



Crystal structure and functional analysis of mycobacterial erythromycin resistance methyltransferase Erm38 reveals its RNA-binding site

Received for publication, October 6, 2021, and in revised form, December 29, 2021 Published, Papers in Press, January 8, 2022,

<https://doi.org/10.1016/j.jbc.2022.101571>

Boon Chong Goh^{1,2,*} , Xinyu Xiang³, Julien Lescar^{1,2,3,*}, and Peter C. Dedon^{1,4,*}

From the ¹Antimicrobial Resistance Interdisciplinary Research Group, Singapore-MIT Alliance for Research and Technology Centre, Singapore, Singapore; ²NTU Institute of Structural Biology, Experimental Medicine Building (EMB), and ³School of Biological Sciences, Nanyang Technological University, Singapore, Singapore; ⁴Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Edited by Craig Cameron

Erythromycin resistance methyltransferases (Erms) confer resistance to macrolide, lincosamide, and streptogramin antibiotics in Gram-positive bacteria and mycobacteria. Although structural information for ErmAM, ErmC, and ErmE exists from Gram-positive bacteria, little is known about the Erms in mycobacteria, as there are limited biochemical data and no structures available. Here, we present crystal structures of Erm38 from *Mycobacterium smegmatis* in apoprotein and cofactor-bound forms. Based on structural analysis and mutagenesis, we identified several catalytically critical, positively charged residues at a putative RNA-binding site. We found that mutation of any of these sites is sufficient to abolish methylation activity, whereas the corresponding RNA-binding affinity of Erm38 remains unchanged. The methylation reaction thus appears to require a precise ensemble of amino acids to accurately position the RNA substrate, such that the target nucleotide can be methylated. In addition, we computationally constructed a model of Erm38 in complex with a 32-mer RNA substrate. This model shows the RNA substrate stably bound to Erm38 by a patch of positively charged residues. Furthermore, a π - π stacking interaction between a key aromatic residue of Erm38 and a target adenine of the RNA substrate forms a critical interaction needed for methylation. Taken together, these data provide valuable insights into Erm-RNA interactions, which will aid subsequent structure-based drug design efforts.

The emergence of antimicrobial resistance (AMR) in virtually every clinically important bacterial pathogen represents a true crisis with major societal and economic impact (1) (https://amr-review.org/sites/default/files/AMR%20Review%20Paper%20-%20Tackling%20a%20crisis%20for%20the%20health%20and%20wealth%20of%20nations_1.pdf). One major mechanism of AMR in Gram-positive and mycobacterial pathogens affects

three broad classes of front-line antibiotics — macrolides, lincosamides, and streptogramins (MLS) — and is conferred by a family of horizontally and vertically transmitted genes that encode the so-called erythromycin resistance methyltransferases (Erms) (2–6). Using SAM as a cofactor, Erms transfer a methyl group to adenosine at position 2058 (A2058) in the ~3000 nucleotide-long 23S ribosomal RNA. This subtle posttranscriptional modification blocks the antibiotic-binding site in the ribosomal peptide exit tunnel (7). Horizontally transferred Erms are emerging in methicillin-resistant *Staphylococcus aureus* and vancomycin-resistant enterococci, for example, whereas endogenous, inducible Erms obviate the use of MLS antibiotics for *Mycobacterium tuberculosis* and limit their utility for *Mycobacterium abscessus* (5, 6, 8–10). Even at the level of pediatric acute otitis media caused by *S. aureus*, nearly 90% of clinical isolates in one study possessed Erms A, B, and C (11), which greatly limits available antibiotics in patients with β -lactam allergies.

Given the importance of Erms in MLS resistance, it is not surprising that there is significant genetic information about the diversity of types and distributions of Erms. The amino acid sequences and lengths can be broadly distinguished in mycobacteria and Gram-positive bacteria (12). Mycobacterial *erm* genes, such as *erm37* in *M. tuberculosis*, *erm38* in *Mycobacterium smegmatis*, and *erm41* in *M. abscessus*, as well as *ermE* of *Saccharopolyspora erythraea*, are genomically encoded, whereas the *erm* genes of Gram-positive bacteria, including *ermA* and *ermC* of methicillin-resistant *S. aureus* and *ermB* of vancomycin-resistant enterococci, are found on highly mobile plasmids involved in horizontal gene transfer. In terms of size, Erms A, B, and C share similar molecular weights (~30 kDa), whereas Erm37 and Erm41 are smaller (~20 kDa) because of the lack of a C-terminal domain (6, 10). These trends are broken by Erm38 and ErmE (~45 kDa), which possess a longer C-terminal region that is predicted to be disordered (10, 13).

Despite the role of Erms in AMR, there is little structural information about the most clinically important Erms. Several Erm structures are available for Gram-positive bacteria, namely ErmAM (PDB access code: 1YUB) (14), ErmC (PDB access code:

* For correspondence: Peter C. Dedon, pcdedon@mit.edu; Julien Lescar, julien@ntu.edu.sg; Boon Chong Goh, boonchong@smart.mit.edu. Present address for Xinyu Xiang: Department of Structural Biology, Stanford University School of Medicine, Stanford, California 94305, USA.

Mycobacterial Erm38 and its RNA-binding site

1QAO) (15, 16), and ErmE (PDB access code: 6NVM) (13). However, there is no structural information for any mycobacterial Erm proteins, even though Erms confer intrinsic and inducible MLS resistance in mycobacteria (5), and Erm41 acquisition is limiting the efficacy of clarithromycin for *M. abscessus* infections (6, 17). Here, we determined the crystal structure of Erm38 of *M. smegmatis* and analyzed its biochemical and biophysical properties. Mutagenesis data and computational tools enabled us to build an atomic model for the Erm38-RNA complex, which will aid future drug discovery efforts.

Results

Overall structure of Erm38

Erm38 comprises 386 amino acids (NCBI WP_063844518; Uniport Q79N53), with analysis using the Database of Disordered Protein Predictions (18) predicting that the N-terminal 12 residues and C-terminal 124 residues are disordered

regions. Expressing full-length Erm38 in *Escherichia coli* resulted in a heterogeneous population of partially degraded proteins with molecular weights ranging from 30 to 45 kDa, in spite of length variations (Fig. S1). The computational prediction of a long, unstructured C terminus in Erm38 may account for this translational difficulty. We therefore expressed a truncated version of Erm38 spanning residues 13 to 261. This truncated Erm38, referred to as Erm38 throughout, has a molecular mass of 30 kDa and a size similar to most Erms from Gram-positive bacteria (Figs. S2–S4). The highly purified Erm38 formed square plate crystals (Fig. S5) that diffracted to a maximum resolution of 1.9 Å and tolerated up to 20% dimethylsulfoxide, which makes them suitable for ligand-soaking experiments. Hence, cocrystals with either SAM or the isosteric sinefungin inhibitor were obtained by soaking into native Erm38 crystals, and diffraction data to 2.25 Å resolution were collected as summarized in Table 1.

Table 1
Crystallographic data collection and refinement statistics

Parameter	Erm38 unliganded	Erm38 + SAM	Erm38 + SFG
Wavelength (Å)	0.98	0.98	0.98
Resolution ^a range	21.54–1.9 (1.968–1.9)	19.94–2.25 (2.33–2.25)	19.91–2.25 (2.33–2.25)
Space group	P 4 ₂ 2 ₁ 2	P 4 ₂ 2 ₁ 2	P 4 ₂ 2 ₁ 2
Unit cell (Å, °)	77.663 77.663 101.116 90 90 90	78.02 78.02 101.54 90 90 90	78.018 78.018 101.536 90 90 90
Total reflections	657923 (66731)	400975 (34914)	400950 (34914)
Unique reflections	25025 (2449)	15428 (1483)	15427 (1483)
Multiplicity	26.3 (27.2)	26.0 (23.5)	26.0 (23.5)
Completeness (%)	99.88 (100.00)	99.64 (98.87)	99.64 (98.80)
Mean I/sigma(I)	30.08 (2.63)	34.02 (4.31)	34.01 (4.31)
Wilson B-factor (Å ²)	37.56	40.72	42.93
R-merge ^b	0.06656 (1.321)	0.07489 (0.7235)	0.07489 (0.7235)
R-meas ^c	0.06792 (1.346)	0.07639 (0.7395)	0.07639 (0.7395)
R-pim ^d	0.01331 (0.2561)	0.01491 (0.1506)	0.01491 (0.1506)
CC1/2	1 (0.838)	1 (0.926)	1 (0.926)
CC	1 (0.955)	1 (0.981)	1 (0.981)
Reflections used in refinement	25022 (2449)	15422 (1483)	15423 (1482)
Reflections used for R-free	1279 (129)	758 (72)	758 (72)
R-work ^e	0.1921 (0.2501)	0.1882 (0.2390)	0.1942 (0.2398)
R-free ^f	0.2184 (0.2950)	0.2262 (0.3226)	0.2267 (0.3071)
CC(work) ^g	0.958 (0.842)	0.961 (0.874)	0.956 (0.829)
CC(free)	0.962 (0.723)	0.927 (0.764)	0.932 (0.850)
Number of nonhydrogen atoms	2078	2092	2047
Macromolecules	1906	1924	1885
Ligands		16	35
Solvent	172	152	127
Protein residues	245	246	244
RMS (bonds) (Å)	0.013	0.014	0.013
RMS (angles) (°)	1.55	1.75	1.63
Ramachandran favored (%)	99.16	98.35	98.32
Ramachandran allowed (%)	0.84	1.24	1.26
Ramachandran outliers (%)	0.00	0.41	0.42
Rotamer outliers (%)	3.12	4.21	4.76
Clashscore	0.52	2.05	1.81
Average B-factor (Å ²)	43.44	45.91	47.23
Macromolecule	42.50	45.24	46.49
Ligand		63.04	65.08
Solvent	53.87	52.66	53.23

^a The values presented in parentheses are for the highest resolution shell.

^b $R_{merge} = \sum |I_{obs} - I_{avg}| / \sum I_{avg}$.

^c $R_{meas} = \sum hkl \sqrt{\frac{n}{n-1} \sum_{i=1}^n |I_i(hkl) - \bar{I}(hkl)|} / \sum_{hkl} \sum_{i=1}^n I_i(hkl)$, where R_{meas} is the precision indicator of individual observation in unmerged data.

^d $R_{pim} = \sum hkl \sqrt{\frac{1}{n} \sum_{i=1}^n |I_i(hkl) - \bar{I}(hkl)|} / \sum_{hkl} \sum_{i=1}^n I_i(hkl)$, where R_{pim} is the precision indicator of the merged data.

^e $R_{work} = \sum |F_{obs} - F_{calc}| / \sum F_{obs}$.

^f $R_{free} = \sum |F_{obs} - F_{calc}| / \sum I_{calc}$, where R_{free} was randomly sampled from 5% reflections.

^g Correlation Coefficient, $CC = \sum (x - \langle x \rangle)(y - \langle y \rangle) / [\sum (x - \langle x \rangle)^2 \sum (y - \langle y \rangle)^2]^{1/2}$, where $CC_{(1/2)}$ is of two half datasets.

Erm38 features at its N-terminal catalytic domain the classic Rossmann fold, an α/β sandwich that contains the SAM cofactor-binding site and a helical C-terminal domain (Fig. 1A). Generally, Erm38 shares highly a similar structural fold with all published Erm structures and KsgA, a 16S rRNA methyltransferase

(Fig. S6). A comparison between Erm38 and ErmC (Fig. 1B) reveals that both structures are highly similar with RMSD of 2.3 Å for 237 superimposed alpha carbon atoms, despite sharing only 22% sequence identity. Compared to ErmC, the N-terminal region of Erm38 has a longer $\alpha 1$ helix and longer loops connecting

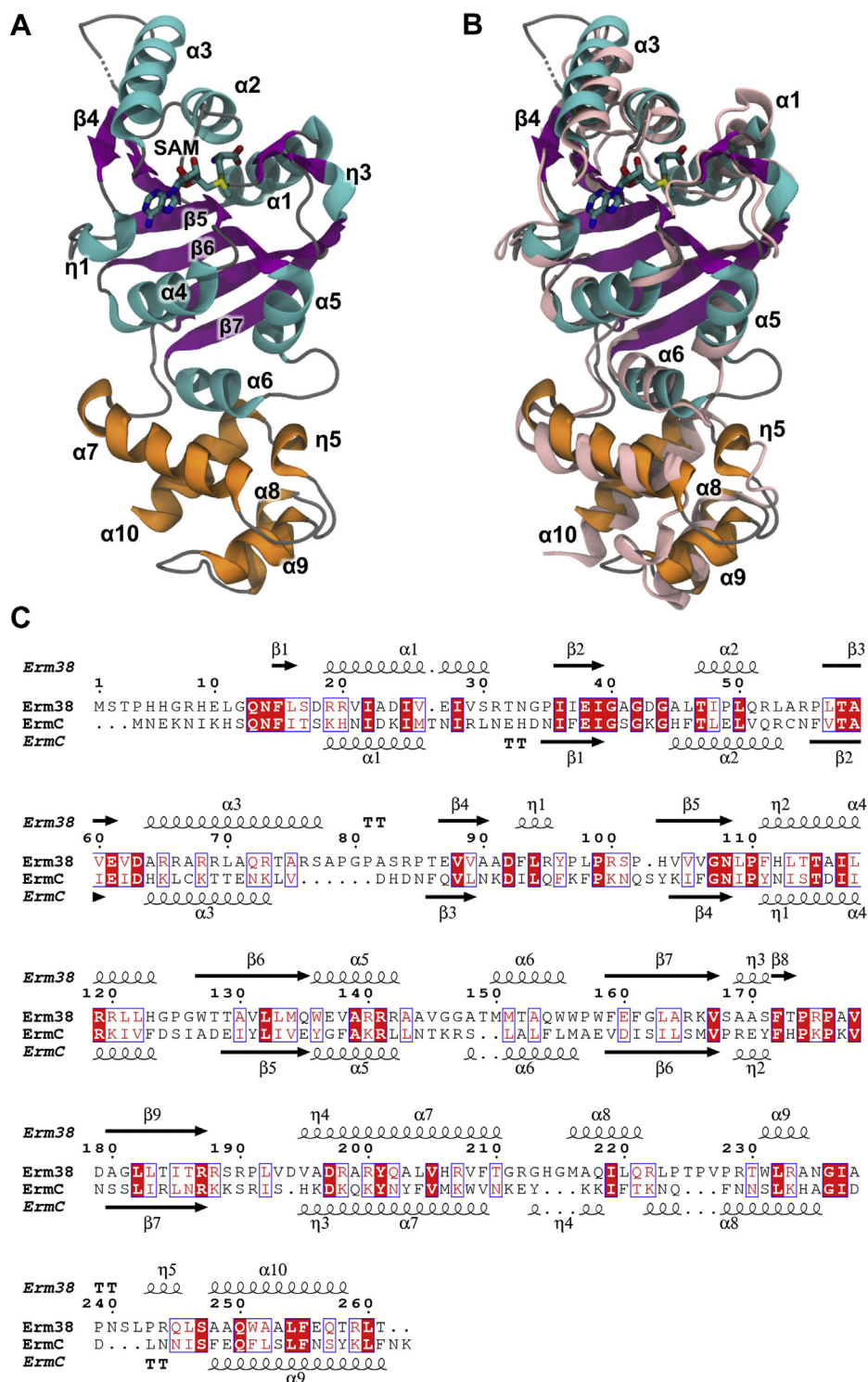


Figure 1. Comparison between Erm38 and ErmC. A, crystal structure of Erm38 liganded with SAM. The β -sheets are colored in purple, helices at the N- and C-terminal are colored in cyan and orange, respectively. B, structural overlay of Erm38 and ErmC shown in light pink ribbon. C, sequence alignment of Erm38 and ErmC. Conserved residues are highlighted with white text on red background, and similar residues are shown in red text. The alignment was carried out with Clustal Omega (48) and visualized with ESPript 3.0 (49). Erm, Erythromycin resistance methyltransferase.

Mycobacterial Erm38 and its RNA-binding site

$\alpha 3$ – $\beta 4$ and $\alpha 5$ – $\alpha 6$. In the C-terminal helical domain, Erm38 possesses a longer $\alpha 8$ helix but shorter $\alpha 9$ and $\alpha 10$ helices, along with an extra 3_{10} helix-labeled $\eta 5$.

Cofactor recognition at the SAM-binding pocket

When analyzed using LigPlot (19), the binding of SAM and sinefungin inhibitor had almost identical interaction networks, with sinefungin forming an additional hydrogen bond with R66 (Fig. 2B). This was expected given the high chemical homology between SAM and sinefungin. The SAM-binding pocket of Erm38 was also compared with ErmC (Fig. 2, A and C). Despite sharing only 22% sequence identity, residues at the SAM-binding pocket are highly conserved between Erm38 and ErmC proteins with minor differences found at V62/I60, A41/S39, and F93/I85, whereas residues F15, G40, G42, E61, D92, N108, and P110 are identical for the two enzymes. Particularly, G40, G42, E61, D92, and P110 of Erm38 are strictly conserved among Erms A, B, C, E, 37, 38, and 41 (Fig. S7), suggesting an important functional role conserved during evolution.

Erm38 methylation activity

Erm38 in *M. smegmatis* confers resistance to lincosamides and macrolides, but not streptogramin B antibiotics, which suggested that it was not a dimethyltransferase that produces double methylation at position N6 on A2058 (m66A) to cause

full MLS (20). However, mass spectrometry-based methylation analysis revealed that Erm38 is indeed a dimethyltransferase, though with apparent low efficiency that results in mainly unmethylated and monomethylated sites (20). We thus compared Erm38 enzyme activity to Erms A, B, and C with the 32-mer RNA substrate and SAH production as the end point. As shown in Table 2 and Fig. S8, the Michaelis–Menten kinetic constants for Erm38 are very similar to those for the other Erms, which suggests that the “reluctance” observed by Madsen *et al.* involves factors other than the innate methylation activity, such as levels of enzyme induction following antibiotic exposure, availability of SAM, or accessibility of A2508 in the 23S rRNA.

Putative RNA-binding site

The electrostatic potential of Erm38 includes a highly positively charged region that comprises residues R119, R140, R141, R142, and K166 located at the protein surface (Fig. 3, A and B). As these solvent-exposed basic residues are likely to be involved in binding the RNA substrate, we produced an alanine mutant for each site and compared the enzymatic activities of the mutant and WT enzymes. R31A was selected as a negative control based on the fact that it is distant from the SAM- and RNA-binding sites, and we concluded that it does not participate in any intramolecular interaction needed to maintain the fold of the protein. E61K is known to reduce

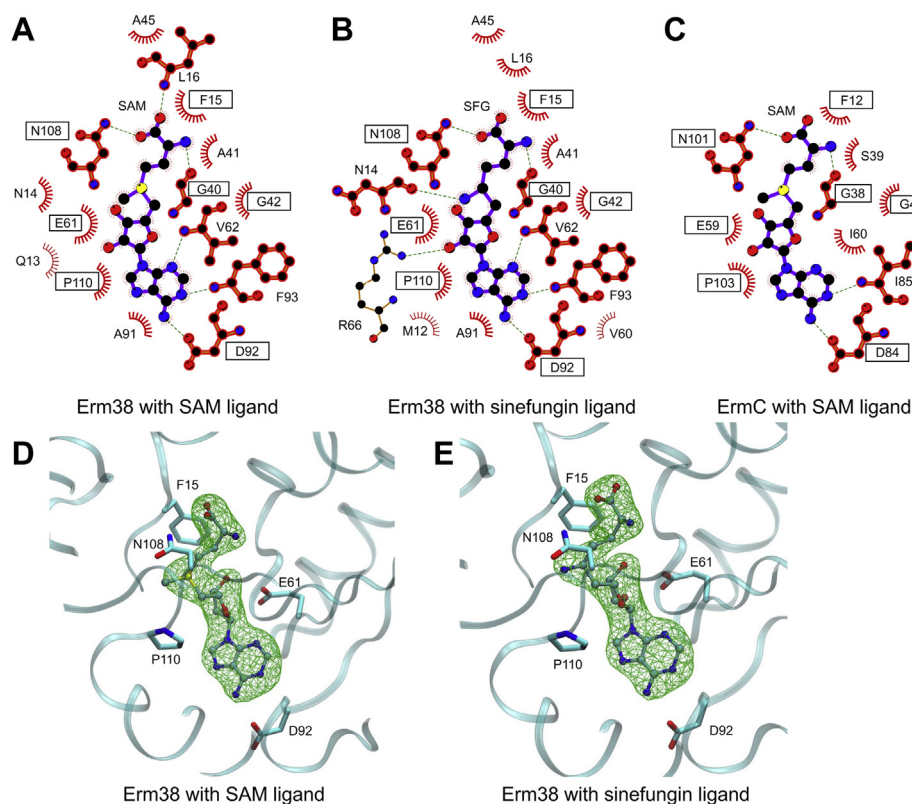


Figure 2. 2D and 3D representation of the protein-ligand interactions at the SAM-binding pocket. A–C, the 2D diagram was produced using Ligplot-v2.2 (19). The broken lines indicate hydrogen bonds and salt-bridges. The residues involved in van der Waals interactions with the ligand are also shown. The red crescents with the bristles represent the hydrophobic interactions. Equivalent side chains are shown in bold, and strictly conserved residues are highlighted in black boxes. D and E, unbiased omit Fo-Fc electron density maps of the ligands contoured at 4 σ . Conserved residues F15, E61, D92, N108, and P110 were shown to facilitate cross-referencing with the 2D interaction diagram. Erm, Erythromycin resistance methyltransferase.

Table 2
Michaelis–Menten kinetic parameters of various Erm enzymes

Enzymes	K _m (μM)	V _{max} (nM/min)
ErmA	5.9 ± 0.8	84 ± 6
ErmB	5.5 ± 0.6	93 ± 4
ErmC	5.1 ± 1.0	90 ± 9
Erm38	6.8 ± 0.6	124 ± 6

methylation activity by disrupting the SAM-binding site (21), so we selected this mutant as a control for loss of enzyme activity. As shown in Figure 3C, the E61K, R119A, R140A, and R141A mutations resulted in >90% reduction in methylation activity, whereas the R142A mutation resulted in 75% loss of enzymatic activity. Conversely, R31A and K166A did not affect the enzyme activity. Owing to their location at the surface of the protein and the absence of contact with SAM, the complete loss in activity observed upon loss of R119, R140, R141, and R142 is likely due to a disruption of electrostatic interactions with the RNA substrate.

We next sought to quantify the effect of the mutations on the interaction between Erm38 and its RNA substrate, using

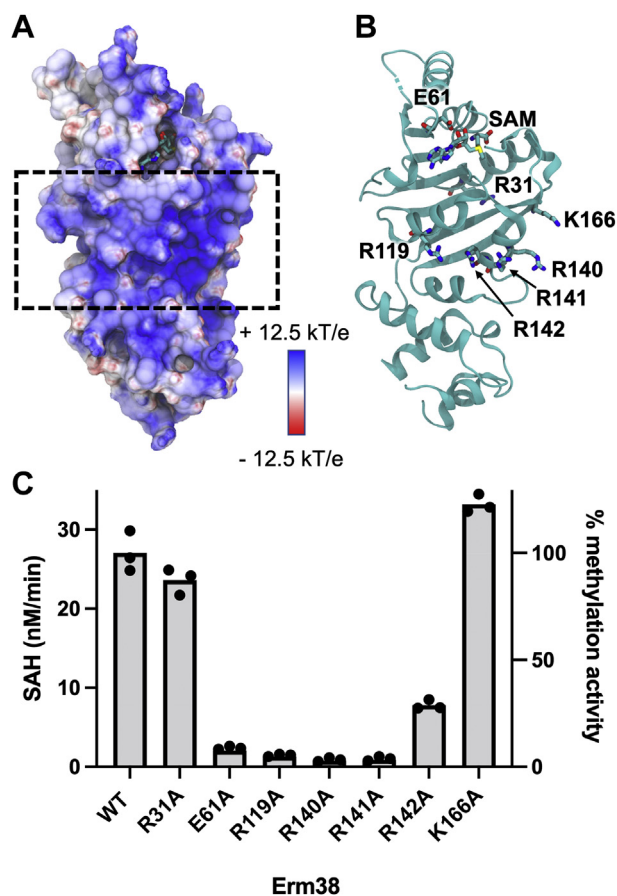


Figure 3. Identification and validation of the residues that form the putative RNA-binding site. A, the putative RNA-binding site was identified by the highly positively charged region shown in the electrostatic potential of Erm38, which was calculated using APBS (50). B, positively charged residues in the putative RNA-binding site are identified and subjected to mutagenesis study. C, the methylation activity of Erm38 with various single mutations. The methylation assay was carried out in technical triplicates. Erm, Erythromycin resistance methyltransferase.

biolayer interferometry to measure the kinetics of protein binding to a surface-bound biotinylated 32-mer RNA substrate (Fig. 4A). Surprisingly, the RNA-binding profiles for WT and mutants were similar, with an interaction model was best represented by 2:1 heterogeneous binding and two dissociation constants (K_D) (Table 3). The micromolar K_{D2} values likely represent a small amount of nonspecific binding, whereas the higher affinity K_{D1} of around 50 nM represents the binding affinity of Erm38 for its RNA substrate.

Structural model for the Erm38–RNA complex

Guided by the mutagenesis data and the knowledge that Erm38 binds strongly to the 32-mer RNA substrate, we created an atomic model of the Erm38–RNA complex. We first performed molecular docking of the RNA substrate on Erm38. Of 100 docked structures, one was most consistent with the electrostatic potential analysis and mutagenesis data, with adenine at position 13 (equivalent to A2058 in *E. coli* 23S rRNA) lying in the vicinity of F111 of Erm38, an aromatic residue known to interact with the adenine to facilitate methylation (21).

The docked structure was then subjected to molecular dynamics (MD) simulation to further refine the binding interface. Figure 5 shows a significant rearrangement of the RNA substrate in the first 10 ns of the simulation. At 3 ns, the stem loop of the RNA substrate relaxes and extends to form tighter interactions with the C-terminal domain of Erm38, whereas the A13 methylation site extends away from the catalytic site. From 5 ns onwards, the A13 reinserts into the catalytic pocket by the formation of a stable π - π stacking interaction with the phenyl ring of F111. After the initial refinement of the binding interface, the RMSD value of the RNA substrate remains at ~10 Å, indicating that the system has reached an equilibrium.

By analyzing the root-mean-square fluctuation of the RNA substrate (Fig. 6), we found that the portion of the RNA substrate that interacts with the patch of positively charged residues is most stable, validating the importance of these surface-exposed residues for RNA binding. In contrast, the A13-containing stem loop and the substrate termini remain relatively flexible. From the pool of 100 docked structures, we also simulated another docked Erm38–RNA model that is closest to the relative orientation of protein–RNA observed in KsgA–RNA model (PDB accession number 3FTF). The Erm38–RNA complex is stable in the 10-ns simulation, whereas this alternate model lacks the π - π stacking interaction and it does not interact with R140 of Erm38, which was shown to be important according to our mutagenesis data (Fig. S9). Therefore, the structural model of Erm38 complexed with the RNA substrate shown in Figures 5 and 6 provides the most accurate view of how Erm binds to its RNA substrate.

Discussion

Despite sharing low amino-acid sequence identity with the other Erms whose structures were reported previously (13–16), Erm38 shares high structural similarity. This conclusion is validated by DALI distance-matrix structural

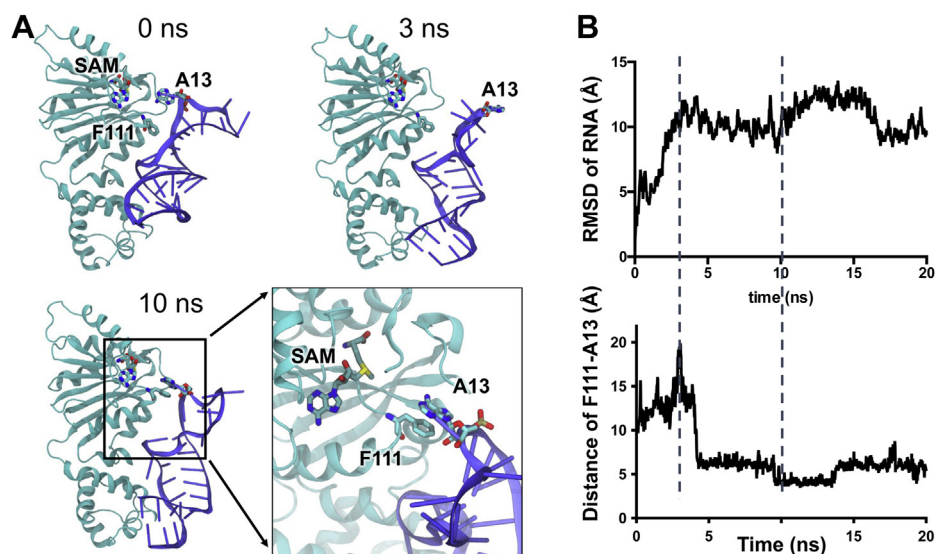


Figure 5. Molecular dynamics simulation of the Erm38+RNA complex. *A*, the protein–RNA interaction was refined where the RNA substrate undergoes a conformational change to bind to the C-terminal of Erm38, and a π - π stacking interaction is formed between F111 of Erm38 and A13 of RNA. Note that A13 is equivalent to A2058 in *E. coli* 23S ribosomal RNA numbering. *B*, trajectories of the RMSD values of the RNA substrate and the distance between F111 and A13 over the 20 ns of simulation. Erm, Erythromycin resistance methyltransferase.

apoprotein. Such precipitation was also observed when using full length ErmB and ErmC (data not shown). This rules out X-ray crystallography, cryo-EM, and solution-state NMR techniques for structure determination. Solid-state NMR offers a solution to analyze the precipitate, whereas the 40 kDa size of the Erm38–RNA complex poses a problem for NMR resolution.

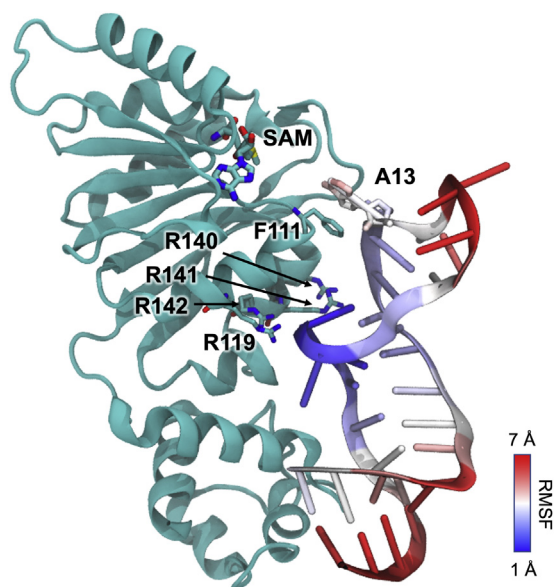


Figure 6. The four arginine residues bind stably to the RNA substrate. The degree of fluctuation of the RNA substrate was measured in the course of simulations by root-mean-square fluctuation (RMSF). The residues in red map to regions with relatively large fluctuation (~ 7 Å), whereas those colored blue map to regions with small fluctuation (~ 1 Å). The dark blue region at the middle segment of the RNA substrate highlights the strong RNA binding of the R119, R140, R141, and E142 of Erm38. Erm, Erythromycin resistance methyltransferase.

Here, we took the alternative approach of integrating biochemical data, unbiased molecular docking, and MD simulations to provide a plausible atomic model of the Erm38–RNA complex (Figs. 6 and 7). The resulting model suggests that, after RNA binding to the patch of positive-charged residues (Fig. 7), A13 of the 32-mer RNA undergoes π - π stacking with F111, an amino acid known to be essential for Erm38 catalytic activity (21). This π - π interaction that was organically formed in the simulation appears to be a key event that facilitates the methylation reaction. The Erm38–RNA model proposed here not only substantiates the important role of the positively charged residues at the N-terminal domain of Erm38 in RNA binding, but it also shows a binding contribution for the C-terminal region of Erm38. Specifically, highly conserved residues R207, R212, and R222 were found to form transient hydrogen bonds with the RNA substrate. That the C-terminal region of Erm plays a role in RNA binding and determines its methylation specificity is supported by the observations of Madsen *et al.* (10) with Erm37 of *M. tuberculosis*, which lacks a C-terminal domain and promiscuously methylates neighboring adenosines in the ribosomal RNA. The RNA-binding interactions of the conserved C-terminal residues observed in the simulation support the critical role of the C-terminal domain in determining the methylation activity of Erm proteins.

In conclusion, the crystal structure of Erm38 of *M. smegmatis* and the atomic model of Erm38 in complex with a 32-mer RNA substrate presented here can be used for structure-based drug design to target the putative RNA-binding site. Nonconventional design approaches such as creating compounds that mimic the footprint of the RNA substrate can now be tested. Furthermore, the high tolerance for dimethylsulfoxide of the Erm38 crystals obtained in this work enables hit identification and hit-to-lead optimization in following drug screening campaigns.

Mycobacterial Erm38 and its RNA-binding site

Table 4
Conserved spatial locations of conserved amino acids in Erms

Location in Erm	Erm38		ErmAM			ErmC			ErmE		
	AA	#	AA	#	Δ , Å ^a	AA	#	Δ , Å	AA	#	Δ , Å
SAM binding	GLY	40	GLY	37	3.9	GLY	38	0.8	GLY	69	1.5
	GLY	42	GLY	39	2.5	GLY	40	1.2	GLY	71	2.1
	GLU	61	GLU	58	1.3	GLU	59	1.6	GLU	90	1.1
	ASP	92	ASP	83	3.2	ASP	84	1.7	ASP	115	0.5
	ASN	108	ASN	100	2.9	ASN	101	0.9	ALA	131	0.7
	PRO	110	PRO	102	1.4	PRO	103	1.5	PRO	133	1.0
Catalytic RNA binding	PHE	111	TYR	103	1.6	TYR	104	1.8	TYR	134	1.2
	ARG	119	LYS	111	3.0	ARG	112	2.2	ASP	142	1.4
	ARG	140	LYS	132	3.5	LYS	133	1.3	ARG	163	0.4
	ARG	141	ARG	133	3.6	ARG	134	0.9	LYS	164	0.4
	ARG	142	THR	134	2.5	LEU	135	1.1	ARG	165	0.8

^a Δ , Å: C _{α} -C _{α} distance in Angstroms.

Experimental procedures

Expression and purification of Erm38 and its mutants

The *erm38* gene was synthesized and cloned into pNIC28-Bsa4 vector by Bio Basic Inc. Nucleotide sequences were codon-optimized to improve the efficiency of soluble expression in *E. coli*. The plasmid-containing truncated Erm38 (residues 13–261) was produced by the NTU Protein Production Platform (Singapore). The truncated Erm38 is referred to throughout as Erm38. Site-directed mutagenesis was performed on Erm38 to introduce the single mutations R31A, R119A, R140A, R141A, R142A, and K166A, using primers listed in Table S1. The mutation sites were confirmed by Sanger sequencing. After transformation into *E. coli* BL21(DE3)-T1R Rosetta strain, Erm38 proteins were expressed by growing in LB broth containing 34 μ g/ml chloramphenicol and 50 μ g/ml kanamycin at 37 °C, with expression induced with 0.2 mM IPTG for 18 h at 16 °C. Cell pellets were lysed by sonication in 50 mM Na HEPES pH7.5, 500 mM NaCl, 10% glycerol, 0.5 mM DTT, and 10 mM Imidazole. After removal of cell debris by centrifugation at 16,000g, the supernatant was sterile filtered and subjected to a series of purification steps. For WT Erm38, high-purity protein was critical for crystallization studies, which necessitated a three-step purification involving immobilized metal affinity chromatography (IMAC), cation-exchange chromatography, and size-exclusion chromatography (see Supplemental Methods). For the Erm38 mutants, a single-step IMAC purification was performed to obtain catalytically active proteins. The IMAC eluates were subjected to buffer exchange by diafiltration to reduce the imidazole concentration to <10 mM. All purified proteins were concentrated to >5 mg/ml and stored at –80 °C in 20 mM HEPES pH 7.5, 300 mM NaCl, 10% glycerol, and 2 mM DTT.

Crystallization of Erm38

The purified WT Erm38 at 15 mg/ml was subjected to initial crystallization screening. The screen was set up using the sitting drop vapor diffusion method with Morpheus and JCSG-plus crystallization kits (Molecular Dimensions) on Intelli-plates 96-3 (Art Robbins Instruments) (26–28). Using a Mosquito HTS (TTP Labtech), crystallization trials were set with protein-to-reservoir volume ratios of 2:1, 1:1, and 1:2.

Plates were sealed and kept at 20 °C in FORMULATRIX for monitoring the crystallization process. Small square crystals grew in 0.8 M sodium succinate pH 7.0. Upon further optimization, large square plate crystals grew in 1.0 M succinic acid and 5% glycerol within 5 days. To obtain the structure of Erm38 with SAM and sinefungin ligands, a few well-formed crystals were selected and soaked for 24 h with 2 mM SAM or 2 mM sinefungin. Before X-ray diffraction data collection, the crystals were protected from freezing damage using the corresponding crystallization buffer supplemented with 20% (v/v) glycerol before being flash-frozen in liquid nitrogen.

X-ray diffraction data were collected at 0.98 Å wavelength at the beamline PROXIMA 2A at the synchrotron SOLEIL (29, 30). The datasets were processed with the program XDS (31) and the CCP4 program (32). The structure was solved by molecular replacement with BALBES (33) using PDB ID 3FTF and 3FUV as search models. Several rounds of manual model building were then performed using program COOT (34), interspersed with structure refinement with programs REFMAC (35) and BUSTER (36). Some extra peaks in the Fo-Fc residual map at the SAM-binding site were not accounted for and left unmodeled in the liganded Erm38 models, as we only modeled the most occupied conformations of the SAM and sinefungin ligands. These unmodeled electron density suggested the possibility of an alternate conformation of the carboxylic tail. All figures representing structures were made using visual molecular dynamics (VMD) (37). Data collection and refinement statistics can be found in Table 1.

Erm enzyme assay

All assays were performed in 10 μ l total volume in a 384-well plate (Grenier, Item No. 784904; white color). Erm reaction kinetics were quantified using the MTase-Glo methyltransferase assay kit (Promega), which detects the SAH product of SAM demethylation in a coupled luminescence reaction (38). The RNA substrate for the reaction was a 32-mer oligoribonucleotide, CGCGACGGACGGAAA-GACCCC UAUCCGUCGCG, which has been shown to be a universal Erm substrate (39, 40) that was designed to mimic the adenine loop in domain V of 23S rRNA with its methylation site confirmed by mass spectrometry (41). The RNA substrate was prepared by denaturing at 90 °C for 1 min and reannealing by cooling slowly to ambient temperature. Assays

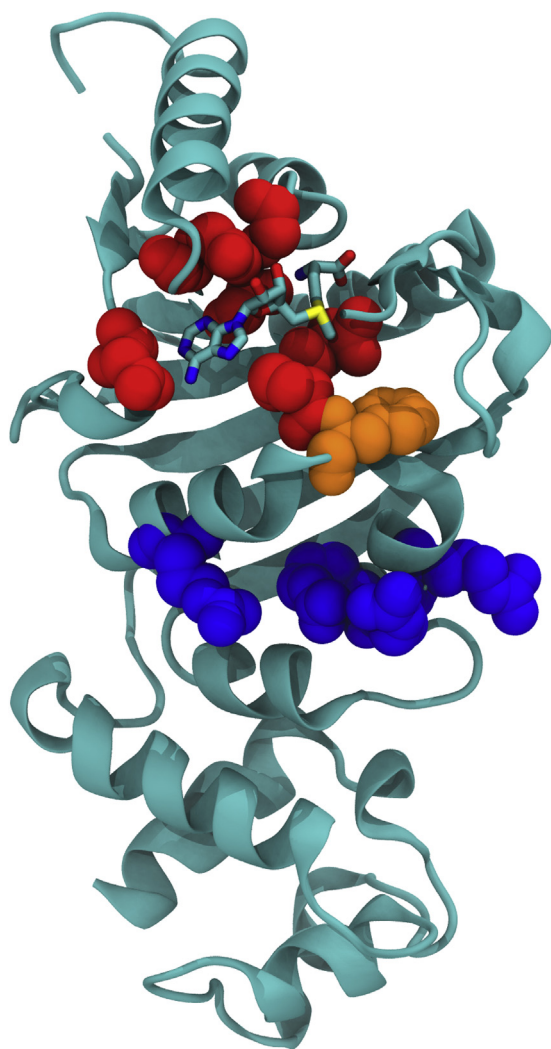


Figure 7. Erm38 structure depicting the locations of key spatially conserved amino acids at the SAM-binding site (red), catalytic site (orange), and RNA-binding site (blue). Table 4 shows the proximity of these conserved amino acids in Erms AM, C, and E. Erm, Erythromycin resistance methyltransferase.

for evaluating the activity of various Erm38 mutants contained 2 μM Erm protein, 2 μM RNA substrate, and 20 μM SAM in a buffer comprised of 50 mM Na HEPES pH 7.5, 40 mM KCl, 1 mM MgCl_2 , and 1 mM DTT (25, 38, 39, 42). To determine the Michaelis–Menten kinetic constants, the RNA substrate was varied from 0.25 μM to 8 μM . The reactions were carried out by mixing RNA and SAM and adding Erm38 and MTase-Glo reagent. After a 30 min incubation at 37 $^\circ\text{C}$, 5 μl of MTase-Glo detection solution was added to initiate the luminescence reaction. The luminescent signal was measured for 30 min using a microplate reader (BioTek Synergy 4 Plate Reader), with the resulting luminescent signal converted to SAH concentration using a SAH standard curve. A reaction without Erm protein was included as a control. The Michaelis–Menten kinetic parameters (maximum rate, V_{max} ; Michaelis constant, K_m) were determined by plotting the SAH concentration as a function of RNA concentration tested and the data were fitted by nonlinear regression with GraphPad Prism version 9 (GraphPad Software, Inc).

Biolayer interferometry to quantify protein–RNA interaction

The binding affinities of WT, R119A, and R130A Erm38 proteins with the 32-mer RNA substrate were measured by biolayer interferometry using an Octet RED96e (ForteBio), with binding data acquired (kinetics mode) and analyzed using built in software. Biosensors were hydrated in phosphate-buffered saline with Tween (PBST) buffer (137 mM sodium chloride, 2.7 mM potassium chloride, 12 mM phosphate, pH 7.4, 0.005% Tween 20, and 0.5 mg/ml bovine serum albumin) for 20 min at ambient temperature preceding data acquisition, and experiments were performed at 25 $^\circ\text{C}$. Erm38 proteins were diluted in PBST buffer. A biotinylated 32-mer RNA substrate was synthesized (Sigma-Aldrich) with biotin appended at the 3'-end of the oligo. The biotinylated RNA was diluted to 100 nM and immobilized on a streptavidin-coated biosensor with a signal threshold of 0.2 nm. A baseline level was established for 60 s before Erm38 proteins were exposed to RNA-loaded biosensors for 120 s and dissociated in PBST buffer for 120 s. To determine K_d values for each Erm38 protein, a reference sensor with loaded RNA but no protein was subtracted from the data before fitting. Curve fitting using a 2:1 interaction model was used to measure the dissociation constant (K_D) for WT and mutants of Erm38.

Molecular docking and MD simulations

Secondary structure of the 32mer RNA substrate was predicted using RNAfold (43), and the 3D model of the 32-mer was constructed computationally using 3dRNA (44). Of the five models predicted by 3dRNA, the 3D model selected for docking has its A13 (equivalent to A2058 in *E. coli* numbering) pointing outwards, which makes this RNA model primed to interact with the methylation site of Erm. Docking of Erm38 and 32-mer RNA was performed using HDOCK (45) without any prior knowledge of the putative-binding site. Of 100 top docked structures, the Erm38 + 32-mer RNA model that best fits the mutagenesis data was selected.

The selected Erm38 + 32mer RNA complex was subjected to all-atom, explicit solvent MD simulation using NAMD (46). The complex was simulated in a water box, where the minimal distance between the solute and the box boundary was 15 Å along all three axes. The charges of the solvated system were neutralized with counter-ions, and the ionic strength of the solvent was set to 150 mM NaCl using VMD (37). The fully solvated system was subjected to conjugate-gradient minimization for 10,000 steps, subsequently heated to 310 K, and a 5 ns equilibration with protein and RNA backbone atoms constrained using a harmonic potential of the form $U(x) = k(x - x_{\text{ref}})^2$, where k is 1 kcal $\text{mol}^{-1} \text{Å}^{-2}$ and x_{ref} is the initial atom coordinates. Finally, 20 ns production simulations were performed without constraints. The simulation was performed under the NPT ensemble assuming the CHARMM36 force field (47) for the protein and RNA molecules and the TIP3P model for the water molecules. All simulation trajectory analysis including RMSD and root mean square fluctuation were performed using VMD.

Data availability

Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data

Mycobacterial Erm38 and its RNA-binding site

bank under accession numbers 7F8A (unliganded Erm38), 7F8B (Erm38 complexed with SAM), and 7F8C (Erm38 complexed with SFG).

Supporting information—This article contains supporting information.

Acknowledgments—We thank Yee Hwa Wong, Swathi Pasunooti, Abirami R, Fu Shan Sherlyn Chia, and Su Shan Khoh for their help at the initial stages of the project. We also thank the scientists in the NTU Protein Production Platform (<https://proteins.sbs.ntu.edu.sg>) for the cloning and expression tests of the protein constructs, the scientists, and beamline staff at the Proxima 2A, SOLEIL (Saint Aubin, France; Proposal 20180290) for their expert assistance, and the staff at the National Supercomputing Centre, Singapore (<https://www.nsc.sg>), for providing access to ASPIRE-1 for the MD simulations. This work was supported by grants from the National Research Foundation of Singapore through the Singapore-MIT Alliance for Research and Technology (SMART) Antimicrobial Resistance Interdisciplinary Research Group and the SMART Innovation Centre (ING-000772 BIO IGN), and by a grant from the A*STAR BMRC Therapeutics Development Review (H19H0G1003/TOR 005–018).

Author contributions—B. C. G., J. L., and P. C. D. conceptualization; B. C. G. and X. X. methodology; B. C. G., X. X., J. L., and P. C. D. data curation; B. C. G., X. X., J. L., and P. C. D. formal analysis; B. C. G. and X. X. visualization; B. C. G. writing—original draft; X. X., J. L., and P. C. D. writing—review and editing; J. L. and P. C. D. supervision.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: AMR, antimicrobial resistance; Erms, Erythromycin resistance methyltransferases; IMAC, immobilized metal affinity chromatography; MLS, macrolides, lincosamides, and streptogramins; PBST, phosphate-buffered saline with Tween; VMD, visual molecular dynamics.

References

- Hofer, U. (2019) The cost of antimicrobial resistance. *Nat. Rev. Microbiol.* **17**, 3
- Chellat, M. F., Raguz, L., and Riedl, R. (2016) Targeting antibiotic resistance. *Angew. Chem. Int. Ed. Engl.* **55**, 6600–6626
- Morić, I., Savić, M., Ilić-Tomić, T., Vojnović, S., Bajkić, S., and Vasiljević, B. (2010) rRNA methyltransferases and their role in resistance to antibiotics. *J. Med. Biochem.* **29**, 165–174
- Wilson, D. N. (2014) Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nat. Rev. Microbiol.* **12**, 35–48
- Buriankova, K., Doucet-Populaire, F., Dorson, O., Gondran, A., Ghnassia, J. C., Weiser, J., and Pernet, J. L. (2004) Molecular basis of intrinsic macrolide resistance in the *Mycobacterium tuberculosis* complex. *Antimicrob. Agents Chemother.* **48**, 143–150
- Nash, K. A., Brown-Elliott, B. A., and Wallace, R. J., Jr. (2009) A novel gene, Erm(41), confers inducible macrolide resistance to clinical isolates of *Mycobacterium abscessus* but is absent from *Mycobacterium chelonae*. *Antimicrob. Agents Chemother.* **53**, 1367–1376
- Vester, B., and Douthwaite, S. (2001) Macrolide resistance conferred by base substitutions in 23S rRNA. *Antimicrob. Agents Chemother.* **45**, 1–12
- Schroeder, M. R., and Stephens, D. S. (2016) Macrolide resistance in *Streptococcus pneumoniae*. *Front. Cell. Infect. Microbiol.* **6**, 98
- Fyfe, C., Grossman, T. H., Kerstein, K., and Sutcliffe, J. (2016) Resistance to macrolide antibiotics in public health pathogens. *Cold Spring Harb. Perspect. Med.* **6**, a025395
- Madsen, C. T., Jakobsen, L., Buriankova, K., Doucet-Populaire, F., Pernet, J. L., and Douthwaite, S. (2005) Methyltransferase Erm(37) slips on rRNA to confer atypical resistance in *Mycobacterium tuberculosis*. *J. Biol. Chem.* **280**, 38942–38947
- Ding, Y. L., Fu, J., Chen, J., Mo, S. F., Xu, S., Lin, N., Qin, P., and McGrath, E. (2018) Molecular characterization and antimicrobial susceptibility of *Staphylococcus aureus* isolated from children with acute otitis media in Liuzhou, China. *BMC Pediatr.* **18**, 388
- Park, A. K., Kim, H., and Jin, H. J. (2010) Phylogenetic analysis of rRNA methyltransferases, Erm and KsgA, as related to antibiotic resistance. *FEMS Microbiol. Lett.* **309**, 151–162
- Stsiapanava, A., and Selmer, M. (2019) Crystal structure of ErmE - 23S rRNA methyltransferase in macrolide resistance. *Sci. Rep.* **9**, 14607
- Yu, L., Petros, A. M., Schnuchel, A., Zhong, P., Severin, J. M., Walter, K., Holzman, T. F., and Fesik, S. W. (1997) Solution structure of an rRNA methyltransferase (ErmAM) that confers macrolide-lincosamide-streptogramin antibiotic resistance. *Nat. Struct. Biol.* **4**, 483–489
- Bussiere, D. E., Muchmore, S. W., Dealwis, C. G., Schluckebier, G., Nienaber, V. L., Edalji, R. P., Walter, K. A., Lador, U. S., Holzman, T. F., and Abad-Zapatero, C. (1998) Crystal structure of ErmC', an rRNA methyltransferase which mediates antibiotic resistance in bacteria. *Biochemistry* **37**, 7103–7112
- Schluckebier, G., Zhong, P., Stewart, K. D., Kavanaugh, T. J., and Abad-Zapatero, C. (1999) The 2.2 Å structure of the rRNA methyltransferase ErmC' and its complexes with cofactor and cofactor analogs: Implications for the reaction mechanism. *J. Mol. Biol.* **289**, 277–291
- Brown-Elliott, B. A., Vasireddy, S., Vasireddy, R., Iakhiaeva, E., Howard, S. T., Nash, K., Parodi, N., Strong, A., Gee, M., Smith, T., and Wallace, R. J. (2015) Utility of sequencing the Erm(41) gene in isolates of *Mycobacterium abscessus* subsp. *abscessus* with low and intermediate clarithromycin MICs. *J. Clin. Microbiol.* **53**, 1211–1215
- Oates, M. E., Romero, P., Ishida, T., Ghalwash, M., Mizianty, M. J., Xue, B., Dosztányi, Z., Uversky, V. N., Obradovic, Z., Kurgan, L., Dunker, A. K., and Gough, J. (2012) D2P2: Database of disordered protein predictions. *Nucleic Acids Res.* **41**, D508–D516
- Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51**, 2778–2786
- Madsen, C. T., Jakobsen, L., and Douthwaite, S. (2005) *Mycobacterium smegmatis* Erm(38) is a reluctant dimethyltransferase. *Antimicrob. Agents Chemother.* **49**, 3803–3809
- Farrow, K. A., Lyras, D., Polekhina, G., Koutsis, K., Parker, M. W., and Rood, J. I. (2002) Identification of essential residues in the Erm(B) rRNA methyltransferase of *Clostridium perfringens*. *Antimicrob. Agents Chemother.* **46**, 1253–1261
- Holm, L., and Laakso, L. M. (2016) Dali server update. *Nucleic Acids Res.* **44**, W351–W355
- Rowe, S. J., Mecaskey, R. J., Nasef, M., Talton, R. C., Sharkey, R. E., Halliday, J. C., and Dunkle, J. A. (2020) Shared requirements for key residues in the antibiotic resistance enzymes ErmC and ErmE suggest a common mode of RNA recognition. *J. Biol. Chem.* **295**, 17476–17485
- Carothers, J. M., Oestreich, S. C., and Szostak, J. W. (2006) Aptamers selected for higher-affinity binding are not more specific for the target ligand. *J. Am. Chem. Soc.* **128**, 7929–7937
- Maravić, G., Bujnicki, J. M., Feder, M., Pongor, S., and Flögel, M. (2003) Alanine-scanning mutagenesis of the predicted rRNA-binding domain of ErmC??? Redefines the substrate-binding site and suggests a model for protein-RNA interactions. *Nucleic Acids Res.* **31**, 4941–4949
- Newman, J., Egan, D., Walter, T. S., Megeed, R., Berry, I., Ben Jelloul, M., Sussman, J. L., Stuart, D. I., and Perrakis, A. (2005) Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: The PACT/JCSG+ strategy. *Acta Crystallogr. D Biol. Crystallogr.* **61**, 1426–1431
- Gorrec, F. (2009) The MORPHEUS protein crystallization screen. *J. Appl. Crystallogr.* **42**, 1035–1042
- Gorrec, F. (2015) The MORPHEUS II protein crystallization screen. *Acta Crystallogr. F Struct. Biol. Commun.* **71**, 831–837

29. Duran, D., Couster, S. L., Desjardins, K., Delmotte, A., Fox, G., Meijers, R., Moreno, T., Savko, M., and Shepard, W. (2013) PROXIMA 2A – a new fully tunable micro-focus beamline for macromolecular crystallography. *J. Phys. Conf. Ser.* **425**, 012005
30. Broennimann, C., Eikenberry, E. F., Henrich, B., Horisberger, R., Huelsen, G., Pohl, E., Schmitt, B., Schulze-Briese, C., Suzuki, M., Tomizaki, T., Toyokawa, H., and Wagner, A. (2006) The PILATUS 1M detector. *J. Synchrotron Radiat.* **13**, 120–130
31. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132
32. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242
33. Long, F., Vagin, A. A., Young, P., and Murshudov, G. N. (2008) BALBES: A molecular-replacement pipeline. *Acta Crystallogr. D Biol. Crystallogr.* **64**, 125–132
34. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
35. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367
36. Smart, O. S., Womack, T. O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vornrhein, C., and Bricogne, G. (2012) Exploiting structure similarity in refinement: Automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 368–380
37. Humphrey, W., Dalke, A., and Schulten, K. (1996) Vmd: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38
38. Jaroensuk, J., Wong, Y. H., Zhong, W., Liew, C. W., Maenpuen, S., Sahili, A. E., Atichartpongkul, S., Chionh, Y. H., Nah, Q., Thongdee, N., McBee, M. E., Prestwich, E. G., DeMott, M. S., Chaiyen, P., Mongkolsuk, S., *et al.* (2019) Crystal structure and catalytic mechanism of the essential m(1)G37 tRNA methyltransferase TrmD from *Pseudomonas aeruginosa*. *RNA* **25**, 1481–1496
39. Bhujbalrao, R., and Anand, R. (2019) Deciphering determinants in ribosomal methyltransferases that confer antimicrobial resistance. *J. Am. Chem. Soc.* **141**, 1425–1429
40. Feder, M., Purta, E., Kosciński, L., Cubrilo, S., Maravic Vlahovicek, G., and Bujnicki, J. M. (2008) Virtual screening and experimental verification to identify potential inhibitors of the ErmC methyltransferase responsible for bacterial resistance against macrolide antibiotics. *ChemMedChem* **3**, 316–322
41. Vester, B., Nielsen, A. K., Hansen, L. H., and Douthwaite, S. (1998) ErmE methyltransferase recognition elements in RNA substrates. *J. Mol. Biol.* **282**, 255–264
42. Zhong, P., Pratt, S. D., Edalji, R. P., Walter, K. A., Holzman, T. F., Shivakumar, A. G., and Katz, L. (1995) Substrate requirements for ErmC methyltransferase activity. *J. Bacteriol.* **177**, 4327–4332
43. Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R., and Hofacker, I. L. (2008) The Vienna RNA websuite. *Nucleic Acids Res.* **36**, W70–W74
44. Wang, J., Wang, J., Huang, Y., and Xiao, Y. (2019) 3dRNA v2.0: An updated web server for RNA 3D structure prediction. *Int. J. Mol. Sci.* **20**, 4116
45. Yan, Y., Zhang, D., Zhou, P., Li, B., and Huang, S.-Y. (2017) Hdock: A web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res.* **45**, W365–W373
46. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802
47. Best, R. B., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M., and MacKerell, A. D. (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone Φ , Ψ and side-chain χ_1 and χ_2 dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273
48. Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D., and Lopez, R. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641
49. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324
50. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10037–10041