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

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

Alessia Ruggiero,[a] Barbara Tizzano,[b] Arie Geerlof,[b] Emilia Pedone,[a] Carlo Pedone,[a] Matthias Wilmanns[b] and Rita Berisio[a]*

[a]Institute of Biostructures and Bioimaging, CNR, Via Mezzocannone 16, I-80134 Napoli, Italy, and [b]European Molecular Biology Laboratory, Hamburg, c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany

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

Received 31 May 2007
Accepted 3 September 2007

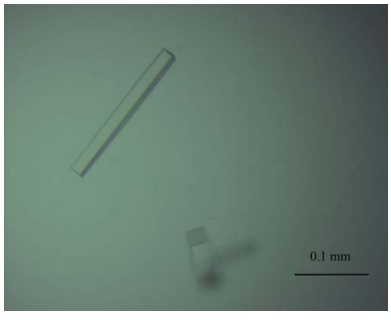
The resuscitation-promoting factor RpfB, the most complex of the five resuscitation-promoting factors produced by *M. tuberculosis*, is devoted to bacterial reactivation from the dormant state. RpfB consists of 362 residues predicted to form five domains. An RpfB fragment containing the protein catalytic domain and a G5 domain has been successfully crystallized using vapour-diffusion methods. This is the first crystallographic study of a resuscitation-promoting factor. Crystals of this protein belong to space group $I422$, with unit-cell parameters $a = 97.63$, $b = 97.63$, $c = 114.87$ Å. Diffraction data have also been collected from a selenomethionine derivative at 2.9 Å resolution. Model building using the phases derived from the multiwavelength anomalous dispersion experiment is in progress.

1. Introduction

The interaction between *Mycobacterium tuberculosis* and the human host after infection may manifest itself as a chronic disease or as a latent (or dormant) infection, a state that is capable of evading host responses. The probability of reactivation from dormancy is strongly affected by the type of host immune response and is significantly enhanced in immunocompromised patients, *e.g.* those suffering from AIDS. Understanding and controlling the entry and exit from dormancy is important in the development of new antitubercular therapies.

Cell growth of dormant *M. luteus* cultures has been shown to be enhanced after the addition of a secreted protein, named resuscitation-promoting factor (Rpf; Mukamolova *et al.*, 1998). Homologues of this protein exist in *M. tuberculosis*, which has five genes encoding Rpfs (*rpfA–E*; Mukamolova *et al.*, 2002; Cole *et al.*, 1998). Whether the existence of five Rpfs is the result of full redundancy or not is an issue that is under debate. Indeed, analysis of deletion mutants has shown that each of the five Rpfs is dispensable for growth (Tufariello *et al.*, 2004). On the other hand, mutants lacking three of the five Rpfs are defective both for growth *in vivo* and for resuscitation *in vitro* (Downing *et al.*, 2005). More recently, *in vivo* studies using a mouse model of infection and reactivation have shown that the deletion of RpfB is the sole requirement to produce delayed reactivation from chronic tuberculosis (Tufariello *et al.*, 2006).

It has been demonstrated that dormant *Escherichia coli* bacteria present a substantial increase in the degree of peptidoglycan cross-linking (Signoretto *et al.*, 2002). In this context, Rpfs have been proposed to act as peptidoglycan hydrolases on the cell wall of dormant bacteria, thus altering the mechanical properties of the cell wall and favouring cell division and/or the release of antidormancy factors (Mukamolova *et al.*, 2006; Telkov *et al.*, 2006; Keep *et al.*, 2006). Consistent with this hypothesis, fold-prediction studies as part of the 'ten most wanted' targets in 'Critical Assignment of Techniques for Protein Structure Prediction' have shown that the conserved Rpf domain exhibits a lysozyme-like fold (Abbott, 2001). This prediction has been confirmed by the NMR structure of the conserved domain of RpfB (Cohen-Gonsaud *et al.*, 2005).



RpfB is the largest and the most complex of the five Rpf proteins encoded by *M. tuberculosis*. In addition to the catalytic domain (about 80 residues), RpfB (362 residues) contains (as defined by the PFAM database; Finn *et al.*, 2006) a G5 domain and three DUF348 domains. It is therefore not surprising that this protein plays a key role in bacterial revival (Tufariello *et al.*, 2006). No functional or structural roles are currently known for the G5 and DUF348 domains, but hypotheses may be possible. G5 domains are found as one to seven copies in a variety of enzymes, such as various glycosyl hydrolases in bacteria. A common feature of proteins containing G5 domains is *N*-acetylglucosamine binding. DUF348 domains normally occur as tandem repeats. As in RpfB, they are found in conjunction with G5 domains. Although both G5 and DUF348 domains are widespread in proteins, their structures are so far not known. Here, we report the cloning, expression, purification, crystallization and preliminary crystallographic investigations of an RpfB construct containing the catalytic domain and the G5 domain. This protein (residues 185–362), lacking the N-terminal DUF348 domains, is denoted here as $_{\Delta DUF}$RpfB. To better understand the structural basis of Rpf functionality, $_{\Delta DUF}$RpfB has also been cocrystallized with tri-*N*-acetylglucosamine.

The results obtained here will provide the first crystallographic study of an Rpf reported to date. In addition to aiming towards better understanding of the mechanism of exit from dormancy in *M. tuberculosis*, comprehension of the structural features associated with Rpf activity/inhibition will provide the basis for the identification of molecules (pro-latency molecules) that are able to restrict bacterial life to the nondangerous latent state.

2. Experimental methods

2.1. Cloning and purification

Two primers, Rv1009F (5′-CATG**CCATGG**AGTCGAAGGCATGCCA-3′) and Rv1009R (5′-CCC**AAGCTT**ATCAGCGCGCACCCGTC-3′), containing *Nco*I and *Hin*dIII restriction sites (in bold), were employed to amplify the *rpfB* coding sequence starting at residue Val185 from the H37Rv strain of *M. tuberculosis*. The PCR product (528 bp) was cloned into the expression vector pETM-11, which gives a protein with a cleavable N-terminal poly-His tag (sequence MKHHHHHHPMSDYDIPTTENLYFEGA). The resulting positive plasmid was used to transform *E. coli* strain BL21(DE3). The transformed *E. coli* cells were grown overnight at 310 K in LB containing 50 µg ml^{-1} kanamycin and then induced overnight with 1 m*M* IPTG at 295 K.

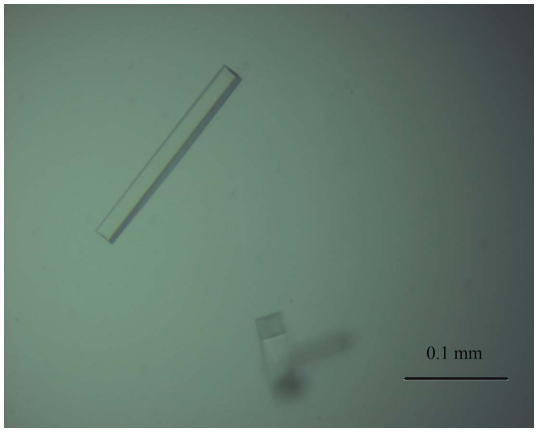


Figure 1
Image of typical $_{\Delta DUF}$RpfB crystals grown using 18%(*v*/*v*) 2-propanol in 60 m*M* sodium cacodylate trihydrate buffer pH 6.5 (see text for details).

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] containing a protease-inhibitor cocktail (Roche Diagnostic). 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 $_{\Delta DUF}$RpfB were pooled and dialyzed at 277 K against 2 l 50 m*M* Tris–HCl, 200 m*M* NaCl, 10%(*v*/*v*) glycerol pH 8.0. After removal of the histidine tag, the protein was further purified by size-exclusion chromatography on Superdex 200 [GE Healthcare; 50 m*M* Tris–HCl, 150 m*M* NaCl, 10%(*v*/*v*) glycerol pH 8.0]. The homogeneity of the protein was tested by SDS–PAGE. The molecular weight obtained by mass spectroscopy, 19 367 kDa, is consistent with the presence of two extra residues (Gly-Ala) at the protein N-terminus, which remained after proteolytic removal of the His tag. Freshly concentrated protein, usually at 5–10 mg ml^{-1}, was used for crystallization experiments.

A selenomethionine derivative of $_{\Delta DUF}$RpfB (SeMet $_{\Delta DUF}$RpfB) was prepared by growing *E. coli* BL21(DE3) cells expressing the recombinant enzyme in 1 l minimal media (M9) containing 0.4% glucose, 1 m*M* $MgSO_4$, 0.1 m*M* $CaCl_2$, 50 µg l^{-1} kanamycin, 100 µg l^{-1} thiamine at 310 K. After reaching an OD_{600} of 0.7, an amino-acid mix (50 mg l^{-1} Ile, Leu and Val and 100 mg l^{-1} Phe, Thr and Lys) was added to the culture at 295 K. After equilibration, 60 mg l^{-1} seleno-L-methionine was added to the culture and induction was performed. A purification protocol similar to that for the native enzyme was used for purification of the selenomethionine (SeMet) derivative.

2.2. Crystallization experiments

Crystallization was performed at 293 K by the hanging-drop vapour-diffusion method. Preliminary crystallization trials were carried out using commercially available sparse-matrix screens (Crystal Screens I and II, Hampton Research). Crystals suitable for X-ray diffraction were obtained by tuning the protein and precipitant concentrations. Drops were prepared by mixing 1 µl protein solution with 1 µl precipitant solution and were equilibrated against 400 µl reservoir solution. The same approach was used to grow crystals of the SeMet derivative. Crystals of $_{\Delta DUF}$RpfB were also grown by cocrystallization with 5 m*M* tri-*N*-acetylglucosamine.

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 data at 3.3 Å were collected for native $_{\Delta DUF}$RpfB at beamline BW7A of the DESY synchrotron (Hamburg, Germany) at 100 K. Crystals were flash-cooled after the addition of 28%(*v*/*v*) glycerol to the crystallization buffer. Multiwavelength anomalous dispersion (MAD) data were collected at beamline BM14 at the ESRF synchrotron (Grenoble, France). Three different data sets were collected from a single crystal using wavelengths determined from the selenium-absorption spectrum. Data processing and scaling were performed using the program *HKL*-2000 (Otwinowski & Minor, 1997).

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

	SeMet derivative (tri-*N*-acetylglucosamine complex)		
	Peak	Inflection point	Remote
Beamline	BM14	BM14	BM14
Space group	*I*422	*I*422	*I*422
Unit-cell parameters (Å)			
a	97.63	97.731	97.507
b	97.63	97.731	97.507
c	114.87	115.055	114.751
Resolution range (Å)	30.00–2.90	30.00–3.10	30.00–2.90
Last shell (Å)	3.00–2.90	3.21–3.10	3.00–2.90
Wavelength (Å)	0.9789	0.9791	0.8856
Average redundancy	9.3 (9.8)	10.8 (11.1)	7.9 (7.8)
Unique reflections	6394	5306	6383
Completeness	99.6 (99.8)	99.8 (99.6)	99.8 (100.0)
$R_{\rm merge}$† (%)	6.0 (45.0)	5.1 (44.8)	5.6 (46.1)
Average $I/\sigma(I)$	47.2 (3.2)	42.7 (3.9)	35.9 (2.9)

† $R_{\rm merge} = \sum_h \sum_i |I(h,i) - \langle I(h)\rangle| / \sum_h \sum_i I(h,i)$, where $I(h,i)$ is the intensity of the ith measurement of reflection h and $\langle I(h)\rangle$ is the mean value of the intensity of reflection h.

2.4. Structure determination

Solution of the RpfB structure using MAD methods is in progress. The program *SOLVE* was used to identify and localize the selenium sites present in the asymmetric unit and to derive the experimental phases (Terwilliger & Berendzen, 1999). Phases were improved by density modification using the program *RESOLVE* (Terwilliger, 2003*b*). Model building using both automatic (Terwilliger, 2003*a*) and manual (Jones, 2004) approaches is in progress.

3. Results and discussion

The initial screenings using commercially available solutions revealed several promising conditions for the crystallization of $_{\Delta\rm DUF}$RpfB. All favourable conditions were characterized by the presence of an alcohol as precipitating agent. The quality of the crystals was improved by fine-tuning the concentration of the protein and of the precipitants. $_{\Delta\rm DUF}$RpfB crystals (Fig. 1) suitable for X-ray diffraction data collection (0.05 × 0.05 × 0.4 mm) were obtained by hanging-drop vapour diffusion using 8.0 mg ml^{-1} protein solution and 18%(*v*/*v*) 2-propanol in 60 m*M* sodium cacodylate trihydrate buffer pH 6.5. The crystals belonged to space group *I*422, with unit-cell parameters a = 97.63, b = 97.63, c = 114.87 Å (Table 1). Matthews coefficient calculations (Matthews, 1968) suggested the presence of either two molecules ($V_{\rm M}$ = 1.77 Å^3 Da^{-1}, 30.4% solvent content) or one molecule per asymmetric unit ($V_{\rm M}$ = 3.53 Å^3 Da^{-1}, 65.2% solvent content). However, no significant peak, apart from that located at the origin, was observed in the self-rotation function.

Several attempts were made to solve the structure by molecular replacement (MR) using the NMR structure of the RpfB catalytic domain as a starting model (PDB code 1xsf; Cohen-Gonsaud *et al.*, 2005) and various MR packages (Storoni *et al.*, 2004; Navaza & Saludjian, 1997; Caliandro *et al.*, 2006). However, as frequently found when NMR models are used (Chen, 2001), all MR trials were unsuccessful. Therefore, an SeMet derivative of the protein was prepared in order to perform MAD experiments. Three methionines are present in the $_{\Delta\rm DUF}$RpfB sequence, which consists of 178 residues. Crystals of SeMet$_{\Delta\rm DUF}$RpfB grew by hanging-drop vapour diffusion using 5.0 mg ml^{-1} protein solution and 18%(*v*/*v*) 2-propanol in 60 m*M* sodium cacodylate trihydrate buffer pH 6.5. The best crystals, which were obtained in cocrystallization experiments with tri-*N*-acetylglucosamine, diffracted to 2.8 Å at the BM14 beamline, ESRF, Grenoble (Fig. 2). In order to determine the peak and inflection wavelengths, a fluorescence scan was recorded on a single SeMet-labelled $_{\Delta\rm DUF}$RpfB crystal (Fig. 3). Using data sets collected at wavelengths optimized for SeMet (Table 1), the program *SOLVE* (Terwilliger & Berendzen, 1999) identified three selenium sites in the asymmetric unit of the protein. This finding supports the hypothesis that only one molecule is present in the asymmetric unit. The program *SOLVE* provided a set of initial phases, which were improved using the solvent-flattening methods implemented in the program *RESOLVE* (Terwilliger, 2003*b*). Electron-density maps after solvent flattening were of sufficient quality to automatically trace nearly 80% of the residues present in the asymmetric unit (Terwilliger, 2003*a*). Manual model-building sessions aimed at

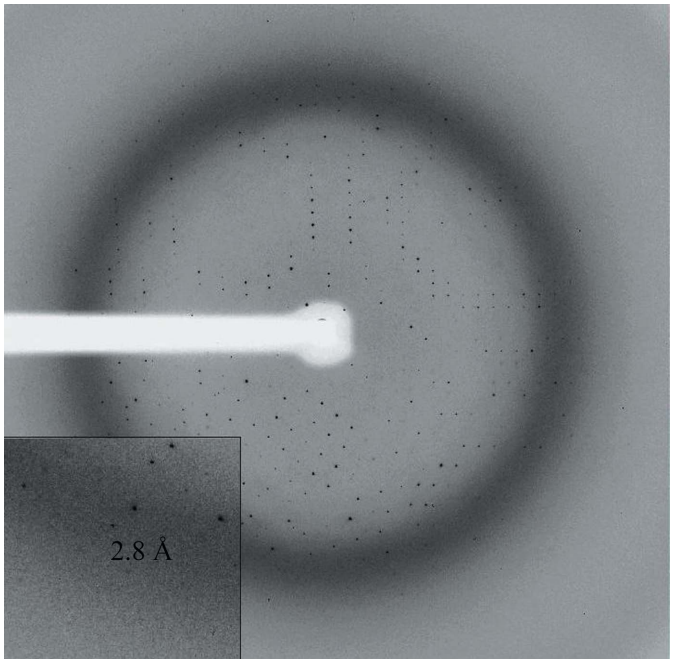


Figure 2
Diffraction pattern of a $_{\Delta\rm DUF}$RpfB crystal (SeMet derivative). Diffraction data are detectable to 2.8 Å resolution.

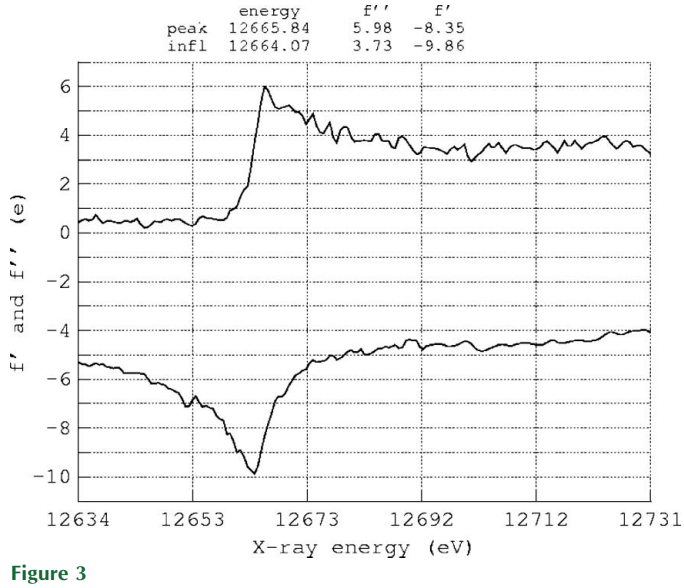


Figure 3
Selenium fluorescence measured from a crystal of $_{\Delta\rm DUF}$RpfB.

defining the complete $_{\Delta DUF}$RpfB structure are in progress (Jones, 2004).

We acknowledge the staff of the Macromolecular Crystallography Group at ESRF (beamline BM14, Grenoble, France) and the DESY staff (beamline BW7A, Hamburg, Germany) for providing the synchrotron-radiation facilities and for valuable assistance during data collection.

References

Abbott, A. (2001). *Nature (London)*, **409**, 4.
Caliandro, R., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C., Mazzone, A. M. & Siliqi, D. (2006). *J. Appl. Cryst.* **39**, 185–193.
Chen, Y. W. (2001). *Acta Cryst.* D**57**, 1457–1461.
Cohen-Gonsaud, M., Barthe, P., Bagneris, C., Henderson, B., Ward, J., Roumestand, C. & Keep, N. H. (2005). *Nature Struct. Mol. Biol.* **12**, 270–273.
Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
Downing, K. J., Mischenko, V. V., Shleeva, M. O., Young, D. I., Young, M., Kaprelyants, A. S., Apt, A. S. & Mizrahi, V. (2005). *Infect. Immun.* **73**, 3038–3043.
Finn, R. D., Mistry, J., Schuster-Bockler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S. R., Sonnhammer, E. L. & Bateman, A. (2006). *Nucleic Acids Res.* **34**, D247–D251.
Jones, T. A. (2004). *Acta Cryst.* D**60**, 2115–2125.
Keep, N. H., Ward, J. M., Cohen-Gonsaud, M. & Henderson, B. (2006). *Trends Microbiol.* **14**, 271–276.
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., Murzin, A. G., Salina, E. G., Demina, G. R., Kell, D. B., Kaprelyants, A. S. & Young, M. (2006). *Mol. Microbiol.* **59**, 84–98.
Mukamolova, G. V., Turapov, O. A., Young, D. I., Kaprelyants, A. S., Kell, D. B. & Young, M. (2002). *Mol. Microbiol.* **46**, 623–635.
Navaza, J. & Saludjian, P. (1997). *Methods Enzymol.* **276**, 581–594.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Signoretto, C., Lleo, M. M. & Canepari, P. (2002). *Curr. Microbiol.* **44**, 125–131.
Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* D**60**, 432–438.
Telkov, M. V., Demina, G. R., Voloshin, S. A., Salina, E. G., Dudik, T. V., Stekhanova, T. N., Mukamolova, G. V., Kazaryan, K. A., Goncharenko, A. V., Young, M. & Kaprelyants, A. S. (2006). *Biochemistry (Moscow)*, **71**, 414–422.
Terwilliger, T. C. (2003*a*). *Acta Cryst.* D**59**, 38–44.
Terwilliger, T. C. (2003*b*). *Methods Enzymol.* **374**, 22–37.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Tufariello, J. M., Jacobs, W. R. Jr & Chan, J. (2004). *Infect. Immun.* **72**, 515–526.
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.