Crystallization and preliminary crystallographic analysis of the major acid phosphatase from *Legionella pneumophila*

Dan Zhou,^{a,‡} Yang Pan,^{a,‡} Xiaofang Chen,^a Nannan Zhang^{a,*} and Honghua Ge^{a,b,*}

Received 18 March 2015

Accepted 25 April 2015

Edited by P. Dunten, Stanford Synchrotron Radiation Lightsource, USA

‡ These authors contributed equally to this work.

Keywords: *Legionella pneumophila* major acid phosphatase; histidine acid phosphatase; acid phosphatase.

^aInstitute of Health Sciences and School of Life Sciences, Anhui University, Hefei 230601, People's Republic of China, and ^bModern Experiment Technology Center, Anhui University, Hefei 230601, People's Republic of China.

*Correspondence e-mail: nancy_zhangnn@foxmail.com, gehonghua@gmail.com

The major acid phosphatase from *Legionella pneumophila* (*Lp*MAP) belongs to the histidine acid phosphatase superfamily. It contains the characteristic histidine acid phosphatase (HAP) sequence motif RHGXRXP responsible for the hydrolysis of a phosphoryl group from phosphate monoesters under acidic conditions. Here, the crystallization and preliminary X-ray analysis of crystals of *Lp*MAP in the apo form and in complex with L-(+)-tartrate are described. By using the hanging-drop vapour-diffusion method, apo *Lp*MAP and *Lp*MAP–tartrate were crystallized in space group $P2_1$, with unit-cell parameters $a = 91.50$, $b = 56.48$, $c = 146.35$ Å, $\beta = 110.01^\circ$, and in space group $P1$, with unit-cell parameters $a = 55.51$, $b = 73.51$, $c = 98.78$ Å, $\alpha = 78.82$, $\beta = 77.65$, $\gamma = 67.73^\circ$, respectively. Diffraction data were collected at 100 K and the phases were determined using the molecular-replacement method.

1. Introduction

The histidine phosphatase superfamily is a large group of proteins. They share a common catalytic core centred on a histidine which is phosphorylated during the catalytic reaction. The superfamily has been classified into two branches sharing very limited sequence similarity. The first consists of functionally diverse enzymes, including various phosphatases and phosphoglycerate mutases. The second is comprised mainly of histidine acid phosphatases and phytases (Rigden, 2008).

Histidine acid phosphatase (HAP) contains the highly conserved RHGXRXP motif (where X represents any residue), which was first identified in human prostatic acid phosphatase (hPAP; Van Etten *et al.*, 1991). The characteristic HAP sequence motif can use an active-site histidine to catalyze the hydrolysis of a phosphoryl group from phosphate monoesters under acidic conditions. The first bacterial HAP crystal structure was solved and reported from *Francisella tularensis* (FtHAP; Singh *et al.*, 2009). FtHAP is a periplasmic acid phosphatase and has been shown to be essential for the intramacrophage survival and virulence of *F. tularensis* (Mohapatra *et al.*, 2008). Meanwhile, crystal structures of hPAP (LaCount *et al.*, 1998; Ortlund *et al.*, 2003; Jakob *et al.*, 2000) and the rat prostatic acid phosphatase (rPAP; Lindqvist *et al.*, 1993; Schneider *et al.*, 1993) are also available in the PDB.

Legionella pneumophila, the causative agent of Legionnaires' disease, is a Gram-negative facultative intracellular pathogen that is capable of multiplying in a wide spectrum of

eukaryotic cells (Fields *et al.*, 2002). Major acid phosphatase (EC 3.1.3.2) from *L. pneumophila* (*LpMAP*) has been identified as an extracellular secreted tartrate-sensitive acid phosphatase (Aragon *et al.*, 2001). It is 354 amino acids in length, with a molecular mass of 39.5 kDa, and is predicted by the *SignalP4.1* software (Petersen *et al.*, 2011) to have an N-terminal 21-residue signal peptide. *LpMAP* contains the highly conserved HAP sequence motif and shares 39% and 29% sequence identity with *FiHAP* and *hPAP*, respectively.

Unlike *FiHAP*, *LpMAP* is not essential for intracellular infection by *L. pneumophila* since *map* mutants grew within macrophage-like U937 cells and *Hartmannella amoebae* to the same degree as wild-type *Legionella* (Aragon *et al.*, 2001). However, it may play an important role in the survival of *Legionella* within activated macrophages or neutrophils. It is believed that *LpMAP* probably promotes extracellular survival and spread *in vivo* (Aragon *et al.*, 2001). In order to better understand the details of the biological functions of *LpMAP*, we purified recombinant *LpMAP* expressed in *Escherichia coli*, crystallized the protein in the apo form and in complex with L-(+)-tartrate and performed preliminary X-ray analysis of the crystals.

2. Materials and methods

2.1. Macromolecule production

DNA for *L. pneumophila* MAP (NCBI code YP_095152) without the 21-amino-acid signal peptide was amplified by PCR from *L. pneumophila* DNA using the following sense and antisense primers containing NdeI and XhoI sites (underlined): 5'-CTTTTTCATATGGAAGATAAGCTTATTTT-3' and 5'-CTTCTCGAGTTACTGAATTTTGAATC-3'. The amplified fragments were then cloned into a modified pET-28a vector (Novagen) by deleting the sequence AGCAGCGGCCTGGTGC CGCGCGGCAGC between the NcoI and NdeI restriction sites. The resultant N-terminal amino-acid sequence of recombinant *LpMAP* is MGHHHHHH**MD**, where the residue in bold is the *LpMAP* protein starting at residue 22. The recombinant plasmid was sequenced to ensure that no mutations had occurred during the polymerase chain reaction. This recombinant plasmid was transformed into *E. coli* strain Rosetta. Cells were grown at 310 K in Luria-Bertani (LB) medium supplemented with 0.1 mg ml⁻¹ kanamycin. Expression of *LpMAP* was induced at an OD₆₀₀ of 0.8–1.0 by adding 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) followed by incubation at 289 K for 16 h. The cells were harvested and sonicated in 20 mM Tris-HCl pH 8.0 buffer containing 200 mM NaCl in an ice-water bath. The lysate was clarified by centrifugation at 15 000g for 30 min at 277 K. The supernatant was loaded onto an Ni²⁺-NTA column (GE Healthcare) pre-equilibrated with binding buffer (20 mM Tris-HCl pH 8.0, 200 mM NaCl). The resin was washed with ice-cold washing buffer (20 mM Tris-HCl pH 8.0, 200 mM NaCl, 50 mM imidazole) and the target protein was eluted with elution buffer (20 mM Tris-HCl pH 8.0, 200 mM NaCl, 250 mM imidazole).

Table 1

Macromolecule-production information.

Source organism	<i>L. pneumophila</i>
DNA source	DNA for <i>L. pneumophila</i> MAP (NCBI entry code YP_095152)
Forward primer†	5'-CTTTTTCATATGGAAGATAAGCTTATTTT-3'
Reverse primer‡	5'-CTTCTCGAGTTACTGAATTTTGAATC-3'
Cloning vector	Modified pET-28a vector§ (Novagen)
Expression vector	Modified pET-28a vector§ (Novagen)
Expression host	<i>E. coli</i> strain Rosetta
Complete amino-acid sequence of the construct produced	MGHHHHHHMEDKLIFAVDIIIRHGRPTIPVALPTV-NYQWQEGGLQTAEGMQQEQYKMGVAFRKKYIE-ESHLLEPHYEYGTIYVRSTDYARTLMSAQSLLMGLYPPGTGPTIPAGTSALPHAFAQPIPVFSAP-SKYDEVIIQQVDRKERELMEQYVFSTREWQQ-KNNELKDKYPLWSRLTGINIDNLGDLETVGHT-LYIHQIHNAPMPEGLASNDIETIINSAEAWAFMAQEKPPQIANVYSSKLMNTIADYLSGSMKKSKLYVLLSAHDTTASVLSFLGAPLEKSPPYASNVNFSLYDNGANYTYVKITYNGNPVSIPACGGSVCLEQLQLINLVHDSKNSV

† The NdeI site is underlined. ‡ The XhoI site is underlined. § Obtained by deleting the sequence AGCAGCGGCCTGGTGC CGCGCGGCAGC between the NcoI and NdeI restriction sites.

The eluate was subsequently loaded onto a Superdex 200 column (Amersham Biosciences) in 20 mM Tris-HCl pH 8.0, 200 mM NaCl. The eluted recombinant *LpMAP* was concentrated to 60 mg ml⁻¹ by centrifugal ultrafiltration (Millipore) for apo *LpMAP* crystallization assays. In order to obtain crystals of *LpMAP* in complex with inhibitor, the protein was buffer-exchanged with 20 mM Tris-HCl pH 8.0, 200 mM NaCl, 10 mM L-(+)-tartrate prior to crystallization assays. The protein concentration was determined by the Bradford method (Bio-Rad protein assay) using bovine serum albumin as a standard. The purity of the fractions was checked by SDS-PAGE (greater than 95% purity). The yield of the soluble recombinant protein was about 50 mg per litre of induced culture. Macromolecule-production information is summarized in Table 1.

2.2. Crystallization

Initial crystallization trials were set up with the Index (Hampton Research) and Proplex (Molecular Dimensions) reagent kits at 285 K using the hanging-drop vapour-diffusion method. Each drop, consisting of 1 µl protein solution (8–60 mg ml⁻¹) and an equal volume of reservoir solution, was equilibrated against 200 µl reservoir solution. Further crystal optimization experiments were performed by systematic variation of the precipitant concentration and protein concentration and by testing the effects of additives. Agarose gels, which are known to prevent convection and crystal sedimentation, provided a microgravity environment for the growth of high-quality protein crystals (Biertümpfel *et al.*, 2002). 1% (m/v) low gelling temperature agarose (*T_g* = 301 K; Takara) was first melted at 315 K and kept liquid at 303 K. This agarose gel solution was then directly added to the reservoir solution to a final agarose concentration of 0.025% (m/v) prior to setting up the drops. The crystals grew to their full dimensions in the final conditions after one week. Crystallization information is summarized in Table 2.

Table 2
Crystallization.

	Apo <i>Lp</i> MAP	<i>Lp</i> MAP–tartrate
Method	Vapour diffusion	Vapour diffusion
Plate type	Hanging drop	Hanging drop
Temperature (K)	285	285
Protein concentration (mg ml ^{−1})	30	15
Buffer composition of protein solution	20 mM Tris–HCl pH 8.0, 200 mM NaCl	2 mM Tris–HCl pH 8.0, 200 mM NaCl,
		10 mM L-(+)-tartrate
Composition of reservoir solution	0.2 M sodium formate, 17%(w/v) PEG 3350,	0.1 M Tris pH 8.5, 21%(w/v) PEG 6000,
	0.025%(w/v) low gelling temperature agarose	0.025%(w/v) low gelling temperature agarose
Volume and ratio of drop	1 µl:1 µl	1 µl:1 µl
Volume of reservoir (µl)	200	200

The crystals were harvested using cryoloops and immersed briefly in a cryoprotectant solution consisting of 80%(v/v) reservoir solution and 20%(v/v) glycerol. The crystals were subsequently flash-cooled and stored in liquid nitrogen, and were transferred to beamline BL17U of the Shanghai Synchrotron Radiation Facility (SSRF) for X-ray diffraction analysis and data collection.

2.3. Data collection and processing

After screening for diffraction quality, X-ray data were collected to 2.2 Å resolution at 100 K from a single crystal of apo *Lp*MAP at a wavelength of 0.97915 Å using an ADSC Quantum 315r CCD detector. A total of 216 images were recorded with 1 s exposure at a crystal-to-detector distance of 250 mm using an oscillation range of 1°. The diffraction data for the *Lp*MAP–tartrate crystal were collected in a similar way as described above, except that the resolution reached 1.9 Å and a total of 200 images were recorded. All data were processed with *HKL-2000* (Otwinowski & Minor, 1997). Table 3 gives a summary of the data-collection statistics. In order to solve the phase problem, the molecular-replacement method was applied. A homologous structure (PDB entry 3it0; Singh *et al.*, 2009) was used as the search model. *Phaser* (McCoy *et al.*, 2007) was used for molecular-replacement calculations.

3. Results and discussion

*Lp*MAP was successfully overexpressed in *E. coli* strain Rosetta and purified to homogeneity. The apoenzyme was crystallized in the substrate-free form. Within one week, Index condition 90 [0.2 M sodium formate, 20%(w/v) polyethylene glycol (PEG) 3350] gave small needle-like crystals. After initial crystal optimization experiments, crystals with rod-like morphology were obtained from a reservoir solution consisting of 0.2 M sodium formate, 17%(w/v) PEG 3350. However, it was difficult to obtain fine monocrystals for diffraction. Therefore, various different additives were tested to increase the diffraction quality of the crystals. Crystals obtained in the presence of low gelling temperature agarose produced diffraction of decent quality. The final crystals of apo *Lp*MAP (Fig. 1*a*) were produced by mixing 1 µl protein solution (30 mg ml^{−1}) with an equal volume of reservoir solution consisting of 0.2 M sodium formate, 17%(w/v) PEG 3350, 0.025%(w/v) low gelling temperature agarose and equilibrating against 200 µl reservoir solution at 285 K.

The best crystals of *Lp*MAP in complex with L-(+)-tartrate (Fig. 1*b*) were yielded by mixing 1.0 µl protein solution (15 mg ml^{−1}) and an equal volume of reservoir solution consisting of 0.1 M Tris pH 8.5, 21%(w/v) PEG 6000, 0.025%(w/v) low gelling temperature agarose and equilibrating against 200 µl reservoir solution at 285 K.

Figure 1
Crystals of apo *Lp*MAP (*a*) and *Lp*MAP–tartrate (*b*).

The diffraction data were collected at 100 K using synchrotron X-rays (Fig. 2). The apo *Lp*MAP crystal belonged to space group $P2_1$, with unit-cell parameters $a = 91.50$, $b = 56.48$, $c = 146.35$ Å, $\alpha = \gamma = 90$, $\beta = 110.01^\circ$. The Matthews coefficient (V_M) of the crystal was 2.3 Å³ Da⁻¹, suggesting that four *Lp*MAP molecules are contained in the crystallographic asymmetric unit, with an estimated solvent content of approximately 46.9%(v/v). From the 69 543 accepted observations to 2.2 Å resolution, the data set contains 31 610 unique reflections and scaled with an overall R_{merge} of 8.2%. The *Lp*MAP–tartrate crystal diffracted to 1.9 Å resolution and belonged to space group $P1$, with unit-cell parameters $a = 55.51$, $b = 73.51$, $c = 98.78$ Å, $\alpha = 78.82$, $\beta = 77.65$, $\gamma = 67.73^\circ$. Assuming the presence of four *Lp*MAP–tartrate molecules in the asymmetric unit, a Matthews coefficient of 2.3 Å³ Da⁻¹ and a solvent content of 46.2%(v/v) were calculated. Crystal parameters, data-collection and data-processing statistics are summarized in Table 1.

The molecular-replacement method was applied to solve the phase problem. A monomer of the *F. tularensis* HAP structure (PDB entry 3it0; Singh *et al.*, 2009), which has 39% identity to the target structure, was used as the molecular-replacement template. In order to generate the best search model, the *Ft*HAP structure was ‘pruned’ using *CHAINS*AW (Stein, 2008). All solvent molecules and heteroatoms [phosphate ion, PEG and 2-(2-methoxyethoxy)ethanol] were stripped from the model. Finally, using the improved model, a single solution obtained using *Phaser* (McCoy *et al.*, 2007) showed a log-likelihood gain (LLG) of 138 and Z-scores of 6.0 for the rotation function (RFZ) and 24.8 for the translation function (TFZ). This initial model was subjected to 20 cycles of rigid-body refinement and ten cycles of restrained refine-

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

	Apo <i>Lp</i> MAP	<i>Lp</i> MAP–tartrate
Diffraction source	SSRF beamline BL17U	
Wavelength (Å)	0.97915	
Temperature (K)	100	
Detector	ADSC Quantum 315r CCD detector	
Crystal-to-detector distance (mm)	250	
Rotation range per image (°)	1	
Total rotation range (°)	216	200
Exposure time per image (s)	1	
Space group	$P2_1$	$P1$
a , b , c (Å)	91.50, 56.48, 146.35	55.51, 73.51, 98.78
α , β , γ (°)	90.00, 110.01, 90.00	78.82, 77.65, 67.73
Resolution range (Å)	50–2.2 (2.24–2.20)	50–1.9 (1.93–1.90)
Molecules in asymmetric unit†	4	4
Solvent content‡ (%)	46.9	46.2
Total No. of reflections	69543	97641
No. of unique reflections	31610	48820
Completeness (%)	96.5 (98.4)	88.9 (85.5)
Multiplicity	2.2	2.0
$\langle I/\sigma(I) \rangle$	11.9 (2.8)	11.6 (2.0)
$R_{\text{merge}}§$ (%)	8.2 (47.9)	7.3 (53.3)
Overall B factor from Wilson plot (Å ²)	31.85	19.36

† The number of molecules in the asymmetric unit was confirmed by molecular replacement using *Phaser* (McCoy *et al.*, 2007). ‡ Solvent content calculated using the Matthews coefficient calculator (Matthews, 1968). The molecular weight was derived from the sequences given in Table 1. § $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl) \rangle$ is the average intensity of multiple observations of symmetry-related reflections.

ment using *REFMAC* (Murshudov *et al.*, 2011), which resulted in R -factor and R_{free} values of 37.1 and 40.7%, respectively.

At present we are completing the refinement of the structure. Structure determination of *Lp*MAP–tartrate is currently in progress, and the presence of a tartrate molecule has been

Figure 2
X-ray diffraction images from crystals of apo *Lp*MAP (a) and *Lp*MAP–tartrate (b). Resolution circles are also shown labelled with the corresponding values.

identified in the electron density. Furthermore, crystallization of *Lp*MAP with bound substrate is advancing.

Acknowledgements

The authors thank Dr Zhongliang Zhu at USTC for generous assistance and Dr Zhaoqing Luo for the *L. pneumophila* genomic DNA. We also thank the staff at SSRF beamline BL17U for assistance with synchrotron data collection. This work was supported by grants from the National Natural Science Foundation of China (31270770), the MOST–TEKES China–Finland International Cooperation Project (2014DFG42290), the National Natural Science Foundation of China (31400641) and the Doctoral Science Foundation of Anhui University (01001905).

References

- Aragon, V., Kurtz, S. & Cianciotto, N. P. (2001). *Infect. Immun.* **69**, 177–185.
- Biertümpfel, C., Basquin, J., Suck, D. & Sauter, C. (2002). *Acta Cryst. D* **58**, 1657–1659.
- Fields, B. S., Benson, R. F. & Besser, R. E. (2002). *Clin. Microbiol. Rev.* **15**, 506–526.
- Jakob, C. G., Lewinski, K., Kuciel, R., Ostrowski, W. & Lebioda, L. (2000). *Prostate*, **42**, 211–218.
- LaCount, M. W., Handy, G. & Lebioda, L. (1998). *J. Biol. Chem.* **273**, 30406–30409.
- Lindqvist, Y., Schneider, G. & Vihko, P. (1993). *J. Biol. Chem.* **268**, 20744–20746.
- 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.
- Mohapatra, N. P., Soni, S., Reilly, T. J., Liu, J., Klose, K. E. & Gunn, J. S. (2008). *Infect. Immun.* **76**, 3690–3699.
- 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.
- Ortlund, E., LaCount, M. W. & Lebioda, L. (2003). *Biochemistry*, **42**, 383–389.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. (2011). *Nature Methods*, **8**, 785–786.
- Rigden, D. J. (2008). *Biochem. J.* **409**, 333–348.
- Schneider, G., Lindqvist, Y. & Vihko, P. (1993). *EMBO J.* **12**, 2609–2615.
- Singh, H., Felts, R. L., Schuermann, J. P., Reilly, T. J. & Tanner, J. J. (2009). *J. Mol. Biol.* **394**, 893–904.
- Stein, N. (2008). *J. Appl. Cryst.* **41**, 641–643.
- Van Etten, R. L., Davidson, R., Stevis, P. E., MacArthur, H. & Moore, D. L. (1991). *J. Biol. Chem.* **266**, 2313–2319.