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

Expression, purification, crystallization and preliminary X-ray analysis of two arginine-biosynthetic enzymes from *Mycobacterium tuberculosis*

Fatemeh Moradian,[a,b] Craig Garen,[a] Leonid Cherney,[a] Maia Cherney[a] and Michael N. G. James[a]*

[a]Protein Structure and Function Group, Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada, and [b]Department of Biochemistry, Faculty of Sciences, Tarbiat Modarres University, PO Box 14115-175, Tehran, Iran

Correspondence e-mail: michael.james@ualberta.ca

Received 17 July 2006
Accepted 28 August 2006

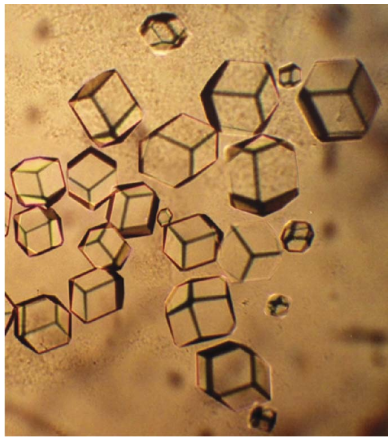
The gene products of two open reading frames from *Mycobacterium tuberculosis* (*Mtb*) have been crystallized using the sitting-drop vapour-diffusion method. Rv1652 encodes a putative *N*-acetyl-γ-glutamyl-phosphate reductase (*Mtb*AGPR), while the Rv1656 gene product is annotated as ornithine carbamoyltransferase (*Mtb*OTC). Both *Mtb*AGPR and *Mtb*OTC were expressed in *Escherichia coli*, purified to homogeneity and crystallized. Native data for each crystal were collected to resolutions of 2.15 and 2.80 Å, respectively. Preliminary X-ray data are presented for both enzymes.

1. Introduction

Tuberculosis (TB) is a serious chronic respiratory disease caused by the pathogen *Mycobacterium tuberculosis* (*Mtb*). It is estimated that eight million people worldwide are added annually to the 2.2 billion *Mtb*-infected individuals (Raviglione, 2003). Despite the existence of a somewhat effective treatment administered globally by the World Health Organization (http://www.who.int/mediacentre/factsheets/fs104/en/), non-compliance with the recommended regimens has given rise to strains resistant to existing antimycobacterial chemotherapies (Espinal, 2003). A synergistic relationship also exists between HIV and TB epidemics: HIV-positive individuals are many times more likely to contract TB and TB has become the leading cause of death among AIDS sufferers (Corbett *et al.*, 2003; Coker, 2004). The result is that approximately 6000 people die everyday from tuberculosis. The *M. tuberculosis* structural genomics consortium was formed in 2000 (http://www.doe-mbi.ucla.edu/TB). The goal of this consortium is to provide a structural basis for the development of new effective therapeutics for tuberculosis (Goulding *et al.*, 2002; Smith & Sacchettini, 2003; Terwilliger *et al.*, 2003).

In this study, we report the cloning, expression, purification and crystallization of two enzymes involved in mycobacterial arginine biosynthesis. The Rv1652 gene encodes *N*-acetyl-γ-glutamyl phosphate reductase (EC 1.2.1.38; *Mtb*AGPR). AGPR catalyzes the third step of eight in the arginine-biosynthetic pathway (Cybis & Davis, 1975) by reversibly converting *N*-acetyl-γ-glutamyl phosphate to *N*-acetyl-γ-glutamate semialdehyde (NAGSA) *via* an NADPH-dependent reductive dephosphorylation. Acting as a NAGSA dehydrogenase (*Mtb*NAGSD), this enzyme catalyzes the $NADP^+$-dependent oxidative phosphorylation of NAGSA (Miura-Ohnuma *et al.*, 2005). There are four reported structures of AGPRs in the Protein Data Bank (PDB; Berman *et al.*, 2000). The amino-acid sequences of each were compared with *Mtb*AGPR using *ClustalW* (Thompson *et al.*, 1994). The two that show the most similarity to our enzyme based on sequence are 1vkn (41% identity) from *Thermotoga maritima* (Joint Center for Structural Genomics, unpublished work) and 2cvo (38% identity) from *Oryza sativa* (Nonaka *et al.*, 2005).

The Rv1656 gene product is annotated as ornithine carbamoyltransferase (*Mtb*OTC; EC 2.1.3.3). This enzyme mediates the sixth step in arginine biosynthesis, specifically the carbamoylation of L-ornithine to produce L-citrulline (Cybis & Davis, 1975; Allewell *et al.*, 1999). Ten OTC structures from five organisms have been deposited in the PDB. The two that are most similar to *Mtb*OTC are 1akm (39% identity) from *Escherichia coli* (Jin *et al.*, 1997) and ovine



OTC (1fb5; De Gregorio *et al.*, 2003), which has 38% sequence identity.

2. Experimental methods

2.1. Cloning, expression and purification

Primers were designed using sequences for directional cloning of inserts into the Gateway cloning system (Invitrogen). They included homologous recombination sites and a region encoding a tobacco etch virus (TEV) protease-cleavage site as published elsewhere (Biswal *et al.*, 2006). The gene-specific sequences used for these two pairs of primers are Rv1652F (5′-ATGCAAAATCGGCAGGTGGCCAATGCG-3′), Rv1652R (5′-TCACGGCGCCACCCCCACAACCGAAAGGCC-3′), Rv1656F (5′-GTGATCAGGCATTTCCTGCGCGACGACG-3′) and Rv1656R (5′-TCATGAGCGCTCCAGCAGCCACACCAGCAGC-3′). *Mtb* strain H37Rv genomic DNA in an ordered bacterial artificial chromosome library from L'Institute Pasteur (Brosch *et al.*, 1998) was used as template to amplify the open reading frames (ORFs) by PCR. Both ORFs were inserted into the cloning vector pDONR-201 and Rv1652 was then transferred into an expression vector encoding an amino-terminal glutathione-*S*-transferase (GST) fusion (pDEST-15) to create the expression vector pGST-1652. Rv1656 was inserted downstream from a region encoding an N-terminal hexahistidine affinity tag in the expression plasmid pDEST-17, yielding pHIS-1656. Both expression plasmids were confirmed by DNA-sequence analysis (DNA core facility, Department of Biochemistry, University of Alberta, Canada).

GST-AGPR was expressed in *E. coli* BL21 (DE3) pLysS cells (Novagen). Transformed cells were incubated at 310 K until the $OD_{600\,nm}$ = 0.8; the medium then was shifted to 295 K and expression was induced by adding isopropyl thiogalactopyranoside (IPTG) to a final concentration of 1 m*M*. After 18 h incubation, the cells were harvested by centrifugation at 2150*g* for 20 min. The bacterial pellet was resuspended in phosphate-buffered saline (PBS) with complete protease inhibitor (Roche), 2 m*M* DTT and 250 µg ml^{-1} lysozyme pH 8.0. Cells were lysed by freeze–thaw and then subjected to ultrasonication. The lysate was cleared by centrifugation (30 min, 20 000*g*) and the supernatant was loaded onto a 5 ml GSTrap cartridge (GE Healthcare). His_6-Rv1656 was expressed in the same manner; however, harvested cells were resuspended in 20 m*M* Na_2HPO_4, 30 m*M* imidazole, 500 m*M* NaCl, 1 m*M* DTT and 250 µg ml^{-1} lysozyme pH 7.4. Following lysis and clarification, the supernatant was loaded onto an Ni^{2+}-charged 5 ml IMAC column (GE Healthcare). GST-*Mtb*AGPR and His_6-*Mtb*OTC were competitively eluted with either a step gradient with 10 m*M* glutathione or a linear gradient of imidazole from 30 to 500 m*M*, respectively. The GST tag was cleaved using recombinant TEV protease (Invitrogen), resulting in a full-length enzyme (352 residues) preceded by a single glycine from the expression construct. The cleaved fusion protein was loaded onto a second GST cartridge and the flowthrough was collected. *Mtb*OTC was not proteolytically cleaved from its affinity tag; therefore, it retained 29 amino acids (MSYYHHHHHHLESTSLYKKAGSENLYFQG) from the fusion construct at the N-terminus of the full-length (307 residues) native enzyme. *Mtb*AGPR and His_6-*Mtb*OTC were dialyzed to change the respective buffer conditions to 20 m*M* Tris–HCl and 500 m*M* NaCl pH 7.4 for size-exclusion chromatography (SEC). The proteins were separately loaded onto a Superdex 16/60 75pg SEC column (GE Healthcare) and each protein's purity was determined by 12% SDS–PAGE analysis (Fig. 1). *Mtb*AGPR and *Mtb*OTC were further confirmed by MALDI–TOF MS peptide identification following in-gel proteolysis of the excised bands. Fractions from each SEC run containing purified *Mtb*AGPR or *Mtb*OTC were dialyzed into 5 m*M* Tris–HCl pH 7.4, 50 m*M* NaCl and then concentrated using an Amicon Ultra 10 kDa cutoff concentrator (Millipore). The final protein concentrations were 7.5 mg ml^{-1} and 12.5 mg ml^{-1} for *Mtb*AGPR and *Mtb*OTC, respectively, as determined using a Bradford protein assay kit (Biorad) with BSA (Pierce) as a standard.

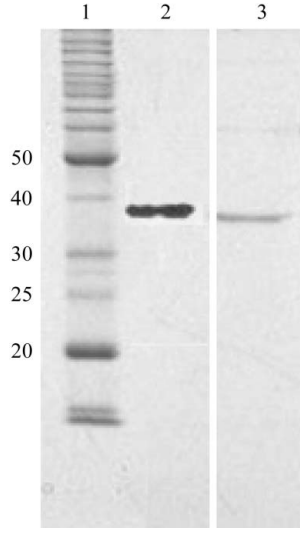


Figure 1
12% SDS–PAGE analysis revealing the expected molecular weights for both *Mtb*AGPR (36 kDa) and *Mtb*OTC (33 kDa). Lane 1, molecular-weight ladder (values shown are in kDa); lane 2, purified His_6-*Mtb*OTC; lane 3, purified *Mtb*AGPR.

2.2. Crystallization

Initial crystallization conditions for *Mtb*AGPR and *Mtb*OTC were found using a Hydra Plus 1 robot (Matrix Technology) to set up trials with the sitting-drop vapour-diffusion method in 96-well format (Intelliplates, Hampton Research). Sparse-matrix screening (Jancarik & Kim, 1988) was performed using Crystal Screens I and II and the Index Screen (Hampton Research). 0.5 µl of each concentrated protein was mixed with an equal volume of reservoir solution and then incubated over 100 µl reservoirs at 273 K for 2–3 d. Optimization of crystallization conditions for *Mtb*AGPR gave a final condition of 0.1 *M* Bis-Tris, 20% PEG MME 5000 pH 6.5. Diffraction-quality *Mtb*OTC crystals were grown in 0.1 *M* MES, 1.6 *M* magnesium sulfate pH 6.5.

2.3. Data collection and processing

Crystals for X-ray data collection were first rinsed in cryoprotectant (25% glycerol in mother liquor) and then flash-frozen in liquid nitrogen. Native data sets for *Mtb*AGPR and *Mtb*OTC crystals were collected on beamline 8.3.1 at the Advanced Light Source at Lawrence Berkeley National Laboratory. Raw data from both were reduced, merged and scaled using the *HKL*-2000 program suite (Otwinowski & Minor, 1997). The crystallographic statistics of the data from the native crystals are summarized in Table 1.

3. Results and discussion

The purifications of both *Mtb*AGPR and *Mtb*OTC were analyzed by 12% SDS–PAGE (Fig. 1), revealing bands of the expected molecular weight for both *Mtb*AGPR and *Mtb*OTC (36.3 and 33.0 kDa,

Table 1
Crystal parameters and data-collection statistics for native data sets for *Mtb*AGPR and *Mtb*OTC.

Values in parentheses refer to the highest resolution shell.

Crystal	*Mtb*AGPR (native)	*Mtb*OTC (native)
Space group	$P2_12_12_1$	$P4_12_12$
Unit-cell parameters		
a (Å)	89.70	136.14
b (Å)	104.03	136.14
c (Å)	138.44	141.66
$\alpha = \beta = \gamma$ (°)	90.0	90.0
Z†	16	24
Mosaicity (°)	0.80	0.32
Data collection		
Temperature (K)	100	100
Detector	ADSC Q210	ADSC Q210
Wavelength (Å)	1.11587	0.97946
Resolution (Å)	2.15 (2.23–2.15)	2.80 (2.90–2.80)
Unique reflections	69868 (6426)	32903 (3231)
Multiplicity	3.5 (2.3)	7.7 (7.8)
$I/\sigma(I)$	15.5 (1.7)	17.8 (3.3)
Completeness (%)	98.9 (92.3)	99.2 (99.8)
R_{sym} (%)‡	7.3 (46.0)	10.6 (71.6)

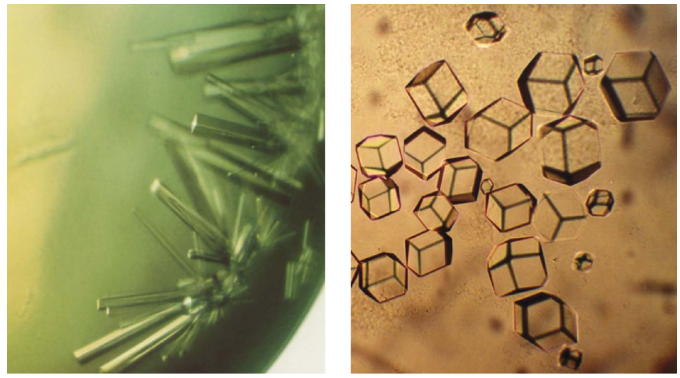
† No. of molecules per unit cell. ‡ $R_{sym} = \sum_{\mathbf{h}} \sum_i |I_{\mathbf{h}i} - \langle I_{\mathbf{h}} \rangle| / \sum_{\mathbf{h}} \sum_i \langle I_{\mathbf{h}} \rangle$, where $I_{\mathbf{h}i}$ is the ith observation of refection $\mathbf{h}$ and $\langle I_{\mathbf{h}} \rangle$ is the weighted average intensity for all observations i of reflection $\mathbf{h}$.

respectively). Gel slices were submitted for in-gel proteolysis followed by MALDI–TOF MS identification of the resulting peptides and these confirmed the presence of the Rv1652 and Rv1656 gene products as expected.

The orthorhombic *Mtb*AGPR crystals (Fig. 2*a*) belong to space group $P2_12_12_1$, with unit-cell parameters a = 89.7, b = 104.0, c = 138.4 Å. The Matthews coefficient (Matthews, 1968) of 2.22 Å^3 Da^{-1} and solvent content of 44.69% assume the presence of four subunits of *Mtb*AGPR per asymmetric unit. A tetramer is the active biological unit in related structures from other organisms (PDB codes 1vkn and 2cvo) determined to date. The program *MOLREP* (Vagin & Teplyakov, 1997) from the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994) was used to generate a solution for *Mtb*AGPR using the atomic coordinates of PDB entry 1vkn (*T. maritima*, 38% identity) as a search model. This solution confirmed the presence of four subunits in the asymmetric unit.

The tetragonal *Mtb*OTC crystals (Fig. 2*b*; space group $P4_12_12$) have unit-cell parameters $a = b$ = 136.1, c = 141.7 Å. The Matthews coefficient for three molecules in the asymmetric unit is 3.31 Å^3 Da^{-1}, corresponding to a solvent content of 62.87%. Other OTC crystal structures determined to date (PDB codes 1akm and 1fb5) have a trimer as the active biological unit. The *Mtb*OTC asymmetric unit contains density for three subunits as determined by molecular replacement using PDB entry 1fb5 as a search model (ovine OTC, 38% identity). Crystallographic refinement and molecular-structure analysis for both of these enzymes will be published elsewhere in the future.

X-ray diffraction data were collected at beamline 8.3.1 of the Advanced Light Source (ALS) at Lawrence Berkeley Laboratory under an agreement with the Alberta Synchrotron Institute (ASI). The ALS is operated by the Department of Energy and supported by the National Institutes of Health. Beamline 8.3.1 was funded by the National Science Foundation, the University of California and Henry Wheeler. The ASI synchrotron-access program is supported by grants from the Alberta Science and Research Authority (ASRA) and the Alberta Heritage Foundation for Medical Research (AHFMR). Many thanks to the members of the James laboratory for helpful discussions and for critical reading of this manuscript and to Perry D'Obrenan for aid in the preparation of figures. Research in the laboratory of MNGJ is supported by the Canadian Institutes of Health Research (CIHR) and the Alberta Heritage Foundation for Medical Research (AHFMR); MNGJ is the holder of a Canada Research Chair.



(*a*) (*b*)

Figure 2
Diffraction-quality crystals of each enzyme. (*a*) *Mtb*AGPR of approximate maximum dimensions of 200 × 20 × 20 µm grown in 0.1 *M* bis-tris, 20% PEG MME 5000 pH 6.5. (*b*) *Mtb*OTC crystals grown in 0.1 *M* MES, 1.6 *M* magnesium sulfate pH 6.5. The approximate average crystal dimensions are 50 × 50 × 50 µm.

References

Allewell, N. M., Shi, D., Morizono, H. & Tuchman, M. (1999). *Acc. Chem. Res.* **32**, 885–894.
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.
Biswal, B. K., Garen, G., Cherney, M. M., Garen, C. & James, M. N. G. (2006). *Acta Cryst.* F**62**, 136–138.
Brosch, R., Gordon, S. V., Billaut, A., Garnier, T., Eiglmeier, K., Sorvito, C., Barrell, B. G. & Cole, S. T. (1998). *Infect. Immun.* **66**, 2221–2229.
Coker, R. J. (2004). *Trop. Med. Int. Health*, **9**, 25–40.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Corbett, E. L., Watt, C. J., Walker, N., Maher, D., Williams, B. G., Raviglione, M. C. & Dye, C. (2003). *Arch. Intern. Med.* **163**, 1009–1021.
Cybis, J. & Davis, R. H. (1975). *J. Bacteriol.* **123**, 196–202.
De Gregorio, A., Battistutta, R., Arena, N., Panzalorto, M., Francescato, P., Valentini, G., Bruno, G. & Zanotti, G. (2003). *Org. Biomol. Chem.* **1**, 3178–3185.
Espinal, M. A. (2003). *Tuberculosis*, **83**, 44–51.
Goulding, C. W. *et al.* (2002). *Curr. Drug. Targets Infect. Disord.* **2**, 121–141.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jin, L., Seaton, B. A. & Head, J. F. (1997). *Nature Struct. Biol.* **4**, 622–625.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Miura-Ohnuma, J., Nonaka, T., Katoh, S., Murata, K., Kita, A., Miki, K. & Katoh, E. (2005). *Acta Cryst.* F**61**, 1058–1061.
Nonaka, T., Kita, A., Miura-Ohnuma, J., Katoh, E., Inagaki, N., Yamazaki, T. & Miki, K. (2005). *Proteins*, **61**, 1137–1140.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Raviglione, M. C. (2003). *Tuberculosis*, **83**, 4–14.
Smith, C. V. & Sacchettini, J. C. (2003). *Curr. Opin. Struct. Biol.* **13**, 658–664.
Terwilliger, T. C. *et al.* (2003). *Tuberculosis*, **83**, 223–249.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.