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

Alessia Ruggiero,[a]‡ Flavia Squeglia,[a,b]‡ Luciano Pirone,[a,b] Stefania Correale[a] and Rita Berisio[a]*

[a]Institute of Biostructures and Bioimaging, CNR, Via Mezzocannone 16, I-80134 Napoli, Italy, and [b]Università degli Studi di Napoli 'Federico II', I-80134 Napoli, Italy

‡ These authors contributed equally to this work.

Correspondence e-mail: rita.berisio@unina.it

Received 22 October 2010
Accepted 29 November 2010

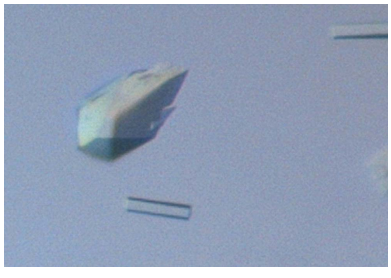
Expression, purification, crystallization and preliminary X-ray crystallographic analysis of a major fragment of the resuscitation-promoting factor RpfB from *Mycobacterium tuberculosis*

RpfB is required for the virulence and the resuscitation from dormancy of *Mycobacterium tuberculosis*, the bacterium responsible for tuberculosis. This protein is a cell-wall glycosidase that acts by cleaving peptidoglycans of the bacterial cell wall and therefore stimulates both bacterial growth and resuscitation from latency. RpfB consists of 362 residues organized into five domains. A long portion of RpfB, including its C-terminal catalytic domain, the G5 domain and one of its three DUF348 domains, which are of hitherto unknown structure and function, has been successfully crystallized using vapour-diffusion methods and seeding techniques. The crystals diffracted to 2.55 Å resolution and belonged to space group $C222_1$, with unit-cell parameters a = 102.3, b = 126.2, c = 85.87 Å. Model building using phases derived from the combined use of multiwavelength anomalous dispersion and molecular replacement is in progress. The results obtained here will provide the first structural characterization of a DUF348 domain reported to date and will shed light on the functional role of the noncatalytic domains of RpfB.

1. Introduction

Tuberculosis is one of the most life-threatening infectious diseases worldwide (Kaufmann, 2008; Kaufmann & McMichael, 2005) and can act both intra-pulmonarily and extra-pulmonarily (Kaufmann, 2008; Esposito *et al.*, 2008). Its causative agent, *Mycobacterium tuberculosis*, is able to survive in humans without producing apparent symptoms (Kaufmann, 2008). However, this apparent dormancy can develop into active disease even decades after initial infection when the immune response weakens (Kaufmann, 2008; Keep *et al.*, 2006). Since about one-third of the world's population is infected with dormant *M. tuberculosis*, the risk of disease reactivation is troublesome.

Over the years, it has become clear that the reactivation from dormancy, growth and division of bacteria require cleavage of the cell-wall peptidoglycan. The main participants in the *M. tuberculosis* reactivation process are cell-wall hydrolases named resuscitation-promoting factors (Rpfs; Mukamolova, Yanopolskaya *et al.*, 1998; Mukamolova, Kaprelyants *et al.*, 1998). These protein factors have been shown to be able to enhance cell growth when added to dormant cultures (Mukamolova *et al.*, 2002; Davies *et al.*, 2008; Kana *et al.*, 2008; Wu *et al.*, 2008; Russell-Goldman *et al.*, 2008). *M. tuberculosis* expresses five Rpfs (RpfA–E; Fig. 1). However, only one of these enzymes, RpfB, is indispensable for resuscitation *in vivo* (Tufariello *et al.*, 2006). Of the five Rpfs produced by *M. tuberculosis*, RpfB exhibits the highest structural complexity. Indeed, the sole structured domain present in RpfA, RpfC, RpfD and RpfE is a lysozyme-like catalytic domain which acts by cleaving cell-wall glycosidic bonds. In contrast, RpfB contains a further four domains in addition to the catalytic domain; these are a G5 domain and three DUF348 (domains of unknown function) domains (Fig. 1). It is therefore not surprising that this protein plays a key role in bacterial revival (Tufariello *et al.*, 2006).



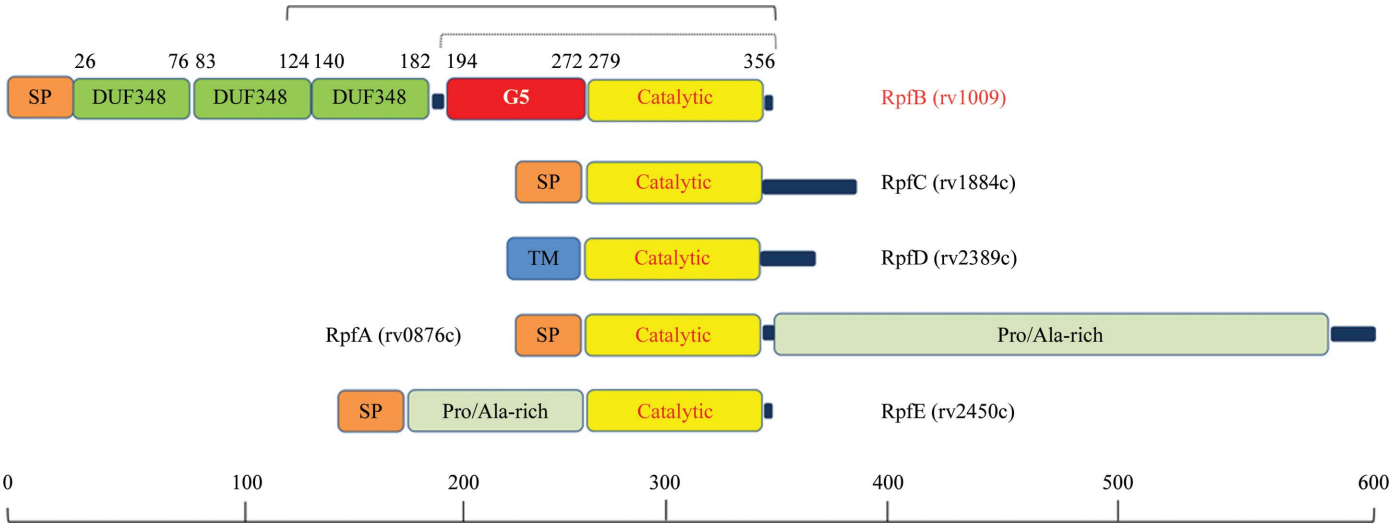


Figure 1
Domain organization of *M. tuberculosis* Rpf proteins as defined by the PFAM database. The catalytic domain (yellow) is the conserved domain in Rpf proteins. The regions of RpfB analysed in this study (amino-acid residues 115–362) and in previous work (Ruggiero *et al.*, 2009; residues 185–362) are indicated as a continuous line and a dashed line, respectively.

In a previous study, we determined the crystal structure of a portion of RpfB including the catalytic and G5 domains (Ruggiero *et al.*, 2009). These studies provided evidence for a novel fold of the G5 domain, which was constituted of two β-sheets connected by a small triple-helix motif and named β-TH-β. The structural features of the β-TH-β fold, with a large amount of exposed backbone hydrogen-bond donors and acceptors, which is a typical feature of adhesive proteins, suggested that G5 domains may be important for cell-wall adhesion (Ruggiero *et al.*, 2009), although this suggestion has yet to be proved experimentally. DUF348 domains normally occur as tandem repeats and, as in RpfB, they are found in conjunction with G5 domains. Although DUF348 domains are widespread in proteins, neither their functional role nor their structure is presently known. Here, we report the cloning, expression, purification, crystallization and preliminary crystallographic investigation of a biologically relevant portion of RpfB ($RpfB_{115–362}$) which contains its catalytic domain, the G5 domain and one DUF348 domain. The results obtained here will allow us to present the first structural characterization of a DUF348 domain and to gather clues on the functional role of the RpfB extra-domains that distinguish RpfB from its homologues.

2. Experimental methods

2.1. Cloning, expression and purification

The primers Rv1009F (5′-CATGCCATGGGAACGGCGCCGGCCGCG-3′) and Rv1009R (5′-AGCCGGATCCTCAGCGCGCACCCGCT-3′), containing *Nco*I and *Hin*dIII restriction sites, were used to amplify the *rpfB* coding sequence starting at residue Thr115 from the H37Rv strain of *M. tuberculosis*. The PCR (∼760 bp) product was cloned into expression vector pETM-11, giving a protein with a TEV-cleavable N-terminal poly-His tag named $RpfB_{115–362}$. The resulting positive plasmid was used to transform *Escherichia coli* BL21 (DE3) strain. The overnight culture was used to inculate 1 l LB medium containing 50 µg ml^{-1} kanamycin; protein induction was performed by the addition of 1 m*M* IPTG at 295 K when an OD_{600} value of 0.7 was reached. After approximately 16 h the cells were harvested and the protein was isolated by sonicating cell pellets resuspended in 20 ml binding buffer [5 m*M* imidazole, 300 m*M* NaCl, 50 m*M* Tris–HCl, 10%(*v*/*v*) glycerol pH 8.0] in the presence of a protease-inhibitor cocktail (Roche Diagnostics). The lysate was cleared by centrifugation at 18 000 rev min^{-1} and the supernatant was loaded onto a 5 ml Ni–NTA column (Pharmacia) equilibrated with binding buffer. After washing with ten volumes of binding buffer, a linear gradient of imidazole (5–300 m*M*) was applied to elute the protein. The fractions containing $RpfB_{115–362}$ were pooled and dialysed against 2 l 50 m*M* Tris–HCl, 150 m*M* NaCl, 10%(*v*/*v*) glycerol pH 8.0 with one exchange at 277 K. After removal of the His tag using TEV protease, the cleaved protein was purified by a second Ni–NTA affinity run. A final purification step was carried out using a Superdex 200 16/60 (Pharmacia) column [50 m*M* Tris–HCl, 150 m*M* NaCl, 10%(*v*/*v*) glycerol pH 8.0]. All purification steps were carried out at 293 K. The homogeneity of the protein was tested by SDS–PAGE (Fig. 2). The molecular mass calculated by mass spectrometry was 26.1 kDa. Freshly concentrated protein at 8 mg l^{-1} was used for crystallization experiments.

A selenomethionine derivative of $RpfB_{115–362}$ ($_{SeMet}RpfB_{115–362}$) was prepared by growing *E. coli* BL21 (DE3) cells expressing the recombinant enzyme in 1 l minimal medium (M9) containing 0.4% glucose, 1 m*M* $MgSO_4$, 0.1 m*M* $CaCl_2$, 50 µg ml^{-1} kanamycin and 100 µg ml^{-1} thiamine at 310 K. After reaching an OD_{600} of 0.7, an

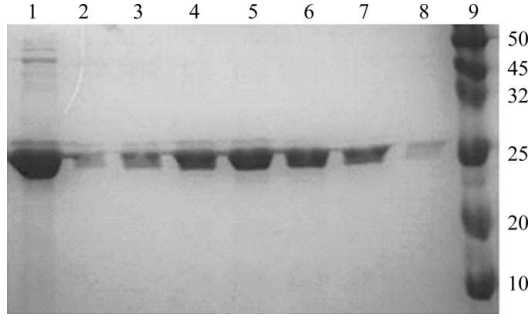


Figure 2
SDS–PAGE of eluted fractions (lanes 2–8) after S200 size-exclusion chromatography. Lane 1 contains the loaded sample after Ni–NTA affinity chromatography. Lane 9 contains molecular-weight markers (kDa).

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

	SeMet derivative			
	Peak	Inflection point	Remote	Native
Beamline	BM14	BM14	BM14	BM14
Space group	$C222_1$	$C222_1$	$C222_1$	$C222_1$
Unit-cell parameters (Å)				
a	102.31	102.24	102.24	102.38
b	126.96	126.41	126.35	126.99
c	85.97	86.03	86.16	86.29
Resolution range (Å)	50.00–2.95 (3.15–2.95)	50.00–3.05 (3.16–3.05)	50.00–3.02 (3.11–3.02)	50.00–2.55
Wavelength (Å)	0.9784	0.9788	0.9464	0.9464
Mosaicity (°)	0.6	0.8	0.8	0.5
Average multiplicity	9.2 (3.8)	5.1 (3.2)	6.3 (4.9)	16.2 (8.7)
No. of observed reflections	106763	49471	68066	283983
Unique reflections	11630	9780	10804	17505
Completeness (%)	90.5 (65.8)	90.8 (71.3)	90.0 (70.1)	93.7 (71.0)
$R_{\rm merge}$† (%)	9.0 (38.0)	7.0 (33.1)	9.0 (44.8)	7.7 (28.8)
Average $I/\sigma(I)$	13.8 (2.0)	12.5 (1.9)	9.8 (2.0)	31.5 (3.0)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the *i*th measurement of reflection *hkl* and $\langle I(hkl)\rangle$ is the mean value of the intensity of reflection *hkl*.

amino-acid mixture (50 mg l^{-1} Ile, Leu and Val and 100 mg l^{-1} Phe, Thr and Lys) was added to the culture, which was then shifted to 295 K. After equilibration, 60 mg l^{-1} seleno-L-methionine was added and induction was performed. The labelled protein was purified as described above.

2.2. Crystallization experiments

Crystallization was performed at 293 K by the hanging-drop vapour-diffusion method. Preliminary crystallization trials were carried out using Crystal Screen HT reagents (Hampton Research). Optimization of the crystallization conditions was performed by fine-tuning the protein and precipitant concentrations using a drop consisting of 1 µl protein solution and 1 µl precipitant solution and a reservoir volume of 400 µl. As a result, only tiny microcrystals could be obtained. Therefore, both the macroseeding and the microseeding techniques were adopted.

For macroseeding, initial microcrystals were isolated and partially dissolved in their mother liquor. The microcrystals were transferred to an array of equilibrated crystallization drops in different conditions. For microseeding, microcrystals were ground with a mortar to produce a stock mixture of crystal seeds (Luft & DeTitta, 1999). This mixture was diluted tenfold, 100-fold and 500-fold in a stabilizing solution consisting of 50% reservoir solution, 10% protein solution and 40% water. 1 µl of each seeded solution was mixed with 99 µl purified protein solution at a concentration of 10 mg ml^{-1}. The resulting protein solutions were again screened against the Crystal Screen HT reagents.

2.3. Data collection and processing

Preliminary diffraction data were collected in-house at 100 K using a Rigaku MicroMax-007 HF generator producing Cu $K\alpha$ radiation and equipped with a Saturn944 CCD detector.

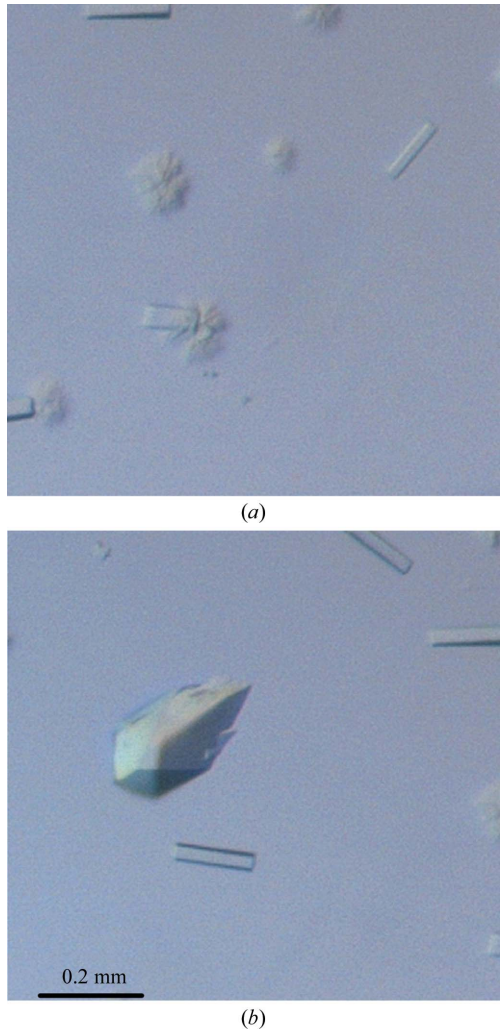
Higher resolution diffraction data for both native $RpfB_{115–362}$ and $_{SeMet}RpfB_{115–362}$ were collected on synchrotron beamline BM14 at the ESRF (Grenoble, France) at 100 K. Cryoprotection of the crystals was achieved by rapid soaking (1–2 s) in a solution consisting of 14%(*v*/*v*) ethanol, 50 m*M* Tris–HCl buffer pH 8.5, 30%(*v*/*v*) glycerol. An oscillation range of 1° and an X-ray dose corresponding to about 4 s exposure were adopted for all experiments. The data sets were scaled and merged using the *HKL*-2000 program package (Otwinowski & Minor, 1997). The multiwavelength anomalous diffraction (MAD) method was performed using data collected at three different wavelengths determined from a selenium absorption spectrum and the native data set (Table 1).

2.4. Structure determination

The structure of the enzyme was solved using a combination of the anomalous dispersion and molecular-replacement techniques as implemented in the *Auto-Rickshaw* pipeline (Panjikar *et al.*, 2005). The phases derived by *SHELXD* (Sheldrick, 2008) were improved by the solvent-flattening, phase-extension and noncrystallographic symmetry-averaging methods as implemented in the program *RESOLVE* (Terwilliger, 2003*b*).

3. Results and discussion

The gene for $RpfB_{115–362}$ consists of 760 bp encoding 248 amino-acid residues (115–362). These contain the predicted catalytic domain and the G5 domain as well as a novel domain, DUF348, as identified in the PFAM database. This construct has been successfully cloned and expressed and was crystallized using vapour-diffusion and seeding



(*a*)

(*b*)

Figure 3
Image of typical $RpfB_{115–362}$ crystals grown (*a*) by vapour diffusion and (*b*) by using macroseeding techniques (see text for details).

methods. The purified $RpfB_{115–362}$ showed a single band of approximately 26 kDa on SDS–PAGE, which is in good agreement with the molecular mass of 26.1 kDa derived by mass spectrometry. Initial screenings using commercially available solutions revealed preliminary indications of crystallization in several conditions containing alcohol. Fine-tuning of the crystallization conditions using the hanging-drop method produced microcrystals that were not suitable for X-ray diffraction studies (Fig. 3*a*). The use of additive screens (Hampton Research) did not improve their morphology. The best crystals of $RpfB_{115–362}$ were obtained using the macroseeding technique. Microcrystals (Fig. 3*a*) were partially dissolved in their mother liquor and transferred into several solutions containing 1:1 mixtures of 10 mg ml^{-1} $RpfB_{115–362}$ and dilutions of commercially available solutions (Hampton Research). Crystals of dimensions 0.05 × 0.05 × 0.2 mm were obtained after the transfer of microcrystals to a pre-equilibrated solution containing 5 mg ml^{-1} $RpfB_{115–362}$, 18%(*v*/*v*) ethanol in 50 m*M* Tris–HCl buffer pH 8.5 (Fig. 3*b*).

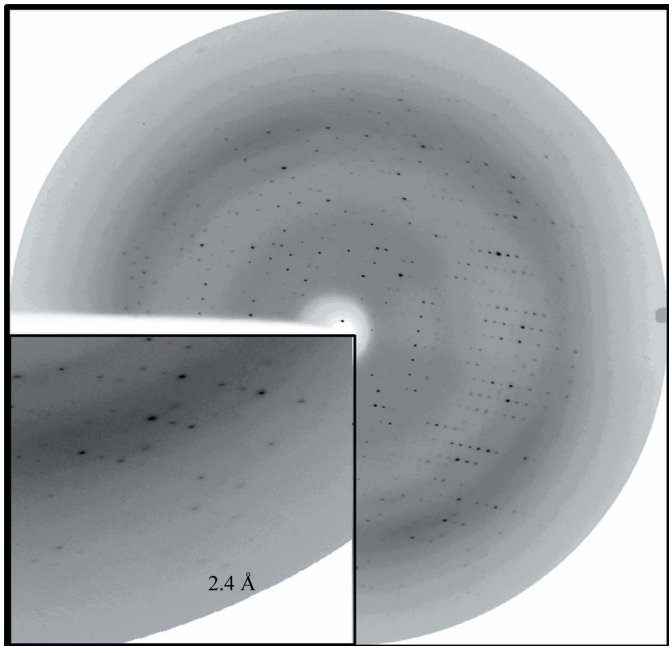


Figure 4
Diffraction pattern of a native RpfB crystal. Diffraction data were detectable to 2.4 Å resolution.

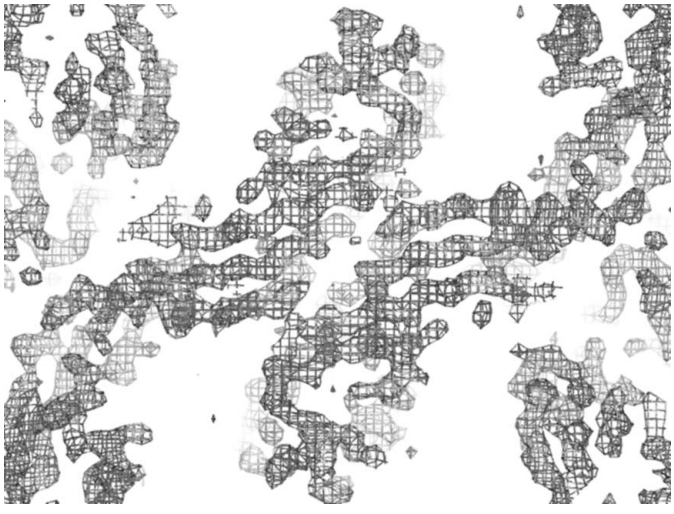


Figure 5
Experimental electron-density map contoured at 1.5σ after density modification. The density clearly shows β-sheet regions of the protein.

The crystals diffracted to 2.4 Å resolution (Fig. 4) and belonged to space group $C222_1$, with unit-cell parameters a = 102.38, b = 126.99, c = 86.29 Å (Table 1). Matthews coefficient calculations suggested the presence of two molecules (V_M = 2.7 Å^3 Da^{-1}, 54% solvent content; Matthews, 1968) in the asymmetric unit. Several attempts were made to solve the structure by molecular replacement (MR) using either the entire structure of $_{\Delta DUF}$RpfB (residues 194–362) or its isolated domains as a starting model (PDB code 3eo5; Ruggiero *et al.*, 2009) and various MR packages (Storoni *et al.*, 2004; Navaza & Saludjian, 1997; Caliandro *et al.*, 2006). However, all MR trials were unsuccessful. Therefore, an SeMet derivative of the protein was prepared in order to perform MAD experiments. Crystals of $_{SeMet}RpfB_{115–362}$ were obtained using the same procedure as adopted for the native protein. The best crystals grew from 3 mg ml^{-1} protein solution and 18%(*v*/*v*) ethanol in 60 m*M* Tris–HCl buffer pH 8.5. These crystals diffracted to 2.95 Å resolution on the BM14 beamline at the ESRF, Grenoble. For determination of the peak and inflection wavelengths, a fluorescence scan was recorded from a single SeMet-labelled $RpfB_{115–362}$ crystal.

Using data sets collected at wavelengths optimized for SeMet (Table 1) and the data set from a native crystal, the *Auto-Rickshaw* pipeline was adopted, combining phases derived by the four-wavelength MAD and molecular-replacement (MR) methods (Panjikar *et al.*, 2005). The best results were obtained using the structure of the sole G5 domain (PDB code 3eo5) as a search model. Interestingly, the sole use of MAD techniques did not produce readily interpretable maps. Using the combined MAD–MR approach, the program *SHELXD* identified ten selenium sites in the asymmetric unit of the protein (Sheldrick, 2008). The initial set of phases was improved using solvent-flattening, phase-extension and noncrystallographic symmetry-averaging methods as implemented in the program *RESOLVE* (Terwilliger, 2003*b*). After this procedure, the electron-density maps at 2.95 Å resolution (Fig. 5) were of sufficient quality to automatically trace about 80% of the residues present in the asymmetric unit (Terwilliger, 2003*a*). Manual model-building sessions aimed at defining the complete structure of $RpfB_{115–362}$ are in progress. This work will lead to the first structure of a DUF348 domain. In addition, the use of homology-modelling techniques to derive the structures of the remaining two DUF348 domains of RpfB will allow us to derive structural information on the entire enzyme. This information will be important for understanding the role in RpfB function of the noncatalytic domains that make RpfB indispensable for resuscitation (Tufariello *et al.*, 2006).

This work was funded by the MIUR (FIRB, contract No. RBRN07BMCT). We acknowledge the staff of the Macromolecular Crystallography Group at the ESRF (beamline BM14, Grenoble, France) for providing the synchrotron-radiation facilities and for valuable assistance during data collection. We also thank Ms Luisa Palamara for her help during sample preparation.

References

Caliandro, R., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C., Mazzone, A. M. & Siliqi, D. (2006). *J. Appl. Cryst.* **39**, 185–193.
Davies, A. P., Dhillon, A. P., Young, M., Henderson, B., McHugh, T. D. & Gillespie, S. H. (2008). *Tuberculosis*, **88**, 462–468.
Esposito, C., Pethoukov, M. V., Svergun, D. I., Ruggiero, A., Pedone, C., Pedone, E. & Berisio, R. (2008). *J. Bacteriol.* **190**, 4749–4753.

Kana, B. D., Gordhan, B. G., Downing, K. J., Sung, N., Vostroktunova, G., Machowski, E. E., Tsenova, L., Young, M., Kaprelyants, A., Kaplan, G. & Mizrahi, V. (2008). *Mol. Microbiol.* **67**, 672–684.
Kaufmann, S. H. (2008). *Nature (London)*, **453**, 295–296.
Kaufmann, S. H. & McMichael, A. J. (2005). *Nature Med.* **11**, S33–S44.
Keep, N. H., Ward, J. M., Cohen-Gonsaud, M. & Henderson, B. (2006). *Trends Microbiol.* **14**, 271–276.
Luft, J. R. & DeTitta, G. T. (1999). *Acta Cryst.* D**55**, 988–993.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mukamolova, G. V., Kaprelyants, A. S., Young, D. I., Young, M. & Kell, D. B. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 8916–8921.
Mukamolova, G. V., Turapov, O. A., Young, D. I., Kaprelyants, A. S., Kell, D. B. & Young, M. (2002). *Mol. Microbiol.* **46**, 623–635.
Mukamolova, G. V., Yanopolskaya, N. D., Kell, D. B. & Kaprelyants, A. S. (1998). *Antonie Van Leeuwenhoek*, **73**, 237–243.
Navaza, J. & Saludjian, P. (1997). *Methods Enzymol.* **276**, 581–594.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2005). *Acta Cryst.* D**61**, 449–457.
Ruggiero, A., Tizzano, B., Pedone, E., Pedone, C., Wilmanns, M. & Berisio, R. (2009). *J. Mol. Biol.* **385**, 153–162.
Russell-Goldman, E., Xu, J., Wang, X., Chan, J. & Tufariello, J. M. (2008). *Infect. Immun.* **76**, 4269–4281.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* D**60**, 432–438.
Terwilliger, T. C. (2003*a*). *Acta Cryst.* D**59**, 45–49.
Terwilliger, T. C. (2003*b*). *Methods Enzymol.* **374**, 22–37.
Tufariello, J. M., Mi, K., Xu, J., Manabe, Y. C., Kesavan, A. K., Drumm, J., Tanaka, K., Jacobs, W. R. Jr & Chan, J. (2006). *Infect. Immun.* **74**, 2985–2995.
Wu, X., Yang, Y., Han, Y., Zhang, J., Liang, Y., Li, H., Li, B. & Wang, L. (2008). *J. Appl. Microbiol.* **105**, 1121–1127.