was exchanged for the cryoprotectant in five steps over a period of 10–20 min without moving the crystals. At each step, the concentration of PEG 200 was increased by 5% and the drop was gently mixed by aspiration without disturbing the crystals. Finally, the crystals were picked up with Hampton mounting loops and plunged into liquid nitrogen.

2.3. Data collection and processing

Initial characterization of X-ray diffraction was performed using an R-AXIS IV image-plate detector coupled to a Rigaku copper rotating-anode generator. Autoindexing of diffraction images using *CrystalClear* (Pflugrath, 1999) suggested a primitive tetragonal lattice with unit-cell parameters a = 62.0, c = 211.0 Å. Frozen crystals were transported to Lawrence Berkeley National Laboratory for high-resolution data collection at the Advanced Light Source (ALS).

Diffraction data were collected at ALS beamline 4.2.2 using a NOIR-1 detector with the detector distance and angle set to 150 mm and 17°, respectively. A total of 180° of data were collected using an oscillation angle of 0.5° and an exposure time of 8 s per image. The data were integrated and scaled to 1.75 Å resolution using *d*TREK* (Pflugrath, 1999). The refined unit-cell parameters were determined to be a = 61.96, c = 210.78 Å. Analysis of the data with *dtcell* (Pflugrath, 1999) suggested space group $P4_12_12$. Matthews calculations suggested that this crystal form has one molecule in the asymmetric unit, 53% solvent content and a Matthews coefficient of 2.7 Å^3 Da^{-1} (Matthews, 1968). See Table 1 for data-processing statistics.



(*a*) (*b*)

Figure 2
Crystals of *F. tularensis* HAP. (*a*) Initial crystals grown from Hampton Research Index Screen reagent 63. These crystals grew to approximate dimensions of 0.5 × 0.2 × 0.2 mm within 48 h. (*b*) Crystals of FtHAP obtained after optimizing the crystallization condition to 10%(*v*/*v*) Tacsimate, 0.1 *M* HEPES pH 7.0 and 19%(*w*/*v*) PEG 3350. This crystal has dimensions of 0.4 × 0.4 × 0.7 mm. The smallest division of the ruler in both panels corresponds to 0.02 mm.

Since the apparent Laue class is 4/*mmm*, the possibility of merohedral twinning was considered. A plot of the cumulative intensity distribution for acentric reflections did not display the sigmoidal shape characteristic of twinned data (Rees, 1980). The average value of $\langle I^2(h)\rangle/\langle I(h)\rangle^2$ was 2.2 for acentric reflections, which is close to the value of 2.0 expected for untwinned data (Redinbo & Yeates, 1993). For reference, a value of 1.5 is expected in the case of perfect hemihedral twinning (Redinbo & Yeates, 1993). Based on these results, we do not anticipate difficulties arising from twinning during structure determination.

The FtHAP sequence was compared with sequences in the Protein Data Bank (PDB; Berman *et al.*, 2000) to identify a search model for molecular-replacement calculations. The closest relative in the PDB is human prostatic acid phosphatase (PDB code 1cvi; Jakob *et al.*, 2000), which shares 28% global amino-acid sequence identity with FtHAP. Molecular-replacement trials were performed with *MOLREP* (Vagin & Teplyakov, 1997) using 1cvi as the search model. All possible space groups with Laue symmetry 4/*mmm* were tested. The top solution had $R > 0.6$ and correlation coefficient < 0.12, which indicated that molecular replacement is not a suitable phasing method. Structure determination using crystals of selenomethionyl FtHAP is in progress. 12 Met residues are expected in the asymmetric unit (Fig. 1).

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 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.
Aragon, V., Kurtz, S. & Cianciotto, N. P. (2001). *Infect. Immun.* **69**, 177–185.
Baron, G. S., Reilly, T. J. & Nano, F. E. (1999). *FEMS Microbiol. Lett.* **176**, 85–90.
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.
Dennis, D. T., Inglesby, T. V., Henderson, D. A., Bartlett, J. G., Ascher, M. S., Eitzen, E., Fine, A. D., Friedlander, A. M., Hauer, J., Layton, M., Lillibridge, S. R., McDade, J. E., Osterholm, M. T., O'Toole, T., Parker, G., Perl, T. M., Russell, P. K. & Tonat, K. (2001). *JAMA*, **285**, 2763–2773.
Ellis, J., Oyston, P. C., Green, M. & Titball, R. W. (2002). *Clin. Microbiol. Rev.* **15**, 631–646.
Felts, R. L., Reilly, T. J. & Tanner, J. J. (2005). *Biochim. Biophys. Acta*, **1752**, 107–110.
Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. M. Walker, pp. 571–607. Totowa, NJ, USA: Humana Press.

Jakob, C. G., Lewinski, K., Kuciel, R., Ostrowski, W. & Lebioda, L. (2000). *Prostate*, **42**, 211–218.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Oyston, P. C., Sjostedt, A. & Titball, R. W. (2004). *Nature Rev. Microbiol.* **2**, 967–978.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Redinbo, M. R. & Yeates, T. O. (1993). *Acta Cryst.* D**49**, 375–380.
Rees, D. C. (1980). *Acta Cryst.* A**36**, 578–581.
Reilly, T. J., Baron, G. S., Nano, F. E. & Kuhlenschmidt, M. S. (1996). *J. Biol. Chem.* **271**, 10973–10983.
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 Exp. Purif.* **45**, 132–141.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Van Etten, R. L., Davidson, R., Stevis, P. E., MacArthur, H. & Moore, D. L. (1991). *J. Biol. Chem.* **266**, 2313–2319.