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

Expression, purification, crystallization and preliminary X-ray analysis of Rv3117, a probable thiosulfate sulfurtransferase (CysA3) from *Mycobacterium tuberculosis*

Sarah J. Witholt, Ramasamy Sankaranarayanan, Craig R. Garen, Maia M. Cherney, Leonid T. Cherney and Michael N. G. James*

Protein Structure and Function Group, Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada

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

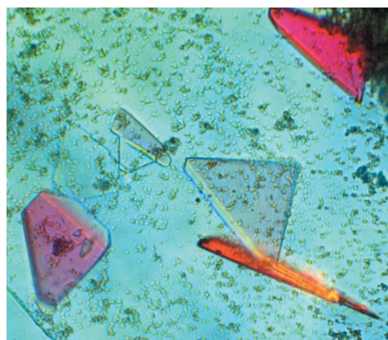
Received 14 March 2008
Accepted 13 May 2008

The gene product of open reading frame Rv3117 from *Mycobacterium tuberculosis* (*Mtb*) strain H37Rv is annotated as encoding a probable rhodanese-like thiosulfate sulfurtransferase (*Mtb*CysA3). *Mtb*CysA3 was expressed and purified and then crystallized using the sitting-drop vapour-diffusion method. X-ray diffraction data were collected and processed to a maximum resolution of 2.5 Å. The crystals belong to the monoclinic space group $P2_1$, with unit-cell parameters a = 38.86, b = 91.43, c = 83.57 Å, β = 96.6°. Preliminary diffraction data shows that two molecules are present in the asymmetric unit; this corresponds to a $V_{\rm M}$ of 2.4 $Å^3$ Da^{-1}.

1. Introduction

Tuberculosis (TB) is a common and serious contagious disease caused by *Mycobacterium tuberculosis* (*Mtb*). One third of the world's population is infected with the TB bacillus, leading to an estimated 4400 human deaths per day (World Health Organization, 2007). If not treated, each person with active TB will infect 10–15 further people each year (World Health Organization, 2007). Numerous drug-resistant strains of TB have emerged owing to inconsistent or partial drug treatments and limited drug availability. The TB Structural Genomics Consortium was formed in 2000 with the goal of developing a basis for new TB drug development by determining the three-dimensional structures of *Mtb* proteins (http://www.doe-mbi.ucla.edu/TB/mission.php) with the view of eliminating TB as a public health problem.

As members of the *Mtb* structural genomics consortium, we have conducted molecular cloning and preliminary X-ray analysis on a rhodanese-like thiosulfate sulfurtranferase protein (*Mtb*CysA3). Classical rhodanese proteins (EC 2.8.1.1) catalyze the transfer of a sulfane sulfur from a donor substrate to the enzyme active site and then in turn to a thiophilic acceptor (Westley, 1973; Nandi & Westley, 1998). It was initially believed that the primary function of this enzyme was to convert cyanide to the less toxic thiocyanate (Lang, 1933). In fact, the best studied reaction *in vitro* has thiosulfate as the donor to form a sulfur-substituted rhodanese-enzyme intermediate (and sulfite), followed by transfer of the sulfane sulfur to cyanide, which acts as the acceptor substrate. The products of this displacement mechanism are thiocyanate and sulfur-free rhodanese. It has been proposed that dihydrolipoate can act as a possible sulfur acceptor that yields lipoate and inorganic sulfite as final products (Villarejo & Westley, 1963; Volini & Westley, 1966). Other studies have shown that by using reduced lipoate, rhodanese may be used to reconstitute redox centres to aid in the formation of the characteristic prosthetic group of iron–sulfur proteins (Pagani *et al.*, 1984; Bonomi *et al.*, 1985). 24 rhodanese-like structures have been deposited in the Protein Data Bank (PDB; Berman *et al.*, 2000). These include nine single rhodanese homology domain structures and 13 two-domain structures. One three-domain mercaptopyruvate sulfur transferase



© 2008 International Union of Crystallography
All rights reserved

structure that contains similar N-terminal and central domains to thiosulfate sulfurtransferase rhodanese has also been solved. Four of the two-domain structures were used to compare sequence alignment with *Mtb*CysA3 (Fig. 1). The two sequences that show the greatest similarities are 1uar (52% identity) from *Thermus thermophilus* (RIKEN Structural Genomics/Proteomics Initiative, unpublished work) and 1urh (26% identity) from *Escherichia coli* (Spallarossa *et al.*, 2004). Sequence alignments were performed using *ClustalW* (Thompson *et al.*, 1994).

2. Experimental methods

2.1. Cloning, expression and purification

The entire genome of the H37Rv strain was cloned into a bacterial artificial chromosome (BAC) library at L'Institut Pasteur (Brosch *et al.*, 1998; Gordon *et al.*, 1999). Polymerase chain reaction (PCR) was used to amplify the Rv3117 protein, using the BAC library as the template. PCR primers were designed for directional cloning of inserts into the Gateway cloning system (Invitrogen). The primer sequences were Rv3117F (5′-GGGACAAGTTTGTACAAAAAAGCAGGCTCC*GGTCTGGTTCCGCGTGGT*ATGGCACGCTGCGATGTCCTGGTCT-3′) and Rv3117R (5′-GGGACCACTTTGTACAAGAAAGCTGGGTCTCAGCTTCCCAACTCGATCGGGGCG-3′). The open reading frame encoding residues 1–277 of Rv3117 was cloned into the Gateway entry vector pDONR 221 (Invitrogen) and then into the expression plasmid containing an amino-terminal hexahistadine (His_6) fusion tag (pDEST-17, Invitrogen). The presence of the gene insert was confirmed by DNA-sequence analysis (DNA core facility, Department of Biochemistry, University of Alberta, Canada).

His_6-Rv3117 was expressed in *E. coli* BL21 (DE3) cells (Novagen). Incubation of the transformed cells at 310 K was continued until the $OD_{600\,nm}$ reached 0.5. Subsequently, the temperature was shifted to 295 K and protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (Fisher Scientific) to a final concentration of 0.5 m*M*. After overnight incubation, the cells were harvested by centrifugation for 15 min at 9380*g*. Bacterial pellets were resuspended in 20 m*M* NaH_2PO_4 pH 7.4, 300 m*M* NaCl, 10 m*M* imidazole, 2 m*M* β-mercaptoethanol containing Complete protease inhibitor (Roche), 1 m*M* phenylmethylsulfonyl fluoride (Bioshop) and 10 μg ml^{-1} hen egg-white lysozyme (Sigma). For purification, the cells were lysed by freeze–thawing and then subjected to ultrasonication in the resuspension buffer. The lysate was cleared by

Figure 1
Sequence comparison of selected sulfurtranferase proteins. Abbreviations and accession numbers are as follows: BovRHD, bovine liver rhodanese (gi:135823); EcSseA, *E. coli* strain K12 (gi:401186); TtRHD, *T. thermophilus* strain HB8 (gi:81600441); Rv3117, *M. tuberculosis* strain H37Rv (gi:15610254); AvRHD, *A. vinelandii* (gi:1729961). *ClustalW* was used to perform sequence alignment (Thompson *et al.*, 1994) and the figure was generated using the program *ESPript* (Gouet *et al.*, 1999).

Table 1
Crystal parameters and data-collection statistics for Rv3117.

Values in parentheses are for the highest resolution shell.

Crystal parameters	
Space group	$P2_1$
Unit-cell parameters	
a (Å)	38.86
b (Å)	91.43
c (Å)	83.57
α (°)	90.00
β (°)	96.6
γ (°)	90.00
No. of molecules per unit cell	4
Data collection	
Temperature (K)	100
Detector	ADSC Q315
Wavelength (Å)	1.214591
Resolution (Å)	50.0–2.5 (2.59–2.50)
Unique reflections	20078 (2006)
Multiplicity	3.6 (3.4)
Mosaicity (°)	1.3–1.7
$I/\sigma(I)$	11.8 (2.1)
Completeness (%)	98.8 (98.0)
R_{merge}†	0.104 (0.546)
Wilson B factor (Å^2)	53.7

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith observation of reflection hkl and $\langle I(hkl)\rangle$ is the weighted average intensity for all observations i of reflection hkl.

centrifugation (30 min, 20 000g) and the supernatant was then passed through a 0.45 µM syringe filter (Millipore). The cleared supernatant was loaded onto a 5 ml HisTrap FF column (GE Healthcare) pre-equilibrated with 20 mM NaH_2PO_4 pH 7.4, 300 mM NaCl, 10 mM imidazole and 1 mM β-mercaptoethanol. The His_6-Rv3117 fusion protein was eluted with a linear gradient of imidazole from 10 to 300 mM. The His_6 tag and the N-terminal recombination site were removed by proteolytic cleavage using thrombin (Amersham Biosciences), leaving an additional glycine residue at the N-terminus. The site recognized by thrombin is encoded in the forward primer (italicized in the primer sequence given). After dialysis against 20 mM NaH_2PO_4 pH 7.4, 300 mM NaCl, 10 mM imidazole and 1 mM β-mercaptoethanol, the cleaved protein mixture was once again loaded onto a HisTrap column to remove the His_6 tag. The flow-through fractions containing the digested *Mtb*CysA3 were concentrated to 8 mg ml^{-1} using an Amicon Ultra (10 kDa cutoff; Millipore). The protein was then dialyzed overnight against 5 mM Tris–HCl pH 7.4, 100 mM NaCl and 1 mM dithiothreitol (Fisher Scientific). The whole process of purification was performed at 277 K and the results of each step were monitored using 16% SDS–PAGE.

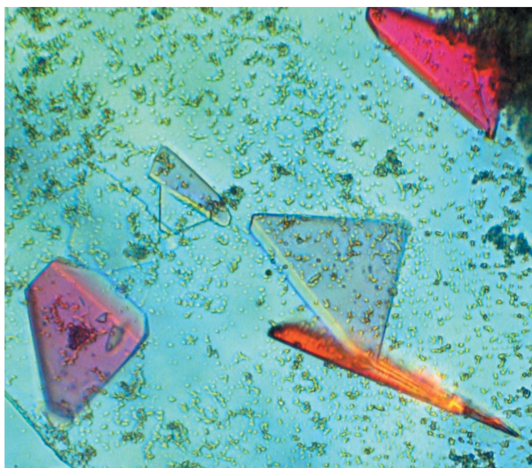


Figure 2
Preliminary *Mtb*CysA3 crystals. Diffraction-quality crystals grew to approximate dimensions of 200 × 100 × 30 µm in 25% PEG 3350, 0.1 M Tris–HCl pH 8.5, 0.2 M $MgCl_2$.

2.2. Crystallization

Crystallization of native full-length *Mtb*CysA3 was performed by screening with various crystallization conditions using the sitting-drop vapour-diffusion method in 96-well Intelliplates (Hampton Research). Index Screen and Crystal Screens I and II (Hampton Research) were used by mixing equal volumes (0.5 µl) of concentrated protein and precipitating solutions. After 3 d, preliminary crystals were obtained from a variety of conditions. After optimization of the best screening condition, X-ray diffraction-quality crystals were grown in hanging drops in 24-well VDX plates (Hampton Research) containing 0.5 µl protein solution at 8 mg ml^{-1} and 0.5 µl precipitating solution and the drops were equilibrated against 1 ml reservoir solution (25% PEG 3350, 0.1 M Tris–HCl pH 8.5, 0.2 M $MgCl_2$). The average dimensions of the *Mtb*CysA3 crystals were 200 × 100 × 30 µm (Fig. 2).

2.3. Data collection

Crystals for synchrotron data collection were first rinsed in cryoprotectant (25% glycerol in mother liquor) and then flash-cooled by immersion in liquid nitrogen. Native data sets were collected on beamline 8.3.1 at the Advanced Light Source (ALS) at the Lawrence Berkeley National Laboratory, revealing a diffraction pattern to 2.5 Å resolution (Fig. 3). The *HKL*-2000 program suite (Otwinowski & Minor, 1997) was used to reduce, integrate and scale the collected data. Crystallographic statistics of the native data are summarized in Table 1.

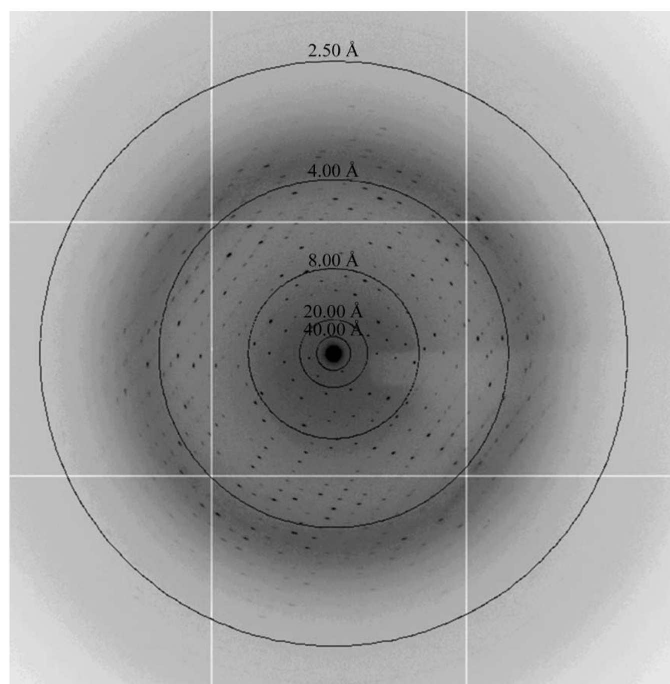


Figure 3
X-ray diffraction-pattern frame collected from *Mtb*CysA3 crystals. The pattern displays a maximum resolution of 2.5 Å and space group $P2_1$.

3. Results

Native data sets were collected from *Mtb*CysA3 crystals. Based upon the expected molecular weight of Rv3117 (31 014.2 Da), the Matthews coefficient V_M (Matthews, 1968) is calculated to be 2.4 $Å^3$ Da^{-1}, with an estimated solvent content of 48.3%. These results suggest the presence of two molecules of Rv3117 per asymmetric unit. In related structures, it has been found that this enzyme exists in a dimeric form when biologically active (PDB codes 1dp2 and 1h4k). The program *Phaser* (McCoy *et al.*, 2005) from the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994) confirmed the presence of a dimer; *Phaser* was used to determine the structure solution by molecular replacement for *Mtb*CysA3 using the coordinates of PDB entry 1uar (*T. thermophilus*, 52% identity). Crystallographic refinement and molecular-structure analysis will be published in a future communication.

X-ray diffraction data were collected on beamline 8.3.1 at the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory under agreements with the Alberta Synchrotron Institute (ASI). The ALS is supported by the National Institutes of Health and operated by the Department of Energy. 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). Research in the laboratory of MNGJ is supported by Alberta Heritage Foundation for Medical Research (AHFMR); MNGJ is the holder of a Canada Research Chair in Protein Structure and Function.

References

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.
Bonomi, F., Pagani, S. & Kurtz, D. M. (1985). *Eur. J. Biochem.* **148**, 67–73.
Brosch, R., Gordon, S. V., Billault, A., Garnier, T., Eiglmeier, K., Soravito, C., Barrell, B. G. & Cole, S. T. (1998). *Infect. Immun.* **66**, 2221–2229.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Gordon, S. V., Brosch, R., Billault, A., Garnier, T., Eiglmeier, K. & Cole, S. T. (1999). *Mol. Microbiol.* **32**, 643–655.
Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.
Lang, K. (1933). *Biochem. Z.* **259**, 243–256.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Nandi, D. L. & Westley, J. (1998). *Int. J. Biochem. Cell Biol.* **30**, 973–977.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pagani, S., Bonomi, F. & Cerletti, P. (1984). *Eur. J. Biochem.* **142**, 361–366.
Spallarossa, A., Forlani, F., Carpen, A., Armirotti, A., Pagani, S., Bolognesi, M. & Bordo, D. (2004). *J. Mol. Biol.* **335**, 583–595.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Villarejo, M. & Westley, J. (1963). *J. Biol. Chem.* **238**, 4016–4020.
Volini, M. & Westley, J. (1966). *J. Biol. Chem.* **241**, 5168–5176.
Westley, J. (1973). *Adv. Enzymol. Relat. Areas Mol. Biol.* **39**, 327–368.
World Health Organization (2007). *The World Health Organization Global Tuberculosis Program.* http://www.who.int/topics/tuberculosis/en/.