Transcription factor Rv0081 from *Mycobacterium tuberculosis*: purification, crystallization and initial crystallographic analysis

Shishang Dong,^{a,b} Zhenzhen Ding,^{b,c} Yu Wang,^{b,c} Yan Yang,^{b,c} Yonghong Mao^{b*} and Ying Wang^{b*}

Received 26 January 2017

Accepted 2 April 2017

Edited by R. L. Stanfield, The Scripps Research Institute, USA

Keywords: Rv0081; transcription factor; *Mycobacterium tuberculosis*.

^aCollege of Life Sciences, Nankai University, Tianjin 300071, People's Republic of China, ^bTianjin International Joint Academy of Biotechnology and Medicine, Tianjin 300457, People's Republic of China, and ^cCollege of Pharmacy, Nankai University, Tianjin 300071, People's Republic of China. *Correspondence e-mail: yonghong_87@126.com, ying.wang@htmdc.org

Because of its high infectivity and pathogenicity, *Mycobacterium tuberculosis* is a serious threat to human health. While the transcription-regulatory system of *M. tuberculosis* remains incompletely understood, Rv0081, an essential regulatory hub, is known to mediate the initial response to hypoxia in the long-term survival of *M. tuberculosis*. Here, the production, crystallization and initial X-ray crystallographic analysis of Rv0081 are reported. The crystals of Rv0081 belonged to space group $P6_2$, with unit-cell parameters $a = 67.48$, $b = 67.48$, $c = 40.84$ Å, $\gamma = 120^\circ$. The Matthews coefficient is $2.09 \text{ Å}^3 \text{ Da}^{-1}$, assuming the presence of one molecule in the asymmetric unit, with a corresponding solvent content of 41.27%. Phasing of the native crystal form of Rv0081 was performed by molecular replacement. Currently, the structure has been refined to 2.00 Å resolution with an R_{work} of 25.99% and an R_{free} of 30.88%.

1. Introduction

Tuberculosis (TB) has become one of the top ten causes of death worldwide. In 2015, the World Health Organization reported that 10.4 million people fell ill with TB and 1.8 million people died from the disease (including 0.4 million people with HIV; World Health Organization, 2015). Its ability to persist in infected individuals without producing any symptoms for very long periods of time makes *Mycobacterium tuberculosis* (MTB) an extremely successful pathogen. Therefore, it is important to understand the mechanisms that regulate the survival of latent MTB.

Latent MTB can adapt to potentially harmful stress within the host by regulating specific gene expression. Transcriptome analysis has identified critical regulators of gene expression, such as the cholesterol regulator KstR (Kendall *et al.*, 2007), the virulence lipid anabolism regulator WhiB3 (Singh *et al.*, 2009) and the well known hypoxia regulator DosR (Voskuil *et al.*, 2004; Park *et al.*, 2003). While the regulatory mechanisms underlying these and other adaptations are largely unknown, adaptations to hypoxia play a prominent role in the survival and pathogenesis of MTB (Galagan *et al.*, 2013).

The locus *Rv0081–Rv0088* encodes the components of a predicted formate hydrogenylase enzyme complex that is directly regulated by MprA and DosR/DevR. Rv0081, the product of the first gene in this operon, negatively regulates its own expression by binding to an inverted repeat element in the upstream region of the locus (He *et al.*, 2011). An MTB

© 2017 International Union of Crystallography

Table 1
Macromolecule-production information.

Source organism	<i>M. tuberculosis</i> H37Rv
DNA source	Genome
Forward primer†	CGCGGATCCATGGAGTCCGAACCGCTGTAC
Reverse primer‡	CCGCTCGAGTTACGTGGCCGACCGCCGGC
Cloning vector	pET-28a-SUMO derived from pET-28a
Expression vector	The same as the cloning vector
Expression host	<i>E. coli</i> BL21(DE3)
Complete amino-acid sequence of the construct produced§	MGSSHHHHHSSGLVPRGSHMASMSDSEVNQEAKE- PEVKPEVKPETHINLKVSDDGSEIFFKIKKTT- PLRLMEAFKAKRQKEMDSLRLFLYDGIRIQAD- QTPEDLDMEDNDIIEAHREQIGGSEMESEPLYK- LKAFFKTLAHPARIRILELLVERDRSVGELL- SSDVGLESSLNSQQLGVLRAGVVAARRDGN- MIYSTIAPDIAELLAVARKVLARVLSRVAVL- EDLRAGGSAT
No. of amino acids	114
Molecular weight (Da)	12356.32

† The BamHI site is underlined. ‡ The XhoI site is underlined. § The tag residues removed by the SUMO protease are shown in italics. The residues of native Rv0081 are underlined.

regulatory network study reported that Rv0081 is the only regulator in the initial hypoxic response apart from DosR. Rv0081, which is linked to lipid metabolism, cholesterol metabolism, protein degradation and the initial hypoxic response, forms the largest hub and interacts with other hubs in the network model (Galagan *et al.*, 2013). Additionally, some experiments have verified Rv0081 as an important transcription factor that may be involved in the regulation of hypoxia (Ruan *et al.*, 2016). Azizi and coworkers illustrated that Rv0081 and DosR are important in the formation of feed-forward loops (FFL), mediating the induction of a second hierarchy of regulators with a time delay and leading to a later hypoxic response (Azizi *et al.*, 2014).

Owing to social systems, levels of economic development and the lack of an effective vaccine for TB (Ottenhoff & Kaufmann, 2012), diagnosis of TB remains a challenge in resource-constrained areas. After stimulating whole blood in pulmonary TB patients and household contacts (HHCs) with Rv0081, some specific immune response cytokines (such as IL-12, IP-10, IL-10 and TNF- α) show discriminating levels between TB cases and HHCs. Rv0081-specific levels of specific immune response cytokines were the most promising candidates for TB diagnosis (Chegou *et al.*, 2012). The expression levels of IL-6, IL-10 and TNF- α were distinct after stimulating whole blood with four latency antigens (Rv0081, Rv2629, Rv2006 and Rv1733c) in pulmonary TB patients (Kassa *et al.*, 2012), whereas on stimulation with Rv0081 in TB and non-TB cases only tissue plasminogen activator (tPA) responses were significantly different (Awoniyi *et al.*, 2016). Rv0081-specific levels of specific immune response cytokines require further research as promising candidates for TB diagnosis.

Currently, the function of Rv0081 is largely unknown in the context of *M. tuberculosis* gene regulation, pathogenesis and physiology, and no structure of Rv0081 is available in the Protein Data Bank (PDB). The protein is predicted to be a member of the helix–turn–helix (HTH) superfamily according to the NCBI database (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Here, we present the cloning, expression, purification,

crystallization and X-ray diffraction analysis of Rv0081 from *M. tuberculosis* H37Rv.

2. Materials and methods

2.1. Macromolecule production

Wild-type cDNA for Rv0081 was amplified by PCR using the *M. tuberculosis* H37Rv genome as a template and was cloned into the plasmid pET-28a-SUMO derived from pET-28a (Novagen). The recombinant plasmid for Rv0081 was transformed into *Escherichia coli* strain BL21 (DE3) (TransGen Biotech, People’s Republic of China) and over-expressed. Freshly transformed cells were cultured at 310 K in 800 ml LB medium containing 50 mg l⁻¹ kanamycin. Once the OD₆₀₀ had reached 0.5, the culture was transferred to 289 K. Protein expression was then induced by the addition of 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) for 17 h. The cells were harvested by centrifugation, resuspended in 10 ml lysis buffer (20 mM HEPES, 500 mM NaCl, 5% glycerol pH 7.0) per gram of pellet and frozen at 253 K. The cells were disrupted using a low-temperature ultrahigh-pressure cell disrupter (JNBIO, People’s Republic of China). The lysate was centrifuged at 12 000 rev min⁻¹ for 30 min at 277 K to remove cell debris. The supernatant was then loaded three times onto a Ni-NTA column (GE Healthcare, USA) pre-equilibrated with lysis buffer. The resin was washed four times with 60 ml lysis buffer. The 6 $\times$ His-SUMO tag was cleaved by incubating the protein with SUMO protease overnight at 277 K. Recombinant Rv0081 was eluted with lysis buffer and buffer-exchanged into initial ion-exchange buffer (20 mM HEPES, 40 mM NaCl pH 7.0) using an Amicon Ultra-15 centrifugal filter unit (Millipore; 10 kDa molecular-weight cutoff). The sample was then further purified using a Resource S column (1 ml; GE Healthcare, USA). The purified protein was exchanged into low-salt buffer (20 mM HEPES, 200 mM

Figure 1
Chromatogram representing the purification of Rv0081 using a Superdex 200 column (GE Healthcare, USA). The 15% SDS-PAGE gel stained with Coomassie Brilliant Blue dye shows purified Rv0081.

Table 2
Crystallization.

Method	Sitting-drop vapour diffusion (screening), hanging-drop vapour diffusion (production)
Plate type	48-well plates (screening), 16-well plates (production)
Temperature (K)	289
Protein concentration (mg ml ⁻¹)	3
Buffer composition of protein solution	20 mM HEPES, 200 mM NaCl, 5 mM DTT pH 7.0
Composition of reservoir solution	200 mM zinc acetate dehydrate, 100 mM sodium cacodylate trihydrate pH 6.5, 6% (w/v) PEG 8000
Volume and ratio of drop	2 µl; 1:1 ratio of protein:reservoir
Volume of reservoir (µl)	100 (screening), 200 (production)

NaCl, 5 mM DTT pH 7.0), concentrated to 10 mg ml⁻¹ and stored at 193 K. Protein purity was assessed by 15% SDS-PAGE and the protein homogeneity was analysed by size-exclusion chromatography using a Superdex 200 10/300 GL column (GE Healthcare, USA; Fig. 1). Macromolecule-production information is summarized in Table 1.

2.2. Crystallization

Initial crystallization trials were performed in a 96-well format using a 1:1 ratio of well solution to protein solution at 6 mg ml⁻¹ by screening at 289 K with commercial crystal

(a)

(b)

Figure 2
Crystals of Rv0081 (a) before and (b) after optimization.

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	BL18U1, SSRF
Wavelength (Å)	0.97778
Temperature (K)	100
Detector	PILATUS CBF
Crystal-to-detector distance (mm)	350.00
Rotation range per image (°)	0.5
Total rotation range (°)	100
Exposure time per image (s)	0.2
Space group	<i>P</i> 6 ₂
<i>a</i> , <i>b</i> , <i>c</i> (Å)	67.48, 67.48, 40.84
α , β , γ (°)	90, 90, 120
Mosaicity (°)	0.59
Resolution range (Å)	50.00–2.00 (2.03–2.00)
Total No. of reflections	35366
No. of unique reflections	7096 (377)
Completeness (%)	97.3 (100.0)
Multiplicity	5.0 (4.8)
$\langle I/\sigma(I) \rangle$	14.50 (2.48)
<i>R</i> _{r.i.m.}	0.076 (0.655)
<i>R</i> _{pim}	0.038 (0.295)
<i>R</i> _{merge}	0.065 (0.580)
Overall <i>B</i> factor from Wilson plot (Å ²)	18.69

screening kits, including Index, Crystal Screen and PEG/Ion from Hampton Research. Small crystals of Rv0081 (Fig. 2*a*) first appeared after 2 d in 200 mM zinc acetate dehydrate, 100 mM sodium cacodylate trihydrate pH 6.5, 18% (w/v) PEG 8000. Further optimization by adjusting the pH and the concentrations of PEG 8000 and protein was performed, and the final optimized crystallization conditions were 200 mM zinc acetate dehydrate, 100 mM sodium cacodylate trihydrate pH 6.5, 6% (w/v) PEG 8000 with 3 mg ml⁻¹ Rv0081 using the hanging-drop method (Fig. 2*b*). Crystallization information is summarized in Table 2.

2.3. Data collection and processing

Crystals were flash-cooled in liquid nitrogen with reservoir solution containing 20% glycerol as a cryoprotectant. A data set at a resolution of 2.00 Å was collected on BL18U1 at Shanghai Synchrotron Radiation Facility (SSRF). Diffraction data sets were indexed, integrated and scaled using the *HKL*-3000 suite (Minor *et al.*, 2006). Data-collection and processing statistics are shown in Table 3.

3. Results and discussion

Recombinant Rv0081 was expressed and purified from *E. coli* BL21 (DE3) cells using a two-step purification protocol (affinity and ion-exchange chromatography). Protein purity was assessed by 15% SDS-PAGE and protein homogeneity was analysed by size-exclusion chromatography. Rv0081 eluted from the Superdex 200 column at 15.9 ml, which is between the elution volumes of β -lactoglobulin (molecular mass 35 000 Da) and ribonuclease A (molecular mass 13 700 Da), indicating that Rv0081 is most likely to be a dimer in solution (Fig. 1). Initial small crystals were observed in several conditions. After optimization, diffraction-quality crystals (Fig. 2*b*) of Rv0081 were obtained in 200 mM zinc

acetate dehydrate, 100 mM sodium cacodylate trihydrate pH 6.5, 6% (w/v) PEG 8000 by the hanging-drop method. To verify that the protein had indeed crystallized, several crystals were looped out and washed with the same reservoir. These crystals were then dissolved in low-salt buffer and subjected to 15% SDS-PAGE analysis with Coomassie Brilliant Blue staining, which showed one band corresponding to Rv0081 (Fig. 3).

The best diffraction-quality crystals diffracted to a resolution of 2.00 Å (Fig. 4) on BL18U1 at SSRF and belonged to space group $P6_2$, with unit-cell parameters $a = 67.48$, $b = 67.48$,

Figure 3
15% SDS-PAGE analysis of Rv0081 crystals. The left lane contains stained protein molecular-mass marker (TransGen Biotech, People's Republic of China; molecular masses are labelled in kDa).

Figure 4
X-ray diffraction image from an Rv0081 crystal.

$c = 40.84$ Å, $\gamma = 120^\circ$. Analysis of the crystal content using programs from the CCP4 suite (Winn *et al.*, 2011) indicated that there was one molecule in the asymmetric unit, with a Matthews coefficient of 2.09 Å³ Da⁻¹ and a corresponding solvent content of 41.27%. When the sequence was submitted to a BLASTP protein database search at NCBI, no significant similarity was found in the results. However, Phyre² (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) predicted the three-dimensional structure to be extremely similar to that of the ArsR-like transcriptional regulators. We are currently solving the crystal structure of Rv0081 by molecular replacement (MR) using the truncated monomer structure of SmtB from *Synechococcus elongatus* (PDB entry 1r22; Eicken *et al.*, 2003) as a model with programs from the CCP4 suite (Winn *et al.*, 2011). After running Phaser-MR, we replaced the model sequence with the object sequence and built the C-terminal amino acids of Rv0081 with 'autobuild/refine' in Buccaneer. Initial refinement resulted in an R_{work} and R_{free} of 32.78 and 43.94%, respectively. The Rv0081 structure was further manually built into the modified experimental electron density using Coot (Emsley & Cowtan, 2004) and refined in PHENIX (Adams *et al.*, 2010). Currently, the structure has been refined to 2.00 Å resolution with an R_{work} of 25.99% and an R_{free} of 30.88%. Further refinement and model building of the structure is in progress.

Acknowledgements

We are grateful to the beamline staff at Shanghai Synchrotron Radiation Facility. We are also grateful to Zuokun Lu and Shiyu Che for their help with data collection and processing.

Funding information

Funding for this research was provided by: National Natural Science Foundation of China (award No. 31670731); Ministry of Science and Technology of China '973' Project (award Nos. 2014CB542800, 2013CB911103); Tianjin Municipal Natural Science Foundation (award Nos. 14JCQNJC10100, 14ZCZDSY00039, 13ZCZDSY04200).

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Awoniyi, D. O. *et al.* (2016). *J. Infect.* **73**, 219–230.
- Azizi, E., Airolidi, E. M. & Galagan, J. E. (2014). *JMLR Workshop Conf. Proc.* **32**, 1440–1448.
- Chegou, N. N., Essone, P. N., Loxton, A. G., Stanley, K., Black, G. F., van der Spuy, G. D., van Helden, P. D., Franken, K. L., Parida, S. K., Klein, M. R., Kaufmann, S. H., Ottenhoff, T. H. M. & Walzl, G. (2012). *PLoS One*, **7**, e38501.
- Eicken, C., Pennella, M. A., Chen, X., Koshlap, K. M., VanZile, M. L., Sacchettini, J. C. & Giedroc, D. P. (2003). *J. Mol. Biol.* **333**, 683–695.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Galagan, J. E. *et al.* (2013). *Nature (London)*, **499**, 178–183.
- He, H., Bretl, D. J., Penoske, R. M., Anderson, D. M. & Zahrt, T. C. (2011). *J. Bacteriol.* **193**, 5105–5118.
- Kassa, D., Ran, L., Geberemeskel, W., Tebeje, M., Alemu, A., Selase, A., Tegbaru, B., Franken, K. L., Friggen, A. H., van Meijgaarden, K. E., Ottenhoff, T. H., Wolday, D., Messele, T. & van Baarle, D. (2012). *Clin. Vaccine Immunol.* **19**, 1907–1915.

- Kendall, S. L., Withers, M., Soffair, C. N., Moreland, N. J., Gurcha, S., Sidders, B., Frita, R., Ten Bokum, A., Besra, G. S., Lott, J. S. & Stoker, N. G. (2007). *Mol. Microbiol.* **65**, 684–699.
- Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. (2006). *Acta Cryst. D* **62**, 859–866.
- Ottenhoff, T. H. & Kaufmann, S. H. (2012). *PLoS Pathog.* **8**, e1002607.
- Park, H.-D., Guinn, K. M., Harrell, M. I., Liao, R., Voskuil, M. I., Tompa, M., Schoolnik, G. K. & Sherman, D. R. (2003). *Mol. Microbiol.* **48**, 833–843.
- Ruan, J., Wu, W., Sun, X., Gao, Y., Jiang, J., Cui, Z., Ge, B., Wu, H., Zhang, L. & Li, Y. (2016). *PLoS One*, **11**, e0148965.
- Singh, A., Crossman, D. K., Mai, D., Guidry, L., Voskuil, M. I., Renfrow, M. B. & Steyn, A. J. (2009). *PLoS Pathog.* **5**, e1000545.
- Voskuil, M. I., Visconti, K. C. & Schoolnik, G. K. (2004). *Tuberculosis*, **84**, 218–227.
- Winn, M. D. *et al.* (2011). *Acta Cryst. D* **67**, 235–242.
- World Health Organization (2015). *Global Tuberculosis Report 2015*. Geneva: World Health Organization. http://www.who.int/tb/publications/global_report/.