doi:10.1016/j.jmb.2009.02.058 J. Mol. Biol. (2009) 387, 1250–1260

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

# Structure and Function of the Macrolide Biosensor Protein, MphR(A), with and without Erythromycin

**Jianting Zheng[1]†, Vatsala Sagar[1,2]†, Adam Smolinsky[1], Chase Bourke[1], Nicole LaRonde-LeBlanc[1,2]* and T. Ashton Cropp[1]***

[1]*Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20742, USA*

[2]*Center for Biomolecular Structure and Organization, University of Maryland, College Park, MD 20742, USA*

*Received 25 November 2008; received in revised form 12 February 2009; accepted 16 February 2009*
Available online 2 March 2009

***Edited by K. Morikawa***

The regulatory protein MphR(A) has recently seen extensive use in synthetic biological applications, such as metabolite sensing and exogenous control of gene expression. This protein negatively regulates the expression of a macrolide 2′-phosphotransferase I resistance gene (*mphA*) via binding to a 35-bp DNA operator upstream of the start codon and is de-repressed by the presence of erythromycin. Here, we present the refined crystal structure of the MphR(A) protein free of erythromycin and that of the MphR(A) protein with bound erythromycin at 2.00- and 1.76-Å resolutions, respectively. We also studied the DNA binding properties of the protein and identified mutants of MphR(A) that are defective in gene repression and ligand binding in a cell-based reporter assay. The combination of these two structures illustrates the molecular basis of erythromycin-induced gene expression and provides a framework for additional applied uses of this protein in the isolation and engineered biosynthesis of polyketide natural products.



*Keywords:* macrolide antibiotic; repressor; biosensor; erythromycin; *Streptomyces*

## Introduction

Macrolide antibiotics, such as the natural product erythromycin and its semisynthetic derivatives azithromycin and clarithromycin, are among the most widely used treatments of bacterial infections today. Resistance to these compounds can be mediated by several different mechanisms, including target modification and efflux pumps.[1–3] Another strategy that bacteria use is the direct modification of the compounds by a macrolide 2′-phosphotransferase, encoded by the *mphA* gene, which renders the antibiotics inactive by phosphorylation of the 2′-hydroxyl group on the desosamine sugar.[4] This modification destroys a key interaction that the molecules have with the drug target, the ribosome. It has been shown that the expression of this resistance gene is controlled at the transcriptional level by a repressor protein, MphR(A), which is de-repressed in the presence of erythromycin.[5] The genes encoding these resistance and regulatory proteins are clustered together on a gene cassette that has been isolated from several clinical and environmental sources[6,7] and appears to be a mobile resistance element. Thus, bacteria harboring this gene cluster have the ability to "sense" the presence of macrolide antibiotics and then respond to this threat by enzymatic inactivation.

Based on sequence homology, the MphR(A) protein is predicted to be a member of the TetR class of transcriptional regulators.[8] Proteins in this class typically contain an N-terminal helix–turn–helix (HTH) motif responsible for DNA binding and repression followed by a C-terminal ligand binding domain. Upon ligand binding, the DNA operator is released, allowing free transcription of the downstream genes. Members of this family of proteins show similar, homodimeric, and entirely α-helical folds, but with the sequence homology confined to the N-terminus of the protein. Intriguingly, the ligand binding domains of this class of proteins display similar folds despite showing little, if any, sequence homology. This sequence diversity of the ligand binding domains may explain the extremely wide range of small molecules that serve as ligands for the TetR family of proteins. Previous

**Corresponding authors.* E-mail addresses: nlaronde@umd.edu; acropp@umd.edu.

† J.Z. and V.S. contributed equally to this work.

Abbreviations used: HTH, helix–turn–helix; SeMet, selenomethionine; SAD, single-wavelength anomalous diffraction; TEV, tobacco etch virus.

DNA footprinting experiments have indicated that MphR(A) binds 2173 bp upstream of its own start codon to a 35-bp region that completely overlaps the P*mphA* promoter[5] and therefore blocks transcription in the absence of ligand but activates expression in the presence of erythromycin. The genes in the cluster are all transcribed in the same direction with an apparent organization such that MphR(A) will repress its own expression in a negative feedback loop[5] (Fig. 1).

Recently, the MphR(A) protein has seen extensive use in synthetic biological applications. For example, Weber *et al.*[9] and Kramer *et al.*[10] used chimeras of this protein as inducible gene control systems for mammalian tissue culture and in mice. These and related hybrid systems have also been used *in vitro* as biological tests for the presence of antibiotics (erythromycin) in milk with very sensitive and rapid readouts.[11,12] Finally, it has been utilized in its native context expressed in *Escherichia coli* as a cellular biosensor to detect the presence of macrolide antibiotics in *Streptomyces* fermentation extracts.[13] This approach has the potential to benefit the drug discovery process by allowing researchers to rapidly screen soil and culture extracts for this class of antibiotics prior to full structural identification. Interestingly, the protein shows a versatile response to a broad range of ligands and is activated by 12-, 14-, 15-, and 16-membered macrolactone antibiotics, yet this activation is independent of the biological action of these molecules. Among the compounds that have been shown to serve as ligands for MphR (A) are erythromycin (**1**), azithromycin (**2**), methymycin (**3**), and josamycin (**4**) (Fig. 1), all with varying degrees of activation. While these molecules share some structural similarities, such as a common desosamine sugar residue and macrocyclic lactones, there are obvious size and shape differences. To understand the structural basis for this plasticity of macrolide recognition and to set the stage for the directed evolution of new protein biosensors and related synthetic applications, we solved the crystal structure of the MphR(A) protein with and that of the MphR(A) protein without the prototypical ligand erythromycin A. The protein was first expressed as a selenomethionine (SeMet) derivative, and the structures were determined using the single-wavelength anomalous diffraction (SAD) method. The data were then used to solve the native structures by molecular replacement. In addition to the structures, we examined the properties of the protein using *in vitro* biochemical experiments and an *in vivo* genetically engineered reporter circuit.

## Results

### Structure of MphR(A) with erythromycin

We purified the MphR(A) protein to homogeneity as an N-terminally His-tagged protein that was then proteolytically removed to generate protein ready for crystallization trials. Attempts to crystallize MphR(A) were initially more successful when the protein solution contained the ligand erythromycin, with diffraction-quality crystals growing in approximately 3 days. Therefore, we first obtained the structure of SeMet-substituted MphR(A) complexed with erythromycin by using SAD phasing. The phases obtained from the SeMet structure were combined with a higher-resolution data set obtained from native crystals for refinement. Crystal and refinement statistics are shown in Table 1. We observed a single dimer per asymmetric unit. Electron density was observed and modeled for residues 5–189 in one subunit and for residues 6–189 in the second subunit, out of a total of 194 residues each. In addition to the protein, the ligand-bound structure shows clearly defined electron density for erythromycin A located in a binding pocket. Each monomer of the MphR(A) dimer (Fig. 2) consists of nine α-helices and thus resembles the architecture of prototypical proteins in the class—TetR,[14] EthR,[15] and QacR,[16] among others. The helices of MphR span residues 9–22 (α1), 30–37 (α2), 41–48 (α3), 51–72 (α4), 80–90 (α5), 99–111 (α6), 118–133 (α7), 145–160 (α8), and 165–181 (α9). A

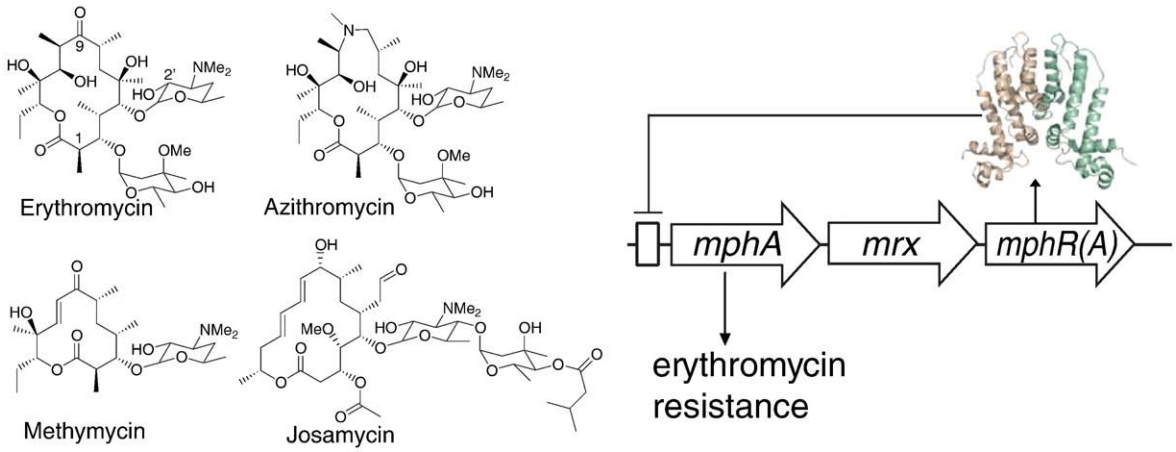


**Fig. 1.** Structures of known MphR(A) ligands and organization of the erythromycin resistance gene cassette. MphA is an erythromycin 2′-phosphotransferase. Mrx is of unknown function.

**Table 1.** Data collection and refinement statistics

| | MphR(A)–SeMet | MphR(A)–erythromycin | MphR(A)–APO |
|---|---|---|---|
| Protein Data Bank code | | 3FRQ | 3G56 |
| *Data collection* | | | |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ | $P2_1$ |
| Cell dimensions | | | |
| *a*, *b*, *c* (Å) | 44.4, 61.1, 134.7 | 44.3, 61.2, 134.6 | 37.4, 113.8, 43.4 |
| α, β, γ (°) | 90.0, 90.0, 90.0 | 90.0, 90.0, 90.0 | 90.0, 93.9, 90.0 |
| Molecules per asymmetric unit | 2 | 2 | 2 |
| Wavelength (Å) | 0.9795 | 0.9795 | 1.000 |
| Resolution (Å) | 50–2.0 | 50–1.76 | 50–2.1 |
| $R_{sym}$ (last shell) | 0.070 (0.321) | 0.064 (0.419) | 0.040 (0.467) |
| $I/\sigma I$ (last shell) | 51.2 (8.1) | 30.0 (4.6) | 30.2 (2.1) |
| Completeness (%) (last shell) | 99.7 (97.4) | 99.5 (97.9) | 98.5 (90.7) |
| Redundancy (last shell) | 7.3 (7.0) | 7.8 (7.4) | 4.1 (3.6) |
| No. of reflections | 45,208 | 35,182 | 20,802 |
| *Refinement* | | | |
| Resolution (Å) | | 30–1.76 | 57–2.1 |
| $R_{work}/R_{free}$ (%) | | 18.1/22.1 | 22.0/28.4 |
| No. of protein residues in monomers A/B/C/D/E/F/G/H | | 184/183 | 176/179 |
| No. of non-protein atoms | | | |
| Ligand | | 102 | 0 |
| Solvent | | 273 | 84 |
| Mean *B*-factors (Å$^2$) | | 30.0 | 48.1 |
| RMSD | | | |
| Bond lengths (Å) | | 0.017 | 0.015 |
| Bond angles (°) | | 1.569 | 1.419 |

structural alignment of the coordinates using Secondary Structure Matching search algorithm[17] shows the highest homology (RMSD of 2.49 Å for 156 matched $C^\alpha$ atoms in the monomer) to a recently reported regulator, IcaR from *Staphylococcus epidermidis*[18] (Protein Data Bank entry 2ZCM), despite very limited primary sequence identity (14.8%). IcaR is also reported to be involved in "antibiotic sensing," albeit on a different class (gentamicins), and is implicated in controlling biofilm formation. Similar alignment scores are obtained with other members in the class, such as TetR (3.26 Å, 135 matched $C^\alpha$ atoms), EthR (2.98 Å, 147 matched $C^\alpha$ atoms), and QacR (2.76 Å, 144 matched $C^\alpha$ atoms).

The N-terminal portion of MphR(A) is the presumed DNA binding domain and consists of an HTH motif that spans residues 8–50. The distance between the two HTH motifs across the dimeric interface is 46.3 Å, as measured by the distance between the two $C^\alpha$ atoms of L44, a residue that is conserved among many TetR family members. This distance is comparable with the same distances in the other liganded structures of QacR and ActR, which are 44 and 46 Å, respectively. This distance is assumed to be larger when ligand is bound and thus is a main contributing factor to a decrease in affinity for the DNA operator in the presence of ligand. The bottom portion of the HTH motif consists of hydrophobic and conserved positively charged residues, including K35, R41, and R51, consistent with an interaction with the negatively charged phosphate backbone of DNA (Fig. 3).

MphR(A) binds one molecule of erythromycin per monomer in a large hydrophobic binding pocket. This space is lined by residues donated from helices α4–α8 and the dimeric interface of the other monomer. This pocket accepts the entire ligand molecule, including most of the desosamine and cladinose sugar residues. The entrance to this ligand binding pocket is a short tunnel (Fig. 2) that opens to a cavity lined with 17 residues, most of which are hydrophobic. These residues make intimate contact with the bound erythromycin ligand and contain side chains within 4 Å (Fig. 3). While these contacts are mostly hydrophobic interactions, there is a small patch of positive charge, derived from R122, N123, and H147, that neutralizes a region of partial negative charge of erythromycin. Direct hydrogen bonds inside this pocket are formed between N123 and the C-9 and C-11 oxygen atoms of erythromycin, H147 and the C-9 carbonyl, and R122 and the C-1 carbonyl. In addition, further stabilization is provided by water-mediated interactions between the C-12 hydroxyl group of erythromycin and the backbone carbonyl of A151 and the hydroxyl of Y103. A chloride ion was placed at a position of high positive electron density (>10 σ in difference density map) that interacts between the desosamine moiety of the erythromycin and the backbone amide of N94. Hydrogen bonds between K21 and the erythromycin cladinose O-6 and 2′-hydroxyl and between S106 and the C-1 carbonyl are seen in only one of the two monomers. Intriguingly, the ε-nitrogen of K21 is located 4.6 Å from the 2′-hydroxyl of the desosamine residue of erythromycin in the second monomer. This 2′-hydroxyl is the site of modification by the MphA phosphotransferase protein that provides the actual erythromycin resistance. Based on this structure, it is plausible that the *product* of the resistance protein is actually a better ligand for the

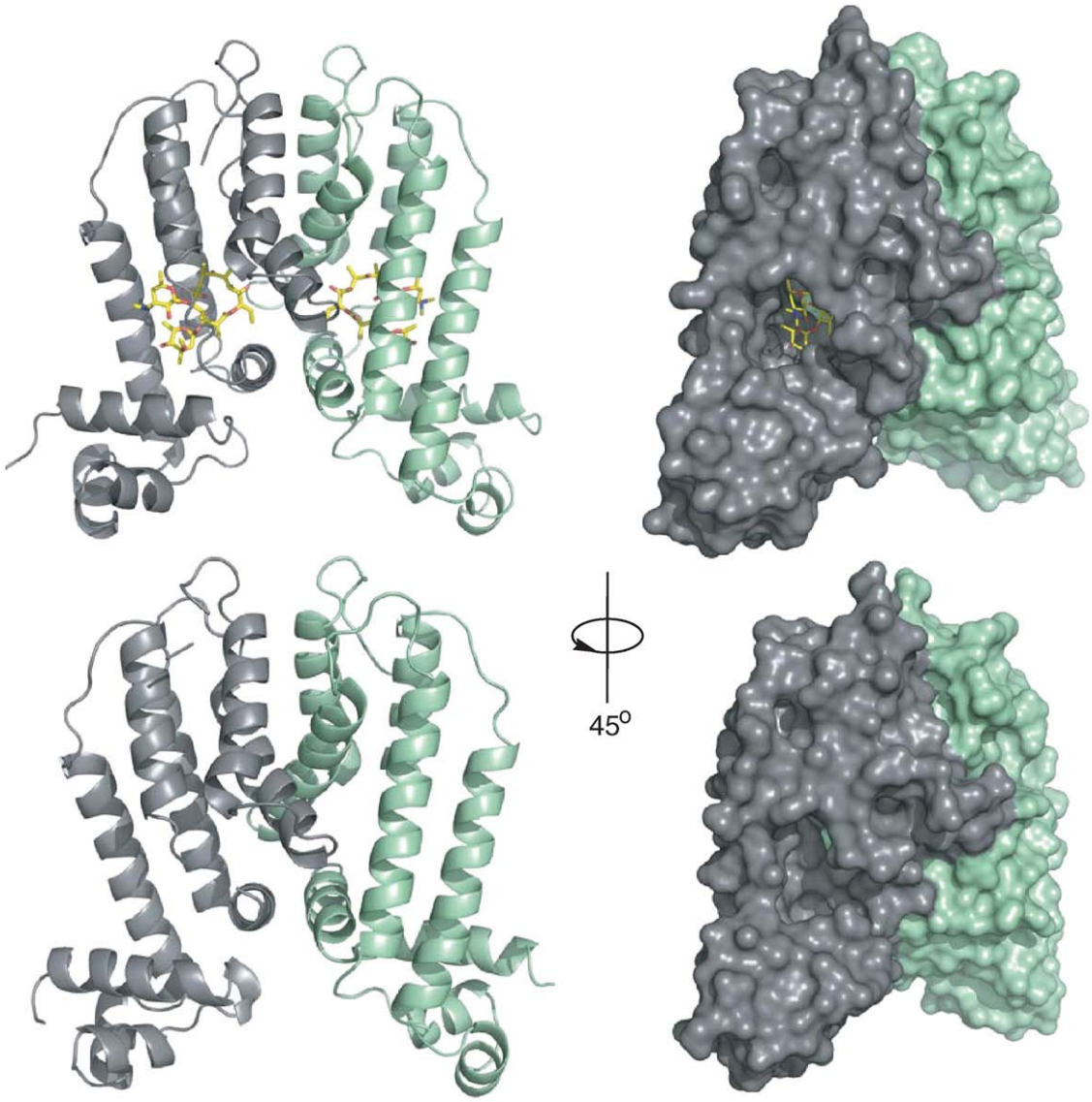


**Fig. 2.** Structures of MphR(A) with (top) and without (bottom) erythromycin. Surface rendering shows the entrance tunnel to the ligand binding pocket. Figures were made using PyMOL.

MphR protein than erythromycin due to a favorable electrostatic interaction between a 2′-O-phosphate and K21 via ionic interaction.

### Structural effects of ligand binding

The structure of MphR(A) without ligand was solved by molecular replacement using the liganded structure. The unliganded structure was modeled for residues 8–94 and 101–190 for one monomer and for residues 6–95 and 99–186 for the second monomer. Residues 95–100 for the first monomer and residues 96–98 for the second monomer could not be modeled due to poor electron density. The structures of the erythromycin-bound and -free MphR(A) aligned with an RMSD of 0.86 Å for 298 $C^{\alpha}$ atoms when the dimer was aligned by least-squares methods and with an RMSD of 1.028 Å for 173 atoms if only monomer A was aligned. Less $C^{\alpha}$ atoms per monomer are compared in the overall dimer comparison due to default rejection of atoms with RMSD values greater than 4 Å, which occurs with greater frequency when differences are minimized over two molecules instead of one. The observed structural difference between the liganded and unliganded structures is due to a significant conformational shift that results in changes in the positioning of one monomer relative to the other in the dimer and a large shift in the relative positioning of the HTH DNA binding domains. When measured from the distance between the two $C^{\alpha}$ atoms of L44 as above, the HTH domains are separated by a distance of 42.9 Å, shorter than that observed with ligand bound (Fig. 4). This suggests that, as in the case of homologous proteins, the change in distance of the DNA binding domain results in concomitant change in affinity for the DNA and thereby de-repression. The conformational differences between the two molecules also result in a larger cavity volume in the unliganded state compared with the

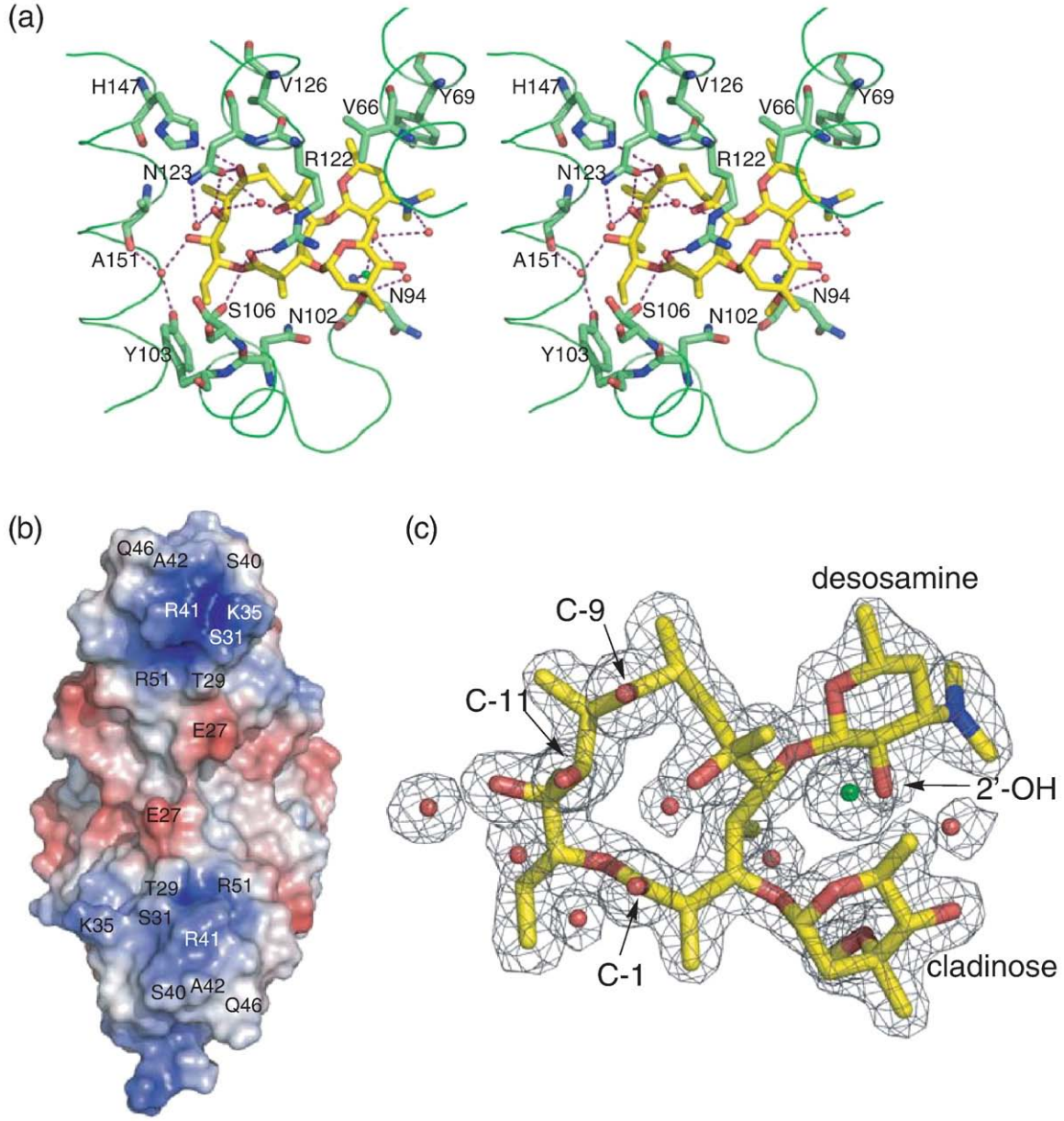


**Fig. 3.** Interactions of MphR(A). (a) Stereo image of the ligand binding site of MphR(A) with bound erythromycin. Polar contacts within 4 Å of erythromycin are shown as dotted lines. (b) Rendering of surface charge of the MphR(A) molecule as viewed from the bottom of the HTH motif with surface residues labeled. (c) The $2F_o - F_c$ electron density of the bound erythromycin ligand contoured at 1 σ is shown. Numbering of erythromycin atoms is mentioned in the text. Water molecules and chloride ion are shown as red and green spheres, respectively.

erythromycin-bound MphR(A). In the liganded MphR(A) dimer, the volumes of the two cavities, calculated with the protein atoms only using the program VOIDOO (see Experimental Procedures for reference), are 1183 and 1234 Å$^3$. In the unliganded dimer, the two cavities have volumes of 2273 and 1756 Å$^3$. These differences are consistent with the cavity closing around the ligand (709 Å$^3$) upon binding, which, in turn, results in a movement of the DNA binding domains toward the ligand binding sites and away from each other. There are several molecular interactions that appear to mediate this change. On helix 1, there is a hydrophobic patch facing the ligand binding domain composed of T17 and L20. Upon ligand binding, this collapses to a favorable interaction with the hydrophobic regions of the cladinose sugar residue. The movement of helix 1 upon ligand binding is also mediated by a new favorable ionic interaction formed between E14 and R61. Both of these residues are ordered and in close proximity in the presence of ligand; however, R61 is disordered in the apo structure, and therefore, this interaction is less significant in the absence of ligand.

### Correlation of *in vitro* and *in vivo* MphR(A) function

While the DNA binding properties of MphR(A) have been observed,[5] accurate thermodynamic

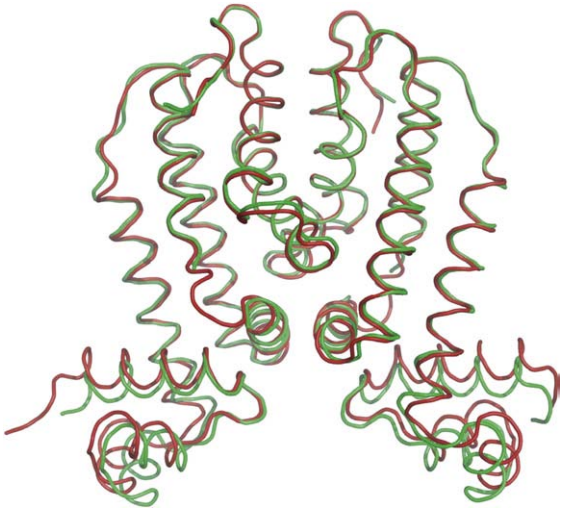

**Fig. 4.** Superimposition of the backbone structures of MphR(A) with (red) and without (green) erythromycin.

characterization including dissociation constants has not been reported. We chose to use fluorescence polarization to determine the $K_d$ of the protein for a segment of DNA harboring the binding site that was covalently labeled with fluorescein. As can be seen in the binding isotherm shown in Fig. 5, a titration experiment indicated that MphR binds its operator with a dissociation constant $K_d$ of 574±29 nM when treating the dimer as the biological unit. This value is a surprisingly low-affinity interaction when compared with other members of the protein class. Based on this rather low observed interaction for the DNA operator, we decided to create a synthetic gene circuit in which we could observe the repressor activity *in vivo*. In this system, MphR(A) is overexpressed from the strong, constitutive P(Lac) IQ promoter and then, in turn, controls expression of a reporter gene, *lacZ*, expressed from the P*mphA* promoter. This system would be expected to create an excess of repressor protein in comparison with the template plasmid DNA to compensate for a rather weak affinity.

In these cells [Genehogs with pBAD/MphR(A)/P*mphA*-lacZ], we could monitor the functionality of the MphR(A) protein based on de-repression of *lacZ*, resulting in galactosidase activity that is easily assayed (Fig. 5).[19] In the absence of erythromycin, this system shows ~100% repression of the reporter gene to background levels (30 Miller units), indicating that despite the apparent low affinity of MphR(A) for its operator, when constitutively expressed in cells, it is sufficient for inhibition of transcription. Reporter gene activation increased in an erythromycin-dependent fashion and *plateaued* to 1250 Miller units, or a 41-fold activation of gene expression, at an erythromycin concentration of 125 μM. Indeed, a plate-based assay shows that filter disks containing erythromycin show a "halo" effect of gene expression according to the concentration gradient. We also observed that in this synthetic gene circuit, it is not possible to achieve 100% activation when compared with a control circuit lacking the *mphR(A)* gene, which displayed 4000 Miller units under the same conditions (data not shown). This is most likely due to the inability to achieve high intracellular concentrations of the antibiotic in this system without toxicity. Cells that are engineered to express an alternative erythromycin resistance gene might alleviate this problem and allow circuit performance over expanded ranges of ligand concentrations.

## Mutations that affect MphR(A) function

To investigate the consequence of mutations of MphR(A), we first generated an error-prone PCR library targeted to the *mphR(A)* gene and then screened that library in the context of the β-galactosidase reporter strain for mutations that destroy gene repression. Colonies harboring misfolded MphR proteins or those that cease to bind the promoter are expected to be blue when grown on X-gal, even in the absence of erythromycin. We picked some of the perpetually blue colonies, sequenced the corresponding mutants, and then analyzed gene repression using Miller assays. As can be seen in Fig. 5, L30H, L44S, and A153T display high galactosidase activity across all ranges of erythromycin concentrations (>3000 Miller units). These activities correspond to a nonrepressed control plasmid that lacks the *mphR(A)* gene and indicate that the protein

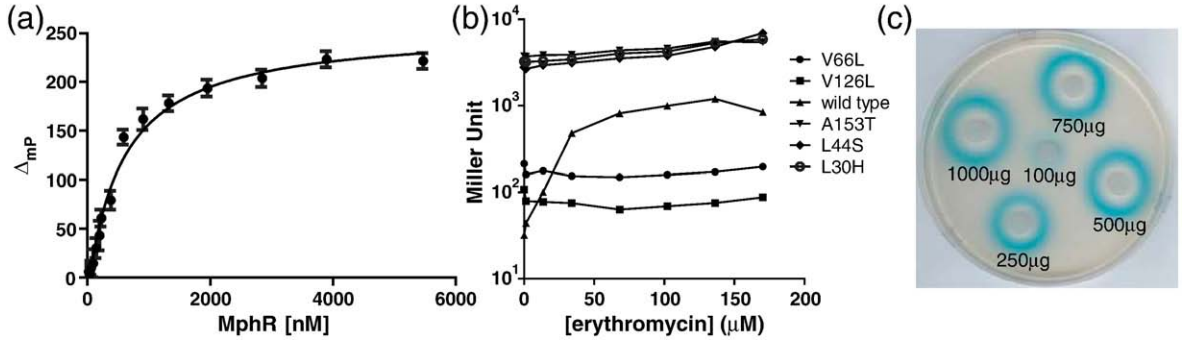


**Fig. 5.** *In vitro* and *in vivo* analyses of MphR(A). (a) Binding isotherm of MphR(A) as determined by fluorescence polarization using fluorescein-labeled DNA. Error bars represent 1 SD of each concentration. (b) Plot of reporter gene expression as a function of erythromycin concentration. (c) Plate-based assay of the biosensor strain.

is nonfunctional. Based on the crystal structure, the location of these mutations is sensible as both L30 and L44 are located in the interior hydrophobic core of the HTH motif thought to bind DNA. A153 is located at the dimeric interface of the two MphR(A) monomers, and therefore, larger residues that disrupt this interface may not be tolerated. Clearly, these mutations represent a select few of the many possibilities that could render the protein unable to block transcription of the reporter gene.

Mutations with altered ligand binding properties that retain repression activity are less frequent. We therefore chose a more conservative approach and, with the use of the crystal structure, identified two amino acid residues (V66 and V126) that might potentially be mutated to disrupt erythromycin binding. We chose to mutate each of these residues to leucine with the rationale that this would be a conservative hydrophobic mutation, but the larger side chain would sterically block the binding site. In contrast to the above mutations, the colonies are perpetually white, and, indeed, both the V66L and V126L mutants show activities of 149 and 63 Miller units, respectively, in the presence of 125 μM erythromycin. This indicates that rationally designed mutations do not affect the ability to repress gene expression but rather the affinity of the protein for erythromycin.

## Discussion

Initially, members of the TetR family of proteins were identified in association with antibiotic resistance genes. Recently, there is increasing evidence that these proteins can also control aspects of secondary metabolism and are even activated by metabolic intermediate and precursor ligands.[20] Intriguingly, a BLAST-P analysis of MphR(A) yields homologues that are classified as "TetR family members" derived from *Myxococcus*, *Salinispora*, and *Streptomyces* species, all of which are known to devote approximately 10% of their genome to secondary metabolism. As can be seen in Fig. 6, an alignment with these homologues suggests that they may function in a similar fashion, too. Nearly all homologues, including the prototypical member TetR itself, contain a hydrophobic residue that corresponds to L20 of MphR(A), which presumably mediates an interaction between helix 1 and the ligand binding pocket. The proteins from *Myxococcus* and *Rhodopseudomonas* share the conserved residues E14 (or D14) and R61, which form the favorable ionic interaction between helices 1 and 4 upon ligand binding. In addition, residues R122 and N123, which, in the case of MphR(A), mediate important interactions with the macrolide ligand, are conserved. It is possible that these proteins recognize the same or similar ligands to erythromycin. Clearly, in addition to a wealth of biosynthetic capability, these organisms appear to have diverse sensory systems that serve to regulate secondary metabolism and environmental defenses. In nature, the production of secondary metabolites and concomitant resistance are surely linked, and the structure and function of TetR-like proteins are a prime example.

The MphR(A) protein is reported to recognize macrolide antibiotics of different sizes, including those that are substantially larger than erythromycin. The crystal structure presented here suggests that those larger ligands might be accommodated with the sugar residues extending from the entrance tunnel, and, indeed, there is no clear polar interaction with the desosamine residue of erythromycin that would suggest otherwise. It is not clear that the sugar residues on these macrolides are critical for ligand

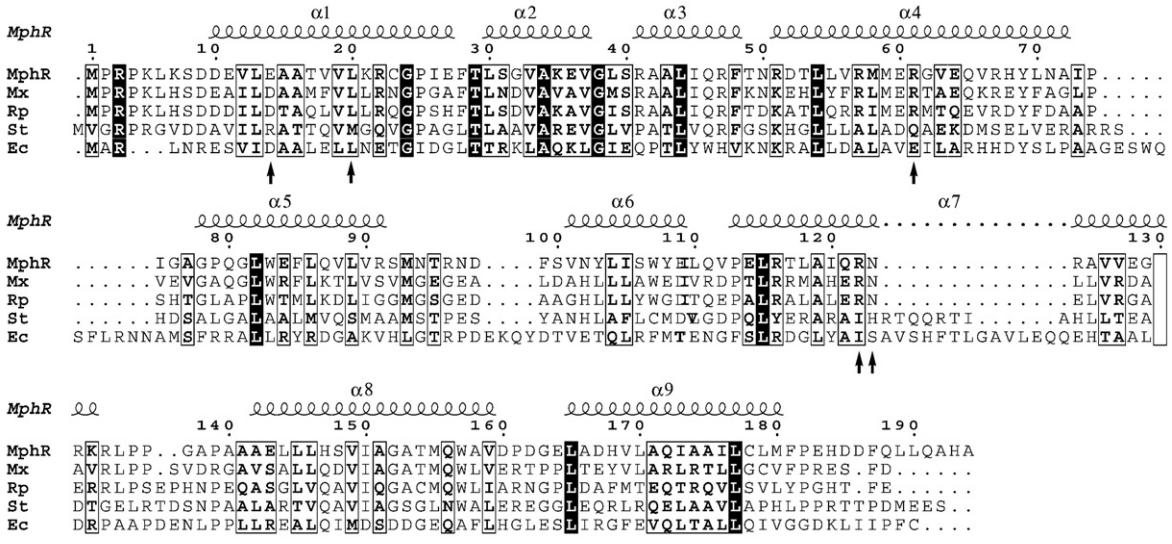


**Fig. 6.** CLUSTALW alignment of homologues of MphR. Mx, *Myxococcus xanthus* (accession no. YP_633146); Rp, *Rhodopseudomonas palustris* (accession no. YP_ABE40088); St, *Salinispora tropica* (accession no. YP_001159116); Ec, TetR from *E. coli*. Black boxes indicate complete identity across all proteins; open boxes indicate similarity. The residues mentioned in the text (E14, L20, R61, R122, and N123) are indicated by arrows.

recognition. The polar interactions seen in the ligand binding site suggest that ligands must contain oxygen atoms at C-1, C-9, and C-11 for proper recognition. Based on the structure, the biosynthetic precursor aglycone of erythromycin, 6-deoxyerythronolide B, may also serve as a ligand. Likewise, the product of the resistance enzyme, erythromycin 2′-O-phosphate, could serve as a ligand, or an *improved* ligand, which would lead to a cascade activation of the operon. The biosensor systems such as those reported here and elsewhere[13] allow a rapid assay of molecules that do or do not activate MphR (A), and experiments to test these possibilities are underway and will be reported in due course.

In addition to illustrating the natural system, the crystal structures and genetic screen presented here set the stage for altering the specificity and sensitivity of the MphR(A) genetic circuit. In our initial biosensor circuit, the *mphR(A)* gene is constitutively expressed and the reporter protein is controlled by the promoter of the natural circuit. One could envision creating synthetic promoter organizations that contain, for example, multiple MphR(A) binding sites to modulate sensitivity or cascade systems that amplify the readout of MphR(A)-based gene control, creating a more "digital" rather than graded response.[21,22] Recently, Resch *et al*. reported a double-sieved selection process through which the ligand preference of TetR was modified, including an interesting mutant that shows an inverted phenotype in which tetracycline actually serves as a co-repressor.[23] Typically, these approaches are based on creating protein libraries based on randomization methods such as error-prone PCR and DNA shuffling. However, the illustration here of the MphR(A) protein ligand binding site makes it possible to create site-saturation libraries in which those residues that line this pocket (Fig. 3) are randomized using doped codons, such as NNK. We are currently investigating both of these approaches in an attempt to create protein variants that can discriminate between different macrolide antibiotics and different concentration ranges. Such protein sensors will have applied uses while providing fundamental information on erythromycin sensing in the wild-type system. It is important to note that polyketide natural products such as erythromycin have recently been the subject of a large body of research aimed at taking advantage of modular biosynthetic pathways. While there are successful examples of combinatorial design of polyketide biosynthesis,[24–26] these approaches are often fraught with low productivity, presumably due to imperfect hybrid proteins that have not benefited from evolutionary pressure. Some methods have sought to combine directed evolution with these approaches to mimic the natural process through which secondary metabolic pathways are created and show promise.[27,28] Using genetically encoded biosensors, such as MphR(A), for the products of these systems might allow researchers to use directed evolution to improve upon their design and generate novel "unnatural" products.

## Experimental Procedures

### Protein expression and purification

The *mphR(A)* gene was amplified from pTZ3509[5] using primers 5′-AAA*CCATGG*GG**CATCATCATCATCATCAT**TTGGAAAACCTATACTTTCAAGGCCCCCGCCCCAAGCTCAAG-3′ (for CL699) and 5′-TTT*GAATTC*CGGCCGCTACTAGTTTACG-CATGTGCCTGGAGG-3′ (for CL343), which allowed incorporation of an N-terminal 6× histidine tag (bold), a tobacco etch virus (TEV) protease cleavage site (underlined), and NcoI-EcoRI sites (italics) for cloning into the same sites in pET28b (Novagen). The resulting plasmid was sequence verified and transformed into *E. coli* BL21 (DE3). The His-TEV-MphR(A) protein was overexpressed in LB medium containing 30 μg/mL of kanamycin. A 4-L culture was grown at 37 °C to an $OD_{600}$ (optical density at 600 nm) of 0.5 and induced with IPTG (1 mM) for 5 h at 30 °C. The cells were then harvested by centrifugation at 1100*g* for 15 min. The cell pellet was washed and resuspended with lysis buffer containing 50 mM Tris–HCl, 300 mM NaCl, and 10 mM imidazole, pH 8.0. The cell suspension was disrupted by Sonifier 450 and centrifuged at 27,000*g* for 45 min to remove cell debris. The cell-free extract was loaded onto HisLink™ protein purification resin that had been washed with lysis buffer. The column was washed with wash buffer containing 50 mM Tris–HCl, 300 mM NaCl, and 50 mM imidazole, pH 8.0. His-TEV-MphR(A) was then eluted with elution buffer containing 50 mM Tris–HCl, 300 mM NaCl, and 300 mM imidazole, pH 8.0, and dialyzed into TEV protease working buffer containing 50 mM Tris–HCl, 100 mM NaCl, 1 mM DTT, and 1 mM (ethylenediaminetetraacetic acid, pH 8.0. The protein was then digested with TEV protease for 12 h at room temperature. Undigested protein and the TEV protease were removed by HisLink™ protein purification resin as described above. The resulting MphR(A) was dialyzed to buffer containing 10 mM Tris–HCl, 50 mM NaCl, and 1 mM DTT, pH 8.0, and finally concentrated to 10 mg/mL by iCON™ Concentrators (7 mL/9K, Thermo Scientific). The resulting protein is the native sequence of the MphR(A) protein with the N-terminal methionine replaced by glycine.

SeMet-labeled MphR(A) was obtained by the pathway inhibition method. *E. coli* BL21(DE3) containing pET28b/His-TEV-MphR was grown to an $OD_{600}$ of 0.5 in M9 medium containing 0.4% glucose as carbon source at 37 °C. Then, amino acids lysine, phenylalanine, and threonine were added to a final concentration of 100 mg/L each; isoleucine, leucine, and valine were added to a final concentration of 50 mg/L each; and SeMet was added to a final concentration of 60 mg/L. The culture was incubated for another 15 min at 37 °C and then moved to 20 °C and induced by 1 mM IPTG for 12 h. The purification of the seleno-MphR(A) followed the protocol described above with the exception that all buffers were degassed and contained 1 mM DTT to prevent the oxidation of protein.

### Protein crystallization, data, and refinement

For crystallization, 10% glycerol was added to MphR(A) and seleno-MphR(A) stock solution. The protein concentration used for screening was 10 mg/mL. A total of 10 mM erythromycin was added for crystallization of the protein with ligand. An initial screen was performed

robotically through 480 unique conditions at three protein-to-well solution ratios using the sitting-drop vapor-diffusion method. Crystal optimization was performed by mixing 2-μL protein solutions with 1-μL precipitant solution in hanging drops. The crystals of unliganded MphR(A) on which data were collected were grown in precipitant solution containing 18–22% (w/v) polyethylene glycol 3350, 0.2 M ammonium acetate, and Bis-Tris, pH 5.5. The crystals of MphR(A) and seleno-MphR(A) bound to erythromycin were obtained with precipitant solution containing 35–40% (w/v) polyethylene glycol 3350, 0.2 M $MgCl_2$, and Bis-Tris, pH 5.5. Crystals were flash frozen in well solution containing 20% glycerol prior to data collection.

Data were processed using HKL2000,[29] and the MphR (A)–erythromycin structure was solved using the program HKL2MAP to identify the heavy atoms and calculate the phases from selenium SAD. The resulting map was put through heavy-atom position refinement, density modification, and model building using autoSHARP.[30] Additional model building was performed using Coot.[31] The refined liganded MphR(A) model was used to solve the structure for the unliganded MphR(A) in MolRep. Refmac[32] was used for refinement in both cases. TLS refinement was utilized by setting each monomer as a separate TLS group.[33] VOIDOO[34] was used for calculation of solvent-free cavity volumes using a van der Waals probe radius of 1.35 Å, a grid size of 1.00, 10 volume refinement cycles, and an equivalent starting point for all four cavities. Default values were used for all other parameters.

### Determination of MphR(A) DNA binding

For fluorescence polarization, a 35-bp double-stranded DNA containing a 5′-conjugated fluorescein was prepared by annealing complementary oligonucleotides (10 μM each) (forward: 5′-fluor-GATTGAATATAACCGACGTGACTGTTACATTTAGG-3′; reverse: 5′-CCTAAATGTAACAGTCACGTCGGTTATATTCAATC-3′) in 10 mM Tris – HCl, pH 8.0, 100 mM NaCl, and 1 mM ethylenediaminetetraacetic acid by heating the reaction to 85 °C in a water bath for 10 min and allowing it to cool to room temperature slowly. A 1-mL solution of 10 nM DNA in binding buffer [10 mM Mops, pH 7.0, 100 mM NaCl, and 1 μg/mL of poly(dI–dC)] was used for measurements. MphR(A) additions were made from a concentrated stock solution used for crystallography (420 μM MphR), and the components were mixed gently. All measurements were performed using an ISS PC1 spectrofluorimeter (ISS, Champaign, IL) equipped with a VWR Scientific 1160 constant temperature bath (PolyScience, Niles, IL) maintained at 21 °C with an excitation wavelength of 494 nm and a fluorescence signal measured at 518 nm. Each titration point was recorded 10 times, and the data were analyzed using the equation $P = \{(P_{bound} - P_{free})*[protein]/(K_d + [protein])\} + P_{free}$, 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 completely bound DNA, and [protein] is the concentration of MphR(A) as a dimer. Curve fitting was accomplished with the program Prism.

### *In vivo* erythromycin circuit analysis

The plasmid pBAD/mphR-P-lacZ was constructed by PCR assembly using four oligonucleotides: CL818, 5′-AACATCG*TCTAGA*TTGAATATAACCGACGTGACTGTTACATTTAGGTGGGCTAAC**AGGAGG**AATTAACCATGGATCCAC; CL819, 5′-CTAGTACA*TCTAGA*GGTAATACGGTTATCCACAGAATCAGGGGA; CL820, 5′-AAGATCG*TCTAGA*TGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCCCAACCCATAC**AGAAGG**TGAACACTGAT; and CL821, 5′-CTAGTACA*TCTAGA*GTGATCAAGACTTCGATACCACCGACCGTA. Promoters PmphR and PlacIQ are underlined, and XbaI restriction sites and the Shine–Dalgarno sequence are italicized and in boldface, respectively. The 5.6-kb fragment containing the PmphR promoter, pBR322 origin, and *lacZ* open reading frame was amplified by CL818 and CL819 using pBAD/Myc-His/lacZ (Invitrogen) as template, while the 0.75-kb fragment of *mphR* was amplified by CL820 and CL821 using pTZ3509 as template. Digestion with XbaI and ligation of the two fragments resulted in plasmid in which *mphR* was opposite lacZ. For agar plate diffusion assays, a 0.2-mL aliquot of an overnight culture of the sensor strain (Genehogs with pBAD/mphR-P-lacZ) in LB medium with 100 μg/mL of ampicillin was added to 15 mL of soft agar consisting of LB with 0.6% agarose and 100 μg/mL of ampicillin. After filter papers with erythromycin were applied, plates were incubated at 37 °C overnight.

### Mutation of MphR

Generation of the V66L and V126L mutants was accomplished using the GeneTailor Site-Directed Mutagenesis System (Invitrogen) following the manufacturer's instructions. The oligonucleotides used for mutagenesis were as follows: V66L, 5′-GAGCGCGGCGTCGAGCAGCTGCGGCATTACCTGA-3′ and 5′-CTGCTCGACGCCGCGCTCCATCATCCTCACCAG-3′, and V126L, 5′-ATCCAGCGGAACCGCGCGCTGGTGGAGGGGATCC-3′ and 5′-CGCGCGGTTCCGCTGGATCGCAAGCGTGCGTAG-3′. Both mutants were verified by DNA sequencing. To introduce random mutations into MphR, we performed error-prone PCR with 0.2 mM dATP and dGTP, 1 mM dCTP and dTTP, 7 mM $MgCl_2$, and 0.3 mM $MnCl_2$ using Taq DNA polymerase and the following oligonucleotides: CL865, 5′-CTACTAAA*GGTCT*CCCTAG*TGTACA*GTGATCAAGACTTCGATACCACCGACCGTA-3′ (BsaI and BsrGI sites are italicized), and CL820. The PCR products were first cloned into pBAD-mphR-P-lacZ by BsaI and XbaI to introduce BsrGI into the plasmid in which *mphR* was opposite lacZ. Then, BsrGI and XbaI cloning sites were used to clone error-prone PCR products into the plasmid to obtain a library of approximately $10^5$ clones. Ten mutants that were always blue in X-gal assay both with and without erythromycin were sequenced.

### Protein Data Bank accession numbers

The coordinates for these structures have been deposited to the Protein Data Bank with accession codes 3G56 [MphR(A)] and 3FRQ [MphR(A) complexed with erythromycin].

## Acknowledgements

Financial support was provided by the University of Maryland. 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. 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 Prof. Norihisa Noguchi (Tokyo University of Pharmacy and Life Science) for providing pTZ3509 and to Prof. Dorothy Beckett (University of Maryland) for allowing use of the fluorimeter.

## References

1. Weisblum, B. (1995). Erythromycin resistance by ribosome modification. *Antimicrob. Agents Chemother.* **39**, 577–585.
2. Ross, J. I., Eady, E. A., Cove, J. H., Cunliffe, W. J., Baumberg, S. & Wootton, J. C. (1990). Inducible erythromycin resistance in staphylococci is encoded by a member of the ATP-binding transport super-gene family. *Mol. Microbiol.* **4**, 1207–1214.
3. Leclercq, R. & Courvalin, P. (1991). Intrinsic and unusual resistance to macrolide, lincosamide, and streptogramin antibiotics in bacteria. *Antimicrob. Agents Chemother.* **35**, 1273–1276.
4. Noguchi, N., Emura, A., Matsuyama, H., O'Hara, K., Sasatsu, M. & Kono, M. (1995). Nucleotide sequence and characterization of erythromycin resistance determinant that encodes macrolide 2′-phosphotransferase I in *Escherichia coli*. *Antimicrob. Agents Chemother.* **39**, 2359–2363.
5. Noguchi, N., Takada, K., Katayama, J., Emura, A. & Sasatsu, M. (2000). Regulation of transcription of the *mph(A)* gene for macrolide 2′-phosphotransferase I in *Escherichia coli*: characterization of the regulatory gene *mphR(A)*. *J. Bacteriol.* **182**, 5052–5058.
6. Szczepanowski, R., Krahn, I., Bohn, N., Puhler, A. & Schluter, A. (2007). Novel macrolide resistance module carried by the IncP-1beta resistance plasmid pRSB111, isolated from a wastewater treatment plant. *Antimicrob. Agents Chemother.* **51**, 673–678.
7. Szczepanowski, R., Krahn, I., Linke, B., Goesmann, A., Puhler, A. & Schluter, A. (2004). Antibiotic multiresistance plasmid pRSB101 isolated from a wastewater treatment plant is related to plasmids residing in phytopathogenic bacteria and carries eight different resistance determinants including a multidrug transport system. *Microbiology*, **150**, 3613–3630.
8. Ramos, J. L., Martinez-Bueno, M., Molina-Henares, A. J., Teran, W., Watanabe, K., Zhang, X. *et al.* (2005). The TetR family of transcriptional repressors. *Microbiol. Mol. Biol. Rev.* **69**, 326–356.
9. Weber, W., Fux, C., Daoud-El Baba, M., Keller, B., Weber, C. C., Kramer, B. P. *et al.* (2002). Macrolide-based transgene control in mammalian cells and mice. *Nat. Biotechnol.* **20**, 901–907.
10. Kramer, B. P., Viretta, A. U., Daoud-El Baba, M., Aubel, D., Weber, W. & Fussenegger, M. (2004). An engineered epigenetic transgene switch in mammalian cells. *Nat. Biotechnol.* **22**, 867–870.
11. Link, N., Weber, W. & Fussenegger, M. (2007). A novel generic dipstick-based technology for rapid and precise detection of tetracycline, streptogramin and macrolide antibiotics in food samples. *J. Biotechnol.* **128**, 668–680.
12. Weber, C. C., Link, N., Fux, C., Zisch, A. H., Weber, W. & Fussenegger, M. (2005). Broad-spectrum protein biosensors for class-specific detection of antibiotics. *Biotechnol. Bioeng.* **89**, 9–17.
13. Mohrle, V., Stadler, M. & Eberz, G. (2007). Biosensor-guided screening for macrolides. *Anal. Bioanal. Chem.* **388**, 1117–1125.
14. Hinrichs, W., Kisker, C., Duvel, M., Muller, A., Tovar, K., Hillen, W. & Saenger, W. (1994). Structure of the Tet repressor–tetracycline complex and regulation of antibiotic resistance. *Science*, **264**, 418–420.
15. Dover, L. G., Corsino, P. E., Daniels, I. R., Cocklin, S. L., Tatituri, V., Besra, G. S. & Futterer, K. (2004). Crystal structure of the TetR/CamR family repressor *Mycobacterium tuberculosis* EthR implicated in ethionamide resistance. *J. Mol. Biol.* **340**, 1095–1105.
16. Schumacher, M. A., Miller, M. C., Grkovic, S., Brown, M. H., Skurray, R. A. & Brennan, R. G. (2001). Structural mechanisms of QacR induction and multidrug recognition. *Science*, **294**, 2158–2163.
17. Krissinel, E. & Henrick, K. (2004). Secondary-Structure Matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2256–2268.
18. Jeng, W. Y., Ko, T. P., Liu, C. I., Guo, R. T., Liu, C. L., Shr, H. L. & Wang, A. H. (2008). Crystal structure of IcaR, a repressor of the TetR family implicated in biofilm formation in *Staphylococcus epidermidis*. *Nucleic Acids Res.* **36**, 1567–1577.
19. Miller, J. H. (1972). *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
20. Willems, A. R., Tahlan, K., Taguchi, T., Zhang, K., Lee, Z. Z., Ichinose, K. *et al.* (2008). Crystal structures of the *Streptomyces coelicolor* TetR-like protein ActR alone and in complex with actinorhodin or the actinorhodin biosynthetic precursor (*S*)-DNPA. *J. Mol. Biol.* **376**, 1377–1387.
21. Anderson, J. C., Voigt, C. A. & Arkin, A. P. (2007). Environmental signal integration by a modular AND gate. *Mol. Syst. Biol.* **3**, 133.
22. Sayut, D. J., Kambam, P. K. & Sun, L. (2007). Engineering and applications of genetic circuits. *Mol. BioSyst.* **3**, 835–840.
23. Resch, M., Striegl, H., Henssler, E. M., Sevvana, M., Egerer-Sieber, C., Schiltz, E. *et al.* (2008). A protein functional leap: how a single mutation reverses the function of the transcription regulator TetR. *Nucleic Acids Res.* **36**, 4390–4401.
24. McDaniel, R., Thamchaipenet, A., Gustafsson, C., Fu, H., Betlach, M. & Ashley, G. (1999). Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. *Proc. Natl Acad. Sci. USA*, **96**, 1846–1851.
25. Yoon, Y. J., Beck, B. J., Kim, B. S., Kang, H. Y., Reynolds, K. A. & Sherman, D. H. (2002). Generation of multiple bioactive macrolides by hybrid modular polyketide synthases in *Streptomyces venezuelae*. *Chem. Biol.* **9**, 203–214.
26. Rowe, C. J., Bohm, I. U., Thomas, I. P., Wilkinson, B., Rudd, B. A., Foster, G. *et al.* (2001). Engineering a polyketide with a longer chain by insertion of an extra module into the erythromycin-producing polyketide synthase. *Chem. Biol.* **8**, 475–485.
27. Kim, B. S., Sherman, D. H. & Reynolds, K. A. (2004). An efficient method for creation and functional analysis of libraries of hybrid type I polyketide synthases. *Protein Eng. Des. Sel.* **17**, 277–284.
28. Zhang, Y. X., Perry, K., Vinci, V. A., Powell, K.,

Stemmer, W. P. & del Cardayre, S. B. (2002). Genome shuffling leads to rapid phenotypic improvement in bacteria. *Nature*, **415**, 644–646.
29. Otwinoswski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, *307–326.*
30. Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2007). Automated structure solution with autoSHARP. *Methods Mol. Biol.* **364**, 215–230.
31. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
32. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53**, 240–255.
33. Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **57**, 122–133.
34. Kleywegt, G. J. & Jones, T. A. (1994). Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **50**, 178–185.