STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Structure of the MarR family protein Rv0880 from *Mycobacterium tuberculosis*

Yun-Rong Gao,[a,b] Na Feng,[a] Tao Chen,[c] De-Feng Li[a] and Li-Jun Bi[a]*

[a]Key Laboratory of RNA Biology, National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, People's Republic of China, [b]University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China, and [c]Center for Tuberculosis Control of Guangdong Province, Guangzhou 510630, People's Republic of China. *Correspondence e-mail: blj@ibp.ac.cn

Received 3 March 2015
Accepted 11 April 2015

Edited by L. J. Beamer, University of Missouri, USA

Keywords: MarR; *Mycobacterium tuberculosis*; transcription factor; MarR family protein.

PDB reference: Rv0880, 4yif

Supporting information: this article has supporting information at journals.iucr.org/f

Rv0880 from the pathogen *Mycobacterium tuberculosis* is classified as a MarR family protein in the Pfam database. It consists of 143 amino acids and has an isoelectric point of 10.9. Crystals of Rv0880 belonged to space group *P*1, with unit-cell parameters a = 54.97, b = 69.60, c = 70.32 Å, α = 103.71, β = 111.06, γ = 105.83°. The structure of the MarR family transcription regulator Rv0880 was solved at a resolution of 2.0 Å with an R_{cryst} and R_{free} of 21.2 and 24.9%, respectively. The dimeric structure resembles that of other MarR proteins, with each subunit comprising a winged helix–turn–helix domain connected to an α-helical dimerization domain.

1. Introduction

All living organisms have molecular systems that enable them to resist a variety of toxic substances and environmental stresses. Proteins belonging to the multiple antibiotic-resistance regulator (MarR) family have been reported to regulate the expression of proteins conferring resistance to multiple antibiotics and pathogenic factors (Perera & Grove, 2010; Grove, 2013). In *Escherichia coli*, resistance to numerous antimicrobial agents is believed to be determined primarily *via* efflux pumps, the expression of which is controlled by the binding of MarR repressor proteins to their cognate DNA, preventing initiation of gene transcription (Lomovskaya *et al.*, 1995).

Knowledge of the structures of MarR family regulators has contributed to the understanding of their mechanism of action. To date, the structures of more than 20 MarR family proteins have been solved and deposited in the Protein Data Bank. These proteins are homodimers comprising a largely helical dimerization domain linked to a DNA-binding domain that contains a winged helix–turn–helix (wHTH) motif. MarR family proteins repress the activity of their target genes by binding as dimers to upstream pseudopalindromic sequences in the promoters of regulated proteins (Perera & Grove, 2010). MarR proteins typically function as repressors *via* one of two mechanisms: either by disrupting disulfide-bridge formation when MarR cysteines are oxidized or by the binding of a co-inducer such as salicylate. For example, the oxidation of cysteines in the MarR protein OhrR in *Bacillus subtilis* results in structural changes such that it is no longer able to bind to DNA (Soonsanga *et al.*, 2008; Eiamphungporn *et al.*, 2009), and the binding of salicylate to ST170 in *Sulfolobus tokodaii* strain 7 (Yu *et al.*, 2009) leads to a major rearrangement in the dimerization region of this MarR protein such that its DNA-binding domain is significantly altered, preventing DNA recognition.

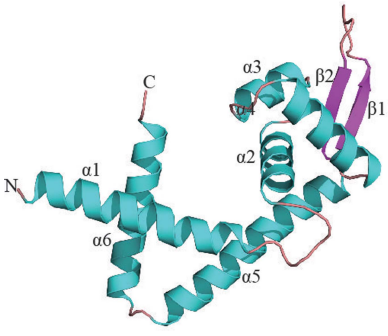


Table 1
Macromolecule-production information.

Source organism	*M. tuberculosis* H37Rv
DNA source	*M. tuberculosis* H37Rv genome
Forward primer	GGAATTCCATATGGTGCTTGACAGCGATGCGCG
Reverse primer	CCCAAGCTTTCACGGGCTTTCGTCGACCAG
Cloning vector	pET-28a
Expression vector	pET-28a
Expression host	*E. coli* BL21(DE3)
Complete amino-acid sequence of the construct produced	MGSSHHHHHHSSGLVPRGSHMVLDSDARLASDLS-LAVMRLSRQLRFRNPSSPVSLSQLSALTTLAN-EGAMTPGALAIRERVRPPSMTRVIASLADMGF-VDRAPHPIDGRQVLVSVSESGAELVKAARRAR-QEWLAERLATLNRSERDILRSAADLMLALVDE-SP

Nine million new cases of tuberculosis, caused by the pathogen *Mycobacterium tuberculosis* (*Mtb*), are reported annually (World Health Organization, 2014), and multidrug-resistant (MDR), extensively drug-resistant (XDR) and most recently totally drug-resistant (TDR) strains of *Mtb* have emerged and are spreading globally. Five MarR family proteins are encoded in the *Mtb* genome; the potential role of these regulatory proteins in drug resistance is unclear and is worthy of investigation. Given the potential importance of the MarR family of proteins in *Mtb*, we have performed structural studies on Rv0880, a previously uninvestigated protein which has been annotated as a MarR protein, and here describe its structure at 2.0 Å resolution.

2. Materials and methods

2.1. Macromolecule production

The Rv0880 gene was amplified from *M. tuberculosis* strain H37Rv genomic DNA by PCR and ligated into the vector pET-28a(+) (Novagen) using NdeI and HindIII restriction sites, generating a plasmid encoding recombinant Rv0880 protein with a 6×His tag at its N-terminus (pET-28aΩRv0880) which was then transformed into *E. coli* BL21(DE3) cells. Large-scale cultures were grown to an OD_{600} of between 0.6 and 0.8; Rv0880 expression was then induced by incubation with IPTG (final concentration of 0.5 m*M*) for 16 h at 16°C.

Bacterial cells were collected, suspended in ice-cold lysis buffer consisting of 20 m*M* Tris pH 7.4, 10 m*M* imidazole, 500 m*M* NaCl and subjected to high-pressure homogenization. Cell debris was removed by centrifugation at 30 000*g* for 40 min at 4°C. The crude lysate was loaded onto an Ni^{2+}-chelating column pre-equilibrated with 20 m*M* Tris pH 7.4, 10 m*M* imidazole, 500 m*M* NaCl, which was then washed with six column volumes of buffer consisting of 80 m*M* imidazole, 500 m*M* NaCl, 20 m*M* Tris pH 7.4. The Rv0880 protein was then eluted with two column volumes of buffer consisting of 300 m*M* imidazole, 500 m*M* NaCl, 20 m*M* Tris pH 7.4. The eluted protein was concentrated to 0.5 ml using an Amicon Ultra-4 10K centrifugal filter device (Millipore) and loaded onto a Superdex 200 (GE Healthcare) size-exclusion column pre-equilibrated with 20 m*M* Tris pH 7.4, 150 m*M* NaCl, 5% glycerol at a flow rate of 0.5 ml min^{-1}. Fractions containing the pure protein were pooled and concentrated to 12 mg ml^{-1}.

Table 2
Crystallization.

Method	Hanging-drop vapour diffusion
Plate type	16-well hanging-drop plate
Temperature (K)	289
Protein concentration (mg ml^{-1})	12
Buffer composition of protein solution	20 m*M* Tris pH 7.4, 150 m*M* NaCl, 5% glycerol
Composition of reservoir solution	0.1 *M* HEPES, 5%(*v*/*v*) (±)-2-methyl-2,4-pentanediol, 12%(*w*/*v*) PEG 6000
Volume and ratio of drop	1.0 µl; 1:1
Volume of reservoir (ml)	0.2

After checking the purity by SDS–PAGE analysis, purified Rv0880 was flash-frozen in liquid nitrogen and stored at −80°C. Macromolecule-production information is summarized in Table 1.

2.2. Crystallization

Purified Rv0880 protein at a concentration of 12 mg ml^{-1} in 20 m*M* Tris pH 7.4, 150 m*M* NaCl, 5% glycerol was used in all crystallization experiments. Crystallization experiments were performed at 16°C using the hanging-drop vapour-diffusion method. Crystallization screening kits from Hampton Research (Index, Crystal Screen and Crystal Screen 2) were used to determine the initial crystallization conditions for the Rv0880 protein. Typically, 1 µl drops of protein solution were mixed with an equal volume of screening solution and equilibrated over a reservoir containing 0.2 ml reservoir solution. Crystallization information is summarized in Table 2.

2.3. Data collection and processing

Native Rv0880 crystals were soaked for 30 s in reservoir solution supplemented with 15% glycerol and flash-cooled in a stream of cold nitrogen at 100 K. X-ray diffraction data were collected at an in-house data-collection facility using a Rigaku R-AXIS IV^{++} image plate and a Rigaku MicroMax-007 copper rotating-anode generator operating at 60 mA and 45 kV. A total of 360 frames were collected with an oscillation step of 1° and an exposure of 3 min per frame. The crystal-to-detector distance was maintained at 150 mm. Diffraction data were processed with *iMosflm* (Battye *et al.*, 2011; Powell *et al.*, 2013) and scaled with *SCALA* (Winn *et al.*, 2011; Cowtan *et al.*, 2011).

2.4. Structure solution and refinement

The structure of Rv0880 was solved by molecular replacement with *Rosetta* (Terwilliger *et al.*, 2012) from the *PHENIX* package (Adams *et al.*, 2002, 2010) using the structure of a putative transcriptional regulator protein from *Pseudomonas aeruginosa* PA01 (PDB entry 2hr3; Midwest Center for Structural Genomics, unpublished work) as a search model. Model building and structural refinement were performed with *Coot* (Emsley & Cowtan, 2004) and *PHENIX*, respectively. Data-collection and refinement statistics are summarized in Tables 3 and 4. All structural figures were rendered in *PyMOL* (http://www.pymol.org).

3. Results and discussion

The Rv0880 gene is 432 bp in length and codes for 143 amino-acid residues. The isoelectric point of the protein is calculated to be 10.9. Purified Rv0880 gave a band of approximately 16 kDa on SDS–PAGE (Fig. 1), which is in good agreement with the calculated molecular weight.

Crystals grew after 7 d at 16°C in Crystal Screen 2 solution No. 30 [0.1 *M* HEPES pH 7.5, 5%(*v*/*v*) (±)-2-methyl-2,4-pentanediol, 10%(*w*/*v*) PEG 6000]. To optimize the crystal, a systematic grid refinement was employed by varying the concentration of the precipitant (PEG 6000) from 5 to 15% in 1% steps and the pH value of the reservoir solution from 7.0 to 8.0 in steps of 0.1. The best crystals were obtained in a final optimized reservoir solution consisiting of 0.1 *M* HEPES pH 7.2, 5%(*v*/*v*) (±)-2-methyl-2,4-pentanediol, 12%(*w*/*v*) PEG 6000.

X-ray diffraction data were indexed and integrated using *iMosflm* (Battye *et al.*, 2011), giving a data set to 2.0 Å resolution that had a completeness of 94.3% and an overall $R_{\rm merge}$ of 7.1%. The crystals belonged to space group *P*1, as determined using *phenix.xtriage* (Terwilliger *et al.*, 2012), and six molecules were present in the unit cell. The solvent content of the Rv0880 crystal was 44.5%, and the calculated Matthews value was 2.2 Å^3 Da^{-1}.

Like other MarR proteins, Rv0880 is present as a dimer and belongs to the α/β family of proteins. It consists of six α-helices and two β-strands arranged in the order $\alpha1$–$\alpha2$–$\alpha3$–$\alpha4$–$\beta1$–$\beta2$–$\alpha5$–$\alpha6$ in the primary structure. Each subunit is composed of two functional domains: the dimerization domain, which includes helices $\alpha1$, $\alpha5$ and $\alpha6$, and the DNA-binding domain, which includes $\alpha2$, $\alpha3$, $\alpha4$, $\beta1$ and $\beta2$. The N- and C-termini of each monomer, located in the $\alpha1$ and $\alpha6$ helices, respectively, are closely intertwined and form the dimer interface, which is stabilized by hydrophobic and hydrogen-bond interactions between residues located within these regions. Residues

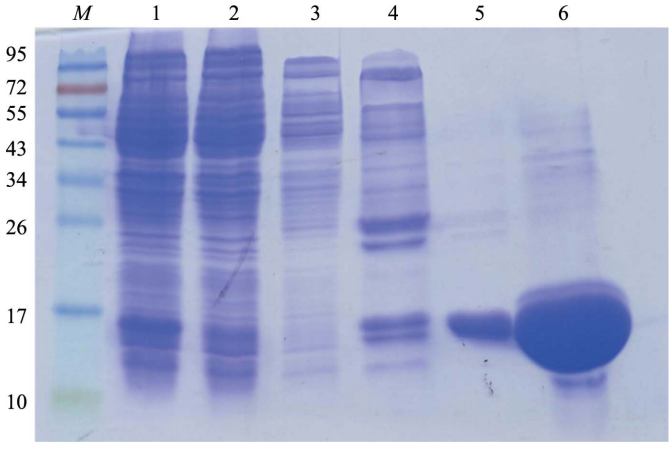


Figure 1
Purification of Rv0880. Monitoring protein purification using SDS–PAGE. Lane *M*, molecular-weight markers (labelled in kDa); lane 1, whole cell lysate after IPTG induction of protein expression; lane 2, flowthrough after the cell lysate was loaded onto an Ni^{2+}-chelating column; lanes 3, 4, 5 and 6, flowthrough after elution with buffer containing 40, 80, 120 or 300 m*M* imidazole, respectively.

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	Rigaku MicroMax-007
Wavelength (Å)	1.5418
Temperature (K)	100
Detector	R-AXIS IV^{++} image plate
Crystal-to-detector distance (mm)	150
Rotation range per image (°)	1
Total rotation range (°)	360
Exposure time per image (s)	180
Space group	*P*1
a, *b*, *c* (Å)	54.97, 69.60, 70.32
α, β, γ (°)	103.71, 111.06, 105.83
Mosaicity (°)	0.5
Resolution range (Å)	41.15–2.00 (2.11–2.00)
Total No. of reflections	218999 (30572)
No. of unique reflections	55124 (7737)
Completeness (%)	94.3 (90.3)
Multiplicity	4.0 (4.0)
$\langle I/\sigma(I)\rangle$	12.3 (2.4)
$R_{\rm r.i.m.}$	0.071 (0.543)
Overall *B* factor from Wilson plot (Å^2)	29.7

Table 4
Structure solution and refinement.

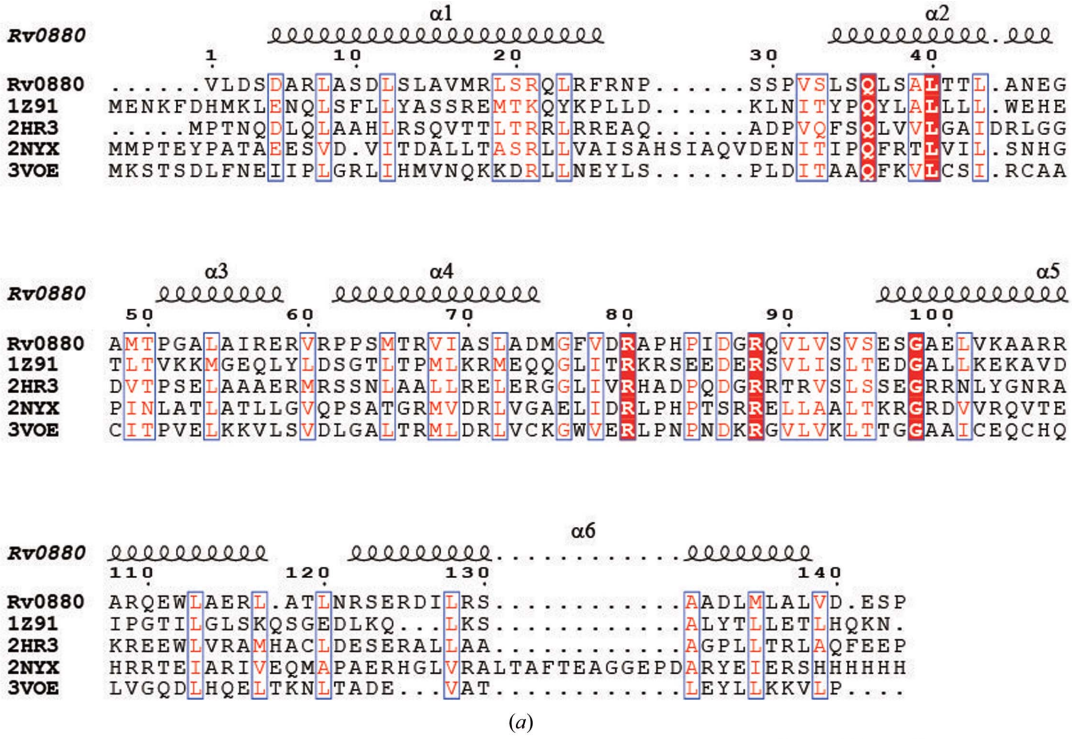
Values in parentheses are for the outer shell.

Resolution range (Å)	37.108–2.002
Completeness (%)	94.01
σ Cutoff	1.96
No. of reflections, working set	52178
No. of reflections, test set	2793
Final $R_{\rm cryst}$	0.212
Final $R_{\rm free}$	0.249
No. of non-H atoms	
Protein	6181
Ligand	0
Water	529
Total	6710
R.m.s. deviations	
Bonds (Å)	0.006
Angles (°)	1.222
Average *B* factors (Å^2)	41.7
Protein	41.7
Water	41.0
Ramachandran plot	
Most favoured (%)	98.0
Allowed (%)	1.0
Disallowed (%)	1.0

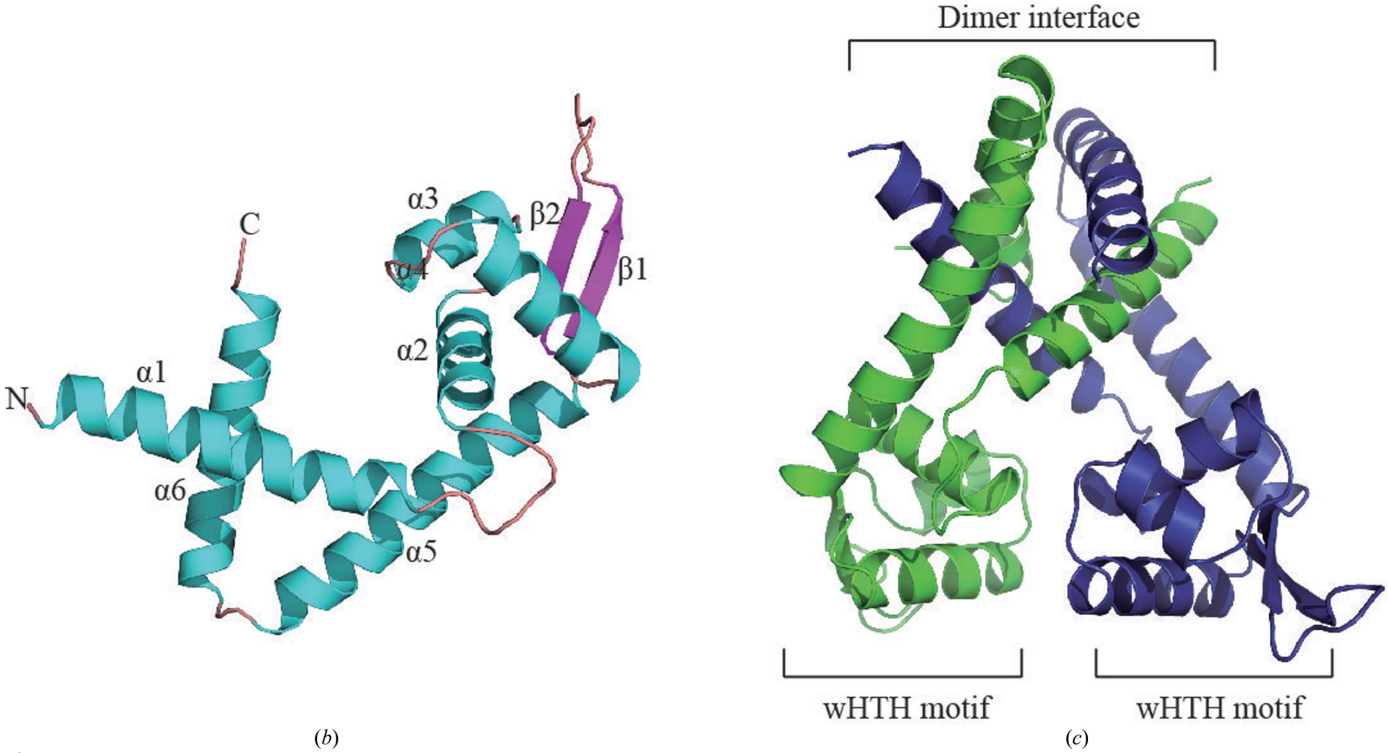
located within the $\alpha2$–$\alpha3$–$\alpha4$–$\beta1$–$\beta2$ structure form a wHTH DNA-binding motif (Fig. 2*c*) which is distal to the dimer interface.

Although Rv0880 shares less than 20% sequence identity with other MarR family proteins (Fig. 2*a*), it possesses the same core fold as other members of this family of regulators. While the structure of these proteins is quite similar, their dimerization is quite different. For example, when the structure of Rv0880 is superimposed onto the MarR family protein from *E. coli* (PDB entry 3voe; H. Lou, R. Zhu & Z. Hao, unpublished work; Fig. 3), it matches it closely with an r.m.s.d. of 2.5 Å. However, the orientation of the monomers in these two dimers is different. Analysis of the dimer interface using *PISA* (Krissinel & Henrick, 2007) showed that in PDB entry 3voe the interface, which contains eight hydrogen bonds and

four salt bridges, buries a total area of 3819 Å^2 and encompasses 46 residues from monomer *A* and 49 residues from monomer *B*. In Rv0880, the total area of the interface, which contains 22 hydrogen bonds and 15 salt bridges, is 5477 Å^2 and encompasses 62 residues from monomer *A* and 63 residues from monomer *B*.

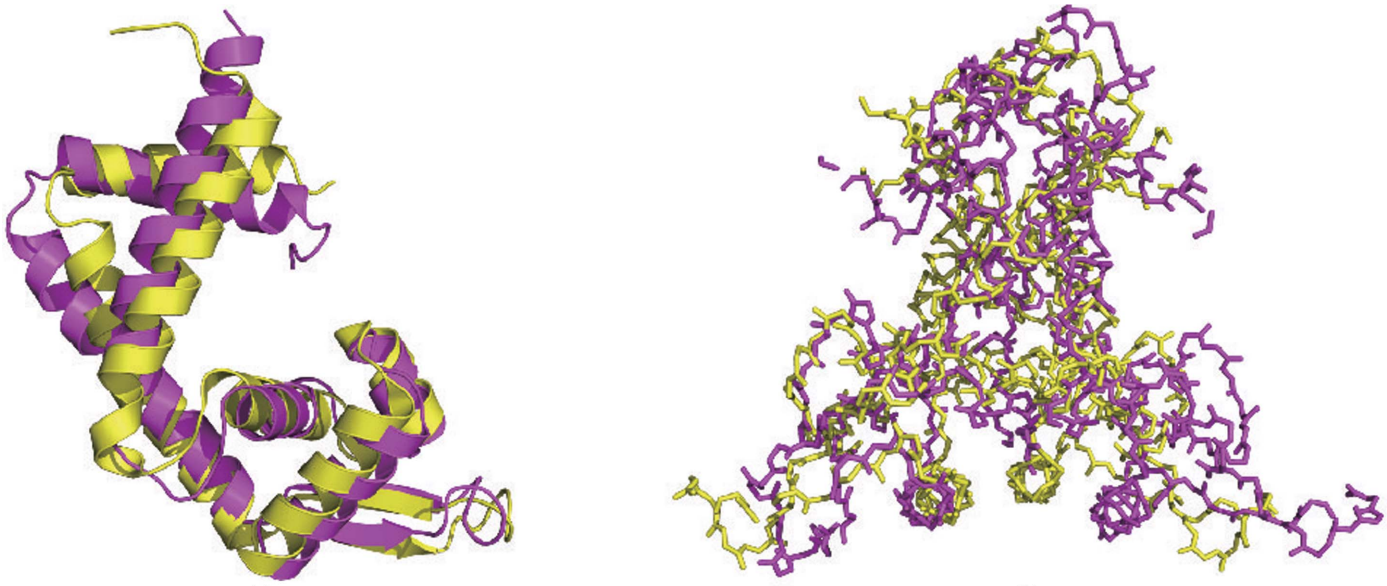


(*a*)



(*b*) (*c*)

Figure 2
Structure of Rv0880. (*a*) Primary sequence alignment of Rv0880 with other structurally characterized MarR family members from *B. subtilis* (PDB entry 1z91; Hong *et al.*, 2005), *P. aeruginosa* (PDB entry 2hr3; Midwest Center for Structural Genomics, unpublished work), *M. tuberculosis* (Rv1404; PDB entry 2nyx; TB Structural Genomics Consortium/Integrated Center for Structure and Function Innovation, unpublished work) and *E. coli* K-12 (PDB entry 3voe; H. Lou, R. Zhu & Z. Hao, unpublished work). The multiple sequence alignment was made with *ClustalW* (Thompson *et al.*, 2002) and displayed with secondary structures using *ESPript* (Gouet *et al.*, 1999). (*b*) Ribbon diagram of the Rv0880 subunit. Secondary-structural elements are labelled and the N- and C-termini are labelled N and C, respectively. (*c*) Ribbon diagram of the Rv0880 dimer.



(*a*) (*b*)

Figure 3
Structural comparison of Rv0880 (magenta) with the MarR family protein from *E. coli* (PDB entry 3voe; yellow). (*a*) Superposition of the Rv0880 monomer with PDB entry 3voe. (*b*) Superposition of the Rv0880 dimer with PDB entry 3voe.

The structural information on Rv0880 that we have obtained supports its annotation as a MarR family protein. It will be interesting to investigate the functional mechanism of this protein and to determine whether it plays a role in drug resistance in *M. tuberculosis*.

Acknowledgements

We would like to gratefully acknowledge the help of the staff of the Protein Science Core Facility Center at the Institute of Biophysics, Chinese Academy of Science. Financial support for this project was provided by the Chinese Ministry of Science and Technology 973 program (2011CB910302).

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Cowtan, K., Emsley, P. & Wilson, K. S. (2011). *Acta Cryst.* D**67**, 233–234.
Eiamphungporn, W., Soonsanga, S., Lee, J. W. & Helmann, J. D. (2009). *Nucleic Acids Res.* **37**, 1174–1181.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.
Grove, A. (2013). *Curr. Biol.* **23**, R142–R143.
Hong, M., Fuangthong, M., Helmann, J. D. & Brennan, R. G. (2005). *Mol. Cell*, **20**, 131–141.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Lomovskaya, O., Lewis, K. & Matin, A. (1995). *J. Bacteriol.* **177**, 2328–2334.
Perera, I. C. & Grove, A. (2010). *J. Mol. Cell Biol.* **2**, 243–254.
Powell, H. R., Johnson, O. & Leslie, A. G. W. (2013). *Acta Cryst.* D**69**, 1195–1203.
Soonsanga, S., Lee, J. W. & Helmann, J. D. (2008). *Mol. Microbiol.* **68**, 978–986.
Terwilliger, T. C., DiMaio, F., Read, R. J., Baker, D., Bunkóczi, G., Adams, P. D., Grosse-Kunstleve, R. W., Afonine, P. V. & Echols, N. (2012). *J. Struct. Funct. Genomics*, **13**, 81–90.
Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2002). *Curr. Protoc. Bioinformatics*, Unit 2.3. doi:10.1002/0471250953.bi0203s00.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
World Health Organization (2014). *Global Tuberculosis Report 2014*. Geneva: World Health Organization. http://www.who.int/tb/publications/global_report/en/.
Yu, L., Fang, J. & Wei, Y. (2009). *Biochemistry*, **48**, 2099–2108.