Adam E. Shahine,^{a,b} Phooi Y.
Chan,^{b,c} Dene Littler,^{a,b}
Julian Vivian,^{a,b} Rajini
Brammananth,^{b,c} Paul K.
Crellin,^{b,c} Ross L. Coppel,^{b,c}
Jamie Rossjohn^{a,b} and Travis
Beddoe^{a,b*}

^aDepartment of Biochemistry and Molecular
Biology, Monash University, Clayton, Victoria,
Australia, ^bAustralian Research Council (ARC)
Centre of Excellence in Structural and Functional
Microbial Genomics, Monash University,
Clayton, Victoria, Australia, and ^cDepartment of
Microbiology, Monash University, Clayton,
Victoria, Australia

Correspondence e-mail:
travis.beddoe@monash.edu

Received 22 March 2013
Accepted 16 April 2013

Cloning, expression, purification and preliminary X-ray diffraction studies of a mycobacterial protein implicated in bacterial survival in macrophages

Mycobacterium species have developed numerous strategies to avoid the antimycobacterial actions of macrophages, enabling them to survive within the generally inhospitable environment of the cell. The recently identified MSMEG_5817 protein from *M. smegmatis* is highly conserved in *Mycobacterium* spp. and is required for bacterial survival in macrophages. Here, the cloning, expression, purification and crystallization of MSMEG_5817 is reported. Crystals of MSMEG_5817 were grown in 1.42 M Li₂SO₄, 0.1 M Tris–HCl pH 7.7, 0.1 M sodium citrate tribasic dihydrate. Native and multiple-wavelength anomalous dispersion (MAD) data sets have been collected and structure determination is in progress.

1. Introduction

Macrophages are key components of host defence pathways that play a role in the control and eradication of foreign microbes, including pathogenic and nonpathogenic mycobacteria. Pathogenic mycobacterial species have developed numerous strategies to avoid the antimycobacterial actions of macrophages and to survive within the environment of the cell, frequently resulting in disease (Chandra *et al.*, 2011). Upon infection, the invading mycobacteria are internalized by host macrophages and taken into phagosomes, where they are able to survive for long periods of time. Mycobacteria create a hospitable environment within phagosomes through the alteration of host signalling pathways that mediate correct vesicular membrane trafficking and the formation of the phagolysosome. The resultant phagosomes fail to mature and do not fuse with late endosomes, and the lysosomes fail to recruit lysosomal hydrolases that are required for the degradation of internalized microbes (Armstrong & Hart, 1971; Fratti *et al.*, 2003). Identification of the mechanisms utilized by invading mycobacteria to disrupt host-macrophage antibacterial activities is paramount for the understanding of mycobacterial virulence and disease. While *Mycobacterium smegmatis* is generally considered to be nonpathogenic, it does have a limited capacity to survive and multiply within macrophages and to delay phagosomal acidification, making it a suitable model system to study intracellular mycobacterial survival (Anes *et al.*, 2006). In previous studies, the members of a random Tn611 transposon mutant library of *M. smegmatis* were screened for a number of atypical phenotypes indicative of cell-wall defects (Billman-Jacobe *et al.*, 1999; Kovacevic *et al.*, 2006; Patterson *et al.*, 2000). One mutant, designated Myco132, exhibited altered colony morphology and an increased ability to absorb dyes from growth media (Pelosi *et al.*, 2012). The gene disrupted in Myco132 was identified by sequencing as MSMEG_5817 (Pelosi *et al.*, 2012). MSMEG_5817 encodes a conserved hypothetical protein of 128 amino-acid residues that has orthologues in a number of pathogenic mycobacteria. Inactivation of the MSMEG_5817 gene led to accelerated cell death of *M. smegmatis* in J774A.1 macrophages within the first 8 h (Pelosi *et al.*, 2012). Complementation of the mutant with an active copy of the gene restored survival to near-wild-type levels. In addition to decreased survival in macrophages, there was also a significant increase in NF- κ B activation in macrophages infected with the MSMEG_5817 mutant compared with those

infected by wild-type (WT) *M. smegmatis* (Pelosi *et al.*, 2012). Since NF- κ B activation plays key roles in regulating host immune responses through cell activation and the induction of cytokine production, MSMEG_5817 may be involved in suppressing NF- κ B activation to promote survival. Although *MSMEG_5817* is implicated in bacterial survival within macrophages, it is not important for growth, with cultures grown *in vitro* in rich medium showing no differences in growth rates between Myco132 and WT *M. smegmatis* (Pelosi *et al.*, 2012). In this paper, we report the expression, purification and crystallization of this protein. The structural determination of MSMEG_5817 will provide valuable insight into the mechanistic pathways that mediate the function of this protein.

2. Materials and methods

2.1. Cloning

The *MSMEG_5817* gene was PCR-amplified from *M. smegmatis* genomic DNA using HotStar HiFidelity polymerase (Qiagen) and specific sense (5'-GGAATTGCATATGGCCAGCCGCCGTAGTG-C-3') and antisense (5'-GGAATTCCTACGGGTCGATACGCACC-ACGGG-3') primers. The highlighted sequences in the primers represent *Nde*I and *Eco*RI restriction sites, respectively. The 567 bp PCR-amplified product was purified using the MO BIO kit (MO BIO Laboratories) and the fragment was digested with *Nde*I and *Eco*RI (NEB) and cloned into pET-28c(+) vector (Novagen), giving a protein with a thrombin-cleavable N-terminal hexahistidine tag. The resultant N-terminal amino-acid sequence of the recombinant MSMEG_5817 is MGSSHHHHHHSSGLVPRGSH**MASRR**, where the sequence in bold is the MSMEG_5817 protein starting at residue 1. The clone obtained was confirmed by DNA sequencing.

2.2. Overexpression of recombinant protein

Recombinant protein was overexpressed in *Escherichia coli* host strain BL21 (DE3). The clone was transformed into competent BL21 (DE3) cells and transformants were selected on LB agar plates containing 34 μ g ml⁻¹ kanamycin. Pre-inoculum cultures were grown overnight in LB-kanamycin medium from a single colony. The starter culture was used at a dilution of 1:100 to inoculate 800 ml fresh LB medium containing 34 μ g ml⁻¹ kanamycin and was grown at 310 K until the OD₆₀₀ reached 0.6. Expression of the protein was induced by the addition of 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) and the cells were allowed to grow for an additional 4 h at 310 K. The cells were harvested by centrifugation and stored overnight at 193 K. Expression of the selenomethionine derivative of MSMEG_5817 was performed in *E. coli* host strain B834 (DE3). Pre-inoculum cultures grown overnight in LB-kanamycin medium from a single colony were then diluted in 6 l fresh LB medium containing 34 μ g mg⁻¹ kanamycin and grown at 310 K until the OD₆₀₀ reached 0.6. The cells were harvested by centrifugation and washed twice with 400 ml M9 minimal salt medium consisting of 47 mM NH₄Cl, 55 mM KH₂PO₄, 120 mM Na₂HPO₄, 43 mM NaCl. The washed cells were resuspended in 400 ml M9 minimal medium containing trace metals, 0.5% glucose, 5.1 μ M biotin, 4.2 μ M thiamine, 2.5 mM MgCl₂, 0.75 mM CaCl₂, 34 μ g ml⁻¹ kanamycin and 255 μ M selenomethionine. The cultures were recovered at 310 K for 1 h before induction with 0.5 mM IPTG. The cells were allowed to grow for an additional 16 h. The cells were harvested by centrifugation and stored overnight at 193 K.

2.3. Purification of recombinant protein

The native protein and selenomethionine-derivatized protein were purified using the same methods. Cells were thawed and resuspended in 100 ml cold lysis buffer consisting of 50 mM sodium phosphate pH 8.0, 300 mM NaCl. Cells were lysed using a French Press on ice in the presence of 0.2 μ M phenylmethanesulfonyl fluoride (PMSF). Cell debris was removed from the lysate by centrifugation. All subsequent purification steps were performed at 293 K. The lysate was added to a column packed with TALON metal-affinity resin (Roche) pre-equilibrated with lysis buffer. The protein was allowed to bind before being washed with one column volume of 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 20 mM imidazole. The protein was eluted using five column volumes of 50 mM sodium phosphate pH 6.0, 300 mM NaCl, 250 mM imidazole. The eluted protein was concentrated to 500 μ l by low-speed centrifugation using a Vivaspin 2 3000 Da molecular-weight cutoff concentrator (Sartorius Stedim Biotech). The protein was further purified by size-exclusion chromatography using an ÄKTA basic fast protein liquid-chromatography (FPLC) system with a Sepharose 75 10/30 column (GE Healthcare Life Sciences). Chromatography was carried out in 20 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) pH 6.0 buffer containing 100 mM NaCl. Purified protein was concentrated to 20 mg ml⁻¹ and stored at 277 K for use in crystallization trials. The concentration of the MSMEG_5817 protein was determined spectrophotometrically (Nanodrop 1000, Thermo Scientific) at 280 nm and was calculated using an extinction coefficient of 16 500 M⁻¹ cm⁻¹. The molecular weight, purity and identity of the protein were confirmed by SDS-PAGE (Fig. 1) and Western blotting with anti-hexahistidine antibody (R&D Systems).

2.4. Crystallization

The initial crystallization experiments involved screening 480 conditions from commercially available kits from Hampton Research (Crystal Screen, Crystal Screen 2, Index, PEG/Ion and SaltRx) and Emerald BioSystems (Wizard I and II) using a CrystalMation integrated robotic workstation (Rigaku) and the sitting-drop vapour-diffusion technique. Initial small spheroid crystals were obtained in 1.50 M Li₂SO₄, 0.1 M Tris-HCl pH 8.5 (SaltRx screen, Hampton

Figure 1
18% SDS-PAGE of purified native MSMEG_5817 (lane 1) and purified selenomethionine-derivatized MSMEG_5817 (lane 2). Molecular weights are indicated in kDa on the left.

Table 1
Crystal and data-collection statistics.
Values in parentheses are for the outer shell.

	Native MSMEG_5817	Selenomethionine-derivatized MSMEG_5817		
		Inflection (λ_1)	Peak (λ_2)	High-energy remote (λ_3)
Diffraction source	MX2, Australian Synchrotron	MX2, Australian Synchrotron	MX2, Australian Synchrotron	MX2, Australian Synchrotron
Detector	ADSC Quantum 315r CCD	ADSC Quantum 315r CCD	ADSC Quantum 315r CCD	ADSC Quantum 315r CCD
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Temperature (K)	100	100	100	100
Unit-cell parameters ($\text{\AA}$, $^\circ$)	$a = 50.5$, $b = 50.5$, $c = 205.5$, $\alpha = \beta = \gamma = 90.0$	$a = 49.9$, $b = 49.9$, $c = 205.0$, $\alpha = \beta = \gamma = 90.0$	$a = 49.9$, $b = 49.9$, $c = 205.0$, $\alpha = \beta = \gamma = 90.0$	$a = 49.9$, $b = 49.9$, $c = 205.0$, $\alpha = \beta = \gamma = 90.0$
Resolution range ($\text{\AA}$)	50.05–2.40 (2.67–2.40)	40.3–2.40 (2.53–2.40)	41.06–2.47 (2.60–2.47)	41.4–2.50 (2.64–2.50)
Wavelength ($\text{\AA}$)	0.95369	0.97927	0.97949	0.95369
Total No. of reflections	20386	42924	38677	37473
No. of unique reflections	1516	1549	1429	1377
$R_{\text{merge}}^\dagger$ (%)	9.9 (56.4)	9.4 (56.1)	10.5 (68.9)	10.3 (51.1)
Mean $I/\sigma(I)^\ddagger$	15.9 (4.6)	28.7 (7.5)	24.0 (6.0)	23.7 (7.7)
Completeness (%)	98.7 (98.2)	100 (100)	100 (100)	100 (100)
Anomalous redundancy		15.1 (15.1)	14.7 (14.8)	14.9 (14.9)

$^\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the scaled intensity of the i th measurement of reflection hkl and $\langle I(hkl) \rangle$ is the mean intensity for that reflection. $^\ddagger I$ is the integrated intensity and $\sigma(I)$ is the estimated standard deviation of that intensity.

Research). Optimization of the native MSMEG_5817 crystallization conditions was performed by fine-tuning the protein and precipitant concentrations in 24-well Linbro plates (Hampton Research) using a hanging drop consisting of 1 μl protein solution and 1 μl precipitant solution and a 500 μl reservoir volume. Additional optimization of the crystallization conditions was performed using Additive Screen (Hampton Research), in which drops consisting of 0.9 μl protein solution, 0.9 μl precipitant solution and 0.2 μl additive solution were equilibrated against a 500 μl reservoir volume consisting of 1.42 M Li_2SO_4 , 0.1 M Tris–HCl pH 7.7. The best crystals appeared after 3 d of equilibration against the crystallization condition and grew to full size in 5 d using 0.1 M sodium citrate tribasic dihydrate as an additive (Fig. 2a). Crystals of selenomethionine-derivatized MSMEG_5817

could only be obtained using the cross-seeding technique (Fig. 2b). Cross-seeding was performed using the streak-seeding method. Hanging-drop experiments were performed with drops consisting of 0.9 μl protein solution, 0.9 μl precipitant solution and 0.2 μl 0.1 M sodium citrate tribasic dihydrate equilibrated against a 500 μl reservoir volume. The drops were allowed to equilibrate for 5 d. A cat whisker was used to transfer seeds of the native MSMEG_5817 crystal into clear equilibrated drops. The best selenomethionine-derivatized MSMEG_5817 crystals grew after 24 h using 1.30 M Li_2SO_4 , 0.1 M Tris–HCl pH 7.7, 0.1 M sodium citrate tribasic dihydrate.

2.5. X-ray data collection

For X-ray diffraction data collection, crystals of the native and the selenomethionine-derivative protein were transferred to a CryoLoop and soaked in a cryoprotectant consisting of 1.5 M Li_2SO_4 , 0.1 M Tris–HCl pH 7.7, 15% glycerol before cooling to 100 K in a stream of nitrogen gas. A complete data set was collected from a single crystal on the MX2 beamline at the Australian Synchrotron using an ADSC Quantum 315r CCD detector. For the native crystal, a total of 360 frames of 0.5 $^\circ$ oscillation were recorded with an exposure time of 1 s per frame. Using a single selenomethionine-derivative crystal, a fluorescence scan selecting the L_{III} edge of selenium was conducted to determine the inflection, peak and high-energy remote levels. Three complete data sets were collected from a single crystal at the required wavelengths; a total of 360 images at 0.5 $^\circ$ oscillation were recorded with an exposure time of 1 s per frame (Fig. 3). The data were processed using MOSFLM (Leslie & Powell, 2007) and various programs from the CCP4 suite (Winn, 2011). The final statistics of data collection and processing are summarized in Table 1.

3. Results and discussion

The *M. smegmatis* gene MSMEG_5817 encoding a 128-residue protein was successfully cloned and the protein was expressed, purified and crystallized using vapour-diffusion methods. The MSMEG_5817 protein was purified to homogeneity as determined by SDS–PAGE (Fig. 1) and routinely produced yields of 5 mg l^{-1} . Crystals suitable for X-ray crystallographic studies were obtained using 1.42 M Li_2SO_4 , 0.1 M Tris–HCl pH 7.7 with 0.1 M sodium citrate tribasic dihydrate as an additive (Fig. 2a). A complete

Figure 2
Crystals of (a) native MSMEG_5817 and (b) selenomethionine-derivatized MSMEG_5817 obtained from 20 mg ml^{-1} protein solution (in 20 mM MES pH 6.0, 100 mM NaCl) using 1.42 M lithium sulfate, 0.1 M Tris pH 7.7, 0.1 M sodium citrate tribasic dihydrate.

Figure 3

A typical 0.50° oscillation image from a selenomethionine-derivatized MSMEG_5817 crystal flash-cooled in 15% glycerol.

diffraction data set was collected to 2.40 Å resolution at 100 K. Preliminary crystallographic analysis indicated that the crystals belonged to space group $P4_32_12$, with unit-cell parameters $a = 50.5$, $b = 50.5$, $c = 205.5$ Å, $\alpha = \beta = \gamma = 90.0^\circ$. The data-collection and processing statistics are summarized in Table 1. Based on Matthews coefficient calculation, two molecules (42.68% solvent content) could be accommodated in the asymmetric unit, with an acceptable V_M of around $2.14 \text{ Å}^3 \text{ Da}^{-1}$ (Matthews, 1968). Molecular replacement was unsuccessful for MSMEG_5817 owing to the existing model (PDB entry 1iw7; *Thermus thermophilus* RNA polymerase holoenzyme; Vassylyev *et al.*, 2002) having relatively low sequence similarity. A selenomethionine derivative was expressed and the protein was purified using the same methods as used for the native protein. High-quality crystals were generated in the same crystallization condition by cross-seeding using crystals of the native protein (Fig. 2*b*). Three data sets were collected from a single crystal at 100 K based on experimentally determined inflection ($\lambda_1 = 0.97927$ Å), peak ($\lambda_2 = 0.97949$ Å) and high-energy remote ($\lambda_3 = 0.95369$ Å) wavelengths. A strong anomalous signal from selenium was detected to 2.5 Å resolution in each of the data sets (Fig. 3). Structural determination by multiple-wavelength anomalous dispersion is currently under way.

We thank the staff of the Australian Synchrotron and Monash Macromolecular Crystallization Facility for assistance with crystallization and X-ray data collection. This work was supported by the Australian Research Council (ARC) Centre of Excellence in Struc-

tural and Functional Microbial Genomics and the National Health and Medical Research Council of Australia. JR is an NHMRC Australia Fellow and TB is a Pfizer Australian Research Fellow.

References

- Anes, E., Peyron, P., Staali, L., Jordao, L., Gutierrez, M. G., Kress, H., Hagedorn, M., Maridonneau-Parini, I., Skinner, M. A., Wildeman, A. G., Kalamidas, S. A., Kuehn, M. & Griffiths, G. (2006). *Cell. Microbiol.* **8**, 939–960.
- Armstrong, J. A. & Hart, P. D. (1971). *J. Exp. Med.* **134**, 713–740.
- Billman-Jacobe, H., McConville, M. J., Haite, R. E., Kovacevic, S. & Coppel, R. L. (1999). *Mol. Microbiol.* **33**, 1244–1253.
- Chandra, N., Kumar, D. & Rao, K. (2011). *Tuberculosis*, **91**, 487–496.
- Frattini, R. A., Chua, J., Vergne, I. & Deretic, V. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 5437–5442.
- Kovacevic, S., Anderson, D., Morita, Y. S., Patterson, J., Haite, R., McMillan, B. N., Coppel, R., McConville, M. J. & Billman-Jacobe, H. (2006). *J. Biol. Chem.* **281**, 9011–9017.
- Leslie, A. G. W. & Powell, H. R. (2007). *Evolving Methods for Macromolecular Crystallography*, edited by R. J. Read & J. L. Sussman, pp. 41–51. Dordrecht: Springer.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Patterson, J. H., McConville, M. J., Haite, R. E., Coppel, R. L. & Billman-Jacobe, H. (2000). *J. Biol. Chem.* **275**, 24900–24906.
- Pelosi, A., Smith, D., Brammananth, R., Topolska, A., Billman-Jacobe, H., Nagley, P., Crellin, P. K. & Coppel, R. L. (2012). *PLoS One*, **7**, e31788.
- Vassylyev, D. G., Sekine, S., Laptchenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S. & Yokoyama, S. (2002). *Nature (London)*, **417**, 712–719.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.