Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Zane Laughlin Date

Mechanism of Ribosomal Subunit Recognition and Modification by the *Mycobacterium tuberculosis* Ribosomal RNA Methyltransferase TlyA

By

Zane T. Laughlin
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Biochemistry, Cell and Developmental Biology

Graeme L. Conn, Ph.D.
Advisor

Anita H. Corbett, Ph.D.
Committee Member

Christine M. Dunham, Ph.D.
Committee Member

Bo Liang, Ph.D.
Committee Member

William M. Shafer, Ph.D.
Committee Member

David S. Weiss, Ph.D.
Committee Member

Accepted:

Kimberly Jacob Arriola, Ph.D, MPH
Dean of the James T. Laney School of Graduate Studies

Date

Mechanism of Ribosomal Subunit Recognition and Modification by the *Mycobacterium tuberculosis* Ribosomal RNA Methyltransferase TlyA

By

Zane Timothy Laughlin
B.S., University of Virginia, 2016

Advisor: Graeme L. Conn, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Biochemistry, Cell and Developmental Biology
2022

Abstract

Mechanism of Ribosomal Subunit Recognition and Modification by the *Mycobacterium tuberculosis* Ribosomal RNA Methyltransferase TlyA

By Zane T. Laughlin

Antibiotics are a vital component of medicine today and have been extensively used since their discovery at the beginning of the 20th century. Since then, however, bacteria have developed resistance mechanisms to counter all classes of antibiotics. In 2019, it was estimated that antibiotic-resistant bacteria and fungi cause almost 3 million infections and 35,900 deaths per year in the United States. One mechanism of resistance to ribosome targeting antibiotics is the expression of ribosomal RNA (rRNA) methyltransferases that chemically modify the antibiotic binding site to reduce its binding. However, in the case of intrinsically-expressed ribosomal methyltransferase TlyA of *Mycobacterium tuberculosis*, modification of the ribosome by this enzyme instead confers susceptibility to tuberactinomycin antibiotics like capreomycin. TlyA incorporates 2'-O methylations on two bacterial ribosome nucleotides: 16S rRNA C1409 of the 30S subunit and 23S rRNA C1920 of the 50S subunit (*E. coli* numbering). The full-length structure of TlyA and how it recognizes and modifies these two sites in either of their different contexts, until now, was unknown.

In this work, we present the cryo-electron microscopy (cryo-EM) structure of a mycobacterial 50S subunit-TlyA complex trapped in a post-catalytic state using a S-adenosyl-L-methionine analog. Complementary functional analyses reveal a conserved surface of residues that spans both TlyA domains which recognizes the 23S rRNA. Specifically, conserved TlyA residues make critical contacts to unique tertiary structure at the base of 23S rRNA Helix 69 as well as surrounding the site of modification. The mechanism of modification determined from these studies positions the TlyA active site over the target nucleotide C1920 whose base is flipped from Helix 69. This work suggests that base flipping may be a common mechanism among rRNA methyltransferase enzymes even for those enzymes like TlyA which do not modify the nucleotide base. Finally, additional functional analyses on the 30S subunit suggest that the same 23S rRNA recognition surface of TlyA engages this second substrate, but with different dependencies on specific critical residues for binding.

Collectively, the work presented here details the mechanism by which TlyA recognizes and modifies its ribosomal substrates, particularly the 50S ribosomal subunit, through the use of structural and activity studies and contributes to the overall knowledge of the molecular mechanism of recognition and modification of other ribosomal methyltransferases.

Mechanism of Ribosomal Subunit Recognition and Modification by the Mycobacterium tuberculosis Ribosomal RNA Methyltransferase TlyA

By

Zane Timothy Laughlin
B.S., University of Virginia, 2016

Advisor: Graeme L. Conn, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Biochemistry, Cell and Developmental Biology
2022

Table of Contents

Table of Figures

Chapter 1

Chapter 2

Chapter 3

Table of Tables

Chapter 3

Chapter 1

Introduction

Antibiotics: Discovery and Usage

Antibiotics are a critical component of modern medicine and have been in heavy use since their discovery in the beginning of 20th century. However, the origin of antibiotics dates much farther back as several species of Streptomyces and other antibiotic-producing soil-dwelling bacteria are estimated to have evolved ~440 million years ago, with antibiotics probably being much older (1). It was initially speculated that these natural products were first produced by bacteria to eliminate or limit the growth of other competitor bacteria living in the soil nearby. However, considering that these compounds are used in this natural setting at significantly lower concentration than seen in the clinic, it is possible that they instead evolved as a means of communication or coordination with other bacteria in building and stabilizing bacterial communities (2).

Even before their discovery in the 20th century, antibiotic-producing organisms had been inadvertently implemented by humans in preventing infection for thousands of years through use of mold and soil for treating wounds (3). However, the first naturally-produced antibiotic, penicillin from *Penicillium rubens*, was not discovered until 1928 by Alexander Fleming (4). It should be noted, however, that the idea of anti-infective treatments had existed some time before Fleming's discovery. In 1910, Paul Ehrlich developed the first synthetic antibiotic, salvarsan, for use in treating syphilis (5) in work that became the genesis of the idea of chemotherapy. Shortly after the discovery of synthetic and natural antibiotics, Selman Waksman defined an antibiotic as "*a compound made by a microbe to destroy other microbes*". Critically, in his own research he identified the producers of many of these compounds as belonging to the genus Streptomyces, a genus which served as the origin of most types of natural antibiotics used in the 20th century, as well as their derivatives (6).

Following the discovery of these compounds and the development of processing for their mass production and purification, antibiotic treatments were implemented around the world for many types of bacterial infections and illnesses and completely revolutionized the field of medicine. Many diseases which were in the past difficult to treat and often fatal, such as tuberculosis and sexually-transmitted infections, were now quickly and easily cured by a single compound. However, in the years following the introduction of these antibiotics, clinicians and researchers noted the appearance of resistance to these drugs in bacteria. In the case of penicillin, clinical resistance to treatment was identified ~10 years after its discovery and this trend of resistance following introduction has continued for all antibiotics in use today (7).

Antibiotic Mechanism of Action

Antibacterials (antibiotics which target bacteria) are classified as either bactericidal or bacteriostatic, meaning they either kill bacteria or inhibit their growth, respectively. This antibacterial action is accomplished through diverse mechanisms, including the targeting of the bacterial cell wall, DNA replication, RNA synthesis, or protein synthesis by the ribosome (**Fig. 1A**)

Antibiotics targeting the bacterial cell wall include classes of drugs such as the uridyl peptides, mannopeptimycins, glycopeptides, and, most notably, beta-lactams (**Fig. 1B**). Beta-lactams, such as penicillin, inhibit peptide bond formation by bacterial transpeptidases which are responsible for building and cross-linking the peptidoglycan units that make up the cell wall. If beta-lactams are present, they prevent the correct building of the cell wall or result in changes in its shape that are detrimental to essential

bacterial activities like cell division; this can lead to cellular stress responses and ultimately to cell lysis (8).

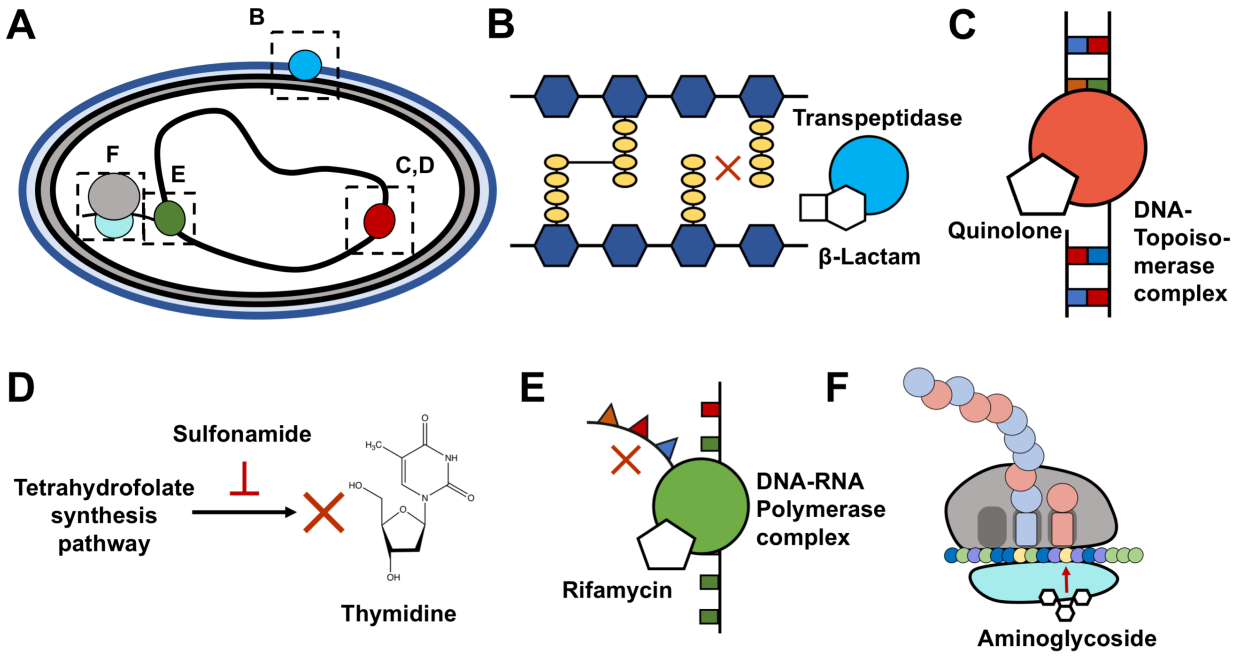


Figure 1. Antibiotics target diverse processes and locations within the bacterial cell. A. Overview of a bacterial cell displaying the cell wall, membrane, and DNA along with the particular targets of antibiotics: transpeptidase, DNA topoisomerase complex and thymidine production, DNA-RNA polymerase complex, and the ribosome. **B.** Beta-lactam antibiotics inhibit the activity of transpeptidase, preventing cross-linking of the cell wall. **C.** Quinolones inhibit the re-ligation activity of DNA-topoisomerase complex, leaving the DNA backbone broken and preventing replication from occurring. **D.** Sulfonamides inhibit the tetrahydrofolate synthesis pathway and block the production of thymidine, a nucleotide needed for normal cell function. **E.** Rifamycins inhibit the transcription initiation process of the DNA-RNA Polymerase complex, blocking the production of mRNA. **F.** Aminoglycosides and many other classes of antibiotic target the ribosome; aminoglycosides typically reduce the accuracy of mRNA decoding, leading to incorrect amino acids being added to proteins.

Antibiotics which inhibit DNA replication generally either interfere with DNA supercoiling by targeting DNA-topoisomerase complexes (aminocoumarins and quinolones) or interfere with nucleic acid synthesis by targeting tetrahydrofolate

synthesis (sulfonamides) (**Fig. 1C,D**). Topoisomerases assist in resolving the over- or underwinding of DNA during replication by cleaving the phosphate backbone to allow DNA to be easily wound or unwound. Antibiotics like quinolones, which target DNA-topoisomerase complexes, allow for the cleavage of the DNA but not for its re-ligation. This action thus leaves the affected region of the DNA inaccessible for replication in cell division, leading to cell death (9). Sulfonamides like sulfamethoxazole inhibit enzymes in the synthetic pathway of tetrahydrofolate which is needed to produce thymidine (10). Without thymidine, new strands of DNA (as well as RNA) cannot be produced, thereby halting DNA replication and leading to cell death.

Antibiotics inhibiting RNA synthesis, such as rifamycins, sorangicin, and streptolydigin, generally target DNA-dependent transcription performed by the DNA-RNA polymerase complex (**Fig. 1E**). For example, rifamycins bind the DNA-bound subunit of the RNA polymerase to sterically inhibit transcription initiation (11). Without the process of transcription, normal cell processes like mRNA and rRNA production cannot continue, leading to cell death.

Finally, the last major target of antibacterials is the bacterial ribosome (**Fig. 1F**). Ribosomes produce proteins required for cell function and structure, and if this process is halted or interfered with, the cell cannot perform many of the functions required for its survival. Protein production by the ribosome is common to all living cells but, despite their functional similarities, bacterial and eukaryotic ribosomes have a number of structural and functional differences that allow antibiotics to bind to and specifically interfere with the activity of bacterial ribosomes. This makes the ribosome an excellent target for antibiotics.

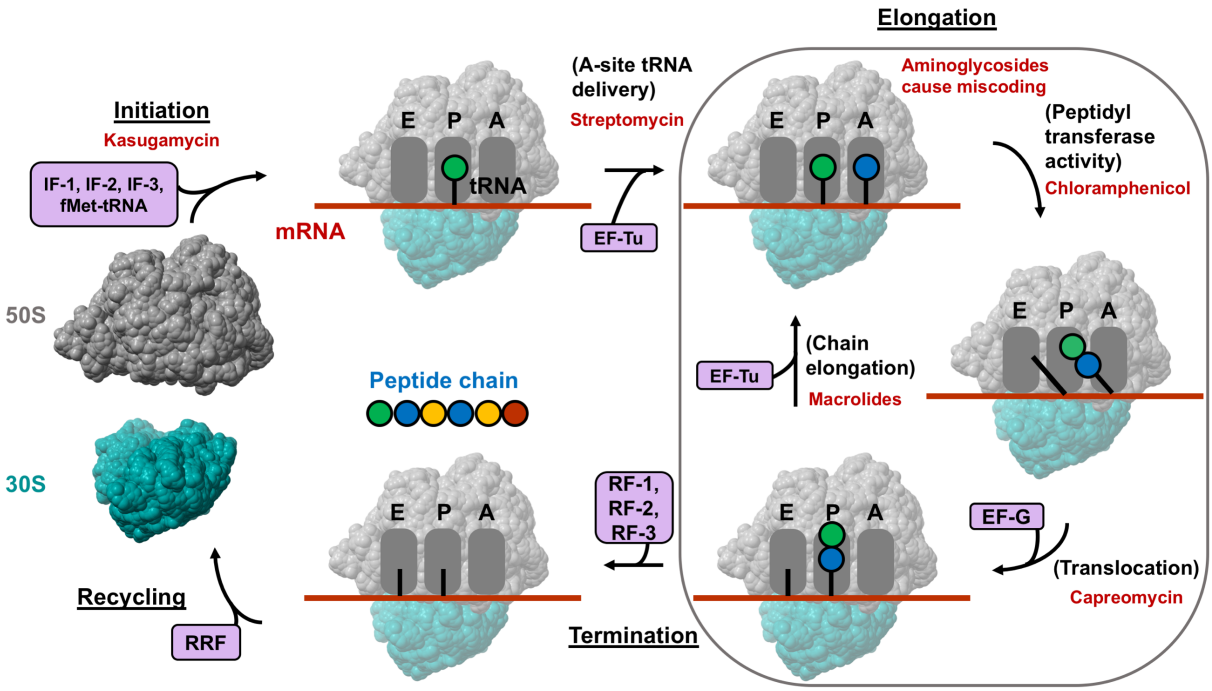


Figure 2. Ribosome-targeting antibiotics can inhibit each step of translation. Ribosomal translation occurs in four major steps: Initiation, Elongation, Termination, and Recycling. Initiation (*left*) begins with the 50S and 30S ribosomal subunits associating on an mRNA message with the help of initiation factors 1, 2, and 3 (IF-1, IF-2, and IF-3) with initiator tRNA (fMet-tRNA^{fMet}) at the start site on the mRNA in the ribosomal P site. Next, elongation involves delivery of an aminoacylated-tRNA to the A site of the ribosome aided by elongation factor Thermounstable (EF-Tu). For a cognate pairing, a peptidyl transfer reaction occurs between the amino acids, disconnecting the chain from the P-site tRNA. The P- and A-site tRNAs then take on the P/E and A/P hybrid states, respectively, before translocation occurs with the aid of elongation factor G (EF-G), fully moving the deacylated P-site tRNA out of the ribosome via the E site, while the amino-acylated tRNA with the elongated peptide chain moves fully from the A to P site. Cycle of elongation continues until termination occurs when a stop codon is reached. The peptide chain is then released with the help of release factors 1, 2, or 3 (RF-1, RF-2, and RF-3) and the ribosomal subunits are dissociated from the mRNA, aided by ribosome recycling factor (RRF), and recycled to be made available for subsequent rounds of translation. Ribosome-targeting antibiotics are known which affect each of these steps (labeled in red at the step of translation they inhibit).

Translation, the protein-producing activity of the ribosome, occurs in four steps: initiation, elongation, termination, and recycling; and antibiotics can act on each of these steps (**Fig. 2**). On the small (30S) ribosomal subunit, drugs primarily bind at the decoding center of the ribosome, proximal to the mRNA and tRNA during translation. For example, aminoglycosides like kanamycin and gentamicin bind at the decoding center and induce miscoding, causing incorrect aminoacyl-tRNAs to bind to the mRNA codon and leading to aberrant protein production and eventually to cell death (12). Among other 30S-binding antibiotics, kasugamycin inhibits translation initiation by preventing the interaction of initiator tRNA and the start codon, while streptomycin inhibits elongation by interfering with the delivery of tRNAs to the ribosomal aminoacyl (A) site (13). On the large (50S) ribosomal subunit, for the most part antibiotics target and bind at or around the peptidyl transferase center (PTC). Chloramphenicol, for example, inhibits peptide bond formation by binding at the A site of the ribosome and blocking the entry of additional tRNA (14). Macrolides, in comparison, do not directly block peptide bond formation but rather bind in the adjacent peptide exit tunnel of the ribosome and block the passage of the nascent peptide chain, preventing elongation (15). Still, other drugs like elfamycins (including kirromycin) inhibit translation by targeting the translation factors which assist in protein synthesis rather than the ribosome itself. Specifically, kirromycin binds elongation factor Tu (EF-Tu), thereby preventing EF-Tu from leaving the ribosome after it assists in the binding of aminoacyl-tRNA to the A site which then blocks translocation (16).

Antibiotic Resistance Mechanisms

As mentioned above, though antibiotics have been extremely effective in treating a variety of infections, resistance inevitably follows their usage. Across all classes of antibiotics, bacteria have developed resistance mechanisms to counter their activity. This can be attributed to both the heavy clinical usage of particular antibiotics skewing selection toward resistant strains of bacteria, and the fact that bacterial antibiotic resistance mechanisms have most likely existed since the advent of antibiotics at least 440 million years ago (1). Bacteria have developed a number of mechanisms to abrogate the bactericidal or bacteriostatic activity of antibiotics, but the major overarching routes to resistance are: **1)** preventing entry or promoting export (efflux) of the antibiotic, **2)** inactivating the antibiotic by modifying or degrading it, **3)** protecting the target, **4)** overexpression of the target, or **5)** mutation or chemical modification of the target site (**Fig. 3**).

To reduce entry of antibiotics, and thereby prevent the drug from having its lethal or inhibitory effect on the cell because it cannot reach its target, certain species of bacteria like *Mycobacterium tuberculosis* and *Pseudomonas aeruginosa* have specialized outer

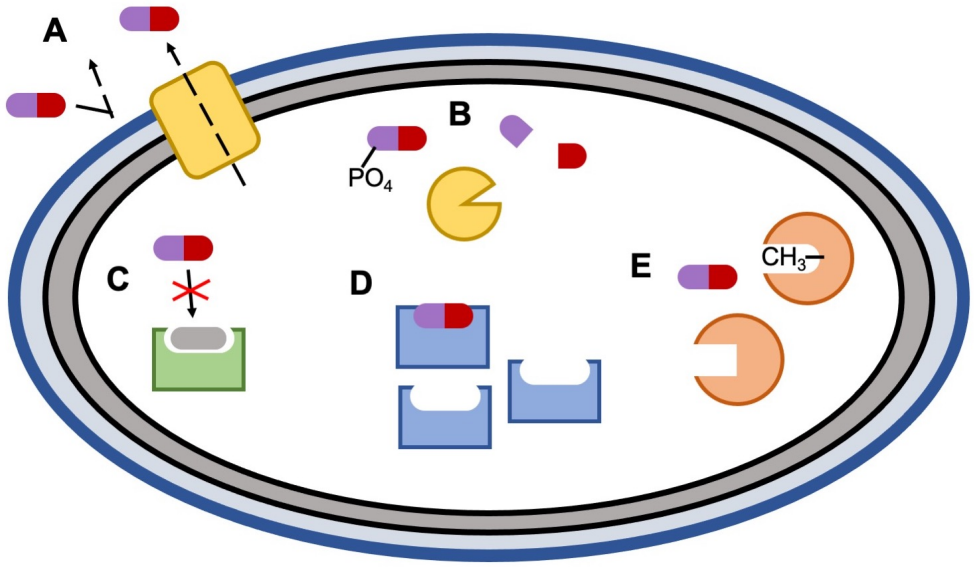


Figure 3. Bacteria employ diverse mechanisms of antibiotic resistance. A. Bacteria can prevent the entry of antibiotics into the cytoplasm using specialized membrane structures or cell wall alterations. Additionally, drugs and toxic compounds within the cell can be transported out using efflux pumps. **B.** Modification of antibiotics (shown as phosphorylation as one example) or degradation of antibiotics by resistance enzymes can decrease the concentration of active antibiotic within the cytoplasm. **C.** Protective groups and compounds can bind at or near the antibiotic target site and even remove antibiotic from its binding site to prevent drug binding and activity. **D.** Bacteria can also overexpress the antibiotic target to keep cell machinery functioning as normal. **E.** The drug target site can be modified via mutation or modification by resistance enzymes to disallow the binding and activity

layers or membranes to restrict the influx of drugs (**Fig. 3A**) (17,18). Additionally, *Acinetobacter baumannii* has been shown to change the molecular makeup of its outer membrane to reduce entry of certain antibiotics (19). To promote exit of antibiotics from within the cell, some species have evolved or acquired efflux pumps such as those of the ATP-binding cassette (ABC), small multidrug resistance (SMR), and resistance-nodulation-cell division (RND) superfamilies of transporters (**Fig. 3A**). These transporters recognize and efflux a large variety of toxic compounds including antibiotics

from the periplasm or cytoplasm, decreasing the overall drug concentration within the cell (20).

Bacteria also use multiple mechanisms to target antibiotics for degradation or modification so that they can no longer bind their intended target or have their intended activity. One of the most well-known resistance enzymes, beta-lactamase, confers resistance to beta-lactam antibiotics by cleaving the eponymous beta-lactam ring of those drugs, rendering them unable to bind their target transpeptidases (**Fig. 3B**) (21). Resistance to aminoglycosides often arises through the action of aminoglycoside N-acetyltransferases and aminoglycoside O-phosphotransferases which add acetyl and phosphate groups to the $-NH_2$ or -OH groups of the aminoglycoside structure, respectively, thereby blocking the drug's ability to bind to the ribosome (**Fig. 3B**) (22).

Prevention of drug binding to its target site can also be achieved through the expression of factors that protect the target site from drug binding or by overexpression of the intended target of the drug to ensure a tolerable level of target remains functional within the cell. TetM and TetO are examples of factors that dislodge tetracycline antibiotics from its binding site, while overexpression of translation factors like EF-Tu can restore ribosome function in the presence of the EF-Tu inhibitor amythiacin (**Fig. 3C,D**) (23,24).

Additionally, modifying or mutating the intended target site of an antibiotic can prevent its binding through a change in charge or structure (**Fig. 3E**). In terms of ribosome-targeting antibiotics, the most common method of antibiotic resistance is mutation or modification of the target site. Ribosome-targeting antibiotics typically bind directly to the ribosomal RNA and interfere with the assembly of the ribosome or translation. Because of this, resistant strains of bacteria often possess mutations at or near

drug-binding sites. For example, rifampin one of the first-line drugs used to treat *M. tuberculosis* (*Mtb*) infections and binds within a highly conserved pocket of the DNA-dependent RNA polymerase (gene *rpoB*) to inhibit its activity, thereby inhibiting transcription. Single point mutants in *rpoB* leading to amino acid substitutions in the polymerase have been found in the clinic which can decrease the affinity of rifampin for this polymerase and yield high levels of rifampin resistance while allowing transcription to continue (25). Additionally, resistance to linezolid, a commonly-used ribosome-targeting antibiotic which disrupts placement of aminoacyl-tRNA in the peptidyl transferase center, is often gained through mutation of nucleotide G2576 within the 23S rRNA which is proximal to the drug's binding site and disrupts its binding to that site (26).

Resistance genes associated with these drugs often encode enzymes like methyltransferases which methylate the ribosomal surface to prevent antibiotic binding (**Fig. 4**). Erythromycin methyltransferases (Erm) are a family of enzymes that methylate (or dimethylate) the N6 amino group of A2058 of the 23S rRNA (*E. coli* numbering) within the nascent peptide exit tunnel. This modification blocks erythromycin from binding in the nascent peptide exit tunnel of the ribosome, thereby allowing the nascent peptide chain to move unhindered during translation and resulting in erythromycin resistance (27). Thiostrepton is a thiopeptide antibiotic that binds the GTPase center of the 50S ribosomal subunit at the interaction surface between ribosomal protein L11 and the 23S

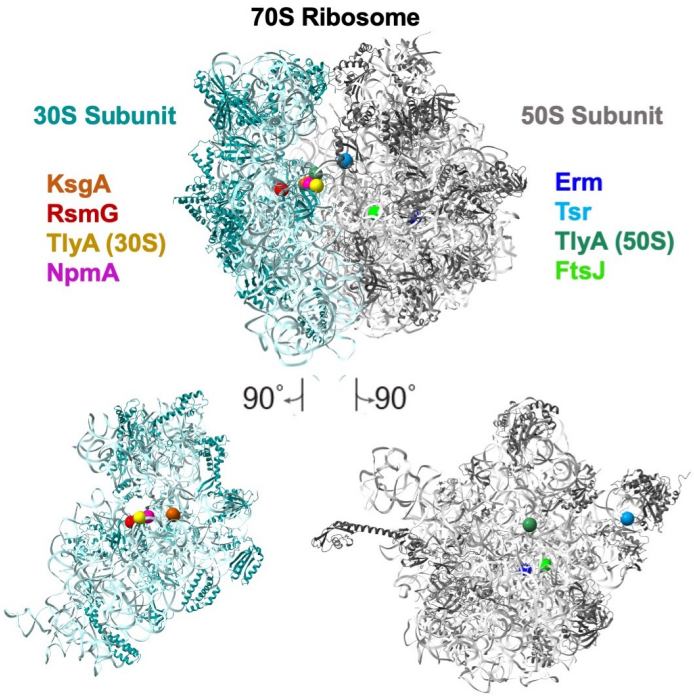


Figure 4. Modification sites of selected ribosomal RNA methyltransferases. The *E. coli* 70S ribosome complex (*top*) is shown with the modification sites indicated (colored spheres) for methyltransferases: 16S rRNA KsgA (A1518 and A1519 in *E. coli* numbering, N6; orange), RsmG, (G527, N7; red), TlyA (C1409, ribose 2'-OH; yellow), NpmA (A1408, N1; magenta), and 23S rRNA Erm (A2058, N6; blue), Tsr (A1067, ribose 2'-OH; light blue), TlyA (C1920, ribose 2'OH; teal), and FtsJ (U2552, ribose 2'-OH; lime). Lower panels show 90˚ rotated views of the individual subunits.

rRNA (28). Resistance to this antibiotic in the thiopstrepton producer *Streptomyces azureus* is given by methyltransferase Tsr which modifies the 2'-O of A1067 (29).

While most modifications to the ribosome are a common pathway of resistance to ribosome-targeting antibiotics, some drugs actually lose efficacy if the ribosome is not modified. In particular, the drugs kasugamycin, streptomycin, and capreomycin require methylation performed by endogenously expressed enzymes at specific sites on the ribosome to retain activity. Kasugamycin is dependent on N6- dimethylation of nucleotides A1518 and A1519 of the 16S rRNA of the 30S ribosomal subunit for its activity

and when methyltransferase KsgA (the enzyme responsible for these modifications) is inactivated, kasugamycin is much less effective (30). Streptomycin resistance can occur when the activity of RsmG (GibD), which incorporates an N7 methylation of G527 (m^7G527) of the 16S rRNA, is inhibited (31). Additionally, capreomycin (used in second-line treatment of drug-resistant tuberculosis) requires methylation at sites C1920 on the 23S rRNA of the 50S ribosomal subunit and C1409 on the 16S of the 30S ribosomal subunit for activity. This drug loses activity when these methylations are no longer performed by methyltransferase TlyA (32).

Drug Development and Strategies to Combat Resistance

Since their introduction in the beginning of the 20th century, antibiotics have remained a critical part of medicine. While commonly used in treating most bacterial infections, antibiotics are also critical for preventing infections following surgery and organ transplants, as well protecting critically-ill individuals from opportunistic infections. However, antibiotic resistance has only grown since the introduction of these compounds and is a serious threat to a large part of our arsenal of medications and treatments in use today. This has led the World Health Organization to name antimicrobial resistance one of the most important public health threats today (33). The increasing incidence and rate of antibiotic resistance in the clinical setting has created a fear amongst scientists of the advent of a "post-antibiotic" era in medicine. In this scenario, the decreased effectiveness of antibiotics leads to increased mortality from infections that were previously easily treated, emphasizing the need for new antibiotics and antibiotic stewardship strategies (34). Unfortunately, very few new antibiotics have been developed in recent years and antibiotic usage can be difficult manage on a worldwide scale (35). An alternate strategy

in fighting resistance that could also prove to be effective is attacking the resistance mechanism itself. If we can understand how bacteria are becoming resistant to antibiotics, we can design and engineer compounds to attack those resistance mechanisms directly (such as inhibitors which block the action of antibiotic efflux pumps or resistance methyltransferases), allowing our current pool of antibiotics to continue to be effective and available for use (36,37).

Mycobacterium tuberculosis

An effective pool of antibiotics is of particular importance for treating one of the most widespread bacteria-caused diseases in the world, tuberculosis (TB). *M. tuberculosis* is an acid-fast bacterium belonging to the phylum Actinobacteria that is the causative agent of TB. *Mtb* is a small bacillus (rod-shaped bacteria) that possesses a membrane similar to gram-positive bacteria but enclosed in an outer layer of mycolic acid. It is estimated that TB appeared about 70,000 years ago in humans and has been a persistent human pathogen since then, evolving alongside humans (38,39). Currently, about one quarter of the world's population is infected with *Mtb* and, as a result, TB is among the world's most widespread and fatal diseases (40). Before the start of the coronavirus pandemic in 2019, TB was the leading cause of death from a single infectious agent (per year) worldwide and lead to the deaths of approximately 1.5 million people in 2020 (41). Additionally, TB is among the leading causes of death among HIV-positive individuals (about 214,000 deaths in 2020) (41).

Mtb infection generally occurs when an individual inhales droplets containing bacteria and these bacteria colonize the lungs. The bacteria then undergo phagocytosis within immune and dendritic cells in the lungs which allows them to evade detection by

the immune system and undergo cell division. From there, *Mtb* is believed to migrate to the lymph nodes and blood stream where it moves to infect other locations within the lungs. In response to this process, the body will form immune cellular structures called granulomas around infected cells to contain the infection (39). Most individuals infected with *Mtb*, however, will not experience symptoms and the bacteria within them become a slowly-replicating, latent infection. However, this infection can become active or contagious under certain conditions (e.g. loss of immune function) and infected individuals can experience symptoms such as chest pain and/or a chronic coughing (39,42). Considering the extent of *Mtb* infection throughout the world and the lethality of TB, treatments for *Mtb* infection are of the utmost importance.

Treatment of *Mtb* infection and drug resistance

Typically, TB is treated with front line antibiotics including rifampin and isoniazid; approximately 85% of people who develop TB disease can be successfully treated with a 6-month regimen of these compounds (41). However, of the 9.9 million people who developed TB in 2020, about 7.5% of those individuals tested for drug resistance showed signs of an infection which was either multi- (MDR), pre-extensively, or extensively drug-resistant (XDR) (41). Though these are a minority of cases and resistance is maintaining a relatively stable incidence, drug resistance in strains of *Mtb* is a serious threat to the health and lives of the millions of people who are infected each year, particularly those with active infections.

In response to the appearance of resistance to drugs in use against *Mtb*, strategies for treatment were created to avoid the onset of pan-resistance (resistance to all drugs) within the worldwide population of *Mtb*. This primarily involves choosing which

antibiotics should be used in which state of treatment. For example, rifampin and isoniazid are common first-line drugs, meaning that they are used from the beginning of treatment. However, if a case appears to be resistant to the first-line treatment, a new drug or combination of new drugs is chosen to treat the infection. This separate set of compounds known as second-line drugs, which include fluoroquinolones and linezolid, are never to be used in a first-line case and are only used in the rarer drug-resistant cases. Capreomycin is a tuberactinomycin antibiotic that was also, until recently, used as a second-line antibiotic in treating drug-resistant TB. Though in use since 1968, in 2018 it was removed from WHO's recommendation list due to the availability of all-oral second-line regimens becoming available (discussed further in Chapter 2) (43). Like other antibiotics, drug resistance has also begun to appear in these second-line antibiotics, including capreomycin. Capreomycin itself is an interesting case for studying drug-resistance as capreomycin resistance can arise from loss of specific modifications on the ribosome. This is in contrast to most other ribosome-targeting drugs which usually lose effectiveness due to the presence of a new resistance modification(s) on the ribosome. As mentioned above, the modifications in question are located on both subunits of the *Mtb* ribosome and are incorporated by the ribosomal methyltransferase, TlyA.

Ribose 2'-OH Ribosomal Methyltransferase TlyA

The *tlyA* gene from *Serpula (Treponema) hyodysenteriae* was first cloned, sequenced, and expressed (in *E. coli*) in 1991 and its encoded protein (TlyA) was originally designated as a hemolysin (cyto**t**oxin hemo**ly**sin **A**) (44). However, mutations in *tlyA* were found to confer resistance to the ribosome-targeting tuberactinomycin antibiotic capreomycin and *in silico* modeling studies suggested that the TlyA protein structure was similar to

methyltransferase FtsJ/RrmJ (45,46). Collectively, these findings pointed to TlyA's activity as a ribosomal methyltransferase.

In 2006, it was found that *Mtb* and *Mycolicibacterium smegmatis* (*Msm*) (gene Rv1694) TlyA are SAM-dependent methyltransferases that methylate the 2'-OH of the ribose sugar of 16S rRNA nucleotide C1409 of the bacterial small (30S) ribosomal subunit and 23S rRNA C1920 of the large (50S) ribosomal subunit (**Fig. 5A-C**) (47). More detailed analysis of the TlyA enzyme family later revealed that bacteria expressing TlyA produce one of two distinct types of the enzyme: $TlyA^{I}$ which possesses shorter N- and C-terminal extensions and which only methylates site C1920, or $TlyA^{II}$ with longer N- and C-terminal extensions and capable of methylating both the C1409 and C1920 sites (32). *Mtb* maintains this gene despite the sensitivity it gives to tuberactinomycins, suggesting *tlyA* plays some critical, but currently unknown role, in cell function or survival. Loss of methylation incorporated by TlyA at these sites shows little effect on bacterial fitness (48), though there is some evidence that they may support the stability of the assembled 70S ribosome complex (49).

Mtb TlyA is 268 amino acids in length with a previously predicted two-domain structure: an N-terminal domain with a ribosomal protein S4-like fold (residues 1-59) and an FtsJ methyltransferase-like C-terminal domain (residues 64-268) (46). The two domains are connected by a short linker with the amino acid sequence RAWV. Previously, our lab determined the structure of the C-terminal domain of TlyA with the short interdomain linker via x-ray crystallography (**Fig. 5D-F**) (50). The C-terminal domain was found to take on the anticipated Class I methyltransferase fold while the four amino-acid long interdomain linker was found to be able to adopt either a loop conformation or to extend the first α-helix of the CTD. Additionally, it was predicted in this study, as

previously suggested, that the NTD takes on an S4-ribosomal protein fold (46,50). One of the more unexpected findings from these studies was that the interdomain linker was critical for SAM binding (50). Taken together with the ability of the linker to adopt

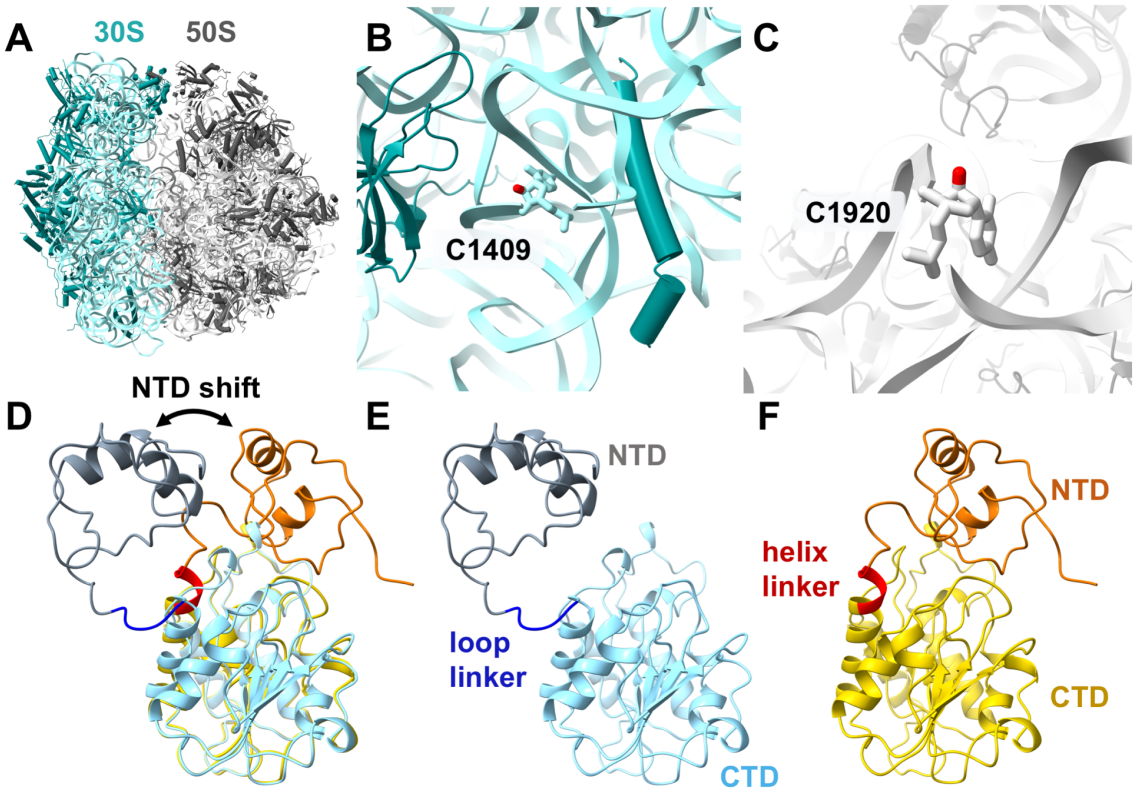


Figure 5. Knowledge of the TlyA modification sites and structure prior to this research. **A.** The *M. smegmatis* ribosome shown with the 30S subunit colored cyan and teal (16S rRNA and ribosomal S-proteins, respectively) and the 50S subunit colored as white and dark grey (23S rRNA and ribosomal L-proteins, respectively). **B.** Zoom-in on the 30S subunit, at the top of 16S rRNA helix 44 in the decoding center, with the target nucleotide, C1409 (*E. coli* numbering), shown in sticks and ribose 2'-OH colored red. **C.** Zoom-in on the 50S subunit centered on 23S rRNA Helix 69 with the target nucleotide, C1920 (*E. coli* numbering), shown in sticks and ribose 2'-OH colored red. **D.** Overlay with NTD positioning shift marked and **E, F.** individual views of the hybrid TlyA crystal structure/homology model highlighting the different interdomain linker conformations in the two crystal forms "loop" and "helix". With the "loop" linker conformation the model is shown as NTD (slate grey), linker (blue), and CTD (light blue); and with the "helix" conformation the model is shown as NTD (orange), linker (red),

different conformations and possibly alter the overall spatial relation between the N- and C-terminal domains, this observation led to speculation that the overall conformation of TlyA can alter the linker conformation and thereby affect SAM-binding, possibly taking on a conformation more amenable for SAM-binding once bound to the ribosome.

TlyA substrate recognition is of particular interest to our lab and the field of ribosomal methyltransferases because this enzyme can recognize and modify two contextually different sites: one on the large ribosomal subunit and the other on the small ribosomal subunit. Most ribosomal methyltransferases modify only one subunit, including the structurally homologous ribosomal methyltransferase FtsJ/RrmJ which modifies the 50S ribosomal subunit (51). One of the most notable differences between these two enzymes is FtsJ's significantly shorter length in both its N- and C-terminal domains (Fig. 6). This, along with the differences in recognition seen in the two types of TlyA (TlyA^{I} and TlyA^{II}) which also differ primarily in N-terminal and C-terminal domain length, suggests that structural motifs present in TlyA^{II} but not in FtsJ and TlyA^{I} may contribute to TlyA^{II}'s ability to target the second site on the 30S subunit (32).

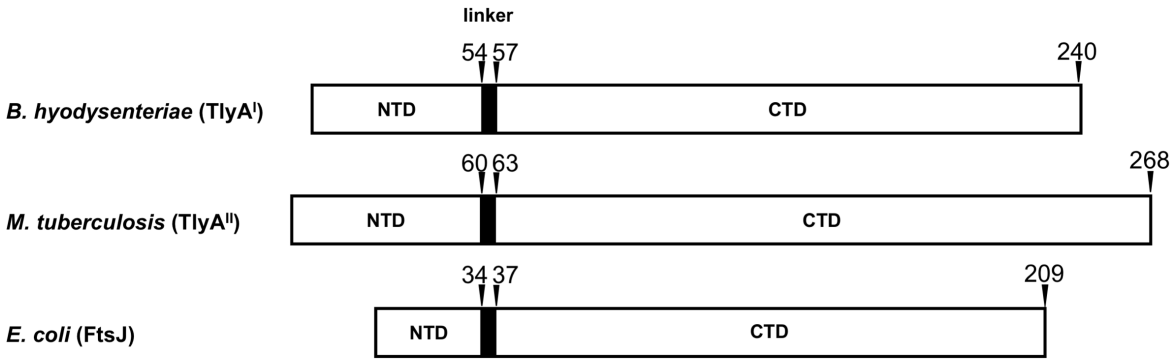


Figure 6. Difference in domain lengths between TlyA types I and II and FtsJ. The length of both N- and C-terminal domains of TlyA^{I} (N- 53 aa, C-183 aa), represented by the TlyA of *Brachyspira hyodysenteriae*, are shorter than those of TlyA^{II} (N- 59 aa, C-205 aa), represented by the TlyA of *Mtb*. Like TlyA^{I}, *E. coli* FtsJ has shorter N- and C-terminal domains (33 and 172 aa, repsectively) as compared to TlyA^{II} and like TlyA^{I}, only modifies a single site on the 50S ribosomal subunit. It is suspected that this difference in domain lengths and difference of motifs between the domains of TlyA types I and II are responsible for their difference in modification sites; i.e., TlyA^{I} only modifies 23S rRNA C1920 of the 50S ribosomal subunit while TlyA^{II} modifies both 23S rRNA C1920 and 16S rRNA C1409 of the 30S subunit.

Goals of this Research

Given the increasing prevalence of antibiotic resistance in *Mtb*, one of the most widespread and fatal pathogens in the world, and the rising antibiotic resistance seen across the spectrum of pathogens, understanding the mechanisms of antibiotic resistance is a critical if we wish to maintain our current arsenal of effective antibiotics. TlyA is an attractive subject of study in this regard due to 1) the widespread presence of resistance methyltransferases that confer resistance to ribosome-targeting antibiotics (52), 2) TlyA's presence in *Mtb* and role in capreomycin resistance when its modifications are lost (45), and 3) TlyA's dual substrate specificity (47). Specifically, our research into TlyA seeks to answer how TlyA recognizes and modifies its two target sites; the complex mechanism of which can be applied to understanding RNA methyltransferases as a whole, including

clinically-relevant resistance ribosomal methyltransferases (both 30S- and 50S-modifying).

Chapter 2 provides a more detailed introduction to the tuberactinomycin family of ribosome-targeting antibiotics which includes the anti-*Mtb* drugs viomycin and capreomycin. Beginning from the discovery of the drugs more than seven decades ago, this Chapter will detail the known tuberactinomycin biosynthetic pathways, their mechanism of action, their clinical use, and the emerging resistance facing this class of drug. While viomycin and capreomycin are no longer used to the same extent as in the past, understanding of their biosynthetic pathway and mechanisms of resistance can guide potential future development of derivatives based on the drugs themselves or their intermediates.

Chapter 3 describes the study of TlyA's mechanism of 50S ribosomal subunit recognition and modification. The structure of full-length TlyA was previously unknown, though our lab was able to determine the Class I methyltransferase structure of its C-terminal domain and interdomain linker, and performed modeling that suggested its N-terminal domain took on an S4-ribosomal protein fold (50). Additionally, how this enzyme recognizes and modifies its two different target sites in the large and small bacterial ribosome subunits was unknown. This chapter describes work that elucidates the full-length structure of TlyA bound to the 50S ribosomal subunit and maps the surface of positively-charged residues which span both the N- and C-terminal domains and are critical for TlyA's recognition of its modification sites. Additionally, and somewhat unexpectedly, TlyA is found to use a base-flipping mechanism of action for its activity which is seen commonly in DNA modifying and repair enzymes and has been observed or speculated for other antibiotic-resistance rRNA methyltransferases. Overall, this study

reveals the first example of an rRNA ribose 2'-OH methyltransferase bound to the bacterial 50S ribosomal subunit, the structural basis for its recognition of its target site, and the mechanism of its modification (53).

Altogether, these studies give us better insight into the mechanisms of how ribosomal modification enzymes, particularly methyltransferases, identify and modify their target sites. The analysis of the past, present, and future role of tuberactinomycin antibiotics, including capreomycin, demonstrates the impact these methyltransferases have on drug susceptibility, particularly ribosomal methyltransferase TlyA. Through our 50S-TlyA research, we have discovered the full-length structure of TlyA and shed a great deal of light on its mechanism of recognition and modification on the 50S ribosomal subunit and how it may methylate its second modification site on the 30S ribosomal subunit as well. This knowledge works to better our understanding of ribosome-targeting antibiotics and their associated resistance-methyltransferases, allowing us to better respond to resistance in the future.

References

1. Hutchings, M.I., Truman, A.W. and Wilkinson, B. (2019) Antibiotics: past, present and future. *Curr Opin Microbiol*, **51**, 72-80.
2. Spagnolo, F., Trujillo, M. and Dennehy, J.J. (2021) Why Do Antibiotics Exist? *mBio*, **12**, e0196621, PMCID: PMC8649755.
3. Haas, L.F. (1999) Papyrus of Ebers and Smith. *J Neurol Neurosurg Psychiatry*, **67**, 578, PMCID: PMC1736639.
4. Fleming, A. (2001) On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929. *Bull World Health Organ*, **79**, 780-790, PMCID: PMC2566493.
5. Gelpi, A., Gilbertson, A. and Tucker, J.D. (2015) Magic bullet: Paul Ehrlich, Salvarsan and the birth of venereology. *Sex Transm Infect*, **91**, 68-69, PMCID: PMC4318855.
6. Waksman, S.A., Schatz, A. and Reynolds, D.M. (2010) Production of antibiotic substances by actinomycetes. *Ann N Y Acad Sci*, **1213**, 112-124.
7. Abraham, E.P.C., E.; Fletcher, C.M.; Gardner, A.D.; Heatley, N.G.; Jennings, M.A.; Florey, H.W. (1941) Further observations on penicillin. *Lancet*, **238**, 177-189.
8. Tomasz, A. (1979) The mechanism of the irreversible antimicrobial effects of penicillins: how the beta-lactam antibiotics kill and lyse bacteria. *Annu Rev Microbiol*, **33**, 113-137.
9. Drlica, K., Malik, M., Kerns, R.J. and Zhao, X. (2008) Quinolone-mediated bacterial death. *Antimicrob Agents Chemother*, **52**, 385-392, PMCID: PMC2224783.
10. Hitchings, G.H. (1973) Mechanism of action of trimethoprim-sulfamethoxazole. I. *J Infect Dis*, **128**, Suppl:433-436 p.
11. Ho, M.X., Hudson, B.P., Das, K., Arnold, E. and Ebright, R.H. (2009) Structures of RNA polymerase-antibiotic complexes. *Curr Opin Struct Biol*, **19**, 715-723, PMCID: PMC2950656.
12. Krause, K.M., Serio, A.W., Kane, T.R. and Connolly, L.E. (2016) Aminoglycosides: An Overview. *Cold Spring Harb Perspect Med*, **6**, PMCID: PMC4888811.
13. Wilson, D.N. (2014) Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nat Rev Microbiol*, **12**, 35-48.
14. Schlunzen, F., Zarivach, R., Harms, J., Bashan, A., Tocilj, A., Albrecht, R., Yonath, A. and Franceschi, F. (2001) Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature*, **413**, 814-821.

15. Lenz, K.D., Klosterman, K.E., Mukundan, H. and Kubicek-Sutherland, J.Z. (2021) Macrolides: From Toxins to Therapeutics. *Toxins (Basel)*, **13**, PMCID: PMC8150546.

16. Barber, J., Carver, J.A., Leberman, R. and Tebb, G.M. (1988) The molecular basis of kirromycin (mocimycin) action; a 1H NMR study using deuterated elongation factor Tu. *J Antibiot (Tokyo)*, **41**, 202-206.

17. Curran, C.S., Bolig, T. and Torabi-Parizi, P. (2018) Mechanisms and Targeted Therapies for Pseudomonas aeruginosa Lung Infection. *Am J Respir Crit Care Med*, **197**, 708-727, PMCID: PMC5855068.

18. Putman, M., van Veen, H.W. and Konings, W.N. (2000) Molecular properties of bacterial multidrug transporters. *Microbiol Mol Biol Rev*, **64**, 672-693, PMCID: PMC99009.

19. Powers, M.J. and Trent, M.S. (2018) Phospholipid retention in the absence of asymmetry strengthens the outer membrane permeability barrier to last-resort antibiotics. *Proc Natl Acad Sci U S A*, **115**, E8518-E8527, PMCID: PMC6130378.

20. Du, D., Wang-Kan, X., Neuberger, A., van Veen, H.W., Pos, K.M., Piddock, L.J.V. and Luisi, B.F. (2018) Multidrug efflux pumps: structure, function and regulation. *Nat Rev Microbiol*, **16**, 523-539.

21. Tooke, C.L., Hinchliffe, P., Bragginton, E.C., Colenso, C.K., Hirvonen, V.H.A., Takebayashi, Y. and Spencer, J. (2019) beta-Lactamases and beta-Lactamase Inhibitors in the 21st Century. *J Mol Biol*, **431**, 3472-3500, PMCID: PMC6723624.

22. Ramirez, M.S. and Tolmasky, M.E. (2010) Aminoglycoside modifying enzymes. *Drug Resist Updat*, **13**, 151-171, PMCID: PMC2992599.

23. Connell, S.R., Tracz, D.M., Nierhaus, K.H. and Taylor, D.E. (2003) Ribosomal protection proteins and their mechanism of tetracycline resistance. *Antimicrob Agents Chemother*, **47**, 3675-3681, PMCID: PMC296194.

24. Gross, S., Nguyen, F., Bierschenk, M., Sohmen, D., Menzel, T., Antes, I., Wilson, D.N. and Bach, T. (2013) Amythiamicin D and related thiopeptides as inhibitors of the bacterial elongation factor EF-Tu: modification of the amino acid at carbon atom C2 of ring C dramatically influences activity. *ChemMedChem*, **8**, 1954-1962.

25. Floss, H.G. and Yu, T.W. (2005) Rifamycin-mode of action, resistance, and biosynthesis. *Chem Rev*, **105**, 621-632.

26. Mendes, R.E., Deshpande, L.M. and Jones, R.N. (2014) Linezolid update: stable in vitro activity following more than a decade of clinical use and summary of associated resistance mechanisms. *Drug Resist Updat*, **17**, 1-12.

27. Svetlov, M.S., Syroegin, E.A., Aleksandrova, E.V., Atkinson, G.C., Gregory, S.T., Mankin, A.S. and Polikanov, Y.S. (2021) Structure of Erm-modified 70S ribosome reveals the mechanism of macrolide resistance. *Nat Chem Biol*.

28. Dunstan, M.S., Hang, P.C., Zelinskaya, N.V., Honek, J.F. and Conn, G.L. (2009) Structure of the thiostrepton resistance methyltransferase.S-adenosyl-L-methionine complex and its interaction with ribosomal RNA. *J Biol Chem*, **284**, 17013-17020, PMCID: PMC2719339.

29. Thompson, J., Schmidt, F. and Cundliffe, E. (1982) Site of action of a ribosomal RNA methylase conferring resistance to thiostrepton. *J Biol Chem*, **257**, 7915-7917.

30. Ochi, K., Kim, J.Y., Tanaka, Y., Wang, G., Masuda, K., Nanamiya, H., Okamoto, S., Tokuyama, S., Adachi, Y. and Kawamura, F. (2009) Inactivation of KsgA, a 16S rRNA methyltransferase, causes vigorous emergence of mutants with high-level kasugamycin resistance. *Antimicrob Agents Chemother*, **53**, 193-201, PMCID: PMC2612157.

31. Mikheil, D.M., Shippy, D.C., Eakley, N.M., Okwumabua, O.E. and Fadl, A.A. (2012) Deletion of gene encoding methyltransferase (gidB) confers high-level antimicrobial resistance in Salmonella. *J Antibiot (Tokyo)*, **65**, 185-192.

32. Monshupanee, T., Johansen, S.K., Dahlberg, A.E. and Douthwaite, S. (2012) Capreomycin susceptibility is increased by TlyA-directed 2'-O-methylation on both ribosomal subunits. *Mol Microbiol*, **85**, 1194-1203, PMCID: PMC3438285.

33. (2021). World Health Organization, https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance.

34. (2014). World Health Organization, Geneva.

35. Dodds, D.R. (2017) Antibiotic resistance: A current epilogue. *Biochem Pharmacol*, **134**, 139-146.

36. Foik, I.P., Tuszynska, I., Feder, M., Purta, E., Stefaniak, F. and Bujnicki, J.M. (2018) Novel inhibitors of the rRNA ErmC' methyltransferase to block resistance to macrolides, lincosamides, streptogramine B antibiotics. *Eur J Med Chem*, **146**, 60-67.

37. AlMatar, M., Albarri, O., Makky, E.A. and Koksal, F. (2021) Efflux pump inhibitors: new updates. *Pharmacol Rep*, **73**, 1-16.

38. Banuls, A.L., Sanou, A., Van Anh, N.T. and Godreuil, S. (2015) Mycobacterium tuberculosis: ecology and evolution of a human bacterium. *J Med Microbiol*, **64**, 1261-1269.

39. Ehrt, S., Schnappinger, D. and Rhee, K.Y. (2018) Metabolic principles of persistence and pathogenicity in Mycobacterium tuberculosis. *Nat Rev Microbiol*, **16**, 496-507, PMCID: PMC6045436.

40. (2019). World Health Organization, Geneva.

41. (2021). World Health Organization, Geneva.

42. (2022). Mayo Clinic, Vol. 2022.

43. (2018). World Health Organization, Geneva.

44. Muir, S., Koopman, M.B., Libby, S.J., Joens, L.A., Heffron, F. and Kusters, J.G. (1992) Cloning and expression of a Serpula (Treponema) hyodysenteriae hemolysin gene. *Infect Immun*, **60**, 529-535, PMCID: PMC257660.

45. Maus, C.E., Plikaytis, B.B. and Shinnick, T.M. (2005) Mutation of tlyA confers capreomycin resistance in Mycobacterium tuberculosis. *Antimicrob Agents Chemother*, **49**, 571-577, PMCID: PMC547314.

46. Arenas, N.E., Salazar, L.M., Soto, C.Y., Vizcaino, C., Patarroyo, M.E., Patarroyo, M.A. and Gomez, A. (2011) Molecular modeling and in silico characterization of Mycobacterium tuberculosis TlyA: possible misannotation of this tubercle bacilli-hemolysin. *BMC Struct Biol*, **11**, 16, PMCID: PMC3072309.

47. Johansen, S.K., Maus, C.E., Plikaytis, B.B. and Douthwaite, S. (2006) Capreomycin binds across the ribosomal subunit interface using tlyA-encoded 2 '-O-methylations in 16S and 23S rRNAs. *Mol Cell*, **23**, 173-182.

48. Freihofer, P., Akbergenov, R., Teo, Y., Juskeviciene, R., Andersson, D.I. and Bottger, E.C. (2016) Nonmutational compensation of the fitness cost of antibiotic resistance in mycobacteria by overexpression of tlyA rRNA methylase. *RNA*, **22**, 1836-1843, PMCID: PMC5113204.

49. Salamaszynska-Guz, A., Taciak, B., Kwiatek, A. and Klimuszko, D. (2014) The Cj0588 protein is a Campylobacter jejuni RNA methyltransferase. *Biochem Bioph Res Co*, **448**, 298-302.

50. Witek, M.A., Kuiper, E.G., Minten, E., Crispell, E.K. and Conn, G.L. (2017) A Novel Motif for S-Adenosyl-l-methionine Binding by the Ribosomal RNA Methyltransferase TlyA from Mycobacterium tuberculosis. *J Biol Chem*, **292**, 1977-1987, PMCID: PMC5290967.

51. Caldas, T., Binet, E., Bouloc, P., Costa, A., Desgres, J. and Richarme, G. (2000) The FtsJ/RrmJ heat shock protein of Escherichia coli is a 23 S ribosomal RNA methyltransferase. *J Biol Chem*, **275**, 16414-16419.

52. Doi, Y., Wachino, J.I. and Arakawa, Y. (2016) Aminoglycoside Resistance: The Emergence of Acquired 16S Ribosomal RNA Methyltransferases. *Infect Dis Clin North Am*, **30**, 523-537, PMCID: PMC4878400.

53. Laughlin, Z.T., Nandi, S., Dey, D., Zelinskaya, N., Witek, M. A., Srinivas, P., Nguyen, H. A., Kuiper, E. G., Comstock, L. R., Dunham, C. M., and Conn, G. L. (In press) 50S subunit recognition and modification by the Mycobacterium tuberculosis ribosomal RNA methyltransferase TlyA. *Proc Natl Acad Sci U S A*.

Chapter 2

Tuberactinomycin antibiotics: Biosynthesis, anti-mycobacterial action and resistance

Zane T. Laughlin and Graeme L. Conn

A version of this review is currently in preparation for submission.

Tuberactinomycin antibiotics: Biosynthesis, anti-mycobacterial action and resistance

Zane T. Laughlin[1,2] and Graeme L. Conn[1,3]*

[1]Department of Biochemistry, Emory University School of Medicine, Atlanta, GA, 30322, USA.

[2]Graduate Program in Biochemistry, Cell and Developmental Biology (BCDB), Graduate Division of Biological and Biomedical Sciences, Emory University, Atlanta, GA, 30322,USA.

[3]Emory Antibiotic Resistance Center (ARC), Emory University, Atlanta, 30322, GA, USA.

*To whom correspondence should be addressed, email: gconn@emory.edu.

Running title: Tuberactinomycin antibiotics

Key Words: ribosome, antibiotic resistance, capreomycin, viomycin, tuberculosis (TB), mycobacteria, methyltransferase

Abstract

The tuberactinomycins are a family of cyclic peptide ribosome-targeting antibiotics with a long history of use as essential second-line treatments for drug-resistant tuberculosis. Beginning from the first identification of viomycin in the early 1950s, in this review we describe the discovery, chemical structure, and current knowledge of tuberactinomycin biosynthesis. Past and present application of these drugs in the treatment of *Mycobacterium tuberculosis* is also described alongside details of their mechanism of action on the ribosome as well as resistance mechanisms that have emerged since their introduction into the clinic. Finally, we discuss future potential applications of these drugs given the context of their recent removal from the World Health Organization's List of Essential Medicines.

Introduction

Antibiotics have been a critical component of modern medicine since their discovery in the beginning of the 20th century, offering effective treatments for otherwise potentially fatal bacterial infections and preventing infection during numerous other medical procedures such as surgery or during immunosuppression (1,2). The bacterial ribosome has provided fertile ground in nature and in the research laboratory for antibiotic development and is one of the primary targets of antibiotics in clinical use today (**Fig. 1**). However, resistance to these essential medicines among diverse human bacterial pathogens has developed against almost all classes of antibiotic since their introduction at the beginning of the 20th century (3). As such, it is widely recognized that improved antibiotics, and strategies for their use, must be developed to counter the increasing prevalence of resistance and to maintain our capacity to effectively treat diseases caused by bacterial infection. Without such action, we face the alternative future of a "post-antibiotic world" where common infections and diseases have a much greater mortality rate due to the ineffectiveness of antibiotics (4).

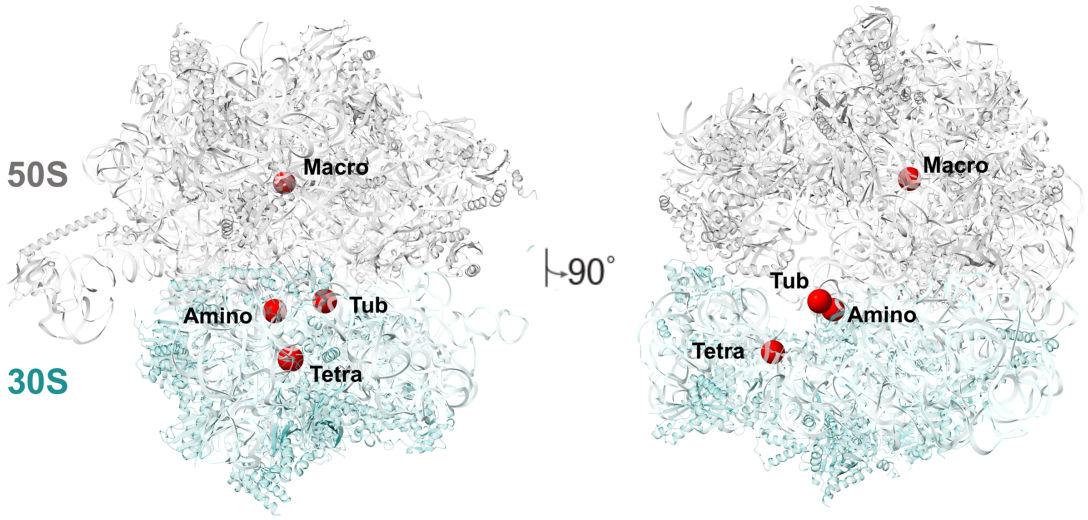


Figure 1. Ribosome targeting antibiotics bind at a variety of sites to inhibit translation. An *E. coli* 70S ribosome with the binding sites of various classes of ribosome targeting antibiotics shown as red spheres. Macrolides (Macro) bind the peptide exit tunnel of the ribosome and inhibit translation by blocking the movement of the nascent peptide chain through said tunnel. Tuberactinomycins (Tub) bind the intersubunit bridge between the 50S and 30S subunits and inhibit translocation (the movement of the ribosome to the next codon on the mRNA) during the elongation phase of translation. Aminoglycosides (Amino) bind at the decoding center of the ribosome and cause miscoding of the protein produced. Tetracyclines (Tetra) bind near the A site and inhibits translation by preventing the binding of aminoacyl-tRNA to the A site of the ribosome.

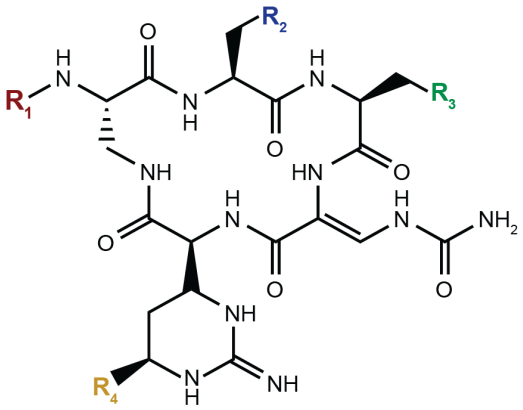
The tuberactinomycins are one example of an important class of ribosome-targeting antibiotics with a long history of clinical use (5). In particular, viomycin and capreomycin are potent anti-mycobacterial agents, effective in treating infection by *Mycobacterium tuberculosis* (*Mtb*), the causative agent of tuberculosis (TB) (6,7). These antibiotics were, until recently, employed primarily as second-line drugs against drug-resistant *Mtb* infections as a strategy designed, in part, to limit development of resistance. Nonetheless, like most other classes of antibiotics resistance to tuberactinomycins has developed, resulting in cases of *Mtb* resistant to multiple classes of antibiotics.

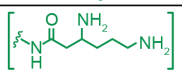
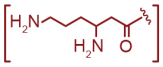
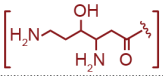
Beginning with their initial discovery and structural characterization in the early 1950s, this review describes current knowledge on tuberactinomycin biosynthesis, mechanisms of action on the ribosome, and clinically relevant resistance mechanisms that have emerged due to their use in treating TB. Finally, viewed through this lens, we speculate on the future potential use(s) of these and similar antibiotics.

Discovery and chemical structure of the tuberactinomycin antibiotics

Like most antibiotics in use today, tuberactinomycins are natural products of soil-dwelling bacteria. Viomycin being the first example isolated from *Streptomyces puniceus* in 1951 and was quickly demonstrated to have potent antibiotic activity against *Mtb* (6,8). Subsequent discoveries completed the currently recognized tuberactintomycin antibiotic family: tuberactinomycins A, B, N (also known as enviomycin) and O from *Streptomyces griseoverticillatus* var. *tuberacticus* in 1968; capreomycin from *Streptomyces capreolus* (now *Saccharothrix mutabilis* subspecies *capreolus*) in 1960; and tuberactinamine N, isolated in 1975 (9-16). Subsequent work to isolate the individual tuberactinomycins A, B, N and O revealed that tuberactinomycin B and viomycin were the same compound (17). Capreomycin was also initially identified as a mixture of four components (capreomycins IA, IB, IIA, and IIB) which were subsequently isolated and the differences in their core ring substituents characterized (10-12) (**Fig. 2**). Notably, however capreomycin went into clinical use two years later (in 1973) as a mixture of all four components (5).

Structurally, the tuberactinomycin antibiotics are defined by their conserved pentapeptide core ring containing at least one L-serine and at least one of the non-proteinogenic amino acids 2,3-diaminopropionate, L-capreomycidine, and β-ureidodehydroalanine (**Fig 2**). As described further in the next section, these cyclic peptides are



	R_1	R_2	R_3	R_4
Capreomycin IA	H	OH		H
Capreomycin IB	H	H		H
Capreomycin IIA	H	OH	NH_2	H
Capreomycin IIB	H	H	NH_2	H
Viomycin*		OH	OH	OH
Tuberactinomycin O		OH	OH	H
Tuberactinomycin A		OH	OH	OH
Tuberactinomycin N		OH	OH	H
Tuberactinamin N	H	OH	OH	H

*Also known as Tuberactinomycin B

Fig. 2. Chemical structures of the tuberactinomycin antibiotics. Tuberactinomycin antibiotics share a common pentapeptide core, with varying substituents at four positions (R1 to R4; indicated on the *right*). Distinguishing the main groups, tuberactinomycins A, B, N (viomycin), and O have a bulkier amine containing group at R1, while capreomycins IA, IB, IIA, and IIB have either an amino or bulkier amine containing substituent at position R3. Other positions (R2 and R3) vary between hydroxyl or hydrogen.

derived from non-ribosomally synthesized pentapeptides, which are cyclized and modified in a series of subsequent reactions to produce the final active compound.

Elucidation of tuberactinomycin biosynthesis

The viomycin biosynthetic gene cluster in *Streptomyces vinaceus* is ~36.3 kb long and contains 20 open reading frames (ORFs) (18). This cluster contains all the genes necessary for the biosynthesis, regulation, export, and activation of viomycin and also includes a viomycin resistance gene (*vph*, encoding viomycin phosphotransferase) to protect the producing bacterium from self-intoxication. A study by Barkei *et al.* has characterized the roles of some of the ORFs in the synthesis of viomycin (19). First, the cyclic pentapeptide core (common to all tuberactinomycins) is assembled by VioA, VioF, VioI, and VioG from various canonical and noncanonical amino acids. VioJ then desaturates part of the ring and additional modifications are added by VioL (carbamoylation), VioM (N-acylation), VioO (N-acylation), and VioQ (hydroxylation). The other genes associated with this biosynthetic cluster may not be directly

associated with the building of viomycin but perform other important tasks like synthesizing specialized amino acids for use in synthesis (VioP).

The capreomycin biosynthetic gene cluster in *Saccharothrix mutabilis* subspecies *capreolus* has also been characterized and comprises 33 ORFs (20). However, of these 33 ORFs, only 19 are proposed to be involved with the production of capreomycin, i.e. regulation of expression, non-ribosomal peptide biosynthesis, modification etc, and the function of the other 14 ORFs remains unknown. These ORFs have no sequence similarity to genes that would play an obvious role in capreomycin biosynthesis and sequence analyses of the viomycin and capreomycin gene clusters show that homology only exists between the already-identified ORFs (20). For example, gene *cmnU* of the capreomycin biosynthetic gene cluster is thought to encode a 16S rRNA m^1A1408 methyltransferase. Why the capreomycin biosynthesis cluster has a redundancy in its resistance genes is unclear. It is possible that this may give some additional protection against ribosome binding of capreomycin, alleviating toxicity that capreomycin or some secondary metabolite produced in the process of synthesizing capreomycin causes that viomycin does not.

Viomycin and capreomycin mechanism of antibiotic action

Tuberactinomycins are ribosome-targeting antibiotics that inhibit the process of translation, i.e. bacterial protein synthesis by the ribosome. Specifically, the bactericidal effect of this class of antibiotics is derived from their capacity to block the process of translocation, or movement of the ribosome on the mRNA to position the next three-nucleotide codon to be decoded within the aminoacyl tRNA binding site (A site) (21). During this step following peptide bond formation, the deacylated tRNA is moved from the peptidyl tRNA site (P site) to the exit site (E site), while the A-site tRNA, now carrying the elongating polypeptide chain, is moved to the P site. The movement of the A-site tRNA is driven by elongation factor G (EF-G) and hydrolysis of GTP but,

in the presence of viomycin, the ribosome is unable to move along the mRNA thereby stalling translation and hindering cell processes.

Structures of viomycin and capreomycin bound to ribosome 70S subunits have been determined via x-ray crystallography and both antibiotics share at least one common binding site at the subunit interface between 16S rRNA helix 44 (h44) on the 30S subunit and 23S rRNA

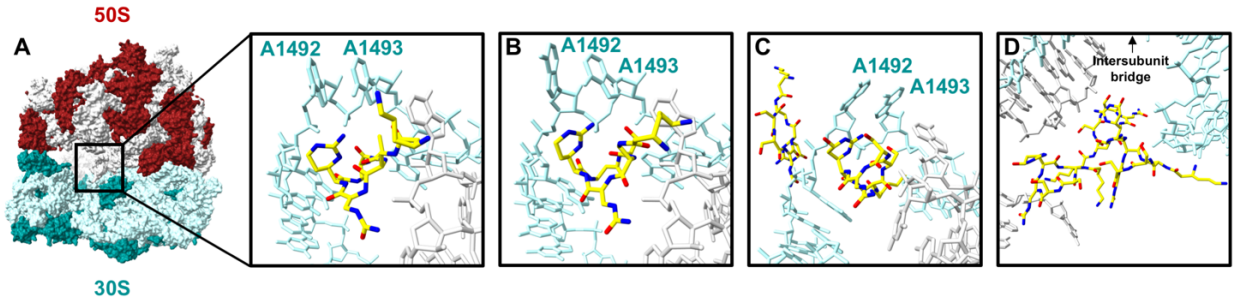


Figure 3. Capreomycin and viomycin bind the 70S ribosome at subunit interface ribosome. *A,* Overview of the *Mtb* 70S structure and a zoomed-in image of capreomycin binding at the 30S and 50S intersubunit bridge (PDB code 5V93). Note that 16S rRNA residues A1492/ A1493 which interrogate the mRNA/ tRNA pairing in the A site as part of the decoding process are flipped out from h44. Additional panels on the top row are views corresponding to *panel A* but for ***B***, capreomycin bound to the *Tth* 70S ribosome (PDB code 4V7M), ***C,*** viomycin bound to the *Tth* 70S *ribosome* (PDB code 4V7L), and ***D***, viomycin bound to the *E. coli* 70S ribosome at additional binding sites apart from the intersubunit bridge (PDB code 6LKQ).

Helix 69 on the 50S subunit (22,23) (**Fig. 3A-C**) Additionally, for capreomycin, the binding site and interactions with rRNA are largely conserved when bound to either *Mtb* or *Thermus thermophilus* (*Tth*) 70S ribosomes (compare **Fig. 3A-C**). Finally, a structural analysis of viomycin binding using *E. coli* ribosomes, which lack both rRNA ribose methylations needed for optimal tuberactinomycin drug binding (see below), identified multiple partially overlapping binding sites for viomycin (**Fig.3D**) (24). However, the concentration of viomycin in the sample preparation of this structure was very high at 0.5 mM and the fact that these structural studies were performed in ribosomes which were not C1409 and C1920 methylated does not give a full picture of what viomycin binding would look like in a clinical setting within Mtb. A future study

using a clinical concentration of viomycin and within Mtb or at least ribosomes of TlyA-expressing organism would perhaps give us better information on the binding sites of viomycin in-vivo.

At the primary binding site shared by capreomycin and viomycin, these antibiotics appear to stabilize the positioning of 16S rRNA nucleotides A1492 and A1493, and the 23S rRNA nucleotide A1913, located near tRNA in the ribosomal A site. A1492 and A1493 are responsible for interrogating tRNA/mRNA anti-codon-codon pairing and with these nucleotides stabilized in that position, it may halt the movement of the ribosome to the next mRNA codon. Additionally, in the drug bound structures, A1913 is positioned to form a hydrogen bond with the tRNA. which would make it more difficult for tRNAs to leave the A site. These interpretations are also supported by studies which show that viomycin dramatically increases tRNA affinity for the A site and can also promote back translocation (25,26). Collectively, these studies suggest that stabilization of aminoacyl tRNA in the A site of the ribosome underpins these drug's inhibition of the ribosomal translocation step.

Notably, capreomycin binding, and thus its antimycobacterial activity, is also reliant on ribose 2'-OH methylation of nucleotides C1409 of the 16S rRNA on the 30S subunit and C1920 of the 23S rRNA on the 50S (27,28). These modifications are incorporated by a single Class I S-adenosyl-L-methionine (SAM)-dependent methyltransferase TlyA. Many species of bacteria, including *Mtb*, express this enzyme but among those which do not express this methyltransferase, like *E. coli*, they are intrinsically less susceptible to tuberactinomycins such as capreomycin (27,28).

Clinical Use

According to the World Health Organization (WHO), tuberculosis was the second-leading cause of death from a single infectious agent worldwide in 2020 (behind COVID-19), resulting in approximately 1.5 million deaths (29). Further, of an estimated 9.9 million individuals who

developed TB in 2020, 7.5% of those tested for drug resistance indicating infection with either a multi- (MDR), pre-extensively (pre-XDR), or extensively (XDR) drug-resistant strain of *Mtb*. MDR TB is defined by the WHO as disease that is resistant to treatment with both rifiampicin and isoniazid, the two most effective first-line drugs in current use for TB. Pre-XDR TB is defined as resistant to rifampicin and any fluoroquinolone (a class of second-line drugs) and XDR TB as resistant to rifampicin, any fluoroquinolone, and at least one of bedaquiline or linezolid.

All of the tuberactinomycins have activity against TB, however, of these various natural products, only viomycin was used clinically where it was commonly implemented as a second-line treatment against drug-resistant *Mtb*. One major drawback to its usage is its associated ototoxicity and nephrotoxicity which is what eventually lead to it being replaced by the less-toxic capreomycin as a second-line treatment against drug-resistant *Mtb*. Capreomycin has a long history in the treatment of MDR TB and rifampin-resistant TB (RR-TB) and until 2018, capreomycin was included in the WHO Model List of Essential Medicines (30). However, because of the adverse side effects of capreomycin, was not recommended for children and those who have mild forms of TB (31). In 2018, the WHO revised guidelines around treatment of RR-TB and MDR-TB and have since recommended against use of injectable agents, including capreomycin, due to all-oral regimens being available to more patients worldwide (30).

Despite this recent decline in the use of tuberactinomycins for TB, with their broadly retained activity they remain in reserve and may find application in treatment of other infectious diseases. For example, viomycin has been shown to be effective vancomycin-resistant enterococci and MRSA (32,33). Additionally, capreomycin has shown to have anti-viral activity against SARS-CoV2 despite its previous use as strictly an antibiotic (34).

Mechanisms of Resistance to Tuberactinomycins

As noted in the previous section, recent estimates suggest that ~7.5% of *Mtb* infections leading to TB were rifampin-resistant, MDR or XDR, highlighting the importance of second-line

treatments like capreomycin (WHO 2020). Tuberactinomycins bind at a similar location to kanamycin and amikacin, aminoglycosides that are also used as second line drugs, so cross-resistance between these drugs in *Mtb* is common (35-37). Resistance to tuberactinomycins typically occurs through one of two potential mechanisms: mutation of the 16S rRNA surrounding the drug binding pocket or mutation to *tlyA*, the gene encoding methyltransferase TlyA (28,35,38).

Mutation of 16S rRNA can either to directly affect the ability of tuberactinomycins to bind to the intersubunit bridge of the 70S complex or to affect the ability of TlyA to modify 16S rRNA nucleotide C1409 on its ribose 2'-OH. Mutations of the 16S or 23S rRNA found in isolates of tuberactinomycin-resistant TB and XDR-TB occur at or near the modification site (C1409). Modification of the C1409 site seems to be more critical for the binding/action of capreomycin (39). As noted earlier, optimal binding of capreomycin to the ribosome, and thus its activity, is reliant on methylation of two specific sites in the 70S ribosome, C1409 and C1920 of the 16S and 23S rRNA, respectively, by the "housekeeping" rRNA methyltransferase TlyA (**Fig 4**).

Bacteria, including *Mtb*, can thus decrease their susceptibility to this antibiotic by reducing the modification of these sites through mutation of the *tlyA* gene. TlyA's role in bacteria is not entirely known and studies have found that removal of endogenous TlyA expression in certain bacteria has little cost to fitness but causes a reduction in the amount of functional 70S in other bacteria (40,41).

Conclusions and Future Perspective

Tuberactinomycins are a valuable class of antibiotics that have played an important historical role in treatment drug-resistant tuberculosis. Despite mounting resistance to these drugs and WHO recommending against the continued use of some of capreomycin, the door is still open to

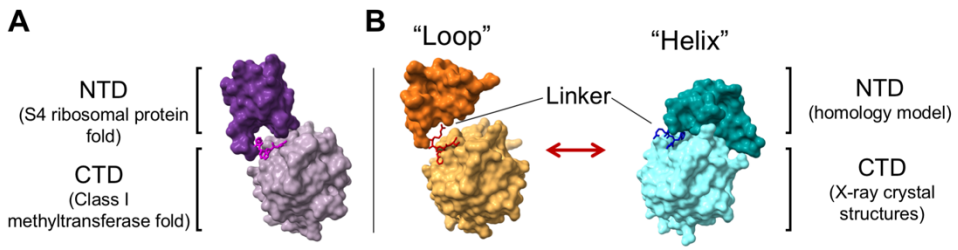


Figure 4. 2'-O methyltransferase TlyA. TlyA is a methyltransferase endogenously expressed by bacteria like *Mtb* and *Msm* which methylates the ribose 2'-OH of nucleotides C1409 of the 16S rRNA of the 30S subunit and C1920 of the 23S rRNA of the 50S subunit. TlyA consists of two folded domains and an interdomain linker. Previous cryo electron microscopy (**A**) and homology/x-ray crystallography (**B**) studies have revealed that the N-terminal domain takes on an S4-ribosome protein fold and is responsible for substrate recognition while the C-terminal domain takes on a class-I methyltransferase fold and is most likely responsible for recognition in addition to substrate modification. The linker between the two domains has been crystallized in two different conformations and is necessary for cosubstrate binding, suggesting it plays some role in TlyA's activity (42).

derivatives based on tuberactinomycins. Previous reports show that the addition of various compounds to solutions of bacteria expressing the biosynthetic gene cluster can alter the products of antibiotic production and can produce compounds based on the tuberactinomycin core (43). Additionally, via synthetic chemistry approaches may access new derivatives by altering current natural products members. New synthetic tuberactinomycin derivatives could also be designed to reduce toxicity, allowing these antibiotics to be used outside of a second-line treatment if necessary. Additionally, as discussed above, tuberactinomycins may find other clinical uses against other diseases, even viral-caused diseases like COVID-19, that would lengthen the lifespan of this important class of antibiotics.

Acknowledgements

We thank Dr. William Shafer for comments on the manuscript draft. Our research on TlyA is supported by NIH/ NIAID awards T32-AI106699 (to ZTL) and R01-AI088025 (to GLC).

References

1. Fleming, A. (2001) On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929. *Bull World Health Organ*, **79**, 780-790, PMCID: PMC2566493.
2. Gelpi, A., Gilbertson, A. and Tucker, J.D. (2015) Magic bullet: Paul Ehrlich, Salvarsan and the birth of venereology. *Sex Transm Infect*, **91**, 68-69, PMCID: PMC4318855.
3. Abraham, E.P.C., E.; Fletcher, C.M.; Gardner, A.D.; Heatley, N.G.; Jennings, M.A.; Florey, H.W. (1941) Further observations on penicillin. *Lancet*, **238**, 177-189.
4. (2014). World Health Organization, Geneva.
5. (1973) Evaluation of a new antituberculous agent. Capreomycin sulfate (capastat sulfate). *JAMA*, **223**, 179-180.
6. Bartz, Q.R., Ehrlich, J., Mold, J.D., Penner, M.A. and Smith, R.M. (1951) Viomycin, a New Tuberculostatic Antibiotic. *Am Rev Tuberc Pulm*, **63**, 4-6.
7. Sutton, W.B., Gordee, R.S., Wick, W.E. and Stanfield, L. (1966) In vitro and in vivo laboratory studies on the antituberculous activity of capreomycin. *Ann N Y Acad Sci*, **135**, 947-959.
8. Finlay, A.C., Hobby, G.L., Hochstein, F., Lees, T.M., Lenert, T.F., Means, J.A., Pan, S.Y., Regna, P.P., Routien, J.B., Sobin, B.A., Tate, K.B. and Kane, J.H. (1951) Viomycin a New Antibiotic Active against Mycobacteria. *Am Rev Tuberc Pulm*, **63**, 1-3.
9. Ando, T., Matsuura, K., Izumi, R., Noda, T., Take, T., Nagata, A. and Abe, J. (1971) Studies on Tuberactinomycin .2. Isolation and Properties of Tuberactinomycin-N, a New Tuberactinomycin Group Antibiotic. *J Antibiot*, **24**, 680-&.
10. Herr, E.B., Jr., Haney, G.E. and Pittenger, G.E. (1961), *140th Meeting of the American Chemical Society*.
11. Herr, E.B., Jr., Haney, G.E., Pittenger, G.E. and Higgens, C.E. (1960) Isolation and Characterization of a New Peptide Antibiotic. *Proceedings Indiana Academy of Sciences*, **69**, 134.
12. Herr, E.B., Jr. and Redstone, M.O. (1966) Chemical and physical characterization of capreomycin. *Ann N Y Acad Sci*, **135**, 940-946.

13. Nagata, A., Ando, T., Izumi, R., Sakakibara, H. and Take, T. (1968) Studies on tuberactinomycin (tuberactin), a new antibiotic. I. Taxonomy of producing strain, isolation and characterization. *J Antibiot (Tokyo)*, **21**, 681-687.

14. Wakamiya, T. and Shiba, T. (1975) Chemical Studies on Tuberactinomycin .8. Isolation of Tuberactinamine-N, Cyclic Peptide Moiety of Tuberactinomycin-N and Conversion of Tuberactinomycin-N to O. *J Antibiot*, **28**, 292-297.

15. Wakamiya, T., Shiba, T., Kaneko, T., Sakakibara, H., Take, T. and Abe, J.N. (1970) Chemical Studies on Tuberactinomycin .1. Structure of Tuberactidine, Guanidino Amino Acid Component. *Tetrahedron Lett*, 3497-+.

16. Yoshioka, H., Aoki, T., Goko, H., Nakatsu, K., Noda, T., Sakakiba.H, Take, T., Nagata, A., Abe, J., Wakamiya, T., Shiba, T. and Kaneko, T. (1971) Chemical Studies on Tuberactinomycin .2. Structure of Tuberactinomycin-O. *Tetrahedron Lett*, 2043-&.

17. Noda, T., Take, T., Nagata, A., Wakamiya, T. and Shiba, T. (1972) Chemical studies on tuberactinomycin. 3. The chemical structure of viomycin (tuberactinomycin B). *J Antibiot (Tokyo)*, **25**, 427-428.

18. Thomas, M.G., Chan, Y.A. and Ozanick, S.G. (2003) Deciphering tuberactinomycin biosynthesis: isolation, sequencing, and annotation of the viomycin biosynthetic gene cluster. *Antimicrob Agents Chemother*, **47**, 2823-2830, PMCID: PMC182626.

19. Barkei, J.J., Kevany, B.M., Felnagle, E.A. and Thomas, M.G. (2009) Investigations into viomycin biosynthesis by using heterologous production in Streptomyces lividans. *Chembiochem*, **10**, 366-376, PMCID: PMC2765823.

20. Felnagle, E.A., Rondon, M.R., Berti, A.D., Crosby, H.A. and Thomas, M.G. (2007) Identification of the biosynthetic gene cluster and an additional gene for resistance to the antituberculosis drug capreomycin. *Appl Environ Microbiol*, **73**, 4162-4170, PMCID: PMC1932801.

21. Modolell, J. and Vazquez. (1977) The inhibition of ribosomal translocation by viomycin. *Eur J Biochem*, **81**, 491-497.

22. Stanley, R.E., Blaha, G., Grodzicki, R.L., Strickler, M.D. and Steitz, T.A. (2010) The structures of the anti-tuberculosis antibiotics viomycin and capreomycin bound to the 70S ribosome. *Nat Struct Mol Biol*, **17**, 289-293, PMCID: PMC2917106.

23. Yang, K., Chang, J.Y., Cui, Z., Li, X., Meng, R., Duan, L., Thongchol, J., Jakana, J., Huwe, C.M., Sacchettini, J.C. and Zhang, J. (2017) Structural insights into species-specific features of the ribosome from the human pathogen Mycobacterium tuberculosis. *Nucleic Acids Res*, **45**, 10884-10894, PMCID: PMC5737476.

24. Zhang, L., Wang, Y.H., Zhang, X., Lancaster, L., Zhou, J. and Noller, H.F. (2020) The structural basis for inhibition of ribosomal translocation by viomycin. *Proc Natl Acad Sci U S A*, **117**, 10271-10277, PMCID: PMC7229676.

25. Peske, F., Savelsbergh, A., Katunin, V.I., Rodnina, M.V. and Wintermeyer, W. (2004) Conformational changes of the small ribosomal subunit during elongation factor G-dependent tRNA-mRNA translocation. *J Mol Biol*, **343**, 1183-1194.

26. Szaflarski, W., Vesper, O., Teraoka, Y., Plitta, B., Wilson, D.N. and Nierhaus, K.H. (2008) New features of the ribosome and ribosomal inhibitors: non-enzymatic recycling, misreading and back-translocation. *J Mol Biol*, **380**, 193-205.

27. Johansen, S.K., Maus, C.E., Plikaytis, B.B. and Douthwaite, S. (2006) Capreomycin binds across the ribosomal subunit interface using tlyA-encoded 2 '-O-methylations in 16S and 23S rRNAs. *Mol Cell*, **23**, 173-182.

28. Monshupanee, T., Johansen, S.K., Dahlberg, A.E. and Douthwaite, S. (2012) Capreomycin susceptibility is increased by TlyA-directed 2'-O-methylation on both ribosomal subunits. *Mol Microbiol*, **85**, 1194-1203, PMCID: PMC3438285.

29. (2021). World Health Organization, Geneva.

30. (2018). World Health Organization, Geneva.

31. WHO. (2016). 2016 Update ed, Vol. October 2016 Revision.

32. Dirlam, J.P., Belton, A.M., Birsner, N.C., Brooks, R.R., Chang, S.P., Chandrasekaran, R.Y., Clancy, J., Cronin, B.J., Dirlam, B.P., Finegan, S.M., Froshauer, S.A., Girard, A.E., Hayashi, S.F., Howe, R.J., Kane, J.C., Kamicker, B.J., Kaufman, S.A., Kolosko, N.L., LeMay, M.A., Linde, R.G., Lyssikatos, J.P., MacLelland, C.P., Magee, T.V., Massa, M.A., Miller, S.A., Minich, M.L., Perry, D.A., Petitpas, J.W., Reese, C.P., Seibel, S.B., Su, W.G., Sweeney, K.T., Whipple, D.A. and Yang, B.V. (1997) Cyclic homopentapeptides .1. Analogs of tuberactinomycins and capreomycin with activity against vancomycin-resistant enterococci and Pasteurella. *Bioorg Med Chem Lett*, **7**, 1139-1144.

33. Linde, R.G., Birsner, N.C., Chandrasekaran, R.Y., Clancy, J., Howe, R.J., Lyssikatos, J.P., MacLelland, C.P., Magee, T.V., Petitpas, J.W., Rainville, J.P., Su, W.G., Vu, C.B. and Whipple, D.A. (1997) Cyclic homopentapeptides .3. Synthetic modifications to the capreomycins and tuberactinomycins: Compounds with activity against methicillin-resistant Staphylococcus aureus and vancomycin-resistant Enterococci. *Bioorg Med Chem Lett*, **7**, 1149-1152.

34. Kumar, S., Singh, B., Kumari, P., Kumar, P.V., Agnihotri, G., Khan, S., Kant Beuria, T., Syed, G.H. and Dixit, A. (2021) Identification of multipotent drugs for COVID-19 therapeutics with the evaluation of their SARS-CoV2 inhibitory activity. *Comput Struct Biotechnol J*, **19**, 1998-2017, PMCID: PMC8025584.

35. Akbergenov, R., Shcherbakov, D., Matt, T., Duscha, S., Meyer, M., Wilson, D.N. and Bottger, E.C. (2011) Molecular basis for the selectivity of antituberculosis compounds capreomycin and viomycin. *Antimicrob Agents Chemother*, **55**, 4712-4717, PMCID: PMC3187005.

36. Jugheli, L., Bzekalava, N., de Rijk, P., Fissette, K., Portaels, F. and Rigouts, L. (2009) High level of cross-resistance between kanamycin, amikacin, and capreomycin among Mycobacterium tuberculosis isolates from Georgia and a

close relation with mutations in the rrs gene. *Antimicrob Agents Chemother*, **53**, 5064-5068, PMCID: PMC2786337.

37. Maus, C.E., Plikaytis, B.B. and Shinnick, T.M. (2005) Molecular analysis of cross-resistance to capreomycin, kanamycin, amikacin, and viomycin in Mycobacterium tuberculosis. *Antimicrob Agents Chemother*, **49**, 3192-3197, PMCID: PMC1196259.

38. Maus, C.E., Plikaytis, B.B. and Shinnick, T.M. (2005) Mutation of tlyA confers capreomycin resistance in Mycobacterium tuberculosis. *Antimicrob Agents Chemother*, **49**, 571-577, PMCID: PMC547314.

39. Monshupanee, T., Gregory, S.T., Douthwaite, S., Chungjatupornchai, W. and Dahlberg, A.E. (2008) Mutations in conserved helix 69 of 23S rRNA of Thermus thermophilus that affect capreomycin resistance but not posttranscriptional modifications. *J Bacteriol*, **190**, 7754-7761, PMCID: PMC2583624.

40. Freihofer, P., Akbergenov, R., Teo, Y., Juskeviciene, R., Andersson, D.I. and Bottger, E.C. (2016) Nonmutational compensation of the fitness cost of antibiotic resistance in mycobacteria by overexpression of tlyA rRNA methylase. *RNA*, **22**, 1836-1843, PMCID: PMC5113204.

41. Salamaszynska-Guz, A., Taciak, B., Kwiatek, A. and Klimuszko, D. (2014) The Cj0588 protein is a Campylobacter jejuni RNA methyltransferase. *Biochem Bioph Res Co*, **448**, 298-302.

42. Witek, M.A., Kuiper, E.G., Minten, E., Crispell, E.K. and Conn, G.L. (2017) A Novel Motif for S-Adenosyl-l-methionine Binding by the Ribosomal RNA Methyltransferase TlyA from Mycobacterium tuberculosis. *J Biol Chem*, **292**, 1977-1987, PMCID: PMC5290967.

43. Morse, B.K., Brown, M.S., Gagne, J.W., McArthur, H.A.I., McCormick, E.L., Murphy, T.K., Narrol, M.H., Perry, D.A., Smogowicz, A.A., Wax, R.G. and Wong, J.W. (1997) Production of tuberactinamine A by Streptomyces griseoverticillatus var. tuberacticus NRRL 3482 fed with (S)-2-aminoethyl-L-cysteine. *J Antibiot*, **50**, 698-700.

Chapter 3
50S subunit recognition and modification by the *Mycobacterium tuberculosis* ribosomal RNA methyltransferase TlyA

Zane T. Laughlin, Suparno Nandi, Debayan Dey, Natalia Zelinskaya, Marta A. Witek, Pooja Srinivas, Ha An Nguyen, Emily G. Kuiper, Lindsay R. Comstock, Christine M. Dunham, and Graeme L. Conn

A version of this manuscript is currently under review for publication in the Proceedings of the National Academy of Sciences of the United States of America.

50S subunit recognition and modification by the *Mycobacterium tuberculosis* ribosomal RNA methyltransferase TlyA

Zane T. Laughlin[1,2], Suparno Nandi[1,‡], Debayan Dey[1,‡], Natalia Zelinskaya[1], Marta A. Witek[1], Pooja Srinivas[1,3], Ha An Nguyen[1,4], Emily G. Kuiper[1], Lindsay R. Comstock[5], Christine M. Dunham[1,6] and Graeme L. Conn[1,6]*

[1]Department of Biochemistry, Emory University School of Medicine, Atlanta, GA, 30322, USA.

[2]Graduate Program in Biochemistry, Cell and Developmental Biology (BCDB), Graduate Division of Biological and Biomedical Sciences, Emory University, Atlanta, GA, 30322,USA.

[3]Graduate Program in Molecular and Systems Pharmacology (MSP), Graduate Division of Biological and Biomedical Sciences, Emory University, Atlanta, GA, 30322,USA.

[4]Department of Chemistry Graduate Program, Emory University, Atlanta, GA, 30322,USA

[5]Department of Chemistry, Wake Forest University, 455 Vine Street, Wake Downtown, NC, 27101, USA.

[6]Emory Antibiotic Resistance Center (ARC), Emory University, Atlanta, 30322, GA, USA.

‡These authors contributed equally

*To whom correspondence should be addressed, email: gconn@emory.edu.

Competing Interest Statement: The authors declare no competing interests.

Classification: Biological Sciences/ Biochemistry

Running title: Structure of the 50S subunit-TlyA complex

Key Words: Ribosome, RNA modification, antibiotic resistance, mycobacteria, methyltransferase.

Abstract

Changes in bacterial ribosomal RNA methylation status can alter the activity of diverse groups of ribosome-targeting antibiotics. These modifications are typically incorporated by a single methyltransferase that acts on one nucleotide target and rRNA methylation directly prevents drug binding, thereby conferring drug resistance. Loss of intrinsic methylation can also result in antibiotic resistance. For example, *Mycobacterium tuberculosis* becomes sensitized to tuberactinomycin antibiotics, such as capreomycin and viomycin, due to the action of the intrinsic methyltransferase TlyA. TlyA is unique among antibiotic resistance-associated methyltransferases as it has dual 16S and 23S rRNA substrate specificity and can incorporate cytidine-2'-O-methylations within two structurally distinct contexts. Here, we report the structure of a mycobacterial 50S subunit-TlyA complex trapped in a post-catalytic state with a S-adenosyl-L-methionine analog using single-particle cryogenic electron microscopy. Together with complementary functional analyses, this structure reveals critical roles in 23S rRNA substrate recognition for conserved residues across an interaction surface that spans both TlyA domains. These interactions position the TlyA active site over the target nucleotide C2144 which is flipped from 23S Helix 69 in a process stabilized by stacking of TlyA residue Phe157 on the adjacent A2143. Base flipping may thus be a common strategy among rRNA methyltransferase enzymes even in cases where the target site is accessible without such structural reorganization. Finally, functional studies with 30S subunit suggest that the same TlyA interaction surface is employed to recognize this second substrate, but with distinct dependencies on essential conserved residues.

Significance Statement

The bacterial ribosome is an important target for antibiotics used to treat infection. However, resistance to these essential drugs can arise through changes in ribosomal RNA (rRNA) modification patterns through the action of intrinsic or acquired rRNA methyltransferase enzymes. How these antibiotic resistance-associated enzymes recognize their ribosomal targets for site-specific modification is currently not well defined. Here, we uncover the molecular basis for large ribosomal (50S) subunit substrate recognition and modification by the *Mycobacterium tuberculosis* methyltransferase TlyA, necessary for optimal activity of the antitubercular drug capreomycin. From this work, recognition of complex rRNA structures distant from the site of modification and "flipping" of the target nucleotide base both emerge as general themes in ribosome recognition for bacterial rRNA modifying enzymes.

Introduction

Ribosome-targeting antibiotics are a structurally and mechanistically diverse group of anti-infectives that comprise a significant proportion of currently used treatments for bacterial infections (1, 2). However, among resistance mechanisms exploited by pathogenic bacteria to evade the effects of these antibiotics, ribosomal RNA (rRNA) drug-binding site methylation is already established or is quickly emerging as a major threat to such treatments (3, 4). For example, diverse human pathogens have acquired resistance modifications that impact the efficacy of aminoglycosides, macrolides and multiple other drug classes targeting the ribosome peptidyl transferase center. These modifications are incorporated by S-adenosyl-L-methionine (SAM)-dependent methyltransferases such as the Class I aminoglycoside-resistance 16S rRNA methyltransferases (e.g. NpmA and ArmA/RmtA-H), the Class I Erm family methyltransferases, and the radical SAM enzyme Cfr (5-7). Less commonly, reduced intrinsic methylation can also lead to resistance, such as for kasugamycin, streptomycin or capreomycin through loss of activity of the 16S rRNA methyltransferases RsmG/GidB (8, 9), KsgA (10), and TlyA (11), respectively.

Capreomycin is a member of the tuberactinomycin class of ribosome-targeting antibiotics and has an important history in the treatment of *Mycobacterium tuberculosis* (*Mtb*) infections resistant to the first-line drugs rifampin and isoniazid (12). Capreomycin binds at the subunit interface of mature 70S ribosomes, adjacent to 16S rRNA helix 44 (h44) of the small (30S) subunit and 23S rRNA Helix 69 (H69) of the large (50S) subunit (13). In a recent study, the tuberactinomycin antibiotic viomycin was also found to bind the 70S ribosome at several other locations (14), suggesting that this class of antibiotics may target multiple ribosomal sites to interfere with translation. Capreomycin's activity is thought to arise via stabilization of tRNA in the A site of the ribosome, thereby halting translation (13). Capreomycin has also been proposed to disrupt the interaction of ribosomal proteins uL10 and bL12, thereby blocking binding of elongation factors during translation (15). However, this mechanism is harder to reconcile with the binding sites of capreomycin and viomycin which are distant from both uL10 or bL12 (13, 14, 16), as well as the impact of changes in rRNA modification status in the A site on their activity.

Capreomycin binding to the *Mtb* 70S ribosome is dependent on 2'-O methylation of two nucleotides at the subunit interface, 16S rRNA C1392 and 23S rRNA C2144 (**Fig. 1A**; corresponding to *E. coli* nucleotides C1409 and C1920, respectively) (17). While the precise role of these modifications in ribosome structure and function is currently unclear, it is thought that they may somehow change the conformation of the rRNA allowing for optimal capreomycin binding (17, 18). Evolutionary maintenance of intrinsic rRNA modifications which increase sensitivity to antibiotics may be driven by their contribution to optimal fitness in the absence of drug or through decreased stability of unmodified 70S, as observed in *Mycolicibacterium smegmatis* (formerly *Mycobacterium smegmatis*; *Msm*) and *Campylobacter jejuni*, respectively (19, 20). Both modifications are incorporated by a single SAM-

dependent ribose 2'-O-methyltransferase, TlyA, encoded by Rv1694 in *Mtb* (18). TlyA has strong substrate preference for intact ribosomal subunits over free 16S or 23S rRNA, and individual modification of isolated subunits occurs prior to 70S assembly due to the target site locations on the interface surfaces of their respective subunits (**Fig. 1A**) (17). The TlyA family of methyltransferases is also divided into two groups based on their substrate specificities: Type I TlyA ($TlyA^{I}$) exclusively methylate 23S rRNA, while the slightly larger $TlyA^{II}$, including the *Mtb* enzyme, possess dual 16S and 23S specificity (17). However, how *Mtb* TlyA and other $TlyA^{II}$ enzymes recognize and modify these two structurally distinct substrates is not currently known.

We previously determined the crystal structure of the C-terminal domain (CTD) of *Mtb* TlyA with and without a four amino acid interdomain linker sequence (21). The TlyA CTD adopts the expected Class I methyltransferase fold but was unexpectedly found to be deficient in SAM binding in the absence of the interdomain linker. A TlyA CTD structure including the linker also revealed that this short motif can either extend the first α-helix of the CTD, or form a loop structure similar to that proposed earlier via homology modeling (21, 22). While a structure of the TlyA N-terminal domain is currently not available, modeling suggests a ribosomal protein S4-like domain (22). Collectively, these findings suggest that the N-terminal domain may be essential for rRNA recognition and binding, with the interdomain linker potentially playing a role in promoting SAM binding and methyltransferase activity in the CTD once bound to the correct substrate (21).

Here, we describe the structure of full-length *Mtb* TlyA bound to the *Msm* 50S subunit (hereafter, 50S-TlyA). The structure reveals the critical role played by the TlyA NTD in recognizing a complex 23S rRNA structure at the base of H69, positioning the TlyA CTD on H69 with its active site over the target nucleotide, C2144 (in *Msm* numbering, which is used exclusively hereafter unless noted). In addition, we find that TlyA uses a mechanism of base flipping for target site recognition and modification despite the accessibility of the C2144 ribose 2'-OH in H69, suggesting that this may be a general strategy for substrate molecular recognition among rRNA methyltransferases.

Results

***Determination of the 50S-TlyA complex structure*–**50S subunits without ribose modification on 23S rRNA nucleotide C2144 were isolated from *Msm* strain LR222 C101A which lacks TlyA activity and *Mtb* TlyA was expressed in *E. coli* and purified as previously described (21). A SAM-analog, "N-mustard 6" (NM6), was used to increase occupancy of TlyA on the 50S subunit (23, 24); NM6 is transferred by TlyA in its entirety to the ribose 2'-OH of C2144 and its covalent attachment to the 23S rRNA thus stabilizes the 50S-TlyA complex by virtue of the enzyme's affinity for both substrate and SAM analog (**Fig. S1**). Using this approach, we determined a 3.05 Å-resolution overall map of TlyA bound to the 50S subunit in a state immediately after catalysis of C2144 ribose modification by single-particle cryogenic electron microscopy (cryo-EM) (**Fig. 1**; **Fig. S2**, **Fig. S3A-C**).

One 23S rRNA feature, H54a (also called the “handle”), was significantly shifted from its position in the previously solved *Msm* 70S structure where it makes extensive interactions with the 30S subunit (PDB code 5O60) (25). This feature was visible in some 3D reconstructions (**Fig. S4**), where H54a lies across the subunit interface surface of the 50S subunit. In contrast, in other reconstructions, the map was weaker, suggesting H54a is dynamic in the free 50S subunit. H54a’s variable position and weak map adjacent to the bound TlyA in most 3D reconstructions suggest that H54a does not contribute to TlyA interaction with the 50S subunit despite its proximity to the enzyme NTD (**Fig. S4**).

An initial model for the 50S-TlyA complex was produced by docking a model of full-length TlyA, previously produced using a combination of NTD homology model and CTD crystal structure (PDB code 5KYG) (21), into unoccupied map surrounding 23S rRNA H69. This TlyA structure was subsequently rebuilt in Coot (26), including a complete rebuilding of the NTD (see Methods and Materials). Although the map was of sufficient quality for initial rebuilding of the NTD, the region corresponding to the TlyA CTD was less well defined. We therefore also performed multibody refinement with H69 and TlyA masked to separate this specific region of interest from the remainder of the 50S subunit (**Fig. S2**). This multibody refinement produced separate maps of the H69:TlyA complex (3.61 Å following post-processing; **Figs. S3D-F**) and the remaining 50S subunit structure lacking H69 (2.99 Å following post-processing; **Figs. S3G-I**). The former map was significantly improved compared to the corresponding region of the original map, with more information on the secondary structure and side chains of TlyA providing insights into how full-length TlyA interacts with its 23S rRNA substrate. Each map from multibody refinement was used for final model building and refinement of its associated structure, and the separate structures combined to generate a complete model of the 50S-TlyA complex (**Fig. 2**; **Fig. S2**).

When bound to the 50S subunit, the TlyA CTD is essentially identical to the previously determined structure of the isolated domain (PDB code 5KYG; 2.45 Å RMSD for 209 CTD Cα atoms), with the exception of a significant movement (~6-8 Å) of the loop containing residues 114-117 that is necessary to avoid clash with the minor groove surface of H69 (**Fig. S5A,B**). TlyA is structurally similar to the putative *S. thermophilus* hemolysin (PDB code 3HP7), but with a significant difference in the relative NTD and CTD orientation as a result of the distinct backbone path at the linker between the domains (**Fig. S5C,D**). The structures thus align well for superpositions based on either individual domain (3.68 and 2.00 Å RMSD for 209 CTD or 59 NTD Cα atoms, respectively), but less well for the full protein (overall 7.13 Å RMSD for 268 Cα atoms). The final NTD model reveals a globular domain with expected similarity to ribosomal protein S4 (2.55 Å RMSD for 59 NTD Cα atoms; **Fig. S5E**), comprising two adjacent short α-helices (residues 6-14 and 20-28) and two short β-strands (residues 32-24 and 52-54) preceding an interdomain linker (residues 60-63).

As described further in the following sections, TlyA binds the 50S subunit on its subunit interaction surface, with both TlyA domains surrounding H69 and the NTD making additional contacts

to the rRNA junction at the base of H69 (**Fig. 2A-D**). Together, the two TlyA domains form a continuous positively charged surface in contact with the 23 rRNA, suggesting that both play an important role in 50S subunit binding and specific substrate recognition (**Fig. 2D,E**). The final model also reveals the CTD of TlyA with bound SAM analog NM6 positioned directly over H69 residue C2144 in a post-catalytic state (i.e., with C2144 modified with NM6) .

TlyA NTD residues Arg6 and Arg20 exploit a complex rRNA structure for specific substrate recognition–Eight residues in the TlyA NTD were identified to make potentially critical interactions with nucleotides at the base of H69 and the adjacent rRNA junction: Arg4, Arg6, Arg18, Ser19, Arg20, Gln21, Gln22, and Lys41 (**Fig. 3A**). Three of these residues, Arg4, Arg6 and Arg20, are clustered around a complex (non-A-form helical) RNA structure formed by nucelotides C2149-G2153 of the 23 rRNA sequence immediately following H69 (**Fig. 3B**; **Fig. S6A**). While Arg4 is positioned to form a single electrostatic interaction with the phosphate group of A2151, Arg6 and Arg20 each form interaction networks with multiple rRNA nucleotides and with each other, likely stabilized by additional interactions with Asp8 (**Fig. 3B**). Specifically, Arg6 recognizes a sharp turn in the rRNA backbone via contacts with the bridging oxygen of A2151 and non-bridging oxygens of A1552, as well as a cation-π stacking interaction on the nucleobase of U2150. Similarly, Arg20 recognizes the phosphate backbone of 23S rRNA via electrostatic interactions with the phosphate groups of C2149 and U2150 (**Fig. 3B**). Consistent with critical roles in 23S rRNA recognition for Arg6 and Arg20, these two residues, as well as Asp8, are almost universally conserved among TlyA homologs (**Fig. 3C**; **Fig. S7**).

To confirm the importance of Arg6 and Arg20 residues in 23S rRNA recognition, individual alanine substitution variants were created, and their proper folding was confirmed using nano-differential scanning fluorimetry (nDSF; **Fig. S8**). Next, enzyme activity was assessed in two complementary methyltransferase activity assays: quantification of 50S subunit ^{3}H incorporation following transfer of a [^{3}H]-methyl group from radiolabeled SAM cosubstrate ([^{3}H]-SAM) and direct visualization of C2144 2'-O-methylation via reverse transcription (RT) primer extension. For the [^{3}H]-SAM assay, we first established optimal conditions using wild-type TlyA and then compared these and all other variants in a single time-point assay under conditions corresponding to ~90% completion of 50S subunit methylation for the wild-type enzyme (**Fig. S9**). Consistent with an essential role in specific 50S subunit substrate recognition, individual substitution of either Arg6 or Arg20 completely eliminated methyltransferase activity (**Fig. 3D**). This result was corroborated in the RT assay in which no methylation above background at C2144 was observed for either TlyA R6A or R20A (**Fig. 3E**). In contrast, TlyA R4A exhibited some activity in the [^{3}H]-SAM assay and more robust methylation via primer extension, suggesting this residue makes a smaller contribution to 50S subunit binding by TlyA, as previously noted (17). While the reason for the difference between the two assays in R4A variant

activity is not immediately obvious, the RT assay is less readily amenable to accurately assessing complete methylation for wild-type TlyA and thus for quantitative comparison with variants.

Four other residues surrounding Arg20 are also positioned to make interactions with H69 nucleotides A2146, U2147 and C2148, as well as the TlyA-bound SAM analog. Arg18 is adjacent to one non-bridging oxygen of the phosphate group of U2147, while Gln21 is located between the second non-bridging oxygen of the same phosphate group, the base O4 atom of U2147, and the terminal carboxyl group of the bound SAM analog (**Fig. 3F**; **Fig. S6B**). Ser19 is also located between these residues and the phosphate of C2148, with Gln22 in a central location within 3-4 Å of all three other TlyA residues as well as the U2147 phosphate group. As before, these residues were substituted with alanine (R18A and S19A) or as a double change with a more conservative asparagine substitution at both glutamine residues (Q21N/Q22N) and assessed in the two activity assays after confirming their correct folding (**Fig. 3D,E**; **Fig. S8B,D**). Consistent with their more modest conservation among TlyA homologs compared to Arg6 and Arg20, all three variant proteins were affected by the amino acid substitution but retained some activity in both assays (**Fig. 3C-E**; **Fig. S7**). These results suggest these residues play supporting, but not individually critical, roles in TlyA substrate recognition

Finally, within the TlyA NTD, the highly conserved Lys41 is positioned to make a single electrostatic interaction with the phosphate group of C2055 which is in a bulge loop at the base of H68, on the strand complementary to the 23S rRNA sequence preceding H69 (**Fig. 3C,G**; **Fig. S6C**, **Fig. S7**). Again, a reduction in both activity assays was observed suggesting an important, but not individually critical, role in 50S subunit binding for Lys41 (**Fig. 3D,E**). Thus, these analyses have identified the NTD residues that contribute collectively to 50S subunit interaction (Arg4, Arg18, Ser19, Gln21, Gln22, and Lys41), including two, Arg6 and Arg20, whose coordinated recognition of a complex 23S rRNA structure adjacent to H69 is critical for specific substrate recognition by TlyA.

TlyA CTD interactions with H69 position the methyltransferase domain for C2144 modification– The TlyA CTD makes extensive contact with the irregular minor groove of H69, from the U2132:A2146 pair near the base to the tip of the helix, with two positive patches on either side of the SAM binding pocket extending the NTD contact surface on the rRNA (**Fig. 2**). Five TlyA residues of moderate to very high conservation are positioned to interact with H69: Arg65, Tyr115, Arg133, Arg137, and Lys189 (**Fig. 4A-D**; **Fig. S6D-F**). Two additional residues at the C2144 target nucleotide, Phe157 and Ser234, and their role in TlyA activity are described further in the next section. As before, each residue was individually substituted with alanine, and additionally to isoleucine in the case of Tyr115 to specifically probe the requirement for an aromatic side chain at this position. The purified variant proteins were assessed using nDSF (**Fig. S8**) which revealed them to be properly folded with only one potential exception, Y115I, which retained an unfolding temperature (T_i) similar to the wild-type protein but with an inverted profile (**Fig. S8E**).

Towards the base of H69, Arg65 recognizes the phosphate backbone of nucleotides A2146 and U2147 with the guanidine head group positioned beneath the phosphate of A2146 and within electrostatic interaction distance of a non-bridging oxygen of the U2147 phosphate (**Fig. 4B**). On the opposite stand of H69, Tyr115 extends into the minor groove, contacting the G2134 ribose and G2133 ribose and base edge, with its hydroxyl group within hydrogen bonding distance of the G2133 nucleobase N3 atom (**Fig. 4B**). Positioning of Tyr115 to make these interactions also depends upon a local but significant loop reorganization between the free and 50S subunit-bound forms of TlyA (**Fig. S5B**). Arg65 is universally conserved and substitution with alanine completely ablates activity in both assays (**Fig. 4E-G**; **Fig. S7**), consistent with a critical role in 50S subunit binding and substrate recognition. Substitution of Tyr115 with either alanine or isoleucine similarly results in fully diminished enzyme activity in both assays (**Fig. 4F-G**). Although Tyr115 is not as highly conserved as Arg65, this position is most commonly aromatic and/ or basic (e.g. tyrosine, histidine or arginine; **Fig. 4E**; **Fig. S7**), suggesting conservation of interactions like those we observe in the structure is essential in other TlyA homologs.

Three basic TlyA residues (Arg133, Arg137, and Lys189) surround the hairpin loop structure at the tip of H69. Arg133 and Arg137 approach the backbone phosphate groups on the minor groove side of nucleotides U2135/A2136 and A2138, respectively, while on the opposite side of H69, Lys189 is positioned alongside the base edges of A2136 and U2141 (**Fig. 4C,D**). Arg137 is highly conserved among TlyA homologs (85-90%), whereas conservation is more modest at the other two positions, though still most commonly a basic residue (~40-77% Arg/Lys; **Fig. 4E**; **Fig S7**). Functional analyses of individual alanine substitution variants at these residues revealed a modest impact on TlyA activity with all three comparably reduced in the [^{3}H]-SAM assay, but only R137A exhibiting significantly diminished activity in the RT assay (**Fig. 4F,G**). While these results suggest that Arg133, Arg137, and Lys189 contribute to H69 binding by TlyA, this may be accomplished through their collective interactions with the tip of the helix.

Together, our structural insights and functional analyses suggest that the TlyA CTD contains at least two residues critical for H69 binding, Arg65 and Tyr115, and several others that collectively recognize features along the length of H69. Further, these residues lie on a contiguous surface with similarly essential NTD residues (Arg6 and Arg20), suggesting coordinated recognition of distinct features of 23S rRNA underpin specific substrate recognition of the 50S subunit by TlyA.

TlyA employs a base flipping mechanism to position C2144 for ribose methylation–Binding of TlyA on the 50S subunit precisely positions the opening to the SAM binding pocket and the TlyA active site directly over the target nucleotide C2144 (**Fig 5A**). As noted earlier, use of NM6 in preparing the 50S-TlyA complex also facilitated capture of the enzyme in a post-catalytic state, with C2144 covalently modified on its 2'-OH and the SAM analog still bound in TlyA's cosubstrate binding pocket. While much

of H69 and the adjacent rRNA junction is structurally unaltered upon TlyA binding, suggesting that the enzyme specifically recognizes the mature 50S subunit, we observe significant local deformations around the target nucleotide in our structure. Most strikingly, C2144 fully flips out from H69, with two TlyA residues, Phe157 and Ser234, positioned to stabilize the altered H69 structure (**Fig. 5B-D**; **Fig. S6G**). Phe157, which is almost universally conserved among TlyA homologs (**Fig. 5C**, **Fig. S7**), stacks on A2143 and partly fills the space normally occupied by C2144. This interaction appears mechanistically critical as removal of the amino acid side chain (F157A substitution) completely abrogates activity (**Fig. 5D-F**). Further, a F157I substitution, which maintains a bulkier side chain but lacks an aromatic nature that would favor stacking on the RNA base, also renders TlyA completely inactive, suggesting that the π-π stacking of aromatic side chain and nucleobase is specifically critical. In contrast, the observed interaction of Ser234 via its hydroxyl group with the NH_2 of C2144 is not essential for activity, consistent with the very low level of conservation at this position (**Fig. S7**). As such, TlyA does not appear to require direct identification of the base identity at C2144 for modification (**Fig. 5D-F**). However, we also note that Ser234 is flanked by two universally conserved glycine residues which likely impart the necessary flexibility in this short loop to intimately sequester the flipped base which may allow some level of discrimination among possible RNA bases.

Insights into 30S subunit recognition and impact of TlyA clinical mutations–With the new structural and functional understanding presented thus far on how TlyA specifically recognizes the 50S subunit for C2144 ribose modification, we addressed two key questions on TlyA's dual substrate specificity and the functional impact of known clinical variants that lead to capreomycin resistance. First, using the collection of TlyA variants already generated, we asked whether the same dependencies on specific NTD and CTD residues also applies to substrate recognition and modification of TlyA's 30S substrate target site (**Fig. 6A**). Most strikingly among the NTD variants, some activity is retained in both the R6A and R20A variants, with the latter exhibiting around 50% activity compared to the wild-type enzyme. This is in stark contrast to modification of the 50S subunit where both amino acid substitutions fully eliminated TlyA activity (**Fig. 3D**). Additionally, the R4A substitution which had more modestly reduced activity on 50S subunit, resulted in an equal reduction in activity to R6A for 30S subunit modification. The activity of CTD variants on 30S subunit also appears to differ with some activity observed for the R65A, Y115A/I, and F157A/I variants which fully abrogated activity on the 50S subunit. However, as for 50S subunit modification, S234 does not appear to play a critical role in substrate recognition. These results suggest that the same molecular surfaces engage with both 50S and 30S subunits, but the critical dependencies on specific residues are distinct for TlyA's interaction with its two substrates.

Clinical resistance to capreomycin can arise through 16S rRNA mutation (27, 28) or via amino acid substitutions in TlyA that eliminate its activity and thus incorporation of the rRNA methylation

required for optimal capreomycin activity (28-31). To directly test whether these TlyA mutations result in enzymes lacking activity due to structural disruption of key interactions with the ribosome subunits, we created additional TlyA variants corresponding to three *Mtb* mutations associated with capreomycin resistance, R14W, A67E and N236K (28-31). Although all three proteins were expressed and soluble, analysis of their folding using nDSF suggested more significant structural perturbations for R14W and N236K (**Fig. S10**), compared to other NTD and CTD variants designed to test specific interactions as described above. Consistent with their role in clinical resistance to capreomycin, all three TlyA variants were unable to methylate either substrate efficiently, with no detectable activity on the 50S subunit for modification of C2144 (**Fig. 6B**).

Discussion

Changes in rRNA modification status can have profound effects on ribosome assembly, function, and sensitivity to ribosome-targeting antibiotics (32). rRNA methylations have been identified which either block antibiotic action or are necessary for optimal drug binding and thus antibacterial activity. In bacteria, these rRNA methylations are incorporated by Class I or Class IV methyltransferases (33, 34), with a single enzyme typically responsible for each individual modification. Exceptions to this strict specificity do exist, such as for TlyA which is capable of incorporating cytidine 2'-O-methyl modifications on both the small and large subunit, within two structurally distinct contexts.

Here, we determined the cryo-EM structure of TlyA bound to the 50S ribosomal subunit, revealing the full-length structure of the enzyme and the detailed molecular mechanism of specific recognition of one of its two ribosomal subunit substrates. To our knowledge, this structure represents the first example of an rRNA 2'-O methyltransferase bound to a bacterial ribosomal subunit. These studies also identified an essential 23S rRNA interaction surface that spans both the NTD and CTD of TlyA and contains a set of residues critical for 50S subunit substrate binding and 2'-O-methylation of C2144. TlyA accomplishes specific 50S subunit recognition via essential interactions of its NTD with a unique tertiary structure at the base of H69 and of the CTD with H69, which position the bound methyl group donor SAM over the target nucleotide. Finally, the nucleobase of C2144 is flipped out of the H69 helical stack, in a conformation stabilized by TlyA Phe157 stacking on the adjacent A2143, placing the 2'-OH of the ribose of C2144 adjacent to SAM and the catalytically important residues of TlyA (22).

The TlyA NTD adopts an S4 ribosomal protein fold that makes critical interactions, primarily by the highly conserved TlyA residues Arg6 and Arg20, with the complex 23S rRNA structure of the junction at the base of H69. In the original report of the S4 structure from *Bacillus stearothermophilus*, the corresponding S4 residues (Arg92 and Arg106 in the *Msm* S4 domain 2; **Fig. S5F,G**) were among several highly conserved basic or aromatic residues proposed to form an extensive rRNA binding surface (35). However, S4 Arg92 and Arg106 do not make extensive interactions with 16S rRNA in the *Msm* 70S ribosome structure and instead these residues appear to play important roles in S4

interdomain interactions (25). In contrast, two other S4-domain containing proteins involved in ribosomal quality control, YabO from *Bacillus subtilis* and the human mitochondrial MTRES1 (36-38), recognize the complex structure at the base of H69 in a manner conserved with TlyA (**Fig. S5H-J**). In particular, the residues corresponding to TlyA Arg6 and Arg20 in the first two α-helices of the YabO and MTRES1 S4 protein folds make essentially the same interactions with the conserved rRNA tertiary structure (**Fig. SJ**). Although MTRES1 appears a little more divergent in sequence, the action of YabO Arg2 (TlyA Arg4) and Arg16 (TlyA Arg20) may also be supported by the conserved Asp4 (TlyA Asp8). Thus, the S4 protein domain thus appears to be a modular unit which has been adopted by diverse proteins for 50S recognition at the base of H69.

Our structure revealed C2144 to be flipped out of H69 in a post-catalytic state captured by use of the SAM analog. Base flipping is a common strategy used by DNA modifying or repair enzymes and has been observed or proposed for other rRNA modifying enzymes (39-42). However, given its relative accessibility in the RNA minor groove and as a common component of all RNA nucleotides, whether this molecular strategy would be employed for ribose 2'-OH methylation was previously less clear. Our structure suggests that base flipping may be a common mechanistic feature regardless of RNA methylation target site. Such a strategy would provide an opportunity to probe the base identity for specific target site recognition and could optimally adjust the 2'-OH geometry for methylation. Ribose methylation can influence sugar pucker and base flipping (43), and may be an important element of the modification process itself.

Comparisons to the well characterized 16S rRNA (m^1A1408; *E. coli* numbering) aminoglycoside-resistance Class I methyltransferases, such as NpmA (42), are also particularly intriguing. NpmA also uses a base-flipping mechanism despite the N1 atom being relatively exposed on the helix 44 surface. Like TlyA, NpmA relies heavily on recognition of complex rRNA structure, distant from the site of modification, to accomplish specific binding to its substrate (42). Further, a single direct base edge contact is made by NpmA to A1408, but the absolute importance of this interaction is unclear given that NpmA retains partial activity against ribosomes with a G1408 nucleotide (44). Similarly, TlyA contacts the nucleobase of C2144 via Ser234 located in the loop linking the sixth and seventh β-strands (β6/7 linker) of its Class I methyltransferase core fold, a region commonly associated with substrate recognition by these enzymes (34, 45). As for NpmA, this specific base contact does not appear critical for TlyA activity based on our functional analyses. However, as noted earlier, the high conservation of the surrounding sequence suggests the TlyA β6/7 linker structure may nonetheless be important for forming the pocket shielding C2144 from exposure to solvent in its flipped conformation. One distinction between TlyA and NpmA appears to be how the flipped conformation is stabilized. In NpmA, a basic residue (Arg207) stabilizes a local distortion of the h44 backbone, and the flipped A1408 is stabilized by stacking between two conserved tryptophan residues. Additionally, the vacated space within helix 44 is left unoccupied and NpmA does not contact or stabilize bases on the complementary strand. In

contrast, TlyA uses the conserved Phe157 to occupy the space vacated by flipping of C2144 via stacking on the adjacent A2143 nucleobase. As such, the mechanism used by TlyA is more akin to DNA methyltransferases which replace DNA base pairing and stacking interactions, normally made by the flipped base, with protein-DNA contacts to the base left unpaired within the DNA double helix (46).

The similarities and distinctions between TlyA and NpmA may also be significant for TlyA's mechanism of recognition of its other target nucleotide in the 30S subunit, C1392 (C1409 in *E. coli*), which immediately follows A1408 in the 16S rRNA. Our speculation is that TlyA may exploit the same complex 16S rRNA tertiary surface used by NpmA and related enzymes, as well as the m^7G1405 aminoglycoside-resistance methyltransferases (47). Our analysis of 30S methylation by the NTD and CTD variants of TlyA suggest that the same surfaces containing these altered residues are also broadly engaged in recognition of the 30S subunit. However, some important differences in dependencies on specific key residues for substrate interaction are apparent: whereas Arg20 is essential for 50S modification on C2144, alteration of this residue only minimally impacts 30S methylation. In contrast, 16S rRNA C1392 modification appears to depend more on additional residues at the very N-terminus (e.g. Arg4). This finding is also consistent with insights gleaned from the existence of two TlyA subtypes, $TlyA^{I}$ and $TlyA^{II}$, of which only the longer $TlyA^{II}$ possesses dual substrate specificity and is able to modify the 30S subunit. $TlyA^{I}$ enzymes lack a short sequence at their N-terminus (containing Arg4) and an entire α-helix that follows the seventh core β-strand in $TlyA^{II}$. Thus, consistent with our functional assays and previous alterations of Arg3 and/ or Arg4 (17), critical elements of 30S subunit recognition appear to reside in these regions. Precisely how TlyA adapts to the two structurally distinct target sites remains to be fully elucidated, but we have also previously proposed that structural plasticity in the short interdomain linker in TlyA may be a mechanism by which the enzyme could accomplish this (21). In further support of this idea, a known clinical capreomycin resistance mutation resulting in an A67E substitution in TlyA (28), which we found to inactivate the enzyme, would likely disrupt the hydrophobic binding pocket which linker residue Trp62 occupies in the 50S subunit-bound TlyA structure. This change, in turn, could prevent SAM binding or correct NTD/ CTD association or interdomain communication. The present work thus reveals common requirements in TlyA for modification of both its substrates, but with some key differences in the residues most critical for individual subunit recognition, and adds support to a mechanism by which TlyA might structurally adapt to these distinct interaction surfaces. However, fully defining the basis of TlyA's dual substrate specificity will require corresponding detailed structural studies of TlyA and its 30S subunit substrate.

Of the two modifications incorporated by TlyA, C2144 methylation most strongly influences the binding of capreomycin (approximately 20 Å away) by a long-range mechanism that is not currently well defined. Comparison of H69 in multiple ribosome structures available in the PDB with and without C2144 modification reveals a small but consistent difference at the tip H69: in unmodified ribosomes, the loop formed by nucleotides A2137, C2138, and U2139 makes a tighter turn than in modified

ribosomes. Our structure now offers a third comparison, with a bulkier modification incorporated but with TlyA still also bound, in which H69 is observed in an structural state between those of ribosomes with unmodified and modified C2144. In the conformation of other unmodified bacterial ribosomes, when H69 is more tightly bent, the bases of A2137 and C2138 are more distant from the capreomycin binding site on the 30S subunit. These observations suggest that modification of C2144 alters the structure of H69 in a manner that changes the position of nucleotides A2137 and C2138, promoting the direct interactions they make with capreomycin.

In addition to the A67E substitution in TlyA noted above, our work offers insight into how capreomycin resistance arises clinically through two other mutations in the gene encoding TlyA. In the TlyA NTD, the mutation resulting in an R14W substitution (28) likely disrupts TlyA NTD folding and its essential contribution to substrate recognition on the 50S subunit. Although it does not directly contact 23S rRNA, Arg14 is positioned directly above Arg6 and Arg20, and interacts with the TlyA backbone at Thr50/Ala51 which are part of a loop that wraps closely around the arginine side chain. Thus, structural changes to accommodate the bulkier tryptophan side chain would disrupt the critical interactions with the rRNA made by Arg6 and Arg20. Another common mutation found in resistant *Mtb* results is a N236K substitution (30) which our structure suggests could impact TlyA activity in several ways. Gln236 immediately follows the β6/7 linker (sequence ^{232}GPSG^{235}) which surrounds the flipped C2144 base. Additionally, this substitution places a lysine residue close to residue Glu238 which has been proposed to play an important role in catalysis (22).

In summary, the present work has revealed the full-length structure of the *Mtb* methyltransferase TlyA and defined the molecular basis for specific recognition of its 50S subunit substrate. While future structural and biochemical studies with the 30S subunit will be necessary for a full understanding of TlyA's dual substrate specificity, these studies have deepened our understanding of rRNA methyltransferase action. In particular, recognition of unusual rRNA structures distant from the site of modification and base flipping both emerge as general themes in substrate molecular recognition for these enzymes.

Materials and Methods

TlyA protein expression, purification and site-directed mutagenesis—An *E. coli* codon-optimized sequence encoding *Mtb* (strain ATCC 25618/H37Rv) TlyA was obtained via chemical synthesis (GeneArt) and subcloned into a pET44a(+) plasmid (pET44-TlyA), as previously described (21). This construct produces TlyA with an N-terminal hexahistidine tag. The TlyA-encoding plasmid was used to transform *E. coli* BL21 (DE3) and cultures were grown at 37 ˚C in Terrific Broth containing 100 μg/mL ampicillin. At mid-log phase (~0.4-0.6 OD_{600}) protein expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside, and growth continued for an additional 3.5 hours. Following harvest via low-speed centrifugation (4,000 x *g*) for 10 minutes at 4 ˚C, the cells were resuspended in lysis buffer (50

mM NaH_2PO_4, pH 8.0, 300 mM NaCl, and 10 mM imidazole containing an EDTA-free SIGMA*FAST*™ Protease Inhibitor Cocktail Tablet) and lysed by sonication (Misonix Sonicator 3000 with microtip: 15 minutes total sonication time, 0.9 s on, 0.6 s off, output level 5.5). Cell lysates were cleared by centrifugation (21,000 x *g*) at 4 ˚C for 40 minutes and filtered before purification of TlyA by sequential Ni^{2+}-affinity (Cytiva HisTrap™ FF crude 1mL or manual His-column using Millipore Ni-NTA His-Bind® Resin) and gel filtration (Cytiva HiLoad™ 16/600 Superdex™ 75) chromatographies on an ÄKTA Purifier 10 system. TlyA variants with single or double amino acid substitutions were created using megaprimer whole-plasmid PCR (48) in pET44-TlyA, and expressed and purified by Ni^{2+}-affinity chromatography as described above for wild-type TlyA. Protein folding and quality control was accomplished using nDSF on a Tycho NT.6 (NanoTemper) which monitors protein thermal unfolding using intrinsic fluorescence at 330 and 350 nm. The unfolding profile and inflection temperature (T_i) of each TlyA variant was determined using the instrument software for comparison to that of wild-type TlyA.

Isolation of Msm 50S and 30S subunits—*Msm* 50S subunits with unmethylated C2144 were isolated from a strain lacking TlyA activity (LR222 C101A) using following established procedures (21, 49). A small culture of Middlebrook 7H9 liquid medium was inoculated with a single colony of *Msm* LR222 C101A and grown overnight at 37 ºC with shaking (100 rpm). Fresh Middlebrook 7H9 medium (0.5-2 L) was inoculated with the overnight culture (1/100 dilution) and the cultures grown for 72 hours at 37 ºC with shaking (100 rpm). Cells were harvested by low-speed centrifugation (4,000 x *g*) for 10 minutes at 4 ˚C and washed (500 mL per L culture) twice with a solution of 10 mM HEPES/KOH (pH 7.6), 10 mM $MgCl_2$, 1 M NH_4Cl, and 6 mM 2-mercaptoethanol, and once with the same solution but with only 0.1 M NH_4Cl. The cells were then resuspended in the same final buffer and lysed using three passages through a French Press. After addition of DNAse I (10U/ ml lysate), the lysate was cleared by centrifuging for 10 and 30 minutes (at 17,300 and 26,900 x *g*, respectively), and the resulting supernatant centrifuged at high speed for 18 hours (277,200 x *g*) to pellet ribosomes. The 70S pellet was resuspended and dialyzed against a solution containing 10 mM HEPES/KOH (pH 7.6), 0.3 mM $MgCl_2$, 100 mM NH_4Cl, and 6 mM 2-mercaptoethanol to split the ribosome subunits. 30S and 50S subunits were then separated by centrifugation (90,200 x *g*) on a 10-30% sucrose gradient for 18 hours at 4 ˚C. The resulting gradient was fractionated using an ÄKTA Purifier 10 system to collect isolated 50S and 30S subunits. Subunits were stabilized by addition of $MgCl_2$ to 10 mM and the solution centrifuged (300,750 x *g*) for 18 hours. The resulting individual subunit pellets were resuspended in a solution of 10 mM HEPES/KOH (pH 7.6), 10 mM $MgCl_2$, 100 mM NH_4Cl, and 6 mM 2-mercaptoethanol, flash frozen, and stored at -80ºC.

Cryo-EM sample preparation, data collection and structure determination—SAM analog NM6 (5'-(diaminobutyric acid)-N-iodoethyl-5'-deoxyadenosine ammoniumhydrochloride) was prepared essentially as previously described (24) and purified by semi-preparative reverse-phase HPLC. A 3.0 µL mixture of purified *Mtb* TlyA, *Msm* 50S subunit, and NM6 (at 0.5 µM, 5 µM, and 10 µM respectively) was applied to glow-discharged Quantifoil Cu R1.2/1.3 300 mesh grids. Grids were blotted at room temperature for 3.0-3.3 s at >90% humidity and frozen in liquid ethane using a CP3 plunger (Gatan). Cryo-EM data (3364 micrographs) were recorded as movies with defocus range of -0.8 to -2.2 μm at 81,000x magnification (1.0691 Å/pixel) on a Titan Krios 300 kV (TEM) with Gatan K3 direct electron detector at the National Center for CryoEM Access and Training (NCCAT). The dose per frame was 1.25 e/Å/frame (total dose of 50.79 $e^-/Å^2$) over a total exposure of 2 s divided over 40 frames (50 ms per frame).

Following the workflow outlined in **Fig. S2**, image alignment and dose-weighting were performed using Motioncor2 (50) and RELION-3.0/3.1 (51) was used for subsequent data processing. The contrast transfer function was estimated using the program Gctf (52). To guide automatic picking, 1094 particles were manually picked and then classified into 2D classes. Automatic picking then selected 1,016,454 particles which were extracted with a box size of 280 Å. Multiple rounds of 2D classifications were performed to remove non-ribosomal particles before 3D refinement using a 60 Å low-pass filtered reference map of the *E. coli* 50S subunit (EMD-3133). Iterative rounds of CTF refinement, 3D refinement, and 3D classification were performed resulting in a 3.05 Å post-processed map (**Fig. S2B***, center*). Analysis of the angular distribution of particles used to generate the map indicated some orientation preference in the data set, but there was good coverage of views containing TlyA (**Fig. S2C**).

Prior to the final post-processing of the complete 50S-TlyA map, multibody refinement was also performed on the remaining particles with separate masks corresponding to TlyA/H69 and the remainder of the 50S subunit, resulting in 3.89 Å and 3.02 Å resolution maps, respectively (**Fig. S2***, right*). The final three maps (complete 50S-TlyA, TlyA/H69, and 50S subunit alone) were then post-processed using Relion resulting in final 3.05 Å, 3.61 Å, and 2.99 Å maps, respectively, based on gold-standard refinement Fourier Shell Correlation (0.143 cutoff) (**Fig. S2, S3**). Local resolution maps were also generated using ResMap 1.1.4 (53).

All three final maps were used for model building and refinement. The 50S subunit model was created by docking an existing *M. smegmatis* 50S subunit structure (PDB code 5O60), after *de novo* modeling of the NM6-modified C2144, into the 50S-TlyA map and using Coot (version 0.9-pre EL, ccpem) (26) and Phenix (version 1.19.2-4158-000) (54, 55). The TlyA model was generated using a TlyA CTD crystal structure (PDB code 5KYG) appended with a homology modeled NTD (21, 22). As initial docking of our hybrid model (21) did not give a satisfactory fit of the NTD into its portion of the map, this ~60 residue domain was manually re-built. The resulting complete full-length TlyA structure

was then used as a search query in the Dali Protein Structure Comparison server (56). This search returned the unpublished structure of a putative hemolysin from *Streptococcus thermophilus* (PDB code 3HP7) as the closest structural homolog which was used to guide further improvement of our TlyA NTD model in regions of the map that were less well resolved. The model was subsequently split into separate TlyA-H69 and remaining 50S subunit models and each separately real-space refined in Phenix (57, 58) using their respective multibody maps (using rigid-body and then to non-rigid body refinement). Finally, the refined models were recombined (without refinement) to create a final complete model of the 50S-TlyA complex and validated using Phenix (54, 55). Complete parameters for data collection and processing, and model building, refinement and validation are summarized in **Table 1**.

RT analysis of 23S rRNA methylation—Extent of methylation of the 50S subunit by wild-type TlyA and mutants was determined using a RT assay. Wild-type or variant *Mtb* TlyA (66 pmol, 2 μM) was incubated for 20 minutes at 37 ˚C with *Msm* 50S subunit (33 pmol, 1 μM) in the presence of SAM (2.1 μM) in 10 mM HEPES-KOH (pH 7.5), 10 mM $MgCl_2$, 50 mM NH_4Cl, and 5 mM 2-mercaptoethanol. The reaction was terminated by phenol/chloroform extraction and the modified rRNA collected by ethanol precipitation. The rRNA modification at C2144 was assessed using an RT primer-extension reaction with a ^{32}P-labeled DNA primer complementary to 23S nucleotides 2188-2204. Modification of the 2'-O was observed only under conditions of complementary (dGTP) depletion (i.e. reactions with 75 μM dATP, dUTP, dCTP; and 0.5 μM dGTP). Controls with no TlyA or no dGTP depletion (i.e. 75 μM dGTP) showed no RT stops corresponding to C2144 ribose modification. Extension products were run on a denaturing (50% urea) 8.6% PAGE sequencing-style gel for 2 hours at 55 W and 50 ˚C. Gels were dried and then imaged using a phosphor storage screen and Typhoon Trio Variable Mode Imager System (GE Healthcare). Extent of modification was estimated by band intensity comparison using ImageQuant TL 1D Version 7.0.

Methyltransferase activity assays using [^{3}H]-SAM—Quantitative extent of methylation of the 50S subunit by wild-type and variant TlyA was determined using a filter-based enzyme assay with ^{3}H-SAM. To establish optimal conditions for comparison to variant proteins, a time-course experiment was performed with wild-type TlyA. TlyA (final concentration 0.76 μM), *Msm* 50S subunits (final concentration 0.38 μM), and ^{3}H-SAM (final concentration 0.8 μM) were added to "Buffer G" (5 mM HEPES-KOH (pH 7.5), 50 mM KCl, 10 mM NH_4Cl, 10 mM MgOAc, and 6 mM 2-mercaptoethanol) to a total reaction volume of 90 μL. The reaction was incubated for 60 minutes at 37 ˚C, with 10 μL aliquots (3.8 pmol 50S subunit) removed and quenched in 140 μL 5% trichloroacetic acid at 0, 1, 2, 5, 10, 20, 40, and 60 minutes. The reaction was then applied to a glass microfiber filter and 50S subunit methylation quantified using scintillation counting of ^{3}H retained on the filter. A 20-minute time point was subsequently selected for comparison of wild-type and variant TlyA proteins using the assay

performed essentially otherwise as described above. Assays using 30S subunit were performed using the same procedures, but with single timepoint measurements taken at 60 minutes as activity was observed to be weaker for this substrate (**Fig. S9B**), as previously noted (17).

Phylogenetic analysis of the TlyA protein family and residue conservation—TlyA homologs were retrieved from InterPro (IPR004538) with conserved sequence feature annotated for Hemolysin A/ rRNA methyltransferase TlyA family. Sequence redundancy was reduced in UniProt using the precalculated UniRef sequence clusters with a cutoff of 50% sequence identity. A total of 223 representative sequences were aligned using Clustal Omega and an unrooted neighbor joining phylogenetic tree was constructed using MEGA X (59) with evolutionary distances computed using the JTT matrix-based method (60). The rate variation among sites was modeled with a gamma distribution (shape parameter = 1) and the residue propensities were calculated using Geneious.

Acknowledgements

We thank Dr. James Posey and colleagues at the Centers for Disease Control and Prevention, Atlanta GA for providing *M. smegmatis* strain LR222 C101A, and Drs. Puneet Juneja and Ricardo Guerrero-Ferreira for assistance with EM data collection and processing. This work was supported by the National Institutes of Health awards R01-AI088025 (GLC and CMD), T32-AI106699 (ZTL and PS), and T32-GM008602 (PS), and the Burroughs Wellcome Fund Investigator in the Pathogenesis of Infectious Disease award (CMD). This study was supported by the Robert P. Apkarian Integrated Electron Microscopy Core (IEMC) at Emory University, which is subsidized by the Emory School of Medicine and Emory College of Arts and Sciences. Some of this work was performed at the National Center for CryoEM Access and Training (NCCAT) and the Simons Electron Microscopy Center located at the New York Structural Biology Center, supported by the NIH Common Fund Transformative High Resolution Cryo-Electron Microscopy program (U24 GM129539), and by grants from the Simons Foundation (SF349247) and NY State Assembly.

References

1. J. Poehlsgaard, S. Douthwaite, The bacterial ribosome as a target for antibiotics. *Nat Rev Microbiol* **3**, 870-881 (2005).
2. D. N. Wilson, Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nat Rev Microbiol* **12**, 35-48 (2014).
3. C. Fyfe, T. H. Grossman, K. Kerstein, J. Sutcliffe, Resistance to Macrolide Antibiotics in Public Health Pathogens. *Cold Spring Harb Perspect Med* **6** (2016).
4. J. I. Wachino, Y. Doi, Y. Arakawa, Aminoglycoside Resistance: Updates with a Focus on Acquired 16S Ribosomal RNA Methyltransferases. *Infect Dis Clin North Am* **34**, 887-902 (2020).

5. K. S. Long, B. Vester, "Resistance to antibiotics in bacteria through modification of nucleosides in 23S ribosomal RNA" in DNA and RNA modification enzymes: structure, mechanism, function and evolution, H. Grosjean, Ed. (Landes Bioscience, Austin, Tex., 2009), pp. 537-549.

6. G. L. Conn, M. Savic, R. Macmaster, "Resistance to antibiotics in bacteria through modification of nucleosides in 16S ribosomal RNA" in DNA and RNA modification enzymes: structure, mechanism, function and evolution, H. Grosjean, Ed. (Landes Bioscience, Austin, Tex., 2009), pp. 524-536.

7. B. Vester, The cfr and cfr-like multiple resistance genes. *Res Microbiol* **169**, 61-66 (2018).

8. D. M. Mikheil, D. C. Shippy, N. M. Eakley, O. E. Okwumabua, A. A. Fadl, Deletion of gene encoding methyltransferase (gidB) confers high-level antimicrobial resistance in Salmonella. *J Antibiot (Tokyo)* **65**, 185-192 (2012).

9. J. Perdigao *et al.*, GidB mutation as a phylogenetic marker for Q1 cluster Mycobacterium tuberculosis isolates and intermediate-level streptomycin resistance determinant in Lisbon, Portugal. *Clin Microbiol Infect* **20**, O278-284 (2014).

10. T. L. Helser, J. E. Davies, J. E. Dahlberg, Change in methylation of 16S ribosomal RNA associated with mutation to kasugamycin resistance in Escherichia coli. *Nat New Biol* **233**, 12-14 (1971).

11. C. E. Maus, B. B. Plikaytis, T. M. Shinnick, Mutation of tlyA confers capreomycin resistance in Mycobacterium tuberculosis. *Antimicrob Agents Chemother* **49**, 571-577 (2005).

12. World Health Organization (2018). Global tuberculosis report 2018. https://apps.who.int/iris/handle/10665/274453

13. R. E. Stanley, G. Blaha, R. L. Grodzicki, M. D. Strickler, T. A. Steitz, The structures of the anti-tuberculosis antibiotics viomycin and capreomycin bound to the 70S ribosome. *Nat Struct Mol Biol* **17**, 289-293 (2010).

14. L. Zhang *et al.*, The structural basis for inhibition of ribosomal translocation by viomycin. *Proc Natl Acad Sci U S A* **117**, 10271-10277 (2020).

15. Y. Lin *et al.*, The antituberculosis antibiotic capreomycin inhibits protein synthesis by disrupting interaction between ribosomal proteins L12 and L10. *Antimicrob Agents Chemother* **58**, 2038-2044 (2014).

16. K. Yang *et al.*, Structural insights into species-specific features of the ribosome from the human pathogen Mycobacterium tuberculosis. *Nucleic Acids Res* **45**, 10884-10894 (2017).

17. T. Monshupanee, S. K. Johansen, A. E. Dahlberg, S. Douthwaite, Capreomycin susceptibility is increased by TlyA-directed 2'-O-methylation on both ribosomal subunits. *Mol Microbiol* **85**, 1194-1203 (2012).

18. S. K. Johansen, C. E. Maus, B. B. Plikaytis, S. Douthwaite, Capreomycin binds across the ribosomal subunit interface using tlyA-encoded 2 '-O-methylations in 16S and 23S rRNAs. *Mol Cell* **23**, 173-182 (2006).

19. A. Salamaszynska-Guz, B. Taciak, A. Kwiatek, D. Klimuszko, The Cj0588 protein is a Campylobacter jejuni RNA methyltransferase. *Biochem Bioph Res Co* **448**, 298-302 (2014).

20. P. Freihofer *et al.*, Nonmutational compensation of the fitness cost of antibiotic resistance in mycobacteria by overexpression of tlyA rRNA methylase. *RNA* **22**, 1836-1843 (2016).

21. M. A. Witek, E. G. Kuiper, E. Minten, E. K. Crispell, G. L. Conn, A Novel Motif for S-Adenosyl-l-methionine Binding by the Ribosomal RNA Methyltransferase TlyA from Mycobacterium tuberculosis. *J Biol Chem* **292**, 1977-1987 (2017).

22. N. E. Arenas *et al.*, Molecular modeling and in silico characterization of Mycobacterium tuberculosis TlyA: possible misannotation of this tubercle bacilli-hemolysin. *BMC Struct Biol* **11**, 16 (2011).

23. V. Stanevich *et al.*, The structural basis for tight control of PP2A methylation and function by LCMT-1. *Mol Cell* **41**, 331-342 (2011).
24. R. L. Weller, S. R. Rajski, Design, synthesis, and preliminary biological evaluation of a DNA methyltransferase-directed alkylating agent. *Chembiochem* **7**, 243-245 (2006).
25. J. Hentschel *et al.*, The complete structure of the Mycobacterium smegmatis 70S ribosome. *Cell Rep* **20**, 149-160 (2017).
26. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).
27. I. B. Olawoye *et al.*, Whole genome sequencing of clinical samples reveals extensively drug resistant tuberculosis (XDR TB) strains from the Beijing lineage in Nigeria, West Africa. *Sci Rep* **11**, 17387 (2021).
28. J. E. Phelan *et al.*, Integrating informatics tools and portable sequencing technology for rapid detection of resistance to anti-tuberculous drugs. *Genome Med* **11**, 41 (2019).
29. Q. Li *et al.*, Mutation and transmission profiles of second-line drug resistance in clinical isolates of drug-resistant Mycobacterium tuberculosis from Hebei Province, China. *Front Microbiol* **10**, 1838 (2019).
30. T. M. Walker *et al.*, A cluster of multidrug-resistant Mycobacterium tuberculosis among patients arriving in Europe from the Horn of Africa: a molecular epidemiological study. *Lancet Infect Dis* **18**, 431-440 (2018).
31. J. Zhao *et al.*, Assessing capreomycin resistance on tlyA deficient and point mutation (G695A) Mycobacterium tuberculosis strains using multi-omics analysis. *Int J Med Microbiol* **309**, 151323 (2019).
32. P. V. G. Sergiev, A. Y.; Prokhorova, I. V.; Sergeeva, O. V.; Osterman, I. A.; Nesterchuk, M. V., "Modifications of ribosomal RNA: from enzymes to function" in Ribosomes: Structure, function, and dynamics*,* M. V. W. Rodnina, W.; Green, R. , Ed. (Springer-Verlag/Wien, Vienna, 2011), pp. 97–110.
33. I. A. Osterman, O. A. Dontsova, P. V. Sergiev, rRNA Methylation and Antibiotic Resistance. *Biochemistry (Mosc)* **85**, 1335-1349 (2020).
34. H. L. Schubert, R. M. Blumenthal, X. Cheng, Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem Sci* **28**, 329-335 (2003).
35. C. Davies, R. B. Gerstner, D. E. Draper, V. Ramakrishnan, S. W. White, The crystal structure of ribosomal protein S4 reveals a two-domain molecule with an extensive RNA-binding surface: one domain shows structural homology to the ETS DNA-binding motif. *EMBO J* **17**, 4545-4558 (1998).
36. S. Filbeck *et al.*, Mimicry of Canonical Translation Elongation Underlies Alanine Tail Synthesis in RQC. *Mol Cell* **81**, 104-114 e106 (2021).
37. C. Crowe-McAuliffe *et al.*, Structural Basis for Bacterial Ribosome-Associated Quality Control by RqcH and RqcP. *Mol Cell* **81**, 115-126 e117 (2021).
38. N. Desai *et al.*, Elongational stalling activates mitoribosome-associated quality control. *Science* **370**, 1105-1110 (2020).
39. C. Hoang *et al.*, Crystal structure of pseudouridine synthase RluA: indirect sequence readout through protein-induced RNA structure. *Mol Cell* **24**, 535-545 (2006).
40. S. Hong, X. Cheng, DNA Base Flipping: A General Mechanism for Writing, Reading, and Erasing DNA Modifications. *Adv Exp Med Biol* **945**, 321-341 (2016).
41. M. Zhao *et al.*, Structural Insights into the Methylation of C1402 in 16S rRNA by Methyltransferase RsmI. *PLoS One* **11**, e0163816 (2016).

42. J. A. Dunkle *et al.*, Molecular recognition and modification of the 30S ribosome by the aminoglycoside-resistance methyltransferase NpmA. *Proc Natl Acad Sci U S A* **111**, 6275-6280 (2014).

43. S. K. Natchiar, A. G. Myasnikov, I. Hazemann, B. P. Klaholz, Visualizing the Role of 2'-OH rRNA Methylations in the Human Ribosome Structure. *Biomolecules* **8** (2018).

44. N. Zelinskaya, M. A. Witek, G. L. Conn, The pathogen-derived aminoglycoside resistance 16S rRNA methyltransferase NpmA possesses dual m1A1408/m1G1408 specificity. *Antimicrob Agents Chemother* **59**, 7862-7865 (2015).

45. K. Vinal, G. L. Conn, Substrate recognition and modification by a pathogen-associated aminoglycoside resistance 16S rRNA methyltransferase. *Antimicrob Agents Chemother* **61** (2017).

46. N. Huang, N. K. Banavali, A. D. MacKerell, Jr., Protein-facilitated base flipping in DNA by cytosine-5-methyltransferase. *Proc Natl Acad Sci U S A* **100**, 68-73 (2003).

47. M. Nosrati *et al.*, Functionally critical residues in the aminoglycoside resistance-associated methyltransferase RmtC play distinct roles in 30S substrate recognition. *J Biol Chem* **294**, 17642-17653 (2019).

48. K. Miyazaki, MEGAWHOP cloning: a method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids. *Methods Enzymol* **498**, 399-406 (2011).

49. A. K. Singh, J. M. Reyrat, Laboratory maintenance of Mycobacterium smegmatis. *Curr Protoc Microbiol* **Chapter 10**, Unit10C.11 (2009).

50. X. Li *et al.*, Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods* **10**, 584-590 (2013).

51. S. H. Scheres, RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-530 (2012).

52. K. Zhang, Gctf: Real-time CTF determination and correction. *J Struct Biol* **193**, 1-12 (2016).

53. A. Kucukelbir, F. J. Sigworth, H. D. Tagare, Quantifying the local resolution of cryo-EM density maps. *Nat Methods* **11**, 63-65 (2014).

54. P. V. Afonine *et al.*, New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr D Struct Biol* **74**, 814-840 (2018).

55. C. J. Williams *et al.*, MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* **27**, 293-315 (2018).

56. L. Holm, DALI and the persistence of protein shape. *Protein Sci* **29**, 128-140 (2020).

57. P. V. Afonine *et al.*, Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531-544 (2018).

58. D. Liebschner *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861-877 (2019).

59. S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* **35**, 1547-1549 (2018).

60. D. T. Jones, W. R. Taylor, J. M. Thornton, The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* **8**, 275-282 (1992).

Tables

Table 1. Cryo-EM data collection, refinement and model validation for the 50S-TlyA complex.

Deposition			
Coordinates (RCSB PDB)	7S0S		
Map (EMDB)	EMD-24792		
Data Collection/ Processing			
Microscope	TFS Titan Krios		
Camera	Gatan K3		
Voltage, kV	300		
Magnification	81,000x		
Electron exposure, $e^-/Å^2$	50.79		
Defocus range	-0.8 to -2.2		
Pixel size, Å	1.069		
Symmetry	C1		
No. particles			
Initial	1,016,454	**Multibody Refinement**	
Final	129,011	**50S subunit**	**TlyA/H69**
Map resolution (FSC 0.143), Å	3.05	2.99	3.61
Refinement and Model			
Model resolution (FSC 0.143), Å	3.0	3.04	3.66
CC_{mask}	0.66	0.84	0.69
Map sharpening, $Å^2$			
50S-TlyA	83.1		
TlyA-H69 (multibody)	109.5		
50S subunit alone (multibody)	72.9		
Non-hydrogen atoms			
Protein residues	3,946	3682	268
RNA residues	3,236	3213	26
Ligand/ modified nucleotide	407	0	1
B factors (min/max/mean), $Å^2$			
Protein	0.21/ 172.4/ 17.1	0.21/ 34.9/ 10.4	50.0/ 172.4/ 114.4
RNA	0.05/ 190.2/ 19.7	0.05/ 88.6/ 19.2	23.1/ 190.2/ 86.7
Ligand/ modified nucleotide	0.78/ 109.5/ 17.3	0.78/ 43.0/ 6.20	109.5/ 109.5/ 109.5
RMS deviations			
Bond lengths, Å	0.008	0.007	0.003
Bond angles, °	0.915	0.808	0.760
Validation			
MolProbity score	1.99	1.90	2.37
Clashscore	9.68	9.81	18.99
Poor rotamers, %	1.21	0.03	0.00
Protein			
Ramachandran plot			
Favored, %	93.79%	94.27	88.35
Allowed, %	6.21%	5.73	10.90
Disallowed, %	0%	0.00	0.75
RNA			
Pucker outliers, %	0%		
Bond outliers, %	0%		
Angle outliers, %	0.1%		

Figures

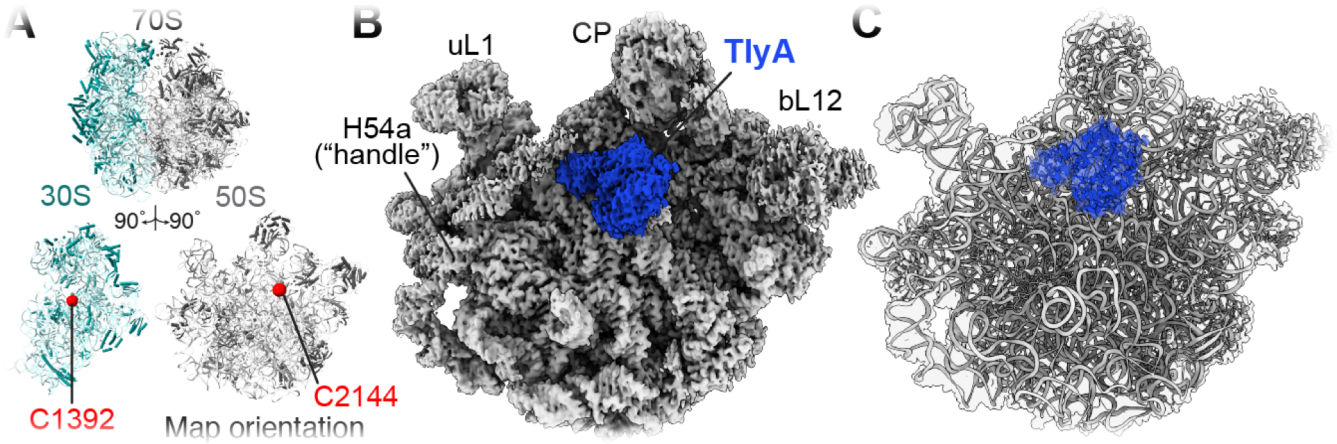


Figure 1. Cryo-EM map at 3.05 Å resolution of the 50S-TlyA complex. ***A,*** Cartoon of the *Msm* 70S ribosome (*top*) and individual 30S and 50S subunits rotated to show their intersubunit interface surfaces (*bottom*), with the nucleotides modified by TlyA indicated (red spheres). ***B,*** Final overall cryo-EM map for the 50S-TlyA complex. TlyA (blue) is bound to the 50S subunit (white) on the subunit interface over H69 containing the modification site (residue C2144). Key 50S subunit features are indicated: uL1 stalk (uL1), central protuberance (CP), the bL12 stalk (bL12) and H54a (also known as the handle). H54a, which extends outward and interacts extensively with 30S subunit in the intact 70S ribosome, was partially observed on the 50S subunit surface in proximity to TlyA, but with weaker map features close to the enzyme. ***C***, Final model of the 50S-TlyA complex shown with semi-transparent map (white behind the model for 50S and blue in front of the model for TlyA).

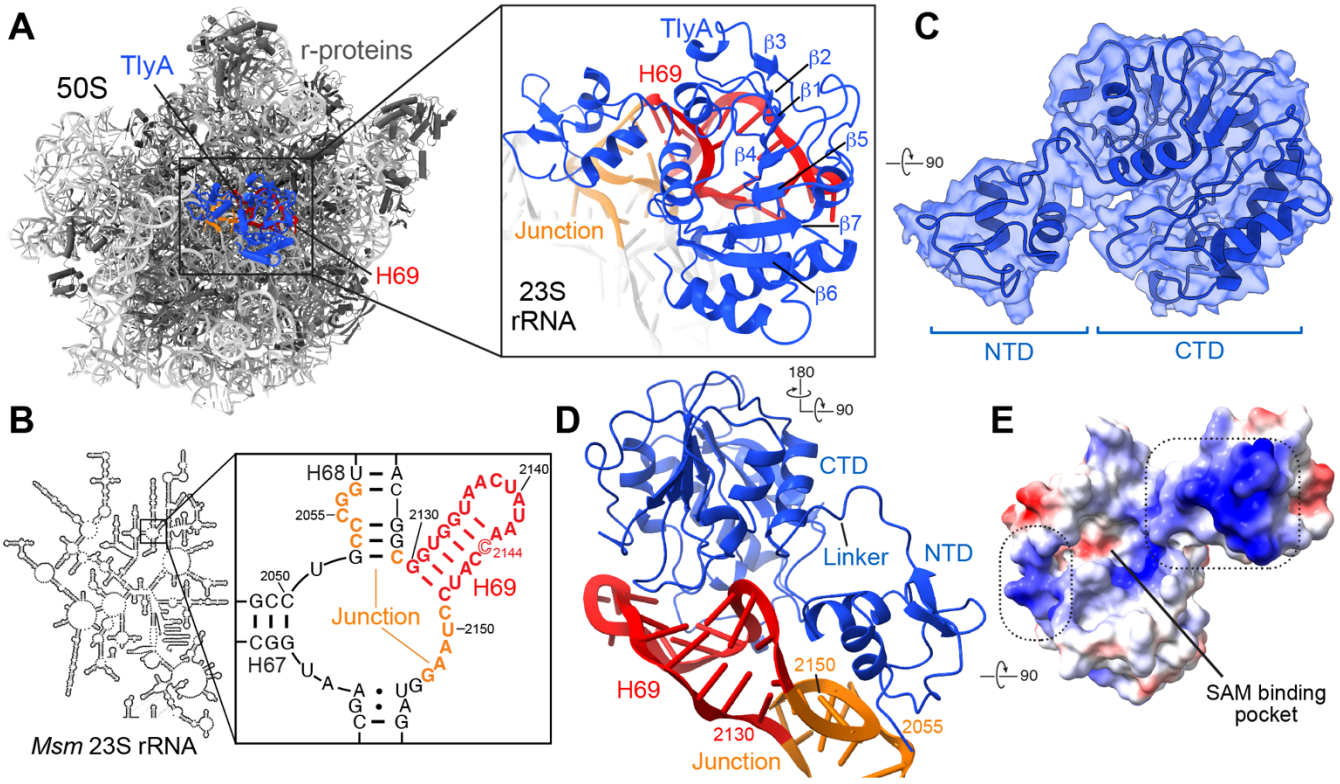


Figure 2. TlyA binds to 23S rRNA H69 and the adjacent rRNA junction via a surface of positively charged residues. *A*, Structure of the 50S-TlyA complex with TlyA (blue cartoon) bound at H69 (red) and adjacent junction (orange) on the 50S subunit interface surface. Ribosomal proteins are shown in dark gray and the remaining 23S rRNA in white. *Inset*, zoomed-in view of TlyA bound to H69 and the adjacent rRNA junction. ***B*,** *Msm* 23S rRNA secondary structure, highlighting the sequence of regions bound by TlyA: H69 and the adjacent junction. ***C***, Modeled structure of full-length TlyA comprising an N-terminal ribosomal protein S4 fold (NTD) and a C-terminal Class I methyltransferase fold with a seven β-strand core (β1-β7, labeled in *panel A*) surrounded by α-helices. The structure is shown in an orthogonal view to *panel A* with surrounding multibody map. ***D***, The TlyA-H69/junction interaction viewed from the 50S subunit surface. The TlyA NTD binds at the base of H69 and the adjacent 23S rRNA junction, while the TlyA CTD interacts exclusively with H69 nucleotides around the modification site. ***E***, Electrostatic surface of TlyA revealing two main patches of positively charged surface (blue, dashed boxes) along the face of TlyA interacting with the rRNA. The negatively charged (red) cosubstrate-binding pocket is indicated between the two positive patches.

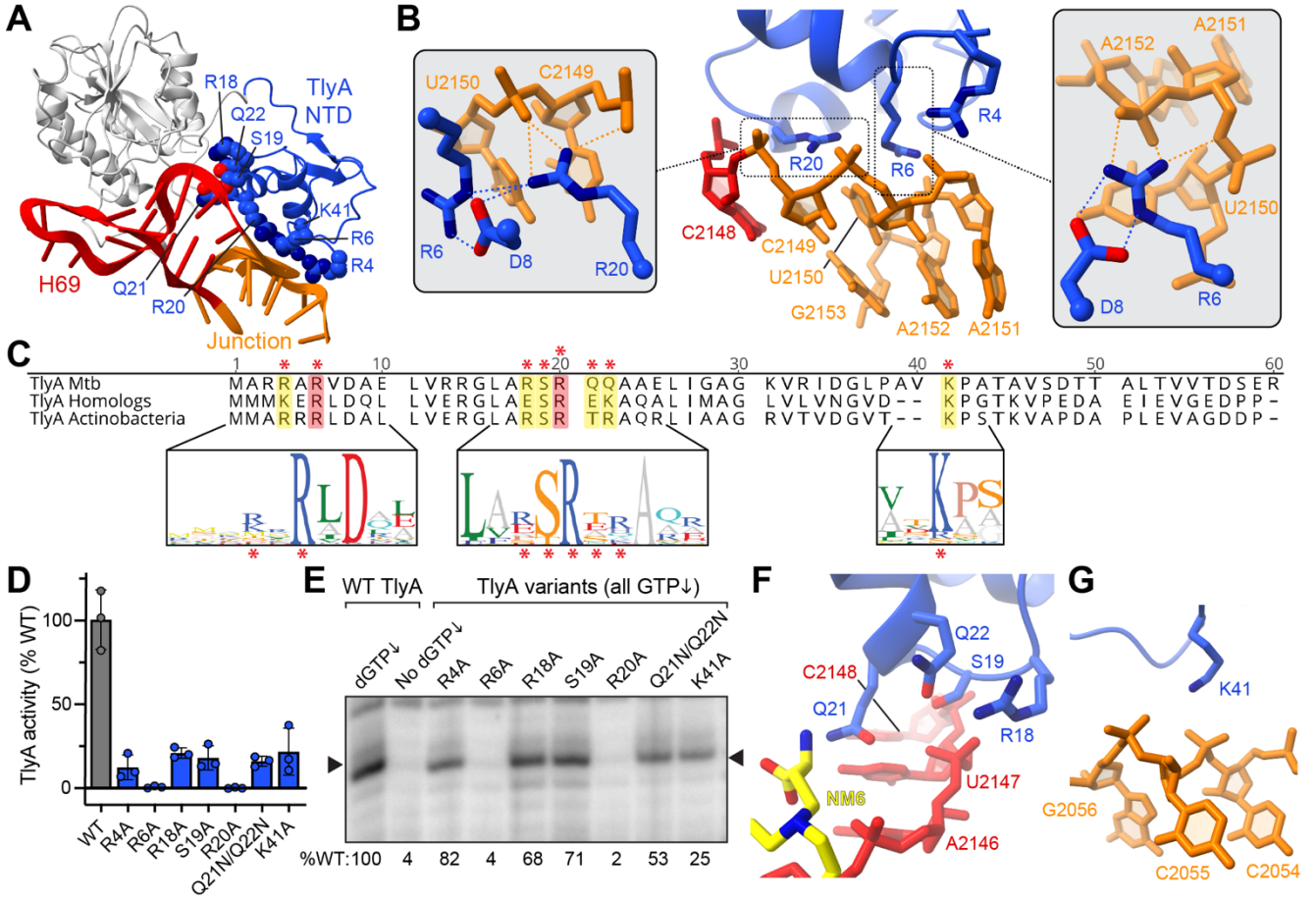


Figure 3. The TlyA NTD recognizes a complex rRNA structure at the base of H69. ***A***, Overview of the TlyA-H69 complex highlighting NTD residues on the TlyA interaction surface for which amino acid substitutions were made. ***B***, Zoomed in view of interactions made by TlyA NTD residues Arg4, Arg6 and Arg20 with nucleotides of the rRNA junction proximal to H69 (orange). Insets: alternate views of Arg6/Arg20 and Arg6 alone with interactions with rRNA and between protein residues indicated with orange and blue dotted lines, respectively. ***C***, Sequence alignment of the *Mtb* TlyA NTD sequence with the consensus sequences for all TlyA homologs and closer homologs from actinobacteria only. Shown below are sequence logo plot representations of sequence conservation for the selected regions among actinobacterial TlyA sequences. The red asterisk denotes sites of amino acid substitutions generated in this work. ***D***, *In vitro* methylation of *Msm* 50S subunit by wild-type (normalized to 100%) and variant TlyA proteins using [^{3}H]-SAM. ***E***, Representative gel showing the results of *Msm* 50S subunit methylation by variant TlyA proteins detected via RT using a radiolabeled DNA primer. Stops at methylated C2144 ribose (arrowhead) are only observed under conditions of depleted dGTP (compare first two lanes with wild-type TlyA). Values below the image are the average band intensity relative to wild-type TlyA for at least two independent experiments. Zoomed views of the interactions made by TlyA NTD residues ***F***, Arg18, Ser19, Gln21 and Gln22, and ***G***, Lys41.

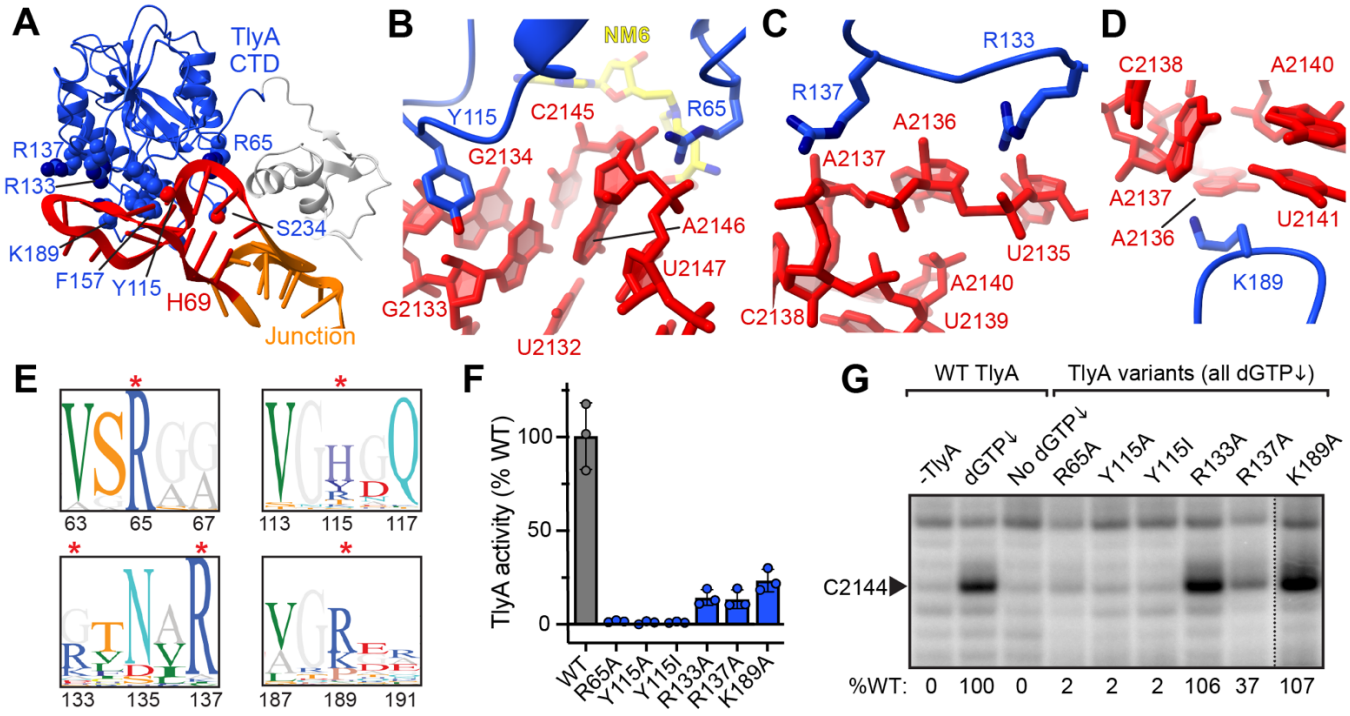


Figure 4. **The TlyA CTD interacts with H69 surrounding the modification site.** ***A***, Overview of the TlyA-H69 complex highlighting CTD residues on the TlyA interaction surface for which amino acid substitutions were made. Zoomed in views of interactions between H69 (red) and TlyA CTD residues ***B*,** Arg65, Tyr115, ***C*,** Arg133, Arg137, and ***D*,** Lys189. ***E*,** Sequence logo plot representations of actinobacterial TlyA sequence conservation for regions surrounding the selected CTD residues. The red asterisk denotes sites of amino acid substitutions made in this work. ***F*,** *In vitro* methylation of *Msm* 50S subunit using [^{3}H]-SAM for wild-type TlyA and CTD variant proteins. ***G*,** Representative gel showing the results of RT analysis of *Msm* 50S subunit methylation by TlyA CTD variants. Values below the image are the average band intensity relative to wild-type TlyA for at least two independent experiments.

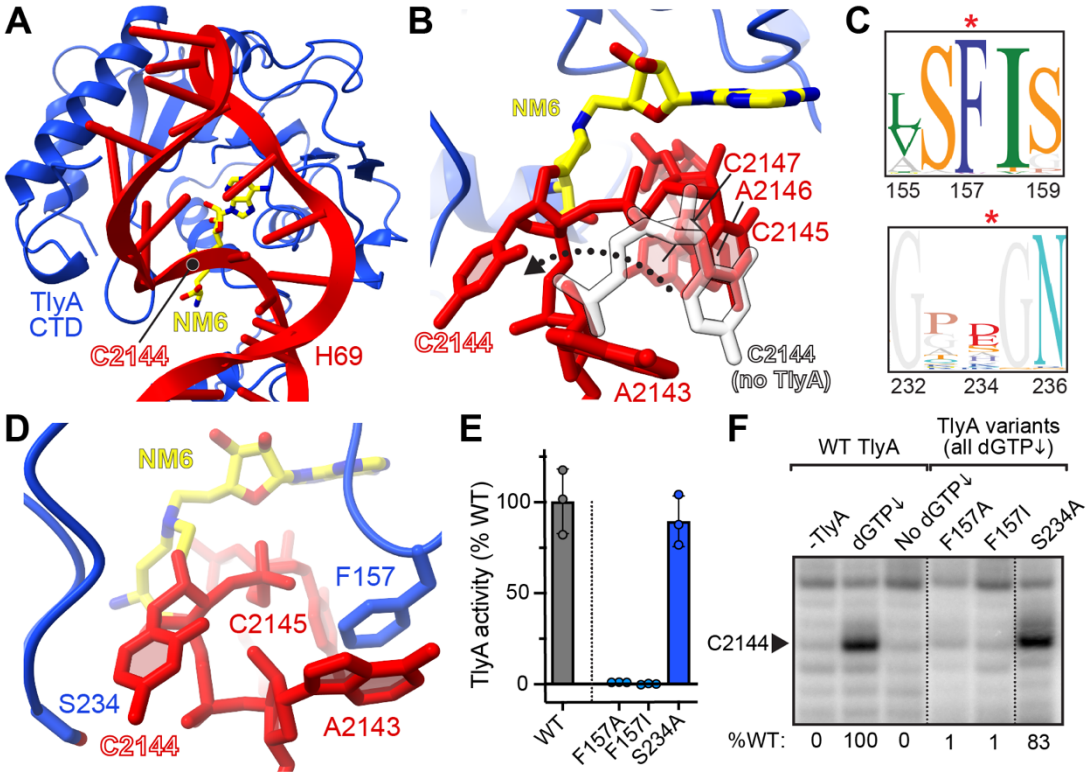


Figure 5. **TlyA uses a base flipping mechanism to position C2144 for 2'-OH modification**. ***A*,** View of TlyA CTD and NM6 cosubstrate positioned over the H69 modification site. ***B*,** The NM6-modified C2144 nucleotide is flipped from H69 compared to its original position (white, semi-transparent sticks). ***C*,** Sequence logo plot representations of actinobacterial TlyA sequence conservation for regions surrounding the selected CTD residues proximal to the flipped nucleotide. The red asterisk denotes sites of amino acid substitutions generated in this work. ***D***, TlyA CTD residues Phe157 and Ser234 interact with H69 nucleotide C2144 and A2143, respectively. ***E*,** *In vitro* methylation of *Msm* 50S subunit using [^{3}H]-SAM for wild-type TlyA and indicated CTD variant proteins. ***F*,** Representative gel showing the results of RT analysis of *Msm* 50S subunit methylation by these TlyA CTD variants. Values below the image are the average band intensity relative to wild-type TlyA for at least two independent experiments. In *panels E* and *F*, data for wild-type TlyA are the same as in **Fig. 4** (dotted lines denote regions removed from the original images).

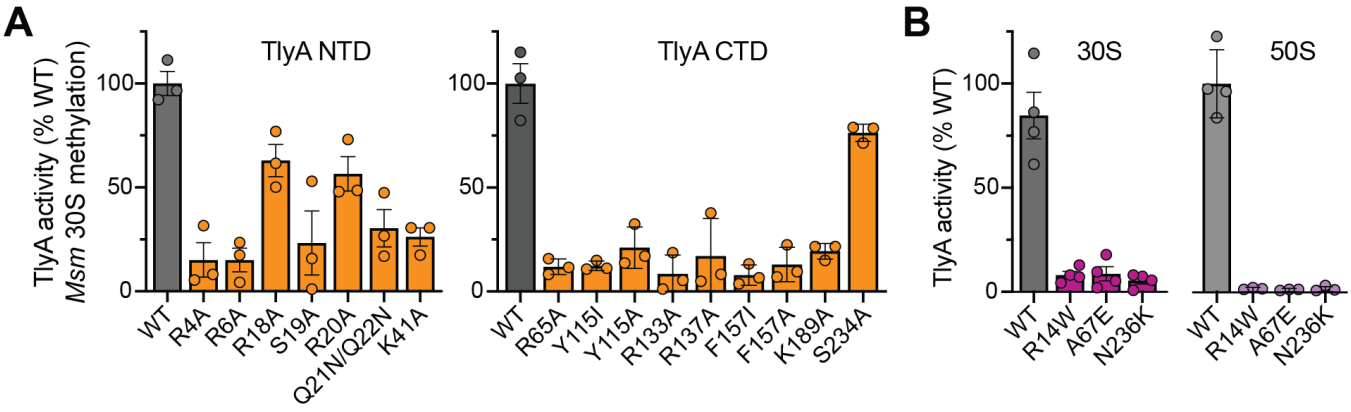


Figure 6. TlyA has distinct residue dependencies for 30S methylation and is inactivated by clinically-identified resistance mutations. ***A***, *In vitro* methylation of *Msm* 30S subunit by wild-type TlyA (normalized to 100%) and the indicated NTD (*left*) and CTD (*right*) variant proteins using [^{3}H]-SAM. ***B***, As in *panel A* but for *Msm* 30S (*left*) or 50S (*right*) subunit with the three indicated TlyA variants associated with clinical resistance to capreomycin.

PNAS
www.pnas.org

Supplementary Information for

50S subunit recognition and modification by the *Mycobacterium tuberculosis* ribosomal RNA methyltransferase TlyA

Zane T. Laughlin[1,2], Suparno Nandi[1], Debayan Dey[1], Natalia Zelinskaya[1], Marta A. Witek[1], Pooja Srinivas[1,3], Ha An Nguyen[1,4], Emily G. Kuiper[1], Lindsay R. Comstock[5], Christine M. Dunham[1,6] and Graeme L. Conn[1,6]*

*To whom correspondence should be addressed: gconn@emory.edu.

This PDF file includes:

Figures S1 to 10

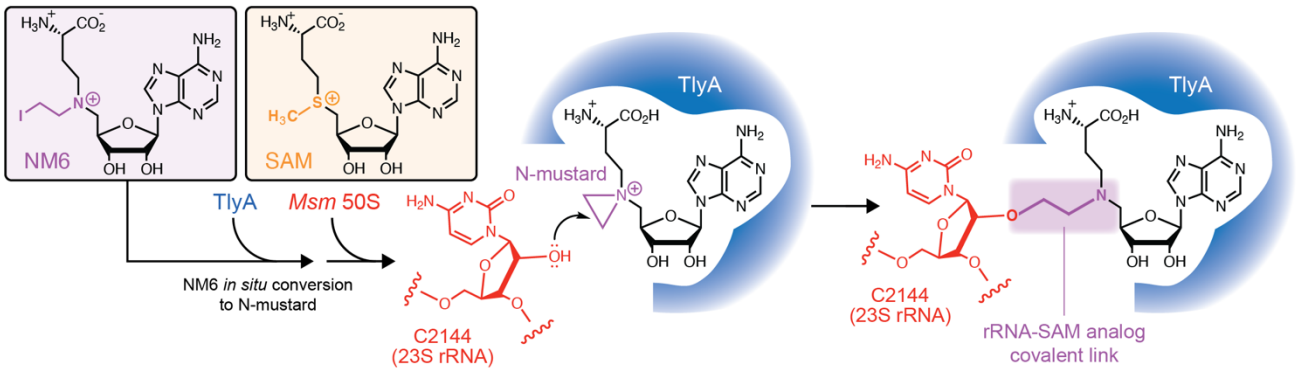


Fig. S1. **Stabilization of the 50S-TlyA complex in a post-catalytic state using a SAM analog (NM6).** N-mustard 6 (NM6; *left box*) is an analog of S-adenosyl-L-methionine (SAM; *right box*) that TlyA can use as a cosubstrate for modification of 23S rRNA nucleotide C2144. In this enzymatic reaction, NM6 becomes covalently attached to the ribose 2' position of C2144. The 50S-TlyA complex is thus stabilized in a state immediately following catalysis by virtue of TlyA's affinity for both 50S subunit and the cosubstrate analog covalently attached to C2144.

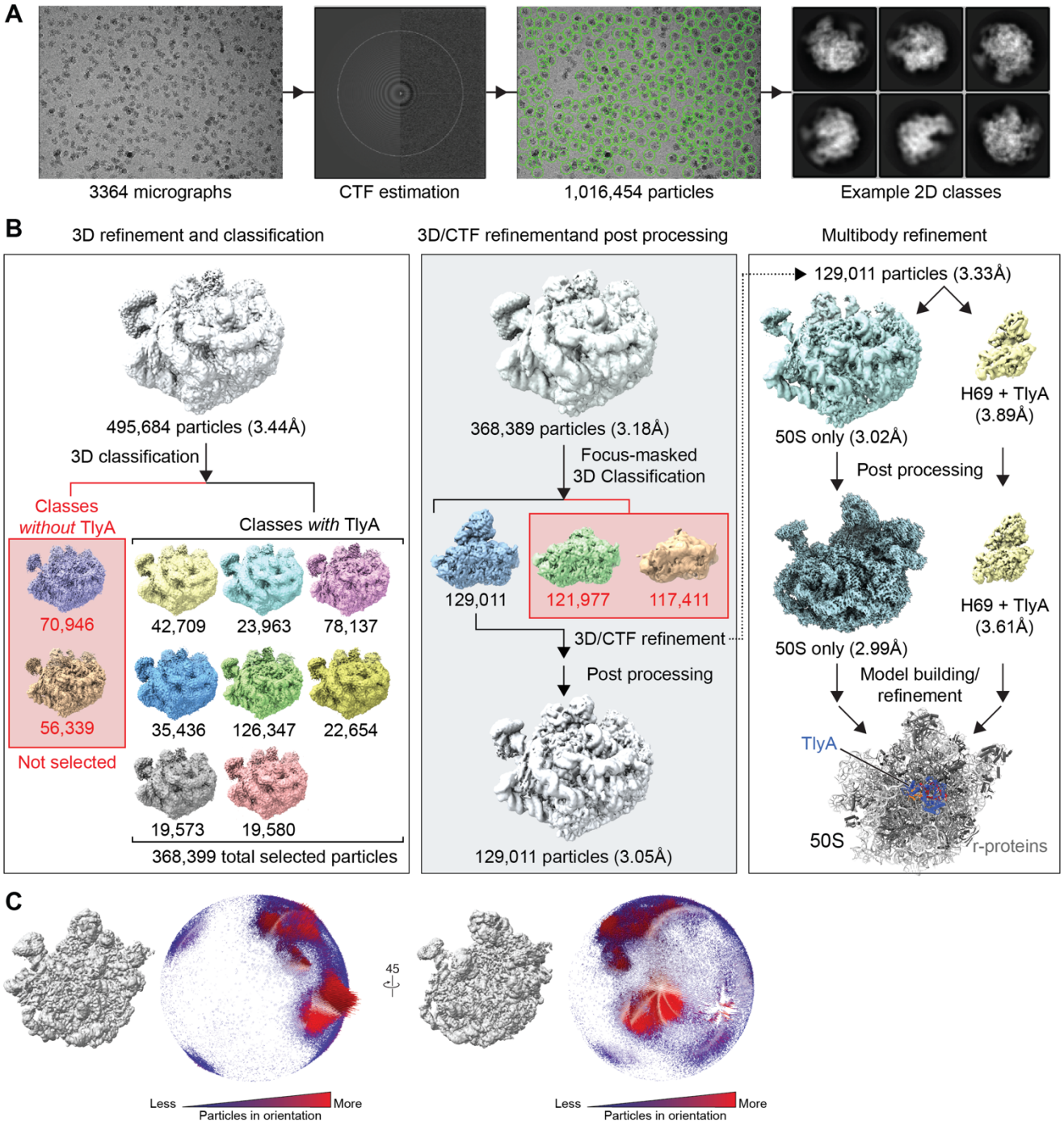


Fig. S2. **Workflow for cryo-EM structure determination.** ***A***, Example micrograph, CTF estimation, particle picking, and final 2D classes. ***B***, The 3D refinement workflow leading to the final 50S-TlyA structure. Following initial 3D refinement from 2D classifications (*upper left*), a 3D classification was used to separate classes lacking TlyA (*lower left*). Following a series of subsequent refinement steps (*upper center*), a focus-masked 3D classification was performed to select only high-resolution map for TlyA and the corresponding particles used in a final refinement and post-processing to produce a final map for the 50S-TlyA complex (*lower center*). In addition, using the same map prior to post-processing, a multibody refinement (separating TlyA/ H69 and the remainder of the 50S) was performed to further boost the map resolution in the region of interest (*upper right*). The resulting maps were post-processed, and used for model building and individual refinement in Phenix, before models were combined to generate the final structure (*lower right*). ***C***, Angular distribution plots of particles contributing to the final map. The final map following focus-masked 3D classification (129,011 particles) and corresponding angular distribution plot have the same orientation in both views; the left view is the same as that shown in other figures. The

height (low to high) and color (blue to red) of the cylinder bars is proportional to the number of particles in those views.

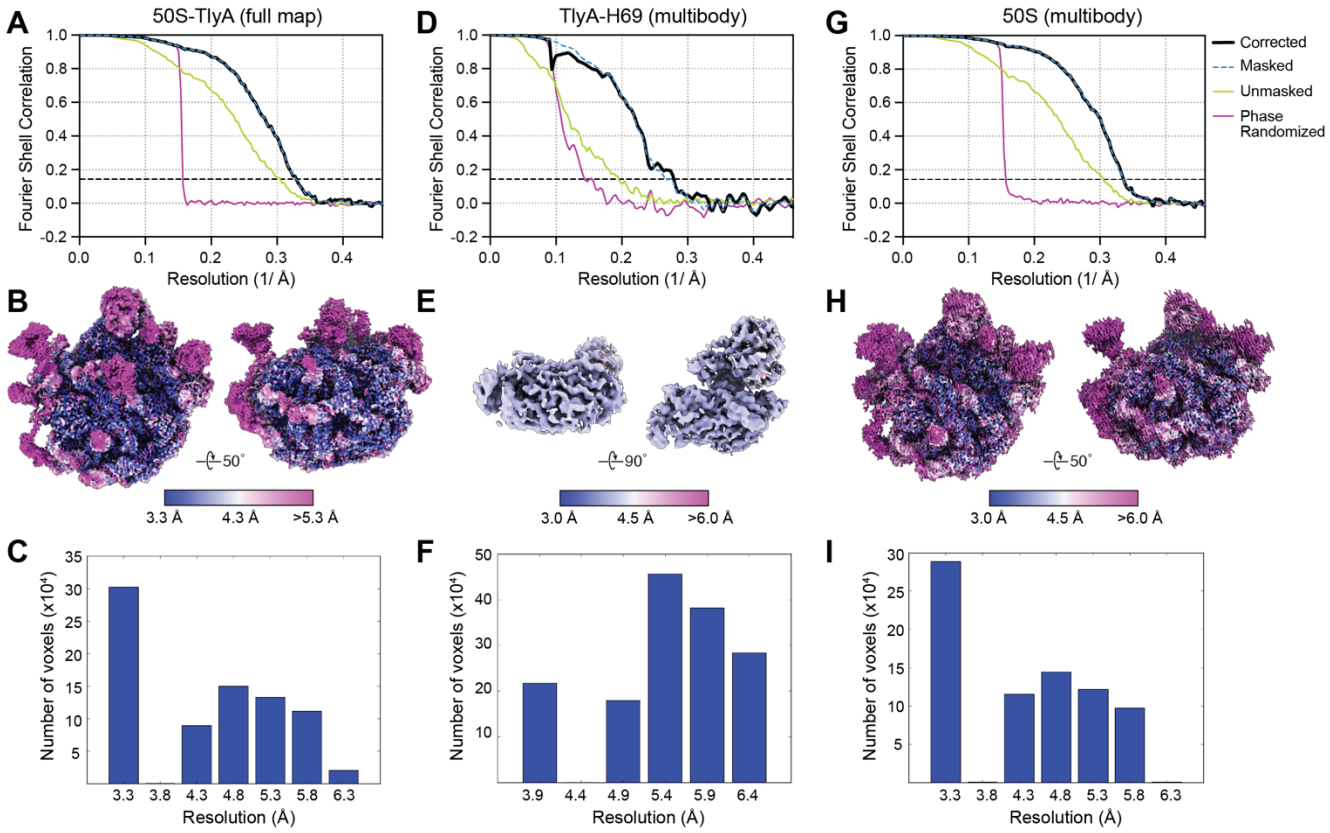


Fig. S3. Cryo-EM map resolution analysis. ***A***, Fourier Shell Correlation (FSC) curves for the full 50S-TlyA map: corrected (black), masked data (blue dashed line; under black line where not visible), unmasked data (green), and phase randomized data (purple). Map resolution was calculated from the 0.143 FSC intercept of the corrected data. ***B***, The local-resolution map of the 50S-TlyA complex with resolution indicated as noted in the scale bar. ***C***, Histogram of resolution distribution across the full 50S-TlyA map. Resolution is grouped in 0.5 Å intervals. ***D-F***, as *panels A-C* but for TlyA-H69 multibody refinement. ***G-I***, as *panels A-C* but for 50S-only multibody refinement.

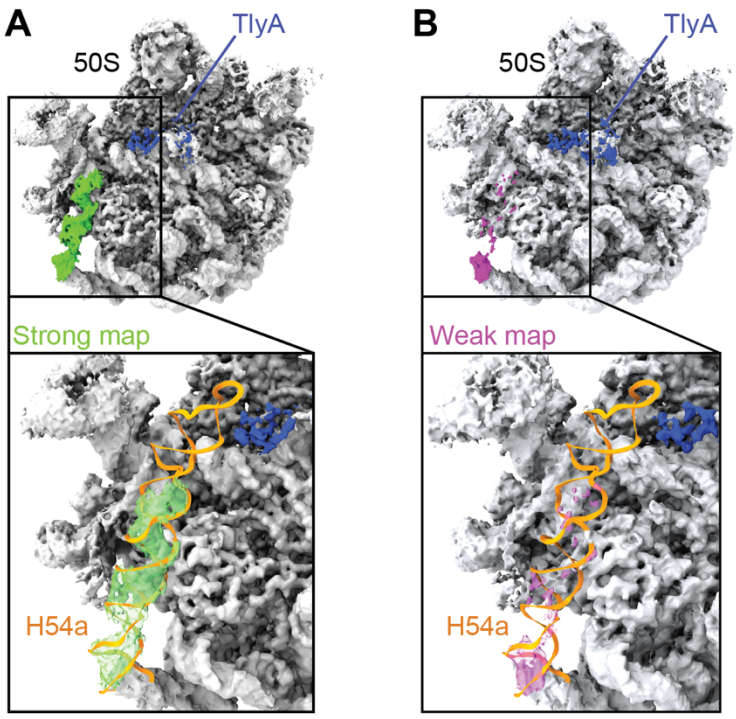


Fig. S4. Map corresponding to H54a differs among 3D classes. *A*, Map corresponding to 3D reconstruction with H54a (the Handle; green map) visible on the subunit interface surface of the 50S subunit (grey), placing it close to the TlyA NTD (blue map). ***B*,** 3D reconstruction with a disordered H54a and weak map features in the same region (magenta map). In both panels, the final model of H54a (orange) is shown for context.

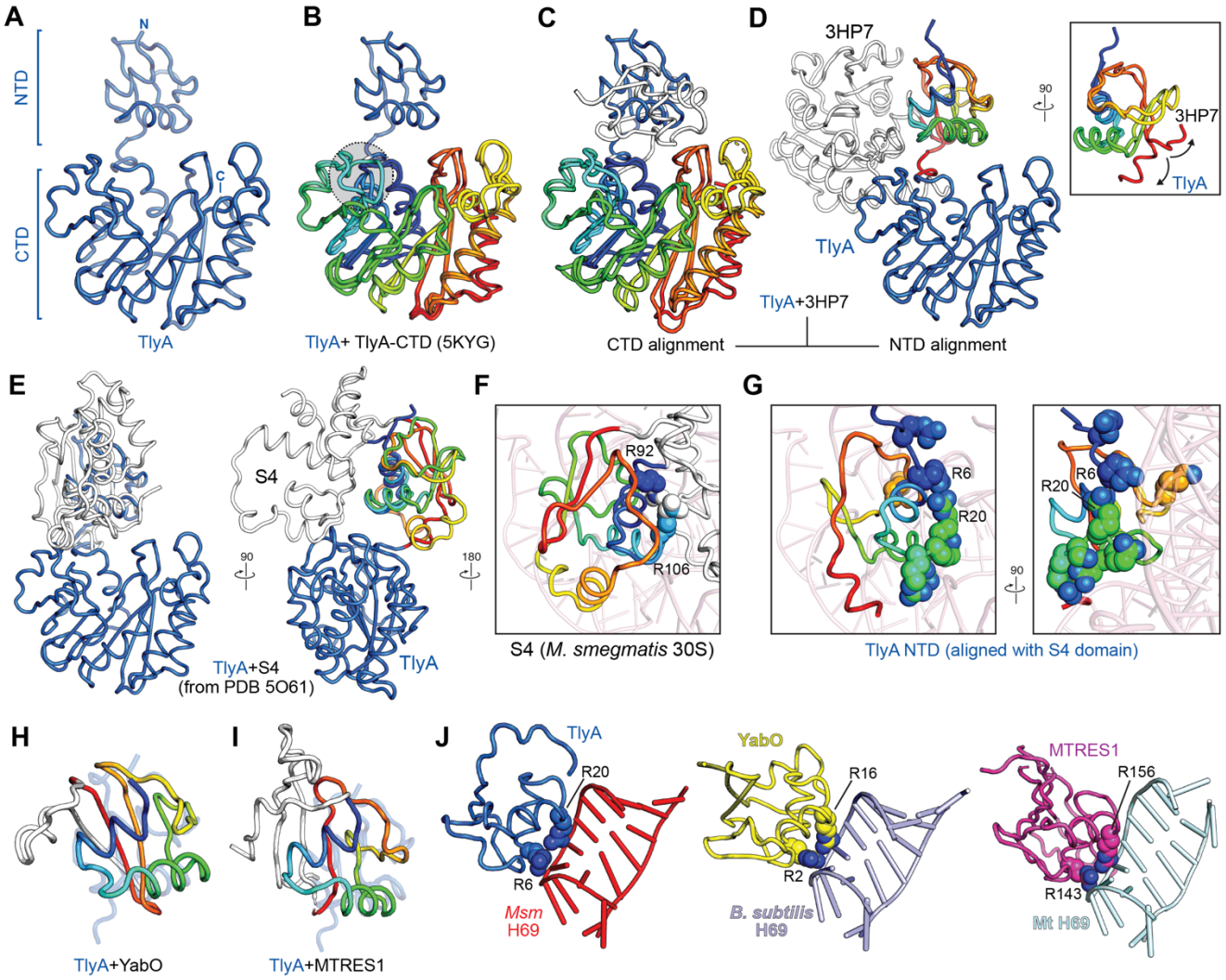


Fig. S5. Superposition of TlyA with structural homologs. ***A***, Structure of TlyA with the CTD and *de novo* modeled NTD indicated. ***B***, Superposition of TlyA with the previously determined TlyA CTD crystal structure (PDB code 5KYG). The regions used for alignment are colored as rainbow from N- (blue) to C-terminus (red). Restructuring of the TlyA loop surrounding residue Tyr115 upon 50S subunit binding is highlighted (circle with gray shading). ***C***, Superposition of TlyA with the putative hemolysin from *S. thermophilus* (PDB code 3HP7) aligned based on the protein CTDs (rainbow coloring). ***D***, As *panel C*, but aligned based on the protein NTDs (rainbow coloring). The inset shows a 90˚ rotated view highlighting the different backbone paths taken at the junction of NTD and CTD in TlyA and the putative hemolysin from *S. thermophilus*, which results in the widely differing relative domain orientations. ***E***, Two views of the superposition of TlyA with *Msm* ribosomal protein S4 (from the structure of the complete *Msm* 70S, PDB code 5O61). Alignment was based on the TlyA NTD and corresponding region in S4 (shown in rainbow coloring on the right image). ***F***, View of S4 on the 16S rRNA with residues corresponding to the functionally critical TlyA Arg6 (S4 Arg92) and Arg20 (S4 Arg106) shown as spheres. ***G***, Two views of the TlyA NTD positioned on 16S rRNA by superposition with S4. The residues of TlyA tested for their potential role in 23S rRNA interaction are shown as spheres, including the two functionally critical residues Arg6 and Arg20. ***H***,***I*** Alignment of the TlyA NTD (semi-transparent blue) with ***H***, *Bacillus subtilis* YabO and ***I***, human mitochondrial MTRES1 shown as rainbow coloring for the S4 fold only. ***J***, Comparison of TlyA, YabO and MTRES1 bound to their respective ribosomal RNA H69; TlyA R6 and R20 and corresponding residues in the other proteins are shown as spheres.

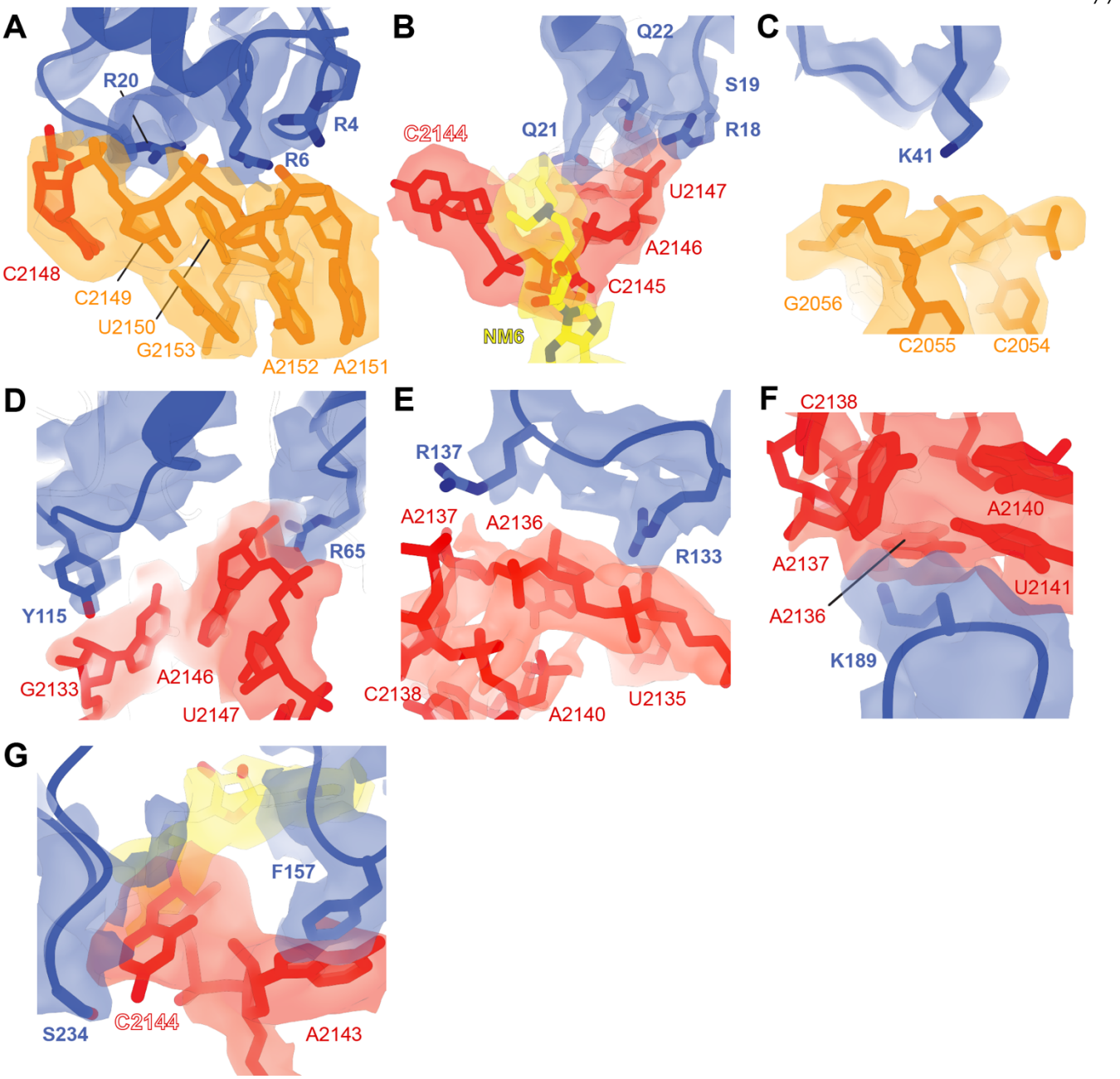
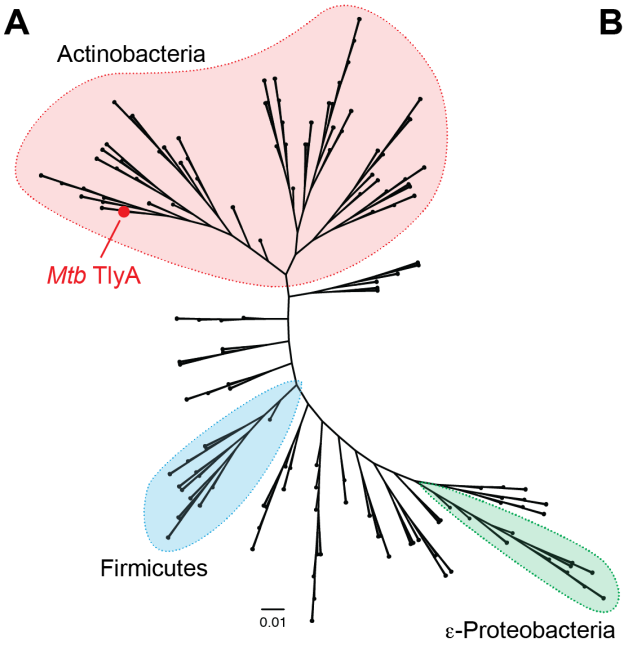


Fig. S6. Interactions of TlyA NTD and CTD residues with 23S rRNA. Views corresponding to main figure panels of the TlyA NTD and CTD interactions with H69 (red) and the adjacent rRNA junction (orange) shown with the final complete 50S-TlyA complex map (EMD-24792): ***A,*** TlyA NTD residues Arg4, Arg6 and Arg20 with nucleotides of the rRNA junction proximal to H69 (also see **Fig. 3B**). ***B***, TlyA NTD residues Arg18, Ser19, Gln21 and Gln22 with H69 and NM6 (yellow) (also see **Fig. 3F**). ***C***, TlyA NTD residue Lys41 with nucleotides of the rRNA junction proximal to H69 (also see **Fig. 3G**). ***D***, TlyA CTD residues Arg65 and Try115 with H69 (also see **Fig. 4B**). ***E***, TlyA CTD residues Arg133 and Arg137 with H69 (also see **Fig. 4C**). ***F***, TlyA CTD residue Lys189 with H69 (also see **Fig. 4D**). ***G***, TlyA CTD residues Phe157 and Ser234 with H69 at the site of C2144 base flipping (also see **Fig. 5D**).



B

TlyA residue	% Conservation (residue identity)	
	All homologs	Actinobacteria
Arg4	74.1 (K/R)	58.8 (K/R)
Arg6	97.3 (R)	100 (R)
Arg18	15.5 (R)/ 35.5 (D/E)	34.0 (R)/ 30.9 (D/E)
Ser19	78.6 (S),15 (T)	85.6 (S), 9.3 (T)
Arg20	90.0 (R)	95.9 (R)
Gln21	14.3 (S), 20.2 (T)	16.5 (S), 30.9 (T)
Gln22	13.0 (Q), 54.7 (K/R)	13.4 (Q), 50.6 (K/R)
Lys41	88.3 (K)	85.4 (K)
Arg65	97.8 (R)	100 (R)
Tyr115	35 (Y), 28.7 (H)	16.5 (Y), 55.7 (H)
Arg133	41.3 (R), 14.3 (K)	30.9 (R), 10.3 (K)
Arg137	84.8 (R)	90.7 (R)
Phe157	97.8 (F)	99 (F)
Lys189	51.6 (R), 23.3 (K)	57.7 (R), 19.6 (K)
Ser234	11.5 (S), 38.7 (D/E)	15.6 (S), 47.9 (D/E)

Fig. S7. Phylogenetic analysis of the TlyA enzyme family and residue conservation. ***A***, Phylogenetic tree of TlyA homologs inferred using the neighbor-joining method, identifying a major Actinobacterial clade (red) which includes the *M. tuberculosis* TlyA, and two smaller clades with TlyA/ hemolysin homologs from the Firmicutes and Proteobacteria. The scale bar denotes the number of amino acid substitutions per site. ***B***, Residue conservation at sites targeted for mutagenesis in this work. The predominant residue(s) as percentage are shown for TlyA proteins among all homologs or in actinobacteria only.

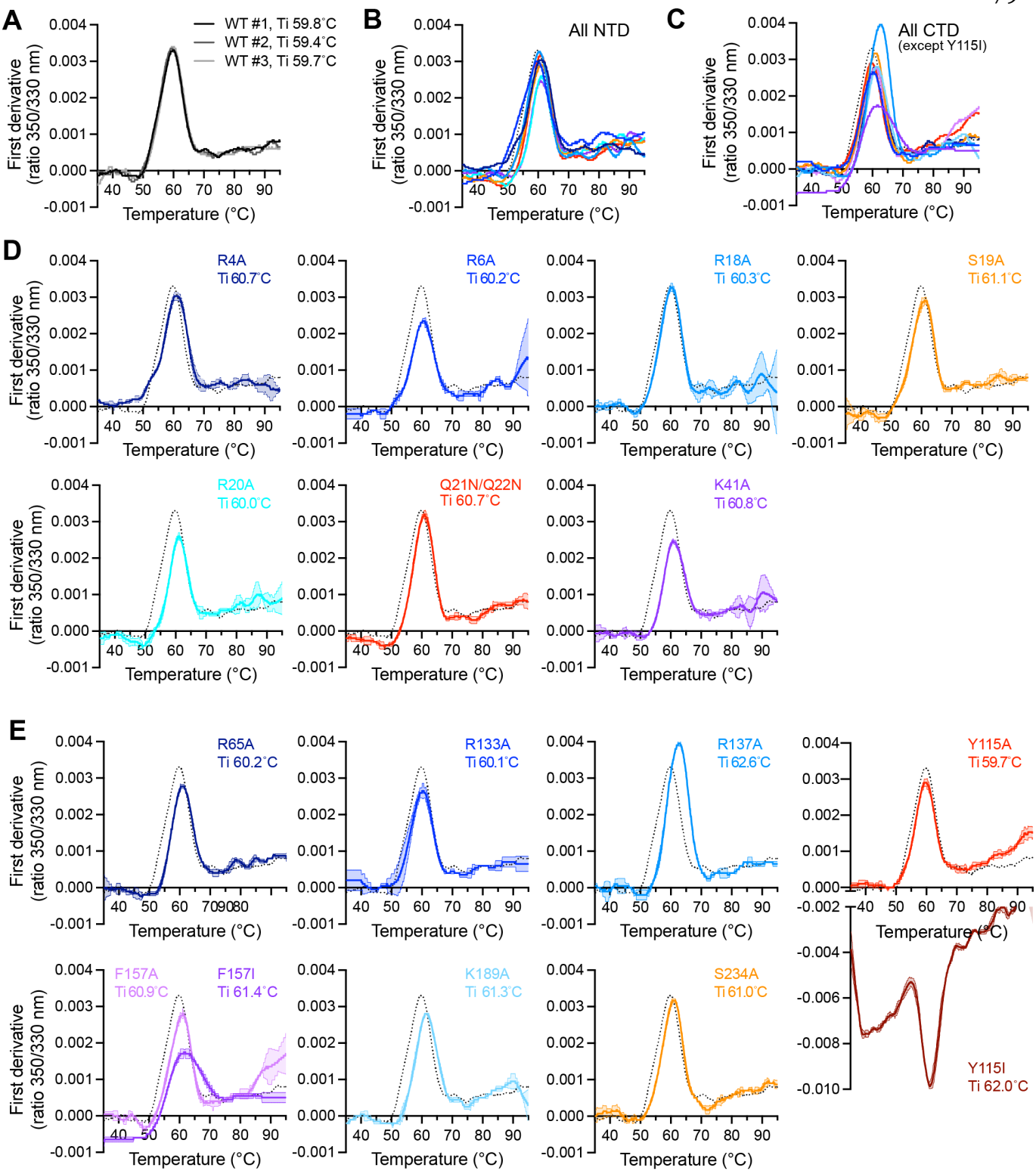


Fig. S8. **Quality control of purified wild-type and variant TlyA proteins by thermal denaturation.** ***A***, Replicate measurements of wild-type (WT) TlyA unfolding monitored using intrinsic fluorescence at 330 and 350 nm (Tycho analysis), illustrating the reproducibility of the method between experiments and protein preparations. First derivative plots are shown for fluorescence ratio (350/330 nm) from which inflection temperatures (T_i) are derived and shown above the plot. Equivalent analysis for: ***B***, all NTD variants, ***C***, all CTD variants (except Y115I), ***D***, individual NTD variants, and ***E***, individual CTD variants. Wild-type TlyA is shown for comparison in all panels (black dotted line; average of all measurements in *panel A*). The shaded regions on individual curves show the standard deviation between measurements.

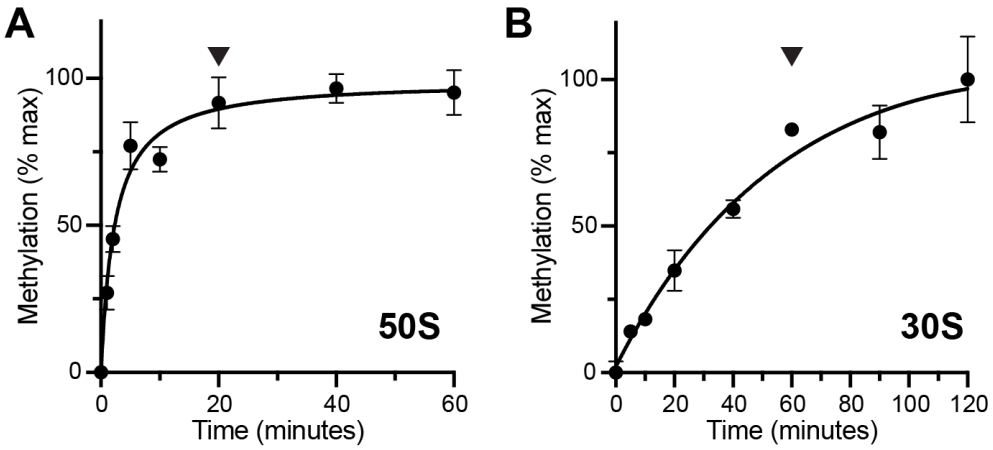


Fig. S9. Optimization of a [^{3}H]-SAM methyltransferase assay for TlyA variants. Time-course activity assays for wild-type TlyA methylation of ***A***, 50S or ***B***, 30S ribosomal subunits isolated from a *Msm* strain lacking TlyA activity. Methylation is ~90% complete at 20 minutes under the conditions used for 50S subunit and this time point (marked with an arrowhead) was used for single-time point activity analyses of all TlyA variants in the [^{3}H]-SAM methyltransferase assay. As 30S subunit is a poorer substrate (see Methods and Methods), a longer time point of 60 minutes (marked with an arrowhead in *panel B*) was used for all analyses of TlyA variants with this substrate.

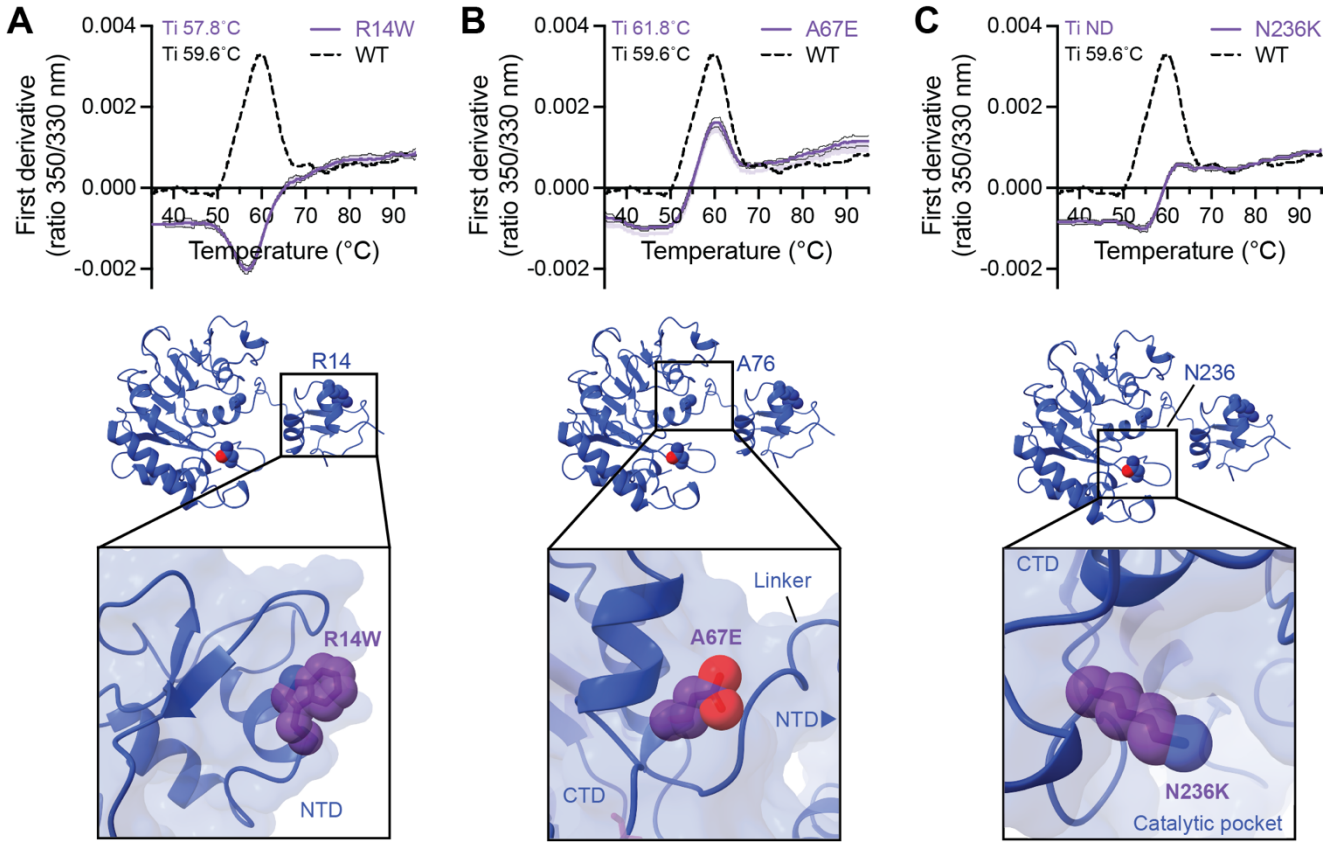


Fig. S10. Analysis of TlyA clinical variant proteins. Replicate nDSF measurements of TlyA variant unfolding for: ***A***, R14W, ***B***, A67E and ***C***, N236K, monitored using intrinsic fluorescence at 330 and 350 nm (Tycho analysis). First derivative plots are shown for fluorescence ratio (350/330 nm) from which inflection temperatures (T_i) are derived and shown above the plot (ND, not determined). Wild-type TlyA is shown for comparison in all panels (black dashed line; same data as shown in **Fig. S8**). The shaded regions on individual curves show the standard deviation between measurements. Below each plot is a view of the TlyA structure and zoomed views highlighting the residue change in each variant (amino acid substitutions were generated in PyMol).

Chapter 4:

TlyA structure, dynamics, and 30S subunit binding

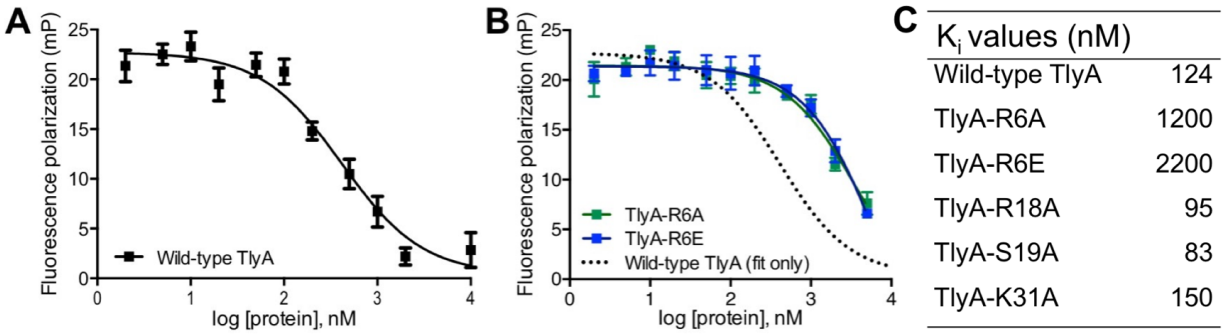
Introduction

Additional preliminary studies were performed to understand various aspects of TlyA's structure and mechanism of action, including the role of protein dynamics in dual substrate recognition. As described in the following sections, this work included adapting a fluorescence polarization (FP) binding assay to quantify TlyA-30S ribosomal subunit interaction, attempts to determine the full-length structure of the free TlyA enzyme via x-ray crystallography, and finally, an analysis of the structural dynamics of TlyA in the presence or absence of cosubstrate using hydrogen-deuterium exchange coupled to mass spectrometry (HDX-MS).

Mechanism of TlyA interaction with 30S subunit

In addition to TlyA *in vitro* activity assays (see Chapter 3), an established competition FP binding assay developed for studies of NpmA (1) and applied to other enzymes in our lab (2) was adapted for studies of TlyA-30S subunit interaction. This assay allowed determination of a binding affinity (as K_i) for TlyA and 30S subunit of 124 nM, comparable to other 16S rRNA methyltransferases (1,2). These FP binding assays were also performed to determine the contribution of individual TlyA residues to binding of the 30S ribosomal subunit. As seen in the TlyA-30S activity assays, Arg6 was found to be an important residue for 30S subunit binding as both amino acid substitutions at this site (to Ala or Glu) significantly reduced the binding affinity of TlyA for the 30S subunit (**Fig. 1**). In contrast, Arg18 and Ser19 appear to contribute less significantly to 30S subunit binding, again consistent with *in vitro* methylation activity (see Chapter 3). These studies set the scene for a detailed analysis of 30S subunit recognition by TlyA

through on-going structural (cryo-EM) and functional analyses, to fully define the enzyme's mechanism of dual substrate recognition.



C

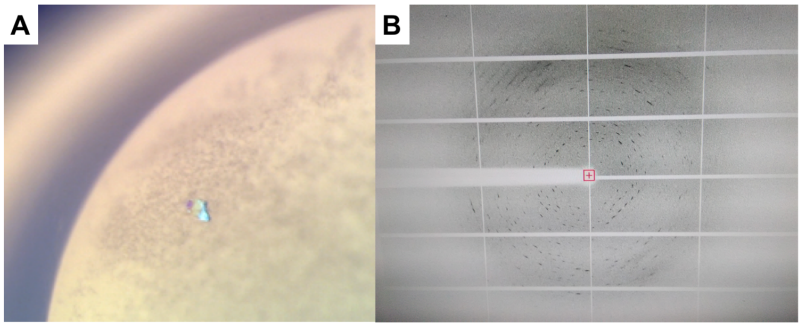
K_i values (nM)	
Wild-type TlyA	124
TlyA-R6A	1200
TlyA-R6E	2200
TlyA-R18A	95
TlyA-S19A	83
TlyA-K31A	150

Figure. 1. TlyA-30S binding assay and analysis of TlyA variants. A. Competition FP binding assay using a fluorescein-labeled NpmA probe (NpmA-E184C*) and *E. coli* 30S subunit at fixed concentrations and varying concentrations of unlabeled wild-type TlyA; the K_i determined from fit in Prism is 124 nM. **B.** Example analysis of variant TlyA proteins for TlyA-R6A and TlyA-R6E. Error bars are SEM in both panels. **C.** K_i values for wild-type and tested TlyA variants for 30S subunit binding.

Towards a structure of *apo*-TlyA

Our 50S-TlyA studies have determined the structure of TlyA bound to the 50S ribosomal subunit, but the structure of full-length TlyA free of a ribosomal substrate is yet to be determined. Determining this structure would reveal into how TlyA's structure changes between the free form and bound to substrate, and could give critical insight into how substrate selection is regulated. Previously, our lab was able to determine the crystal structure of the TlyA CTD plus its short interdomain linker (3) but the NTD remained unsolved. Attempts to determine the crystal structure of full-length TlyA yielded crystals which diffracted to modest resolution but a complete data set could not be obtained due to radiation damage during data collection, and reproducing this crystal form has proved challenging (**Fig. 2**). Therefore, additional crystallization experiments using

these or similar conditions will be required to produce more crystals for future x-ray data collection and high-resolution structure determination.



C

X-ray data collection and processing	
Resolution (Å)	50- 3.10
Space Group	C2
a, b, c (Å)	112.26, 94.167, 55.80
α, β, γ (°)	90.0, 119.51, 90.0
Wavelength (Å)	1.0
Detector	Eiger 16M
Completeness (%)	50
Redundancy	1.4 (1.2)
I/I (sig)	4.48 (1.8)
CC½	0.26 (0.635)

Figure. 2. Crystallization of the free TlyA-SAM complex. A. Crystal of TlyA-SAM grown in 0.095 M sodium citrate tribasic dihyrdrate, pH 5.6, 19% v/v 2-propanol, 19% w/v polyethylene glycol, 5% glycerol at 20°C. **B.** Diffraction image collected at the Southeast Regional Collaborative Access Team (SER-CAT) beamline at the Advanced Photon Source (APS) at Argonne National Laboratory. **C.** Initial x-ray diffraction data information and processing statistics.

Determining the effect of SAM binding on the conformation and dynamics of TlyA.

In earlier studies, our lab discovered that the four-amino acid linker between TlyA's N- and C-terminal domain (NTD and CTD) was critical for SAM-binding (3). Specifically, alteration of residues Trp62 and Val63 decreases binding even though SAM is bound by the Class I methyltransferase structure of the CTD in a site that does not directly contact this linker sequence. The TlyA CTD structure was determined by x-ray crystallography with this short interdomain linker being found to adopt two distinct conformations: one in which the linker was integrated to extend the first α-helix of the CTD and a second where the linker was in a non-helical loop structure (3). This finding suggests that the linker could undergo a conformational change that promotes SAM binding within TlyA,

possibly as a mechanism of communication of correct substrate recognition by the NTD once it is bound to a ribosomal subunit. This mechanism would allow for a common signal to be relayed from the NTD via the interdomain linker to the catalytic center when the NTD engages the two structurally distinct surfaces of the 30S or 50S subunit.

Preliminary HDX-MS studies were performed to compare full-length TlyA with and without SAM bound. These studies revealed that the linker, and specifically Trp62 and Val63, have decreased dynamics when SAM is bound to TlyA, further supporting the idea that a conformational change occurs after binding (**Fig. 3**). Further HDX-MS studies will be required to identify whether the changes in linker dynamics occur due to a direct

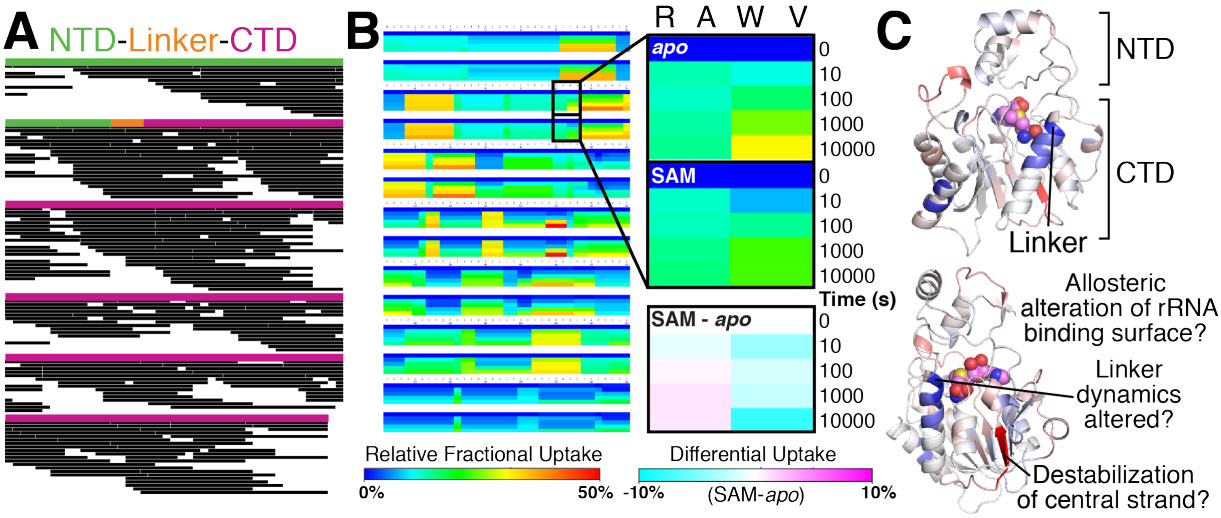


Figure. 3. HDX-MS analysis of TlyA protein dynamics. *A.* Preliminary peptide mapping analysis of TlyA showing 100% coverage with ~12–fold redundancy. These data confirm that the entire sequence of TlyA can be probed for binding-induced changes in conformational dynamics. ***B.*** A preliminary differential HDX-MS analysis of *apo*- and SAM-bound TlyA; the main panel shows the relative exchange in both samples and the zoomed regions show relative and differential (bottom) exchange specifically within the interdomain linker sequence. ***C***, Differential exchange (SAM-bound - *apo*) is mapped on TlyA (NTD homology model/ CTD crystal structure). Preliminary observations from these data are noted.

effect of ligand binding, or, for example, due to altered NTD-CTD interaction (of the type

which might be reversed or further altered when the NTD interacts with a ribosome subunit).

Conclusions

These additional studies set the scene for future studies to define the structure of TlyA in each of its states (free and 30S-subunit bound, in addition to the 50S structure described in Chapter 3), as well as the role of protein conformational dynamics in regulating the activity of the enzyme on its two substrates. How these goals will be accomplished and how this new understanding of TlyA activity can inform more broadly on how ribosomal RNA modification is regulated are described in Chapter 5.

References

1. Vinal, K. and Conn, G.L. (2017) Substrate Recognition and Modification by a Pathogen-Associated Aminoglycoside Resistance 16S rRNA Methyltransferase. *Antimicrob Agents Chemother*, **61**, PMCID: PMC5404524.
2. Nosrati, M., Dey, D., Mehrani, A., Strassler, S.E., Zelinskaya, N., Hoffer, E.D., Stagg, S.M., Dunham, C.M. and Conn, G.L. (2019) Functionally critical residues in the aminoglycoside resistance-associated methyltransferase RmtC play distinct roles in 30S substrate recognition. *J Biol Chem*, **294**, 17642-17653, PMCID: PMC6873201.
3. Witek, M.A., Kuiper, E.G., Minten, E., Crispell, E.K. and Conn, G.L. (2017) A Novel Motif for S-Adenosyl-l-methionine Binding by the Ribosomal RNA Methyltransferase TlyA from Mycobacterium tuberculosis. *J Biol Chem*, **292**, 1977-1987, PMCID: PMC5290967.

Chapter 5:

Discussion

The ribosome is an important target for today's arsenal of antibiotics but ribosome-targeting antibiotics suffer from the same modes of resistance as other antibiotics. For example, resistance to many classes of ribosome-targeting antibiotics can arise from ribosome-methylating resistance enzymes. However, unlike other antibiotics that lose activity when the ribosome is modified, the anti-mycobacterial antibiotic capreomycin requires methylation at two sites on the ribosome for its binding and activity, and loses activity in their absence (1). These methylations on the ribose 2'-OH of 16S rRNA C1409 (*E. coli* numbering) of the 30S ribosomal subunit and 23S rRNA C1920 of the 50S ribosomal subunit were found to be the result of the activity of the intrinsic ribosomal methyltransferase TlyA. While the specific role of these modifications in drug binding are not clear, it was thought that these modifications promote the binding of capreomycin at the intersubunit bridge between the 30S and 50S ribosomal subunits (1,2). Mutations in the gene encoding TlyA or in the rRNA nucleotides at or surrounding the modification sites are commonly observed routes to capreomycin resistance in *Mycobacterium tuberculosis*, the causative agent of tuberculosis (3,4). Because of its ability to modify two distinct substrates and this clinical relevance in tuberactinomycin antibiotic resistance, TlyA is of particular interest for studies that aim to deepen our understanding of rRNA methyltransferase function.

This thesis describes studies focused on the structure of TlyA and its mechanism of recognition and modification of its target sites on the 50S and 30S ribosomal subunits. The new insight generated here will be of particular interest in the field of antibiotic resistance where it could be applied to better understand mechanisms of modification of

other resistance methyltransferases that threaten the clinical usefulness of other important drug classes such as aminoglycosides and macrolides.

Structure of the 50S-TlyA complex and mechanism of modification

The majority of the work on TlyA is presented in Chapter 3. In this study, we present a cryo-EM structure of full-length TlyA bound to the 50S ribosomal subunit with complementary functional assays to probe the role in subunit binding of conserved TlyA amino acids along its interaction surface. These studies confirmed that TlyA's N-terminal domain takes on an S4 ribosomal protein fold and showed that it is critical for substrate binding. In addition to binding SAM and catalyzing the methylation reaction, the C-terminal domain was also found to make essential contributions to substrate recognition and binding. A positively-charged surface that spans both the N- and C-terminal domains recognizes Helix 69 (H69) of the 23S rRNA. Specifically, the N-terminal domain contributes to binding by recognizing the phosphate backbone of a non-A-form helical structure of an rRNA junction at the base of H69. Docking of TlyA to H69 is dictated primarily through interactions with this structure made by residues Arg6 and Arg20, with residues Arg65 and Tyr115 of the C-terminal domain also being critical for recognizing the minor groove face of Helix 69. This study also revealed that TlyA uses a base-flipping mechanism to accomplish ribose 2'-oH modification like other 16S rRNA methyltransferases, which has been proposed for RsmI and demonstrated in the aminoglycoside resistance methyltransferase NpmA (5,6). In TlyA, this mechanism uses the conserved residue Phe157 to stabilize the nucleobase of A1919 when C1920 is flipped from Helix 69. We also performed an analysis on TlyA's ability to bind and recognize the 30S ribosomal subunit and found that binding of this second substrate uses the same residues as in 50S binding but with some critical differences.

Defining the complete 30S-binding surface of TlyA.

Though these studies add to our knowledge of how TlyA modifies it substrates and of rRNA methyltransferase mechanisms more generally, there are still some questions that are left unanswered. To fully understand TlyA's mechanism of dual-substrate recognition and modification, further studies of TlyA and the 30S ribosomal subunit are required. In our studies to date, we determined the critical features of the interaction surface of TlyA with the 50S ribosomal subunit through both the cryo-EM structure of the 50S-TlyA complex and complementary activity assays. To begin addressing the question of how TlyA recognizes its second substrate, we used the same set of TlyA variants to perform activity assays on the 30S subunit. These analyses revealed that residues important for 30S binding overlap extensively with those needed for the 50S (**Fig. 1A,B**). However, some notable differences were also observed. On the N-terminal domain, Arg6 and Arg20 were critically important for 50S subunit binding but for 30S subunit binding, Arg4 and Arg6 are most important. This suggests that the N-terminal-most residues are of relatively greater importance for 30S subunit binding, consistent with the finding that TlyA^{II}, which modifies both ribosomal subunits, has both a longer N- and C-terminal domain than TlyA^{I}, which cannot modify the 30S subunit. The identified critical residues in the C-terminal domain for 50S subunit binding are also important overall for 30S subunit binding, though possibly with a reduced absolutely dependence on a single residue. TlyA variants with substitutions of these residues (particularly Arg65, Tyr115, and Phe157) retained some activity on the 30S subunit whereas modification of the 50S subunit assays was fully abrogated. Overall, this suggests that there is a common binding

surface on TlyA for binding 30S and 50S ribosomal subunits, but with differences in

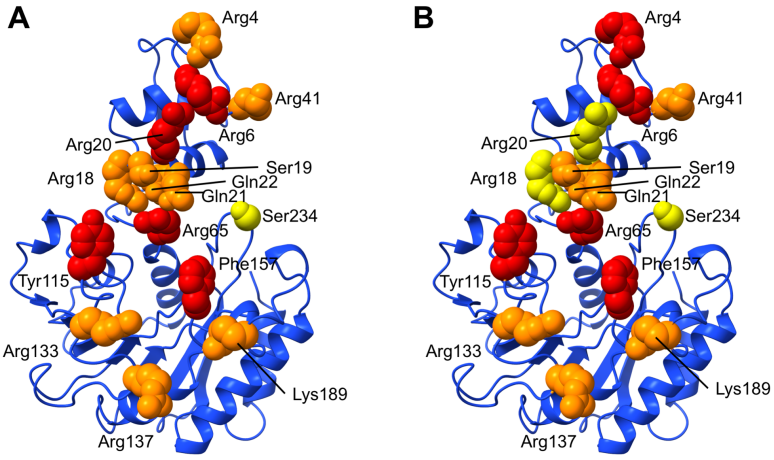


Figure 1. TlyA's recognition surface varies between the 50S and 30S ribosomal subunits. **A.** The recognition surface of TlyA for the 50S ribosomal subunit spans both N-terminal and C-terminal domains and includes residues Arg6, Arg20, Arg65, and Tyr 115 which are the most critical for binding and recognition (all residues are color coded: red = a residue critical for binding, orange = moderately important, yellow = unimportant). Residue Phe157 plays a role in stabilizing base flipping rather than binding. **B.** This surface changes when binding the 30S ribosomal subunit, where residues Arg4 and Arg6 are the most critical for binding with the C-terminal domain binding landscape being relatively the same. Notably, residue Arg20 becomes significantly less important in binding the 30S subunit.

specific dependencies on residues that play the most critical roles in the binding and recognition of each.

A future structure of TlyA bound to the 30S will be critical to fully defining the nature of the binding surface for that substrate. Additionally, complementary activity assays to test the role(s) of the any additional residues of interest from that structure will be necessary to complete mapping of this surface. With both the 50S and 30S interactions defined in molecular details, we could then apply this knowledge to other resistance methyltransferases that modify either the 50S or 30S subunit to help define their binding surfaces and mechanisms as well. For example, a number of ribosomal

methyltransferases, including TlyA and resistance methyltransferases NpmA and the ArmA/RmtA family, modify nucleotides on helix 44 (h44) of the 30S ribosomal subunit (**Fig. 2**) (6). How methyltransferases recognize that region could help guide us in developing ways to block binding and recognition of resistance methyltransferases but also ribosomal methyltransferases that are important for the function of the bacterial ribosome.

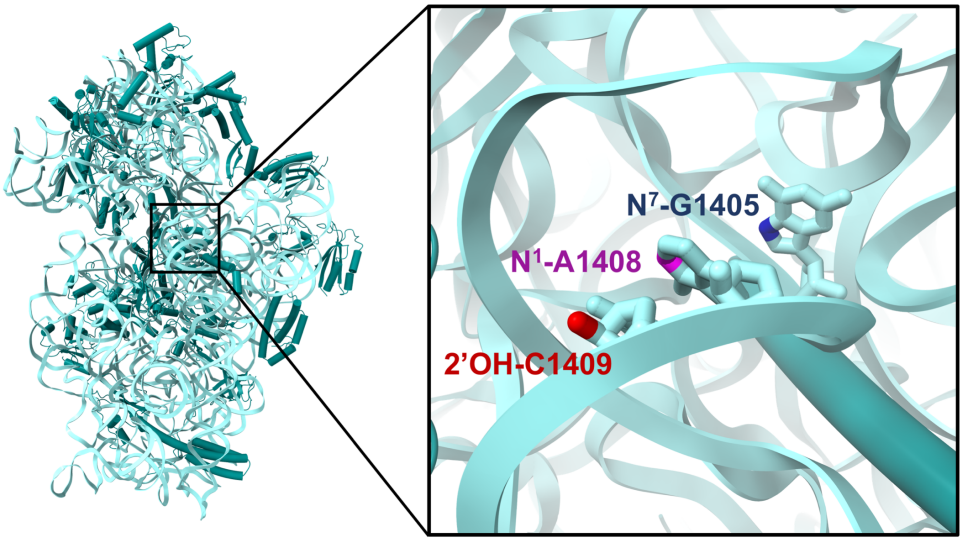


Figure 2. Methyltransferase target sites on h44 of 30S ribosomal subunit. h44 of the 16S rRNA of the 30S ribosomal subunit lies along the 50S subunit interaction surface and is modified by several ribosomal methyltransferases. (Inset) A zoomed in region of h44 showing the modification sites of TlyA (C1409 ribose 2'-OH modification site in red) and two classes of aminoglycoside-resistance rRNA methyltransferases A1408 (N1 modification site in magenta; e.g. NpmA), and G1405 (N7 modification site marked in blue; ArmA and RmtA-H).

Determining the effect of SAM-binding on the conformation and function of TlyA.

Previously, the Conn lab discovered that TlyA's interdomain linker was critical for SAM-binding (specifically Trp62 and Val63) (7). As described in Chapter 4, preliminary

hydrogen-deuterium exchange coupled to mass spectrometry (HDX-MS) studies comparing full-length TlyA with and without SAM bound revealed that the linker had decreased dynamics when SAM was bound, supporting the idea that a conformational change occurs after binding.

In our studies of the 50S-TlyA complex in Chapter 3, the conformation of the linker in the 50S- and NM6-bound structure of TlyA reveals that Trp62 and Val63 form a hydrophobic pocket with surrounding residues that may stabilize the N-terminal domain and C-terminal domain in their specific spatial orientation with respect to each other. This hydrophobic pocket may also restructure the parts of TlyA responsible for SAM binding (e.g. the conserved GASTG motif) which are proximal to the cosubstrate binding pocket. Further studies exploring the structure and dynamics of TlyA with and without SAM would be key to fully understanding the role of the linker and the two domains of TlyA in SAM binding. There are two possible mechanisms regarding TlyA conformation and SAM binding: 1) TlyA binding to the ribosomal subunit causes a change in the conformation of the enzyme (specifically the linker) that promotes SAM-binding, and 2) that SAM-binding causes a change in conformation in the linker (and therefore the conformation between the N- and C-terminal domain) that promotes subunit binding. A direct way to determine which mechanism is correct would be to perform additional HDX-MS experiments that probe the changes in TlyA linker dynamics in the presence of ribosome or SAM. If the linker changes conformation after binding the ribosome, a sample of TlyA and ribosome would show a difference in linker dynamics from a sample of TlyA alone. If the linker changes conformation after binding SAM, a sample of a TlyA N-terminal domain truncation mutant featuring only the linker plus C-terminal domain would show changes in linker dynamics between its *apo* and SAM-bound forms. The N-

terminal domain would be excluded in this last experiment to reveal whether changes the linker dynamics are cause directly by SAM binding or, instead, result directly due to changes in the N-terminal domain and C-terminal domain interaction when SAM is bound.

Base-flipping: a conserved mechanism in rRNA methyltransferases.

Our structural studies of the 50S-TlyA complex revealed that the nucleotide base of TlyA's target 23S rRNA nucleotide, C1920, was flipped out of H69 and this altered helical structure stabilized by the conserved TlyA residue Phe157. This finding was somewhat surprising as the modification performed by TlyA occurs on the ribose ring common to all nucleotides and which is in a location in the H69 minor groove that is relatively accessible to TlyA. Base flipping is a common mechanism used by DNA methyltransferases and repair enzymes (8,9), has been observed in one other structure of an RNA methyltransferase bound to its substrates (6), and has been suggested as the mechanism of modification in other RNA methyltransferases (5,10). More studies into TlyA's base flipping mechanism would be useful to not only understand how TlyA recognizes and modifies its target sites, but also how other ribosomal RNA (rRNA) methyltransferases do so. Base flipping by TlyA could be a mechanism by which the enzyme confirms the identity of its target site, C1920, but this is currently unclear. In our functional studies of 50S-TlyA, we found that TlyA Ser234 plays little role in the enzyme's methyltransferase activity despite being positioned to directly hydrogen bond to the flipped-out base. Considering the flexible nature of the loop containing this residue (two highly conserved glycine residues flank Ser234) and proximity to the flipped base, this loop may play the role of sequestering the base once it is flipped out of the rRNA helix, possibly

discriminating between correct and incorrect nucleotide bases. However, given that no particular TlyA residue in the loop seems to make essential direct interactions with the flipped base, this discrimination may be based on overall sterics, allowing discrimination against the larger purine bases (A or G). Whether activity would be retained with a U1920 variant is less certain.

To fully explore whether base flipping by TlyA confirms the identity of its target site, C1920, and thus contributes to substrate recognition, functional studies on 50S subunits after mutating this cytidine to other nucleotides could be performed. These variant subunits would be testing for changes in the extent of ribose modification to determine whether TlyA is discriminating based on target site identity or will modify any residue at 23S rRNA position 1920. Such studies would define whether substrate recognition on the 50S relies exclusively on more distant contacts, such as the rRNA junction at the base of H69 or also involves direct probing of the base identity. Structural studies by cryo-EM of TlyA bound to the 50S subunit but containing nucleotide mutations at position 1920 of the 23S rRNA could also shed light on how TlyA accommodates such altered 50S substrates should they still be modified. A deeper understanding of these processes this would allow us to speculate whether structurally similar methyltransferases also use this mechanism.

Impact of TlyA's modifications on capreomycin binding.

Though structural studies have been performed on TlyA-modified ribosomes bound with capreomycin (2,11), the contribution of TlyA's methylations at positions C1409 on 16S rRNA and C1920 on 23S rRNA to drug binding is still not fully known. Position C1409 is closest to the capreomycin binding site within the ribosome and may directly interact with

capreomycin to position it within that pocket. However, 23S rRNA C1920 is significantly further (~18 Å) from the drug binding site and makes no direct contact with capreomycin. Comparing the structures of ribosomes from bacteria that do or do not possess TlyA, and therefore possess or lack those modifications, respectively, suggested that the conformation at the tip of H69 changes significantly depending on whether the C1920 2'O-methylation is present (**Fig. 3**). This suggests that the C1920 modification alters the conformation of H69 to thereby allow optimal binding of capreomycin between the bases of nucleotides A2151 and C2152 (*Mtb* numbering) at the tip of H69 (**Fig. 3A,B**). In the structures of ribosomes from bacteria which do not possess TlyA, these two nucleotides and the tip of Helix 69 are much more tightly folded into the hairpin loop in a manner which may reduce capreomycin binding affinity (**Fig. 3C,D**). To determine if this C1920 modification results in a significant change in conformation, cryo-EM structural studies could be performed to compare a TlyA-modified and TlyA-unmodified ribosome, e.g. from the TlyA-deficient *Msm* used in our work and the parental strain. Additionally, the structure of the unmodified ribosome with capreomycin bound would demonstrate the importance of the H69 conformation in its binding.

Role of TlyA modifications in normal cell function.

The role of TlyA's modifications in normal cell function is still not known, but there is some evidence to support the idea that these modifications may stabilize mature 70S ribosomes (12). The two modifications are close to the 50S-30S subunit interaction surface where they may alter binding interactions between the rRNAs of the two subunits and/or the conformation of Helix 69 (as mentioned above). These modifications are also

close to the decoding center of the ribosome, specifically nucleotides A1492 and A1493 (in *E. coli* numbering), which play a key role in ensuring cognate tRNA-mRNA pairing. The change in conformation brought on by these modifications may also adjust the positioning of these critical nucleotides and could thus affect the accuracy of decoding as a result.

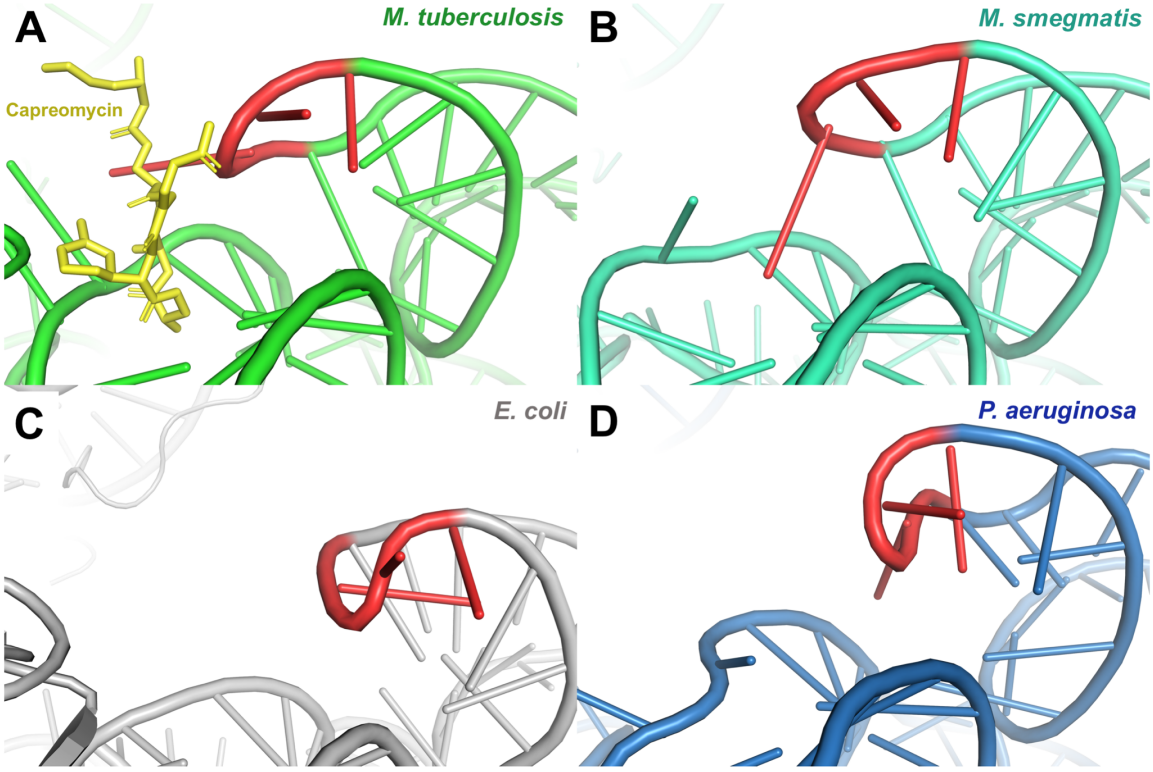


Figure 3. Potential structural consequences of C1920 methylation by TlyA on H69 conformation. **A.** The structure of *Mtb* ribosome with capreomycin (yellow) bound at tip of H69 (equivalent tip nucleotides are marked in red in all panels). **B.** The structure of H69 of the *M. smegmatis* ribosome bears resemblance to the *Mtb* ribosome in that nucleotides of the tip form a more open turn. This is in contrast to the conformation of H69 seen in **C.** *E. coli* and **D**. *P. aeruginosa* which have a tightly curled tip. *E. coli* and *P. aeruginosa* do not possess TlyA and this H69 structural difference may be the result of the lack of modification by this enzyme. This difference in conformation may explain the difference in capreomycin sensitivity seen between bacteria which express TlyA and those which do not, as this "open" conformation of H69 may better accommodate its binding.

To further test the effect of these modifications on the stability of the 70S ribosome, in addition to the previously published polysome profiling analyses (12), an *in vitro* translation assay using methylated or unmethylated ribosomal subunits with protein production as a read out could probe the effect of individual modifications on the overall 70S stability (e.g. methylated 50S with unmethylated 30S and vice versa). If these modifications do contribute to 70S stability, the assay would show greater translation for the modified subunits and less translation (and therefore more unstable 70S) for the unmodified subunits. If these modifications do affect stability, it is possible that TlyA may play a role in adjusting the longevity of the 70S ribosome complex, possibly altering that stability in times of cellular stress.

To test the effect of these modifications on translational fidelity, a different *in vitro* translation assay using single-molecule fluorescence resonance energy transfer (FRET) imaging, similar to that previously used to study the effect of antibiotics on translation (13), could be used to see if these modifications affect the ability of A1492 and A1493 to distinguish cognate pairing between tRNA and mRNA. Two different experiments could be performed, one with a message that cannot be read due to the presence of a codon which requires a tRNA that is not supplied to the system and the other with that tRNA supplied. Translation would be initiated with a donor fluorophore-labeled fMet-tRNA followed by the codon which does or does not have the appropriate tRNA provided. Acceptor fluorophore-labeled tRNAs would be provided to the system and acceptor fluorescence would be measured as a readout of translation (as proximity to the donor in the ribosome would cause acceptor fluorescence). If these modifications affect the translational fidelity of the ribosome by interfering with A1492 and A1493's recognition of cognate tRNA, ribosomes without the modifications would be able to continue

translating by using non-cognate tRNA at the site which would normally stop translation. Such a mechanism to by-pass amino acid limitation would suggest that TlyA plays a role in adjusting protein production in times of starvation, allowing the ribosome to use incorrect amino acids if nutrients are scarce.

Final remarks

Before these studies, knowledge of TlyA was primarily limited to its effect on capreomycin susceptibility in bacteria, the locations of its modifications on the ribosome, the crystal structure of its C-terminal domain and linker, and the predicted structure of its N-terminal domain. With my research, we now know the full-length structure of TlyA, how it binds to the 50S ribosomal subunit, and have a deeper, molecular level insight into its mechanism of recognition and target modification, which makes use of a positive-charged surface spanning both its domains and implements base-flipping.

Overall, this work on TlyA reveals common motifs in recognition and modification shared amongst several other enzymes. The N-terminal domain of TlyA is structurally homologous to ribosomal protein S4 which appears to be a structural motif broadly co-opted for other H69-recognizing enzymes (14-16). At the base of H69, this motif recognizes a specific tertiary structure of the rRNA using highly-conserved positively-charged residues. This suggests that S4-like motifs have been adapted across a variety of enzymes for recognition and binding of H69, specifically the non A-form helical structure located at its base. This is highlighted by the critical importance of the TlyA N-terminal domain in 50S subunit binding as seen in our structural and activity studies. Beyond this (and also seen in our activity studies), the TlyA N-terminal domain is also capable of recognizing the 30S ribosomal subunit, showing the great flexibility of this particular

structural motif in recognizing radically different sites on the ribosome. This flexibility can also be seen in the TlyA C-terminal domain, which takes on a common class I methyltransferase fold, which can recognize and modify sites both on H69 of the 23S rRNA and h44 of the 16S rRNA. Mechanistically speaking, this work suggests that the base-flipping mechanism of nucleotide modification is a common method employed by a variety of DNA- as well as RNA-modifying enzymes, regardless of where the modification lies on the nucleotide.

A couple of things remain to be seen following this research: the purpose of TlyA's modifications in cell function and the difference between these mechanisms for recognizing and modifying the 50S subunit and those for recognizing and modifying the 30S subunit. Understanding how TlyA's modifications affect the function of the ribosome would give us a more complete picture of how translation is regulated in bacteria and may also give us a better understanding of how other modifications (including nearby resistance methylations) affect the ribosome function. Finally, if the mechanisms of recognition and modification by TlyA are conserved across clinically relevant resistance methyltransferases, then these shared sites and structures of recognition on the rRNA, the common folds found in these ribosome-recognizing proteins, and the widespread base-flipping in RNA modification make attractive targets for novel mechanism-based therapeutics against antibiotic-resistant bacteria.

References

1. Johansen, S.K., Maus, C.E., Plikaytis, B.B. and Douthwaite, S. (2006) Capreomycin binds across the ribosomal subunit interface using tlyA-encoded 2 '-O-methylations in 16S and 23S rRNAs. *Mol Cell*, **23**, 173-182.
2. Stanley, R.E., Blaha, G., Grodzicki, R.L., Strickler, M.D. and Steitz, T.A. (2010) The structures of the anti-tuberculosis antibiotics viomycin and capreomycin bound to the 70S ribosome. *Nat Struct Mol Biol*, **17**, 289-293, PMCID: PMC2917106.
3. Walker, T.M., Merker, M., Knoblauch, A.M., Helbling, P., Schoch, O.D., van der Werf, M.J., Kranzer, K., Fiebig, L., Kroger, S., Haas, W., Hoffmann, H., Indra, A., Egli, A., Cirillo, D.M., Robert, J., Rogers, T.R., Groenheit, R., Mengshoel, A.T., Mathys, V., Haanpera, M., Soolingen, D.V., Niemann, S., Bottger, E.C., Keller, P.M. and Consortium, M.-T.C. (2018) A cluster of multidrug-resistant Mycobacterium tuberculosis among patients arriving in Europe from the Horn of Africa: a molecular epidemiological study. *Lancet Infect Dis*, **18**, 431-440, PMCID: PMC5864516.
4. Phelan, J.E., O'Sullivan, D.M., Machado, D., Ramos, J., Oppong, Y.E.A., Campino, S., O'Grady, J., McNerney, R., Hibberd, M.L., Viveiros, M., Huggett, J.F. and Clark, T.G. (2019) Integrating informatics tools and portable sequencing technology for rapid detection of resistance to anti-tuberculous drugs. *Genome Med*, **11**, 41, PMCID: PMC6591855.
5. Zhao, M., Zhang, H., Liu, G., Wang, L., Wang, J., Gao, Z., Dong, Y., Zhang, L. and Gong, Y. (2016) Structural Insights into the Methylation of C1402 in 16S rRNA by Methyltransferase RsmI. *PLoS One*, **11**, e0163816, PMCID: PMC5053481.
6. Dunkle, J.A., Vinal, K., Desai, P.M., Zelinskaya, N., Savic, M., West, D.M., Conn, G.L. and Dunham, C.M. (2014) Molecular recognition and modification of the 30S ribosome by the aminoglycoside-resistance methyltransferase NpmA. *Proc Natl Acad Sci U S A*, **111**, 6275-6280.
7. Witek, M.A., Kuiper, E.G., Minten, E., Crispell, E.K. and Conn, G.L. (2017) A Novel Motif for S-Adenosyl-l-methionine Binding by the Ribosomal RNA Methyltransferase TlyA from Mycobacterium tuberculosis. *J Biol Chem*, **292**, 1977-1987, PMCID: PMC5290967.
8. Neely, R.K., Daujotyte, D., Grazulis, S., Magennis, S.W., Dryden, D.T., Klimasauskas, S. and Jones, A.C. (2005) Time-resolved fluorescence of 2-aminopurine as a probe of base flipping in M.HhaI-DNA complexes. *Nucleic Acids Res*, **33**, 6953-6960, PMCID: PMC1310896.
9. Zhang, H., Gao, Z.Q., Wei, Y., Wang, W.J., Liu, G.F., Shtykova, E.V., Xu, J.H. and Dong, Y.H. (2012) Structural insights into the function of 23S rRNA methyltransferase RlmG (m(2)G1835) from Escherichia coli. *RNA*, **18**, 1500-1509, PMCID: PMC3404371.
10. Hager, J., Staker, B.L. and Jakob, U. (2004) Substrate binding analysis of the 23S rRNA methyltransferase RrmJ. *J Bacteriol*, **186**, 6634-6642, PMCID: PMC516604.
11. Yang, K., Chang, J.Y., Cui, Z., Li, X., Meng, R., Duan, L., Thongchol, J., Jakana, J., Huwe, C.M., Sacchettini, J.C. and Zhang, J. (2017) Structural insights into species-

specific features of the ribosome from the human pathogen Mycobacterium tuberculosis. *Nucleic Acids Res*, **45**, 10884-10894, PMCID: PMC5737476.

12. Salamaszynska-Guz, A., Taciak, B., Kwiatek, A. and Klimuszko, D. (2014) The Cj0588 protein is a Campylobacter jejuni RNA methyltransferase. *Biochem Bioph Res Co*, **448**, 298-302.
13. Jenner, L., Starosta, A.L., Terry, D.S., Mikolajka, A., Filonava, L., Yusupov, M., Blanchard, S.C., Wilson, D.N. and Yusupova, G. (2013) Structural basis for potent inhibitory activity of the antibiotic tigecycline during protein synthesis. *Proc Natl Acad Sci U S A*, **110**, 3812-3816, PMCID: PMC3593886.
14. Crowe-McAuliffe, C., Takada, H., Murina, V., Polte, C., Kasvandik, S., Tenson, T., Ignatova, Z., Atkinson, G.C., Wilson, D.N. and Hauryliuk, V. (2021) Structural Basis for Bacterial Ribosome-Associated Quality Control by RqcH and RqcP. *Mol Cell*, **81**, 115-126 e117.
15. Desai, N., Yang, H., Chandrasekaran, V., Kazi, R., Minczuk, M. and Ramakrishnan, V. (2020) Elongational stalling activates mitoribosome-associated quality control. *Science*, **370**, 1105-1110, PMCID: PMC7116630.
16. Filbeck, S., Cerullo, F., Paternoga, H., Tsaprailis, G., Joazeiro, C.A.P. and Pfeffer, S. (2021) Mimicry of Canonical Translation Elongation Underlies Alanine Tail Synthesis in RQC. *Mol Cell*, **81**, 104-114 e106, PMCID: PMC7796892.