Crystal structure of the transcriptional regulator Rv0678 of *Mycobacterium tuberculosis*

Abhijith Radhakrishnan\(^1\), Nitin Kumar\(^1\), Catherine C. Wright\(^2\), Tsung-Han Chou\(^3\), Marios L. Tringides\(^1\), Jani Reddy Bolla\(^1\), Hsiang-Ting Lei\(^1\), Kanagalaghatta R. Rajashankar\(^4\), Chih-Chia Su\(^3\), Georgiana E. Purdy\(^2\), and Edward W. Yu\(^1\)

\(^1\)From the Department of Chemistry, Iowa State University, Ames, IA 50011, USA

\(^2\)Department of Molecular Microbiology and Immunology, Oregon Health and Sciences University, Portland, OR 97239, USA

\(^3\)Department of Physics and Astronomy, Iowa State University, Ames, IA 50011, USA

\(^4\)NE-CAT and Department of Chemistry and Chemical Biology, Cornell University, Bldg. 436E, Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439, USA

*Running title: Structure of the transcriptional regulator Rv0678*

To whom correspondence should be addressed: Edward W. Yu, Department of Chemistry, Department of Physics & Astronomy, Iowa State University, Ames, IA, USA, Tel: (515) 294-4955; E-mail: ewyu@isatate.edu

**Keywords:** Mycobacterial membrane protein large; Mycobacterial membrane protein small; transcriptional regulation; MarR-family regulators; Rv0678

**Background:** The expression of the *Mycobacterium tuberculosis* MmpS5-MmpL5 transporter is controlled by the MarR-like transcriptional regulator Rv0678.

**Results:** Rv0678 forms a dimeric two-domain molecule with the architecture similar to members of the MarR family of transcriptional regulators.

**Conclusion:** Rv0678 is distinct in that its DNA-binding and dimerization domains cooperate to bind an inducing ligand.

**Significance:** These findings suggest a mechanism for ligand and regulator derepression.

**ABSTRACT**

Recent work demonstrates that the Mycobacterial membrane protein large (MmpL) transporters are dedicated to the export of mycobacterial lipids for cell wall biosynthesis. An MmpL transporter frequently works with an accessory protein, belonging to the Mycobacterial membrane protein small (MmpS) family, to transport these key virulence factors. One such efflux system in *Mycobacterium tuberculosis* is the MmpS5-MmpL5 transporter. The expression of MmpS5-MmpL5 is controlled by the MarR-like transcriptional regulator Rv0678, whose open reading frame is located downstream of the *mmpS5-mmpL5* operon. To elucidate the structural basis of Rv0678 regulation, we have determined the crystal structure of this regulator, to 1.64 Å resolution, revealing a dimeric two-domain molecule with an architecture similar to members of the MarR family of transcriptional regulators. Rv0678 is distinct from other MarR regulators in that its DNA-binding and dimerization domains are clustered together. These two domains seemingly cooperate to bind an inducing ligand that we identified as 2-palmitoylglycerol, which is a fatty acid glycerol ester. The structure also suggests that the conformational change leading to substrate-mediated derepression is primarily caused by a rigid body rotational motion of the entire DNA-binding domain of the regulator towards the dimerization domain. This movement results in a conformational state that is incompatible with DNA binding. We demonstrate using electrophoretic mobility shift assays that Rv0678 binds to the *mmpS5-mmpL5, mmpS4-mmpL4* and the *mmpS2-mmpL2* promoters. Binding by Rv0678 was reversed upon addition of the ligand. These
findings provide new insight into the mechanisms of gene regulation in the MarR family of regulators.

Tuberculosis (TB) is one of the oldest described diseases and remains a significant global problem with more than eight million new cases reported annually (1). The World Health Organization estimates that one-third of the world’s population is infected with *M. tuberculosis*, and most of these individuals have latent TB (2). TB treatments are notoriously difficult and are compromised by the emergence of multiple drug-resistant (MDR), extensively drug-resistant (XDR) and totally drug-resistant (TDR) bacterial strains (3-7). The development of drug-resistant *M. tuberculosis* strains is a major threat that challenges global prospects for TB control.

Although mycobacteria cluster phylogenetically with Gram-positive prokaryotes, they are structurally more similar to Gram-negative bacteria. These mycobacteria are protected by an outer lipid bilayer made of mycolic acids and a cell envelope composed of non-covalently bound lipids and glycolipids. The unique structure and composition of the cell wall differentiates this highly pathogenic microorganism from other prokaryotes. The mycobacterial cell wall plays a crucial role in the host-pathogen interface on several levels (8). First, the thick, greasy cell wall acts as an effective layer of protection, providing intrinsic resistance to antibiotics and bactericidal components of the host immune response. Second, the surface-exposed polyketide and glycoconjugate lipids of the *M. tuberculosis* cell wall are associated with bacterial virulence (9-12).

The genome of *M. tuberculosis* H37Rv contains 15 genes that encode the resistance-nodulation-cell division (RND) proteins designated MmpL transporters (13, 14). Unlike the RND-type efflux pumps of Gram-negative bacteria, MmpL proteins do not usually participate in antibiotic efflux. Instead, there is strong evidence that these MmpL proteins are responsible for exporting fatty acids and lipidic elements of the cell wall (8, 9, 12, 15-17). Five *mmpL* genes are located adjacent to genes coding for proteins involved in fatty acid or polyketide synthesis, suggesting that the MmpL membrane proteins transport these key virulence factors (9, 10). Similar to RND proteins of Gram-negative bacteria, the MmpL transporters of *M. tuberculosis* are believed to work in conjunction with accessory proteins. Specifically, MmpL transporters form complexes with the MmpS-family proteins in order to export cell wall lipid constituents (18). Five genes encoding MmpS proteins are adjacent to genes encoding MmpL proteins (8, 13). Work in the model organism *M. smegmatis* demonstrated that MmpS4 was required for bacterial sliding motility and biofilm formation (19). That the *mmpS4* and *mmpL4* mutants had similar phenotypes underscore a coordinated function for cognate MmpS-MmpL proteins.

Our efforts have focused on elucidating how *M. tuberculosis* transport systems are regulated. We previously crystallized the Rv3066 efflux regulator both in the absence and presence of bound substrate (20). Our data indicated that ligand binding triggers a rotational motion of the regulator, which in turn releases the cognate DNA and induces the expression of the Mrm efflux pump (20). We report here the crystal structure of the Rv0678 regulator, which has been proposed to control the transcriptional regulation of the MmpS5-MmpL5 transport system. Rv0678 belongs to the MarR family of regulators, which are found ubiquitously in bacteria and archaea and control various important biological processes, such as resistance to antimicrobials, sensing of oxidative stress agents and regulation of virulence factors (21). Typically, the MarR-family regulators are dimeric in form and their protein sequences are poorly conserved. However, these proteins share a common fold, consisting of a helical dimerization domain and two winged helix-turn-helix (wHtH) DNA-binding domains within the dimer (22). Our data suggest that fatty acid glycerol esters are the natural ligands of the Rv0678 regulator. Electrophoretic mobility shift assay indicates that Rv0678 binds promoters of the *mmpL2*, *mmpL4* and *mmpL5* operons. These results emphasize the importance of the Rv0678 regulator, which appears to regulate multiple
MmpL transport systems.

EXPERIMENTAL PROCEDURES

Cloning of rv0678 – The rv0678 ORF from genomic DNA of M. tuberculosis strain H37Rv was amplified by PCR using the primers 5’-CCATGGGCAACGTCAACGACCGGGTC-3’ and 5’-GGATCCTCAGTGATGATGATGATGATGGTC GTCTCTCCGGTTCG-3’ to generate a product that encodes a Rv0678 recombinant protein with a 6xHis tag at the C-terminus. The corresponding PCR product was digested with NcoI and BamHI, extracted from the agarose gel, and inserted into pET15b as described by the manufacturer (Merck KGaA, Darmstadt, Germany). The recombinant plasmid (pET15bΔrv0678) was transformed into DH5α cells and the transformants were selected on LB agar plates containing 100 µg/ml ampicillin. The presence of the correct rv0678 sequence in the plasmid construct was verified by DNA sequencing.

Expression and purification of Rv0678 – Briefly, the full-length Rv0678 protein containing a 6xHis tag at the C-terminus was overproduced in E. coli BL21(DE3) cells possessing pET15bΔrv0678. Cells were grown in 6 L of Luria Broth (LB) medium with 100 µg/ml ampicillin at 37°C. When the OD600 reached 0.5, the culture was treated with 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG) to induce Rv0678 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 for 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 Ni2+-chelating column (GE Healthcare Biosciences, Pittsburgh, PA) 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 Rv0678 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% SDS-PAGE 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.

Crystallization of Rv0678 – All crystals of the 6xHis Rv0678 regulator were obtained using hanging-drop vapor diffusion. The Rv0678 crystals were grown at room temperature in 24-well plates with the following procedures. A 2 µl protein solution containing 20 mg/ml Rv0678 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 28% polyethylene glycol (PEG) 1000, 0.1 M Na-acetate (pH 4.0), 0.2 M NaCl and 10% glycerol. The resultant mixture was equilibrated against 500 µl of the reservoir solution. Crystals grew to a full size in the drops within two weeks. Typically, the dimensions of the crystals were 0.2 mm x 0.05 mm x 0.05 mm. Cryoprotection was achieved by raising the PEG 1000 concentration stepwise to 35% with a 3.5% increment in each step. Crystals of the tungsten derivative was prepared by incubating the crystals of Rv0678 in solution containing 28% PEG 1000, 0.1 M Na-acetate (pH 4.0), 0.2 M NaCl and 10% glycerol and 1 mM (NH4)6W6O24([μ-O]6([μ-Cl])6Cl6 for 24 hours at 25°C.

Data collection, structural determination and refinement – All diffraction data were collected at 100K 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 (23). The crystals of Rv0678 belong to the space group P1 (Table 1). Based on the molecular weight of Rv0678 (18.34 kDa), the asymmetric unit is expected to contain four regulator molecules with a solvent content of 45.26%. Six tungsten cluster sites were identified using SHELXC and SHELXD (24) as implemented in the HKL2MAP package (25). Single isomorphous replacement with anomalous scattering (SIRAS) was employed to obtain
experimental phases using the program MLPHARE (26, 27). The resulting phases were then subjected to density modification and NCS averaging using the program PARROT (28). The phases were of excellent quality and allowed for tracing of most of the molecule in PHENIX AutoBuild (29), which led to an initial model with over 90% amino acid residues containing side-chains. The remaining part of the model was manually constructed using the program Coot (30). Then, the model was refined using PHENIX (29) leaving 5% of reflections in Free-R set. Iterations of refinement using PHENIX (29) and CNS (31) and model building in Coot (30) lead to the current model, which consists of two dimers (totally 587 residues in the asymmetric unit) with excellent geometrical characteristics (Table 1).

Identification of fortuitous ligand – To identify the nature of the bound ligand in crystals of Rv0678, we used gas chromatography coupled with mass spectrometry (GC-MS). The Rv0678 crystals were extensively washed with the crystallization buffer and transferred into deionized water. The mixture was then incubated at 100°C for 5 min, then chloroform was added into the mixture to a final concentration of 80% (v/v) to denature the protein and allow for the extraction of ligand. GC-MS analysis indicated that the mass of the bound ligand was octadecanoic acid, 2-hydroxyethyl-1-(hydroxymethyl)ethyl ester, also called 2-palmitoylglycerol.

Virtual ligand screening using AutoDock Vina – AutoDock Vina (32) was used for virtual ligand screening of a variety of compounds. The docking area was assigned visually to cover the internal cavity of the Rv0678 dimer. A grid of 35 Å × 35 Å × 35 Å with 0.375 Å spacing was calculated around the docking area for all atom types presented in the DrugBank (33) and ZINC (34) libraries using AutoGrid. The iterated local search global optimizer algorithm was employed to predict the binding free energies for these compounds.

Isothermal titration calorimetry for ligand binding – We used isothermal titration calorimetry (ITC) to determine the binding affinity of 1-stearoyl-rac-glycerol (an isomer of 2-palmitoylglycerol) to the purified Rv0678 regulator. Measurements were performed on a VP-Microcalorimeter (MicroCal, Northampton, MA) at 25 °C. Before titration, the protein was thoroughly dialyzed against buffer containing 10 mM Na-phosphate pH 7.2, 100 mM NaCl and 0.001% DDM. The protein concentration was determined using the Bradford assay. The dimeric Rv0678 sample was then adjusted to a final concentration of 400 µM and served as the titrant. The ligand solution contained 20 µM 1-stearoyl-rac-glycerol, 10 mM Na-phosphate pH 7.2, 100 mM NaCl and 0.001% DDM. The protein and ligand samples were degassed before they were loaded into the cell and syringe. Binding experiments were carried out with the ligand solution (1.5 ml) in the cell and the protein solution as the injectant. Ten microliter injections of the ligand solution were used for data collection.

Injections occurred at intervals of 300 s, and the duration time of each injection was 20 s. Heat transfer (µcal/s) was measured as a function of elapsed time (s). The mean enthalpies measured from injection of the ligand in the buffer were subtracted from raw titration data before data analysis with ORIGIN software (MicroCal). Titration curves were fitted by a nonlinear least squares method to a function for the binding of a ligand to a macromolecule. Nonlinear regression fitting to the binding isotherm provided us with the equilibrium binding constant ($K_d = 1/K_D$) and enthalpy of binding ($\Delta H$). Based on the values of $K_d$, the change in free energy ($\Delta G$) and entropy ($\Delta S$) were calculated with the equation: $\Delta G = -RT\ln K_d = \Delta H - T\Delta S$, where $T$ is 273 K and $R$ is 1.9872 cal/K mol. Calorimetry trials were also carried out in the absence of Rv0678 in the same experimental conditions. No change in heat was observed in the injections throughout the experiment.

Electrophoretic mobility shift assay – Probes were amplified from the H37Rv genome using the primers listed in Table 2. All probes were labeled with Digoxigenin using the Roche DIG Gel Shift kit. For EMSA analysis, 12 nM Dig-labeled probe and the indicated micromolar concentrations of protein were incubated for 45 minutes at room temperature in the Roche binding buffer modified by the addition of 0.25 mg/mL herring sperm
DNA, and 0.75 mg/mL poly(d[I-C]). For ligand competition studies, 1-steraryl-rac-glycerol (an isomer of 2-palmitoylgllycerol) (Sigma-Aldrich) was resuspended in hot acetone and added to EMSA reactions at 1 µM final concentration. Competition reactions were performed at 37°C.

All reactions were resolved on a 6% native polyacrylamide gel in TBE buffer, transferred to nylon membrane and Dig-labeled DNA-protein complexes detected following the manufacturer’s recommendations. Chemiluminescent signals were acquired using an ImageQuant LAS 4000 (GE).

**Dye primer based DNase I footprint assay** – DNase I footprinting was performed as described by Zianni et al. (35). The 296 bp Rv0678-mpmpS5 probe was generated by PCR using the primers 6FAM-Rv0678-F and HEX-Rv0678-R. Gel purified, fluorescently labeled probe (0.6 pmol) was incubated with either 1 µM Rv0678 or BSA for 30 min at room temperature in standard EMSA binding buffer. After incubation, 10 mM MgCl2 and 5 mM CaCl2 were added to the reaction mixture in a final volume of 50 µl. Then, 0.0025 U of DNase I (Thermo) was added and incubated for 5 minutes at room temperature. Digested DNA fragments were purified with QIAqick PCR Purification columns (Qiagen) and eluted in 20 µl water. Digested DNA samples were analyzed at the Center for Genome Research and Biocomputing at Oregon State University. Purified DNA (2 mL) was mixed with HiDi formamide and GeneScan-500 LIZ size standards (Applied Biosystems) and analyzed using an Applied Biosystems 3730 DNA analyzer.

The 296 bp fragment was sequenced with the primers 6FAM-Rv0678-F and HEX-Rv0678-R, respectively, using the Thermo Sequenase Dye Primer Manual Cycle Sequencing Kit 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. Samples were analyzed using the 3730 DNA analyzer and electropherograms aligned using the GENEMAPPER software (version 5.0, Applied Biosystems).

**Site-directed mutagenesis** – Site-directed point mutations on residues D90 and R92, which are expected to be critical for DNA binding, were performed to generate the single point mutants D90A and R92A. The primers used for these mutations are listed in Table 3. All oligonucleotides were purchased from (Integrated DNA Technologies, Inc., Coralville, IA) in a salt-free grade.

**Fluorescence polarization assay for DNA binding** – Fluorescence polarization assays were used to determine the affinity for DNA binding by Rv0678 and its mutants. Both the 26-bp oligodeoxynucleotide and fluorescein labeled oligodeoxynucleotide were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). These oligodeoxynucleotides contain the consensus 18-bp putative promoter DNA sequence (TTTCAGAGTACAGTGAAA) for Rv0678. The sequences of the oligodeoxynucleotides were 5’- CAGATTTCAGAGTACGAACTTG-3’ and 5’-CAAGTTTCACTGTACTGGAATCTG-3’, where F denotes the fluorescein which was covalently attached to the 5’ end of the oligodeoxynucleotide by a hexamethylene linker. The 26-bp fluoresceinated ds-DNA 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, 5 nM fluoresceinated DNA, and 1 µg of poly(dI-dC) as non-specific DNA. The protein solution containing 2,500 nM dimeric Rv0678 or Rv0678 mutant and 5 nM fluoresceinated DNA was titrated into the DNA binding solution until the millipolarization (mP) become unchanged. All measurements were performed at 25°C using a PerkinElmer LS55 spectrofluorometer equipped with a Hamamatsu R928 photomultiplier. The excitation wavelength was 490 nm, and the fluorescence polarization signal (in ΔP) was measured at 525 nm. Each titration point recorded was an average of 15 measurements. Data were analyzed using the equation, P = ([P_{bound} - P_{free}]/[protein] + [protein]) + P_{free}, where P is the polarization measured at a given total protein concentration, P_{free} is the initial polarization of free fluorescein-labeled DNA, P_{bound} is the maximum.
polarization of specifically bound DNA, and [protein] is the protein concentration. The titration experiments were repeated for three times to obtained the average K_D value. Curve fitting was accomplished using the program ORIGIN (OriginLab Corporation, Northampton, MA).

RESULTS AND DISCUSSION

Overall structure of Rv0678 – M. tuberculosis Rv0678 belongs to the MarR family of regulators. It possesses 165 amino acids, sharing 14% and 15% protein sequence identity with MarR (22) and OhrR (36) (Fig. 1). The crystal structure of Rv0678 was determined to a resolution of 1.64 Å using single isomorphous replacement with anomalous scattering (SIRAS) (Table 1). Four molecules of Rv0678 are found in the asymmetric unit, which assemble as two independent dimers (Fig. 2). Superimposition of these two dimers gives an RMSD of 0.8 Å over 271 C- atoms, indicating that their conformations are nearly identical to each other.

The structure of Rv0678 (Fig. 3) is quite distinct in comparison with the known structures of the MarR-family regulators (22, 36-39). Each subunit of Rv0678 is composed of six α-helices and two β-strands: α1 (17-31), α2 (36-47), α3 (55-62), α4 (66-79), β1 (82-85), β2 (94-97), α5 (101-127) and α6 (132-160) (Fig. 1). The monomer is L-shaped with the shorter side forming a DNA-binding domain. However, the longer side contributes to an extended long arm, creating a dimerization domain for the regulator. Residues 34-99, which include α2, α3, α4, β1 and β2, are responsible for constructing the DNA-binding domain. The dimerization domain of Rv0678 is generated by residues 16-32 and 101-160, which cover α1, α5 and α6 of the protomer. Each protomer of Rv0678 is approximately 55 Å tall, 35 Å wide and 35 Å thick.

As a member of the MarR family of regulators, the DNA-binding domain of Rv0678 features a typical winged helix-turn-helix (wHh) binding motif. The two anti-parallel β1 and β2 strands are found to generate a β-hairpin structure, which also forms the wing of the DNA-binding domain. The crystal structure of the OhrR-DNA complex (36) showed that this β-hairpin directly participates to contact the double-stranded DNA and is critical for repressor-operator interactions. Another important component of the wHh motif for DNA recognition is helix α4. In the OhrR-DNA complex (36), the corresponding α-helix is found to bind within the deep major groove of the B-DNA. Protein sequence alignment suggests that Rv0678 contains three conserved amino acids common among members of the MarR family. These three residues, R84, D90 and R92, are located within the DNA-binding domain of the regulator (Fig. 1), and are likely important for protein-DNA interactions. Among them, D90 and R92 are positioned within the β-hairpin of the wing. The corresponding amino acids located at the winged loop region of the ST1710 regulator play a major role in regulator-promoter interactions (39).

The Rv0678 crystal structure reveals that helices α1, α5 and α6 are involved in the formation of the dimer. Specifically, helices α5, α6, α5’ and α6’ (where the prime denotes the next subunit) form intertwined helical bundles and constitute the dimerization domain. Helices α6 and α6’ are oriented in an anti-parallel fashion and form the scaffold of the dimer (Fig. 3). Extensive hydrophobic interactions are observed at the interface between the two subunits of the regulator. In addition, Y147 and Y159’ and their identical counter pair perform aromatic stacking interactions, securing the dimeric organization. Additional salt bridges between R32 and E115’ and between E106 and R109’ (as well as their counter pairs) stabilize the binding.

Perhaps the most striking difference between the structures of Rv0678 and other MarR-family members is the relative orientation of the DNA-binding and dimerization domains. The structures of MarR (22), OhrR (36) and MexR (37, 38) suggest that helices α4 and α4’ orient approximately perpendicular to the pseudo two-fold axis of the dimeric regulators. However, our crystal structure of Rv0678 depicts these two helices more or less in parallel with the dimer’s pseudo two-fold axis. Similar orientation of helix α4 has also been found in the structure of the
Vibrio cholerae AphA transcriptional activator (40). This conformation is not compatible and does not allow the regulator to interact with the B-form DNA. To bind its cognate DNA, the Rv0678 regulator must undergo a large conformational movement that reorients the DNA-binding domain such that the positions of helices α4 and α4′ can be matched with the two consecutive major groove of the promoter DNA. Based on the OhrR-DNA (36) and ST1710-DNA (39) crystal complexes, we predict that the entire DNA-binding domain of Rv0678, including α2, α3, α4, β1 and β2, has to rotate downward by ~70° with respect to α5 of the dimerization domain before DNA binding (Fig. 4). If this is the case, then the loop region between β2 and α5 form the hinge for this rotational motion.

Rv0678 was liganded – Unexpectedly, a large extra electron density was found at the interface between the DNA-binding and dimerization domains of Rv0678, indicating the existence of a fortuitous bound ligand co-purified and co-crystallized with the regulator (Fig. 5). Thus, this region is also a substrate-binding site. To identify the unknown bound ligand, gas chromatography coupled with mass spectrometry (GC-MS) was applied to investigate the Rv0678 crystals (Fig. 6). The result suggests that the fortuitous ligand is 2-palmitoylglycerol, also called octadecanoic acid, 2-hydroxy-1-(hydroxymethyl)ethyl ester, which contains 21 carbons with the molecular formula C₂₁H₄₂O₄. That this fatty acid glycerol ester is co-purified with the Rv0678 regulator suggests that fatty acid glycerol esters may be the natural substrates for this protein.

The propanetriol of the bound 2-palmitoylglycerol is completely buried within the dimer interface, leaving the tail portion of its elongated octadecanoate hydrophobic carbon chain oriented at the entry point of this binding site. This orientation facilitates the contribution of R32 and E106′ to form two hydrogen bonds with the glycerol head group of the fatty acid. The backbone oxygen of F79′ also participates to create the third hydrogen bond with this glycerol head group. In addition, the carbonyl oxygen of the octadecanoate group contributes to make another hydrogen bond with R109, securing the binding. Interestingly, Rv0678 further anchors the bound fatty acid molecule through hydrophobic interactions with residues F79, F79′ and F81′. Therefore, the binding of 2-palmitoylglycerol in Rv0678 is extensive: Within 4.5 Å of the bound fatty acid glycerol ester, 20 amino acids contact this molecule (Table 4). It should be noted that residues F79, F79′ and F81′ belong to helices α4 and α4′. In the OhrR-DNA structure (36), the corresponding α4 and α4′ helices were buried within the two consecutive major grooves, directly contacting the promoter DNA. Thus, we suspect that helices α4 and α4′ have dual responsibilities in the Rv0678 regulator. They form the DNA-binding site for operator DNA as well as substrate-binding site for inducing ligands.

In the second Rv0678 dimer of the asymmetric unit, it is also found that a 2-palmitoylglycerol molecule is bound within the corresponding substrate-binding site. Residues contributed to form this binding site are nearly identical but with a slightly different subset of amino acids in comparison with those of the first Rv0678 dimer described above (Table 4).

Virtual ligand library screening – Virtual ligand screening was then performed to elucidate the nature of protein-ligand interactions in the Rv0678 regulator. The 2-palmitoylglycerol binding site was chosen as a substrate binding cavity for this docking study. AutoDock Vina (32) was used to screen small molecules listed in the DrugBank (33) and ZINC (34) libraries. Vina utilizes the iterated local search global optimizer algorithm, which results in predicted binding free energies for these compounds ranging from -13.8 to +20 kcal/mol. Of the 70,000 screened compounds, it is predicted that the best substrate for Rv0678 is the heterocyclic compound diethyl-[(5E)-5-(6,8,9,10-tetrahydro-5H-benzo[c]xanthen-11-ylmethylene)-7,8-dihydro-6H-xanthen-3-yl]-yli. Table 5 lists the top three substrates, which have the lowest predicted binding free energies, for the Rv0678 regulator.

As the crystal structure of Rv0678 depicts that a fatty acid glycerol ester is bound within the substrate binding site of this regulator, Vina (32) was also used to examine if these fatty acids are
able to interact with Rv0678. As a positive control, the molecule 2-palmitoylglycerol was docked into the substrate-binding site of this regulator, resulting in a predicted binding free energy of -7.6 kcal/mol. Vina was then used to screen for 2,500 different fatty acids. Based on the lowest predicted binding free energies, the top three compounds in this class was selected and listed in Table 6, where 18-[8-chloro-1-(hydroxymethyl)-6-phenyl-4H-[1,2,4]triazolo[4,3-a][1,4]benzodiazepin-4-yl]octadecanoic acid is the best compound for Rv0678 binding among these fatty acids.

Rv0678-ligand interaction – The binding affinity of 1-stearoyl-rac-glycerol for the Rv0678 regulator was then determined using is ITC, which obtained a binding affinity constant, $K_a$, of $4.9 \pm 0.4 \times 10^5 \text{M}^{-1}$. The titration is characterized by a negative enthalpic contribution, which gives rise to a hyperbolic binding curve (Fig. 7). The thermodynamic parameters of binding of 1-stearoyl-rac-glycerol to Rv0678 display enthalpic ($\Delta H$) and entropic ($\Delta S$) contributions of $-1.0 \pm 0.1$ kcal/mol and 22.5 cal/mol/deg$^2$, respectively. Interestingly, the molar ratio for this binding reaction based on ITC is one Rv0678 dimer per ligand. This ligand-binding experiment confirms that Rv0678 is capable of recognizing fatty acid glycerol esters.

Electrophoretic mobility shift assay – To demonstrate direct transcriptional regulation, we performed EMSAs using a probe corresponding to the intergenic region between mmpS5 and rv0678 (Fig. 8a). This probe shifted in a concentration-dependent manner (Fig. 8b). This result is consistent with previous reports of altered mmpS5/mmpL5 gene expression in M. bovis BCG spontaneous rv0678 mutants (13). Preliminary CHIPSeq data from the TB Systems Biology Consortium suggests Rv0678 regulates expression of additional genes (41). We designed additional probes to experimentally demonstrate binding of Rv0678 to the promoter regions of mmpS2-mmpL2, mmpS4-mmpL4 and rv0991-0992. Probes are depicted schematically in Fig. 8a. We also saw concentration-dependent binding of Rv0678 to these two probes (Fig. 8b). As a control, EMSAs were performed in the presence of non-labeled probes. Release of Dig-labeled probe was observed consistent with specific binding of Rv0678 to the rv0678-mmpS5, rv0505-mmpS2 and mmpL4 probes (Fig. 8c). Using the sequence of the six probes that shifted, we identified a putative consensus binding sequence for Rv0678 using the MEME algorithm (42) (Fig. 8e). Rv0678 co-crystallized with a ligand whose binding renders the protein unable to bind DNA. Addition of 1-stearoyl-rac-glycerol (an isomer of 2-palmitoylglycerol) to the EMSA reaction buffer reduced Rv0678 binding to a target promoter probe (Fig. 8c).

Dye primer based DNase I footprint assay – To further refine the binding site of Rv0678 in the rv0678-mmpS5 intergenic region, DNase I footprint assay was performed on the Rv0678-mmpS5 probe using established methods (35). Electropherograms in Figure 9 show the DNA sequence bound by Rv0678. The control protein BSA did not result in DNA protection at the same concentration. Interestingly, the region bound by Rv0678 includes the start codon of the rv0678 gene (underlined nucleotides, Fig. 9b). The bound sequence contains a potential inverted repeat motif (GAACGTACAGATTTCA $N_8$ ... TGAAACTTGTAGCGTCGCAAC).

Rv0678-DNA interaction – Fluorescence polarization-based assay was carried out to study the interaction between Rv0678 and the 26-bp DNA containing the 18-bp putative promoter DNA sequence (TTCAGAGTACAGTGAAA). Our footprint assay has suggested this promoter DNA sequence was protected by the Rv0678 regulator. Fig. 10a illustrates the binding isotherm of Rv0678 in the presence of 5 nM fluoresceinated DNA. The titration experiment indicated that this regulator binds the 26-bp promoter DNA with a dissociation constant, $K_D$, of $19.6 \pm 3.0$ nM. The binding data also indicate that Rv0678 binds its cognate DNA with a stoichiometry of one Rv0678 dimer per ds-DNA.

In addition, fluorescence polarization was used to determine the binding affinities of this 26-bp DNA by the Rv0678 mutants D90A and R92A. These two residues are located within the β-hairpin of the wHtH motif of the N-terminal DNA-binding domain. In ST1710, the corresponding
two residues are critical for regulator-promoter interactions. Interestingly, our measurements indicate that the $K_D$ values of the D90A-DNA and R92A-DNA complexes are 113.3 $\pm$ 16.8 and 86.0 $\pm$ 7.4 nM (Fig. 10b and c), revealing the DNA binding affinities for these mutants are significantly weaker than that of the native Rv0678 regulator. Like ST1710, our experimental results suggest that residues D90 and R92 are important for DNA recognition. 

With the rising incidence of drug resistant strains of M. tuberculosis, it is increasingly important to understand the molecular mechanisms underlying virulence and drug resistance of this pathogen. This knowledge will inform the development of new strategies to combat TB. In this report, we describe the crystal structure the Rv0678 transcriptional regulator, which controls the expression level of the MmpS5-MmpL5, MmpS4-MmpL4 and MmpS2-MmpL2 transport systems. MmpS4 and MmpS5 contribute to siderophore export, but the substrate of MmpL2 is not known (15). Fortuitously, the structure of Rv0678 was resolved in complex with a 2-palmitoylglycerol molecule, suggesting that fatty acid glycerol esters are the natural substrates for the Rv0678 transcriptional regulator. Further work is required to demonstrate if this ligand is structurally related to the substrate of either efflux system, or how its availability changes in different environments and mycobacterial growth phases. The crystal structure of the 2-palmitoylglycerol-Rv0678 complex probably provides a snapshot of the ligand-binding state of this regulator, whereby both the DNA-binding and dimerization domains are recruited to participate for ligand binding. In this case, the DNA-binding domain must bend upward and shift towards the dimerization domain to accommodate the bound ligand. As crystallized, the regulator is incompatible with the operator DNA. When the inducing ligand is removed from the ligand-binding site, freeing helices $\alpha 4$ and $\alpha 4'$ to rotate downward and shift away from the dimerization domain, this conformational state should be compatible with the B-DNA and allow for DNA binding.
REFERENCES

1. Maartens, G., and Wilkinson R. J. (2007) Tuberculosis. The Lancet 370, 2030-2043.
2. World Health Organization (2010) Fact sheet no. 104: tuberculosis. Available at http://www.who.int/mediacentre/factsheets/fs104/en/index.html.
3. Frieden, T. R., Sterling, T., Pablos-Mendez, A., Kilburn, J. O., Cauthen, G. M., and Dooley, S.W. (1993) The emergence of drug-resistant tuberculosis in New York City. N. Engl. J. Med. 328, 521-556.
4. Pillay, M., and Sturm, A. W. (2007) Evolution of the extensively drug-resistant F15/LAM4/KZN strain of Mycobacterium tuberculosis in KwaZulu-Natal, South Africa. Clin. Infect. Dis. 45, 1409-1414.
5. Goldman, R. C., Plumley, K. V., and Laughon, B. E. (2007) The evolution of extensively drug resistant tuberculosis (XDR–TB): history, status and issues for global control. Infect. Disord. Drug Targets 7, 73-91.
6. Udwadia, Z. F., Amale, R. A., and Rodrigues, C. (2012) Totally drug-resistant tuberculosis in India. Clin. Infect. Dis. 54, 579-581.
7. Iseman M. D. (1993) Treatment of multidrug-resistant tuberculosis. N. Engl. J. Med. 329, 784–791.
8. Domenech, P., Reed, M. B., and Barry, C. E. III. (2005). Contribution of the Mycobacterium tuberculosis MmpL Protein Family to Virulence and Drug Resistance. Infect Immun. 73, 3492-3501.
9. Jain, M., and Cox, J. S. (2005). Interaction between polyketide synthase and transporter suggests coupled synthesis and export of virulence lipid in M. tuberculosis. PLoS Pathog. 1, 12-19.
10. Cox, J. S., Chen, B., McNeil, M., and Jacob, W. R. Jr. (1999) Complex lipid determines tissue-specific replication of Mycobacterium tuberculosis in mice. Nature 402, 79-83.
11. Brennan, P. J., and Nikaido, H. (1995) The envelope of mycobacteria. Annu. Rev. Biochem. 64, 29-63.
12. Converse, S. E., Mougous, J. D., Leavell, M. D., Leary, J. A., Bertozzi, C. R., and Cox, J. S. (2003) MmpL8 is required for sulfolipid-1 biosynthesis and Mycobacterium tuberculosis virulence. Proc. Natl. Acad. Sci. USA 100, 6121-6126.
13. Milano, A., Pasca, M. R., Provvedi, R., Lucarelli, A. N., Manina, G., Ribeiro, A. L., Manganelli, R., and Riccardi, G. (2009) Azole resistance in Mycobacterium tuberculosis is mediated by the MmpS5–MmpL5 efflux system. Tuberculosis 89, 84-90.
14. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E. et al. (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature 393, 537–544.
15. Wells, R. M., Jones, C. M., Xi, Z., Speer, A., Danilchanka, O. et al. (2013) Discovery of a siderophore export system essential for virulence of Mycobacterium tuberculosis PLoS Pathog. 9, 1-14.
16. Pacheco, S., Hsu, F.-F., Powers, K. M., and Purdy, G. E. (2013) MmpL11 transports mycolic acid-containing lipids to the mycobacterial cell wall and contributes to biofilm formation in M. smegmatis. J. Biol. Chem. 288, 24213-24222.
17. Jacobs, W. R., Cox, J. S., Chen, B., and McNeil, M. (1999). Complex lipid determines tissue-specific replication of Mycobacterium tuberculosis in mice. Nature 402, 79-83.
18. Tekaia, F., Gordon, S. V., Garnier, T., Brosch, R., Barrell, B. G., and Cole, S. T. (1999) Analysis of the proteome of Mycobacterium tuberculosis in silico. Tub. Lung. Dis. 79, 329–342.
19. Deshayes, C., Bach, H. Euphrasie, D., Attarian, R., Coureuil, M, Sougakoff, W. Laval, F., Av-Gay, Y., Daffé, M. Etienne, G. and Reyrat, J. M. (2010) MmpS4 promotes glycopeptidolipids biosynthesis and export in Mycobacterium smegmatis. Mol. Microbiol. 78, 989-1003.
20. Bolla, J. R., Do, S. V., Long, F., Dai, L., Su, C.-C., Lei, H. T., Chen, X., Gerkey, J. E., Murphy, D.
Structure of the transcriptional regulator Rv0678

C., Rajashankar, K. R., Zhang, Q., and Yu, E.W. (2012) Structural and functional analysis of the transcriptional regulator Rv3066 of Mycobacterium tuberculosis. Nucleic Acids Res. 40, 9340-9355.

21. Ellison, D. W., and Miller, V. L. (2006) Regulation of virulence by members of the MarR/SlyA family. Curr. Opin. Microbiol. 9, 153-159.

22. Alekshun, M. N., Levy, S. B., Mealy, T. R., Seaton, B.A., and Head, J. F. (2001) The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3Å resolution. Nat. Struct. Biol. 8, 710-714.

23. Otwinowski, Z., and Minor, M. (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307-326.

24. Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD. Acta Crystallogr. D58, 1772–1779.

25. Pape, T., and Schneider, T. R. (2004) HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. J. Appl. Crystallogr. 37, 843-844.

26. Otwinowski, Z. (1991) ML-PHARE, CCP4 Proc. 80-88 (Daresbury Laboratory, Warrington, UK).

27. Collaborative Computational Project No. 4. (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D50, 760-763.

28. Cowtan, K. (2010) Recent developments in classical density modification. Acta Crystallogr. D66, 470-478.

29. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W. et al. (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. 58, 1948-1954.

30. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D60, 2126.

31. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr. D54, 905-921.

32. Trott, O., and Olson, A. J. (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comp. Chem. 31, 455-461.

33. Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., Djoumbou, Y., Eisner, R., Guo, A. C., and Wishart, D. S. (2011) DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. Nucl. Acids Res. 39, D1035-D1041.

34. Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S., and Coleman, R.G. (2012) ZINC: A Free Tool to Discover Chemistry for Biology. J. Chem. Inf. Model. 52, 1757-1768.

35. Ziani, M., Tessanne, K., Merighi, M., Laguna, R. and Tabita, F.R. (2006) Identification of the DNA bases of a DNase I footprint by the use of dye primer sequencing on an automated capillary DNA analysis instrument. J. Biomolecular Techniques 17, 103-113.

36. Hong, M., Fuangthong, M., Helmann, J. D., and Brennan, R. G. (2005) Structure of an OhrR-OhrA Operator Complex Reveals the DNA Binding Mechanism of the MarR Family. Mol. Cell 20, 131-141.

37. Lim, D., Poole, K., and Strynadka, N. C. J. (2002) Crystal structure of the MexR repressor of the
Structure of the transcriptional regulator Rv0678

mexRAB–oprM multidrug efflux operon of Pseudomonas aeruginosa. *J. Biol. Chem.* **277**, 29253-29259.

38. Wilke, M. S., Heller, M., Creagh, A. L., Haynes, C. A., McIntosh, L. P., Poole, K., and Strynadka, N. C. J. (2008) The crystal structure of MexR from *Pseudomonas aeruginosa* in complex with its antirepressor ArmR. *Proc. Natl. Acad. Sci. USA* **39**, 14832-14837.

39. De Silva, R. S., Kovachikova, G., Lin, W., Taylor, R. K., Skorupski, K., and Kull, F. J. (2005) Crystal structure of the virulence gene activator AhpA from *Vibrio cholerae* reveals it is a novel member of the winged helix transcription factor superfamily. *J. Biol. Chem.* **280**, 13779-13783.

40. Kumarevel, T., Tanaka, T., Umehara, T., and Yokoyama, S. (2009) ST1710–DNA complex crystal structure reveals the DNA binding mechanism of the MarR family of regulators. *Nucl. Acids Res.* **37**, 4723-4735.

41. Reddy, T. B., Riley, R., Wymore, F., Montgomery, P., Decaprio, D., Engels, R., Gellesch, M., Hubble, J., Jen, D., Jin H, Koehrsen, M., Larson, L., Mao, M., Nitzberg, M., Sisk, P., Stolte, C., Weiner, B., White, J., Zachariah, Z. K., Sherlock, G., Galagan, J. E., Ball, C. A., and Schoolnik, G. K. (2009) TB database: an integrated platform for tuberculosis research. *Nucl. Acids Res.* **37**, D499-D508.

42. Bailey, T. L., and Elkan, C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28-36.
Acknowledgements – This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines of the Advanced Photon Source, supported by an award GM103403 from the National Institutes of General Medical Sciences. Use of the Advanced Photon Source is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. We are grateful to Louis Messerle at University of Iowa for providing us the (NH₄)₂W₆(µ-O)₆(µ-Cl)₆Cl₂ complex used in this study.

FOOTNOTES

*This work was supported by NIH Grants R01AI087840 (G.E.P.) and R01GM086431 (E.W.Y.). Coordinates and structural factors for the structure of Rv0678 have been deposited at the RCSB Protein Data Bank with an accession code 4NB5.

ψA.R. and N.K. contributed equally to this work.

ςTo whom correspondence should be addressed: Dept. of Chemistry, Dept. of Physics & Astronomy, Iowa State University, Ames, IA 50011. Tel.:515-294-4955; E-mail:ewyu@iastate.edu.

FIGURE LEGENDS

Figure 1. Protein sequence alignment of the MarR family of regulators. Alignment of the amino acid sequences of *Mycobacterium tuberculosis* Rv0678, *Bacillus subtilis* OhrR, *Pseudomonas aeruginosa* MexR, *Escherichia coli* MarR and *Sulfolobus tokodaii* ST1710. The alignment is done using FFAS03. The topology of *M. tuberculosis* Rv0678 is shown at the top. The three conserved amino acids are highlighted with yellow bars.

Figure 2. Stereo view of the experimental electron density maps of Rv0678 at a resolution of 1.64 Å. (a) The electron density maps are contoured at 1.2 σ. The Cα traces of the two Rv0678 dimers in the asymmetric unit are in yellow, light blue, tv red and lime green. Anomalous signals of the six W₆(µ-O)₆(µ-Cl)₆Cl₂²⁺ cluster sites (contoured at 4 σ) found in the asymmetric unit are colored red. (b) Representative section of electron density in the vicinity of helices α1 and α2. The solvent-flattened electron density (50-1.64 Å) is contoured at 1.2 σ and superimposed with the final refined model (green, carbon; red, oxygen; blue nitrogen; yellow, sulfur).

Figure 3. Structure of the *M. tuberculosis* Rv0678 regulator. (a) Ribbon diagram of a protomer of Rv0678. The molecule is colored using a rainbow gradient from the N-terminus (blue) to the C-terminus (red). (b) Ribbon diagram of the Rv0678 dimer. Each subunit of Rv0678 is labeled with a different color (yellow and orange). The bound 2-palmitoyleglycerol within the dimer is shown in sphere form (gray, carbon; red, oxygen). The figure was prepared using PyMOL (http://www.pymol.sourceforge.net).

Figure 4. Rigid-body rotation of the DNA-binding domain of Rv0678. This is a schematic representation illustrating the conformational change of Rv0678 between the ligand bound and unbound structures. Helices α4 and α4’ of the DNA-binding domain are indicated. The ligand is colored blue.

Figure 5. Simulated annealing electron density maps and the 2-palmitoyleglycerol binding site. (a) Stereo view of the simulated annealing electron density map of the bound 2-palmitoyleglycerol within the Rv0678 dimer (the orientation corresponds to the side view of Fig. 1b). The bound 2-palmitoyleglycerol is shown as a stick model (green, carbon; red, oxygen). The simulated annealing 2Fᵣ⁻Fₑ electron density map is contoured at 1.2 σ (blue mesh). The left and right subunits of Rv0678 are shown as orange and
yellow ribbons. (b) The 2-palmitoylglycerol binding site. Amino acid residues within 3.9 Å from the bound 2-palmitoylglycerol (green, carbon; red, oxygen) are included. The side chains of selected residues from the right subunit of Rv0678 in Figure 1b are shown as yellow sticks (yellow, carbon; blue, nitrogen; red, oxygen). Residues from the next subunit of Rv0678 are shown as orange sticks (orange, carbon; blue, nitrogen; red, oxygen). (c) Schematic representation of the Rv0678 and 2-palmitoylglycerol interactions. Amino acid residues within 4.5 Å from the bound 2-palmitoylglycerol are included. Dotted lines depict the hydrogen bonds. The hydrogen-bonded distances are also indicated in this figure.

Figure 6. Identification of the fortuitous ligand by GC-MS. (a) Electron ionization (EI) spectrum of the strongest GC peak at 14.45 min. (b) GC-MS spectrum of octadecanoic acid, 2-hydroxyl-1-(hydroxymethyl)ethyl ester from the internal GC-MS library. The ligand was identified as 2-palmitoylglycerol.

Figure 7. Representative isothermal titration calorimetry for the binding of 1-stearoyl-rac-glycerol to Rv0678. (a) Each peak corresponds to the injection of 10 µl of 400 µM dimeric Rv0678 in buffer containing 10 mM Na-phosphate (pH 7.2), 100 mM NaCl and 0.001% DDM into the reaction containing 20 µM 1-stearoyl-rac-glycerol in the same buffer. (b) Cumulative heat of reaction is displayed as a function of the injection number. The solid line is the least-square fit to the experimental data, giving a $K_D$ of $4.9 \pm 0.4 \times 10^7$ M$^{-1}$.

Figure 8. Rv0678 binds to promoter regions of mmpS2-mmpL2, mmpS4-mmpL4, mmpS5 and rv0991-2c. (a) A schematic depicting the DNA probes used in electrophoretic mobility shift assays (EMSAs) to examine the promoter and intragenic regions of the mmpS2-mmpL2, mmpL3, mmpS4-mmpL4, mmpS5-mmpL5 and rv0991-2c genes. (b) EMSAs were performed using 12 nM Dig-labeled probe and the indicated micromolar concentrations of protein. An arrow denotes the shifted probes. (c) To demonstrate specificity, EMSAs were performed in the presence of non-labeled (“cold”) probe. Reactions were performed with 6 nM Dig-labeled probe, the indicated micromolar concentrations of protein, and 0.6 µM cold probe. The asterisk notes the accumulation of free Dig-labeled probe. (d) EMSAS were performed using 12 µM Dig-labeled probe and 6 µM Rv0678 in the presence or absence of 1 µM 1-stearoyl-rac-glycerol as indicated above the blot. (e) The sequence of the probes bound by Rv0678 in panel b and c were compared using the motif-based sequence analysis tool MEME, yielding a putative Rv0678 binding motif.

Figure 9. Direct binding of Rv0678 to the rv0678-mmpS5 intergenic region by dye primer based DNase I footprint assay. Electropherograms indicating the protection pattern of the Rv0678-mmpS5 probe after digestion with DNase I following incubation with (a) 0, (b) 1 µM Rv0678 and (c) 1 µM BSA are shown. The protected DNA sequence is indicated above the electropherogram in panel b, the predicted start codon of rv0678 is underlined.

Figure 10. Representative fluorescence polarization of Rv0678. (a) The binding isotherm of Rv0678 with the 26-bp DNA containing the 18-bp promoter sequence, showing a $K_D$ of 19.6 ± 3.0 nM. (b) The binding isotherm of mutant D90A with the 26-bp DNA, showing a $K_D$ of 113.3 ± 16.8 nM. (c) The binding isotherm of mutant R92A with the 26-bp DNA, showing a $K_D$ of 86.0 ± 7.4 nM. 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.
Table 1. Data collection, phasing and structural refinement statistics of Rv0678.

| Data set       | Rv0678       | W$_6$(µ-O)$_6$(µ-Cl)$_6$Cl$_6^{2-}$ derivative |
|----------------|--------------|-----------------------------------------------|
| **Data Collection** |             |                                               |
| Wavelength (Å) | 0.98         | 0.98                                          |
| Space group    | $P1$         | $P1$                                          |
| Resolution (Å) | 50 – 1.64    | 50 – 1.90                                    |
|                | (1.70-1.64) | (1.97-1.90)                                  |
| Cell constants (Å) |           |                                               |
| a              | 54.54        | 54.75                                         |
| b              | 57.24        | 57.49                                         |
| c              | 61.44        | 61.42                                         |
| $\alpha$, $\beta$, $\gamma$ (°) | 82.2, 68.4, 72.2 | 82.3, 68.5, 72.4 |
| Molecules in ASU | 4            | 4                                             |
| Redundancy     | 2.0 (2.0)    | 1.9 (1.8)                                    |
| Total reflections | 326,940       | 512,196                                      |
| Unique reflections | 80,449       | 52,208                                       |
| Completeness (%) | 97.5 (95.6)    | 88.4 (90.1)                                  |
| $R_{sym}$ (%)  | 4.4 (39.5)   | 9.1 (35.3)                                   |
| $I / \sigma(I)$ | 17.46 (2.2)  | 14.29 (3.4)                                  |
| **Phasing**    |              |                                               |
| Number of sites |              | 6                                             |
| Phasing power (acentric) | 1.71       |                                               |
| $R_{cullis}$ (acentric) | 0.70        |                                               |
| Figure of merit (acentric) | 0.66     |                                               |
| **Refinement** |              |                                               |
| Resolution (Å) | 50 – 1.64    |                                               |
| $R_{work}$     | 16.28        |                                               |
| $R_{free}$     | 19.44        |                                               |
| Average B-factor (Å$^2$) | 23.85 |                                               |
| RMSD bond lengths (Å) | 0.011  |                                               |
| RMSD bond angles (°) | 1.253 |                                               |
| **Ramachandran plot** |          |                                               |
| most favoured (%) | 96.7    |                                               |
| additional allowed (%) | 3.3   |                                               |
| generously allowed (%) | 0     |                                               |
| disallowed (%) | 0            |                                               |
### Table 2. Primers.

| Probe     | Primer 1         | Primer 2                              |
|-----------|------------------|---------------------------------------|
| Rv0678    | CTTCGGAAACCAAGAAAGTG | CCAACCGAGTCAAAACTCCTG               |
| Rv0505    | GAACACGAGGGTGAGGATG | GCGTCGTCTCGACCGTGAC               |
| Rv0991-2  | GAGCTGGTTGACTTCTCG | CAATGCGGTCGGCGTGTTG               |
Table 3. Primers for site-directed mutagenesis.

|                | Sequence                                                                 |
|----------------|--------------------------------------------------------------------------|
| D90A-forward   | 5'-CGCCTGGGAGCTGGTGCTCGTCGTCACGTATTTTCGTC-3'                              |
| D90A-reverse   | 5'-GACGAAAATACGTGGAGGAGCACCAGGACTGCCAGGCG-3'                              |
| R92A-forward   | 5'-GCAGTGGTGGAGCACCAGGATATAATTATCGTCTGGC-3'                               |
| R92A-reverse   | 5'-GCAGACGAAAATACGTGGCACGATACCCAGGACTG-3'                                |
Table 4. Rv0678-ligand contacts.

| Residue-ligand contacts | Dimer 1  | Dimer 2  |
|-------------------------|----------|----------|
|                         | Distance (Å) | Distance (Å) |
| R32                     | 3.2*      | -        |
| Q78                     | 3.9       | 3.7      |
| F79                     | 3.8       | 4.2      |
| E108                    | 3.4       | 3.2      |
| R109                    | 2.8*      | 3.2*     |
| R111                    | 3.4       | 3.5      |
| A112                    | 4.0       | 3.6      |
| M113                    | -         | 4.4      |
| E115                    | 3.0       | 2.9      |
| L116                    | 4.4       | 3.7      |
| L144                    | -         | 4.4      |
| L145                    | -         | 4.0      |
| Y28'                    | 4.0       | 3.9      |
| F29'                    | 4.4       | 4.3      |
| R32'                    | 3.5       | 3.6      |
| L34'                    | 4.2       | 3.6      |
| F79'                    | 2.8*      | -        |
| F81'                    | 3.4       | 3.5      |
| F102'                   | 4.5       | 2.3      |
| A103'                   | -         | 4.4      |
| G105'                   | 2.9       | 3.0      |
| E106'                   | 3.2*      | 3.1*     |
| E108'                   | 3.9       | -        |
| R109'                   | 3.8       | 3.5      |

* Hydrogen bond distance
Contacts within 4.5Å are listed
Table 5. Top three ligands for the Rv0678 regulator.

| Ligand                                                                 | Structure of ligand                                                                 | Binding affinity (kcal/mol) |
|------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-----------------------------|
| Diethyl-[(5E)-5-(6,8,9,10-tetrahydro-5H-benzo[c]xanthen-11-ylmethylene)-7,8-dihydro-6H-xanthen-3-yl] | ![Structure of ligand](image1.png)                                                  | -13.8                      |
| (5E)-N,N-diethyl-5-(6,8,9,10-tetrahydrochromeno[3,2-c]chromen-11-ylmethylene)-7,8-dihydro-6H-xanthen | ![Structure of ligand](image2.png)                                                  | -13.5                      |
| Dibenzylfluorescein                                                    | ![Structure of ligand](image3.png)                                                  | -13.4                      |
Table 6. Top three fatty acids for the Rv0678 regulator.

| Fatty acid                                                                 | Structure of fatty acid                                                                 | Binding affinity (keal/mol) |
|----------------------------------------------------------------------------|---------------------------------------------------------------------------------------|----------------------------|
| 18-[8-chloro-1-(hydroxymethyl)-6-phenyl-4H-[1,2,4]triazolo[4,3-a][1,4]benzodiazepin-4-yl]octadecanoic acid | ![Structure 1](image1.png)                                                             | -10.6                      |
| 2-[2-oxo-2-[4-[2-(1-phenyltetrazol-5-yl)sulfanylacetyl]anilino]ethyl]octadecanoic acid | ![Structure 2](image2.png)                                                             | -10.5                      |
| 12-(anthracene-1-carbonyl)octadecanoic acid                                | ![Structure 3](image3.png)                                                             | -10.4                      |
Figure 2a
Figure 5b

- Q78
- F79
- R32
- E108
- R109
- G105
- R109'
- R32'
- F79'
- E108'
- F81'

Distances:
- 3.2 Å
- 2.8 Å
Figure 5c
Figure 6b

(mainlib) Octadecanoic acid, 2-hydroxy-1-(hydroxymethyl)ethyl ester

Relative Abundance

m/z
Figure 7

(a) Graph showing Time (min) vs. µcal/sec with data points and a trend line.

(b) Graph showing Molar Ratio vs. kcal/mole of injectant with data points and a trend line.
Figure 9

(a) [Graph A]

(b) [Graph B]

(c) [Graph C]

Sequence: GACGCATACCGAACGTCACAGATTTCAGAGTACAGTGAAACTTGTCGACGGGGTCGATC
Figure 10a
