10-1-2012

Structural and functional analysis of the transcriptional regulator Rv3066 of Mycobacterium tuberculosis

Jani Reddy Bolla
Iowa State University

Sylvia V. Do
Iowa State University

Feng Long
Iowa State University

Lei Dai
Iowa State University, ldai@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/vmpm_pubs

Part of the Bioinformatics Commons, Biological and Chemical Physics Commons, Chemistry Commons, Computational Biology Commons, and the Veterinary Microbiology and Immunobiology Commons

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/vmpm_pubs/143. For information on how to cite this item, please visit http://lib.dr.iastate.edu/howtocite.html.
Structural and functional analysis of the transcriptional regulator Rv3066 of Mycobacterium tuberculosis

Abstract
The Mmr multidrug efflux pump recognizes and actively extrudes a broad range of antimicrobial agents, and promotes the intrinsic resistance to these antimicrobials in Mycobacterium tuberculosis. The expression of Mmr is controlled by the TetR-like transcriptional regulator Rv3066, whose open reading frame is located downstream of the mmr operon. To understand the structural basis of Rv3066 regulation, we have determined the crystal structures of Rv3066, both in the absence and presence of bound ethidium, revealing an asymmetric homodimeric two-domain molecule with an entirely helical architecture. The structures underscore the flexibility and plasticity of the regulator essential for multidrug recognition. Comparison of the apo-Rv3066 and Rv3066–ethidium crystal structures suggests that the conformational changes leading to drug-mediated derepression is primarily due to a rigid body rotational motion within the dimer interface of the regulator. The Rv3066 regulator creates a multidrug-binding pocket, which contains five aromatic residues. The bound ethidium is found buried within the multidrug-binding site, where extensive aromatic stacking interactions seemingly govern the binding. In vitro studies reveal that the dimeric Rv3066 regulator binds to a 14-bp palindromic inverted repeat sequence in the nanomolar range. These findings provide new insight into the mechanisms of ligand binding and Rv3066 regulation.

Disciplines
Bioinformatics | Biological and Chemical Physics | Chemistry | Computational Biology | Veterinary Microbiology and Immunobiology

Comments

Creative Commons License
This work is licensed under a Creative Commons Attribution-Noncommercial 3.0 License

Authors
Jani Reddy Bolla, Sylvia V. Do, Feng Long, Lei Dai, Chih-Chia Su, Hsiang-Ting Lei, Xiao Chen, Jillian E. Gerkey, Daniel C. Murphy, Kanagalaghatta R. Rajashankar, Qijing Zhang, and Edward W. Yu

This article is available at Iowa State University Digital Repository: http://lib.dr.iastate.edu/vmpm_pubs/143
Structural and functional analysis of the transcriptional regulator Rv3066 of *Mycobacterium tuberculosis*

Jani Reddy Bolla, Sylvia V. Do, Feng Long, Lei Dai, Chih-Chia Su, Hsiang-Ting Lei, Xiao Chen, Jillian E. Gerkey, Daniel C. Murphy, Kanagalaghatta R. Rajashankar, Qijing Zhang and Edward W. Yu

1Department of Chemistry, 2Bioinformatics and Computational Biology Interdepartmental Graduate Program, 3Department of Veterinary Microbiology, College of Veterinary Medicine, Iowa State University, Ames, IA 50011, 4NE-CAT and Department of Chemistry and Chemical Biology, Cornell University, Argonne National Laboratory, Argonne, IL 60439 and 5Department of Physics and Astronomy, Iowa State University, Ames, IA 50011, USA

Received February 16, 2012; Revised June 18, 2012; Accepted June 19, 2012

ABSTRACT

The Mmr multidrug efflux pump recognizes and actively extrudes a broad range of antimicrobial agents, and promotes the intrinsic resistance to these antimicrobials in *Mycobacterium tuberculosis*. The expression of Mmr is controlled by the TetR-like transcriptional regulator Rv3066, whose open reading frame is located downstream of the mmr operon. To understand the structural basis of Rv3066 regulation, we have determined the crystal structures of Rv3066, both in the absence and presence of bound ethidium, revealing an asymmetric homodimeric two-domain molecule with an entirely helical architecture. The structures underscore the flexibility and plasticity of the regulator essential for multidrug recognition. Comparison of the apo-Rv3066 and Rv3066–ethidium crystal structures suggests that the conformational changes leading to drug-mediated derepression is primarily due to a rigid body rotational motion within the dimer interface of the regulator. The Rv3066 regulator creates a multidrug-binding pocket, which contains five aromatic residues. The bound ethidium is found buried within the multidrug-binding site, where extensive aromatic stacking interactions seemingly govern the binding. *In vitro* studies reveal that the dimeric Rv3066 regulator binds to a 14-bp palindromic inverted repeat sequence in the nanomolar range. These findings provide new insight into the mechanisms of ligand binding and Rv3066 regulation.

INTRODUCTION

Tuberculosis (TB) is one of the most deadly diseases and was responsible for the death of 1.7 million people in 2009 (1) (http://www.who.int/mediacentre/factsheets/fs104/en/index.html). This disease is caused by the bacterium *Mycobacterium tuberculosis*, which infects an estimated 9 million people each year. TB is very difficult to treat, requiring at least 6 months of a combination of medications. The treatment must continue even long after the symptoms disappear. However, *M. tuberculosis* has developed resistance to commonly used anti-TB agents, such as isoniazid and rifampicin. The development of these drug-resistant strains is mainly due to the mismatch between treatment and symptoms, such as the irregular intake of drugs throughout the course of treatment and inappropriate prescription of medications (2). TB caused by multidrug-resistant (MDR) (3,4), extensively drug-resistant (XDR) (2,5) and most recently totally drug-resistant (TDR) (6) strains of *M. tuberculosis* have emerged and spread globally. Based on our current knowledge, TDR-TB is untreatable. The World Health Organization has predicted that there will be 2 million MDR or XDR cases worldwide by 2012. It is obvious that the emergence of these drug-resistant TB strains has evolved into a major threat and challenges our global
prospects for TB control. Thus, knowledge of the molecular mechanisms underlying drug resistance in M. tuberculosis is essential for the development of new strategies to combat this disease.

Recent evidence suggests that MDR strains of M. tuberculosis are associated with constitutive or inducible expression of multidrug efflux pumps (7). These pumps have been classified into five different families: the ATP-binding cassette (ABC), resistance nodulation division (RND), multidrug and toxic compound extrusion (MATE), major facilitator (MF) and small multidrug resistance (SMR) families (8). It has been found that the genome of M. tuberculosis contains genes encoding efflux pumps from all these families (9,10). In addition, several of these MDR efflux pumps have been identified and characterized (9). One such pump is the Mmr (Rv3065) multidrug efflux pump, which belongs to the SMR family (11). Mmr has been shown to mediate resistance to several toxic compounds, including acriflavine, ethidium bromide, erythromycin, pyronin Y, safranin O, tetraphenylphosphonium and thiordiazine (11,12).

Elucidating the regulatory systems of multidrug efflux pumps in M. tuberculosis should allow us to understand how this bacterium contributes to multidrug resistance and how it adapts to environmental changes. At present, little is known about the regulatory mechanisms modulating the expression of mmr in M. tuberculosis. Here, we report the crystal structures of the Rv3066 efflux regulator both in the absence and presence of bound ethidium, suggesting that ethidium binding triggers a rotational motion of the regulator. This motion results in inducing the expression of the Mmr efflux pump by releasing the Rv3066 regulator from its cognate DNA at the promoter region. The rv3066 gene is located immediately downstream of mmr and encodes a 202 amino acid protein that shares sequence homology to members of the TetR family of regulators (13,14). Our data indicate that Rv3066 is a TetR-family regulator (15), which represses the transcription of mmr by directly binding to the inverted repeat (IR) of the promoter.

MATERIALS AND METHODS

Cloning of rv3066

The rv3066 ORF from genomic DNA of M. tuberculosis strain H37/Rv was amplified by polymerase chain reaction (PCR) using the primers 5'-CCATGGCAACCGCAGGC TCCGACC-3' and 5'-GGATCTCAATTGATGATGAT GATGATGTTCCGAGGTTTCCGCTCCGAT-3' to generate a product that encodes a Rv3066 recombinant protein with a 6xHis tag at the C-terminus. The corresponding PCR product was digested with NeoI and BamHI, extracted from the agarose gel and inserted into pET15b as described by the manufacturer (Merck KGaA, Darmstadt, Germany). The recombinant plasmid (pET15bΩrv3066) was transformed into DH5α cells and the transformants were selected on LB agar plates containing 100 µg/ml ampicillin. The presence of the correct rv3066 sequence in the plasmid construct was verified by DNA sequencing.

Expression and purification of Rv3066

Briefly, the full-length Rv3066 protein containing a 6xHis tag at the C-terminus was overproduced in Escherichia coli BL21(DE3) cells possessing pET15bΩrv3066. Cells were grown in 40 ml of Luria Broth (LB) medium with 100 µg/ml ampicillin at 37°C. When the OD600 reached 0.5, the culture was treated with 1 mM isopropyl-D-thiogalactopyranoside (IPTG) to induce Rv3066 expression, and cells were harvested within 3 h. The collected bacterial cells were suspended in 100 ml ice-cold buffer containing 20 mM Na-HEPES (pH 7.2) and 200 mM NaCl, 10 mM MgCl2 and 0.2 mg DNase I (Sigma-Aldrich). The cells were then lysed with a French pressure cell. Cell debris was removed by centrifugation at 45 min at 4°C and 20,000 rev/min. The crude lysate was filtered through a 0.2 µm membrane and was loaded onto a 5 ml Hi-Trap Ni32-chelating column (GE Healthcare Biosciences, Pittsburgh, PA, USA) pre-equilibrated with 20 mM Na-HEPES (pH 7.2) and 200 mM NaCl. To remove unbound proteins and impurities, the column was first washed with six column volumes of buffer containing 50 mM imidazole, 250 mM NaCl and 20 mM Na-HEPES (pH 7.2). The Rv3066 protein was then eluted with four column volumes of buffer containing 300 mM imidazole, 250 mM NaCl and 20 mM Na-HEPES (pH 7.2). The purity of the protein was judged using 12.5% sodium dodecyl sulfate–polyacrylamide gel electrophoresis stained with Coomassie Brilliant Blue. The purified protein was extensively dialyzed against buffer containing 100 mM imidazole, 250 mM NaCl and 20 mM Na-HEPES (pH 7.5) and concentrated to 20 mg/ml.

For 6xHis selenomethionyl-substituted (SeMet)-Rv3066 protein expression, a 2 ml LB broth overnight culture containing E. coli BL21(DE3)/pET15bΩrv3066 cells was transferred into 40 ml of LB broth containing 100 µg/ml ampicillin and grown at 37°C. When the OD600 value reached 1.2, cells were harvested by centrifugation at 6000 rev/min for 10 min, and then washed twice with 20 ml of M9 minimal salts solution. The cells were re-suspended in 40 ml of M9 media and then transferred into a 41 pre-warmed M9 solution containing 100 µg/ml ampicillin. The cell culture was incubated at 37°C with shaking. When the OD600 reached 0.4, 100 mg/l of lysine, phenylalanine and threonine, 50 mg/l isoleucine, leucine and valine, and 60 mg/l of l-selenomethionine were added. The protein expression was induced with 1 mM IPTG after 15 min. Cells were then harvested within 3 h after induction. The procedures for purifying the 6xHis SeMet-Rv3066 were identical to those of the native protein.

Crystallization of Rv3066

All crystals of the 6xHis Rv3066 regulator were obtained using hanging-drop vapor diffusion. The Form I SeMet-Rv3066 crystals were grown at room temperature in 24-well plates with the following procedures. A 2 µl protein solution containing 20 mg/ml SeMet-Rv3066 protein in 20 mM Na–HEPES (pH 7.5), 250 mM NaCl and 100 mM imidazole was mixed with a 2 µl...
of reservoir solution containing 24% polyethylene glycol (PEG) 4000, 0.1 M Na-acetate (pH 5.0) and 0.2 M MgCl₂. The resultant mixture was equilibrated against 500 μl of the reservoir solution. Crystals of Form I grew to a full size in the drops within 2 weeks. Typically, the dimensions of the crystals were 0.2 mm × 0.05 mm × 0.05 mm. Cryoprotection was achieved by raising the PEG 4000 concentration stepwise to 30% with a 3% increment in each step.

The Form II crystals of Rv3066 were prepared using similar procedures. The reservoir solution for the Form II crystals consists of 24% PEG 4000, 0.1 M Na–HEPES (pH 8.0) and 0.2 M MgCl₂. Crystals of this form grew to a full size in the drops within 2 weeks. The dimensions of the mature crystals were 0.1 mm × 0.1 mm × 0.1 mm. Cryoprotection was achieved by raising the PEG 4000 concentration stepwise to 30% with a 3% increment in each step.

The Rv3066–ethidium complex crystals were prepared by incubating the Form II crystals of apo-Rv3066 in solution containing 24% PEG 4000, 0.1 M Na–HEPES (pH 8.0), 0.2 M MgCl₂ and 0.5 mM ethidium bromide for 48 h at 25°C. Cryoprotection was achieved by raising the PEG 4000 concentration stepwise to 30% with a 3% increment in each step.

### Table 1. Data collection, phasing and structural refinement statistics

<table>
<thead>
<tr>
<th>Data Collection</th>
<th>Form I</th>
<th>Form II</th>
<th>Rv3066-ethidium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>0.978</td>
<td>0.978</td>
<td>0.978</td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁₂₁₂</td>
<td>P2₁₂₁₂</td>
<td>P3₁</td>
</tr>
<tr>
<td>Cell constants (Å)</td>
<td>a = 78.7, b = 118.9, c = 42.1</td>
<td>a = 91.4, b = 119.9, c = 30.9</td>
<td>a = 99.1, b = 99.1, c = 66.5</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>2.32 (40.00–2.32)</td>
<td>1.83 (40.00–1.83)</td>
<td>2.30 (40.0–2.30)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.9 (99.9)</td>
<td>96.0 (98.4)</td>
<td>100.0 (99.6)</td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>556100</td>
<td>524555</td>
<td>670642</td>
</tr>
<tr>
<td>Redundancy</td>
<td>2.4 (2.5)</td>
<td>2.7 (2.9)</td>
<td>3.1 (3.2)</td>
</tr>
<tr>
<td>Rmerge (%)</td>
<td>36.4 (7.7)</td>
<td>39.8 (4.9)</td>
<td>36.4 (3.7)</td>
</tr>
<tr>
<td>Resolution range of data used (Å)</td>
<td>40.00–2.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Figure of merit (acentric/centric)</td>
<td>56.1/42.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rwork (%)</td>
<td>20.28</td>
<td>20.40</td>
<td>20.59</td>
</tr>
<tr>
<td>Rfree (%)</td>
<td>27.06</td>
<td>24.71</td>
<td>26.16</td>
</tr>
<tr>
<td>B-factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall (Å²)</td>
<td>38.5</td>
<td>24.6</td>
<td>48.0</td>
</tr>
<tr>
<td>Protein chain A/B (Å²)</td>
<td>39.0/36.9</td>
<td>25.8/23.3</td>
<td>45.3/30.3</td>
</tr>
<tr>
<td>Ligand chain A/B (Å²)</td>
<td>–</td>
<td>–</td>
<td>37.0/36.6</td>
</tr>
<tr>
<td>Water (Å)</td>
<td>39.2</td>
<td>29.3</td>
<td>41.1</td>
</tr>
<tr>
<td>No. of atoms in protein chain A/B</td>
<td>1329/1329</td>
<td>1321/1321</td>
<td>1322/1329</td>
</tr>
<tr>
<td>No. of ligands</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>No. of waters</td>
<td>77</td>
<td>67</td>
<td>64</td>
</tr>
<tr>
<td>Rms deviations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.1</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>Bond length (Å)</td>
<td>0.008</td>
<td>0.007</td>
<td>0.008</td>
</tr>
<tr>
<td>Ramachandran analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most favored regions (%)</td>
<td>94.9</td>
<td>94.3</td>
<td>94.2</td>
</tr>
<tr>
<td>Allowed regions (%)</td>
<td>5.1</td>
<td>5.7</td>
<td>5.8</td>
</tr>
<tr>
<td>Generously allowed regions (%)</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Disallowed regions (%)</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

**Data collection, structural determination and refinement**

All diffraction data were collected at 100 K at beamline 24ID-E located at the Advanced Photon Source, using an ADSC Quantum 315 CCD-based detector. Diffraction data were processed using DENZO and scaled using SCALEPACK (16). The crystals of Form I belong to space group *P*₂₁₂₁₂ (Table 1). Based on the molecular weight of Rv3066 (22.78 kDa), a single dimer per asymmetric unit with a solvent content of 36.1% is expected. Two selenium sites were identified using SHELXC and SHELXD (17) as implemented in the HKL2MAP package (18). Single-wavelength anomalous dispersion (SAD) phasing using the program PHASER (19) was employed to obtain experimental phases in addition to phases from the structural model of the EbrR regulator (residues 83–176) (PDB code: 3hta). The resulting phases were then subjected to density modification and NCS averaging using the program PARROT (20). The phases were of excellent quality and allowed for tracing of most of the molecule in PHENIX AutoBuild (21), which led to an initial model containing 72% amino acid residues and 54% of which contained side chains. The remaining part of the model was manually constructed using the program.
Coot (22). Then, the model was refined using translation/libration/screw (TLS) refinement techniques adopting a single TLS body as implemented in PHENIX (21) leaving 5% of reflections in Free-R set. Iterations of refinement using PHENIX (21) and CNS (23) and model building in Coot (22) lead to the current model, which consists of 173 residues (residues 12–184) with excellent geometrical characteristics (Table 1).

The Form II and Rv3066-ethidium crystals took the space groups P2₁2₁2 and P3₁2₁, respectively. These two structures were phased using the molecular replacement (MR) program PHASER (19) by using the Form I structure as the search model. Structural refinements were then performed using PHENIX (21) and CNS (23) (Table 1).

**Electrophoretic mobility shift assay**

To determine the binding of Rv3066 to the operator region of *mmr*, electrophoretic mobility shift assay (EMSA) was performed according to the procedure described by Aleksey et al. (24). The 30-bp oligonucleotide, 5′-CGAGGCTCTTTGTGACATTTGTACATGT-3′, containing the hypothetical operator site IR1 was labeled at the 3′-end with digoxigenin-11-ddUTP (DIG-11-ddUTP) using the DIG Gel Shift Kit (Roche Applied Science, Indianapolis, IN, USA). A random 30-bp internal rvs3066 fragment was used as the control DNA for the gel shift assay. This control DNA fragment was also labeled with DIG-11-ddUTP. The DIG-11-ddUTP-labeled DNA fragments (0.02 μM) were incubated with the purified Rv3066 protein at concentrations ranging from 0.04 to 0.64 μM in binding buffer composed of 20 mM HEPES, pH 7.6, 1 mM EDTA, 10 mM (NH₄)₂SO₄, 5 mM dithiothreitol, 0.2% Tween 20 (wt/vol), 30 mM KCl and 0.5 μg of Poly [d(I-C)]. Bovine serum albumin (BSA) was used for the control experiment. After incubation, 10 mM MgCl₂ and 5 mM CaCl₂ were added to the reaction mixture to a final volume of 50 μl. Then, 0.0025 U of DNase I (Worthington Biochemicals, Lakewood, NJ, USA) was added and incubated for 5 min at room temperature. The reaction was stopped by adding 0.25 M EDTA and extracted with phenol:chloroform:isoamylalcohol (25:24:1). Control digestions with the probe were performed in the absence of Rv3066. The digested DNA fragments were purified with the QIAquick PCR Purification kit (Qiagen, Valencia, CA, USA) and eluted in 20 μl distilled water. After, 4 μl of the purified DNA was mixed with 5.98 μl HiDi formamide (Applied Biosystems, Foster City, CA, USA) and 0.02 μl GeneScan-500 LIZ size standards (Applied Biosystems). The samples were analyzed with the 3730 DNA analyzer coupled by G5 dye set, using an altered default genotyping module that increased the injection time to 30 s and the injection voltage to 3 kV.

The 226-bp fragment was sequenced with the primers 6FAM-Mmr-F and HEX-Mmr-R, respectively, using the Thermo Sequenase Dye Primer Manual Cycle Sequencing Kit (USB, Inc., Cleveland, OH, USA) according to the manufacturer’s instructions. Each reaction was diluted 5-fold in water, and 4 μl was added to 5.98 μl HiDi formamide and 0.02 μl GeneScan-500 LIZ size standard. The samples were analyzed using the 3730 DNA analyzer as described above. Electropherograms were analyzed and aligned using the GENEMAPPER software (version 4.0, Applied Biosystems).

**Cloning of the M. tuberculosis mmr-rv3066 operon into Mycobacterium smegmatis**

Primer pairs (FP: CGCGGATCCATCTTTTACCAGCAG, RP: CCCAAGCTTACGCTATCGTTCG) with added BamHI and HindIII sites were used to amplify the complete *mmr-rv3066* operon (including the promoter region) from *M. tuberculosis* H37Rv genomic DNA. The amplified fragment was digested and ligated into E. coli–Mycobacterium shuttle vector pMV261 (kindly provided by Susan T. Howard, The University of Texas Health Science Center) to construct pMMR. pMV261 is a shuttle plasmid commonly used for cloning and expression of genes in mycobacteria. The cloned *mmr-rv3066* operon carried its own promoter sequence. Although there should be constitutive base-level expression from the vector, *mmr* is inducible due to the binding of drug by the regulator. The constructed pMMR in *E. coli* DH5α was then purified and electroporated into *M. smegmatis* mc^2^155 (ATCC700084) (26).
Real-time quantitative reverse transcriptase–PCR analysis of \textit{mmr} induction

\textit{Mycobacterium smegmatis} mc\textsuperscript{2}155 strain containing pMMR was inoculated into 30 ml of antibiotic-free 7H9 broth. The culture was incubated for overnight to the mid logarithmic phase (OD\textsubscript{600} \approx 0.6) at 37°C. The culture was divided into four aliquots. One aliquot was used as the non-treated control, while the other three were added with thioridazine (80 μg/ml), erythromycin (128 μg/ml) and ethidium bromide (10 μg/ml), respectively. The culture was further incubated for 1 h at 37°C. Total bacteria RNA was isolated from the three individual cultures by the Trizol method (27). RNA was further purified by Qiagen RNeasy Column (Qiagen) and treated with the Turbo DNA-free Kit (Life Technologies, Grand Island, NY, USA) to eliminate DNA contamination in each preparation. Before being used for quantitative reverse transcriptase–PCR (qRT–PCR), each RNA template and each primer set were tested with a conventional one-step RT–PCR kit and a regular PCR kit (Life Technologies) to ensure specific amplification from the target mRNA and no detectable DNA contamination in the RNA preparation. Primer pairs 3065F (5’-CCTATACCTCTTGTGCGGAT-3’) and 3065R (5’-CGAAGCGATGCCATAACC-3’), specific for the \textit{mmr} gene were designed for qRT–PCR analyses, which were conducted using the iScript one-step RT–PCR kit with SYBR green (Bio-Rad, Hercules, CA, USA) along with the MyiQ iCycler real-time PCR detection system (Bio-Rad). Triplicate reactions in a volume of 15 μl were performed for each dilution of the RNA template. Thermal cycling conditions were as follows: 10 min at 95°C, 30 s at 60°C followed by 5 min at 95°C and then 40 cycles of 10 s at 95°C and 30 s at 56°C. Melt-curve analysis was performed immediately following each amplification. Samples between treatments were normalized using the \textit{aph} gene which was located on the pMV261 vector as an internal standard. Cycle threshold values were determined with the MyiQ software (Bio-Rad). The relative changes \((n\text{-fold})\) in \textit{mmr} transcription between the antibiotic treated and nontreated samples were calculated using \(2^{-\Delta\Delta CT}\) method as described by Livak and Schmitten (28).

Site-directed mutagenesis

Site-directed point mutations on residues that are expected to be critical for the binding of drugs were performed to generate single point mutants, W80A, Y101A, N112A, Y115A, W31A and D156A, and a double-point mutant W80A–W31A. The primers used for these mutations are listed in Table 2. All oligonucleotides were purchased from (Integrated DNA Technologies, Inc., Coralville, IA, USA) in a salt-free grade.

Fluorescence polarization assay for the DNA-binding affinity

Fluorescence polarization assays were used to determine the DNA-binding affinity of the Rv3066 regulator. Both the 30-bp oligodeoxynucleotide- and fluorescein-labeled oligodeoxynucleotide were purchased from Integrated DNA Technologies, Inc. These oligodeoxynucleotides contain the predicted 14-bp IR1 site for Rv3066 binding. Their sequences were 5’-CGAGCCCTCCATTGTGATCATAAATGTGCTCG-3’ and 5’-F-ACAAGGAGCCGCTCGTCTGGCGCCGCGCTCGTCTG-3’, where F denotes the fluorescein which was covalently attached to the 5’-end of the oligodeoxynucleotide by a hexamethylene linker. The 30-bp fluoresceinated dsDNA was prepared by annealing these two oligodeoxynucleotides together. Fluorescence polarization experiment was done using a DNA-binding solution containing 10 mM Na-phosphate (pH 7.2), 100 mM NaCl, 1 nM fluoresceinated DNA and 1 μg of poly(dI-dC) as non-specific DNA. The protein solution containing 500 nM dimeric Rv3066 and 1 nM fluoresceinated DNA was titrated into the DNA-binding solution until the millipolarization (mP) became unchanged. All measurements were performed at 25°C using a PerkinElmer LS55 spectrophotometer equipped with a Hamamatsu R928 photomultiplier. The excitation wavelength was 490 nm, and the fluorescence polarization signal (in ΔP) was measured at 520 nm. Each titration point recorded was an average of 15 measurements. Data were

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>W131A-forward</td>
<td>5’-GTCTGTCGTCCTGCTGTGGCCGGCGGATGGTGCTGC-3’</td>
</tr>
<tr>
<td>W131A-reverse</td>
<td>5’-GCAGGACATCCCGCCGCACGACCAGCAGC-3’</td>
</tr>
<tr>
<td>S73A-forward</td>
<td>5’-CATCCGGCAAAACTGACGCCAGAT-3’</td>
</tr>
<tr>
<td>S73A-reverse</td>
<td>5’-ATCTGGGCTGTCATTGGCCGCATG-3’</td>
</tr>
<tr>
<td>T159A-forward</td>
<td>5’-GTTTTTCTGACCCGCTGTCGGGACACGTGCCTG-3’</td>
</tr>
<tr>
<td>T159A-reverse</td>
<td>5’-CACTGGTGCAGCAGCCGCGCTGTCGGGACACGTGCCTG-3’</td>
</tr>
<tr>
<td>Y101A-forward</td>
<td>5’-GGTCTGACCTGCAGTGCCATGTCAGTCAG-3’</td>
</tr>
<tr>
<td>Y101A-reverse</td>
<td>5’-GCCCTGAACAGAACTGGAATGGCCGGCCGACATC-3’</td>
</tr>
<tr>
<td>Y115A-forward</td>
<td>5’-GATGTGCGCCGCGCATGTCAGTGCCGACGTGCCTG-3’</td>
</tr>
<tr>
<td>Y115A-reverse</td>
<td>5’-GCTGCTGGCCGCGCATGTCAGTGCCGACGTGCCTG-3’</td>
</tr>
<tr>
<td>W80A-forward</td>
<td>5’-GAGTTACGAGAAGCCAAGCTGACAGTATGATGTC-3’</td>
</tr>
<tr>
<td>W80A-reverse</td>
<td>5’-GAGATACGAGAAGCCAAGCTGACAGTATGATGTC-3’</td>
</tr>
<tr>
<td>D156A-forward</td>
<td>5’-GTCACCGTGTATTTCGAGGCGCCACGCTGCAC-3’</td>
</tr>
<tr>
<td>D156A-reverse</td>
<td>5’-GTGCAGCTGCGGCCGCTGCGAAAACACGGTCGAC-3’</td>
</tr>
<tr>
<td>N112A-forward</td>
<td>5’-GCTTACGCGACCACCTGCGAATGATATGCTG-3’</td>
</tr>
<tr>
<td>N112A-reverse</td>
<td>5’-GCTTACGCGACCACCTGCGAATGATATGCTG-3’</td>
</tr>
</tbody>
</table>

Table 2. Primers for site-directed mutagenesis
analyzed using the equation, $P = \{(P_{\text{bound}} - P_{\text{free}})\} / \{(K_D + [\text{protein}]) + P_{\text{free}}\}$, where $P$ is the polarization measured at a given total protein concentration, $P_{\text{free}}$ is the initial polarization of free fluorescein-labeled DNA, $P_{\text{bound}}$ is the maximum polarization of specifically bound DNA and [protein] is the protein concentration. The titration experiments were repeated for three times to obtain the average $K_D$ value. Curve fitting was accomplished using the program ORIGIN (OriginLab Corporation).

Fluorescence polarization assay for ligand-binding affinity

Fluorescence polarization was used to determine ethidium-binding affinities of Rv3066 and its mutants. This approach was also employed to study the interaction between the Rv3066 regulator and the anti-TB drug thioridazine. The experiment was done using a ligand-binding solution containing 10 mM Na-phosphate (pH 7.2), 100 mM NaCl and 1 μM ethidium bromide. The protein solution consisting of Rv3066 or Rv3066 mutant in 10 mM Na-phosphate (pH 7.2), 100 mM NaCl and 1 μM ethidium bromide was titrated into the ligand-binding solution until the polarization ($P$) was unchanged. As this is a steady-state approach, fluorescence polarization measurement was taken after a 5 min incubation for each corresponding concentration of the protein and ligand to ensure that the binding has reached equilibrium. All measurements were performed at 25°C using a PerkinElmer LS5 spectrophotometer equipped with a Hamamatsu R928 photomultiplier. The excitation and emission wavelengths were 483 and 620 nm for ethidium binding, whereas these wavelengths were 320 and 430 nm for thioridazine measurement. Fluorescence polarization signal (in $\Delta P$) was measured at the emission wavelength. Each titration point recorded was an average of 15 measurements. Data were analyzed using the equation, $P = \{(P_{\text{bound}} - P_{\text{free}})\} / \{(K_D + [\text{protein}]) + P_{\text{free}}\}$, where $P$ is the polarization measured at a given total protein concentration, $P_{\text{free}}$ is the initial polarization of free ligand, $P_{\text{bound}}$ is the maximum polarization of specifically bound ligand and [protein] is the protein concentration. The titration experiments were repeated for three times to obtain the average $K_D$ value. Curve fitting was accomplished using the program ORIGIN (OriginLab Corporation).

RESULTS

Overall structure of Rv3066

In the absence of inducer molecule, the Rv3066 regulator was crystallized in two different Forms, I and II (Supplementary Figure S1). The Form I (SeMet) crystal structure was determined to a resolution of 2.3 Å (Table 1 and Figure 1a). The asymmetric unit contains a single homodimer, suggesting that this regulator is dimeric in nature. Similar to LfrR (30), the left and right subunits of Rv3066 have their own individual roles. The dimeric Form I structure of Rv3066, indicating an all-helical protein, is shown in Figure 1b. Superimposition of both subunits of Rv3066 gives rise to an overall rms deviation of 1.3 Å calculated over the Cα atoms.

The crystal structure of Form II was refined to a resolution of 1.8 Å (Table 1). Like Form I, the structure of Form II indicates that this protein is an asymmetric homodimer. Superimposition of the Cα atoms of the two subunits of Form II results in an rms deviation of 1.8 Å. Surprisingly, the conformations of the structures of Forms I and II are quite distinct from each other, suggesting that these two structures probably represent two different transient states of the regulators. Superimposition of the entire dimer of these two apo-Rv3066 structures (Forms I and II) provides an rms deviation of 3.0 Å (Figure 2). Overall, the architecture of these two Rv3066 structures are in good agreement with those

Molecular modeling

The program AutoDock Vina (29) was used to predict the ligand-binding mode of several Rv3066 drugs, including acriflavine, ethidium, proflavine, pyronin Y, safranin O and thioridazine. A monomer of the structure of Rv3066-ethidium with the bound ethidium removed was used for docking. The protein was set as a rigid structure whereas the conformation of each ligand was optimized during all modeling and docking procedures. For each ligand, the results were ranked on the basis of predicted free energy of binding and the one with the highest binding affinity was recorded.

The model of the DNA-bound form of Rv3066 was generated using AutoDock Vina (29). In brief, the two N-terminal domains of dimeric apo-Rv3066 (Form I structure) were chosen to dock onto the two successive major grooves of the IR1 double helix. During the docking procedure, both the structures of Rv3066 and IR1 were held rigid. After obtaining the docking result, the entire structure of the dimeric Rv3066 regulator was then superimposed onto the docked N-terminal helices at the major grooves of the DNA to generate the final Rv3066–DNA complex model.
of the homolog proteins TetR (31,32), QacR (33,34), CprB (35), EthR (36,37), CmeR (38), AcrR (39), SmeT (40) and LfrR (30). Each subunit of Rv3066 is composed of nine α helices (α1–α9 and α10–α90, respectively) (Figures 1b and 2) and can be divided into two motifs: an N-terminal DNA-binding domain and a C-terminal ligand-binding domain. The helices of Rv3066 are designated numerically from the N-terminus as α1 (15–30), α2 (37–44), α3 (48–54), α4 (58–86), α5 (90–103), α6 (105–120), α7 (122–143), α8 (145–165) and α9 (171–182). In this arrangement, the smaller N-terminal domain includes helices α1 through α3 and the N-terminal end of α4 (residues 58–68), with α2 and α3 forming a typical helix-turn-helix motif. However, the larger C-terminal domain comprises the C-terminal end of helices α4 (residues 69–86) through α9 and helices α6, α8 and α9 are involved in the dimerization of the regulator. The smaller N-terminal domain shares considerably high sequence and structural similarities with the TetR-family regulators, suggesting that Rv3066 belongs to the TetR family. For example, residues 13–68 possess 20% amino acid identity and 68% homology to that of TetR (31). This N-terminal region also shows identities of 29% and 25%, and similarities of 62% and 63% to those of the \( M. \) smegmatis LfrR (30) and \( M. \) tuberculosis EthR repressors (37), respectively. Protein sequence alignment of Rv3066 with the TetR-family members of other \( M. \)ycobacterium species is shown in Supplementary Figure S2. The alignment suggests that Rv3066 is a typical TetR-family regulator.

Conformational flexibility of the Rv3066 regulator

Two distinct conformations of apo-Rv3066 were captured using crystallography, suggesting that this regulatory protein is quite flexible in nature. A comparison of the N-terminal DNA-binding domains of the dimeric structures of Forms I and II indicates that these two structures may depict two different transient states of the regulator. Apparently, these two conformations are related in which an 8° rotational motion of the right subunit (helices α1’–α9’) with respect to the left protomer (helices α1–α9) is attributed to the difference (Supplementary Figure S3).
Indeed, we can easily generate the Form II conformation using the Form I structure as a starting point and then following the rigid-body rotational trajectory based on the TLS parameters. The resulting final B-factors and overall structure to those of Forms I and II apo-Rv3066 result in overall rms deviations of 3.4 and 1.9 Å, respectively.

Although the conformations of the two subunits within the dimer are not identical, their ethidium-binding modes are quite similar. Both ethidium sites utilize the same set of amino acids to accommodate the binding, with a slight difference in the interaction distances. The $F_o - F_c$ electron density maps of the two bound ethidiums within the dimer are illustrated in Figure 4a. Interestingly, the total volume of the two ligand-binding pocket has expanded to 929 Å$^3$ (443 Å$^3$ for the left subunit and 486 Å$^3$ for the right subunit) in the ethidium bound structure.

Each bound ethidium molecule is completely buried in the Rv3066-binding pocket. The ligand-binding pocket is found to be hydrophobic in nature. Five aromatic residues, W80, Y101, Y115, W131 and F155, participate to make aromatic stacking interactions and hydrophobic contacts with the bound ethidium (Figure 4b and c and Supplementary Figure S6). In addition, L76, T98, L111 and T159 are involved to secure the binding through hydrophobic interaction. Furthermore, the N1 and N2 amino group nitrogens of the phenanthridinium system of ethidium are hydrogen bonded to the side chain oxygen of T159 and S73 respectively (Table 3). Additional hydrogen bonds have also been found between the N1 and N2 amino group nitrogens of the bound ethidium and the backbone oxygens of F155 and S73 to secure the binding (Table 3). In addition, a negatively charged residue D156 participates to form the ligand-binding site, and this residue is $\sim$3.3 Å away from the phenyl groups of the bound ethidium in each subunit of the dimer (Figure 4b and Supplementary Figure S6).

Electrophoretic mobility shift assay
Bacterial multidrug efflux regulators usually bind to the palindromic IR sequences at the promoter regions to control the expression of the multidrug efflux pumps. In the mmr-rv3066 operon, it appears that there are two 14-bp palindromic IR sequences located upstream of mnr. These two IRs (IR1 and IR2) are located right next to one another, where their sequences are complementary to each other (Supplementary Figure S7). The sequences of IR1 and IR2 are 5' - TGTACATTGGTACA-3' and 5' - TGTACAAATGTGACA-3', respectively. The presence of these IRs suggests potential binding sites for the Rv3066 regulator. Thus, EMSA was performed using a 30-bp dsDNA containing the IR1 sequence and purified Rv3066 protein to detect if the regulator specifically binds IR1. As shown in Figure 5a and b, the shift of the labeled DNA band was dependent upon the protein concentration as well as the addition of the unlabeled specific 30-bp dsDNA. The data indicate that the IR1 sequence potentially forms the specific binding site for Rv3066.

EMSA was further performed with the purified Rv3066 protein and 30-bp DNA fragment containing the IR1 site in the presence of ethidium, which is a substrate of the Mmr efflux pump and is found to be bound in the ligand-binding site of the Rv3066 regulator. As shown in Figure 5c, the addition of ethidium to the Rv3066–DNA complex resulted in the loss of the retarded band,
indicating the separation of the protein and DNA components. The result suggests that ethidium is a substrate of Rv3066 and that the binding of ethidium triggers significant conformational change to the regulator, which in turn renders it unable to bind its cognate operator DNA.

Dye primer based DNase I footprint assay

To further confirm the binding site of Rv3066 in the \textit{mmr-rv3066} promoter region, DNase I footprint assay was performed using the method of dye primer sequencing (25). In comparison with different electropherograms at various concentrations of the dimeric Rv3066 protein, we were able to uncover the specific DNA sequence, TTGTGTACATTTGTACACAAAGG, which was protected by the regulator (Figure 6 and Supplementary Figure S8). Interestingly, the IR1 sequence (TGTACA TTTGTACA) was found within this protected region, suggesting that Rv3066 is likely to specifically bind IR1.

Impact of drugs on \textit{mmr} transcription in \textit{M. smegmatis}

We cloned the \textit{mmr-rv3066} operon into \textit{M. smegmatis} to assess its function and induction by antimicrobials. It has

Figure 3. Crystal structure of the Rv3066–ethidium complex. (a) Ribbon diagram of the Rv3066–ethidium complex. The bound ethidiums are shown as sticks (magenta, carbon; blue, nitrogen). (b) Structural comparison of Form I and Rv3066–ethidium. This is a superimposition of the dimeric structures of Form I (orange) and Rv3066–ethidium (green). The bound ethidiums are in yellow sticks. Helices \(\alpha_1–\alpha_9\) (left subunit) and \(\alpha_{1'}–\alpha_{9'}\) (right subunit) are labeled. The arrow indicates a change in orientation of the right subunit of dimeric Rv3066–ethidium when compared with the Form I structure. This conformational change can be interpreted as a rotational motion of the right subunit of Rv3066 with respect to the horizontal axis passing through the two ligand-binding pockets of the dimer upon ethidium binding.
been reported that the anti-TB drug thioridazine was used to cure 10 XDR-TB patients (41). Thus, this drug was chosen to investigate its effect on the expression of the Mmr multidrug efflux pump using qRT–PCR. We also studied the effect of ethidium and erythromycin on mmr transcription using the same approach. After 1 h induction with thioridazine, the mmr gene was up-regulated by 2.70 ± 0.09-fold (P-value = 0.0001). No obvious induction of the mmr gene was observed in the presence of ethidium or erythromycin at the same time-point.

Fluorescence polarization assay

Presumably, Rv3066 suppresses the expression of the Mmr multidrug efflux pump by directly binding to its target DNA. Fluorescence polarization-based assay was carried out to study the interaction between Rv3066 and the 30-bp DNA containing the IR1 sequence. Figure 7a illustrates the binding isotherm of Rv3066 in the presence of 1 nM fluoresceinated DNA. The titration experiment indicated that this regulator binds the 30-bp IR1 operator with a dissociation constant, \( K_D = 4.4 \pm 0.3 \) nM. This value is similar to that of the QacR regulator where it binds DNA with the \( K_D \) of 5.7 nM (42). The binding data also indicate that Rv3066 binds its cognate DNA with a stoichiometry of one Rv3066 dimer per IR1.

To investigate whether the presence of ethidium affects the binding of Rv3066 with IR1, fluorescence polarization was also carried out to study the interaction between Rv3066 and IR1 in the presence of this drug. The experimental results suggest that the \( K_D \) of Rv3066–IR1 becomes 10.7 ± 0.9 nM in the presence of 1 \( \mu \)M ethidium (Supplementary Figure S9), indicating that ethidium significantly weakens the binding affinity between Rv3066 and IR1 by 2.4 times.

In addition, fluorescence polarization was used to determine the binding affinity of ethidium and the anti-TB drug thioridazine by Rv3066. The measurements indicate that the \( K_D \) values of the Rv3066–ethidium and

![Figure 4](https://example.com/fig4.png)
Rv3066–thioridazine complexes are $2.9 \pm 0.2$ and $211.8 \pm 34.2 \mu M$, respectively (Figure 7b and c). These binding data suggest that the protein employs a simple binding stoichiometry with a 1:1 monomeric Rv3066-to-drug molar ratio. This molar ratio is indeed in good agreement with the crystal structure where each monomer of Rv3066 binds one ethidium molecule.

The structure of Rv3066–ethidium indicates that residues W80, Y101, N112, Y115, W131 and D156 are involved in ligand binding. These residues were mutated into alanines (W80A, Y101A, N112A, Y115A, W131A and D156A). The corresponding mutant regulators were then expressed and purified. The ability of these mutant regulators to bind ethidium was tested using fluorescence polarization assay (Table 4). The results demonstrate that several of these point mutants show a significant decrease in the binding affinity for ethidium when compared with the wild-type Rv3066 regulator. Particularly, mutant D156A increases the dissociation constant, $K_D$, of ethidium binding by 20-fold, suggesting that D156 is a critical residue in the multidrug-binding site of Rv3066. Moreover, mutants W80A, N112A and W131A decrease the affinity for ethidium binding by three to four times. A double-point mutant W80A–W131A was then produced to investigate how these tryptophans affect the binding of ethidium. Similar to the D156A mutant, the W80A–W131A double mutant indicates a significantly weaker binding affinity (20-fold decrease) when compared with that of the wild-type Rv3066, suggesting these two tryptophans, W80 and W131, are important residues for drug recognition.

Gel filtration

To confirm the stoichiometry of one Rv3066 dimer bound to one IR1 operator site, gel-filtration experiment was performed using the purified Rv3066 protein pre-incubated with the purified, complementary, annealed 30-bp oligonucleotides containing the IR1 sequence. The result suggests an average molecular weight of 67.3 ± 3.8 kDa for the Rv3066–DNA complex (Figure 8). This value is in good agreement with the theoretical value of 66.0 kDa for two Rv3066 molecules bound to the 30-bp DNA. Thus, the stoichiometry of the Rv3066–IR1 complex is 1:1 dimeric Rv3066-to-DNA molar ratio.

Docking of ligands into the multidrug drug binding site

To understand how Rv3066 binds different ligands, we used the program AutoDock Vina (29) to identify potential binding modes for a variety of drug molecules. We first predicted an ethidium-binding site in Rv3066. We found that the predicted bound ethidium molecule was completely overlapped with the bound ethidium identified from the crystal structure and their binding modes are nearly identical (Supplementary Figure S10), suggesting that AutoDock Vina is sufficiently precise for identification of potential drug-binding pockets. When Vina was used to search for binding sites for different Rv3066 drugs, including acriflavine, proflavine, pyronin Y, safranin O and thioridazine, it was found that...
all these drug molecules were bound within the ethidium-binding site identified by X-ray crystallography (Figure 9a and Table 5).

**Predicting the structure of DNA-bound form of Rv3066**

AutoDock Vina (29) was also used to generate a model of the Rv3066–DNA complex structure (Figure 9b and Table 5). This model reveals that the two N-terminal domains of Rv3066 within the dimer are able to fit well into the two successive major grooves of the target DNA. Extensive interactions between the regulator and DNA have also been found to stabilize this protein–DNA complex structure. Specifically, T52 is found to make a contact with one base of IR1. In addition, Y53 participates to interact directly with two different bases. Additional stabilization of the complex comes from R15, H37 and R38, in which these residues make phosphate contacts with the DNA.

**DISCUSSION**

With the rising incidences of multidrug resistant strains of TB, it has become increasingly important to understand how individual proteins function to recognize and confer resistance to multiple antibiotics in this pathogen. The crystal structures of Rv3066 both in the absence and presence of the bound ethidium provide direct information about how this regulator controls the expression of the Mmr multidrug efflux pump, which mediates the resistance of several antimicrobial agents. It appears that ethidium binding triggers a series of cooperativity motions of the C-terminal helices, including the horizontal shifts of helices α5, α6 and α7 toward the dimer interface, and an upward movement of helix α8 within one subunit of the regulator. These conformational changes initiate a rotational motion of the second subunit of Rv3066 with respect to the horizontal axis passing through the two ligand-binding pockets of the dimer (Figures 3b, 5a and b), presumably making the relative orientation of the two N-terminal DNA-binding domains no longer compatible with the two consecutive major grooves of the operator B-DNA. The net result is that this dimeric regulator is released from the promoter, initiating the expression of the Mmr multidrug efflux pump.

We generated a speculative model of DNA-bound Rv3066 using AutoDock Vina (29). The model suggests that the two N-terminal DNA-binding domains of the dimeric apo-Rv3066 regulator (Form I) can easily be docked onto the two consecutive major groove of IR1, suggesting that the structure of the regulator in the Rv3066–IR1 complex may be very similar to the Form I structure of apo-Rv3066. By comparing with the apo-Rv3066, Rv3066–ethidium and the speculative Rv3066–DNA structures, it appears that the induction mechanism of Rv3066 seems to be attributed to the rigid body rotational motion of the two subunits triggered by ethidium binding (Figure 5c). As mentioned above, this motion changes the relative orientation of the two DNA-binding domains of Rv3066, which makes the dimer incompatible with the two successive major grooves.
grooves of the IR and results in disallowing the regulator to bind the DNA duplex. This induction mechanism is quite distinct from those of QacR (33,34) and TetR (31,32). However, a similar induction mechanism, which is triggered by the movement as rigid bodies, can be found in the SimR regulator (43), where rigid body rotation within subunits of the dimer in relation to one another contributes to the induction.

In comparison with the Form I and ethidium bound structures, the shift in position of helices $a_5$, $a_6$ and $a_7$ toward the dimer interface results in the formation of four new hydrogen bonds. These hydrogen bonds are found within the dimer interface between H120 and R109, and between R121 and Q113. In addition, the backbone oxygens of R129 and T165 also contribute to the dimer interface. These backbone oxygens interact with I164 and R129 to form two hydrogen bonds. The formation of these new hydrogen bonds is presumed to stabilize the tertiary structure of the ligand bound form of the Rv3066 dimer.

Coupled with the movements of these C-terminal helices, residues located at the binding site are also found to readjust their positions to accommodate for the binding of ethidium. Noticeably, Y115 shifts upward by 2.4 Å presumed to make an interaction with the phenyl group of the bound ethidium, whereas F155 is found to swing upward by 90° to enlarge the volume of the binding site. In addition, W131 adjusts its position to shift upward by 2.3 Å and makes an aromatic stacking interaction with the three-ring system of the phenanthridinium group of ethidium. W80 and Y101 also slightly switch in location by 1.1 and 2.0 Å to contribute aromatic $\pi$-$\pi$ and stacking interactions with the bound ligand (Figure 4c). One common characteristic of multidrug-binding proteins is their flexibility to accommodate for different ligands in a single ligand-binding pocket. This plasticity and flexibility can easily be seen in the Rv3066 regulator in which many of these residues lining the binding site are found to participate and relocate their side chain positions upon ethidium binding.

To investigate if Rv3066 can accommodate different drugs, we used the program AutoDock Vina (29) to study how Rv3066 binds a variety of drugs. The docking study shows that the large cavity of the multidrug-binding site of each Rv3066 monomer can accommodate many other drugs, such as acriflavine, proflavine, pyronin Y, safranin O and thioridazine. In each case, the bound

Table 4. Dissociation constants for ethidium binding

<table>
<thead>
<tr>
<th>Protein</th>
<th>$K_D$ (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type Rv3066</td>
<td>2.9 ± 0.2</td>
</tr>
<tr>
<td>W80A</td>
<td>8.5 ± 0.6</td>
</tr>
<tr>
<td>Y101A</td>
<td>2.1 ± 0.2</td>
</tr>
<tr>
<td>N112A</td>
<td>10.3 ± 0.7</td>
</tr>
<tr>
<td>Y115A</td>
<td>6.1 ± 0.6</td>
</tr>
<tr>
<td>W131A</td>
<td>12.9 ± 1.0</td>
</tr>
<tr>
<td>D156A</td>
<td>41.0 ± 5.9</td>
</tr>
<tr>
<td>W80A-W131A</td>
<td>43.2 ± 5.2</td>
</tr>
</tbody>
</table>

Figure 7. Representative fluorescence polarization of Rv3066. (a) The binding isotherm of Rv3066 with the 30-bp DNA containing the IR1 sequence, showing a $K_D$ of 4.4 ± 0.3 nM. (b) The binding isotherm of Rv3066 with ethidium, showing a $K_D$ of 2.9 ± 0.2 μM. (c) The binding isotherm of Rv3066 with thioridazine, showing a $K_D$ of 211.8 ± 34.2 μM. Fluorescence polarization is defined by the equation, $FP = (V - H)/(V + H)$, where $FP$ equals polarization, $V$ equals the vertical component of the emitted light, and $H$ equals the horizontal component of the emitted light of a fluorophore when excited by vertical plane polarized light. $FP$ is a dimensionless entity and is not dependent on the intensity of the emitted light or on the concentration of the fluorophore. mP is related to FP, where 1 mP equals one thousandth of a FP.
drug was completely buried in the Rv3066 molecule, and strong interaction was observed between the bound drug and the regulator (Table 5). The dockings suggested that all these ligands are bound in the same multidrug-binding site with a similar binding mode, suggesting that the process of induction by these drugs is similar.

A distinguishing feature of multidrug-binding proteins that bind cationic drugs is the presence of buried acidic glutamates or aspartates in the ligand-binding pockets. This was clearly demonstrated by the structures of the QacR–ligand complexes (33). Indeed, a completely buried negatively charged aspartate, which is essential for cationic ligand recognition, has also been found in the ligand-binding pocket of the AcrR multidrug efflux regulator (39). A similar characteristic for the TetR-family regulators that recognize negatively charged antimicrobials has also been observed through x-ray crystallography (38,44). In this case, positively charged histidines or lysines within the ligand-binding pockets are critical for interacting with the negatively charged drugs. For Rv3066, it is found that there is one negatively charged residue (D156) inside the multidrug-binding pocket. D156 is critical for ethidium binding as shown using fluorescence polarization and mutagenesis studies, although this residue does not seem to contribute a significant electrostatic interaction to the cationic ligand. A mutation of D156 with an alanine drastically decreases the Rv3066–ethidium-binding affinity by 20-fold. Instead of contributing hydrogen-bonded interaction to neutralize the formal positive charge of the bound ethidium, the D156 residue, together with N112, is responsible for forming hydrogen bonds with the aromatic residues, Y101, Y115 and W131 (Figure 4b). These three residues, in addition to W80, form an aromatic cage-like binding pocket to position the bound ethidium (Supplementary Figure S6), suggesting that D156 is important for the organization of the ligand-binding site. The importance of the cage-like arrangement of these aromatic residues to bind ethidium is further demonstrated using mutagenesis and fluorescence polarization studies, in which the W80–W131 double mutant has been found to bind ethidium 20 times weaker than the wild-type Rv3066.

It is found that the Rv3066 dimer binds two drug molecules, with each subunit of the regulator contributing to bind one drug. Gel-filtration experiment confirmed that the Rv3066–IR1 complex should be in the form of 1:1 dimeric Rv3066-to-DNA molar ratio. Real-time quantitative RT–PCR analysis suggested that the \textit{mmr} gene was significantly up-regulated by thioridazine induction. This result is indeed in good agreement with a recent finding that the treatment of thioridazine on \textit{M. tuberculosis} causes a profound increase in the expression of the \textit{mmr} multidrug efflux pump (12). The structures of the Rv3066 regulator both in the absence and presence of bound ethidium, together with the experimental data from electrophoretic mobility shift, footprinting analysis, qRT–PCR, mutagenesis, gel filtration and fluorescence polarization, support the role of Rv3066 in regulating the expression level of the multidrug efflux pump \textit{Mmr} in \textit{M. tuberculosis}.

**Figure 8.** Representative gel-filtration experiment. The experiment demonstrated that two Rv3066 molecules are bound to one 30-bp DNA containing the IR1 operator. The y-axis values were defined as: $K_{av} = (V_e - V_0)/(V_T - V_0)$, where $V_T$, $V_e$ and $V_0$ are the total column volume, elution volume and void volume of the column, respectively. Standards used were: A, cytochrome C ($M_r$ 12 400); B, carbonic anhydrase ($M_r$ 29 000); C, albumin bovine serum ($M_r$ 66 000); D, alcohol dehydrogenase ($M_r$ 150 000); and E, \(\beta\)-Amylase ($M_r$ 200 000). The void volume was measured using blue dextran ($M_r$ 2000000).
PROTEIN DATA BANK ACCESSION CODES

Coordinates and structural factors for the structures of Rv3066 have been deposited at the RCSB Protein Data Bank with accession codes 3T6N (Form I, apo-Rv3066), 3V6G (Form II, apo-Rv3066) and 3V78 (Rv3066–ethidium).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–10.

ACKNOWLEDGEMENTS

The authors are grateful to Susan T. Howard at the University of Texas Health Science Center (Tyler) for providing the pMV261 plasmid used in this study. This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines of the Advanced Photon Source, supported by award RR-15301 from the National Center for Research Resources at the National Institutes of Health.

FUNDING

National Institutes of Health (NIH) [DK063008 to Q.Z. and GM086431 to E.W.Y.]; Use of the Advanced Photon Source is supported by the U.S. Department of Energy, Office of Basic Energy Sciences [DE-AC02-06CH11357]; Community College Institute of Science Internship (CCI) and Science Undergraduate Laboratory Internship (SULI) from the DOE (to J.E.G. and D.C.M.). Funding for open access charge: NIH [GM086431].

Conflict of interest statement. None declared.

REFERENCES