STRUCTURAL STUDIES OF MYCOBACTERIUM TUBERCULOSIS RIBOSOME:

FROM TRANSLATIONAL REGULATION TO DRUG RESISTANCE

A Dissertation

by

ZHICHENG CUI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Michael Polymenis
Committee Members,	Junjie Zhang
	Hays Rye
	Yang Shen
Head of Department,	Josh Wand

May 2021

Major Subject: Biochemistry

ABSTRACT

Nearly half of clinically used antibiotics target ribosome, the vital translational machinery for bacterial survival. In this dissertation, we studied the translational regulation and antibiotic resistance in *Mycobacterium tuberculosis (Mtb)* to better understand persistency and provide insights for novel drug design. Firstly, we used cryo-electron microscopy (cryo-EM) and X-ray crystallography to systematically study an ATP-binding cassette F (ABCF) protein MtbEttA, which is possibly involved in translational control and antibiotic resistance. High-resolution cryo-EM structures of the *Mtb* ribosome in complex with MtbEttA at the pre-hydrolysis and transition states of the protein's ATPase cycle were solved. Surprisingly, an asymmetric engagement of conserved motifs around the two ATP binding sites and increased flexibility around the peptidyl transferase center (PTC) were observed in the pre-hydrolysis state. Different degrees of ribosomal intersubunit rotation were found between the pre-hydrolysis state and transition state. A crystal structure of a domain-swapped MtbEttA dimer bound to ADP illustrated the open conformation of the protein, showing a conformational change of nucleotide-binding domains (NBDs). These structures collectively shed light on the interplay between the ribosome and ABCF along the ATP hydrolysis trajectory. Secondly, we focused on the hibernation promoting factor Mpy, which is involved in ribosomal inhibition during stationary growth only when zinc is depleted in the cell. We found that the C-terminal domain of Mpy aggregated upon binding to zinc molecules, suggesting a zinc-controlled switch-mechanism. Finally, in collaboration with Sanofi,

we performed cryo-EM studies that revealed the structural mechanism employed by the novel antibiotic sequanamycin (SEQ-9) to overcome A2296-methylation induced drug resistance. A pivotal modification of the desosamine sugar in SEQ-9 was revealed, which avoided steric clashes with the methyl group of A2296, permitting high binding affinity to be retained. Our atomic-resolution structure of *Mtb* ribosome should also provide the foundation for future drug design.

DEDICATION

I dedicate this dissertation to those in the path of exploring the universe, who are curious, objective, resilient, and persistent.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Junjie Zhang, and my committee members, Dr. Hays Rye, Dr. Michael Polymenis, and Dr. Yang Shen, for their guidance and support throughout my graduate research.

Thanks also go to my previous advisors Dr. Baolei Jia and Dr. Yao Cong, for walking me through the starting line of the marathon of science. I thank my friends and colleagues at every stage of my science endeavor for their kindness and support.

Special thanks to Dr. Xiaojun Li for teaching me X-ray crystallography, performing ribosome purification, and in vitro translation assays. Also, to Dr. James Sacchettini and Dr. Ry Young for fruitful collaborations.

I also thank Dr. Pingwei Li for his guidance in processing X-ray crystallography data.

Finally, but most importantly, thanks to my parents for their encouragement and my wife Stephanie for her tremendous support, patience, and love.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professors Junjie Zhang, Hays S. Rye, Michael Polymenis of the Department of Biochemistry and Biophysics, and Professor Yang Shen of the Department of Electrical and Computer Engineering.

M. tuberculosis ribosome purification and all the in vitro translation assays were performed by Dr. Xiaojun Li. Dr. Li also helped with all aspects of X-ray crystallography.

Drs. Ya-Ming Hou and Howard Gamper from Thomas Jefferson University provided the initiator tRNA.

Funding Sources

The work was supported by Welch Foundation Grants [A-1863, A-0015]; National Institutes of Health TB Structural Genomics Grant [P01AI095208]; National Institutes of Health Grants [P41GM103832, U24GM116787]; Department of Biochemistry and Biophysics Start-up Fund; Center for Phage Technology at Texas A&M University Start-up Fund.

TABLE OF CONTENTS

Page

ABSTRACT ii

DEDICATION iv

ACKNOWLEDGEMENTS v

CONTRIBUTORS AND FUNDING SOURCES vi

TABLE OF CONTENTS vii

LIST OF FIGURES x

LIST OF TABLES xiii

CHAPTER I INTRODUCTION 1

I.1 Tuberculosis, a long-lasting pandemic 1
I.1.1 The treatment burden for TB 2
I.1.2 Morphology and genetics of *Mycobacterium tuberculosis* 4
I.2 Translational regulation in *Mycobacterium tuberculosis* 6
I.2.1 Unique features of Ribosome from *Mtb* 6
I.2.2 Innate properties of *Mtb* ribosome 8
I.2.3 Stress-induced factors involved in translational regulation in *Mtb* 10
I.3 Antibiotic resistance in *Mtb* 12
I.3.1 Anti-TB drugs and resistance mechanism 12
I.3.2 Ribosome is a successful antibiotic target 14
I.3.3 Intrinsic resistance to macrolides in *Mtb* 15
I.3.4 Direct ribosome protection proteins against antibiotics 16
I.4 Introduction and development of cryo-electron microscopy 18

CHAPTER II THE MECHANISM-OF-ACTION OF AN ABCF PROTEIN MTBETTA 25

II.1 Introduction 25
II.2 Materials and Methods 29
II.2.1 Protein expression and purification 29
II.2.2 In-vitro translation assay 31
II.2.3 Cryo-EM sample preparation and data collection 32

II.2.4 Image Preprocessing 33
II.2.5 Resolution Estimation and Post Processing 38
II.2.6 Crystallization, data collection, and structure determination 38
II.2.7 Modeling and Visualization 39
II.3 Results 39
II.3.1 Cryo-EM structures of *Mtb* ribosome with MtbEttA at pre-hydrolysis (ADPNP) and transition (ADP-VO4) states 39
II.3.2 Interactions between MtbEttA and *Mtb* ribosome 45
II.3.3 Remodelling of P-site tRNA and 50S in the course of ATP hydrolysis of MtbEttA 48
II.3.4 Structural variations of MtbEttA during ATP hydrolysis and asymmetric engagement of the nucleotides in pre-hydrolysis state 53
II.3.5 MtbEttA at post-hydrolysis state (ADP) adopts an open conformation 59
II.4 Discussion 65

CHAPTER III CHARACTERIZATION OF HIBERNATION PROMOTING FACTORS IN *MTB* 70

III.1 Introduction 70
III.2 Materials and methods 74
III.2.1 Protein expression and purification 74
III.2.2 Negative staining electron microscopy 75
III.2.3 Dynamic light scattering 75
III.2.4 On-column cross-linking 76
III.2.5 Sucrose gradient ultracentrifugation 76
III.3 Results 77
III.3.1 RafH does not interact with the ribosome in vitro 77
III.3.2 MsmMpy can only solubilize with C-terminal His tag in buffer with high salt and EDTA 80
III.3.3 MsmMpy inhibits translation in vitro 85
III.3.4 In vitro reconstitution and cryo-EM studies of ribosome-Mpy complex 86
III.3.5 Absence of interactions between Mpy and Mrf 89
III.3.6 Full-length and C-terminal domain of untagged MtbMpy can be purified by Ni-NTA resin 91
III.3.7 Tetrameric state is the dominant configuration of MtbMpy C-terminal domain 92
III.3.8 Reversible Zinc-mediated polymerization and aggregation of C-terminal domain of MtbMpy 95
III.3.9 Attempts to identify Zinc-binding residues in the C-terminal domain of MtbMpy 97
III.4 Discussion 99

CHAPTER IV NATURAL PRODUCT DERIVED SEQUANAMYCINS OVERCOME METHYLATION-INDUCED MACROLIDE RESISTANCE IN *MTB* RIBOSOME 101

IV.1 Introduction 101
IV.2 Materials and methods 104
IV.2.1 *Mtb* S30 and S100 preparation 104
IV.2.2 Methylated *Mtb* ribosome purification 104
IV.2.3 IC50 measurement of different compounds 105
IV.2.4 Cryo-electron microscopy and Data processing 106
IV.2.5 Molecular modeling and refinement 107
IV.3 Results 107
IV.3.1 A2296-methylated *Mtb* ribosome is resistant to macrolides 107
IV.3.2 SEQ-9 inhibits the A2296-methylated ribosome in vitro 110
IV.3.3 Cryo-EM structure of SEQ-9 in complex with A2297-methylated *Mtb* ribosome 111
IV.3.4 Structural variability around PTC in different states of *Mtb* ribosome 119
IV.3.5 The cryo-EM map of the 50S contains an RsfS density 120
IV.4 Discussion 122

CHAPTER V CONCLUSIONS AND FUTURE DIRECTIONS 124

V.1 Potential regulations and additional functions of MtbEttA 124
V.1.1 The sequence of events of MtbEttA's action on the ribosome 124
V.1.2 Dimer-monomer transition of MtbEttA 125
V.1.3 The potential role of MtbEttA in leaderless mRNA translation 126
V.2 Structural studies of Mpy and its regulation in vivo 126
V.2.1 Structural studies of Mpy C-terminal domain 127
V.2.2 Verification of the zinc-induced polymerization of Mpy in vivo 128
V.3 Insights of drug design toward methylated *Mtb* ribosome 128

REFERENCES 130

APPENDIX A ASYMMETRIC CRYO-EM STRUCTURE OF THE CANONICAL ALLOLEVIVIRUS QB REVEALS A SINGLE MATURATION PROTEIN AND THE GENOMIC SSRNA IN SITU 156

APPENDIX B STRUCTURES OF QB VIRIONS, VIRUS-LIKE PARTICLES, AND THE QB–MURA COMPLEX REVEAL INTERNAL COAT PROTEINS AND THE MECHANISM OF HOST LYSIS 175

LIST OF FIGURES

Page

Figure 1 TB incidents in the leading eight countries in 2019 2

Figure 2 Cell wall structure of *Mtb* 5

Figure 3 Structure of *Mtb* ribosome 7

Figure 4 Ribosomal stress response in bacteria 11

Figure 5 Structural differences between EttA and VmlR (antibiotic resistance ABCF) 17

Figure 6 Cryo-EM sample preparation using vitrobot 21

Figure 7 Effect of MtbEttA on erythromycin IC50 using in vitro translation assay 29

Figure 8 Cryo-EM data processing workflow for 70SIC-MtbEttA with ADPNP. 35

Figure 9 Cryo-EM data processing workflow for 70SIC-MtbEttA with ADP-VO_4. 37

Figure 10 Analysis of the condition for forming stable MtbEttA-ribosome complex. 41

Figure 11 Examples of cryo-EM density. 42

Figure 12 Cryo-EM structures of *Mtb* ribosome in complex with MtbEttA in pre-hydrolysis and transition states. 44

Figure 13 Electrostatic surface potential of MtbEttA. 46

Figure 14 Structure of MtbEttA and its interaction with *Mtb* Ribosome in Pre_R0 state. 47

Figure 15 Comparison of P-site tRNAs. 49

Figure 16 Comparison of the PtIM of MtbEttA in Pre_R0 state with *E.coli* EttA, VmlR, and MsrE. 50

Figure 17 Remodeling of the 50S in Pre_R1 state due to 30S rotation and conformational change of MtbEttA. 51

Figure 18 Increased flexibility of PTC and uL16 in Pre_R1 state 52

Figure 19 Remodelling of the 50S is unique in the Pre_R1 state. 53

Figure 20 Cryo-EM density of ATP analogues in nucleotide-binding sites. 55

Figure 21 Structural plasticity of MtbEttA in pre-hydrolysis and transition states.....55

Figure 22 Structural plasticity and asymmetric nucleotide engagement in MtbEttA 57

Figure 23 Structural plasticity and asymmetric nucleotide engagement in MtbEttA 58

Figure 24 Dimer monomer transition of MtbEttA 61

Figure 25 Conformational change of MtbEttA and steric clash with *Mtb* Ribosome in post-hydrolysis state. 63

Figure 26 Conformational changes of NBDs and Arm domain in apo and post-hydrolysis states. 65

Figure 27 Schematic model of the functional cycle of MtbEttA in the course of ATP hydrolysis.....69

Figure 28 Scheme of different mechanisms of the formation of 100S 72

Figure 29 Sequence alignment between RafH and LHPFs 78

Figure 30 RafH purification and pull-down experiments with ribosome 80

Figure 31 Purification attempts for MsmMpy.....82

Figure 32 C-terminal His tag purification of MsmMpy 83

Figure 33 Gel filtration and in vitro translation analysis of MsmMpy 86

Figure 34 Sucrose gradient analysis of MsmMpy-70S complex 87

Figure 35 Preliminary cryo-EM structures of MsmMpy-70S complex 88

Figure 36 Gel filtration analysis of purified MsmMrf using His-sumo and MBP tags....90

Figure 37 Co-purification of MsmMpy and MsmMrf.....90

Figure 38 Purification of full-length and different domains of MtbMpy.....92

Figure 39 Gel filtration of purified C-terminal domain of MtbMpy.....93

Figure 40 On-column cross-linking C-terminal domain of MtbMpy.....94

Figure 41 Structural characterization of MtbMpy C-terminal domain using NMR.....95

Figure 42 Dynamic light scattering profile of MtbMpy C-terminal domain. 96

Figure 43 Negative staining analysis of the aggregated MtbMpy C-terminal domain 97

Figure 44 Potential histidine cluster and mutagenesis study.....98

Figure 45 Potential Zinc-mediated switch mechanism of Mpy 100

Figure 46 Methylated *Mtb* ribosome is resistant to different generations of macrolides.....108

Figure 47 Methylated *Mtb* ribosome is NOT resistant to other antibiotics.....109

Figure 48 SEQ-9 can overcome methylation mediated resistance.....111

Figure 49 Cryo-EM maps of different subpopulations in the dataset of methylated *Mtb* ribosome bound with SEQ-9 112

Figure 50 Density of SEQ-9 at different contour levels.....113

Figure 51 Cryo-EM density of 3 potentially methylated RNAs 114

Figure 52 Cryo-EM densities of RNA modifications in *Mtb* ribosome.....115

Figure 53 Structures of the three generations of macrolides.....116

Figure 54 Macrolide resistance due to the methylation of A2296 in *Mtb* ribosome.....117

Figure 55 Structures of the different sequanamycin lead compounds 118

Figure 56 Modified desosamine sugar of SEQ-9 avoids steric clash with the methyl group of A2296.....119

Figure 57 Variant PTC conformations in different ribosome subpopulations revealed by cryo-EM.....120

Figure 58 RafS density is observed on the 50S.....122

LIST OF TABLES

Page

Table 1 Anti-TB drugs 14

Table 2 Antibiotics that target the ribosome 15

Table 3 Antibiotic resistance ABCFs 18

Table 4 X-ray crystallography statistics of MtbEttA-ADP 60

Table 5 Solubility test for MsmMpy with C-terminal His tag in different buffers 84

Table 6 Summary of the IC50 of different antibiotic toward unmethylated and A2296-methylated *Mtb* ribosome 110

CHAPTER I

INTRODUCTION

I.1 Tuberculosis, a long-lasting pandemic

In the midst of a global pandemic, what the year 2020 left us was a grave reminder. It is about the vulnerability of human life, the necessity of scientific communication, and the inequality of our society. Although the development of vaccines against COVID-19 was at an unprecedented pace, the death toll has surpassed 1.7 million worldwide as of December 2020 (1). With more than 3 million projected deaths (2), COVID-19 is neither the first nor the last pandemic to test humankind. Historically, pandemics caused by viruses usually last for one to three years. However, bacteria-induced pandemics can last for decades and recur even centuries later (3).

Tuberculosis (TB), a highly infectious disease caused by the aerosol transmissible bacteria *Mycobacterium tuberculosis* (*Mtb*), has been documented since three thousand years ago (4). Even today, 1.4 million people died from TB, and an estimated 10 million fell ill with TB worldwide in 2019 alone (5). TB remains one of the top 10 causes of death and the leading cause of death from a single infectious agent worldwide. TB is also the leading killer of people with HIV/AIDS, claiming over 25% of the lives of people with HIV (5) (6) (7). Although the occurrence of TB in Europe and North America is low nowadays, it was responsible for almost 25% of all the deaths during the 1600s and 1800s (8). In 2019, 87% of new TB cases were from the 30 high TB burden countries. The leading eight countries responsible for two-thirds of the total

cases are India, Indonesia, China, the Philippines, Pakistan, Nigeria, Bangladesh, and South Africa (5) (Fig 1).

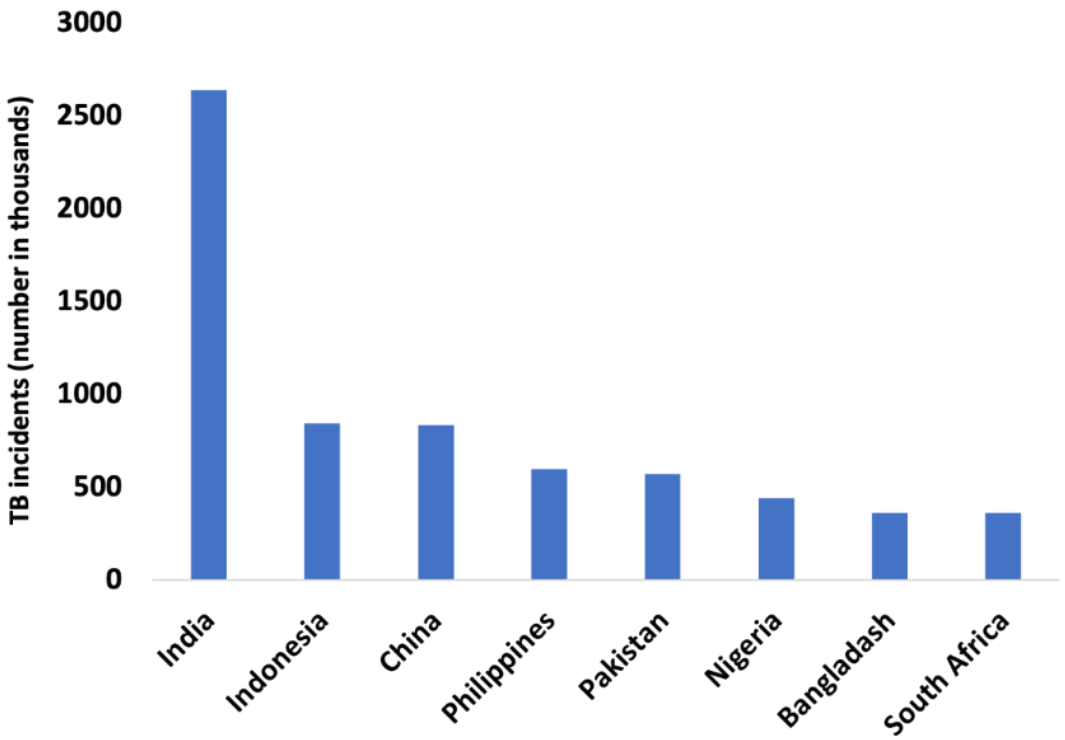


Figure 1 TB incidents in the leading eight countries in 2019

I.1.1 The treatment burden for TB

Although there were tremendous efforts in eliminating TB since Dr. Robert Koch identified the bacterium *Mycobacterium tuberculosis* as the cause in 1882 (9) (10), the prevention and treatment for TB remain challenging even today. The development of vaccines against TB started in the early 20th century. However, only one vaccine, Bacille Calmette-Guérin (BCG), is available with controversial results regarding its efficacy

(11). Diagnoses of TB have evolved from looking for the bacilli under a light microscope to more convenient and robust tests, including chest X-ray, skin and blood test (12) (13) (14). Currently, the treatment for drug-susceptible active TB involves a 6-9 months regimen using first-line antibiotics, including isoniazid (15), rifampin (16), ethambutol (EMB), and pyrazinamide (PZA). Isoniazid and rifampicin are used throughout the regimen, with ethambutol and pyrazinamide in the first 18 weeks of treatment (17). However, the outcome could be compromised by poor adherence and antibiotic resistance, mainly due to the inadequate treatment of active TB (18). Such a prolonged treatment also imposes a financial burden on patients, especially those in developing countries (19).

There are two types of TB, Latent TB Infection (LTBI) and TB disease (20). The world health organization (WHO) estimated that nearly one-quarter of the world's population (1.8 billion) have LTBI. Although people with LTBI neither show any symptoms nor spread active *Mtb* bacilli, there is a 5-10% chance for those individuals to develop active TB (21) (22). Due to the large base number of people with LTBI, the potential number for active TB could be devastating.

One of the standard features of successful pathogens is their ability to escape from the human immune system and then resurface decades later (23) (24). Upon infection, active *Mtb* bacilli are phagocytized by alveolar macrophage and resident dendritic cells. Granuloma, a focal collection of inflammatory cells with specific architectural structures in humans, is formed after the bacilli-containing macrophages and lymphocytes migrate to the primary infection site. Bacterial growth increases

logarithmically in the granuloma and then reaches a plateau coincident with T cell response (25) (26). Maintenance of the granulomas is considered a dynamic process of continual immunologic control and a physical barrier of bacterial replication (27).

Approximately 5% of infected individuals will disseminate active *Mtb* bacilli from granuloma in the first year or two after infection, and another 5% will develop the disease sometime later in life (28). Around 90% of the time, *Mtb* bacilli will be contained in the granuloma in a dormant non-infectious form. The treatment is recommended to prevent latent TB from developing into active TB (29) (30). Current remedies for LTBI include prolonged regimens: 6 or 9-months isoniazid; 12-weeks rifapentine plus isoniazid; 3-4 months isoniazid plus rifampicin; or 3-4 months rifampicin alone. However, due to isoniazid's lengthy treatment and side effects, the effectiveness has been limited by low treatment acceptance and completion (31) (32). Regular monitoring of isoniazid treatment is required due to the severe but rare risk of hepatotoxicity. Other less severe side effects, including nausea, fatigue, and headaches, could also hinder the treatment (33) (34).

I.1.2 Morphology and genetics of *Mycobacterium tuberculosis*

M. tuberculosis belongs to the diverse family of Actinobacteria, a rod-shaped bacterium about 2-4 μm in length and 0.2-0.5 μm in width. Although *Mtb* is classified as a Gram-positive bacterium based on 16S ribosomal RNA sequence comparison (35), it stains very weakly Gram-positive or not at all when a Gram stain is performed (16). Morphologically, it is neither Gram-positive nor Gram-negative due to the unique cell

wall composition (Fig 2). The waxy cell wall of *Mtb* contains peptidoglycan, while over 60% are complex lipids, including mycolic acids and cord factor (36) (37).

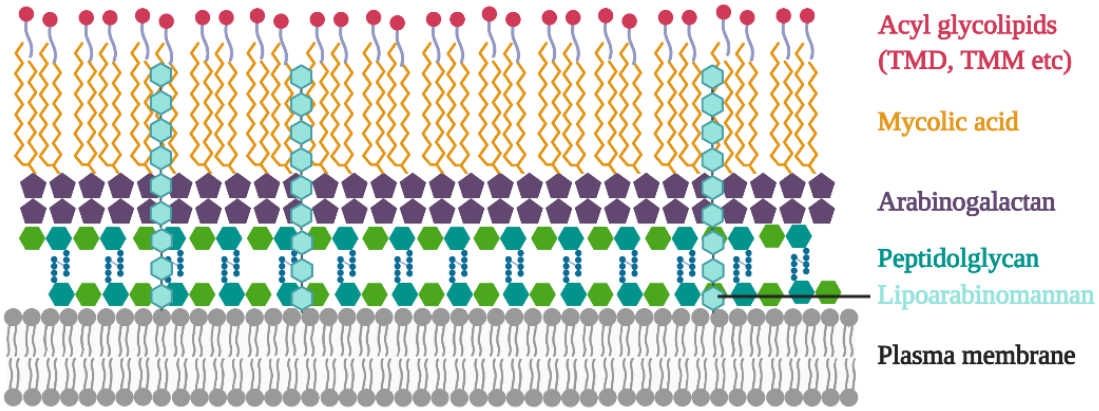


Figure 2 Cell wall structure of *Mtb*

Mycolic acids are unique long-carbon-chain (60-90 carbon atoms) alpha-branched lipids, which make up to 50% of the mycobacterial cell envelope's dry weight (38) (39). Due to the strong hydrophobic properties of mycolic acids, the cell surface of *Mtb* shows low permeability. It contributes to the resistance of cationic proteins, lysozyme, and oxygen radicals in the phagocytic granule (40) (41). Trehalose-6,6-dimycolate (TDM), the mycobacterial cord factor, is the primary cell wall lipid that has long been known to contribute to the virulence of *Mtb* (42). TDM is an inhibitor of polymorphonuclear leukocyte migration and sufficient to cause granuloma formation by itself (43). The unique properties of mycolic acid and cord factor in the cell wall have been associated with the persistency of *Mtb*, including resistance to many antibiotics, acid-fastness, and survival inside of macrophages (44) (45).

The whole-genome sequence of Mycobacterium tuberculosis H37Rv was obtained and analyzed in 1998 (46). As the best-characterized laboratory strain, the genome of H37Rv contains 4,411,529 base pairs, with an average of 65.61% GC content. Initially, around 4,000 genes were annotated. However, the growing knowledge of the unique properties of translation in mycobacteria may give rise to many unannotated genes (47). Genetic analysis revealed many unique aspects of *Mtb*. For example, ~250 distinct enzymes involve fatty acid metabolism in *Mtb*, five times more than those in *Escherichia coli* (*E. coli*). There is only one *rrn* operon in *Mtb*, while seven are present in *E.coli*. Remarkably, 79 toxin-antitoxin systems are found in the *Mtb* chromosome (46).

I.2 Translational regulation in *Mycobacterium tuberculosis*

I.2.1 Unique features of Ribosome from *Mtb*

Ribosome, the vital molecular machinery consisted of RNAs and dozens of proteins, is responsible for protein synthesis in all three domains of life. Although ribosomes' central function in prokaryotes and eukaryotes is the same, their structures and regulations are diverse (48). Species-specific features of ribosomes, even among prokaryotes, are often observed (49) (50) (51). Mycobacterial ribosomes have many unique components, such as ribosomal RNA (rRNA) expansion segments (ES), ribosomal proteins (rProteins), and ribosomal factors (rFactors), whose structures and functions are not yet fully characterized. Previously, we used cryo-EM to explore the unique features of the *Mtb* ribosome (49).

The prominent unique feature of *Mtb* ribosome is a 100-nt long ribosomal RNA (rRNA) expansion segment (ES), named H54a or "handle." (Fig. 3) The handle spans between the L1 stalk and the small subunit (SSU) in the 70S and has interactions with ribosomal proteins bL9 and bS6. The significant flexibility of handle on the 70S was observed based on cryo-EM structures. Moreover, its motion is found to be correlated with the movement of the SSU and L1 stalk. Interestingly, the handle on the 50S rotates almost 40^0 toward the 30S-interacting interface, potentially blocking the formation of 70S. Furthermore, it has stable interactions with H68 of 23S rRNA and ribosomal protein uL2, with little structural variability.

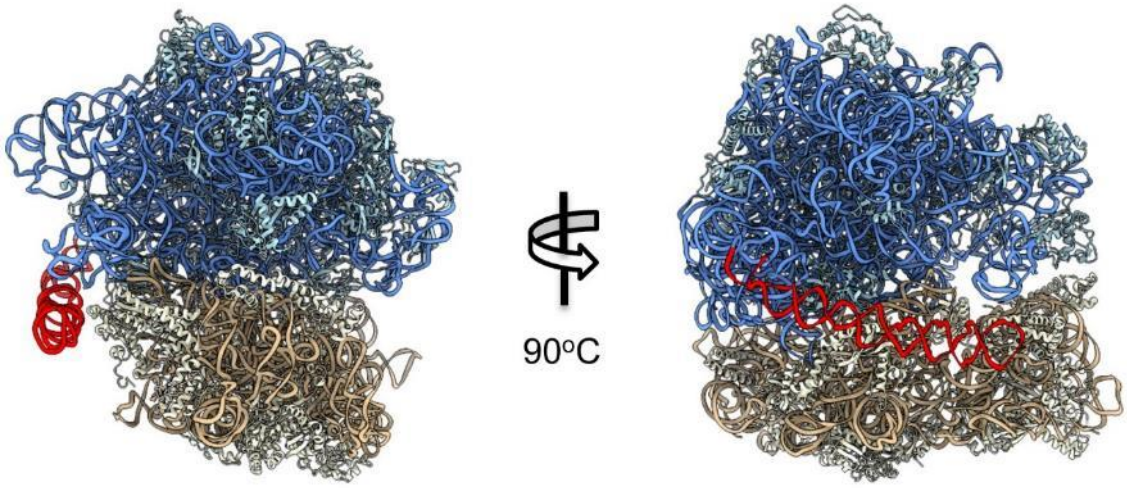


Figure 3 Structure of *Mtb* ribosome
The large subunit (50S) is colored in cornflower blue (23S rRNA and 5S rRNA) and light blue (50S rProteins). The small subunit (30S) is colored in tan (16S rRNA) and light yellow (30S rProteins). The handle is colored in red.

The handle's function needs further investigation, while some speculations can be made according to the structure. Firstly, translational initiation could be unique as the

handle on the 50S could potentially prevent 30S from joining to form 70S. Secondly, the handle's position on the 70S is close to the mRNA exit site, which may interfere with the ribosome's elongation cycle and polysome formation.

Other unique features of *Mtb* ribosome are the newly identified ribosomal proteins: bS22 and bL37. The first new protein, bS22, is about 30 amino acids long. It is located in a pocket, formed by h44 and h45 of the 16S and H67 and H69 of the 23S. While the highly conserved pocket is empty in ribosomes from *E.coli* and other bacteria, an α-helical protein, eL41 or mL41, is present in eukaryotic cytosolic ribosomes or mitoribosomes, respectively. There is a high structural similarity between mL41 and bS22. The second protein, bL37, locates close to the peptidyl transferase center (PTC), in a position that no protein has been observed in any bacterial, archaeal, yeast, or mammalian ribosomes. One end of the helix in bL37 is close to H89, a critical component of the PTC.

I.2.2 Innate properties of *Mtb* ribosome

M. tuberculosis belongs to the slow-growing bacteria, with a doubling time of around 20 hours (52). Its non-pathogenic relative, *Mycobacterium smegmatis* (*M. smegmatis*), has a doubling time of around 2 hours (53). The growth rate for mycobacteria is notably slower compared to *E. coli*, which has a doubling time of around 20 minutes. It is speculated that one of the contributions to the slow growth rate is the lower translational efficiency of mycobacterial ribosomes. However, *M. smegmatis* ribosome's biochemical characterization showed comparable initiation and elongation

kinetics to *E. coli* ribosome (54). Mtb ribosome's translational efficiency is also comparable to *E. coli* ribosome based on our in vitro translation assays. A more plausible explanation regarding the growth rate could point to the different numbers of ribosomes present in a cell. In *E. coli*, around 72,000 ribosomes are present per cell, which account for over 1/3 of the cell's dry mass (55). However, only on average 700 ribosomes are estimated per cell in *Mtb* (56). The significantly smaller number of ribosomes *per Mtb* cell could devote to its slow growth rate.

Despite the slow growth rate, ribosomal profiling of *Mtb* and *M. smegmatis* revealed that almost 25% of all the transcripts are leaderless (47), which by definition do not contain Shine-Dalgarno (SD) sequence and a 5' untranslated region (57). The results indicate that ribosomes from *Mtb* and *M. smegmatis* have the intrinsic ability to translate leaderless mRNA. Canonically, a 30S pre-initiation complex is formed by joining mRNA, the three initiation factors (IFs), and the initiator tRNA (N-formylmethionine-charged tRNA) to the small ribosomal subunit during translational initiation (58). In prokaryotes, the binding of mRNA to 30S is mediated by the direct interactions between the SD sequence of mRNA and the anti-Shine-Dalgarno (aSD) sequence of 16S rRNA (59). However, translating leaderless mRNA involves an alternative mechanism. In *E. coli*, pre-formed 70S ribosomes are required to bind leaderless mRNA first and then joined by initiator fMet-tRNA^{fMet} and initiation factors IF2 and IF3 (60). The ratio of IF2 and IF3 is the determining factor in the efficiency of initiating leaderless mRNAs, with higher relative concentrations of IF2 promoting leaderless translation (61). Nevertheless, the mechanism of leaderless mRNA translation in *Mtb* is yet unexplored.

I.2.3 Stress-induced factors involved in translational regulation in *Mtb*

Since *Mtb* is contained in the granuloma in the latent infection, its ability to transform to and maintain the dormant form is critical for survival (Fig). To study the metabolic state of *Mtb* during latent infection, an in vitro method, namely the "Wayne model" pioneered by Lawrence Wayne, is widely used (62). In this model, *Mtb* cells are inoculated at a low density in a sealed tube. A non-replicating and dormant state of *Mtb* is achieved with stirring and slowly consuming oxygen until the culture is anaerobic (63). And then, the constant-hypoxia model is used to look at dormant *Mtb* that is maintained at a 0.2% oxygen tension in culture flasks (64). An important dormancy survival (Dos) regulon consisted of at least 48 genes appears to be essential for hypoxic survival in mycobacteria (65). It is regulated by the response regulator DosR. Previous studies have found that the ribosome-associated factor under hypoxia (RafH) controlled by DosR is the primary factor responsible for the ribosomal stability in hypoxic mycobacteria (66).

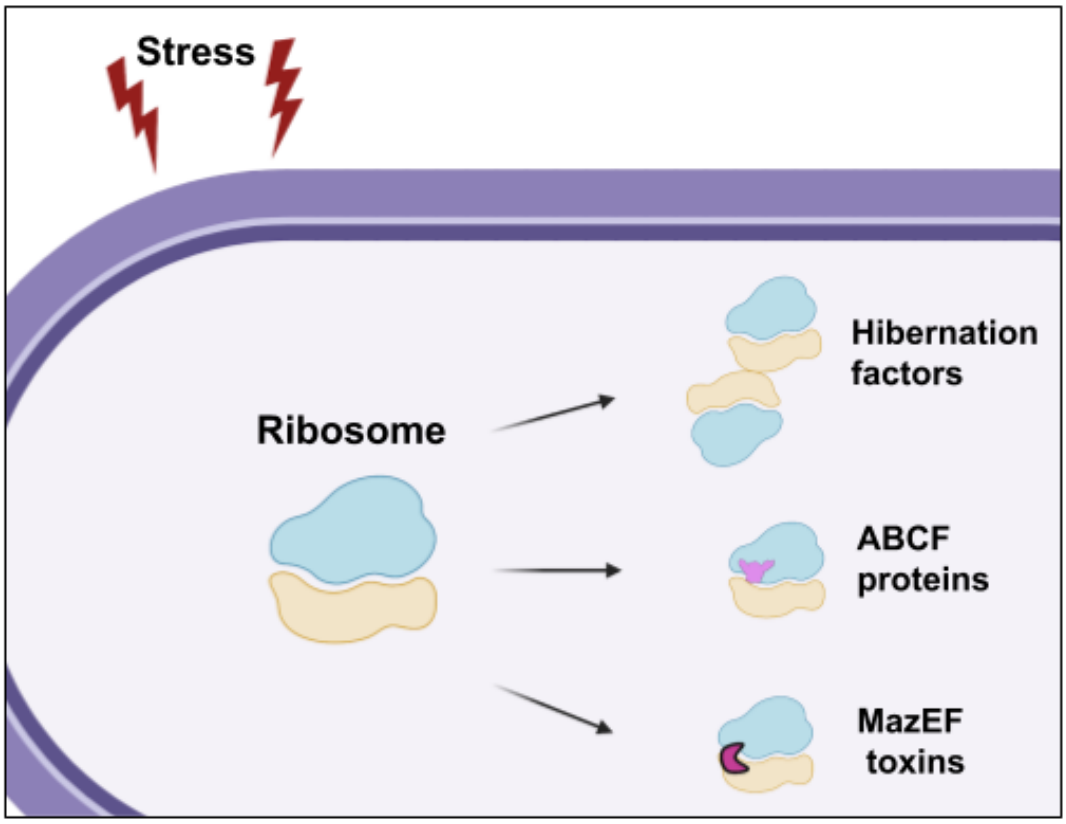


Figure 4 Ribosomal stress response in bacteria

Like RafH, which contains the conserved ribosome-interacting S30AE domain, mycobacteria-specific protein Y (Mpy) shows a unique aspect of translational regulation in *Mtb*. Mpy was previously named ribosome-associated factor during stasis (RafS), indicating its role during stationary growth (66). The dimerization of 70S ribosome (100S), often observed in stationary growth (67), is promoted by the binding of hibernation factors, which are homologs of Mpy. However, in mycobacteria, 100S is not present in the stationary growth phase, suggesting a different ribosomal inhibition mechanism by Mpy (66). Previous research reported that Mpy only binds the ribosome

when cellular zinc is depleted (68). Its function also requires the Mpy recruitment factor (Mrf). A follow-up study revealed that the hibernation of mycobacterial ribosome requires an active caseinolytic protease (Clp) system, where the Clp protease system destabilizes the zinc-bound Mrf. Therefore, zinc-depletion is required to maintain the presence of Mrf (69). However, further research is needed to explore Mpy and Mrf's biophysical properties and the detailed mechanism of Mrf-mediated recruitment of Mpy to the ribosome.

Mtb genome possesses an unusually high number of toxin-antitoxin modules, involving many different cellular processes (46). Toxin-antitoxin modules that belong to *the mazEF* family canonically act on the cleavage of selective mRNA, known as "mRNA interferases." (70) However, the intracellular toxin MazF from *Mtb* has an additional role as a translational inhibitor that cleaves 23S rRNA at the ribosomal A site (71). Also, the antitoxin gene *mazE-mt6* encoding the antitoxin MazE is up-regulated during nutrient starvation (72).

I.3 Antibiotic resistance in *Mtb*

I.3.1 Anti-TB drugs and resistance mechanism

An estimated 60 million lives were saved because of improved diagnosis and treatment since the WHO determined to end TB by 2030. However, drug resistance strains emerged due to inappropriate antibiotic usage, incorrect prescription, low quality, and premature treatment termination (73). Multidrug-resistant tuberculosis (MDR-TB) is a form of TB caused by bacteria resistant to at least rifampicin and isoniazid, the two

most effective first-line drugs. Extensively drug-resistant (XDR)-TB has also been identified as resistant to any fluoroquinolone and at least one of the injectable second-line drugs (74). Besides, some *Mtb* strains have been found resistant to all antibiotics available for testing, which are labeled as totally drug-resistant (TDR)-TB (75). About 206,030 cases of MDR-TB were reported in 2019, around a 10% increase from 2018. With only a 57% treatment success rate, MDR-TB remains a public health crisis and a health security threat (76).

First-line and second-line anti-TB drugs have been used clinically, targeting various molecular machinery *of Mtb*. Unfortunately, resistance has emerged toward each one of them (77) (Table 1).

Table 1 Anti-TB drugs

First-line drugs	Drug target	Resistance mechanism
Isoniazid	Enoyl reductase	Mutations in *katG* and *inhA* or its promoter region
Rifampicin	RNA polymerase	Mutations on 507–533 of the *rpoB* gene
Pyrazinamide	Pyrazinamidase	Mutations in the gene *pncA*
Ethambutol	Arabinosyl transferase	mutations at position *embB*306
Streptomycin	Ribosome	Mutations in *rpsL* and rrs
Second-line drugs	Drug target	Resistance mechanism
Quinolones	DNA gyrase	Mutations in *gyrA* or *gyrB*
Capreomycin	Ribosome	Mutations in the *tlyA* gene
Kanamycin/Amikacin	Ribosome	Mutations in 1400 and 1401 of the *rrs*
Ethionamide	Enoyl-ACP reductase	Mutations in *etaA*/*ethA*, *ethR, and inhA*
Para-aminosalicylic acid	Thymidylate synthase A	Mutations in *folC*

I.3.2 Ribosome is a successful antibiotic target

In the first-line anti-TB drugs, only streptomycin targets ribosomes. The resistance quickly appeared since streptomycin was first introduced, resulting in its absence in the routine treatment regime. However, the ribosome is a widely used drug

target. Almost half of the clinically prescribed antibiotics target different translational processes (78) (Table 2).

Table 2 Antibiotics that target the ribosome

Antibiotics targeting ribosome	Mechanism of action
Macrolides, Clindamycin, Streptogramins	Block the polypeptide exit tunnel on the 50S to prevent chain elongation
Chloramphenicol	Bind to PTC and inhibit the formation of peptide bond
Linezolid	Bind to 50S and prevent the formation of 70S
Tetracyclines	Bind to 30S and interfere with the binding of tRNA to the ribosome
Aminoglycosides	Bind to 30S and cause misreading of mRNA; Interfere with the interaction between mRNA and 30S in the initiation complex

I.3.3 Intrinsic resistance to macrolides in *Mtb*

The unsuccessful application of ribosome-targeting antibiotics to combat *Mtb* is owing to its complex cell wall structure and the intrinsic resistance mechanism. Notably, *Mtb* is intrinsically resistant to macrolides, a class of cornerstone drugs used in bacterial infections (79). The expression of gene *erm*(37), encoding a methyltransferase that functions on the macrolides binding pocket, is induced upon *Mtb* is incubated with macrolides. Interestingly, the induction mechanism of *erm*(37) in *Mtb* is different from

that of erm(C) in *Bacillus subtilis*. The mRNA of erm(C) contains a sequence encoding a leader peptide, which will trap the ribosome when the subinhibitory concentration of erythromycin is used. It will then change the secondary structure because of the stalled ribosome on its leader sequence, resulting in the exposure of the start codon of *erm*(C). Therefore, another ribosome will initiate the translation of protein Erm(C) (80). Nevertheless, no leader sequence is identified in the upstream of *erm*(37) mRNA, indicating a different mechanism. It is suggested that the whiB7 regulates the expression of *erm*(37). However, the detailed mechanism is unclear (81).

The specificities of both enzymes are also different. Erm(C) catalyzes the dimethylation of A2058 in 23S rRNA with high specificity (82). In contrast, Erm(37) primarily catalyzes the monomethylation of A2296 (A2058 in *E.coli* numbering) while potentially attaches additional methyl groups to the neighboring nucleotides A2295 and A2297 (83). Nonetheless, the consequence is that the induction of both methyltransferases leads to macrolides and ketolides resistance.

I.3.4 Direct ribosome protection proteins against antibiotics

In recent years, antibiotic resistance by the ATP-binding cassette F (ABCF) proteins through direct ribosome protection was brought into the limelight (84) (57) (85) (86) (87) (88). This previously recognized but underappreciated mechanism is fundamentally different from the antibiotic resistance mediated by the efflux pumps and ABC transporters. ABCF proteins belong to the ABC superfamily but lack the transmembrane domains. Different ABCF proteins are responsible for the resistance to a

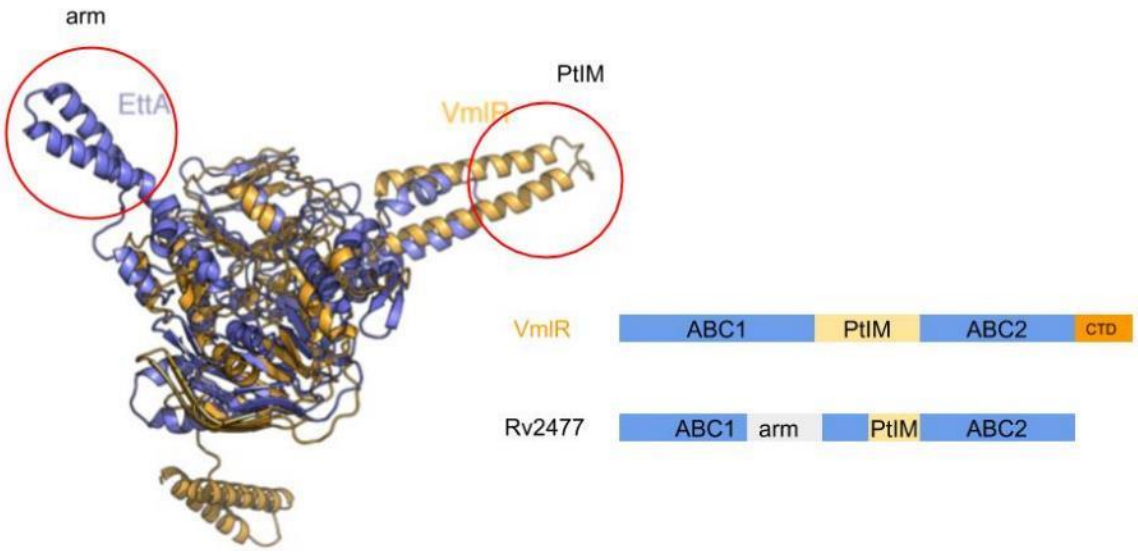


Figure 5 Structural differences between EttA and VmlR (antibiotic resistance ABCF)

The structure of EttA is superimposed with the structure of VmlR. The domain architecture is shown on the left. Circles indicate the general difference between EttA and antibiotic resistance ABCF.

specific class of antibiotics in many pathogenic bacteria (Table 3). Several biochemical and structural studies of antibiotic resistance ABCFs emphasize the importance of the P-tRNA interacting motif (PtIM), which inserts into the peptide exit tunnel to directly displace the bound antibiotics (85) (86).

Three ABCF proteins are found in *Mtb*, including Rv2477c, Rv1437, and Rv1667c/Rv1668c. However, they are homologs to energy-dependent translational throttle A (EttA) protein (89) (90). The significant difference between antibiotic resistance ABCF protein and EttA-like protein is the length of PtIM. Antibiotic resistance ABCFs have significantly longer PtIM than EttA-like ABCFs (Fig. 5). Therefore, whether EttA-like ABCFs also have antibiotic resistance properties is in question.

Table 3 Antibiotic resistance ABCFs

ABCF protein	Species	Resistance
Vga(A)	*Staphylococcus aureus*	streptogramins (group A), lincosamides, pleuromutilins
Vga(A)v	*Staphylococcus aureus*	
Vga(C)	*Staphylococcus aureus*	
Vga(E)	*Staphylococcus aureus*	
Vga(B)	*Staphylococcus aureus*	
Vga(D)	*Enterococcus faecium*	
Msr(A)	*Staphylococcus aureus*	macrolides, ketolides, streptogramins (group B)
Msr(C)	*Enterococcus faecium*	
Msr(D)	*Streptococcus pyogenes*	
Msr(E)	*Pasteurella multocida*	
VmlR	*Bacillus subtilis*	streptogramins (group A), lincosamides, pleuromutilins
Lsa(A)	*Enterococcus faecalis*	
Eat(A)v	*Enterococcus faecium*	
Lsa(C)	*Streptococcus agalactiae*	
Lsa(B)	*Streptococcus sciuri*	
Lsa(E)	*Streptococcus aureus*	
Sal(A)	*Streptococcus sciuri*	
OptrA	*Enterococcus faecium*	oxazolidinones, phenicols
LmrC	*Streptomyces lincolnensis*	specific to the antibiotic produced by each species?
VarM	*Streptomyces virginiae*	
OleB	*Streptomyces antibioticus*	
TlrC	*Streptomyces fradiae*	
SrmB	*Streptomyces ambofaciens*	
CarA	*Streptomyces themotolerans*	

I.4 Introduction and development of cryo-electron microscopy

We employed cryo-electron microscopy (cryo-EM) to study the *Mtb* ribosome's structure and its interacting factors to shed light on the molecular mechanism of translational regulation. We also provided the blueprint for structure-based drug design by solving high-resolution cryo-EM structures of *Mtb* ribosome bound with novel

antibiotics. Unlike X-ray crystallography and nuclear magnetic resonance (NMR), cryo-EM directly visualize macromolecules and computationally reconstruct them into three-dimensional (3D) structures. Just as Richard Feynman once said, “it is very easy to answer many of these fundamental biological questions; you just *look at the thing*!” (91)

Ever since the invention of the light microscope in the 17th century, tremendous scientific discoveries have been achieved, from discovering microscopic pathogens to super-resolution light microscopy (92). However, the resolution of a light microscope is limited by the light wavelength, which is hundreds of nanometers. Although super-resolution light microscopy can exceed the physical resolution of a light microscope to a couple of nanometers, it is still far behind the resolution required and not suitable for studying the 3D structures of macromolecules inside a cell. In contrast, electron microscopes, which use accelerated electrons to produce images rather than light, provide sufficient physical resolution to visualize biomolecules at atomic resolution.

In order to image macromolecules under an electron microscope, several methodological barriers have to be overcome. First, the biological samples are susceptible to high-energy electrons, potentially damaged by the electron beam. An early attempt to address this problem was using heavy metal solutions to stain samples before imaging (93). The staining agent like uranyl acetate will cover and cast the sample, generating a high-contrast image that appears reversed. This method is referred to as negative staining, which reduces the radiation damage problem. However, due to the heavy atom salt's grain size, the resolution limit of negative staining is around 20 Å (94), far from resolving even the secondary structure. Ken Taylor and Robert Glaeser

demonstrated the revolutionized approach to address the radiation problem in 1974 (95). By preserving samples in the liquid nitrogen temperature, the frozen-hydrated samples are more resistant to radiation damage and preserved with high-resolution features. Although the idea of reducing radiation damage using ultra-low temperature was pioneered in X-ray crystallography (96), this work initiated the field of cryo-EM. Following the discovery, Dubochet and colleagues developed a practical method for vitrification of samples (97) (Fig 6). A sample is first applied on the cryo-EM grids and then blotted away by filter papers, leaving a thin layer of the sample. The thin layer of the sample is then rapidly plunged into liquid ethane. Due to the large heat capacity and close to melting point temperature (-188 ^{0}C) of liquid ethane, the water in the thin layer of the sample could not form crystalline ice, resulting in a vitrified state. The sample's vitrification is critical for imaging, as it will prevent the sample from being damaged by crystal ice formation and significantly decrease the background noise. Although there are different commercially available cryo-EM sample preparation apparatuses nowadays, the concept of rapid-freezing follows Dubochet and colleagues' research 30 years ago. The improvement of modern cryo-EM sample preparation focuses on the sample deposition methods, high-throughput, and time-resolving capability (98) (99) (100).

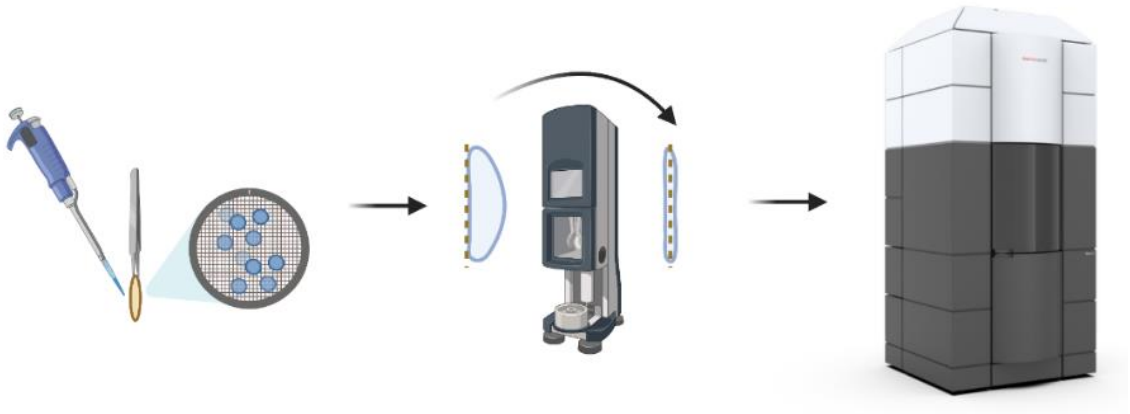


Figure 6 Cryo-EM sample preparation using vitrobot

The second obstacle was the low-contrast image obtained using the electron microscope. Due to biological samples' inherent properties as "soft matters," they inefficiently diffract the electrons. Although the samples are preserved at liquid nitrogen temperature, they are still vulnerable to the electron beam. Therefore, only a low electron dose can be used to image the sample before any radiation damage, resulting in even lower contrast. In the 1970s-1980s, Richard Henderson and his colleagues worked around this problem by studying the 2D crystals of bacteriorhodopsin, using electron diffraction patterns to calculate one of the first 3D atomic structures membrane protein (101). However, this method is not generally applicable as the majority of proteins cannot form 2D crystals. Another way to improve the quality of an EM image is to develop better recording devices. Historically, EM images were recorded on photographic films, with labor-intensive post processes like film development and scanning. The usage of CCD cameras later greatly improved productivity, however, at the expense of image quality. The direct electron detectors (DEEs) (102) were recently

developed with high detective quantum efficiency (DQE) to enhance the signal-to-noise ratio. The DEEs can also record images at an ultra-fast frame rate, about 1,500 frames per second with the latest K3 camera. With the capability to record a movie instead of a single image, the motion correction algorithm was first introduced by Niko Grigorieff in 2012 to improve the quality of EM images (103).

The third challenge was the development of a robust 3D reconstruction algorithm. The first 3D reconstructed structure from EM images was the T4 bacteriophage tail, by David DeRosier and Aaron Klug using helical reconstruction in 1968 (104). In the late 1970s, Joachim Frank and colleagues proposed the idea of "single-particle" reconstruction for the protein samples displaying randomized orientations on an EM grid (105) (106). The key to single-particle cryo-EM 3D reconstruction was developing computational tools to define the relative orientations of the 2D projection images. Different algorithms have been implemented to tackle the problem. Pawel Penczek and Frank proposed the "projection matching" approach, where each 2D particle image is compared with different views of a 3D reference based on a cross-correlation criterion (107). Another method to determine the orientation of 2D projections is based on the “common line” theorem, in which each pair of 2D projections shares a “common line” in the 3D Fourier space. Since then, different software packages have been developed, including Spider, Imagic, and EMAN (108) (109) (110). Frealign focuses on the 3D refinement procedure (111). More recently, RELION (112) and cryoSPARC (113) are the dominant software in the field due to their 3D classification

capability to separate heterogeneous conformations and GPU-acceleration to significantly reduce the computing time.

Apart from overcoming the bottlenecks mentioned above, another significant improvement is the high-throughput data collection for single-particle cryo-EM. With the ultra-fast camera and well-implemented data collection software like SerialEM (114), Leginon (115), and EPU (Thermo Fisher Scientific), the daily outcome could reach ~8,000 movie stacks. With current software and hardware set up, getting thousands of micrographs for different specimens in a day is within reach. Therefore the pre-screening of cryo-EM grids in a low-end electron microscope remains necessary for efficient data collection. The newly released 100kV electron microscope Tundra is suitable for the need and provides a path to democratize cryo-EM.

As single-particle cryo-EM is gaining popularity, rapidly overtaking X-ray crystallography as a mainstream method for protein structure determination, the development of cryo-electron tomography (cryo-ET) is also substantial (116) (117). In cryo-ET, the same area of the specimen is imaged at different tilting angles to obtain information from various orientations. 3D reconstruction of the imaged area is then performed by back projecting all the 2D images. All the hardware developed for single-particle cryo-EM applies to cryo-ET, except the sample preparation procedure. As cryo-ET is mainly used to provide the structural information in situ, the thickness of a thin-sliced and frozen-hydrated cell is critical for the images' quality. Nowadays, cryo-focused ion beam (cryo-FIB) milling is used to generate the lamellae (118). Although it is still labor-intensive and a work-of-art to prepare lamellae, the field has already started

implementing automated cryo-FIB (119)—for example, the Aquilos 2 from Thermo Fisher Scientific. The cryo-ET's data collection routine has also been updated, including using a tilted beam to rapidly collect tomograms (120). Although it is well-accepted that resolving atomic structures in situ is the future of structural biology, many technical barriers are needed to be overcome.

CHAPTER II

THE MECHANISM-OF-ACTION OF AN ABCF PROTEIN MTBETTA

II.1 Introduction

ATP-binding cassette F (ABCF) proteins are widely spread among bacteria and eukaryotes, with dozens of distinct groups (57) (121). Although they belong to the universal ABC transporter superfamily, the lack of transmembrane domains makes them unlikely to function as transporters (122) (123) (87). Indeed, a variety of ABCF proteins are found to be involved in different aspects of protein synthesis. The early search of ABCFs revealed that one ABCF, the eukaryotic Elongation Factor 3 (eEF3), has potential roles in E-site tRNA release and ribosome recycling (124) (125). Later, more ribosome-associated ABCFs were discovered in eukaryotes, many of which have been proposed to be involved in the stress response (126), translation initiation, ribosome biogenesis, and quality control (127) (128) (86). In prokaryotes, the first structurally determined ABCF protein, EttA from *Escherichia coli*, was found to bind to the ribosome and "throttle" the translation depending on the ATP/ADP ratio (90) (89). Interestingly, a subset of bacterial ABCF proteins has been identified to reduce the efficacy of a diverse range of clinically used antibiotics and are classified into different groups according to the types of antibiotics they resist (123) (87). In recent years, direct protection of ribosome against antibiotics by antibiotic resistance (ARE) ABCFs has stepped into the limelight, which is supported by biochemical and structural studies of

various AREs, disfavoring the previous idea that they function as efflux pumps (123) (84) (86) (85).

ABCFs are generally composed of two tandem nucleotide-binding domains (NBDs) connected by a linker of ~60-100 amino acids (129) (57). Despite various lengths in the linker region, distinct variations among ABCFs include an extra N-terminal domain, the insertion of an "Arm" domain within the first NBD (NBD1), as well as a C-terminal extension (57). Like other members in the ABC superfamily, most ABCFs are ATPases with highly conserved ATP binding/hydrolysis motifs, albeit with variations of the aromatic residue at the A-loop within NBD1 (90). The two NBDs interact with each other in a head-to-toe fashion, sandwiching two ATP molecules. Movement between NBDs during ATP hydrolysis, referred to as a "clamping motion," is anticipated based upon available structures of different ABCs in different nucleotide-bound states (130) (131) (132). The ATPase activity is essential for their functions (133) (134) (84). However, the knowledge of the cooperativity between the two ATP binding sites and the effect on its function is limited.

Structural investigations about the mechanism-of-action of ABCFs have been primarily focused on AREs, such as MsrE (resistant to macrolides, streptogramin B, ketolides) and VmlR (virginiamycin M and lincomycin resistance) (86) (85), and the non-ARE ABCF, the *E. coli* EttA (90) (89). The mechanism of antibiotic resistance by MsrE is proposed as direct dispersion of the antibiotics from the ribosome, due to the steric clash between the tip of the linker and the antibiotics, as well as the conformational change of CCA tail and acceptor stem of P-site tRNA upon the binding

of MsrE, both emphasizing the importance of the length of the linker. However, structural and mutational studies of VmlR also provided evidence of allosteric effect toward drug binding, which could serve as an alternative mechanism for reducing the drug efficacy (85). Of note, the linker of most AREs is significantly longer than that of EttA-like ABCFs. A crystal structure of EttA dimer in the nucleotide-free state and a low-resolution cryo-EM structure of a monomeric EttA mutant in complex with the ribosome in the ATP-bound state are available (89) (90). Nevertheless, the detailed understanding of the structure-function relationship of EttA is limited due to the lack of critical structures during its ATP hydrolysis. Therefore, high-resolution structures of EttA at different states during ATP hydrolysis and its interaction with the ribosome are needed to better understand its function.

Rv2477c from *Mycobacterium tuberculosis* (*Mtb*) is an EttA-like ABCF protein (135), which was previously mistaken as an efflux pump (136) (137). Unlike EttA from *E. coli*, the *Mtb* Rv2477c (termed as MtbEttA from now on) is essential based on saturating transposon mutagenesis (138) and has been found to respond to antibiotic treatment (139) (140). MtbEttA could be a potential drug target in *Mtb*, the pathogen leading to ~2 million deaths worldwide annually (5). We used *in vitro* translation assay to evaluate the function of MtbEttA (Fig. 3). Although MtbEttA cannot provide erythromycin resistance, the increased translational activity at a low concentration of erythromycin indicates translational enhancement. Indeed, studies of EttA from *E. coli* suggest that it can promote the first peptide bond formation on the ribosome. To better understand how MtbEttA enhances the translation, we conducted the structural analyses

between the *Mtb* ribosome and MtbEttA along its trajectory of ATP hydrolysis. We determined high-resolution cryo-EM structures of *Mtb* ribosomes in complex with MtbEttA in the pre-hydrolysis (ADPNP) and transition (ADP-VO_4) states. We also solved a crystal structure of MtbEttA alone in the post-hydrolysis (ADP) state. We observed an asymmetric engagement of conserved motifs around the two ATP-binding sites and the alternation of stereochemistry around peptidyl transferase center (PTC), both of which are correlated with the intersubunit rotation in the pre-hydrolysis state. Different degrees of ribosomal intersubunit rotation is observed between the pre-hydrolysis and the transition states. Moreover, the crystal structure of domain-swapped MtbEttA dimer with ADP is in an "open" conformation, in which the two NBDs are apart from each other, not able to tightly encapsulate the nucleotides. These structures reveal the interplay between the ribosome and MtbEttA during its ATPase cycle.

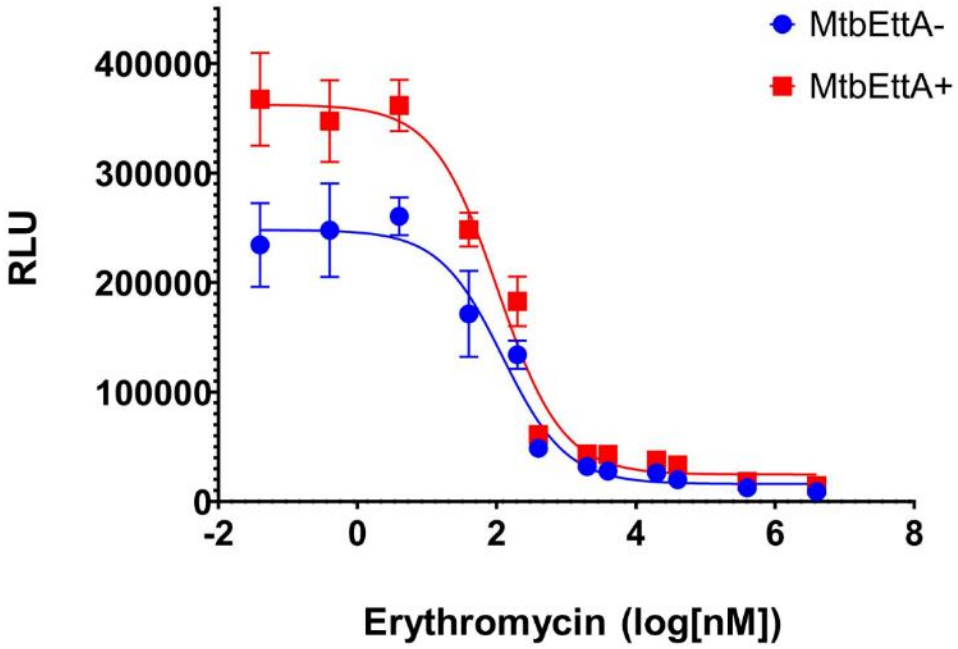


Figure 7 Effect of MtbEttA on erythromycin IC50 using in vitro translation assay

Relative light units (RLU) are recorded due to the production of nanoluciferase in the in-vitro tranlsation system. A series of erythromycin concentration (0.04nM-4mM) is used with (red) or without (blue) MtbEttA to estimate the IC50. IC50 of erythromycin with or without MtbEttA is around 104nM and 122nM, respectively.

II.2 Materials and Methods

II.2.1 Protein expression and purification

The full-length sequence of MtbEttA was cloned into a modified pET28(a) vector with an N-terminal His6-SUMO tag. We then overexpressed full-length MtbEttA in Rosetta™ 2(DE3) cells and purified using Ni-NTA affinity chromatography. The

SUMO tag was then cleaved by SUMO protease and removed by Ni-NTA column, leaving MtbEttA with one extra serine at N terminus. Further purification was done using gel filtration chromatography by Superdex 200 (16/60 GL) column (GE Healthcare) to separate dimer and monomer populations.

Purification of *Mtb* ribosome was done according to a previously established protocol (141). Briefly, *Mtb* cells MC27000 were grown in 7H9 medium supplemented with 10% oleic albumin dextrose catalase (BD), 0.5% glycerol, 0.05% Tween-80, and 50 μg/ml pantothenic acid at 37°C until an OD600 of 1.0. The following procedures were performed at 4°C. *Mtb* ribosome 70S was purified according to modified protocols. After the cells were lysed in buffer (20 mM Tris-HCl [pH 7.5], 100 mM NH_4Cl, 10 mM $MgCl_2$, 0.5 mM EDTA, 6 mM 2-mercaptoethanol), it was clarified by centrifugation at 30,000 × g for 1 hr. The supernatant was pelleted in sucrose cushion buffer (20 mM HEPES [pH 7.5], 1.1 M sucrose, 10 mM $MgCl_2$, 0.5 M KCl, and 0.5 mM EDTA) at 40,000 rpm in a Beckman Type 45Ti rotor for 20 hr. The pellet was resuspended in the buffer containing 20 mM Tris-HCl (pH 7.5), 1.5 M $(NH_4)_2SO_4$, 0.4 M KCl, and 10 mM $MgCl_2$. The suspension was then applied to a hydrophobic interaction column (Toyopearl Butyl-650S) and eluted with a reverse ionic strength gradient from 1.5 M to 0 M $(NH_4)_2SO_4$ in the buffer containing 20 mM Tris-HCl (pH 7.5), 0.4 M KCl, and 10 mM $MgCl_2$. The eluted ribosome peak was changed to re-association buffer (5 mM HEPES-NaOH [pH 7.5], 10 mM NH_4Cl, 50 mM KCl, 10 mM $MgCl_2$, and 6 mM 2-mercaptoethanol) and concentrated before loading on top of a 10%–40% linear sucrose

gradient centrifuged in a Beckman SW28 rotor at 19,000 rpm for 19 hr. The 70S fraction was concentrated to about A260 = 300 after the removal of sucrose.

II.2.2 In-vitro translation assay

The assay used to measure in vitro ribosome activity relied on the production of nanoluciferase in an *Mtb*-based cell-free system. *Mtb* S30 cell-free extract was prepared from *Mtb* MC27000, and equilibrated in S30 buffer (10 mM tris-acetate pH 8.2, 14 mM magnesium acetate, 60 mM potassium acetate, 1 mM DTT). 10 μL of S30 extract (containing 200 nM ribosome) was mixed with 5 μL 10X salt buffer (2 M potassium glutamate, 0.8 M ammonium acetate, and 0.16 M magnesium acetate), 1 mM each of the 20 amino acids and 33 mM phosphoenolpyruvate to a final volume of 41 μL. We then added the erythromycin of various concentrations (2 μL) to the 41 μL mixture, and after the incubation for 10 minutes at RT, the reaction was supplemented with MtbEttA, nanoluciferase mRNA (200 ng), and 5 μL 5X master mix (286 mM HEPES-KOH, pH7.5, 6 mM ATP, 4.3 mM GTP, 333 μM folinic acid, 853 μg/mL tRNA). The final volume was 50 μL in each well of the 384-well plate. The reaction proceeded for 40 mins at 37 degrees and then was terminated by the addition of 80 μM Chloramphenicol. The luminescent signal was detected by the addition of 20 μL of the nanoluciferase substrate Furimazine (Promega). Nanoluciferase mRNA was prepared from an in vitro transcription assay.

II.2.3 Cryo-EM sample preparation and data collection

Mtb ribosome in complex with MtbEttA at pre-hydrolysis state was prepared with 0.2 μM 70S, 0.2 μM modified Z4C mRNA (AGAAAGGAGGUAAAACAUGUUCAAAA) (142), 0.8 μM fMet-tRNA^{fMet}, 33 μM MtbEttA, and 2 mM ADPNP. Firstly, the 70S initiation complex (70SIC: 70S-fMet-tRNA^{fMet}-mRNA) was formed by incubating 70S with mRNA in buffer A (50mM HEPES pH 7.5, 50mM KCl, 10mM NH4Cl, 5mM MgCl2, 6mM β-mercaptoethanol) at 37 °C for 30 min, and then supplied with fMet-tRNA^{fMet} for another 30 min at 37 °C. In parallel, MtbEttA was mixed with ADPNP at 37 °C for 1 h. The final complex was formed by combining two mixtures and bringing the Mg^{2+} concentration to 10mM, incubated at 37 °C for 1 h, and then kept on ice. *Mtb* ribosome in complex with MtbEttA at transition state was prepared similarly, except that 4 mM ADP and 4 mM NaVO4 were used instead of ADPNP. Cryo-EM specimens were then prepared by applying 3 μL of the freshly reconstituted complex to a glow-discharged Quantifoil 2/1 200-mesh Holey Carbon Grid coated with 2-nm continuous carbon and vitrified using a Vitrobot Mark III (FEI Company) at 22 °C with 100% relative humidity.

Cryo-EM images of the 70SIC-MtbEttA-ADPNP complex were recorded under a Titan Krios microscope (FEI Company) operated at 300 kV (SLAC). Data were collected using EPU on a K2 Summit direct detection camera (Gatan) in the electron counting mode with a pixel size of 1.06 Å. Beam shift was enabled to encompass 5 exposures per hole. The beam intensity was adjusted to a dose rate of 5 e^- and per pixel per second on the camera. A 30-frame movie stack was recorded for each exposure with

0.2 s per frame for a total exposure time of 6 s. An in-column energy filter was used with a slit width of 20 eV.

Similarly, 70SIC-MtbEttA-ADP-VO_4 images were recorded under a Titan Krios microscope (FEI Company) operated at 300 kV (UTHSC). Data were collected using EPU on a K2 Summit direct detection camera (Gatan) in the electron counting mode with a pixel size of 1.063 Å. Beam shift was enabled to encompass 4 exposures per hole. The beam intensity was adjusted to a dose rate of 6.5 e^- per pixel per second on the camera. A 30-frame movie stack was recorded for each exposure with 0.2 s per frame for a total exposure time of 6 s. An in-column energy filter was used with a slit width of 20 eV.

II.2.4 Image Preprocessing

Drift correction of collected movie stacks was done using MotionCor2. The defocus value of each aligned micrograph was determined using Gctf (143). Micrographs with visible contamination and poor power spectrum were discarded. Automatic particle picking was done by gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/). 3D reconstruction was done by following the pipeline of Relion3 (144).

The data processing for the 70SIC-MtbEttA-ADPNP sample was as follows (Fig. 4): In total, 1,175,176 particles were selected from 8,949 micrographs and binned by 8 before subjecting to 2D classifications in Relion to remove bad particles. After 2D classifications, 1,016,405 particles were selected and binned by 4. The initial 3D

refinement step was performed with all clean particles to get a consensus map. 3D classification using the consensus map as an initial model and skip_align option was used to classify different subpopulations in the dataset. Two out of eight classes were found with MtbEttA density. 196,100 particles were selected after an additional cleaning step, using focused 3D classification without alignment around the MtbEttA and P-site tRNA region.

Further 3D classification of these particles with skip_align option and 30S mask yielded 3 subpopulations, including Pre_R0 state with 126,715 particles and Pre_R1 state with 34,158 particles. We continued to refine Pre_R0 and Pre_R1 states with non-binned data to the resolution of 2.97 Å and 3.23 Å, respectively. Different subpopulations, including 70S-P/PtRNA, 70S-P/EtRNA, 70S-P/PtRNA-E/EtRNA, were also refined to 2.76 Å, 2.8 Å, and 2.71 Å, respectively.

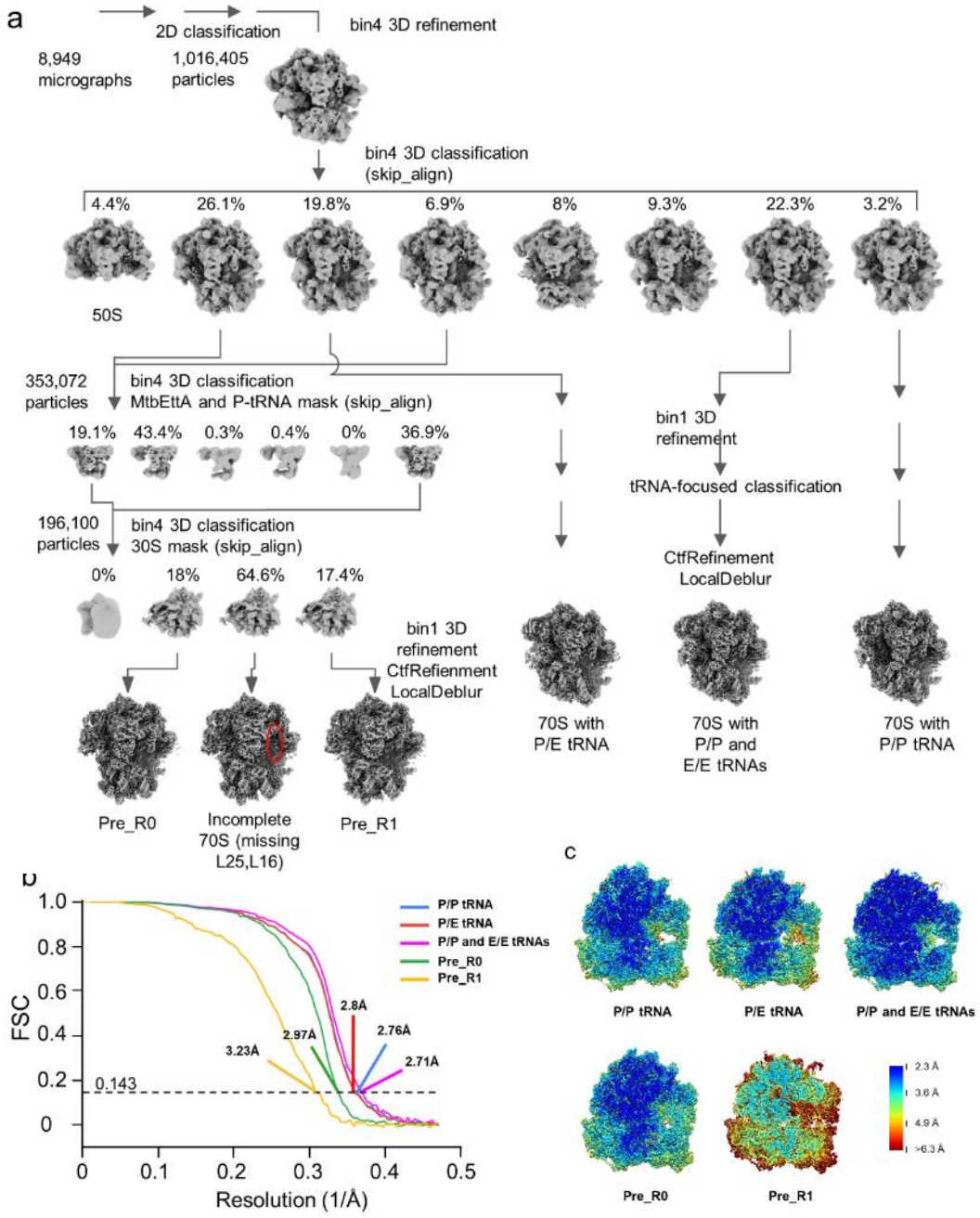


Figure 8 Cryo-EM data processing workflow for 70SIC-MtbEttA with ADPNP.

a, Single-particle data processing strategy for 70SIC-MtbEttA with ADPNP. Red cycle indicates missing bL25 and uL16 density. b, FSC curves of the reconstructed cryo-EM maps. Gold-standard was used to estimate the resolution (FSC=0.143). c, Cutaway view of the cryo-EM maps, which are colored by local resolution.

Similar to the procedure above, 70SIC-MtbEttA-ADP-VO4 data was processed accordingly (Fig. 5): After 2D classifications, 742,504 particles from 8,666 micrographs were selected. 3D classification without alignment was performed using a consensus map obtained from an initial 3D refinement step. Two out of six classes were found with ambiguous MtbEttA density. 125,420 particles were selected after an additional cleaning step, using focused 3D classification without alignment around the MtbEttA and P-site tRNA region. Further 3D classification of these particles with skip_align option and 30S mask yielded 3 subpopulations, including Trans_R0 state with 86,692 particles and Trans_R1 state with 32,731 particles. We continued to refine Trans_R0 and Trans_R1 states with non-binned data to the resolution of 2.79 Å and 3.1 Å, respectively.

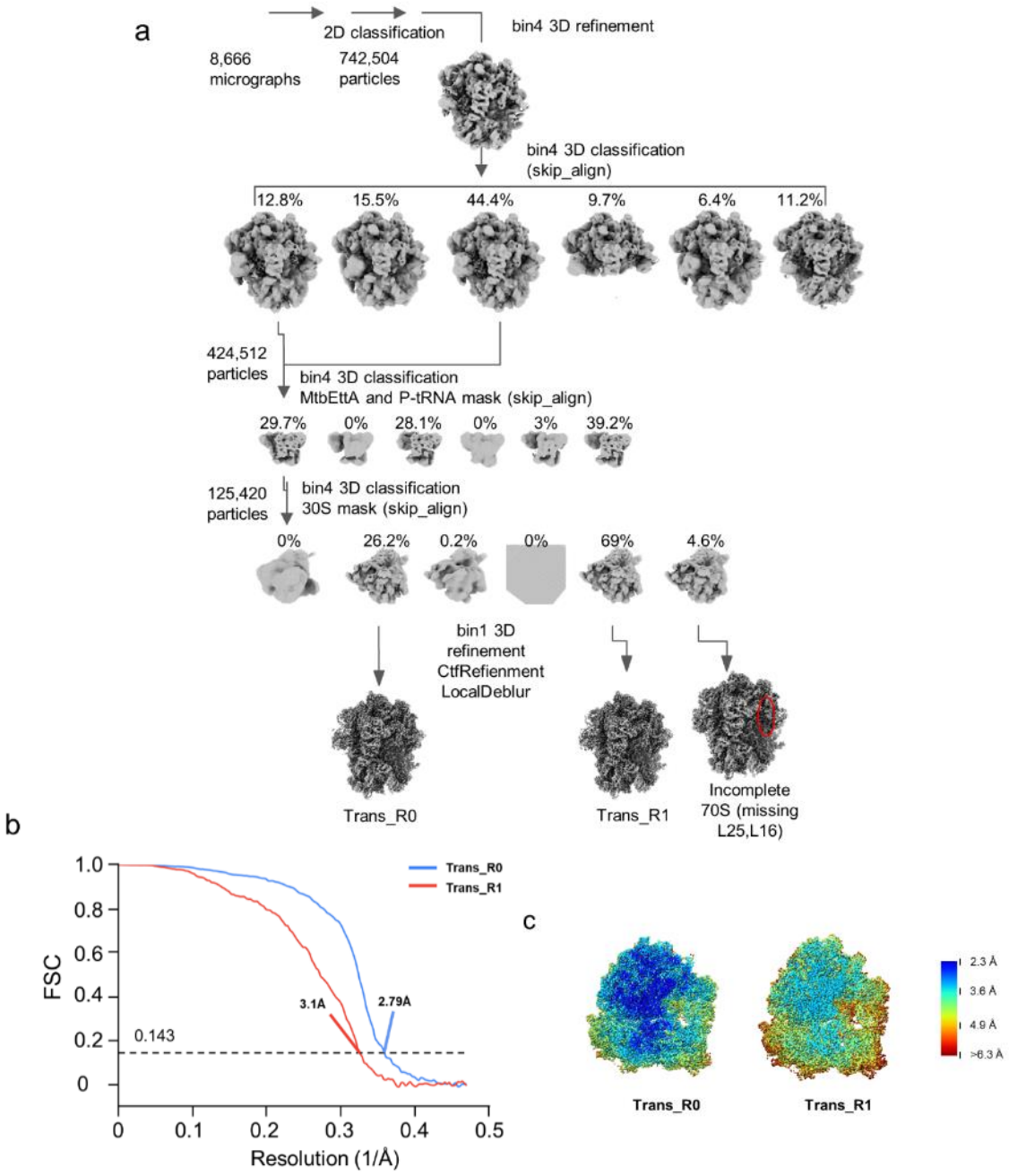


Figure 9 Cryo-EM data processing workflow for 70SIC-MtbEttA with ADP-VO_4.

a, Single-particle data processing strategy for 70SIC-MtbEttA with ADP-VO_4. Red cycle indicates missing bL25 and uL16 density. b, FSC curves of the reconstructed cryo-EM maps. Gold-standard was used to estimate the resolution (FSC=0.143). c, Cutaway view of the cryo-EM maps, which are colored by local resolution.

II.2.5 Resolution Estimation and Post Processing

The overall resolution was assessed using the gold-standard criterion of Fourier shell correlation (145), with a cutoff at 0.143, between two half-maps from two independent half-sets of data. Local resolutions were estimated using Resmap (146). Post-processing was done by LocalDeblur (147).

II.2.6 Crystallization, data collection, and structure determination

Purified full-length MtbEttA was concentrated to ~46 mg/ml before crystallization. The final concentration of 4 mM ADP and 4 mM $MgCl_2$ was mixed with the protein for 1 h at room temperature. Crystals were observed at multiple conditions at 16 ^oC with initial screening against ~600 conditions, using the sitting-drop vapor diffusion set by a Mosquito Crystal liquid handler (TTP Labtech Inc). The diffraction quality of crystals was checked at the in-house x-ray source. Further optimization was performed with the hanging drop vapor diffusion method, yielding The best crystal was produced by hanging drop at condition 0.1 M MES pH 6.7, 0.2 M $MgCl_2$, 10% v/v PEG 4000, at 16 ^oC.

Diffraction data were collected at the Advance Photon Source, Argonne National Laboratory in Chicago. The crystal was diffracted to ~2.5 Å and was processed by HKL2000 (148). The diffraction data were scaled and truncated to 2.85 Å using imosfilm (149). Molecular replacement was performed with AutoMR (150)in the PHENIX package, using two NBDs models of MtbEttA calculated from SWISS-

MODEL (151) based on E. coli EttA crystal structure (PDB 4FIN). The model was iteratively refined and manually built with PHENIX(152) and COOT (153).

II.2.7 Modeling and Visualization

To build the atomic model for 70SIC-MtbEttA-ADPNP and 70SIC-MtbEttA-ADP-VO_4 complexes, we first fit our previous *Mtb* 70S ribosome structure (PDB 5V93) and MtbEttA monomer structure obtained from SWISS-MODEL into the high-resolution cryo-EM map as the rigid body using University of California San Francisco (UCSF) Chimera (154). Model refinement was performed by real-space refinement in PHENIX. The RNA geometry optimization was done by ERRASER. Manual model building was done with COOT to inspect and improve local fitting. Moreover, the iterative process involving refinement and manual building was conducted to achieve the best model. The same model building procedure was done for 70S-P/PtRNA, 70S-P/EtRNA, and 70S-P/PtRNA-E/EtRNA complexes. All of the figures and movies were made using UCSF Chimera and ChimeraX.

II.3 Results

II.3.1 Cryo-EM structures of *Mtb* ribosome with MtbEttA at pre-hydrolysis (ADPNP) and transition (ADP-VO4) states

To obtain a stable complex between the *Mtb* ribosome and MtbEttA, we examined different conditions involving the presence of mRNA, tRNA, and different nucleotide analogs (Fig. 6). Similar to previous results, P-site tRNA and non-

hydrolyzable ATP analog ADPNP are necessary for forming the complex (89) (85). Also, we discovered that MtbEttA can still bind to 70S in the presence of ADP-VO_4, an analog mimicking the transition state of ATP-hydrolysis. To investigate the Mtb ribosome structures in complex with MtbEttA at the pre-hydrolysis and transition states, we reconstituted 70SIC (70S Initiation Complex)-MtbEttA-ADPNP and 70SIC-MtbEttA-ADP-VO_4 complexes *in vitro,* respectively. For the pre-hydrolysis state of MtbEttA, 3D classification and refinement yielded two distinct subpopulations, namely Pre_R0 and Pre_R1, to the resolution of 2.97 Å and 3.23 Å, respectively (Fig. 4). Different subpopulations, including the 70S with P/P tRNA, P/E tRNA, or P/P and E/E tRNAs, were also resolved to 2.76-Å, 2.8-Å, 2.71-Å resolution, respectively. Like the pre-hydrolysis state, two distinct subpopulations were revealed at the transition state of MtbEttA, namely Trans_R0 and Trans_R1, to the resolution of 2.79 Å and 3.1 Å, respectively (Fig. 5). Local resolution analysis showed even higher resolution at the core region of the 70S, which enabled us to identify several conserved RNA modifications in 23S rRNA and 16S rRNA (Fig. 7).

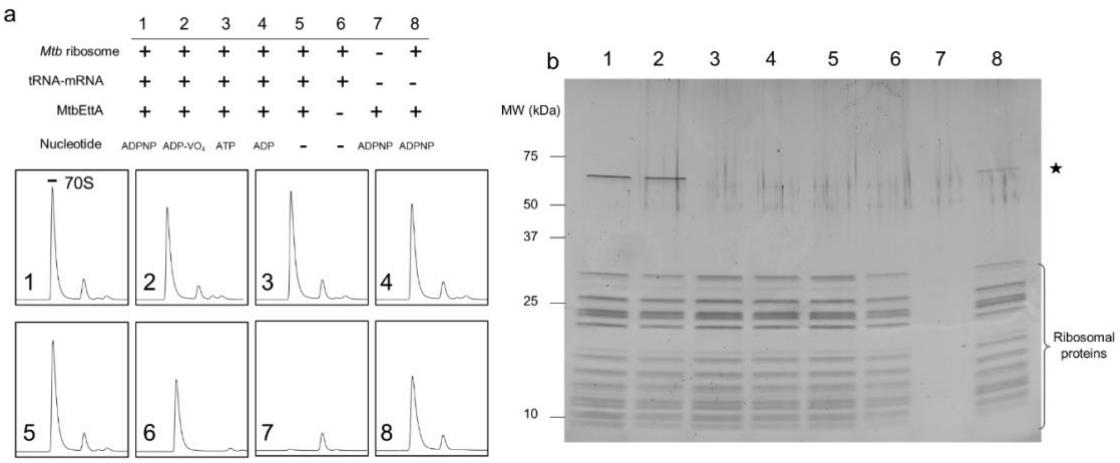


Figure 10 Analysis of the condition for forming stable MtbEttA-ribosome complex.

a, Experiment conditions are listed in the table, with "+" and "–" as included and excluded, respectively. Chromatogram for each condition is shown in the box, with 70S peak labelled in 1. b, SDS-PAGE showing the composition of 70S peak in corresponding condition. The gel is stained by SYPRO™ Ruby. MtbEttA is indicated by the star.

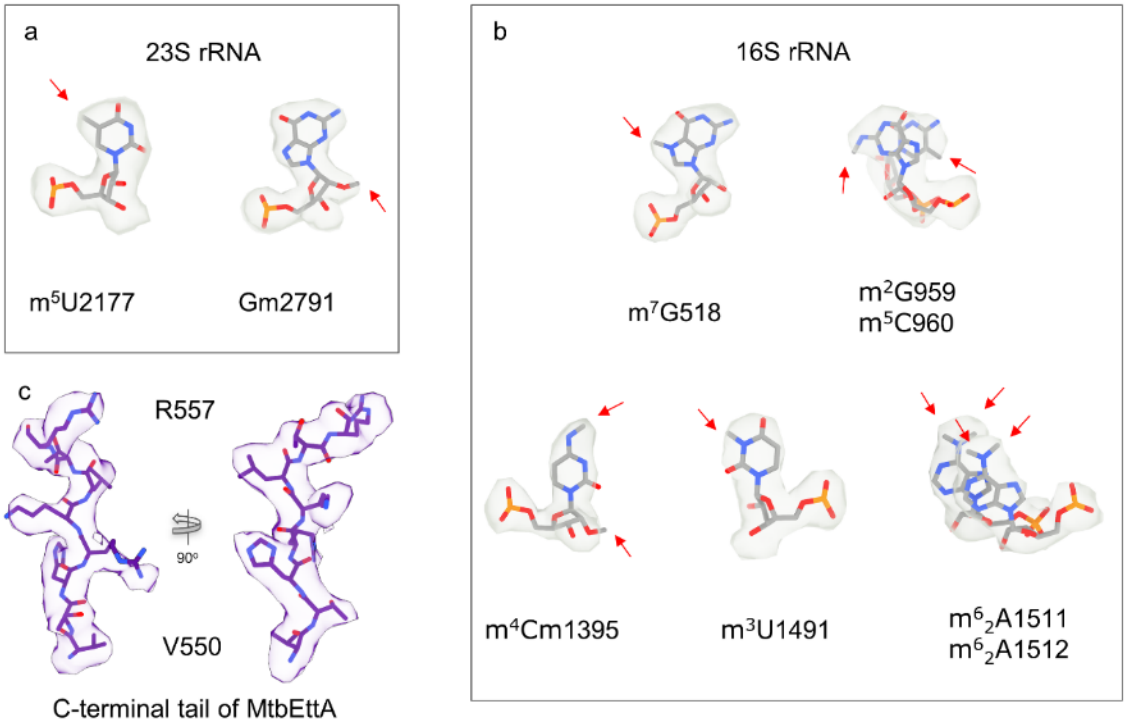


Figure 11 Examples of cryo-EM density.
a,b, RNA modifications in ribosomal 23S rRNA and 16S rRNA in 70S with P/P and E/E tRNAs. Cryo-EM density is shown in transparent lightgray and modification sites are labelled with red arrows. c, Transparent indigo cryo-EM density of C-terminal tail of MtbEttA in Pre_R0 state.

MtbEttA resides at the E-site of the 70S and interacts with several ribosomal proteins, RNAs, and the P-site tRNA. The linker region, termed P-site tRNA interaction motif (PtIM) as for EttA Boel, 2014 #153] (89), has extensive interactions with the P-site tRNA and 23S (Fig. 8a,b, 2b). The global structural difference between pre-hydrolysis and transition states lies in the degree of intersubunit motion between the 30S and 50S. To assess the relative motion between the two subunits, we aligned MtbEttA-

ribosome complex structures to the structure of 70S with P/P tRNA, a classic non-rotated conformation, based on the 23S rRNA. For pre-hydrolysis states, the Pre_R0 30S has a subtle body rotation of ~0.2° and head swivel of ~0.8° relative to the non-rotated state, while the Pre_R1 30S has a body rotation and head swivel of ~4.7° and ~4.1°, respectively (Fig. 8c). In the transition states, the Trans_R1 30S shows a body rotation of ~3.4° and a head swivel of ~1.7°, which shows a larger head movement comparing to the body of the 30S when MtbEttA transits from pre-hydrolysis to transition state. The Trans_R0 state shared a similar motion as the Pre_R0 state, with body rotation of ~0.3° and head swivel of ~1.0°. Nevertheless, the rotational movement of the 30S in the presence of MtbEttA both at pre-hydrolysis and transition states is modest comparing the spontaneously fully rotated state as shown in the 70S with P/E tRNA, which displays ~10.4° body rotation and ~11.6° head swivel (Fig. 8c).

The tip region of PtIM extends into the proximity of the PTC and shows stable interactions with P-site tRNA when the 30S has slight movement as in the Pre_R0 and Trans_R0 states (Fig 8c). However, destabilized interaction and increased flexibility between PtIM and P-site tRNA are observed and related to the extent of 30S movement. The tip region of PtIM and CCA tail of the tRNA display considerable flexibility in the Pre_R1 state, inferred from the cryo-EM maps' resolvability (Fig 8b). However, only subtle PtIM and tRNA variations are detected in the Trans_R1 state, presumably due to the smaller head swivel in the 30S. Therefore, it shows a correlation between the head domain's motion and the plasticity of the PtIM and the P-site tRNA.

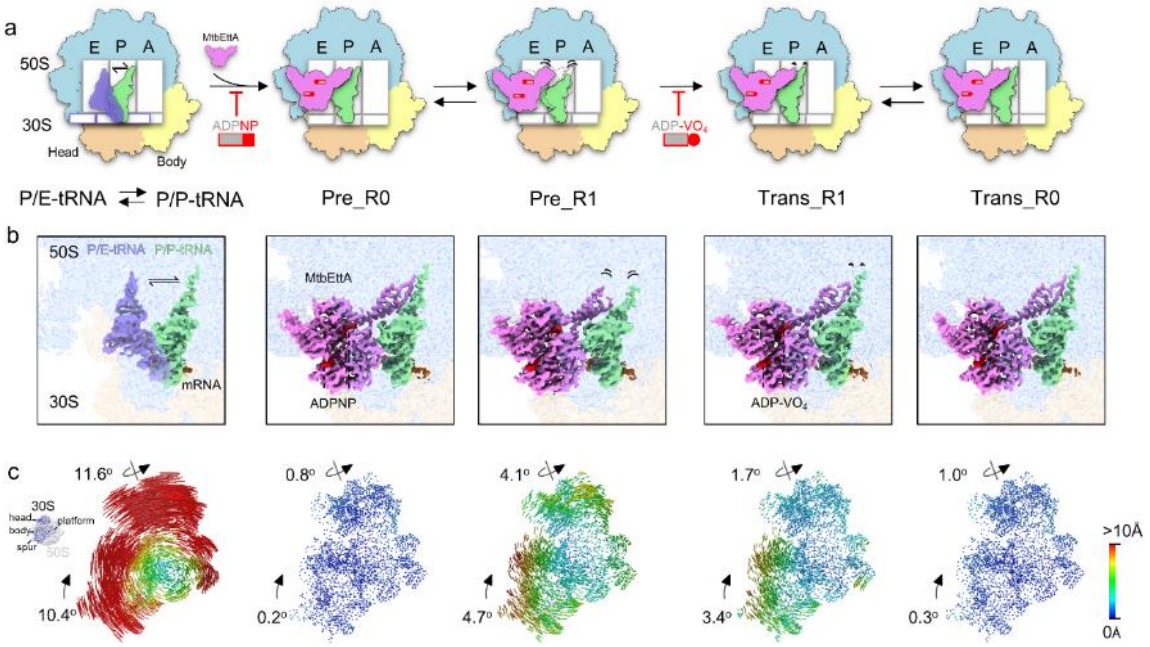


Figure 12 Cryo-EM structures of *Mtb* ribosome in complex with MtbEttA in pre-hydrolysis and transition states.

a, Cartoon representations of *Mtb* ribosome in different states, with E, P, and A sites as well as 50S and 30S (head and body) labelled. First column shows the transition of tRNA from P/P (light green) to P/E (transparent medium slate blue) position. Second and third columns show that MtbEttA (violet) resides at E-site and interacts with P-site tRNA at pre-hydrolysis state (ADPNP), with two distinctive conformations, namely Pre_R0 and Pre_R1, respectively. Fourth and fifty columns show the two conformations of *Mtb* ribosome in complex with MtbEttA in transition state (ADP-VO_4), namely Trans_R1 and Trans_R0, respectively. Dashed outline at E and P sites are the position of MtbEttA in Pre_R0 state and tRNA in P/P-tRNA state, respectively. Flexibility is indicated by curved lines. b, Cryo-EM maps showing the density of tRNAs and MtbEttA with 50S (light blue) and 30S (light yellow) in the background. P/P-tRNA, P/E-tRNA, mRNA, MtbEttA, and nucleotide analogues are labelled accordingly and correspond to the cartoon scheme. c, Analysis of 30S rotation in P/E-tRNA and MtbEttA-bound states, with the classic P/P-tRNA non-rotated state as the reference. (Inset) View of 30S subunit and key features are labelled. Vectors showing the differences between equivalent C4' or Cα atoms of the 30S subunits are colored as rainbow, with arrows indicating the degree and direction of the rotation.

II.3.2 Interactions between MtbEttA and *Mtb* ribosome

We built atomic models with high-resolution cryo-EM maps to elucidate the detailed interaction pattern between MtbEttA and *Mtb* ribosome. MtbEttA has a conserved ABCF architecture with two tandem nucleotide-binding domains (NBDs), the arm domain, PtIM, and basic C-terminal tail (Fig. 10a). Using the Pre_R0 state as an example, the arm domain of MtbEttA interacts with the L1 stalk of the 50S (Fig. 10b). Surprisingly, contrary to the low-resolution cryo-EM structure of the EttA bound *E. coli* ribosome, the basic C-terminal tail of the MtbEttA is inserted between the L1 stalk and NBD1 (Fig. 10c). The MtbEttA C-terminus density is resolved, showing the trace of the residues (Fig. 7). Interactions between MtbEttA and 30S involve a salt bridge between Glu526 of NBD2 and Arg112 of uS7 protein, as well as electrostatic interactions between the basic residues Arg316 and Arg346 of NBD2 and the h41-h42 RNA backbones of the 16S (Fig. 10d). As opposed to the E. coli EttA-ribosome complex structure, PtIM of MtbEttA is more extended than previously demonstrated (Fig. 12). The first α-helix of PtIM (PtIM-α1), which connects to the NBD1, has extensive interactions with H68 of 23S rRNA (Fig. 10e). Moreover, the positively-charged tip of PtIM is stretched and inserted into the region surrounded by H74, H75, H93, and the acceptor arm of the P-site tRNA (Fig. 10b,f). The electrostatic potential of MtbEttA confirms the positive surface where it is in contact with ribosomal rRNAs and tRNA (Fig. 9). Surprisingly, residue Arg 280 at the tip of the PtIM is in parallel with the backbone of the CCA tail of the P-site tRNA and reaches the vicinity of PTC (Fig. 10f).

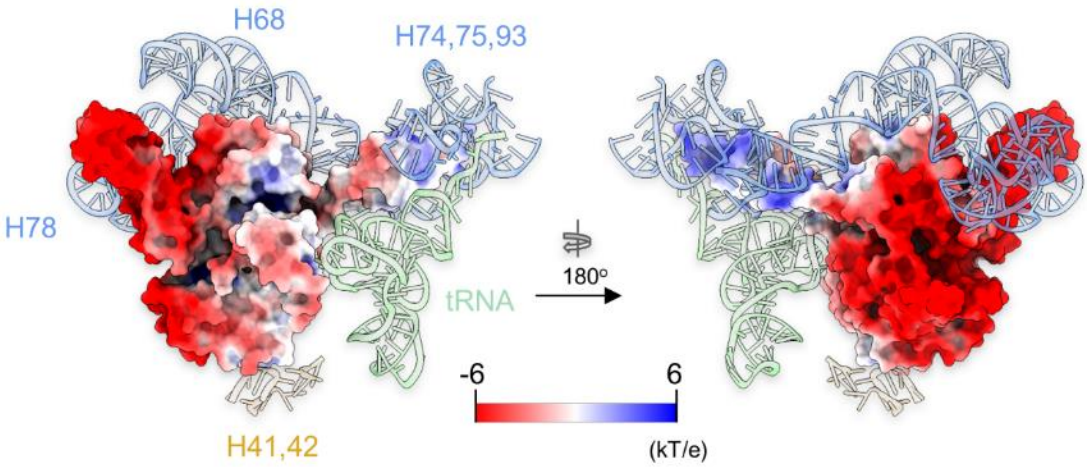


Figure 13 Electrostatic surface potential of MtbEttA.

MtbEttA is colored by the electrostatic potential. H68, H74-75, H78, H93 of 23S rRNA and tRNA are shown in transparent ribbon model.

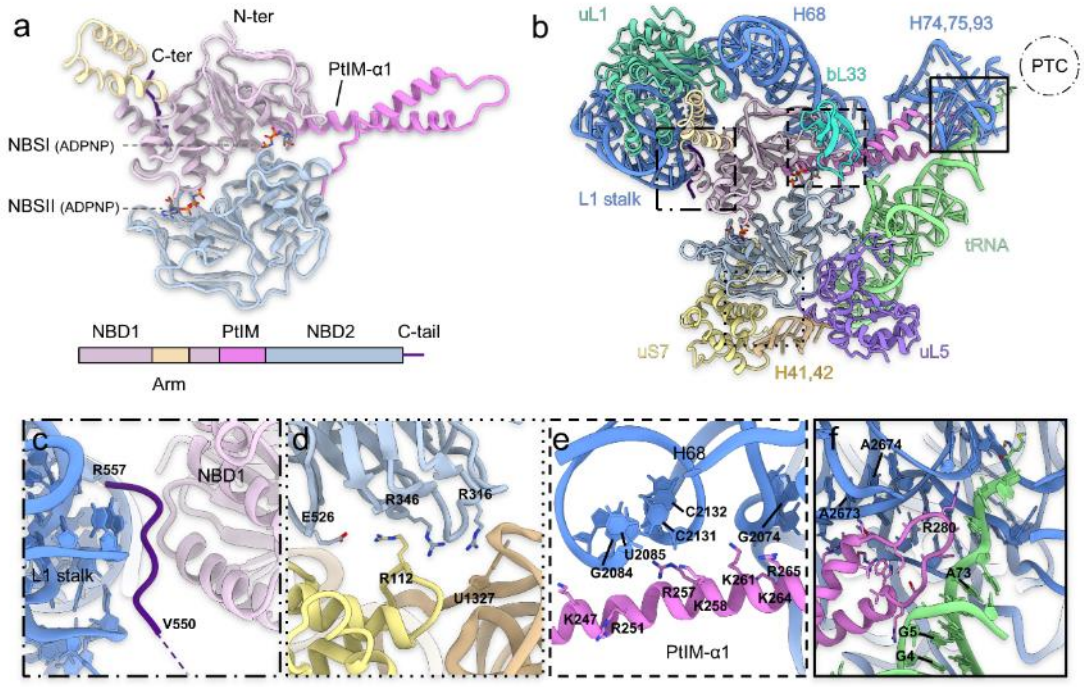


Figure 14 Structure of MtbEttA and its interaction with *Mtb* Ribosome in Pre_R0 state.

a, The structure and domain composition of MtbEttA. Nucleotide-binding domain 1 (NBD1, residues 1-93 and 137-240) and 2 (NBD2, residues 308-545), arm domain (residues 94-136), P-tRNA interacting motif (PtIM, residues 239-307), and the C-terminal basic tail (residues 550-557) are colored in thistle, light steel blue, wheat, violet, and indigo, respectively. Nucleotide-binding sites (NBSI and NBSII) are indicated by dashed lines. b, The overall interactions between MtbEttA and surrounding ribosomal RNAs and proteins. Peptidyl transferase center (PTC) is outlined by a dashed cycle. Detailed view of squared regions is described in c-f. c. The C-terminal tail is inserted between L1 stalk and NBD1. d, Interactions between MtbEttA and 30S. A salt bridge is observed between E525 (MtbEttA) and R112 (uS7). Basic residues in MtbEttA (R316 and R346) are reaching to H41 and H42 of 16S rRNA. e, PtIM-α1 is in close contact with H68 of 23S rRNA. f, The tip of PtIM reaches the proximity of PTC, and interacts with the CCA-stem of P-site tRNA, H74, H75 and H93 of 23S rRNA.

II.3.3 Remodelling of P-site tRNA and 50S in the course of ATP hydrolysis of MtbEttA

The intersubunit movement of the 70S is often accompanied by the shift of P-site tRNA, resulting in various intermediate P-site tRNA positions due to the interactions with different ribosome-interacting factors (155) (156). Previous structures of the ribosome in complex with MsrE and VmlR revealed a novel P-site tRNA conformation, in which the acceptor stem was displaced by the tip of the PtIM and redirected from the PTC toward the A site (86) (85). The PtIM from MtbEttA will unlikely cause such a drastic displacement due to the short PtIM tip (Fig. 12). Indeed, by superimposing structures to 70S with P/P tRNA according to the 23S rRNA, we find that the acceptor stem of the P-site tRNA remains in the same position as the classic P/P-tRNA in the Pre_R0 and Trans_R0 states, while the elbow region shifts ~6 Å toward the E site, presumably due to the interactions with the bound MtbEttA (Fig. 11). In the Pre_R1 state, the elbow of P-site tRNA moves ~14 Å toward the E site, comparable to that in AREs bound ribosomes (Fig. 11). However, the CCA tail and acceptor stem remain at P-site but with increased flexibility.

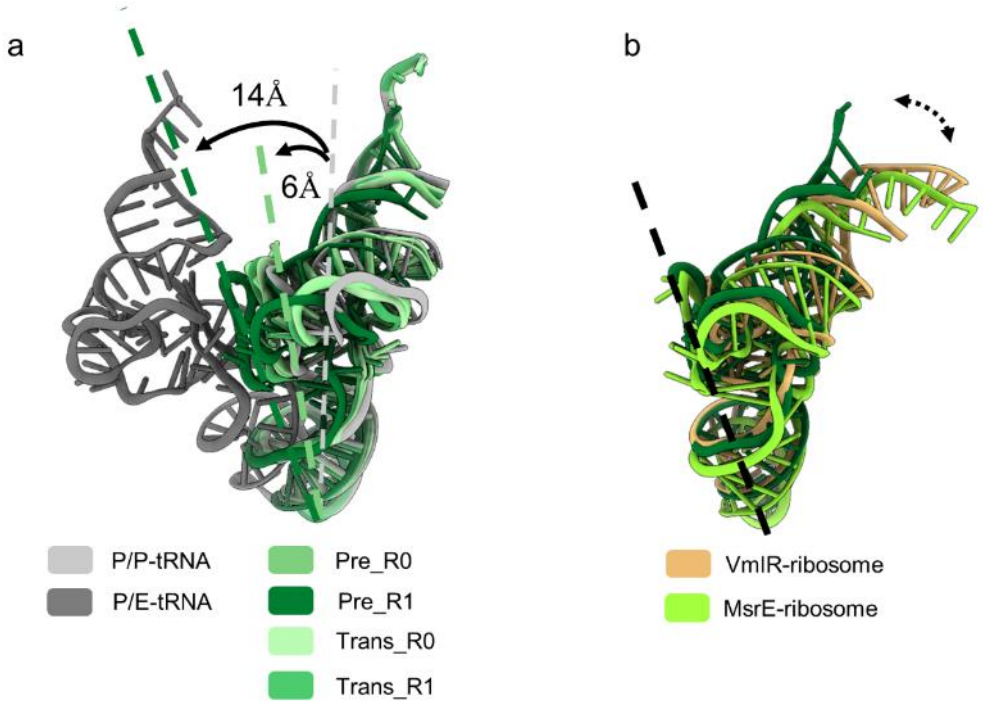


Figure 15 Comparison of P-site tRNAs.

a, tRNA positions in Pre_R0, Pre_R1, Trans_R0, and Trans_R1 states, compared to P/P-tRNA and P/E-tRNA. Dashed lines indicate the elbow region of the tRNA. Distance and direction of the movement in Pre_R0 and Pre_R1 states relative to P/P-tRNA position are labelled by arrows. b, Comparison of tRNA positions in Pre_R1 state, VmlR-ribosome and MsrE-ribosome complexes. Dashed line and double-headed arrow indicate the elbow region and CCA tail of tRNA, respectively.

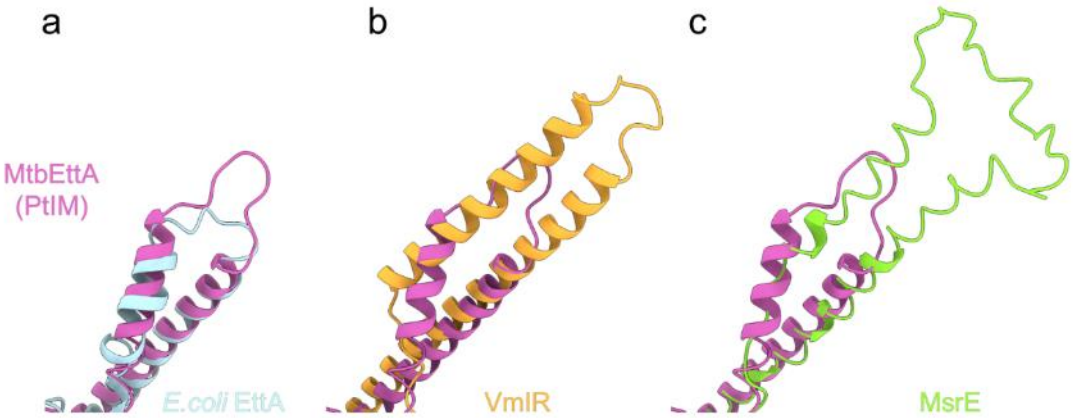


Figure 16 Comparison of the PtIM of MtbEttA in Pre_R0 state with *E.coli* EttA, VmlR, and MsrE.

In coordination with the movement of 30S in the Pre_R1 state, MtbEttA shifts toward the same direction as 30S to ~6 Å. (Fig. 13a,b). Interestingly, the 50S protein bL31, which is responsible mainly for forming the ribosomal intersubunit bridge 1b (B1b), and protein uL5 move ~2 Å to the opposite direction of the 30S head swiveling (Fig. 13a,b). In addition, the 5S rRNA and H84 of the 50S show an upward shift (Fig. 13a). The movement around central protuberance (CP) in the Pre_R1 state results in a splitting interface between the 50S and 30S (Fig. 13b). Although there is a similar observation in the 70S with P/E tRNA, in which the 30S is at a fully-rotated state, it does not reach the same extent as in the Pre_R1 state (Fig. 15b). It suggests that this is not a sole result of 30S rotation but rather the combination of the occupation of MtbEttA at the E-site and the rotation of 30S. Indeed, the movement of CP in transition states is not observed due to subtle 30S rotation (Fig. 15a).

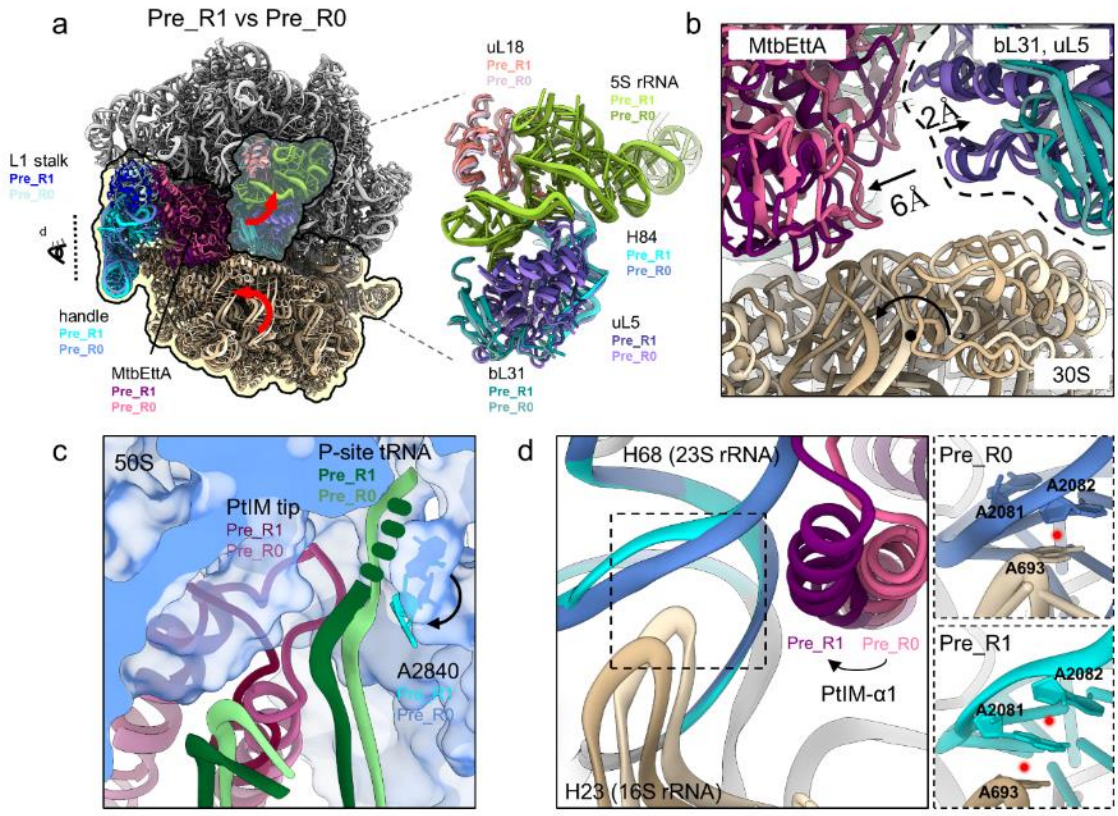


Figure 17 Remodeling of the 50S in Pre_R1 state due to 30S rotation and conformational change of MtbEttA.

a, Structural comparison between Pre_R0 and Pre_R1 state. Models of 70SIC-MtbEttA complex in Pre_R0 and Pre_R1 states were superimposed based on 23S rRNA. Outlines are used to include two colored groups demonstrating the opposite directions of movement. MtbEttA, tRNA, handle and L1 stalk are in line with the movement of 30S. Proteins bL31, uL5, uL18 and 5S rRNA are in the opposite direction of 30S movement and showed as the close-up view in the inset. b, Zoom-in view of the splitting interface in 70S. The dashed line marks the separation point. Distance and direction of movement for MtbEttA, bL31 and uL5 are indicated by arrows. c, Remodeling of PtIM and P-site tRNA near PTC in Pre_R1 state. 50S in Pre_R0 state is shown as a transparent surface. Nucleotide A2840 is shown in filled rings. Disordered CCA tail of tRNA in Pre_R1 state is drawn as dashed line. d, Rearranged 50S-30S interface in Pre_R1 state. Direction of the view is indicated by the eye cartoon in a. Reformed nucleotide stacking is shown with red dot in the inset.

We then examined the effect on the 50S due to the observed unconventional movement in the Pre_R1 state. Increased flexibility around PTC, specifically H89 and protein uL16, is observed (Fig. 14). Also, nucleotide A2840 flips 180^{o}, pointing at the A-site (Fig. 13c). Furthermore, the interface between the 50S and the platform region of 30S is found to be remodeled in the Pre_R1 state (Fig. 13d). Nucleotide stacking between A693 from 16S rRNA and A2082 in H68 from 23S rRNA is maintained in Pre_R0 state. However, due to the movement of 30S and PtIM-α1 in MtbEttA, 23S rRNA nucleotide A2081 slides in between A693 and A2082 and forms stable stacking with both nucleotides. The newly formed stacking interaction may contribute to stabilizing the specific intermediate rotation of the 30S in the Pre_R1 state, as such interaction is not present in the classic non-rotated structure, transition states, and the fully rotated ribosome (Fig. 15c).

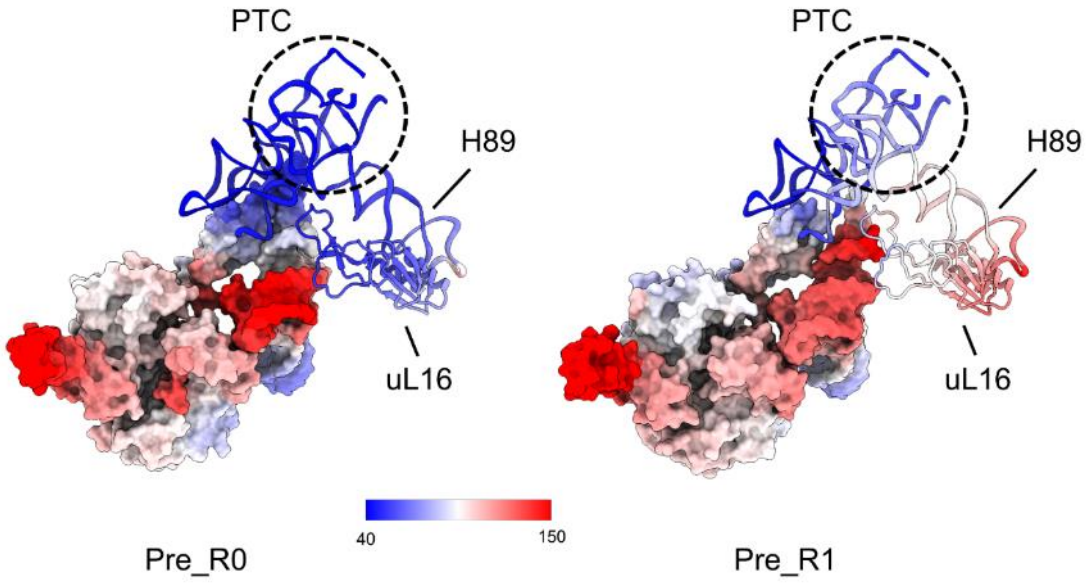


Figure 18 Increased flexibility of PTC and uL16 in Pre_R1 state
Models are colored with b factor values after refinement

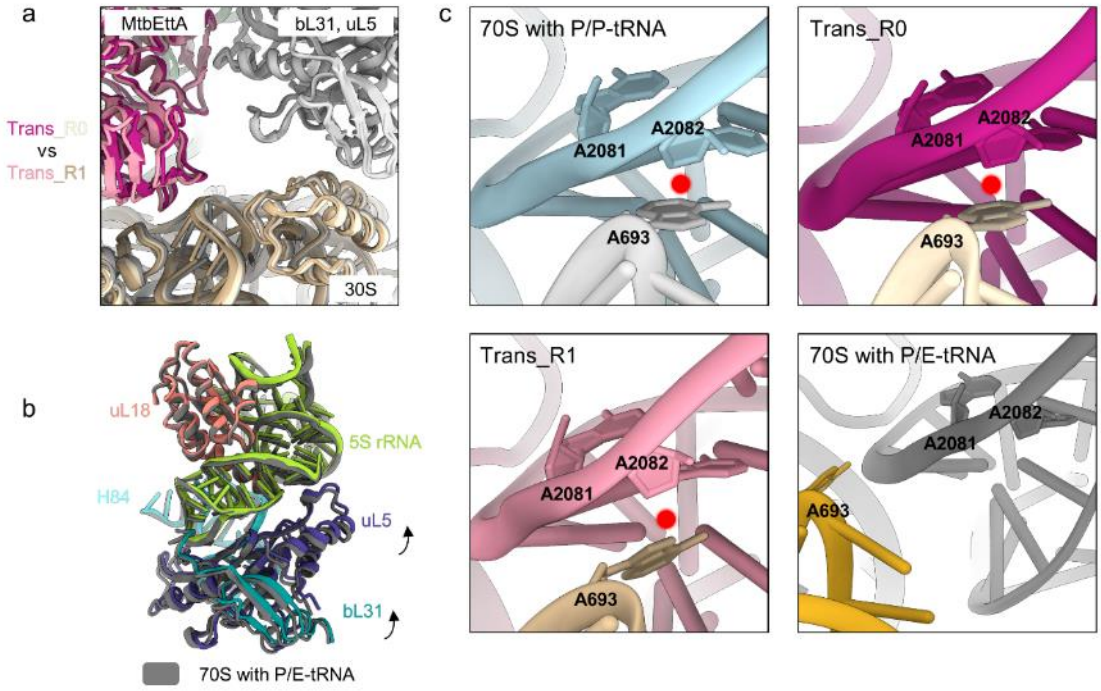


Figure 19 Remodelling of the 50S is unique in the Pre_R1 state.

a, Comparison of Trans_R0 and Trans_R1 states in the same view as in Fig. 3b. 30S of 70S with P/P-tRNA is colored in lightgray. No obvious movement is observed for proteins bL31 and uL5, which are colored in lightgray. b, Structural overlapping between Pre_R1 state (colored model) and 70S with P/E-tRNA (darkgray model). The arrows indicate the direction of movement in Pre_R1 state. c, 50S-30S interfaces in 70S with P/P-tRNA, 70S with P/E-tRNA, Trans_R0 and Trans_R1 states are showed as in Fig. 3c.

II.3.4 Structural variations of MtbEttA during ATP hydrolysis and asymmetric engagement of the nucleotides in pre-hydrolysis state

ATPase activity is essential for the function of ABCFs (90) (89). Moreover, ATP hydrolysis by EttA significantly increases upon adding ribosomes *in vitro* (90). Although much was focused on the effect of ABCFs toward ribosomes, including antibiotic resistance and translational regulation, little is known about the ribosome's role in

stimulating ABCFs' ATPase activities. With high-resolution cryo-EM structures of MtbEttA at pre-hydrolysis and transition states, we can characterize its unique structural variations during the trajectory of ATP hydrolysis.

Conserved motifs of ABC ATPases are present in the MtbEttA at both NBDs, including Walker A, Walker B, H-switch, Q-loop, A-loop, and signature motif (Fig. 16a). When in complex with ribosome at both pre-hydrolysis state and transition state, the MtbEttA is in a "closed" conformation, in which the nucleotides are "sandwiched" between the two NBDs and engaged by the conserved motifs. ADPNP, ADP-VO_4, and the corresponding Mg^{2+} ions were unambiguously identified at both nucleotide-binding sites (NBSs) for the pre-hydrolysis states and transition states, respectively (Fig. 16b). To explore the flexibility of the MtbEttA, we superimposed the structures at different states according to their NBD1s. A prominent variation was observed between Pre_R0 and Pre_R1 states. The tip region of PtIM in the Pre_R1 state shows a relaxed conformation compared to that of the Pre_R0 state, with Cα distances of corresponding residues at a range of 5-10 Å (Fig. 17a). PtIM-α1 also undergoes deformation, showing a disconnected and bent α-helix (Fig. 18b). The movement of PtIM-α1 corresponds to the shift of the H-switch loop at NBS-I (Fig. 18a). Furthermore, the NBD2 of the Pre_R1 state rotates ~5^{o} more toward the NBD1 (Fig. 18b).

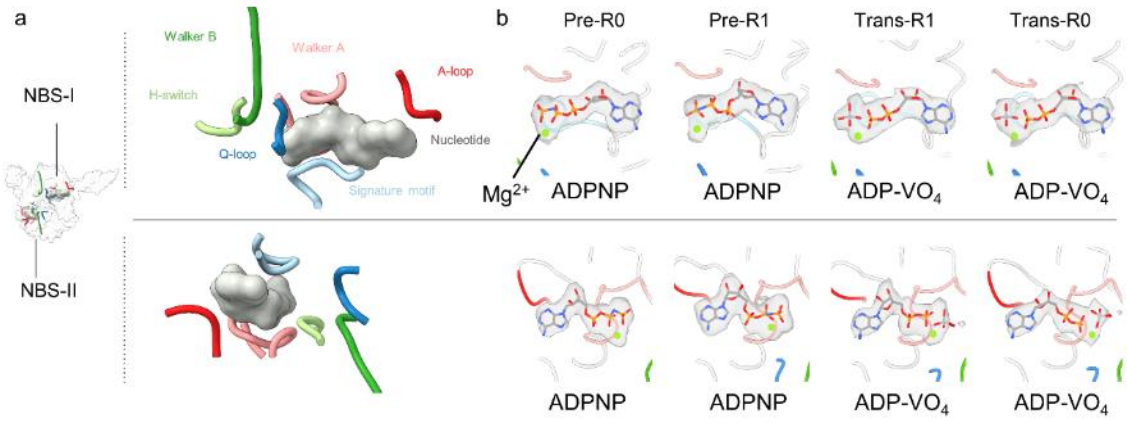


Figure 20 Cryo-EM density of ATP analogues in nucleotide-binding sites.

a, Zoom-in view of conserved motifs in NBSs. Nucleotides are shown as gray molecular surface. b, Cryo-EM densities of ADPNP and ADP-VO_4 are shown in transparent lightgray. Mg^{2+} molecules are shown in cyan spheres.

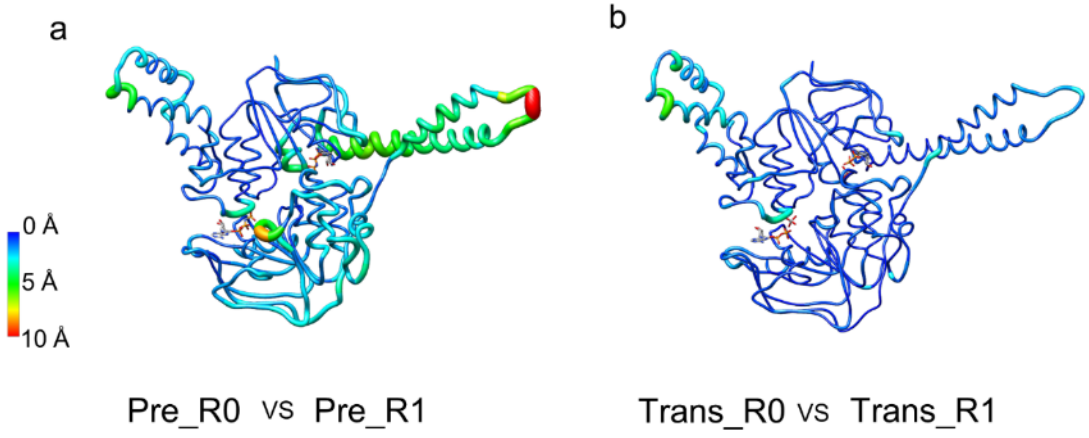


Figure 21 Structural plasticity of MtbEttA in pre-hydrolysis and transition states.

Tube representation of Cα distances between Pre_R0 and Pre_R1 in a, Trans_R0 and Trans_R1 in b. Red color and wide tubes indicate larger Cα distances. MtbEttA shows high plasticity in pre-hydrolysis state.

Surprisingly, an asymmetric engagement of the nucleotides by the conserved motifs in the two NBSs was discovered at the pre-hydrolysis stage (Fig. 18e,f). Upon nucleotide binding, denoted in the Pre_R0 state as the initial stage of interaction between MtbEttA and the ribosome, ADPNP is fully enclosed by the conserved motifs at NBS-I, while the Q-loop at NBS-II is flexible and away from ADPNP, leading to a potentially deficient catalysis site (157) (158)(Fig 18d,f). However, the Q-loop at NBS-II flips inward at a distance of about 4.7 Å to interact with the ADPNP in the Pre_R1 state, creating a functional canonical configuration. While the A-loop (His15) and H-switch (His214) at NBS-I moves outward compared to the canonical positions to 2.5 Å and 3.5 Å, respectively. In the ATP hydrolysis transition states, the 30S rotates back toward a non-rotated position as shown in Trans_R1 state, the H-switch and A-loop at NBS-I return to the canonical conformation but with a slight rotation, while motifs at NBS-II remain in similar positions compared to that of the Pre_R1 state. Then with further movement of the 30S toward the classic non-rotated state, as present in the Trans_R0 state, the H-switch and A-loop at NBS-I the Q-loop at NBS-II are in canonical conformation.

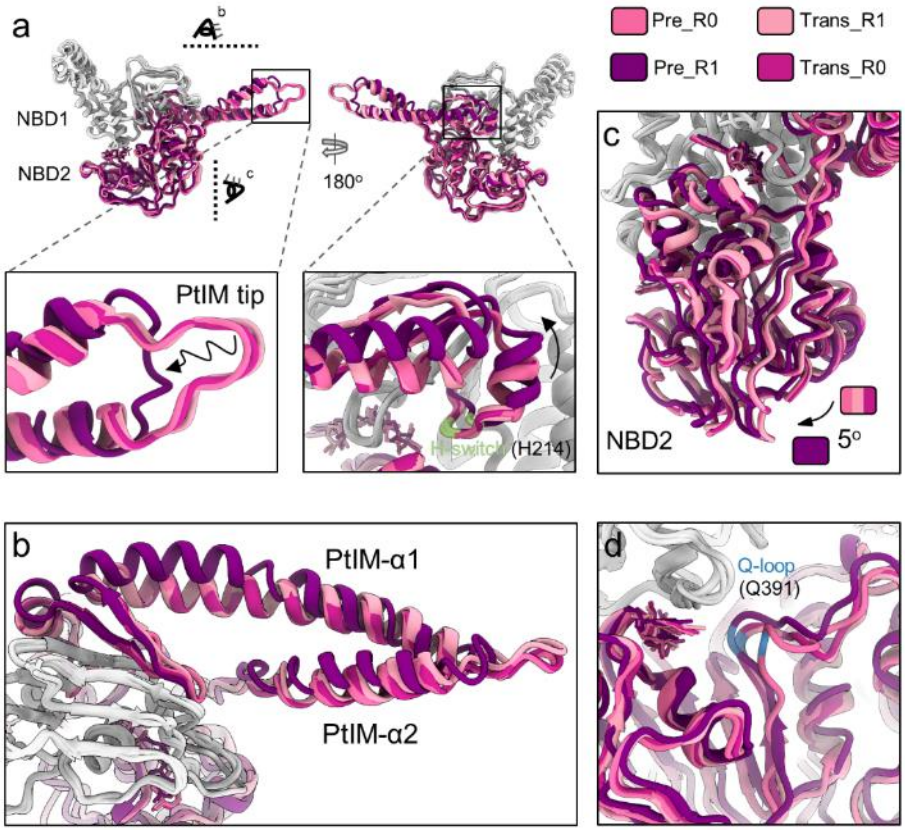


Figure 22 Structural plasticity and asymmetric nucleotide engagement in MtbEttA

a, Overall structural comparison of MtbEttA at different states. Models are aligned based on the NBD1. NBD1 and arm domain are colored as lightgray. PtIM and NBD2 are colored accordingly. Insets show the close-up view of the tip of PtIM and the base of PtIM-α1. H214 of the H-switch in NBD1 is colored in pale green. b, Top view of MtbEttA from the direction of the eye cartoon in a. Focused view of NBD2 and Q-loop in NBD2 are shown in c and d, respectively. Curved arrow indicates the rotation direction in c, and Q391 is colored in dodger blue in d. e,f, Comparison of nucleotide analogues and interacting motifs, which are shown as stick and ribbon models, respectively. Models are superimposed based on NBD1 for e, and NBD2 for f, and are shown in a vertical fashion. Different motifs are colored the same way as in Figure 20. Magnesium is in ball representation and colored in cyan. Cartoons in between e and f indicate different states of 70SIC-MtbEttA complex and different degrees of 30S rotation. Dashed lines represent the Cα positions of corresponding residues in Pre_R0 state. Distances between Cα atoms are labelled, compared to those in Pre_R0 state.

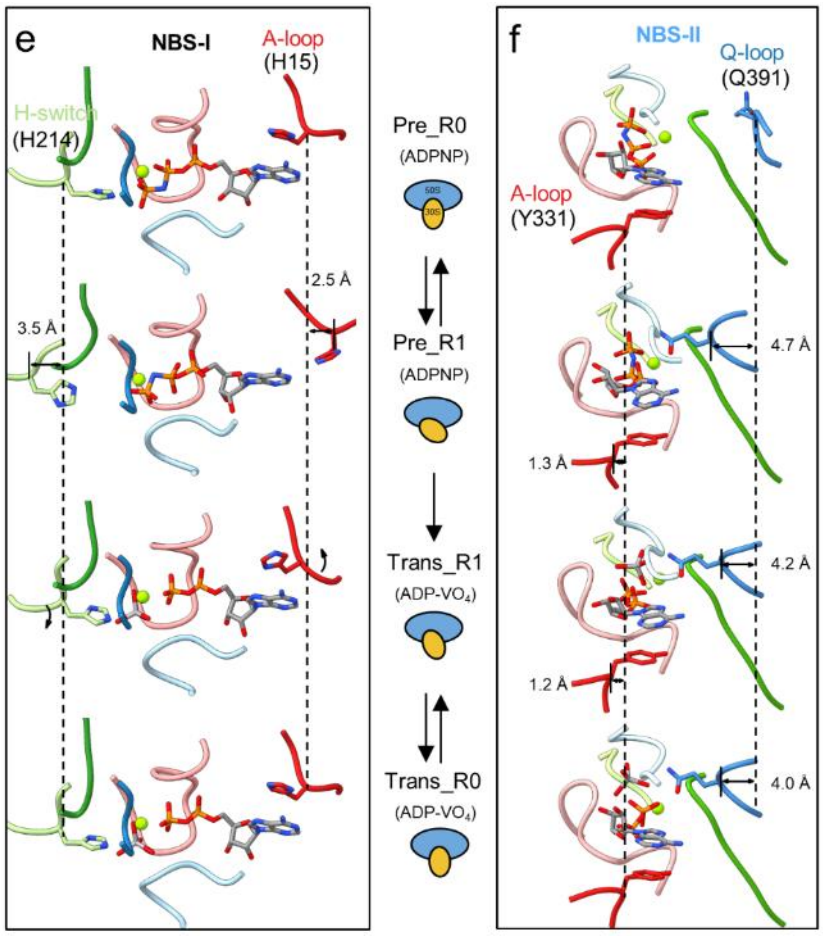


Figure 23 Structural plasticity and asymmetric nucleotide engagement in MtbEttA

a, Overall structural comparison of MtbEttA at different states. Models are aligned based on the NBD1. NBD1 and arm domain are colored as lightgray. PtIM and NBD2 are colored accordingly. Insets show the close-up view of the tip of PtIM and the base of PtIM-α1. H214 of the H-switch in NBD1 is colored in pale green. b, Top view of MtbEttA from the direction of the eye cartoon in a. Focused view of NBD2 and Q-loop in NBD2 are shown in c and d, respectively. Curved arrow indicates the rotation direction in c, and Q391 is colored in dodger blue in d. e,f, Comparison of nucleotide analogues and interacting motifs, which are shown as stick and ribbon models, respectively. Models are superimposed based on NBD1 for e, and NBD2 for f, and are shown in a vertical fashion. Different motifs are colored the same way as figure 16. Magnesium is in ball representation and colored in cyan. Cartoons in between e and f indicate different states of 70SIC-MtbEttA complex and different degrees of 30S rotation. Dashed lines represent the Cα positions of corresponding residues in Pre_R0 state. Distances between Cα atoms are labelled, compared to those in Pre_R0 state.

II.3.5 MtbEttA at post-hydrolysis state (ADP) adopts an open conformation

Others have made speculations that the two NBDs of ABCFs could undergo conformational changes after ATP hydrolysis, as discovered in the ABC superfamily (130). However, the lack of structure of ABCF at post-hydrolysis state limits our understanding of its potential conformational changes. Based on our result (Fig. 6), the MtbEttA in the ADP state does not form a stable complex with the ribosome. Therefore we co-crystallized MtbEttA in complex with ADP and determined the X-ray crystal structure at 2.85 Å (Table 4). The asymmetric unit of MtbEttA-ADP in the crystal contains a domain-swapped dimer, similar to the crystal structure of *E. coli* EttA in the nucleotide-free state. In contrast to *E. coli* EttA, the dimer form of MtbEttA is dominant even at ~10 μM concentration (Fig. 19a). Moreover, the dimer form persists even with different nucleotides and analogs (Fig. 20a). The ADP molecules and Mg^{2+} ions are observed at NBSs in the domain-swapped dimer (Fig. 20b). Only subtle conformational differences are observed between the two NBDs in each monomer (Fig. 19b). To understand the conformational change of MtbEttA at post-hydrolysis state, we focused on half of the domain-swapped dimer composed of the NBD1 from protomer A and NBD2 from protomer B.

Table 4 X-ray crystallography statistics of MtbEttA-ADP

	MtbEttA-ADP
Data collection	
Space group	P 2 21 21
Cell dimensions	
a, *b*, *c* (Å)	50.06, 106.96, 269.19
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)	49.22-2.9 (3.004-2.9)
R_{sym} or R_{merge}	0.0525 (0.4825)
$I / \sigma I$	8.16 (1.85)
Completeness (%)	99.81 (99.72)
Redundancy	2.0 (2.0)
Refinement	
Resolution (Å)	49.22-2.9 (3.004-2.9)
No. reflections(total/unique)	65870/33131
R_{work} / R_{free} (%)	20.39/24.51
No. atoms	8685
Protein	8573
Ligand/ion	112
Water	0
B-factors	
Protein	77.45
Ligand/ion	92.07
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.74

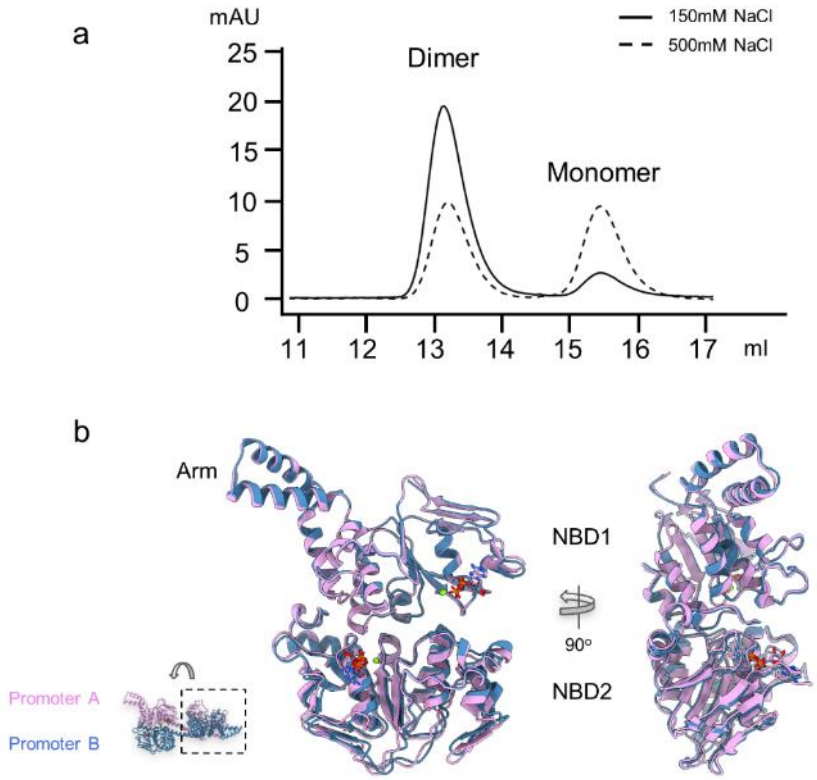


Figure 24 Dimer monomer transition of MtbEttA

a, Size exclusion chromatography of apo MtbEttA at 10μM concentration. Different sodium chloride concentration is plotted in different line style. b, Similar domain organization between two halves of the domain-swapped dimer with ADP molecules. The second half outlined by dashed square (residues 306-550 in protomer A and 2-241 in protomer B) is superimposed to the first half (residues 2-241 in protomer A and 306-550 in protomer B) based on NBD1.

When superimposed the structures of MtbEttA-ADP and MtbEttA in Trans_R0 state according to the NBD1, an opening ~40° rotation of NBD2 in MtbEttA-ADP is observed, confirming the predicted domain movement after ATP hydrolysis (Fig. 20c). The rotational movement accompanies the separation of the LSGGE signature motif, Q-

loop, and H-switch from the nucleotide, showing a classic configuration in the ABC superfamily after ATP hydrolysis (Fig. 20d). To evaluate the effects of such domain movement of MtbEttA at post-hydrolysis state, we superimposed the structure of MtbEttA-ADP to MtbEttA-ribosome complex in Trans_R0 state based on the NBD1 domain (Fig. 20e). Due to the conformational change of MtbEttA at post-hydrolysis state, the NBD2 and arm domain show evident steric clashes with uS7 and uL1, respectively. The rotation of NBD2 and resulting steric clashes would likely facilitate the dissociation of MtbEttA from the ribosome after ATP hydrolysis.

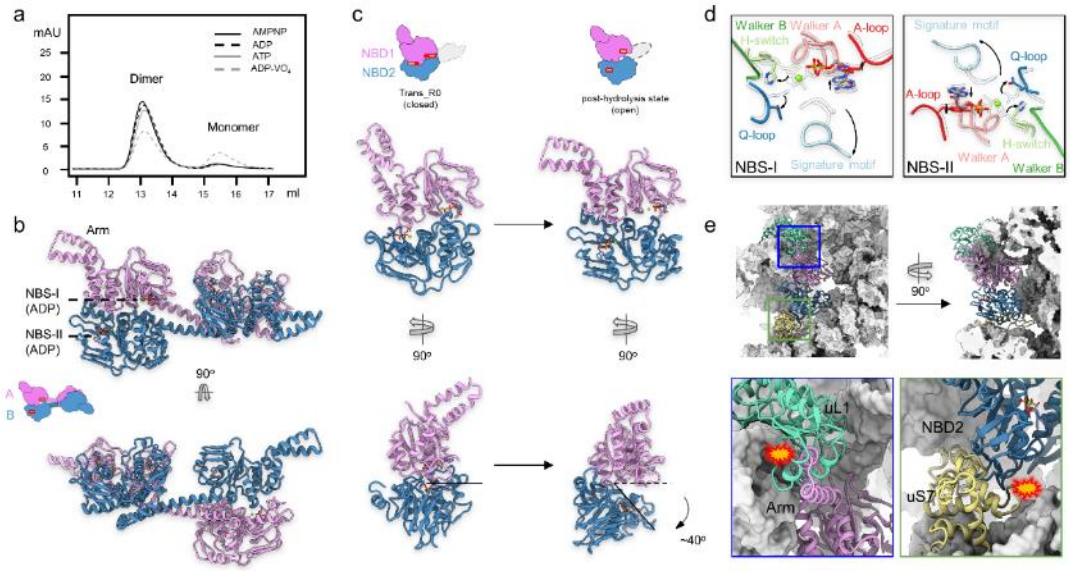


Figure 25 Conformational change of MtbEttA and steric clash with *Mtb* Ribosome in post-hydrolysis state.

a, Distribution of dimer and monomer MtbEttA with different nucleotides and analogues. Size exclusion chromatography graph indicates the distribution of dimer and monomer. Different conditions are represented with different line styles. b, Crystal structure of MtbEttA bound with ADP in the asymmetric unit. The protomer A (plum) and protomer B (steel blue) form a domain-swapped dimer. ADP molecules are labelled in both NBSs. c, Domain opening in the ADP-bound state. Side by side comparison between the conformation of NBD2 (steel blue) in Trans_R0 state and post-hydrolysis state, which is aligned based on NBD1 (plum). PtIM is not shown for the comparison and colored in gray in the cartoon. First column shows the conformation of two NBDs (residues 242-305 are omitted from the model) in Trans_R0 state. Second column represents the model in post-hydrolysis state, which is composed of residues 2-241 in protomer A and 306-550 in protomer B. The direction and degree of NBD2 opening relative to NBD1 is labelled. d, Zoom-in view of the comparison of the two NBSs. Transition of the conserved motifs from Trans_R0 state (transparent) to post-hydrolysis state (colored) is indicated by the arrows. e, Steric clashes between MtbEttA in post-hydrolysis state and *Mtb* Ribosome. MtbEttA in post-hydrolysis state is superimposed with MtbEttA in Trans_R0 state based on NBD1. *Mtb* ribosome is shown as surface in lightgray. uL1 and uS7 are shown as ribbon models and colored in green and yellow, respectively. The stars indicate the clashes between uL1 and the Arm domain, uS7 and NBD2 in the close-up view.

Although the post-hydrolysis state adopts an open conformation, the structural difference is unknown compared to the nucleotide-free state. Since the initial trial of obtaining the nucleotide-free crystal was not successful, we then predicted the structure of the nucleotide-free MtbEttA using homology modeling based on the crystal structure of *E. coli* EttA. By superimposing two structures according to NBD1, we observed that NBD2 from the apo state moves ~8^{o} toward NBD1, compared to the post-hydrolysis state (Fig. 21a). Besides, the different conformations of the arm domain among the transition state, post-hydrolysis state, and apo state indicate high plasticity (Fig. 21b).

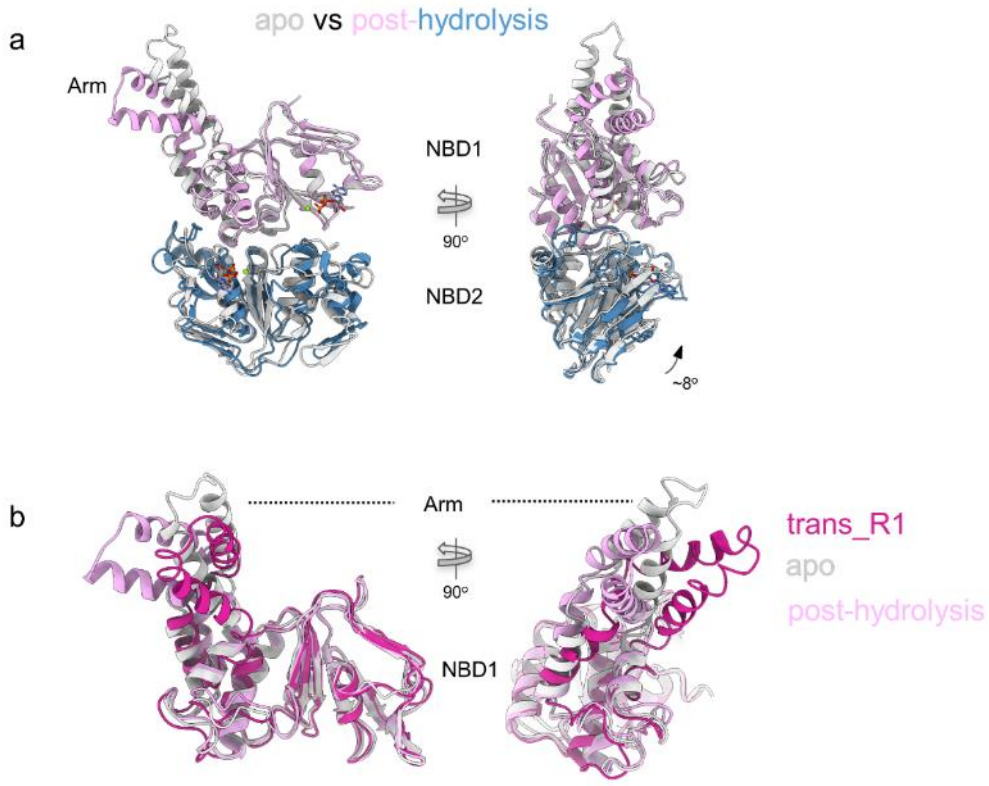


Figure 26 Conformational changes of NBDs and Arm domain in apo and post-hydrolysis states.

a, Structural comparison between post-hydrolysis state and apo state (homology model based on PDB ID: 4FIN), superimposed by NBD1. NBD1 and NBD2 in post-hydrolysis state are colored in plum and steel blue, respectively. Apo state is colored in light gray. b, Arm domain plasticity in transition state, post-hydrolysis state and apo state.

II.4 Discussion

We combined cryo-EM and x-ray crystallography to study the structures of MtbEttA and its interaction with *Mtb* ribosome in the course of ATP hydrolysis. Previous studies suggest that ribosomes can stimulate EttA's ATPase activity, and in return, EttA can promote the first peptide bond formation on the ribosome (90). Through

the investigation of the interplay between MtbEttA and *Mtb* ribosome, we propose a general model for its mechanism-of-action (Fig 22), primarily focusing on two questions: (1) What is the role of the ribosome in affecting the ATPase activities of MtbEttA; (2) How does MtbEttA stimulate the translational efficiency of *Mtb* ribosome. Although the functional state of MtbEttA is the monomer form, the domain-swapped dimer persists even at low concentration, which could serve as an autoinhibitory mechanism. The transition from dimer to monomer in vitro is not promoted by the nucleotide-binding but rather by high salt concentration (Fig. 19a), which suggests the electrostatic interactions stabilize the dimer formation. When transiting from dimer to monomer, the two NBDs would undergo extensive domain rearrangement to form the compact configuration, as indicated by the molecular size on the size exclusion chromatography. However, the structure of PtIM is ambiguous due to the lack of structure of the monomer state alone.

In pre-hydrolysis state, MtbEttA could still maintain an open conformation after ATP-binding, but before interacting with the *Mtb* ribosome, as suggested by the structure of human ABCF1 at pre-hydrolysis state (159). Nevertheless, MtbEttA displays a closed conformation after it binds to the ribosome, indicating that the ribosome could facilitate the closure of NBDs. Surprisingly, an asymmetric engagement of nucleotides at the two NBSs is observed and correlated with the rotation of 30S. In the Pre_R0 state, in which the 30S resembles classic non-rotated conformation, ADPNP at NBS-I is fully engaged with all functional motifs, while it is engaged by all motifs but the Q-loop at NBS-II. As the 30S rotates to an intermediate position shown in Pre_R1 state, the

ADPNP at NBS-II is now fully engaged, but the H-switch and A-loop at NBS-I move away from the ADPNP. The conformational change of H-switch and A-loop at NBS-I is likely mediated by the distortion of PtIM-α1, which serves like a cantilever to transmit the effect of 30S rotation. Although the effect of ATP hydrolysis due to the movement of these residues is unknown, mutations at these positions in other ABC proteins impair their ATPase activities (160) (161) (162). In summary, it appears that in order to accommodate sufficient engagement of nucleotides in pre-hydrolysis state at both NBSs, MtbEttA has to undergo a substantial conformational change, which is facilitated by the rotation of 30S. By contrast, nucleotides are enclosed by all conserved motifs at both NBSs in the transition states, presumably due to smaller head swelling of the 30S or more rigid domain association in MtbEttA when in complex with ADP-VO_4, as increased monomer population is observed on size exclusion chromatography (Fig. 20a).

In agreement with the single-molecule fluorescence resonance energy transfer (smFRET) experiment on *E. coli* EttA (89), as MtbEttA occupies the E-site on the 70S in both pre-hydrolysis and transition states, the movement of 30S is limited although not stopped to prevent the spontaneous rotation which will result in a P/E tRNA conformation. Therefore the translational efficiency of ribosomes could increase due to the favorable conformation for incorporating A-site tRNA. During the accommodation of A-site tRNA, the elbow and 3′-CCA end of the tRNA have to be navigated through the accommodation corridor to transit from A/T to A/A position (163) (164). Dynamic simulation has pointed out that the displacement of H89 of 23S rRNA during elbow accommodation could avoid steric clashes (163). Interestingly, increased flexibility

around PTC, especially H89, is observed in the Pre_R1 state. A reasonable speculation is that a more flexible H89 could facilitate the accommodation of A-site tRNA, the rate-limiting step of peptide formation, by avoiding steric clashes. However, a thorough dynamic simulation that includes the flexible property of H89 and kinetic experiment of the accommodation rate is needed to validate the idea. Importantly, whether MtbEttA interacts with 70 initiation complex before or after EF-Tu-aa-tRNA binding needs to be explored to better understand the role of MtbEttA.

The physiological role of MtbEttA still needs to be addressed, as it is essential in *Mtb* but not in *E. coli*. The essentiality of MtbEttA could be due to the necessity for *Mtb* to survive in a stressful environment, as knockout of the *ettA* gene in *E. coli* results in a severe fitness defect in the long-term stationary phase (90). However, we cannot rule out the possibility that MtbEttA could be involved in a different essential pathway in addition to the translational regulation. Furthermore, our results could provide a possible explanation in the aspect of EttA functioning as a translational throttle based on the ATP/ADP ratio. Due to the nucleotide's asymmetric engagement during the pre-hydrolysis state, ATP could be more exposed and more accessible to be exchanged by ADP when there is a high ADP concentration in the cell. Substitution of ATP to ADP will disrupt the normal functional cycle of EttA-like protein, therefore impose a different remodeling on the ribosome that would potentially slow down translation. Nevertheless,

single molecular experiments are needed to elucidate the kinetics and at different ATP/ADP ratios.

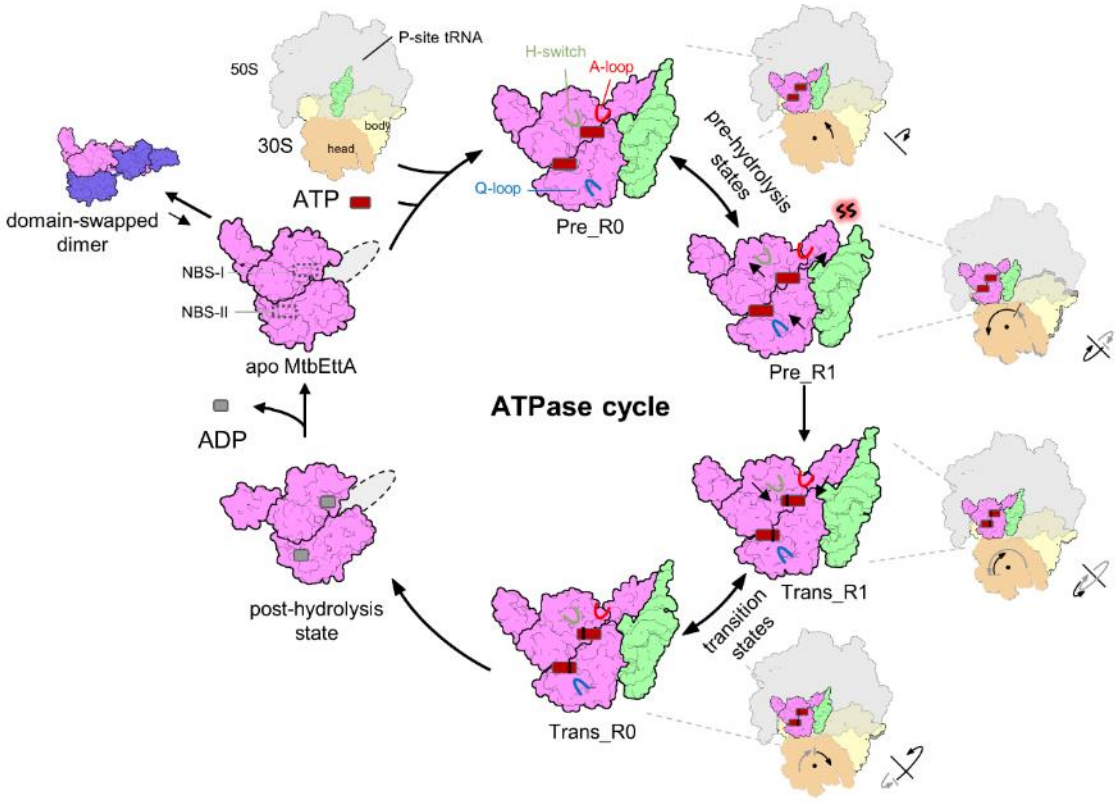


Figure 27 Schematic model of the functional cycle of MtbEttA in the course of ATP hydrolysis

Monomer formation involves the dissociation the domain-swapped dimer (the favored configuration) and rearrengement of the two NBDs to form the open conformation. The monomeric MtbEttA transits to the closed conformation after binding to ATP molecues and the E-site of 70S with mRNA and tRNA in the P-site. At pre-hydrolysis stage, the engagement of Q-loop at NBS-II requires the conformational change of H-switch and A-loop in NBS-I, which is mediated by PtIM-α1 due to the 30S rotation. In addition, the tip of PtIM and CCA-stem of tRNA become distorted. Meanwhile, remodelling of 50S occurs as 30S is in an intermidiate state, potentially facilitating the translocation of A-site tRNA. In the transition state, MtbEttA and P-site tRNA maintain stable interactions due to more restrictred 30S movment. At post-hydrolysis stage, MtbEttA is in an open conformation and disassociated from the ribosome, due to potential steric clash with the ribosome. ADP dissocaition and ATP reloading are necessary for the next cycle.

CHAPTER III

CHARACTERIZATION OF HIBERNATION PROMOTING FACTORS IN *MTB*

III.1 Introduction

Tuberculosis, caused by slow-growing Gram-positive bacillus Mycobacterium tuberculosis (*Mtb*), is one of the top 10 causes of death worldwide (5). About one-fourth of the world's population has latent TB, in which *Mtb* is in a non-replicating persistent or dormancy-like state (165). During latent infection, *Mtb* bacilli have to conquer the nutrient and oxygen limitation by slowing down the metabolic activity, especially protein production (166). Studies involving the M. tuberculosis metabolic state during latent infection are primarily based on the in vitro "Wayne model" (62), wherein a low-density culture is sealed and allowed to slowly consume oxygen until the culture is anaerobic, resulting in a dormancy-like state. With this model, a set of at least 48 tightly regulated genes are identified in the dormancy survival (Dos) regulon controlled by a response regulator, DosR (167) (65). There are two hypothetical ribosome-associated factors among many uncharacterized genes in DosR regulon, RafS, and RafH (66). Understanding how ribosomes are regulated is one of the critical aspects of revealing the molecular mechanism during latency.

The ribosome, an apparatus responsible for protein biosynthesis, comprises two subunits: the 30S and 50S in prokaryotes (168). The small subunit 30S, required for fidelity, contains 16S rRNA and about twenty ribosomal proteins (rProteins). The large subunit 50S, containing the peptidyl transferase active site, includes 5S rRNA, 23S

rRNA, and about thirty rProteins. In bacteria, instead of degrading and regenerating this super complex when switching from stressful to a favorable environment, stabilizing and inhibiting ribosome is the energy-efficient strategy for rapid response to the stress (169). Escherichia coli (*E. coli*), the best-studied Gram-negative bacteria regarding ribosomal stabilization during stress, employs three mechanisms to stabilize and inhibit ribosomes (170). YfiA (previously named as protein Y or RaiA) occupies the A and P tRNA site in the 30S, creating a stable but inactive 70S ribosome. Ribosome modulation factor (RMF) interacts with Shine-Dalgarno (SD) sequence and causes the conformational change of the 30S to facilitate the intimate interaction between two 30S. Therefore the 90S complex, which is made of 70S and 30S, is formed because of the dimerization of 30S. After RMF binds the ribosome, short hibernation promoting factor (SHPF), a homolog to YfiA, also binds to tRNA recognition sites to further form a 100S complex, the dimer of 70S (171) (Fig. 23).

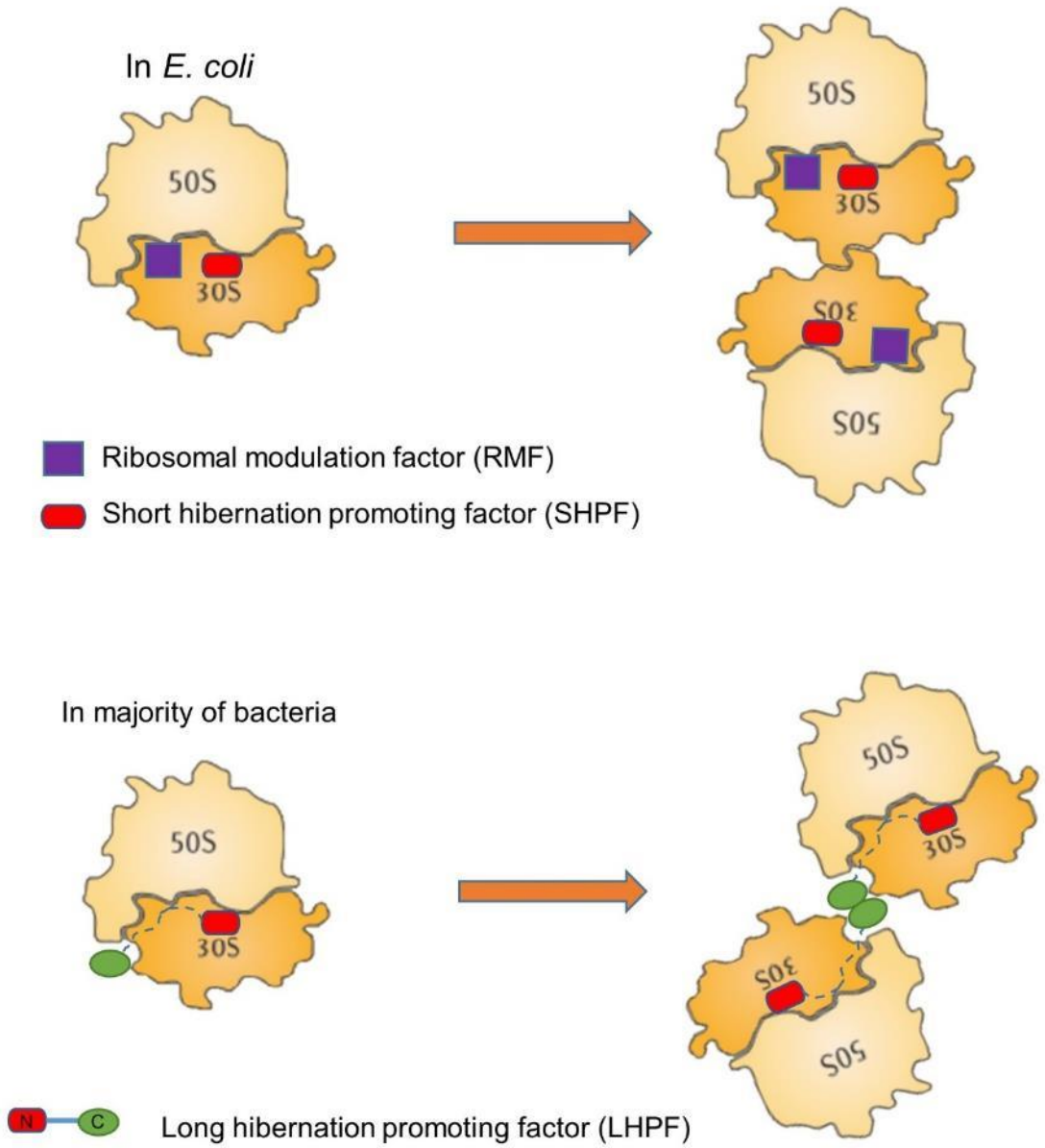


Figure 28 Scheme of different mechanisms of the formation of 100S

The appearance of 100S is also typical for most other bacteria, however, with a distinguished mechanism. They lack both RMF and SHPF, which are only present in a subset of Gammaproteobacteria, including E. coli. Instead, they contain a long-form HPF (LHPF), consisting of an N-terminal domain homologous to SHPF, a C-terminal

domain, and a long linker between them (67). Cryo-EM structures of 100S from four different species, *Staphylococcus aureus* (172) (173), *Lactococcus lactis* (174), *Bacillus subtilis* (175), and *Thermus thermophilus* (176), revealed a more universally conserved mechanism of the formation of 100S in bacteria. The N-terminal domain blocks A and P tRNA sites in the 30S, similar to its homolog SHPF. However, the C-terminal domain can form a dimer through hydrophobic interactions, therefore mediating the formation of 100S, rather than changing the conformation of 30S like in E. coli. Interactions between these two 70S monomers also involve proteins and RNAs from 30S. For example, the interaction between s18, s2, and h26 forms a head-to-head pattern in the dimerization interface. Moreover, no 100S was observed in versions of LHPF with a shorter linker, showing the importance of the length of the linker (175).

Surprisingly, even though RafS and RafH from Mycobacterium are homologs to LHPF, there were no 100S found during the stationary growth and hypoxia stress (66). Indeed, CryoEM structures of the ribosome-RafS complex from *Mycobacterium smegmatis* confirmed that it adopted monomer configuration (68) (177). Similar to other LHPFs, the N-terminal domain binds across A and P sites on the 30S. However, the C-terminal domain is invisible, potentially due to its flexibility. Interestingly, Li et al. showed that RafS only bound to ribosome when the cell was grown at zinc-depleted media and is mediated by Mrf (68). Nonetheless, the detailed molecular mechanism of how zinc is involved in the binding of RafS to the ribosome in *Mtb* is unknown.

Here we present biophysical characterizations of MtbRafH (Rv0079) and *Mtb*Mrf (Rv3241) from *Mtb*. We also explore the interaction between MsmMpy

(MSMEG_1878) and its potential recruitment protein MsmMrf (MSMEG_6069) from *Msm*. Our results suggest that MtbRafH cannot bind to the ribosome in vitro. We also show that there are no stable interactions between Mpy and Mrf. Interestingly, Mpy alone can bind to zinc molecules due to its C-terminal domain, despite lacking the conserved zinc-binding motifs. Besides, the C-terminal domain of Mpy shows reversible zinc-dependent aggregation in vitro. Furthermore, Instead of forming a dimer, the C-terminal domain of Mpy displays as a tetramer, which could potentially prevent ribosomes from forming 100S in Mycobacteria.

III.2 Materials and methods

III.2.1 Protein expression and purification

The full-length MsmMpy (MSMEG_1878), MsmMrf (MSMEG_6069), and MtbRafH (Rv0079), were cloned into a modified pET28(a) vector with an N-terminal His-SUMO tag. MBP-tagged MsmMpy and MsmMrf were constructed with pDB-His-MBP vectors. Untagged versions of *Mtb*Mpy were cloned with pET-22b(+) vectors. We then overexpressed proteins in Rosetta™ 2(DE3) cells and purified using either Ni-NTA or amylose resin depending on the affinity tag. The SUMO tag was then cleaved by SUMO protease and removed by Ni-NTA column, leaving proteins of interest with one extra serine at the N terminus. MBP tag was removed with TEV protease. Further purification was done using gel filtration chromatography by Superdex 200 (16/60 GL) column (GE Healthcare). Buffer exchange was performed using Spectrum™ Labs

Spectra/Por™ dialysis tubing with a different molecular weight cut-off. Protein concentration was done with Sartorius ultrafiltration centrifugal concentrators.

III.2.2 Negative staining electron microscopy

5 μL protein samples (~50nM) were applied on the glow discharged 300 mesh carbon film grid and incubated for 1 min. Protein samples were then blotted away with a filter paper, followed by adding 5 μL 0.75% Uranyl Formate (UF) staining solution onto the grid and blotting away immediately. Another 5 μL 0.75% UF staining solution was applied onto the grid and incubated for 1 min. The staining solution was then blotted away by a filter paper, and the grid was left on a petri dish to dry out for at least 30 mins before imaging under an electron microscope. Images were recorded with Tecnai G2 Spirit BioTWIN operated at 200kV, using K2 Summit direct detection camera (Gatan) in the electron counting mode.

III.2.3 Dynamic light scattering

Particle diameter estimation was done with the Zetasizer Nano-S dynamic light scattering instrument, which provides measurements of particles ranging from 0.6 nm to 6 microns and molecular weight determinations from 1000 to 2 x 10^7 Daltons. A 500 μL sample was used for the measurement.

III.2.4 On-column cross-linking

To mildly cross-link the C-terminal domain of MtbMpy, we performed the "on-column" cross-linking (178). First, 200 µL of 0.25% glutaraldehyde was injected into a pre-equilibrated Superdex 200 (10/300 global in 20mM HEPES, pH 7.4, 500mM NaCl) column and run at 0.25 ml/min for 20 min (i.e., a total of 5ml buffer). Subsequently, the flow was paused, and the injection loop was flushed with equilibration buffer, followed by injection of purified complex (200µl volume, at 5 µM concentration). The column was then run at 0.25 ml/min, and 0.3 ml fractions were collected. Cross-linking product was visualized by running the individual fractions on an 18% SDS gel with Coomassie blue staining.

III.2.5 Sucrose gradient ultracentrifugation

Four different concentrations of sucrose solution (10%, 20%, 30%, 40%) were prepared with the ribosome re-association buffer (5 mM HEPES-NaOH [pH 7.5], 10 mM NH_4Cl, 50 mM KCl, 10 mM $MgCl_2$, and 6 mM 2-mercaptoethanol). The sucrose gradient was prepared as a step gradient first. To easily handle the gradient and increase reproducibility, we used the flash-freezing method to prepare a step sucrose gradient. Each sucrose layer was flash-frozen with liquid nitrogen before adding another layer. A linear 10%-40% sucrose gradient was formed by thawing the step gradient first at room temperature and left at 4 ^{o}C for at least 24 hrs. 100 µL sample was applied on top of the linear gradient and centrifuged with a Beckman SW41 Ti rotor at 40,000 rpm for 2.5 hr. Sample distribution in the gradient was collected with Model EP-1 Econo Pump and

analyzed using Bio-Rad BioLogic LP System. Fractions were collected at every 0.4 ml and visualized by running on an 18% SDS gel with Coomassie blue staining.

III.3 Results

III.3.1 RafH does not interact with the ribosome in vitro

Ribosome-associated factor during Hypoxia (RafH) is proposed to interact with Ribosome and regulate its stability (66). RafH has an N terminal S30AE domain based on sequence analysis, which typically occupies the A and P tRNA sites on the 30S. However, the C terminal domain shows extra ~50 residues than the ribosome hibernation factors (Fig. 24). Although the deletion of RafH affects the ribosome's stability in the hypoxia condition, the characterization of the interactions between RafH and ribosome has not been investigated.

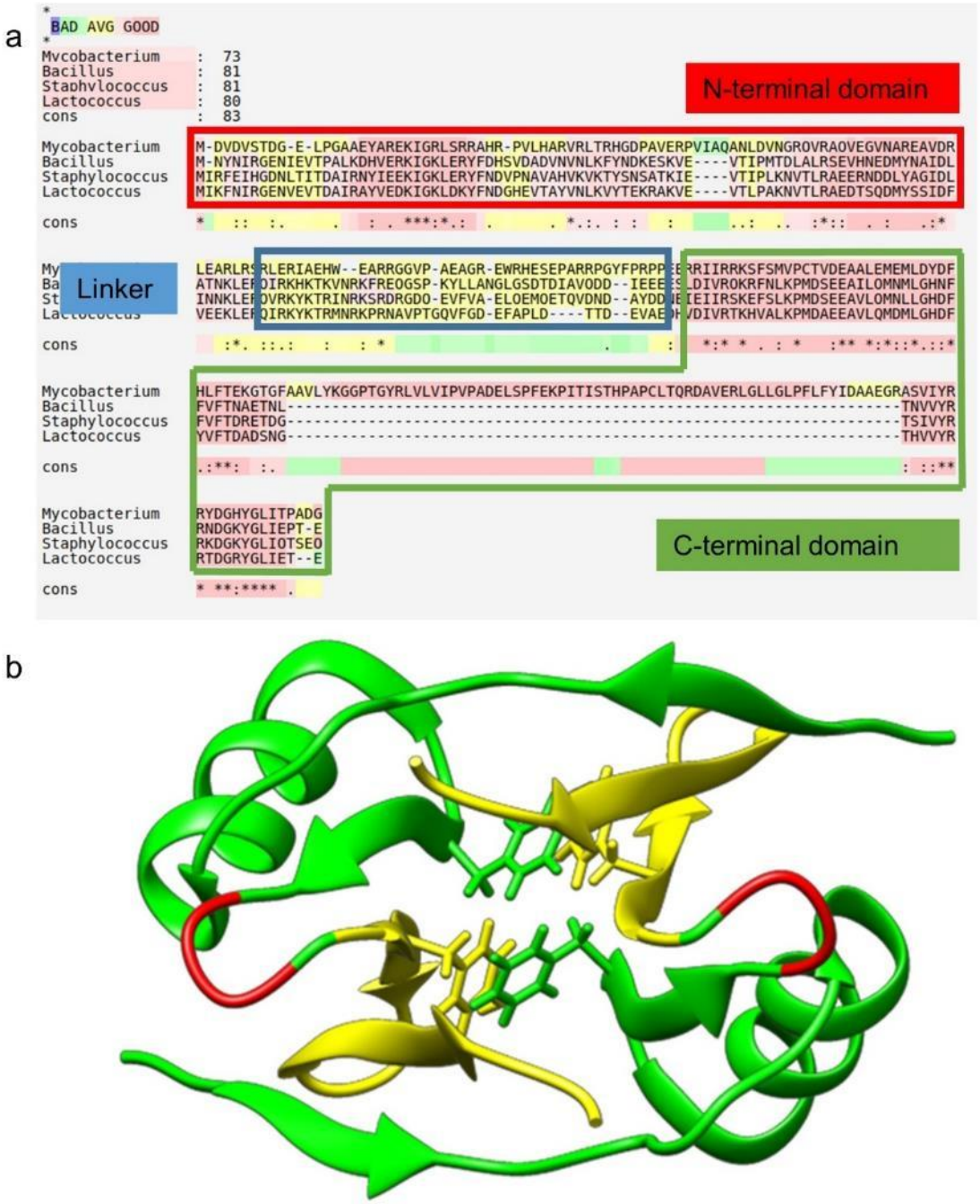


Figure 29 Sequence alignment between RafH and LHPFs

a, the full-length sequence alignment between RafH from *Mtb*, and LHPFs from Staphylococcus and Lactococcus. Different domains are labelled with different color. b, C-terminal domain dimer structure. Red loop shows where the insertion is located in RafH.

We overexpressed MtbRafH in *E.coli* and purified it with His-sumo tag. It reached high purity with one-step purification and was further purified when the sumo tag was removed (Fig. 25a). In order to test the ribosome-binding property of MtbRafH, we performed pull-down experiments with His-sumo-tagged MtbRafH. However, upon mixing His-sumo tagged MtbRafH with either 30S, 50S, or 70S, we observed precipitation. SDS-PAGE analysis showed that the aggregation is mostly His-sumo tagged MtbRafH, while the supernatant still contains soluble MtbRafH (Fig. 25b). We then proceeded with the pull-down experiment with Ni-NTA resin using the His-tag on MtbRafH. Unfortunately, the elution from Ni-NTA resin showed that the His-sumo tagged MtbRafH was unable to pull down 30S, 50S, and 70S (Fig. 25c). Although it did not show stable interaction between MtbRafH and *Mtb* ribosome, we cannot rule out that this could be a false negative due to the sumo tag on MtbRafH. However, MtbRafH, with only His-tag, experienced solubility issues.

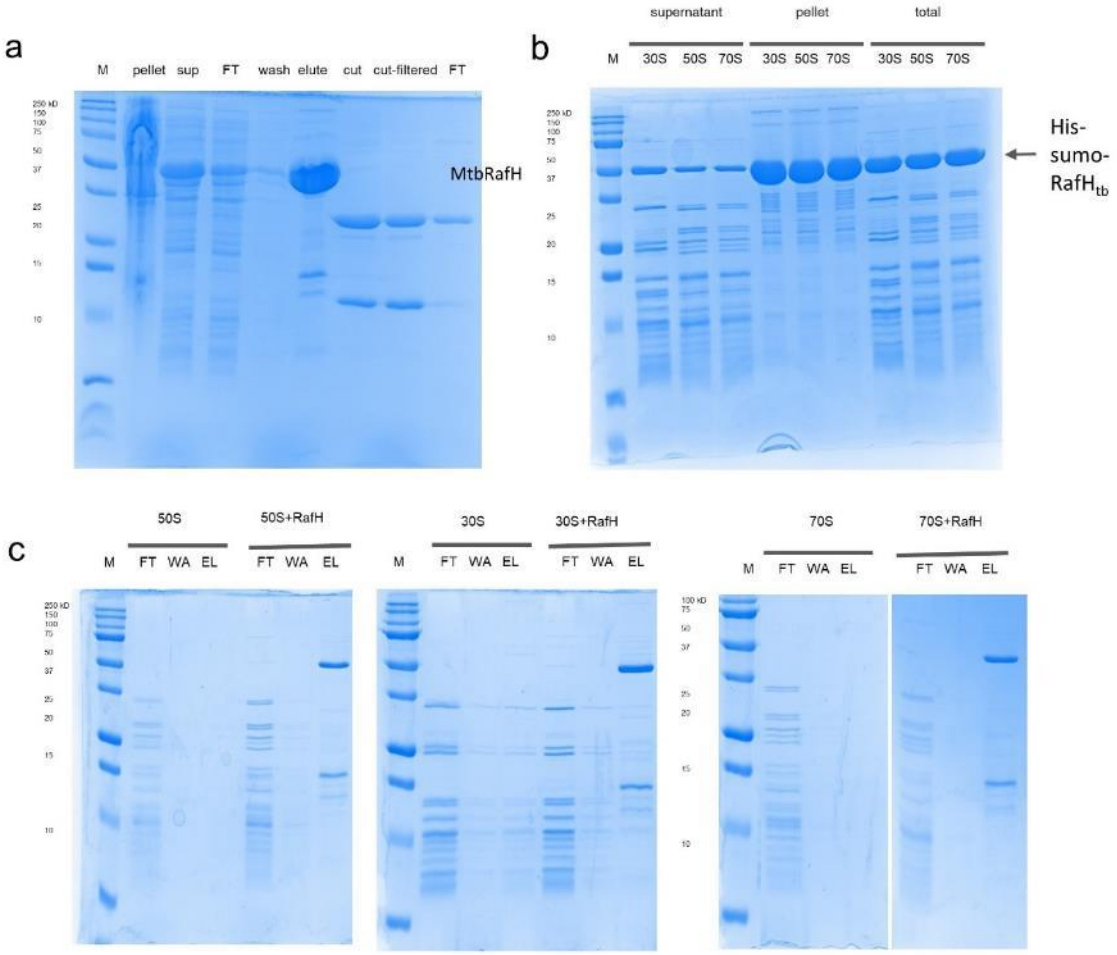


Figure 30 RafH purification and pull-down experiments with ribosome a, SDS-PAGE of the purification of MtbRafH. b, His-sumo-MtbRafH aggregates upon incubating with 30S, 50S, and 70S. c, Pull-down experiments showing MtbRafH is unable to pull down either 30S, 50S, or 70S.

III.3.2 MsmMpy can only solubilize with C-terminal His tag in buffer with high salt and EDTA

Mpy, previously named Ribosome-associated factor during stasis (RafS), is a long hibernation promoting factor. The structural information of Mpy in mycobacteria is still unknown, and so is its interaction with the ribosome. Specifically, the architecture of its C-terminal domain is of great interest. Since there are no 100S observed in

Mycobacteria, the conformation of the C-terminal domain of Mpy could be different from other LHPFs.

Initially, we started with Mpy from *Msm* due to the less time-consuming process of getting ribosomes. An N-terminal His-tag was introduced to the full-length MsmMpy. Although the production is sufficient, there are solubility issues before and after purification (Fig. 26a). In order to address the solubility issue, we tried different buffer conditions during the purification. In the beginning, the high salt (500 mM NaCl) sodium phosphate buffer helped to solubilize the MsmMpy after cell lysis (Fig. 26b) while precipitated after dialysis against low salt buffer (Fig. 26c). Since high salt was necessary to keep the protein soluble, we then tested whether different buffer reagents or pH would keep MsmMpy soluble. Unfortunately, sodium phosphate buffer (pH 7.5) or 5mM HEPES buffer (pH 7.4) failed to keep the protein soluble during concentration (Fig. 26e), as well as HEPES at pH 8.0 (Fig. 26f). We also tested 500 mM NH_4Cl as the salt component, which helped with the solubility of MsmMpy (Fig. 26g). Besides, we found out that it significantly reduced the solubility when removing the 500 mM imidazole in the buffer (Fig. 26h).

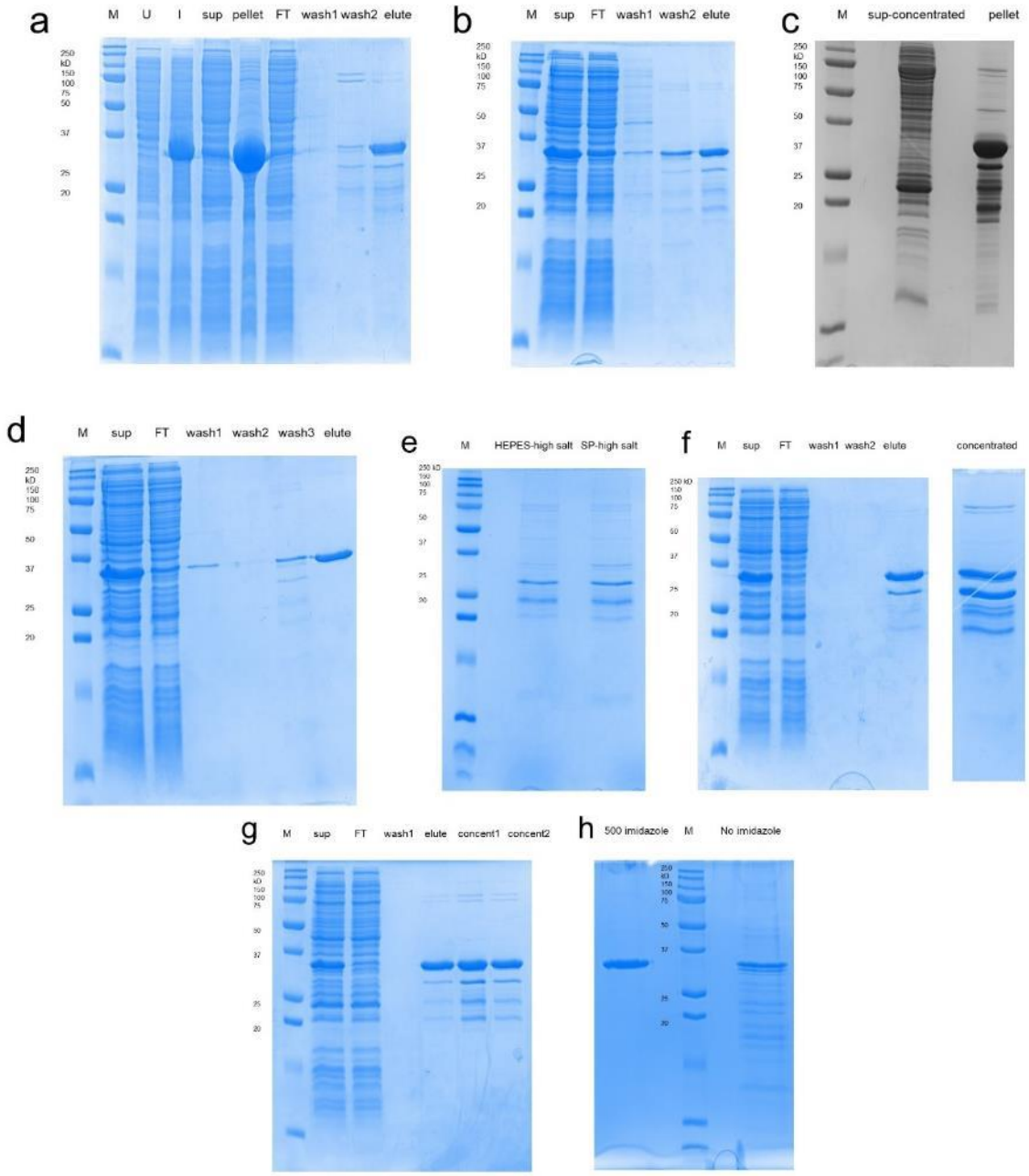


Figure 31 Purification attempts for MsmMpy

a, His-MsmMpy purification with low salt buffer. b, 100mM sodium phosphate low pH (pH 7.0) with high salt buffer. c, Dialysis against Tris low slat buffer (pH 7.7). d,100mM sodium phosphate normal pH (pH 7.5) with high salt buffer e, Buffer exchange using concentrator against 20mM sodium phosphate high salt buffer and 5mM HEPES high salt buffer. f, 500mM sodium chloride in HEPES buffer pH (pH 8.0). g, 500mM Ammonium chloride in HEPES buffer pH (pH 8.0) h, Buffer exchange to remove imidazole

Due to the high salt concentration (500 mM NH_4Cl and 500 mM imidazole) required to stabilize MsmMpy, we did not proceed to study its interaction with the ribosome. Instead, we experimented with a different construct, which has a His-tag at the C-terminus of MsmMpy, to test whether the new construct would provide better stability. The supernatant after cell lysis contains more soluble protein than the previous construct, and the yield significantly increased (Fig. 27). However, aggregation occurred again when we tried to remove imidazole from the elutions either through dialysis or concentrator.

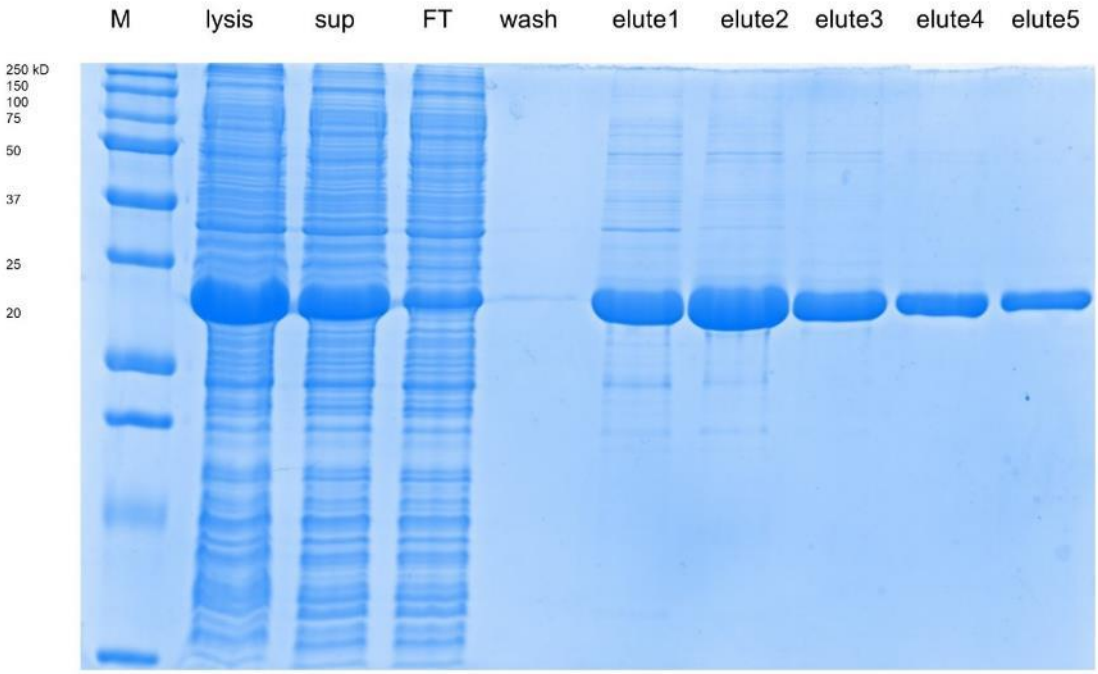


Figure 32 C-terminal His tag purification of MsmMpy

To systematically test how to maintain MsmMpy solubilized upon removal of imidazole, we diluted concentrated ~10mg/ml MsmMpy, which contains 500 mM imidazole, into variant buffers to test its solubility at a concentration of ~1mg/ml (Table 5). Interestingly, only buffer that either has a high NH_4Cl concentration or high

imidazole concentration (>150 mM) can keep MsmMpy soluble. Surprisingly, if the buffer contains 500mM NH_4Cl and 1mM EDTA, it also can help MsmMpy stay soluble. One possible explanation is that the elution from the Ni-NTA column has a small trace of nickel ions, which could induce protein precipitation. However, due to the high concentration of imidazole in the elution, its competition for nickel ions maintained the protein's solubility. Once imidazole is removed during dialysis or the concentration is lowered during buffer exchange using concentrator, the protein starts to aggregate due to binding to nickel ions with high affinity. On the other hand, EDTA neutralizes the effect of the small trace of nickel ions in the buffer completely.

Table 5 Solubility test for MsmMpy with C-terminal His tag in different buffers

	buffer1	buffer2	buffer3	buffer4	buffer5	buffer6	buffer7	buffer8	buffer9	buffer10	buffer11	buffer12	buffer13
HEPES (mM)	20	20	20	20	20	20	20	20	20	20	20	20	20
NH_4Cl (mM)	100	200	100	500	500	100	200	100	100	200	500	200	500
$MgCl_2$ (mM)	20	20	20	20	20	20	20	20	20	20	20		
KCl (mM)	50	50	50	50	50	50	50	50	50	50	50	50	50
Glycerol			10%	10%	20%								
Imidazole (mM)						50	50	150	250	50	50		
EDTA (mM)												1	1

III.3.3 MsmMpy inhibits translation in vitro

The final MsmMpy purification strategy was to use the C-terminal His-tag construct and 500mM NH_4Cl in the buffer. Elution from the Ni-NTA column was collected and titrated into a 10 ml buffer containing 10mM EDTA to prevent aggregation. Removal of imidazole was done using a protein concentrator. The concentrated MsmMpy was further purified using HiLoad 16/600 Superdex 200 pg (Fig. 28a). Two major peaks corresponding to the molecular weight around 320kDa and 160kDa are two different oligomeric states of MsmMpy. The long linker between the N terminal and C terminal domains in MsmMpy could alter its retention volume on the gel filtration column. However, the homolog from *Staphylococcus aureus* showed expected molecular weight (172). Therefore, MsmMpy could be both tetramer and octamer according to the molecular weight. Although the oligomeric state of MsmMpy still needs further validation, we proceeded with the functional analysis of MsmMpy using in vitro translation assay. MsmMpy was shown to inhibit the translation with an IC_{50} of about 2 μM (Fig. 28b).

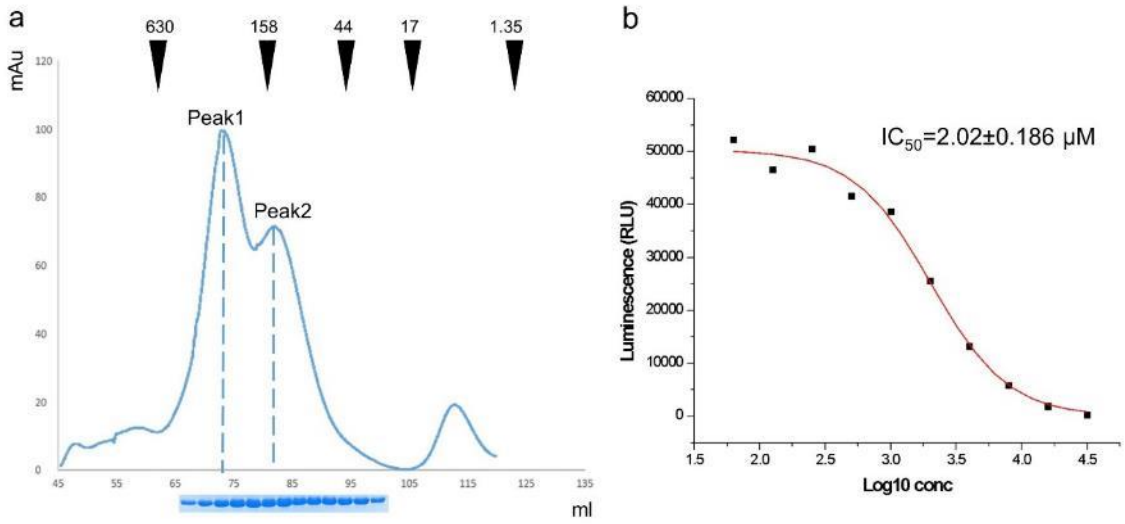


Figure 33 Gel filtration and in vitro translation analysis of MsmMpy
a, Gel filtration result showing two major peaks corresponding to around 240kDa and 120kDa. b, MsmMpy inhibits translation with an IC_{50} ~2 μM

III.3.4 In vitro reconstitution and cryo-EM studies of ribosome-Mpy complex

As MsmMpy displayed active inhibition toward translation in vitro, we set to study the ribosome-Mpy complex structure. We evaluated the interaction between MsmMpy and ribosome using sucrose gradient ultracentrifugation after reconstituting the complex in vitro. In agreement with the in vivo studies, reconstituting the ribosome-Mpy complex in vitro did not result in the formation of 100S (Fig. 29a). SDS-PAGE analysis of the ribosome peak showed that MsmMpy co-migrated with *Msm* ribosome (Fig. 29b), while MsmMpy alone was on the near top of the sucrose gradient (Fig. 29a), indicating a stable complex was formed.

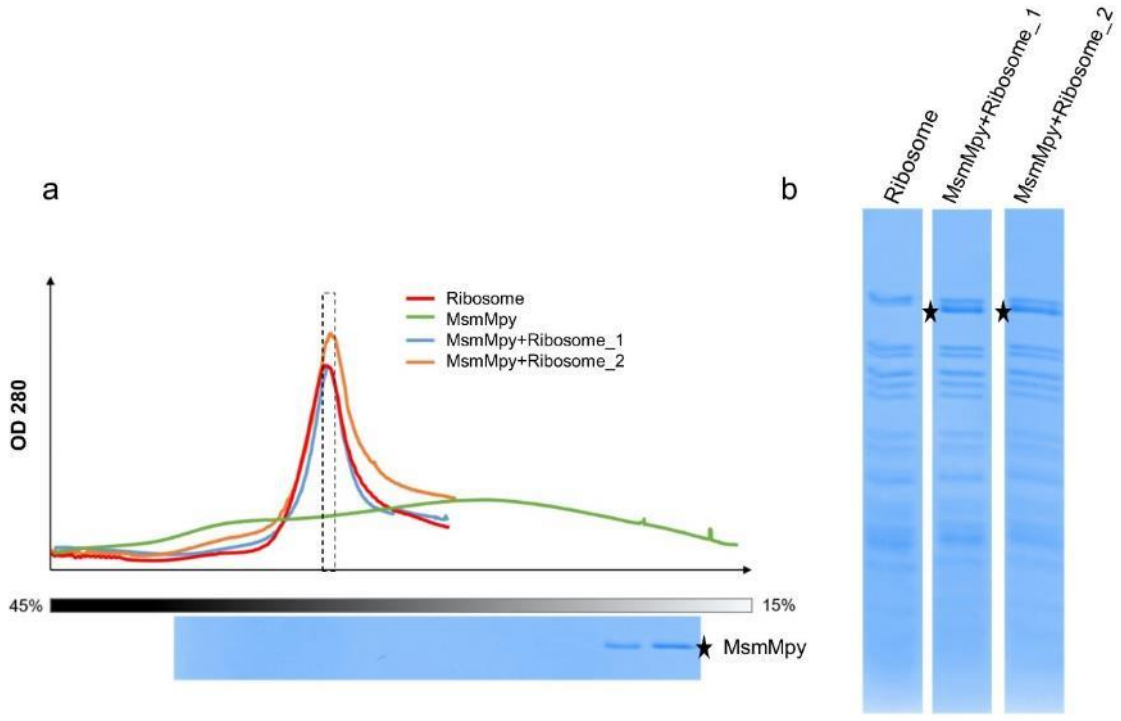


Figure 34 Sucrose gradient analysis of MsmMpy-70S complex
a, chromatograms of different experimental conditions. B, SDS-PAGE analysis of the ribosome fraction. The star indicates the band of MsmMpy.

Samples for cryo-EM studies were made from the sucrose gradient fraction that contains *Msm* ribosome and MsmMpy. However, the preliminary structures of the ribosome did not reveal the density of MsmMpy. 3D classification results showed different states of Msm ribosome (Fig. 30a), while the density of MsmMpy was not observed in the A and P tRNA sites on the 30S, where the N-terminal domain was expected to be. Instead, tRNA density is present at the P tRNA site (Fig. 30c).

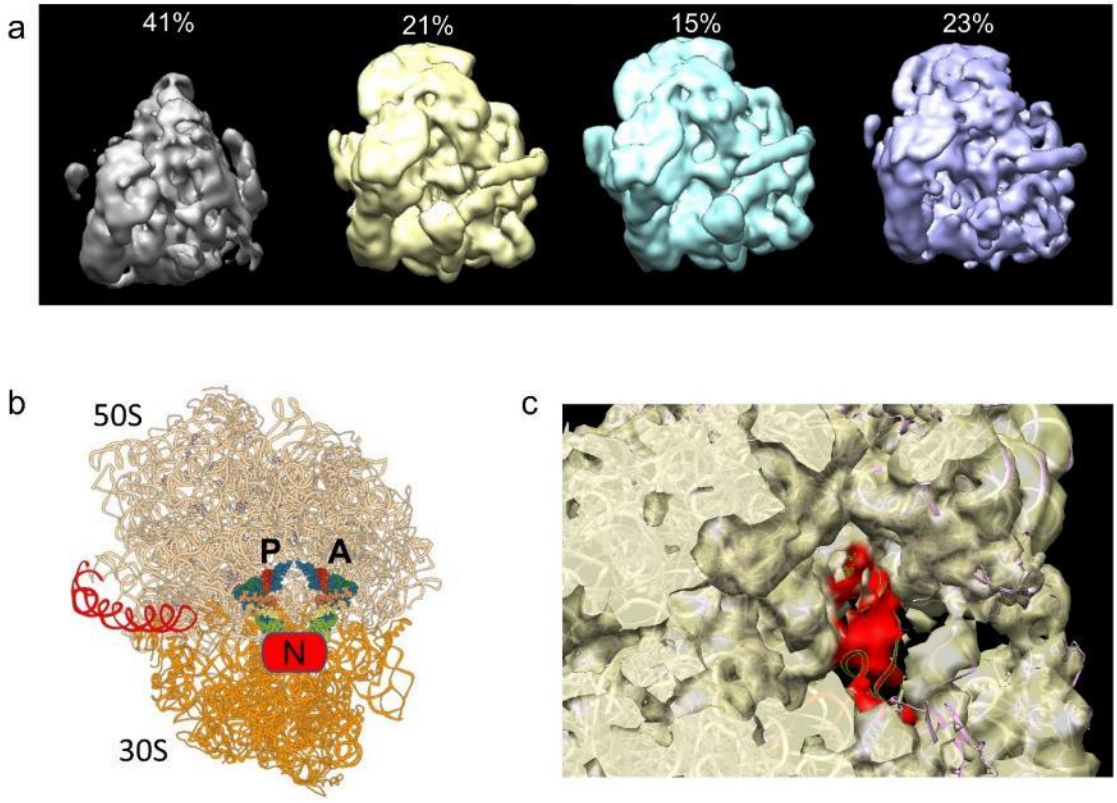


Figure 35 Preliminary cryo-EM structures of MsmMpy-70S complex

a, Different subpopulations of ribosome. b, Potential MsmMpy binding site on ribosome. c, Density map showing the MsmMpy binding pocked, however the density colored in red resembles the P site tRNA instead of N-terminal domain of MsmMpy.

III.3.5 Absence of interactions between Mpy and Mrf

While we were analyzing our initial cryo-EM results, the cryo-EM structure of MsmMpy with Msm ribosome was published. More importantly, Li et al. showed that RafS (renamed as Mpy) only bound to ribosome when the cell was grown at zinc-depleted medium and is mediated by Mrf. However, the detailed molecular mechanism involving zinc in the binding of Mpy to the ribosome in *Mtb* is unknown. Besides, the C-terminal domain of MsmMpy is not resolved in the published structure.

Hence, we changed course to investigate the potential interaction between Mpy and Mrf and zinc ions' possible roles in the interaction between Mpy and ribosome. We tried to purify MsmMrf with His-sumo tag or MBP tag at the N-terminus. Although both are soluble after elution and concentration, SEC results revealed that they are soluble aggregates, displaying more than 600kDa molecular weight (Fig. 31). We then tested co-purification of MsmMrf and MsmMpy, by mixing the cells expressing His-sumo-tagged MsmMrf and untagged MsmMpy before lysis (Fig. 32a). Untagged MsmMpy is shown to be purified together with His-sumo-tagged MsmMrf. Interestingly, a similar amount of MsmMpy and MsmMrf were present in the early elution, while more MsmMpy was eluted later. It also appears that MsmMrf can prevent MsmMpy from aggregating since the early elution did not precipitate. After removing the His-sumo tag on MsmMrf, we ran a gel filtration to examine the interactions between MsmMrf and MsmMpy (Fig. 32b). Unfortunately, MsmMrf and MsmMpy did not co-migrate on the gel filtration column (Fig. 32c). Nevertheless, MsmMrf still showed soluble aggregates, which could potentially compromise its interactions with MsmMpy.

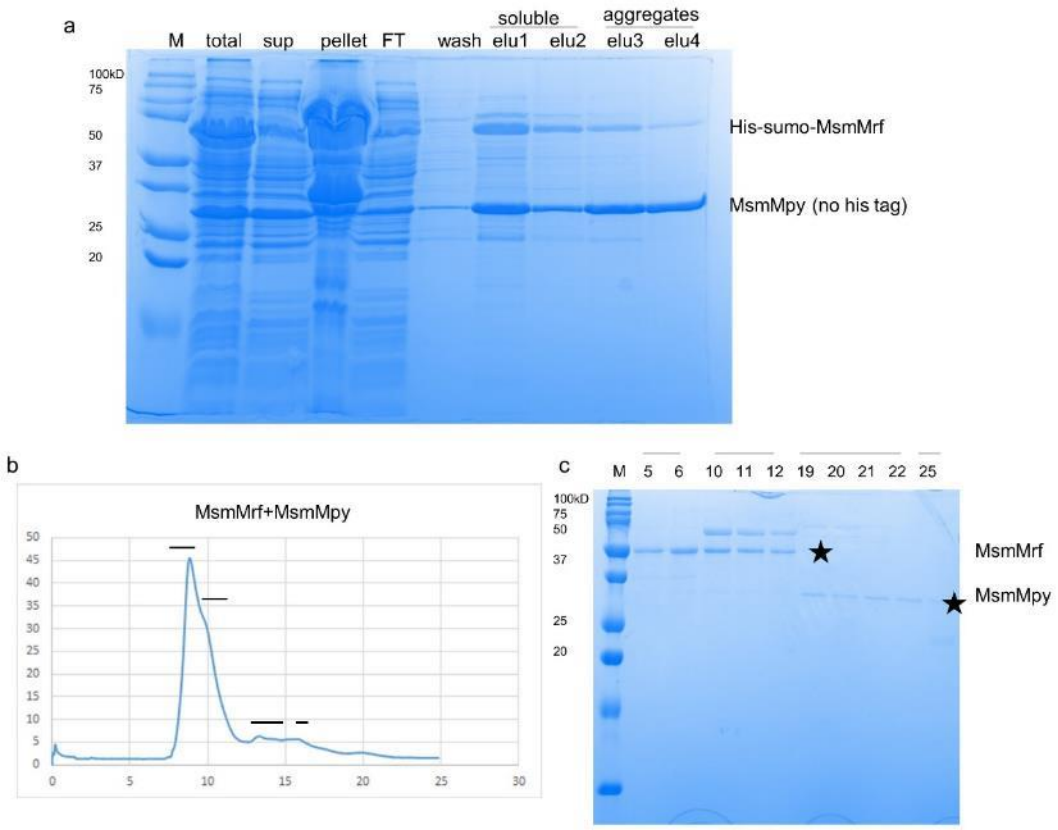


Figure 37 Co-purification of MsmMpy and MsmMrf

a, Untagged MsmMpy is co-purified with His-sumo tagged MsmMrf. b, Gel filtration of co-purified MsmMpy and MsmMrf. c, SDS-PAGE analysis of the peaks on the chromatogram. Co-migration of MsmMrf and MsmMpy is not observed MBP tags.

III.3.6 Full-length and C-terminal domain of untagged MtbMpy can be purified by Ni-NTA resin

From the observation that untagged MsmMpy was eluted after MsmMrf fully eluted from the column during co-purification above, we deduce that untagged Mpy could potentially bind Ni-NTA as well. Also, considering zinc molecules' role in regulating the interactions between Mpy and ribosome, we set out to explore the zinc-binding property of Mpy from *Mtb*. Due to the promiscuity of many zinc-binding proteins in distinguishing zinc and nickel molecules, we initiated the experiment to test whether untagged MtbMpy can be purified by Ni-NTA resin, which could be an indicator of zinc-binding property. The full-length untagged MtbMpy construct was overexpressed in Rosetta 2(DE3) cells. Surprisingly, it can be purified by Ni-NTA resin (Fig. 33), indicating that MtbMpy could bind zinc molecules as well. In order to pinpoint the potential zinc-binding motif, we overexpressed different constructs of MtbMpy, including the N-terminal domain, N-terminal domain with linker region, C-terminal domain with linker region, and C-terminal domain. Only the C-terminal domain is present after purification using Ni-NTA resin, suggesting that the C-terminal domain of MtbMpy is responsible for potential zinc-binding (Fig).

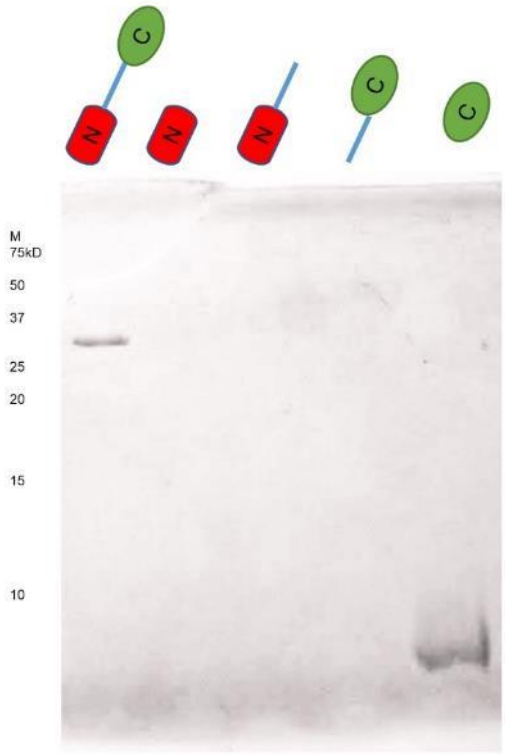


Figure 38 Purification of full-length and different domains of MtbMpy

III.3.7 Tetrameric state is the dominant configuration of MtbMpy C-terminal domain

C-terminal domains of LHPFs often form a dimer, therefore mediating the formation of 100S. However, the absence of 100S in mycobacteria raises the question of the configuration of the C-terminal domain of Mpy. According to the size exclusion chromatography result of the Mpy C-terminal domain, we observed two peaks corresponding to the molecular weight around 24 kDa and 48 kDa, respectively (Fig. 34). The estimated molecular weight indicates tetramer and octamer formation, which is

in agreement with the full-length MsmMpy. We further investigated the oligomeric state of MsmMpy using an on-column cross-linking method. Different oligomeric states were detected on the SDS-PAGE gel, including tetramer and higher-order oligomer population, potentially due to the mild nature of the cross-linking method (Fig. 35).

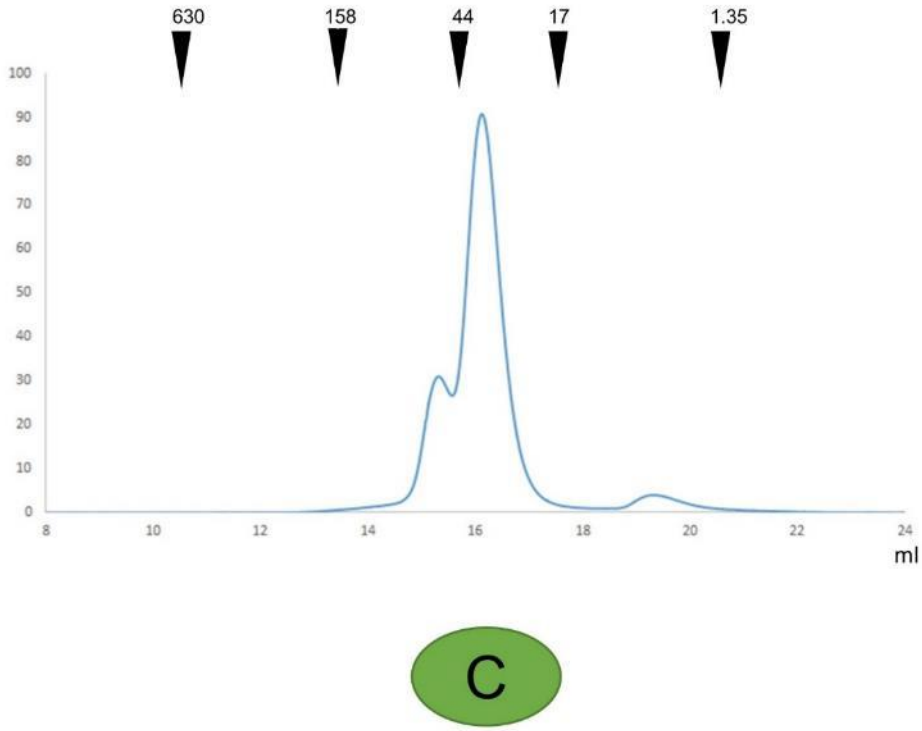


Figure 39 Gel filtration of purified C-terminal domain of MtbMpy

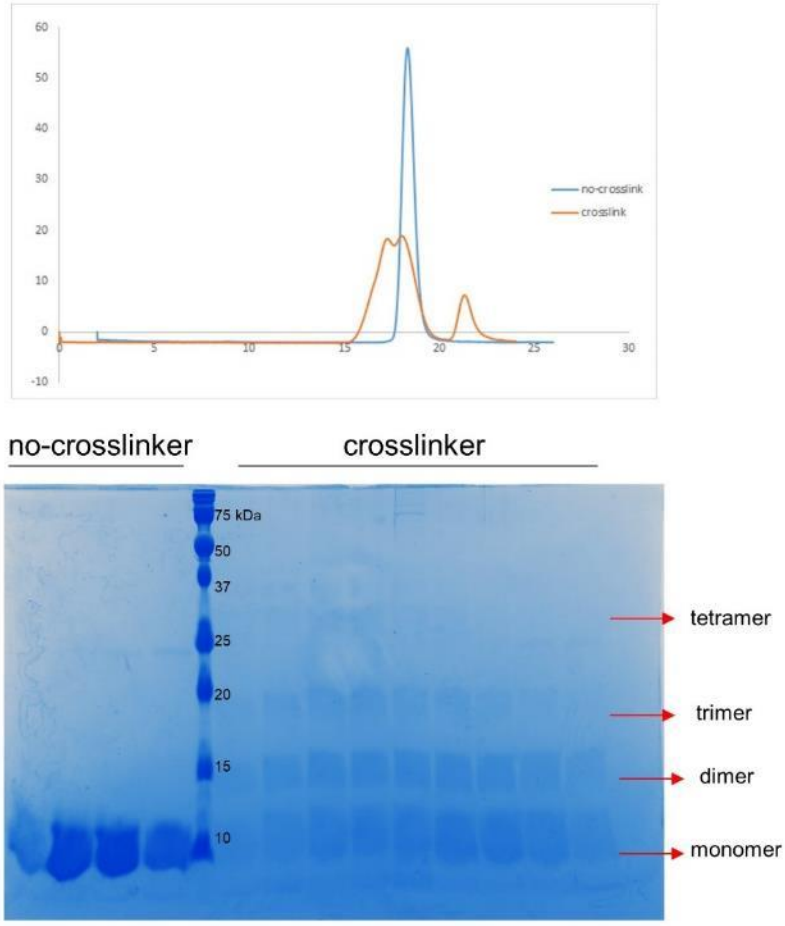


Figure 40 On-column cross-linking C-terminal domain of MtbMpy
Gel filtration profile and SDS-PAGE revealed high oligomerization states.

Structural characterization of the C-terminal domain of MtbMpy was unsuccessful by either X-ray crystallography or NMR. No crystals were observed by the hanging drop method in nearly 600 conditions. Although we can detect the secondary structures using NMR, we did not get 3D information due to the protein's instability in the required low salt buffer (Fig. 36).

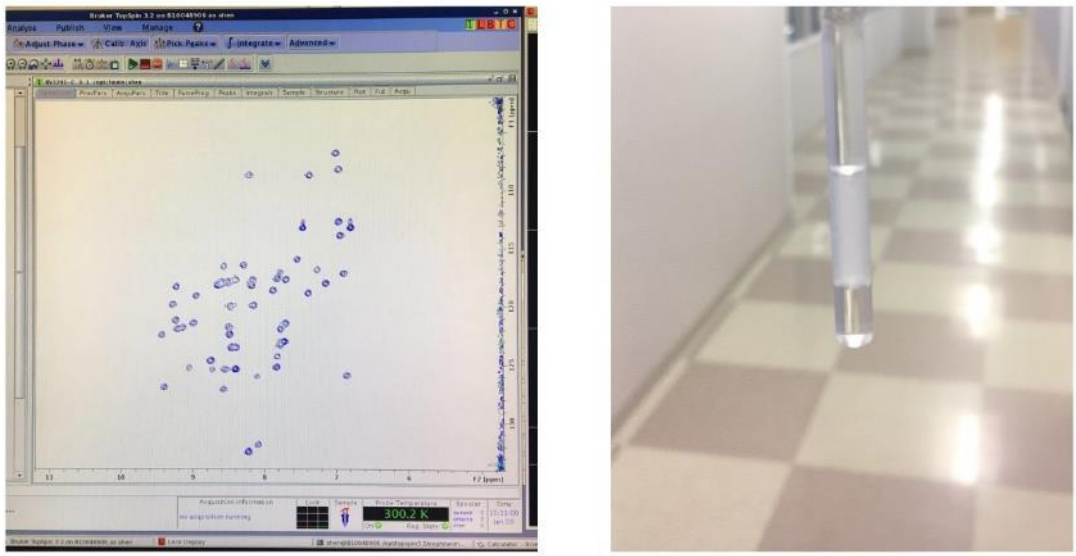


Figure 41 Structural characterization of MtbMpy C-terminal domain using NMR

Although the it shows stable secondary structure, the signal is week and protein aggregated in less than 24 hrs.

III.3.8 Reversible Zinc-mediated polymerization and aggregation of C-terminal domain of MtbMpy

Since the C-terminal domain of MtbMpy can be purified with Ni-NTA resin without a His-tag, we set to investigate the effect of zinc molecules on MtbMpy. We employed dynamic light scattering to estimate the distribution of protein particle

diameter. The average particle diameter of the purified C-terminal domain of MtbMpy is around 6 nm (Fig. 37). However, upon adding zinc ions, visible aggregation was observed, and the averaged particle diameter was more than 2 μm. The precipitation is not due to denaturation but instead mediated by zinc ions since it is reversible upon adding EDTA to the aggregates.

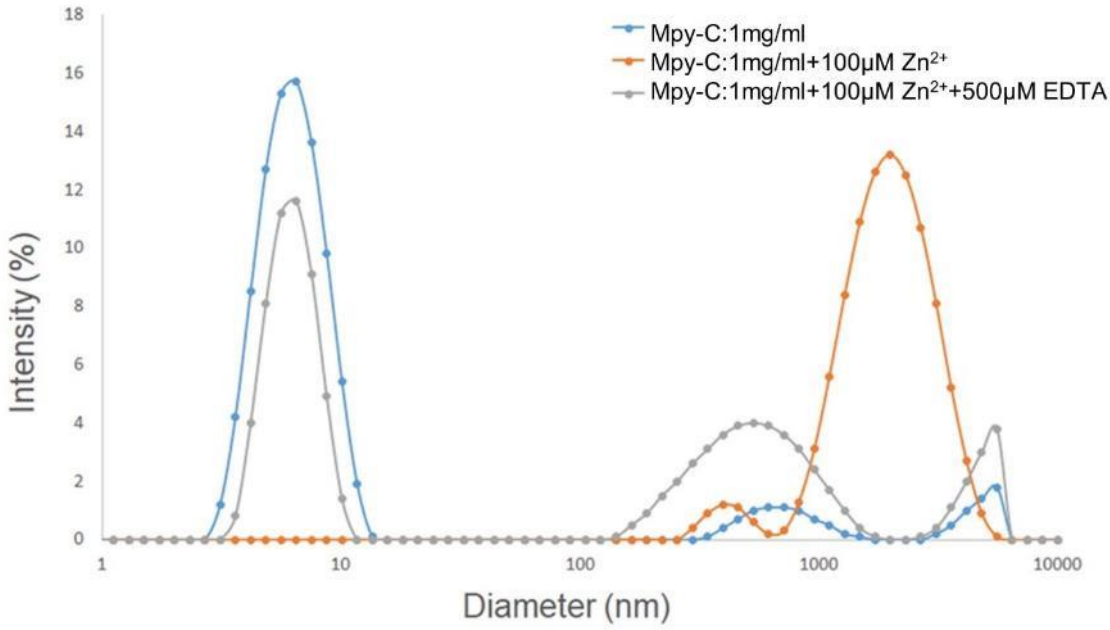


Figure 42 Dynamic light scattering profile of MtbMpy C-terminal domain.

We also used negative staining electron microscopy to image the aggregation induced by Zinc molecules. Although most aggregates were clumps of proteins (Fig. 38), we observed filament-like structures, ranging from 1μm to 3 μm long and more than 50 nm wide. Different high-order structures could form based on the different molar ratios of protein and zinc molecules.

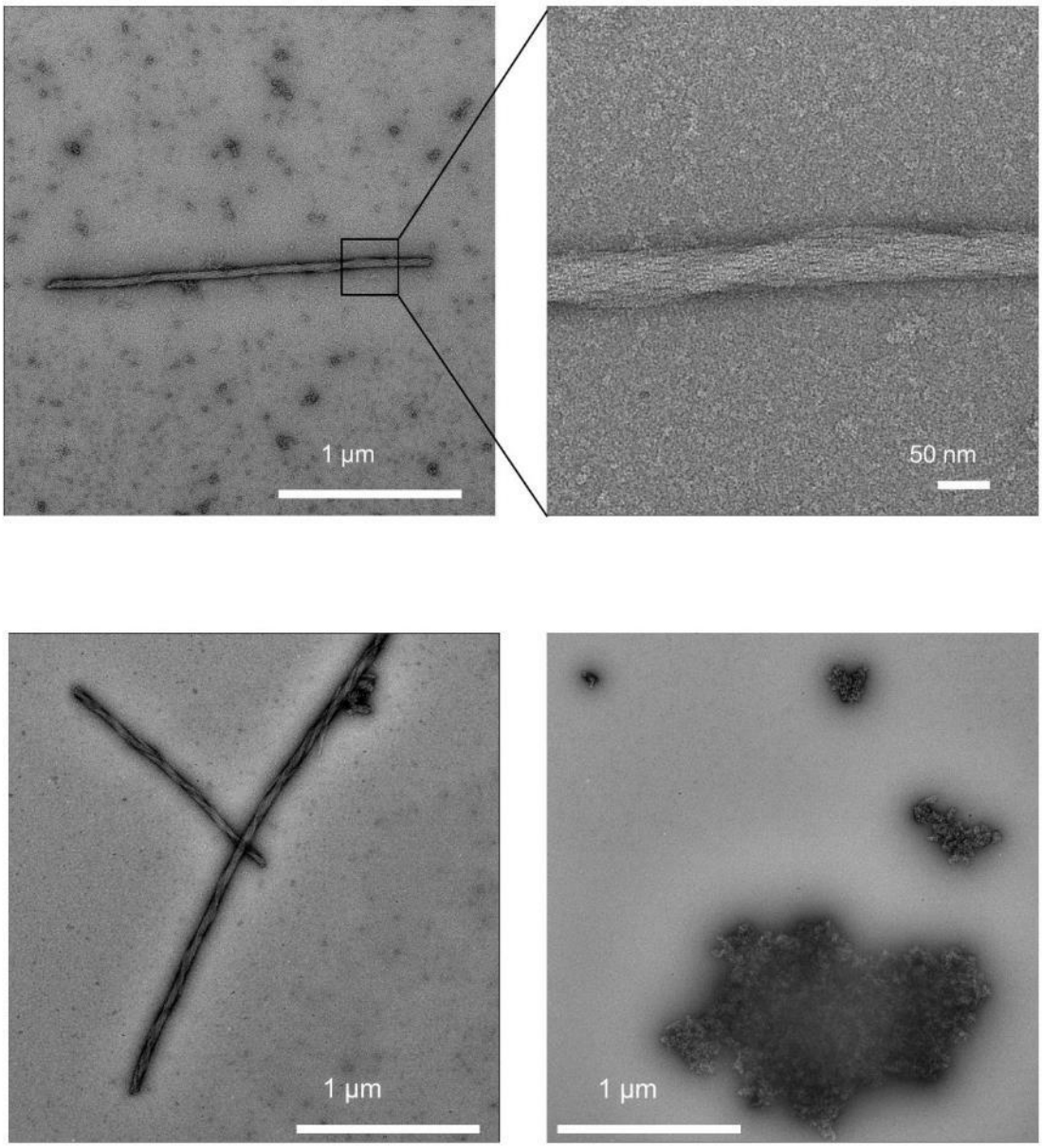


Figure 43 Negative staining analysis of the aggregated MtbMpy C-terminal domain

III.3.9 Attempts to identify Zinc-binding residues in the C-terminal domain of MtbMpy

By building a dimer structure of the C-terminal domain of MtbMpy based on the available homologous structures using the SWISS model, we suspect that possible

histidine clusters are present in the dimer interface (Fig. 39a). Mutant H170F and a double mutant H170F/P171S were expressed. However, both were still able to be purified by Ni-NTA resin (Fig. 39b). Further mutations, including other histidines, are needed to address the zinc-binding residues in the C-terminal domain of MtbMpy.

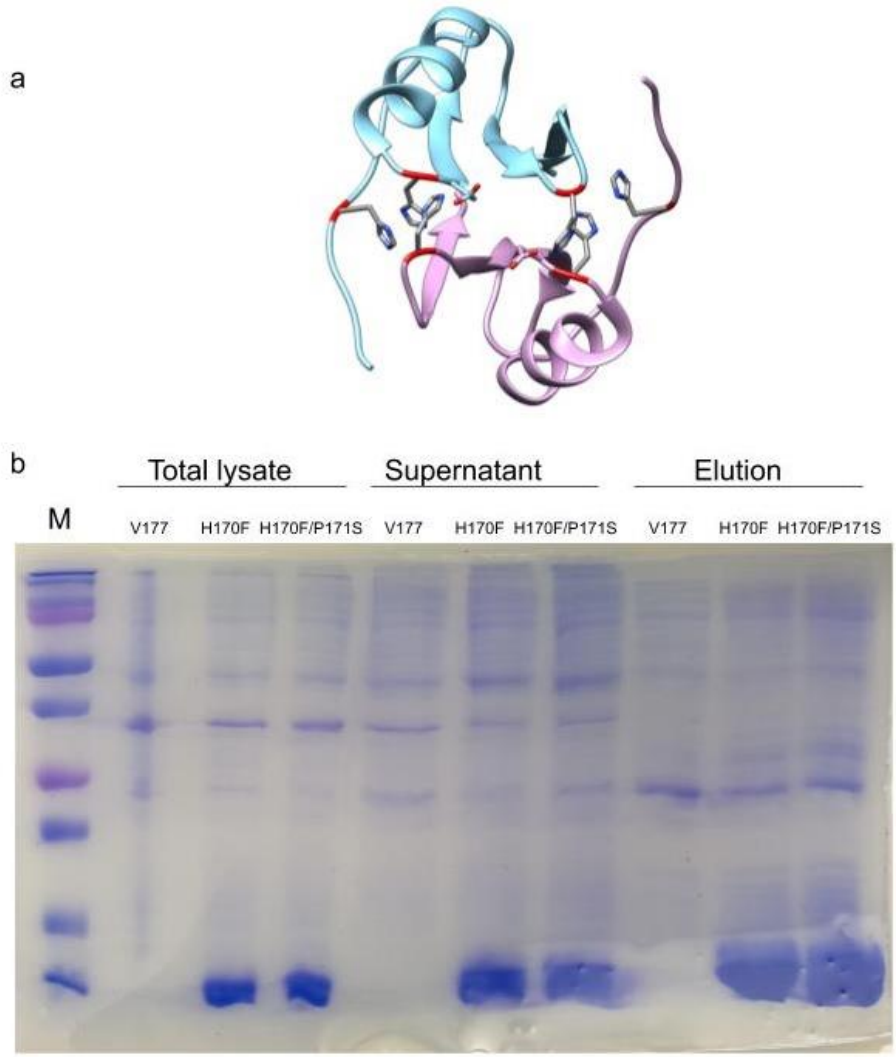


Figure 44 Potential histidine cluster and mutagenesis study a, Structure of predicted C-terminal dimer indicates the potential histidine at the dimer interface. b, preliminary mutagenesis studies shows that H170 is not critical for Zinc-binding.

III.4 Discussion

Here we explored the translational regulation related to the persistency of *Mtb*, primarily focused on the factors during hypoxia and stationary phase. According to our preliminary results, whether there is a direct interaction between RafH and ribosome is ambiguous. However, as stated previously, more thorough experiments are needed to fully address the question. For example, different purification tags could interfere with the results. In addition, inferred from the studies of Mpy, whether RafH needs other factors to facilitate its interaction with ribosome is still unknown and to be explored.

As for the investigation of Mpy, we initially spent extensive efforts purifying and characterizing its interaction with the ribosome without success. Li et al. showed two necessary conditions for Mpy to bind to ribosome, Mrf, and zinc-depletion in the medium. We then changed the course to study its interaction with Mrf and accidentally found out its zinc-binding ability. Although the direct interaction between Mpy and Mrf was not detected, it still needed further declaration as Mrf were soluble aggregates in vitro under our experimental conditions.

Surprisingly, Mpy can be purified by Ni-NTA resin without any tag, which indicates that it can also bind zinc molecules. Further investigations revealed that the C-terminal domain of Mpy is responsible for the zinc-binding. It is reasonable to speculate that the C-terminal domain of Mpy serves as a regulatory domain since the N-terminal domain is the functional domain that binds to the ribosome. The zinc molecule could act as a trigger for a switch-mechanism. Mpy would polymerize when an excess amount of zinc molecules are present in the cell, preventing it from binding to the ribosome. When

there are fewer zinc molecules present in the cell, potentially through the help of Mrf, Mpy returns to tetrameric form, which will interact with the ribosome (Fig. 40).

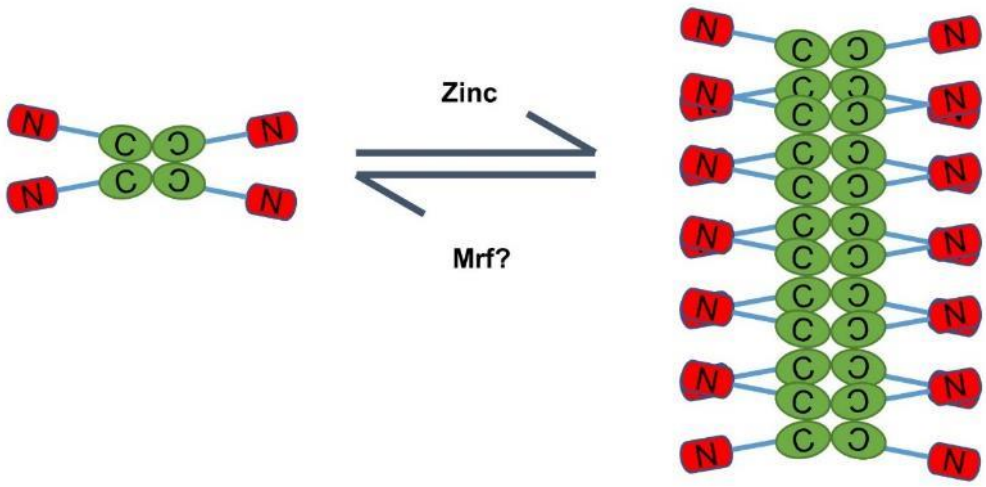


Figure 45 Potential Zinc-mediated switch mechanism of Mpy

The unique property of the C-terminal domain of Mpy compared to its homologs in other bacteria could be inferred from the fact that no 100S is present in mycobacteria, which is typically mediated by the dimerization of the C-terminal domain. Indeed, the C-terminal domain of Mpy appears to be dominantly tetrameric, potentially blocking the formation of 100S.

In conclusion, we propose a direct zinc-mediated switch mechanism of Mpy in this study, which provides an alternative explanation for regulating ribosome hibernation in mycobacteria. We also reveal the possible explanation for the absence of 100S in mycobacteria, which is due to differences in the oligomerization states of the C-terminal domain of Mpy.

CHAPTER IV

NATURAL PRODUCT DERIVED SEQUANAMYCINS OVERCOME METHYLATION-INDUCED MACROLIDE RESISTANCE IN *MTB* RIBOSOME

IV.1 Introduction

As one of the most successful human pathogens, *Mycobacterium tuberculosis* not only can escape the immune system to remain dormant for decades but also show resistance against many commonly used antibiotics (73). The currently recommended treatment for new cases of drug-susceptible TB (DS-TB) is a six-month regimen of four first-line drugs: isoniazid (H), rifampicin (R), ethambutol (E), and pyrazinamide (Z) (17). It is even more challenging to treat TB-HIV co-infection. The alarmingly increasing burden of drug-resistant-TB cases, defined as resistant to at least isoniazid and rifampicin (multi-drug resistance, MDR) or resistant to even more TB drugs (extensively drug-resistant, XDR), has been documented in the past decade (5). With limited therapeutic options to efficiently target drug-resistant TB, the treatment requires much longer, more expensive, and more toxic drugs, while with only a 50% success rate toward MDR-TB. A new 9-month treatment regimen, with a potential success rate of 80%, has been recommended by WHO since 2016 (5).

Due to the inevitably increasing burden of MDR-TB and XDR-TB, there is a substantial need to develop novel anit-TB drug combinations to maximize efficacy, limit resistance to the individual components, and shorten the treatment. Several new TB treatment regimens for both drug-susceptible and drug-resistant TB combining new or

repurposed anti-TB drugs are being tested in Phase II or Phase III trials (179). However, only three novel classes without pre-existing resistance are included in these regimens: nitroimidazoles (Pretomanid or Delamanid), diarylquinolines (Bedaquiline), and oxazolidinones (Linezolid, Sutezolid) (http://www.newtbdrugs.org/).

Although nearly half of clinically used antibiotics target ribosomes, it has not been successful in treating TB by targeting the *Mtb* ribosome (180). However, repositioning the old antibiotic Linezolid has been recently attempted to compensate for the lack of TB drugs that target the *Mtb* ribosome. Despite its well-known toxicity, Linezolid demonstrated a clear impact on XDR treatment in a clinical trial conducted in Korea (181) (182). Besides, recent data in preclinical models using Linezolid and Sutezolid demonstrated the potential of oxazolidinones to reduce the treatment duration needed for the cure when given in combination with other novel compounds (183). Taken together, it suggests that ribosome is not only a traditionally attractive target for classic bacterial infections but also for TB.

Macrolides are a widely-used class of antibiotics that bind at the nascent peptide exit tunnel of the ribosome to block protein biosynthesis (184). Although this class of antibiotics is often preferred prescribed drugs against many bacterial infections, it has been established that the 14-membered macrolides do not work against *Mtb* due to intrinsic resistance (185) (81). When *Mtb* is treated with macrolide antibiotics such as clarithromycin and erythromycin, the *erm37* gene is induced, resulting in a significant reduction in their activity. Unlike *erm* genes in other bacteria, the detailed induction mechanism in *Mtb* is unknown as no lead sequence is found upstream of the start codon

of *erm37*. Gene *erm37* encodes a methyltransferase responsible for the methylation of the *Mtb* ribosome in the 23S rRNA at position A2296, a residue critical for macrolide binding, and also potentially at neighboring nucleotides A2295 and A2297, which could further reducing the binding affinity for macrolides (83). Therefore, the usage of macrolides against *Mtb* is less than desirable. Although clarithromycin is occasionally used in salvage regimens for highly drug-resistant TB, it has never demonstrated significant efficacy against TB in the clinical setting or preclinical models. Aso, chemical optimization of the 14-membered ring macrolides has not met with success (186) (187).

Most efforts to discover anti-TB drugs were stopped in the 1970s when rifampicin was approved for TB therapy (188). With the urgent need to find novel anti-TB drugs in mind, the Sanofi chemical patrimony containing former TB drug programs was re-explored (15). The natural product Sequanamycin A (SEQ-503), belonging to a 14-member ring erythromycin family, was considered a good starting point for a medicinal chemistry optimization program (189). Here, we describe the in vitro activities of the advanced lead compound SEQ-9, demonstrating its efficacy to inhibit the A2296-methylated *Mtb* ribosome, as well as the cryo-EM structure of SEQ-9 in complex with A2296-methylated *Mtb* ribosome to provide detailed molecular information for structure-based drug design (SBDD)

IV.2 Materials and methods

IV.2.1 *Mtb* S30 and S100 preparation

The purified S30 fraction from *Mtb* is prepared according to a previous publication (190) with modifications. Briefly, cell lysates are centrifuged twice at 30,000 *g* for 30 min at 4˚C. The supernatant is collected and 3 mL of pre-incubation buffer (0.3 M Tris-acetate, pH 8.2, 14.0 mM $Mg(Ac)_2$, 12.0 mM ATP, 4.4 mM DTT, 0.04 mM 20 amino acids, 6.7 U/mL pyruvate kinase, 84.0 mM PEP) is added to every 10 mL of cell extract. The mixture is wrapped in aluminum foil and incubated at 37°C on a rotary shaker (120 rpm) for 80 min. Following the pre-incubation step, the S30 extract is dialyzed against four exchanges of 20 volumes of S30 buffer (10 mM tris-acetate pH 8.2, 14.0 mM $Mg(Ac)_2$, 60 mM KAc, 1 mM DTT) for 45 min each. The dialyzed extract is finally centrifuged at 4000 *g* for 10 min at 4°C. Fraction S100 is made from S30 by centrifugation at 100,000 *g* for 2 hours at 4 ˚C on Beckman Type Ti50.2 rotor to remove ribosome.

IV.2.2 Methylated *Mtb* ribosome purification

Mtb ribosomes are purified as previously described, except that methylated ribosomes are prepared from *Mtb* culture with 1 μg/ml of erythromycin. Briefly, cell lysates are clarified by centrifugation at 30,000 *g* for 1 h. The supernatant is pelleted in sucrose cushion buffer (20 mM HEPES pH 7.5, 1.1 M sucrose, 10 mM $MgCl_2$, 0.5 M KCl, and 0.5 mM EDTA) at 185,511 *g* in a Beckman Type 45Ti rotor for 20 h. The pellet is resuspended in a buffer containing 20 mM Tris–HCl pH 7.5, 1.5 M $(NH4)_2SO_4$,

0.4 M KCl, and 10 mM $MgCl_2$. The suspension is then applied to a hydrophobic interaction column (Toyopearl Butyl-650S) and eluted with a reverse ionic strength gradient from 1.5 M to 0 M $(NH_4)_2SO_4$ in a buffer containing 20 mM Tris–HCl pH 7.5, 0.4 M KCl, and 10 mM $MgCl_2$. The eluted ribosome peak is changed to reassociation buffer (5 mM HEPES-NaOH, pH 7.5, 10 mM NH_4Cl, 50 mM KCl, 10 mM $MgCl_2$, and 6 mM β-mercaptoethanol) then concentrated before applying to a 10–40% linear sucrose gradient centrifuged in a Beckman SW28 rotor at 64,921 *g* for 19 h. The 70S fractions are concentrated to about A260 = 150 after removal of the sucrose.

IV.2.3 IC50 measurement of different compounds

The *Mtb* S100 extract (15 μL, containing translation factors such as initiation, elongation, termination, recycling factors, and aminoacyl tRNA synthetase) mixed with 10 μL 10X salt buffer (2 M potassium glutamate, 0.8 M ammonium acetate, and 0.16 M magnesium acetate), 0.5 mM each of the 20 amino acids, 16.7 mM PEP and 1% poly (ethylene glycol) 8000. This mixture (41 μL) is added to each well and mixed with wild-type or erythromycin-treated ribosome to the final concentration at 100 nM.

We next add 2 μL compound of each concentration (40 μM, 4 μM, 2 μM, 1 μM, 400 nM, 200 nM, 100 nM, 40 nM, 20 nM, 10 nM, 4 nM, 0) to the 41 μL mixture, and allow incubating for 10 minutes at room temperature prior to adding master mix and mRNA. The reaction is started by adding nanoluciferase mRNA (100 ng in 2 μL) and 5 μL 5X master mix (286 mM HEPES-KOH, pH7.5, 6 mM ATP, 4.3 mM GTP, 333 μM folinic acid, 853 μg/mL tRNA). The final volume is 50 μL. The reaction is allowed to

proceed for one hour at 37 degrees and then terminated by the addition of 80 μM Chloramphenicol. The luminescent signal is detected by the addition of 20 μL of the nano-luciferase substrate Furimazine in S30 buffer.

IV.2.4 Cryo-electron microscopy and Data processing

SEQ-9-bound A2296-methylated *Mtb* 70S was prepared by incubating 800 nM methylated 70S with 16 μM SEQ-9 on ice for 30 min. A volume of 3 μl of the sample was applied to a glow-discharged Quantifoil R2/2 holey carbon grid (300 mesh) and vitrified using a Vitrobot Mark III (FEI company, The Netherlands) at 22°C with 100% relative humidity. Cryo-EM data were collected under a Titan Krios electron microscope (89) operating at 300 kV, at a nominal magnification of 130,000, which yields a pixel size of 1.06 Å/pixel. Image stacks were recorded on a Gatan K2 Summit (Gatan, Pleasanton CA, USA) direct detection camera in the electron counting mode. A total exposure time of 8.0 s with 0.2 s intervals, and a dose rate of ~6.0 electrons/$Å^2$/s was used, resulting in 40 frames per image stack and an accumulated total dose of ~48 electrons/$Å^2$.

Drift correction of collected image stacks was done by MotionCor2 with dose weighting (191). Aligned and summed image stacks were subjected to CTF estimation using gCTF (143). Images showing ice contamination and low resolution according to the estimation from gCTF were discarded, resulting in a total of 6,481 micrographs. Particle picking and reference-free 2D class average were done by Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/) and RELION-3.0 (144), respectively. A total

of 1,065,544 clean particles were selected and refined into one consensus map. Then 3D classification was used with the option '--skip_align' to classify ribosomes' different conformations. Furthermore, each state was processed according to the pipeline of RELION-3.0. The overall resolution was estimated according to the gold-standard Fourier shell correlation (145).

IV.2.5 Molecular modeling and refinement

To build the atomic model for methylated 70S bound with SEQ-9, we used our previously published structure from Mycobacterium tuberculosis (PDB 5V93) (49), which was fitted into the calculated cryo-EM map using rigid body fitting in Chimera (154). The docked model was then manually adjusted to refine into the cryo-em map using COOT (153). ERRASER (102) was used to improve the RNA backbone geometry. Ligand restraints of SEQ-9 were generated by eLBOW (192). The SEQ-9 molecule was manually docked into the extra density around the PTC region and refined using COOT. Iterative refinement using real-space refinement in PHENIX (152) and COOT was performed to improve the geometry and local fitting.

IV.3 Results

IV.3.1 A2296-methylated *Mtb* ribosome is resistant to macrolides

To purify the A2296-methylated *Mtb* ribosome, we cultured *Mtb* with erythromycin to induce the expression of Erm37, which will methylate A2296 required for the intrinsic resistance. In order to test whether the purified ribosome is resistant to

macrolide, we used in vitro translation assay. As expected, the IC_{50} values of a different generation of macrolides were greatly increased with the ribosome purified from cells cultured with erythromycin, indicating that ribosomes were methylated (Fig. 41). IC_{50} values of clarithromycin and erythromycin increased 100 to 200 folds from 353.9 nM to 88,042.6 nM and 670.5 nM to 1399422.9 nM, respectively. We also tested the IC_{50} of other non-macrolides, including capreomycin, chloramphenicol, linezolid, and streptogramin. Only a modest difference between unmethylated and methylated ribosomes was observed, confirming that the macrolide-resistance is due to A2296-methylation (Fig. 42) (Table 6).

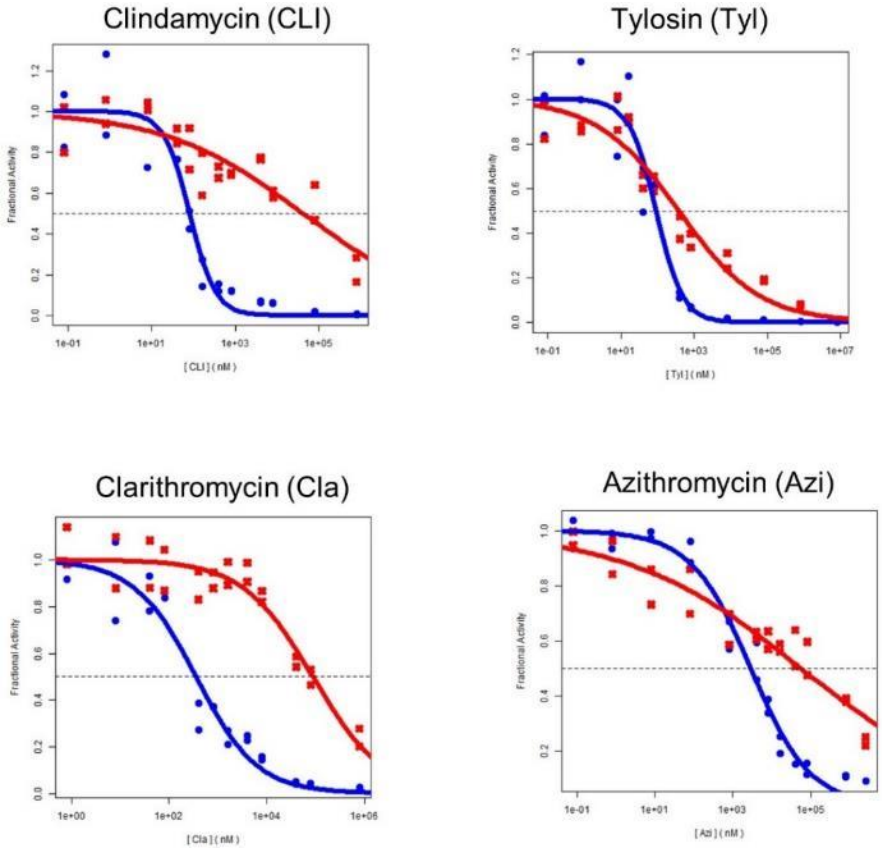


Figure 46 Methylated *Mtb* ribosome is resistant to different generations of macrolides.

In vitro translation assay with wild type ribosome (blue) and methylated ribosome (red)

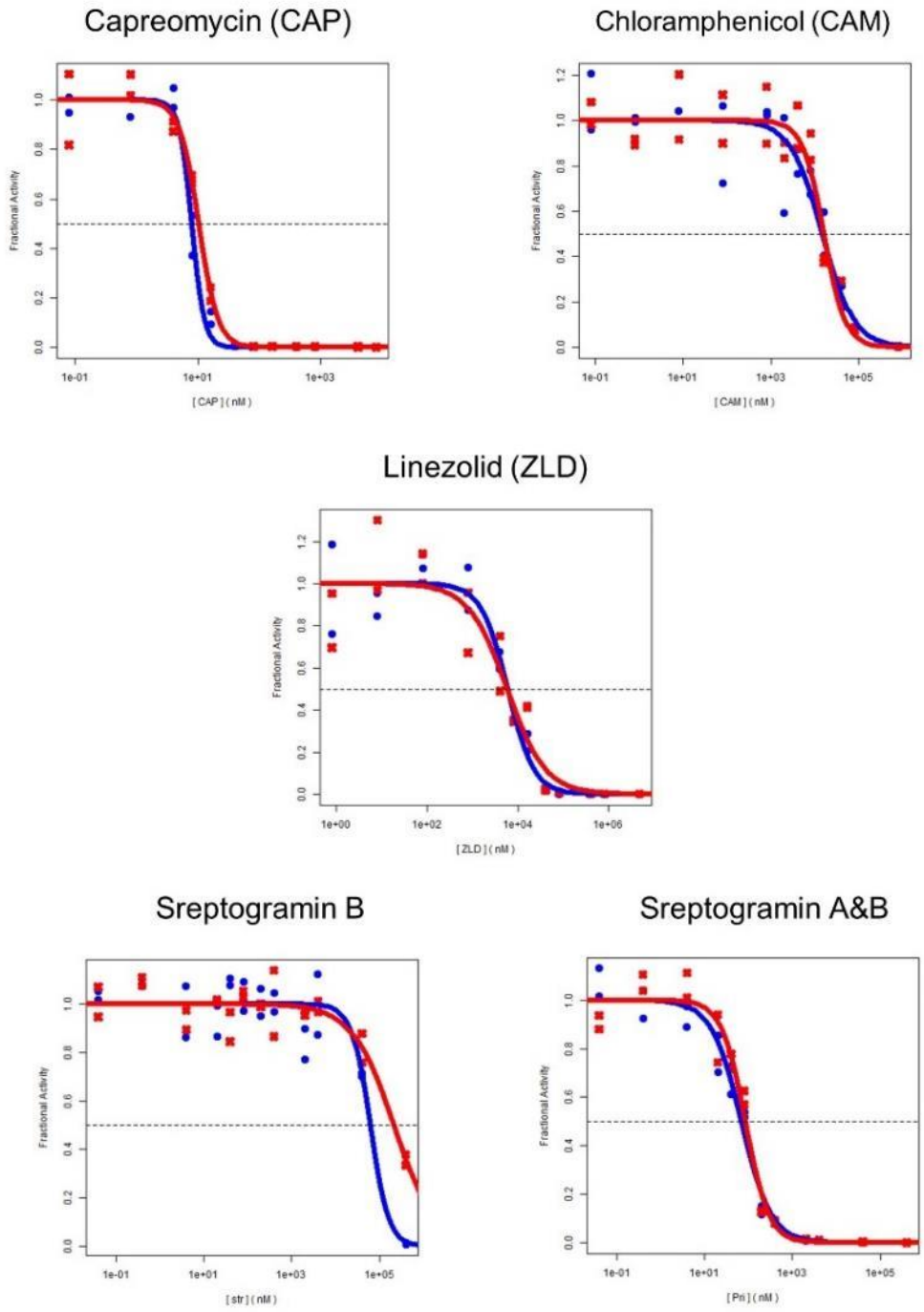


Figure 47 Methylated *Mtb* ribosome is NOT resistant to other antibiotics.

In vitro translation assay with wild type ribosome (blue) and methylated ribosome (red)

Table 6 Summary of the IC50 of different antibiotic toward unmethylated and A2296-methylated *Mtb* ribosome

	unmethylated	A2296-methylated
Clarithromycin (nM)	353.9	88042.6
Azithromycin (μM)	2.9	
Tylosin (nM)	89.2	377.1
Erythromycin (nM)	670.5	139942.9
Capreomycin (nM)	7.9	10.2
Chloramphenicol (μM)	14.9	16.3
Linezolid (μM)	5.9	5.8
Streptogramin (nM)	68.6	84.3

IV.3.2 SEQ-9 inhibits the A2296-methylated ribosome in vitro

According to the in vivo efficacy analysis of different Sequanamycin A derivatives by Sanofi, SEQ-9 shows promising activity, which could serve as an advanced lead compound. To quantitatively test its activity against translation in *Mtb*, including the methylated ribosome, we employed in vitro translation assay to evaluate SEQ-9 (Fig. 43). The IC_{50} value of SEQ-9 toward unmethylated *Mtb* ribosome is around 170 nM, demonstrating potent inhibition. To our surprise, the IC_{50} value of SEQ-9 toward methylated *Mtb* ribosome is almost unaltered, suggesting its potency against A-2296 methylated *Mtb* ribosome.

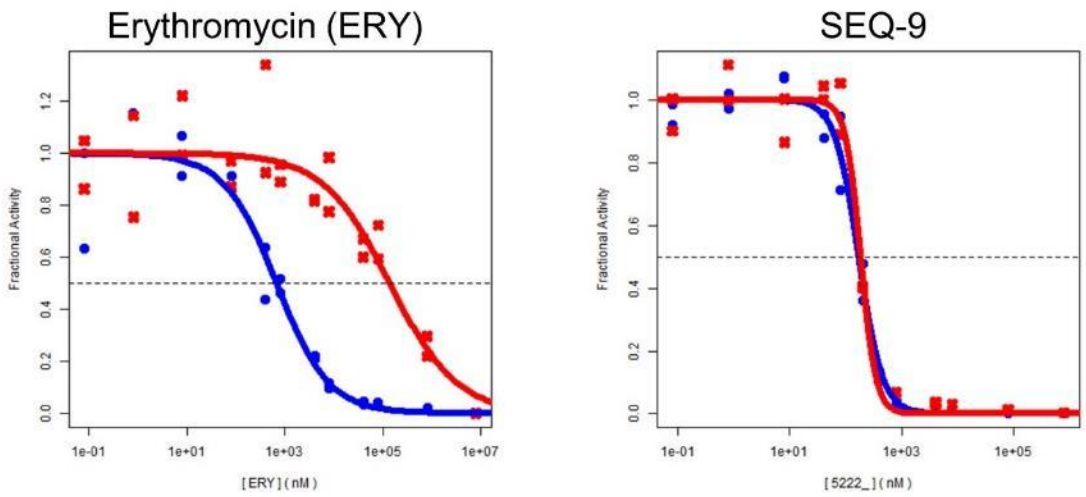


Figure 48 SEQ-9 can overcome methylation mediated resistance
In vitro translation assay with wild type ribosome (blue) and methylated ribosome (red)

IV.3.3 Cryo-EM structure of SEQ-9 in complex with A2297-methylated *Mtb* ribosome

To better understand the mechanism by which sequanamycins overcome the resistance due to the methylation of A2296 on the *Mtb* ribosome, we performed cryo-EM studies on the structure of the A2296-methylated *Mtb* ribosome bound with SEQ-9. A total number of 6,481 micrographs were collected, yielding final cryo-EM maps of variant states of the ribosome, including the 50S, vacant 70S, 70S with P site tRNA, and 70S with P/E site tRNA to the resolutions of 2.9 Å, 2.9 Å, 2.6 Å, and 2.8 Å, respectively (Fig. 44).

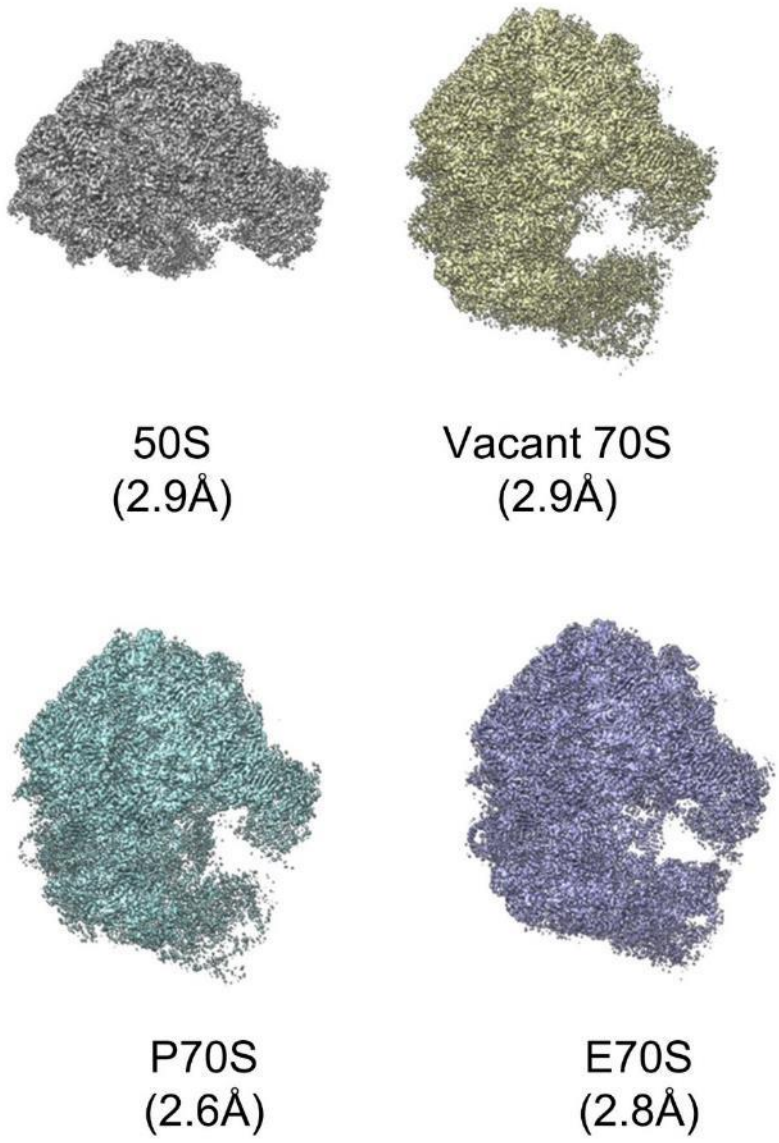
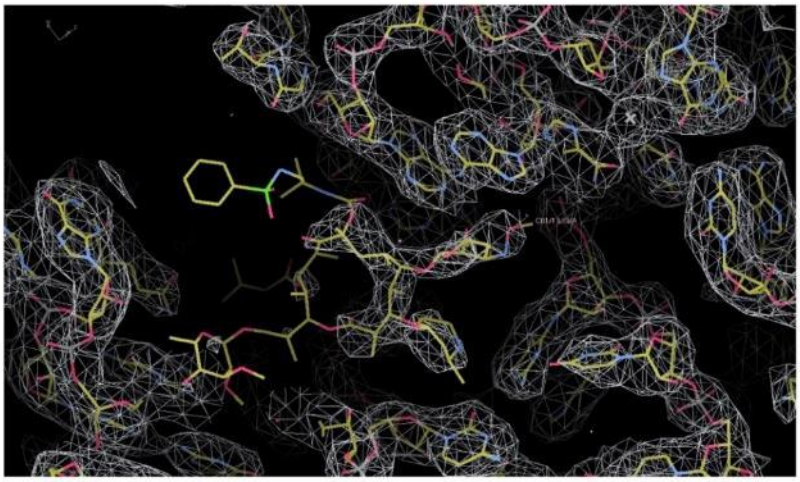


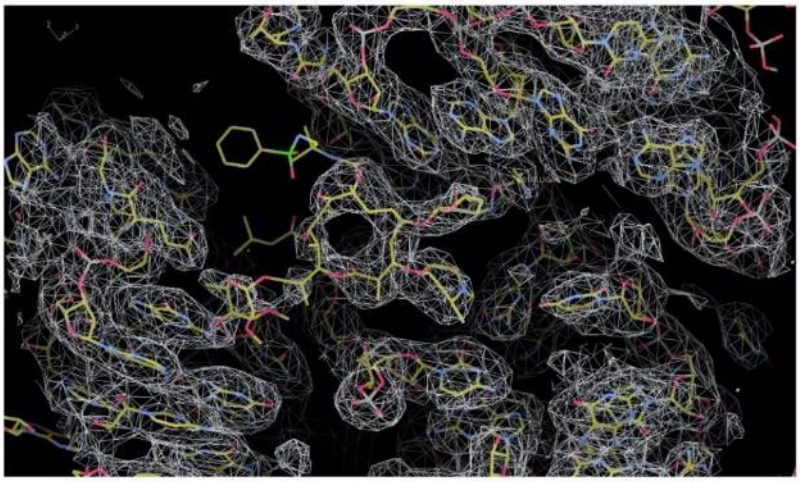
Figure 49 Cryo-EM maps of different subpopulations in the dataset of methylated *Mtb* ribosome bound with SEQ-9

By examining the peptidyl transferase center (PTC), only 70S with P site tRNA state shows apparent density for SEQ-9 (Fig. 45). However, the C13 and C8 side chains' weak cryo-EM densities suggest these side chains may be quite flexible and interact with peptide tunnel walls through nonspecific interaction. It has also been observed on

ketolides, which possess 11, 12-linked carbamate and an extended alkyl-aryl side chain. The alkyl-aryl side-chain stretches down the exit tunnel away from the peptidyl transferase center, protecting A752 in the *E. coli* and *S. aureus* ribosomes, but not on the *D. radiodurans* or *H. halobium* ribosomes. The orientations of ketolides' pharmacophoric side-chain vary when the drug binds to ribosomes from different species (193).



0.022 e^-/A^3



0.017 e^-/A^3

Figure 50 Density of SEQ-9 at different contour levels

The density of an extra methyl group of A2296 was visible in the 70S with P site tRNA state, while no extra density was found for methylated A2295 and A2297 (Fig. 46). Although unambiguously identifying RNA modifications on ribosome from cryo-EM map may need atomic resolution (< 2 Å), some conserved modifications on *Mtb* ribosomal RNA are already visible based on our 2.6-Å resolution map (Fig. 47).

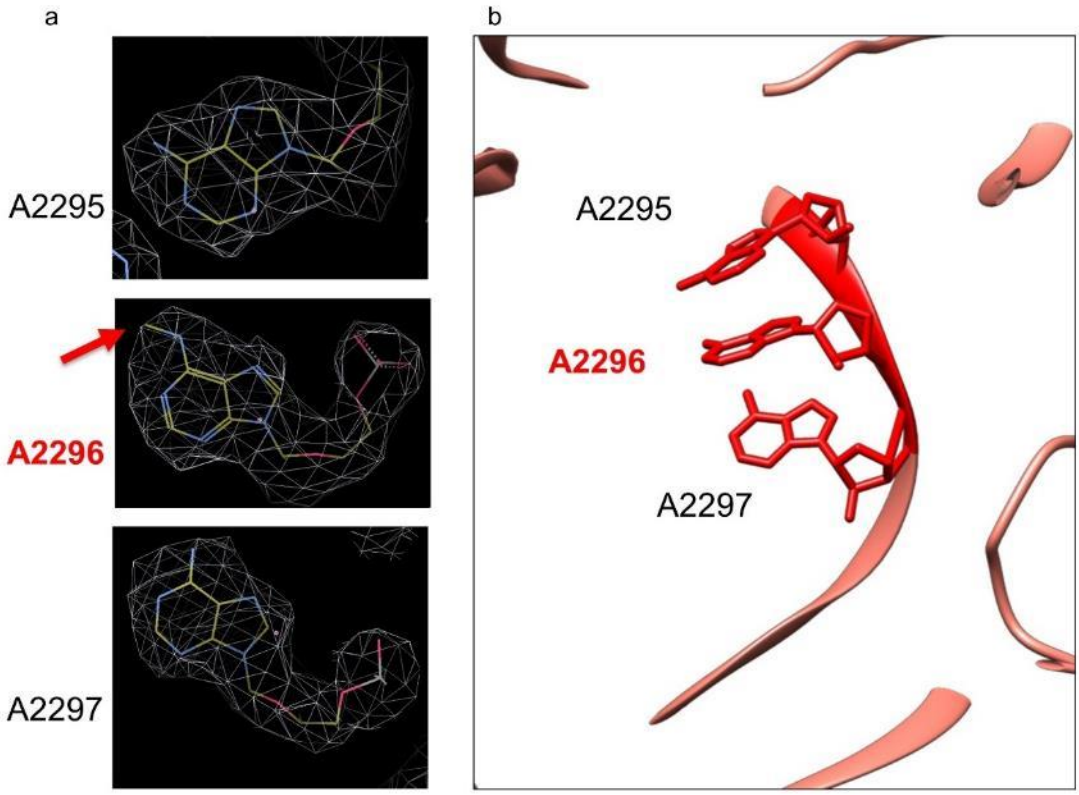


Figure 51 Cryo-EM density of 3 potentially methylated RNAs

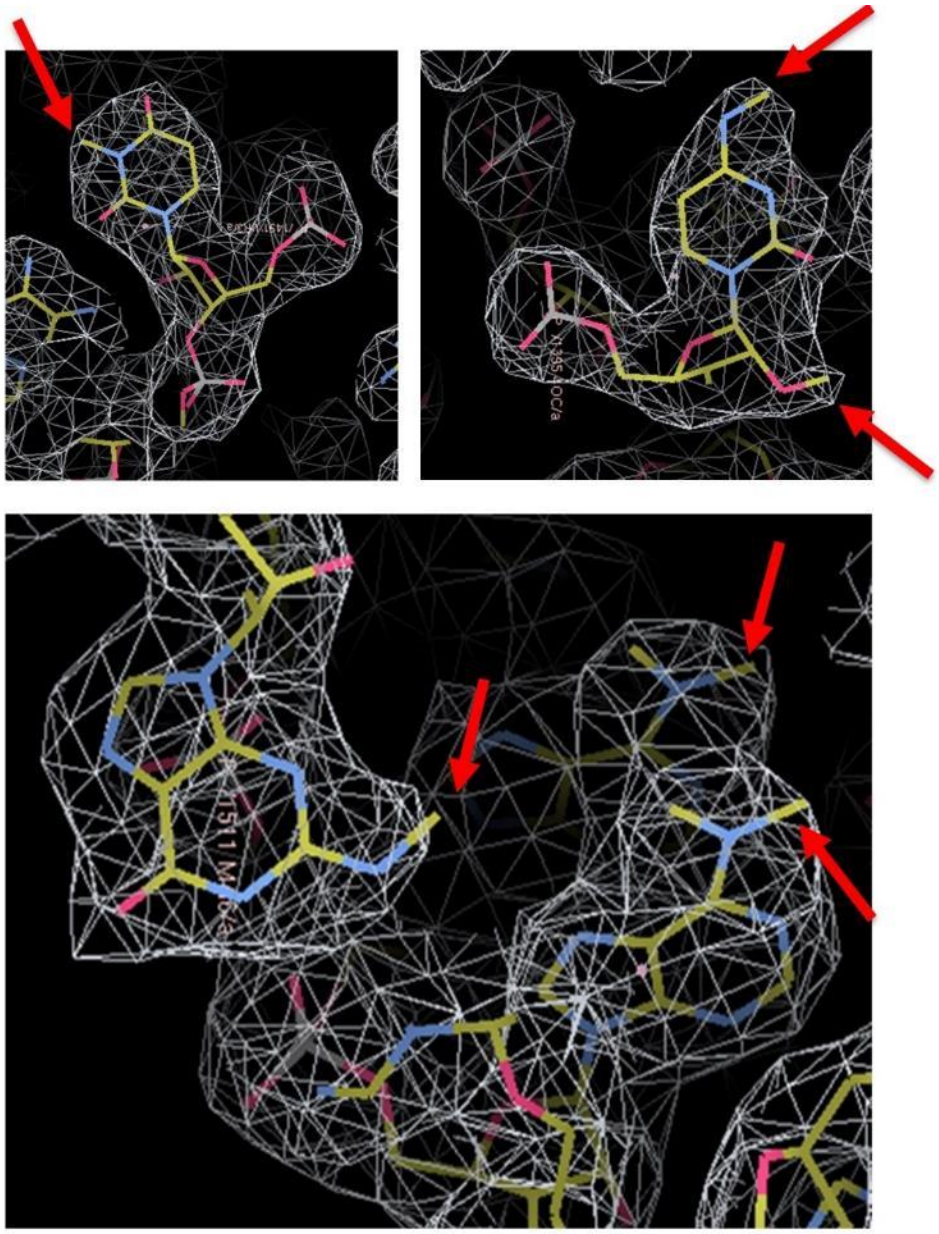


Figure 52 Cryo-EM densities of RNA modifications in *Mtb* ribosome

The previous structure suggests that macrolide resistance is caused by steric clashes between the methyl group of N6 on A2058 and the dimethylamine group of the

desosamine sugar, conserved for the three generations of macrolides (Fig. 48) (194). We superimposed the erythromycin structures (PDB 4V7X) with the SEQ-9 bound A2296-methylated *Mtb* ribosome structure to examine the interaction toward mono-methylated A2296 *Mtb* ribosome (Fig. 49). Despite the observed hydrogen bond between the 3'-hydroxyl group and the N1 of A2296, the distances between the methyl group on N6 of A2296 and either the hydroxyl group or the methyl group on erythromycin are 2.9 Å and 2.4 Å, respectively. These hypothetical close distances indicate potential steric clashes, contributing to macrolides' resistance in the methylated *Mtb* ribosome.

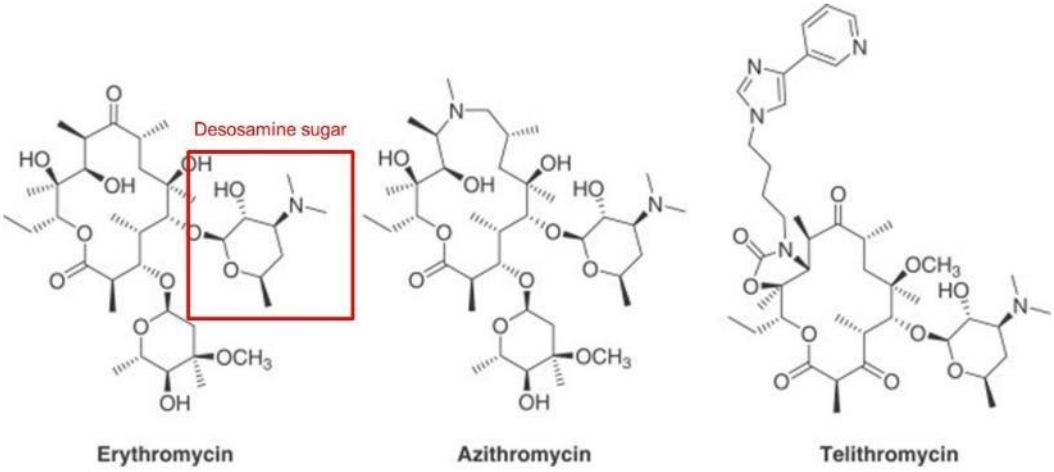


Figure 53 Structures of the three generations of macrolides
Red box highlights the desosamine sugar

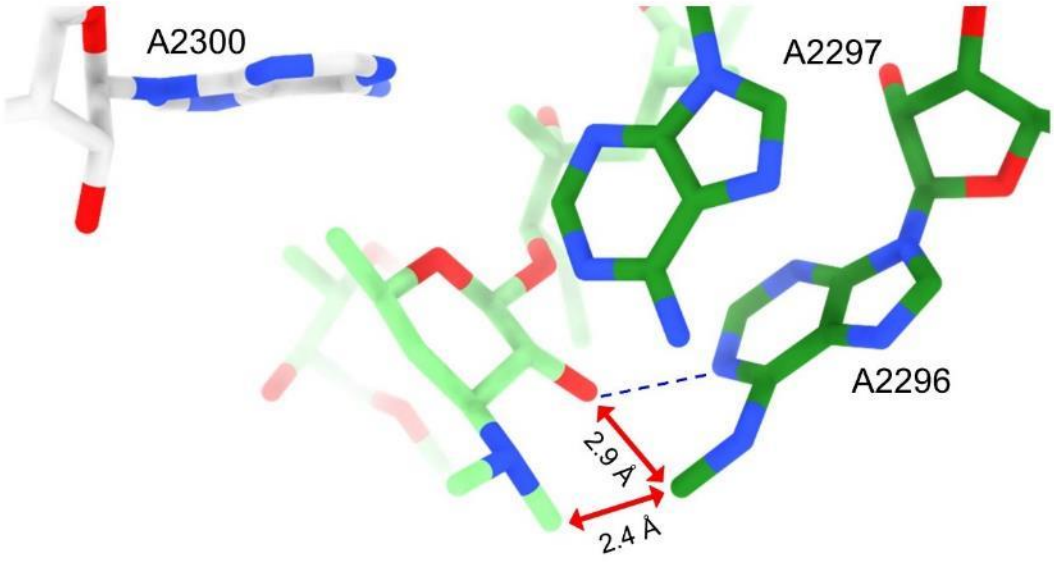


Figure 54 Macrolide resistance due to the methylation of A2296 in *Mtb* ribosome

Cyan model is the desosamine sugar of erythryomucin

For the methylated *Mtb* ribosome, SEQ still showed potent inhibition of protein synthesis, suggesting that the methyl group on N6 of A2296 does not interfere with the interaction between the *Mtb* ribosome and sequanamycins. Comparison between the crystal structure of the erythromycin-bound ribosome and the cryo-EM structure of SEQ-9 bound *Mtb* methylated ribosome indicates macro-lactone and ketoallose groups are located in the same pockets. However, the 3'-hydroxyl and 4'-Methyloxime of ketoallose in SEQ-9 protrude differently in methylated *Mtb* structure compared to the crystal structure (Fig. 51). Besides, A2300 in the SEQ9-bound methylated ribosome structure flips 90 degrees away from SEQ-9. SEQ-9 has a modified desosamine sugar, in which the ketoallose contains one sp^2 methyloxime group, compared to erythromycin with sp^3 dimethylamine groups (Fig. 50). The distances between the methyl group on N6

of A2296 and either the modified ketoallose group or the hydroxyl group on SEQ-9 are 4.1 Å and 3.4 Å, respectively. While it is hard to predict the interactions between ketoallose and nearby nucleic acids, the conformational change of SEQ-9 reduces the clash between ketoallose and the N6 methyl group on A2296, therefore maintaining the anti-tubercular potency.

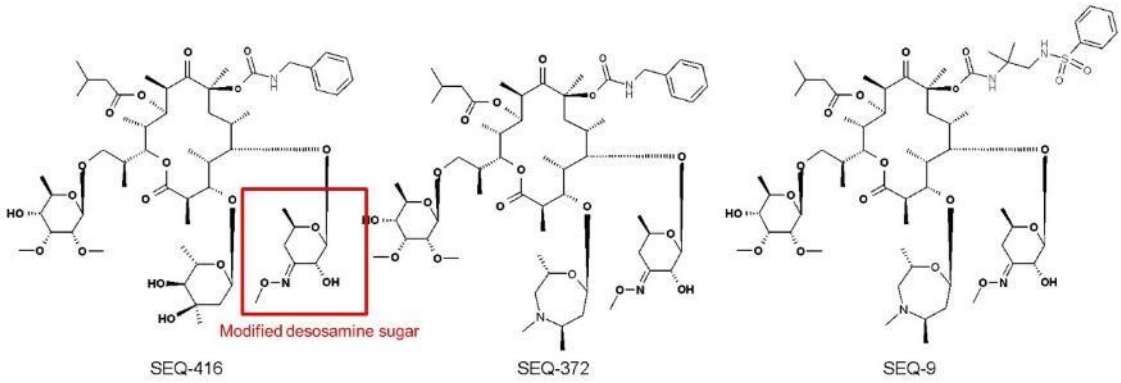


Figure 55 Structures of the different sequanamycin lead compounds Red box highlights the modified desosamine sugar

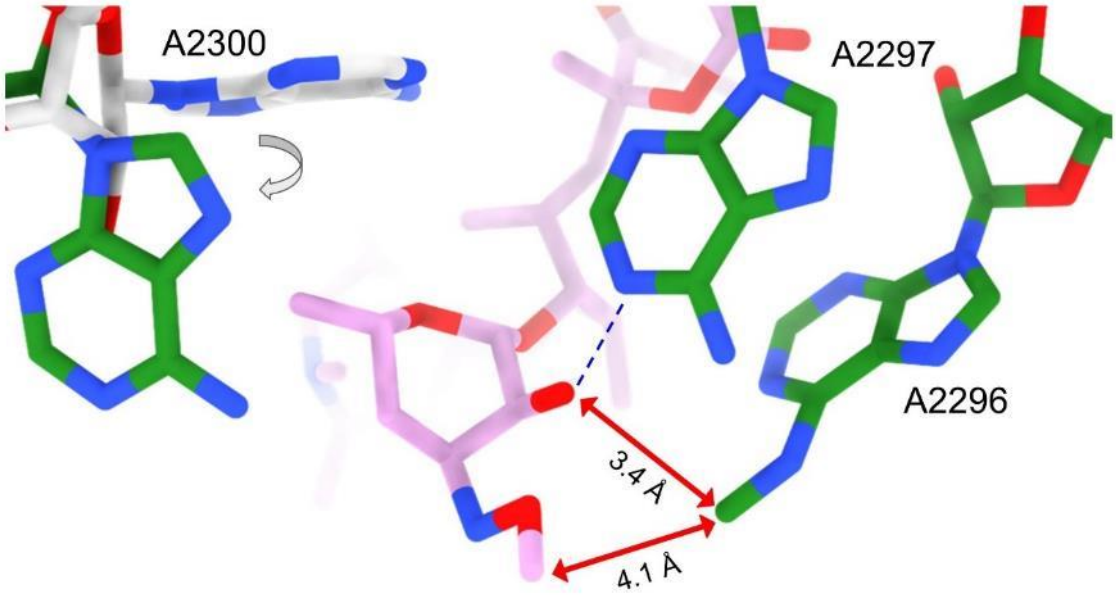


Figure 56 Modified desosamine sugar of SEQ-9 avoids steric clash with the methyl group of A2296
Pink model is the desosamine sugar of SEQ-9

IV.3.4 Structural variability around PTC in different states of *Mtb* ribosome

When we were building the model of the PTC region of A2296-methylated *Mtb* ribosome, we noticed that the densities of the RNA loop around A2296 had different resolvability in different states (Fig. 52). The 70S with P site tRNA has the most defined density that shows the loop's canonical conformation. However, vacant 70S, 70S with P/E site tRNA, and 50S have week RNA loop densities that shift away from the canonical conformation. Although the ribosome's PTC is inherently flexible, coordinating with numerous translation steps, such variations in previously reported Mtb ribosome structures were not observed. Since the data set was collected with methylated *Mtb* ribosome and SEQ-9 compound, flexible PTC in different ribosome states may be due to methylation of A2296, binding of SEQ-9, or the combination of both. However,

we cannot rule out the possibility that the conformation of PTC in ribosomes varies at different stages of cell growth (195) since the methylated ribosome was purified when *Mtb* cells were cultured in the late log phase. Nevertheless, understanding the dynamic of PTC could provide valuable insights for design and optimizing the new generations of antibiotics that target the region around PTC.

IV.3.5 The cryo-EM map of the 50S contains an RsfS density

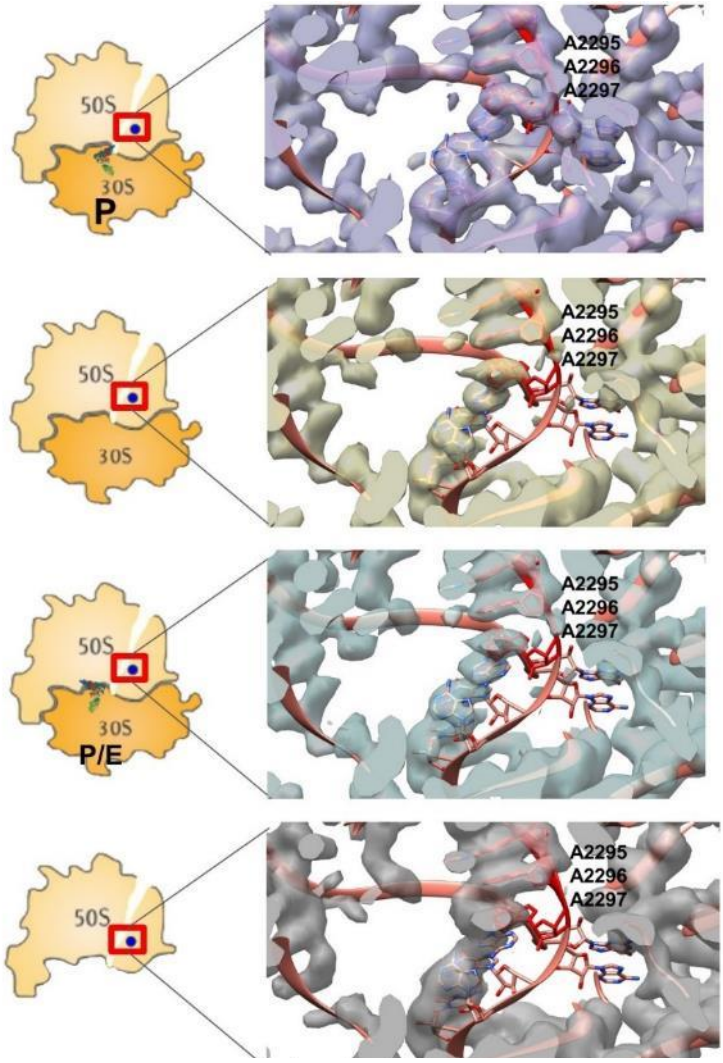


Figure 57 Variant PTC conformations in different ribosome subpopulations revealed by cryo-EM

Ribosomal silencing factor during starvation or stationary phase (RsfS) was initially proposed as the regulatory factor for the 50S under nutrient depletion stress (196). However, the expression of RsfS was also observed during the log phase of cell growth (196). More evidence suggests that RsfS may also function as a 50S biogenesis factor, as a homolog in mitochondria was identified and co-purified with mitoribosome assembly intermediates (197) (198).

Surprisingly, an extra density was found in the cryo-EM map of the 50S, which was later identified as RsfS due to the perfect fit for the atomic model (Fig. 53). Previously, Li et al. solved the cryo-EM structure of *Mtb* 50S in complex with RsfS to 9-Å resolution (141). The complex was reconstituted in vitro using 50S that was purified from the disassociated 70S. However, the 50S subpopulation in our data set was the contamination of the 70S fraction, which represented the 50S population in situ. In the case of *Mtb*, RsfS was found to be bound with mature 50S, indicating its role in maintaining the population distribution among 70S, 50S, and 30S. Our high-resolution cryo-EM data will provide a detailed description of the interactions between RsfS and *Mtb* 50S.

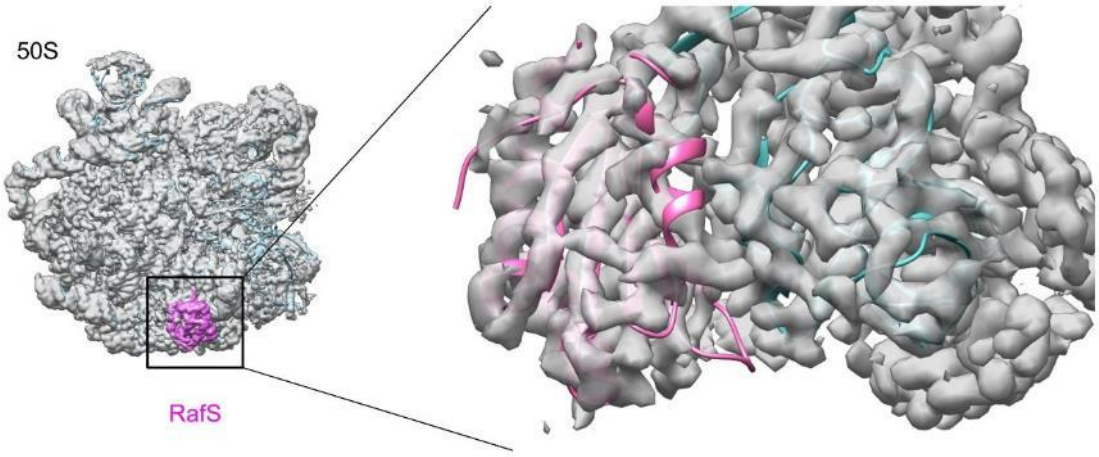


Figure 58 RafS density is observed on the 50S

IV.4 Discussion

Rediscovering macrolide-like antibiotics against TB is an unchartered area, primarily due to the intrinsic resistance to macrolides in *Mtb*. However, here we provided a functional and structural characterization of Sanofi advanced lead compound SEQ-9. Macrolide resistance contributed by the methylation of A2296 is overcome by SEQ-9, verified in vivo and in vitro. The cryo-EM structure shed light on the interactions between A2296-mediated ribosome and SEQ-9, supplying the blueprint for possible further SEQ-9 optimization.

Additionally, structural information about methylated *Mtb* ribosomes purified from a medium supplied with erythromycin was shown for the first time. Although the exact percentage of the methylated ribosome still needs to be confirmed, structures of different ribosome conformations were solved by cryo-EM. Interestingly enough, the RNA loop at PTC around A2296 remained canonical conformation in the 70S with P site

tRNA while showed alternative conformation in the states including 50S, vacant 70S, and 70S with P/E site tRNA. The cause of the observed differences needs further investigation.

Another interesting observation is that the solved 50S structure shows the density of RsfS, a factor that is supposed to be induced under nutrient depletion stress. Although the methylated ribosomes were purified from cells grown in the late log phase, the 50S bound with RsfS was also observed when cells were in the log phase. The physiological role of RsfS remains ambiguous, as its expression is throughout the entire growth curve.

CHAPTER V

CONCLUSIONS AND FUTURE DIRECTIONS

V.1 Potential regulations and additional functions of MtbEttA

We combined cryo-EM and X-ray crystallography to study the interactions between *Mtb* ribosome and MtbEttA during its ATP hydrolysis. We solved the ribosome-MtbEttA complex structures in pre-hydrolysis and transition states and MtbEttA alone in the post-hydrolysis state. With detailed analyses of our structures, we proposed thorough descriptions of the interplay between MtbEttA and ribosome. For instance, 1. MtbEttA destabilizes the PTC of 50S in the pre-hydrolysis state; 2. The rotation of 30S is responsible for promoting the ATP hydrolysis function of MtbEttA. However, there are o aspects of MtbEttA that need further investigation.

V.1.1 The sequence of events of MtbEttA's action on the ribosome

Early research indicates EttA functions on the initiation process of the translation and depends on the ATP/ADP ratio in the cell. Our structures confirm the binding site of EttA-like proteins is at the E site on the 70S, which will conflict with E site tRNA. However, the detailed investigation about the sequence of events during the MtbEttA's function is absent. The first question involves whether MtbEttA binds to the ribosome before or after elongation factor thermo unstable (EF-Tu) brings tRNA to the A site, or even when EF-Tu is still present on the ribosome. Another question regards the Elongation factor G (EF-G) whether it competes or facilitates the binding of MtbEttA to

the ribosome. Although the minimum translation assay was used to investigate the time of action of EttA on the ribosome, it involved using the ATP-hydrolysis deficient mutant. Since there are different conformations of ribosome when bound with MtbEttA at a different stage of ATP-hydrolysis, the ATP-hydrolysis deficient mutant may function differently as the wild type in terms of the sequence of interactions. The most robust way to study whether MtbEttA co-exists with EF-Tu or EF-G is two-color single-molecule experiments, potentially revealing the sequence of interactions.

V.1.2 Dimer-monomer transition of MtbEttA

Several ribosomal factors we have studied undergo dimer-monomer transition, including ribosome silencing factor (Rsf) and MtbEttA. The functional form of EttA-like proteins is the monomer state, while the dimer configuration of EttA-like protein is observed in both *E. coli* EttA and MtbEttA. It is suspected that the dimer form can not interact with the ribosome according to the available structures. Although *E. coli* EttA can shift the equilibrium toward monomer configuration at low concentration, the dimer-monomer transition of MtbEttA requires high salt conditions even at low concentration. In analogy to GTPase, where the GTPase-activating proteins (GAPs) are often found to couple and accelerate the GTPase's activity, it would be intriguing to investigate whether such activating proteins toward MtbEttA are present. Although ribosomes play a role in increasing the ATPase activity of EttA-like proteins, certain conditions or

factors may be first required to break the stable dimer of MtbEttA into functional monomer form in vivo.

Another unique aspect of MtbEttA is its essentiality, which is different from *E. coli* EttA. One possible explanation is that MtbEttA might be essential in maintaining translation homeostasis in *Mtb* since EttA-like proteins are responsible for preserving bacteria's fitness under stress. However, an out-of-box hypothesis is that the dimer form of MtbEttA has additional essential functions in the cell.

V.1.3 The potential role of MtbEttA in leaderless mRNA translation

One prominent feature of translation in mycobacteria is the high efficiency of translating leaderless mRNAs. It is estimated that a quarter of the transcripts are leaderless in mycobacteria (47). Due to the essentiality of MtbEttA, it would be exciting to investigate the role of MtbEttA in leaderless mRNA translation. Although we did not observe that MtbEttA increased the translational efficiency of leaderless mRNA using in vitro translation assay with the preliminary data, more thorough experiments are needed, including different types of mRNAs to verify its function.

V.2 Structural studies of Mpy and its regulation in vivo

The unique hibernation factor Mpy in mycobacteria can inhibit translation without forming the 70S dimer (100S), which is often observed in most bacteria containing homologs of hibernation factor. Based on recent studies, Mpy can only bind

to ribosomes when zinc is depleted inside the cells. However, the detailed mechanism of how zinc molecules prohibit the interaction of Mpy and ribosome is unknown. We discover that Mpy can bind to zinc molecules. Moreover, we pinpointed out that the C-terminal domain is responsible for the zinc-binding property. Interestingly, polymerization of purified C-terminal domain of Mpy occurs upon supplying with zinc molecules. This observation gave rise to a hypothesis that Mpy is self-inhibited as polymers when there is an excess amount of zinc molecules inside cells, which will prevent its interaction with the ribosome. However, the polymerization of Mpy is reversible, resulting in a functional form when Zinc molecules are depleted, potentially with the help of Mrf. Furthermore, the functional form of Mpy is a tetramer instead of a dimer like its homologs, which potentially blocks the formation of 100S.

V.2.1 Structural studies of Mpy C-terminal domain

To verify our observation, including the tetramer formation and zinc-binding property of the Mpy C-terminal domain, we are obligated to solve the 3D structure. However, the attempts to reveal the structure of the Mpy C-terminal domain were not successful. Other efforts like co-crystallization with Mrf could provide a different result. If the structure is ultimately unavailable, other technics like size-exclusion chromatography with multi-angle light scattering (SEC-MALS) could be used to verify the tetramer formation, as well as an alanine scanning experiment to pinpoint the zinc-binding motif.

V.2.2 Verification of the zinc-induced polymerization of Mpy in vivo

Because all the experiments were performed in vitro using purified Mpy proteins, it is crucial to verify the observations in vivo. The straightforward experiment could be using fluorescent microscopy to examine genetically modified Mpy with a fluorescent protein tag under different zinc concentrations. Also, the co-localization of Mpy and Mrf could be detected with fluorescence resonance energy transfer (FRET) experiments. Due to the small size of the bacteria, super-resolution fluorescent microscopy might be needed.

V.3 Insights of drug design toward methylated *Mtb* ribosome

We employed cryo-EM to study how SEQ-9 overcome the macrolide resistance due to the A2296-methylation. Steric clashes between the macrolides' desosamine sugar and the methylated RNA are the proposed vital factors contributing to the resistance. The unique modification of the desosamine sugar of SEQ-9 avoids the steric clashes with methylated A2296 on *Mtb* ribosome. This result provides structure-based guidance for future drug designs that bypasses the intrinsic macrolide resistance in *Mtb*. The modified desosamine sugar and 14-carbon backbone should be kept while changing other groups on the molecule to increase efficacy or reduce toxicity.

A possible direction derived from this study is to investigate the methyltransferase Erm(37)’s mechanism-of-action. There is no structural information

available on how a methyltransferase functions near the PTC. Whether Erm(37) functions on the fully assembled ribosome or the intermediates also need further clarification.

REFERENCES

1. CDC. Global COVID-192020.
2. CDC. Forecasts of COVID-19 Deaths2020.
3. Huremović D. Brief History of Pandemics (Pandemics Throughout History). Psychiatry of Pandemics. 2019:7-35. doi: 10.1007/978-3-030-15346-5_2. PubMed PMID: PMC7123574.
4. Russell DG. Mycobacterium tuberculosis: Here today, and here tomorrow. Nat Rev Mol Cell Bio. 2001;2(8):569-77. PubMed PMID: WOS:000170274500013.
5. WHO. Global tuberculosis report 20202020.
6. Canetti D, Riccardi N, Martini M, Villa S, Di Biagio A, Codecasa L, Castagna A, Barberis I, Gazzaniga V, Besozzi G. HIV and tuberculosis: The paradox of dual illnesses and the challenges of their fighting in the history. Tuberculosis. 2020;122. PubMed PMID: WOS:000538119900001.
7. Sharma SK, Mohan A, Kadhiravan T. HIV-TB co-infection: Epidemiology, diagnosis & management. Indian J Med Res. 2005;121(4):550-67. PubMed PMID: WOS:000229008200028.
8. Barberis I, Bragazzi NL, Galluzzo L, Martini M. The history of tuberculosis: from the first historical records to the isolation of Koch's bacillus. J Prev Med Hyg. 2017;58(1):E9-E12. PubMed PMID: 28515626.
9. Koch R. The Etiology of Tuberculosis. Rev Infect Dis. 1882;4(6):1270-4. PubMed PMID: WOS:A1982QL39600013.
10. Cambau E, Drancourt M. Steps towards the discovery of Mycobacterium tuberculosis by Robert Koch, 1882. Clin Microbiol Infec. 2014;20(3):196-201. PubMed PMID: WOS:000331024400011.

11. Fletcher HA, Schrager L. TB vaccine development and the End TB Strategy: importance and current status. T Roy Soc Trop Med H. 2016;110(4):212-8. PubMed PMID: WOS:000374573000003.
12. Brodie D, Schluger NW. The diagnosis of tuberculosis. Clin Chest Med. 2005;26(2):247-+. PubMed PMID: WOS:000228823000008.
13. Moore DAJ, Evans CAW, Gilman RH, Caviedes L, Coronel J, Vivar A, Sanchez E, Pinedo Y, Saravia JC, Salazar C, Oberhelman R, Hollm-Delgado MG, LaChira D, Escombe AR, Friedland JS. Microscopic-observation drug-susceptibility assay for the diagnosis of TB. New Engl J Med. 2006;355(15):1539-50. PubMed PMID: WOS:000241160400005.
14. Richeldi L. An update on the diagnosis of tuberculosis infection. Am J Resp Crit Care. 2006;174(7):736-42. PubMed PMID: WOS:000240891900004.
15. van der Paardt AF, Wilffert B, Akkerman OW, de Lange WCM, van Soolingen D, Sinha B, van der Werf TS, Kosterink JGW, Alffenaar JWC. Evaluation of macrolides for possible use against multidrug-resistant Mycobacterium tuberculosis. Eur Respir J. 2015;46(2):444-55. PubMed PMID: WOS:000358795800021.
16. Trifiro S, Bourgault AM, Lebel F, Rene P. Ghost Mycobacteria on Gram Stain. J Clin Microbiol. 1990;28(1):146-7. PubMed PMID: WOS:A1990CE00800034.
17. Horsburgh CR, Barry CE, Lange C. Treatment of Tuberculosis. New Engl J Med. 2015;373(22):2149-60. PubMed PMID: WOS:000365354800009.
18. Lange C, Abubakar I, Alffenaar JWC, Bothamley G, Caminero JA, Carvalho ACC, Chang KC, Codecasa L, Correia A, Crudu V, Davies P, Dedicoat M, Drobniewski F, Duarte R, Ehlers C, Erkens C, Goletti D, Gunther G, Ibraim E, Kampmann B, Kuksa L, de Lange W, van Leth F, van Lunzen J, Matteelli A, Menzies D, Monedero I, Richter E, Rusch-Gerdes S, Sandgren A, Scardigli A, Skrahina A, Tortoli E, Volchenkov G, Wagner D, van der Werf MJ, Williams B, Yew WW, Zellweger JP, Cirillo DM. Management of patients with multidrug-resistant/extensively drug-resistant tuberculosis

in Europe: a TBNET consensus statement. Eur Respir J. 2014;44(1):23-63. PubMed PMID: WOS:000338906700008.

19. Ukwaja KN, Modebe O, Igwenyi C, Alobu I. The economic burden of tuberculosis care for patients and households in Africa: a systematic review. Int J Tuberc Lung D. 2012;16(6):733-9. PubMed PMID: WOS:000304580400006.

20. Blumberg HM, Ernst JD. The Challenge of Latent TB Infection. Jama-J Am Med Assoc. 2016;316(9):931-3. PubMed PMID: WOS:000382460200015.

21. Esmail H, Barry CE, Young DB, Wilkinson RJ. The ongoing challenge of latent tuberculosis. Philos T R Soc B. 2014;369(1645). PubMed PMID: WOS:000336094600011.

22. Scanga CA, Mohan VP, Joseph H, Yu KM, Chan J, Flynn JL. Reactivation of latent tuberculosis: Variations on the Cornell murine model. Infect Immun. 1999;67(9):4531-8. PubMed PMID: WOS:000082138600030.

23. Chinen J, Shearer WT. Molecular virology and immunology of HIV infection. J Allergy Clin Immun. 2002;110(2):189-98. PubMed PMID: WOS:000177509800001.

24. Finlay BB, McFadden G. Anti-immunology: Evasion of the host immune system by bacterial and viral pathogens. Cell. 2006;124(4):767-82. PubMed PMID: WOS:000237240900020.

25. de Martino M, Lodi L, Gaffi L, Chiappini E. Immune Response to Mycobacterium tuberculosis: A Narrative Review. Front Pediatr. 2019;7. PubMed PMID: WOS:000482826200001.

26. Ahmad S. Pathogenesis, Immunology, and Diagnosis of Latent Mycobacterium tuberculosis Infection. Clin Dev Immunol. 2011. PubMed PMID: WOS:000286250900001.

27. Lin PL, Flynn JL. Understanding Latent Tuberculosis: A Moving Target. J Immunol. 2010;185(1):15-22. PubMed PMID: WOS:000278933800003.

28. Dutta NK, Karakousis PC. Latent Tuberculosis Infection: Myths, Models, and Molecular Mechanisms. Microbiol Mol Biol R. 2014;78(3):343-71. PubMed PMID: WOS:000341639400001.

29. Sterling TR, Bethel J, Goldberg S, Weinfurter P, Yun L, Horsburgh CR, Studies TE. The scope and impact of treatment of latent tuberculosis infection in the United States and Canada. Am J Resp Crit Care. 2006;173(8):927-31. PubMed PMID: WOS:000236770500018.

30. Toma T. New guidelines for treatment of latent tuberculosis infection. B World Health Organ. 2000;78(5):710-1. PubMed PMID: WOS:000087564100019.

31. Ikram S, O'Brien K, Rahman A, Potter J, Burman M, Kunst H. Barriers and Facilitators to Delivering Latent Tuberculosis Infection (Ltbi) Screening and Treatment to Recent Migrants: A Survey of Providers in a High Prevalence Tb Setting in the Uk. Thorax. 2019;74:A199-A200. PubMed PMID: WOS:000572473500346.

32. Kane M, Waters R, McLaughlin A, McDonald C, Korn B, Walsh C, Saukkonen J, Keane J. Barriers to accepting treatment for latent tuberculosis infection (LTBI). Irish J Med Sci. 2010;179:S469-S. PubMed PMID: WOS:000284214500046.

33. Singla R, Sharma SK, Mohan A, Makharia G, Sreenivas V, Jha B, Kumar S, Sarda P, Singh S. Evaluation of risk factors for antituberculosis treatment induced hepatotoxicity. Indian J Med Res. 2010;132(1):81-6. PubMed PMID: WOS:000279803700015.

34. Singh J, Arora A, Garg PK, Thakur VS, Pande JN, Tandon RK. Antituberculosis Treatment-Induced Hepatotoxicity - Role of Predictive Factors. Postgrad Med J. 1995;71(836):359-62. PubMed PMID: WOS:A1995RB40700009.

35. Fu LM, Fu-Liu CS. Is Mycobacterium tuberculosis a closer relative to Gram-positive or Gram-negative bacterial pathogens? Tuberculosis. 2002;82(2-3):85-90. PubMed PMID: WOS:000178487500007.

36. Brennan PJ. Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. Tuberculosis. 2003;83(1-3):91-7. PubMed PMID: WOS:000183516300016.
37. Maitra A, Munshi T, Healy J, Martin LT, Vollmer W, Keep NH, Bhakta S. Cell wall peptidoglycan in Mycobacterium tuberculosis: An Achilles' heel for the TB-causing pathogen. Fems Microbiol Rev. 2019;43(5):548-75. PubMed PMID: WOS:000493064500005.
38. Barkan D, Liu Z, Sacchettini JC, Glickman MS. Mycolic Acid Cyclopropanation is Essential for Viability, Drug Resistance, and Cell Wall Integrity of Mycobacterium tuberculosis. Chem Biol. 2009;16(5):499-509. PubMed PMID: WOS:000266620800005.
39. Ghazaei C. Mycobacterium tuberculosis and lipids: Insights into molecular mechanisms from persistence to virulence. J Res Med Sci. 2018;23. PubMed PMID: WOS:000444510600005.
40. Marrakchi H, Laneelle MA, Daffe M. Mycolic Acids: Structures, Biosynthesis, and Beyond. Chem Biol. 2014;21(1):67-85. PubMed PMID: WOS:000330082200005.
41. Hong X, Hopfinger AJ. Molecular modeling and simulation of Mycobacterium tuberculosis cell wall permeability. Biomacromolecules. 2004;5(3):1066-77. PubMed PMID: WOS:000221315200059.
42. Lima VMF, Bonato VLD, Lima KM, Dos Santos SA, Dos Santos RR, Goncalves EDC, Faccioli LH, Brandao IT, Rodrigues-Junior JM, Silva CL. Role of trehalose dimycolate in recruitment of cells and modulation of production of cytokines and NO in tuberculosis. Infect Immun. 2001;69(9):5305-12. PubMed PMID: WOS:000170540000012.
43. Lang R. Recognition of the mycobacterial cord factor by Mincle: relevance for granuloma formation and resistance to tuberculosis. Front Immunol. 2013;4. PubMed PMID: WOS:000209374100005.

44. Singh P, Rameshwaram NR, Ghosh S, Mukhopadhyay S. Cell envelope lipids in the pathophysiology of Mycobacterium tuberculosis. Future Microbiol. 2018;13(6):689-710. PubMed PMID: WOS:000436536200009.

45. Meena LS, Rajni. Survival mechanisms of pathogenic Mycobacterium tuberculosis H(37)Rv. Febs J. 2010;277(11):2416-27. PubMed PMID: WOS:000277800600005.

46. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Conner R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence (vol 393, pg 537, 1998). Nature. 1998;396(6707):190-8. PubMed PMID: WOS:000077013300058.

47. Shell SS, Wang J, Lapierre P, Mir M, Chase MR, Pyle MM, Gawande R, Ahmad R, Sarracino DA, Ioerger TR, Fortune SM, Derbyshire KM, Wade JT, Gray TA. Leaderless Transcripts and Small Proteins Are Common Features of the Mycobacterial Translational Landscape. Plos Genet. 2015;11(11). PubMed PMID: WOS:000366179000023.

48. Duc KD, Batra SS, Bhattacharya N, Cate JHD, Song YS. Differences in the path to exit the ribosome across the three domains of life. Nucleic Acids Res. 2019;47(8):4198-210. PubMed PMID: WOS:000473754300033.

49. Yang KL, Chang JY, Cui ZC, Li XJ, Meng R, Duan LJ, Thongchol J, Jakana J, Huwe CM, Sacchettini JC, Zhang JJ. Structural insights into species-specific features of the ribosome from the human pathogen Mycobacterium tuberculosis. Nucleic Acids Res. 2017;45(18):10884-94. PubMed PMID: WOS:000413107400045.

50. Eyal Z, Matzov D, Krupkin M, Wekselman I, Paukner S, Zimmerman E, Rozenberg H, Bashan A, Yonath A. Structural insights into species-specific features of

the ribosome from the pathogen Staphylococcus aureus. P Natl Acad Sci USA. 2015;112(43):E5805-E14. PubMed PMID: WOS:000363458100008.

51. Sohmen D, Chiba S, Shimokawa-Chiba N, Innis CA, Berninghausen O, Beckmann R, Ito K, Wilson DN. Structure of the Bacillus subtilis 70S ribosome reveals the basis for species-specific stalling. Nat Commun. 2015;6. PubMed PMID: WOS:000353704500004.

52. James BW, Williams A, Marsh PD. The physiology and pathogenicity of Mycobacterium tuberculosis grown under controlled conditions in a defined medium. J Appl Microbiol. 2000;88(4):669-77. PubMed PMID: WOS:000086276600014.

53. Stephan J, Bender J, Wolschendorf F, Hoffmann C, Roth E, Mailander C, Engelhardt H, Niederweis M. The growth rate of Mycobacterium smegmatis depends on sufficient porin-mediated influx of nutrients. Mol Microbiol. 2005;58(3):714-30. PubMed PMID: WOS:000232537100009.

54. Bruell CM, Eichholz C, Kubarenko A, Post V, Katunin VI, Hobbie SN, Rodnina MV, Boettger EC. Conservation of bacterial protein synthesis machinery: Initiation and elongation in Mycobacterium smegmatis. Biochemistry-Us. 2008;47(34):8828-39. PubMed PMID: WOS:000258579700003.

55. Bremer H, Dennis PP. Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates. EcoSal Plus. 2008;3(1). doi: 10.1128/ecosal.5.2.3. PubMed PMID: 26443740.

56. Yamada H, Yamaguchi M, Chikamatsu K, Aono A, Lina Y, Igarashi Y, Sakashita K, Takaki A, Mitarai S. Structome analysis based on direct enumeration of virulent Mycobacterium tuberculosis with TEM examination of serial ultrathin sections. Febs J. 2015;282:351-. PubMed PMID: WOS:000362570606133.

57. Murinat V, Kasari M, Takada H, Hinnu M, Saha CK, Grimshaw JW, Seki T, Reith M, Putrins M, Tenson T, Strahl H, Hauryliuk V, Atkinson GC. ABCF ATPases Involved in Protein Synthesis, Ribosome Assembly and Antibiotic Resistance: Structural

and Functional Diversification across the Tree of Life. J Mol Biol. 2019;431(18):3568-90. PubMed PMID: WOS:000484872800012.

58. Rodnina MV. Translation in Prokaryotes. Csh Perspect Biol. 2018;10(9). PubMed PMID: WOS:000443588100009.

59. Kaminishi T, Wilson DN, Takemoto C, Harms JM, Kawazoe M, Schluenzen F, Hanawa-Suetsugu K, Shirouzu M, Fucini P, Yokoyama S. A snapshot of the 30S ribosomal subunit capturing mRNA via the Shine-Dalgarno interaction. Structure. 2007;15(3):289-97. PubMed PMID: WOS:000245617400006.

60. Moll I, Hirokawa G, Kiel MC, Kaji A, Blasi U. Translation initiation with 70S ribosomes: an alternative pathway for leaderless mRNAs. Nucleic Acids Res. 2004;32(11):3354-63. PubMed PMID: WOS:000222399200016.

61. Brock JE, Pourshahian S, Giliberti J, Limbach PA, Janssen GR. Ribosomes bind leaderless mRNA in Escherichia coli through recognition of their 5 '-terminal AUG. Rna. 2008;14(10):2159-69. PubMed PMID: WOS:000259568700018.

62. Wayne LG, Hayes LG. An in vitro model for sequential study of shiftdown of Mycobacterium tuberculosis through two stages of nonreplicating persistence. Infect Immun. 1996;64(6):2062-9. PubMed PMID: WOS:A1996UN18900026.

63. Wayne LG, Sohaskey CD. Nonreplicating persistence of Mycobacterium tuberculosis. Annu Rev Microbiol. 2001;55:139-63. PubMed PMID: WOS:000171732600007.

64. Rustad TR, Harrell MI, Liao RL, Sherman DR. The Enduring Hypoxic Response of Mycobacterium tuberculosis. Plos One. 2008;3(1). PubMed PMID: WOS:000260504200015.

65. Leistikow RL, Morton RA, Bartek IL, Frimpong I, Wagner K, Voskuil MI. The Mycobacterium tuberculosis DosR Regulon Assists in Metabolic Homeostasis and Enables Rapid Recovery from Nonrespiring Dormancy. J Bacteriol. 2010;192(6):1662-70. PubMed PMID: WOS:000274891300021.

66. Trauner A, Lougheed KEA, Bennett MH, Hingley-Wilson SM, Williams HD. The Dormancy Regulator DosR Controls Ribosome Stability in Hypoxic Mycobacteria. J Biol Chem. 2012;287(28):24053-63. PubMed PMID: WOS:000306511300073.
67. Yoshida H, Wada A. The 100S ribosome: ribosomal hibernation induced by stress. Wires Rna. 2014;5(5):723-32. PubMed PMID: WOS:000344452700010.
68. Li YL, Sharma MR, Koripella RK, Yang Y, Kaushal PS, Lin QS, Wade JT, Gray TA, Derbyshire KM, Agrawal RK, Ojha AK. Zinc depletion induces ribosome hibernation in mycobacteria. P Natl Acad Sci USA. 2018;115(32):8191-6. PubMed PMID: WOS:000440982000052.
69. Li YL, Corro JH, Palmer CD, Ojha AK. Progression from remodeling to hibernation of ribosomes in zinc-starved mycobacteria. P Natl Acad Sci USA. 2020;117(32):19528-37. PubMed PMID: WOS:000561788700003.
70. Harms A, Brodersen DE, Mitarai N, Gerdes K. Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology. Mol Cell. 2018;70(5):768-84. PubMed PMID: WOS:000434791400006.
71. Schifano JM, Edifor R, Sharp JD, Ouyang M, Konkimalla A, Husson RN, Woychik NA. Mycobacterial toxin MazF-mt6 inhibits translation through cleavage of 23S rRNA at the ribosomal A site. P Natl Acad Sci USA. 2013;110(21):8501-6. PubMed PMID: WOS:000320328700049.
72. Betts JC, Lukey PT, Robb LC, McAdam RA, Duncan K. Evaluation of a nutrient starvation model of Mycobacterium tuberculosis persistence by gene and protein expression profiling. Mol Microbiol. 2002;43(3):717-31. PubMed PMID: WOS:000174058300015.
73. Kurz SG, Furin JJ, Bark CM. Drug-Resistant Tuberculosis Challenges and Progress. Infect Dis Clin N Am. 2016;30(2):509-+. PubMed PMID: WOS:000378466900011.

74. Seung KJ, Keshavjee S, Rich ML. Multidrug-Resistant Tuberculosis and Extensively Drug-Resistant Tuberculosis. Csh Perspect Med. 2015;5(9). PubMed PMID: WOS:000364532900003.
75. Velayati AA, Farnia P, Masjedi MR, Ibrahim TA, Tabarsi P, Haroun RZ, Kuan HO, Ghanavi J, Farnia P, Varahram M. Totally drug-resistant tuberculosis strains: evidence of adaptation at the cellular level. Eur Respir J. 2009;34(5):1202-3. PubMed PMID: WOS:000271571700030.
76. Harding E. WHO global progress report on tuberculosis elimination (vol 8, pg 19, 2020). Lancet Resp Med. 2020;8(1):E3-E. PubMed PMID: WOS:000503397100003.
77. Da Silva PEA, Palomino JC. Molecular basis and mechanisms of drug resistance in Mycobacterium tuberculosis: classical and new drugs. J Antimicrob Chemoth. 2011;66(7):1417-30. PubMed PMID: WOS:000291536100001.
78. Yonath A. Antibiotics targeting ribosomes: Resistance, selectivity, synergism, and cellular regulation. Annu Rev Biochem. 2005;74:649-79. PubMed PMID: WOS:000231235100022.
79. Andini N, Nash KA. Intrinsic macrolide resistance of the Mycobacterium tuberculosis complex is inducible. Antimicrob Agents Ch. 2006;50(7):2560-2. PubMed PMID: WOS:000238721200045.
80. Bechhofer DH, Zen KH. Mechanism of Erythromycin-Induced Ermc Messenger-Rna Stability in Bacillus-Subtilis. J Bacteriol. 1989;171(11):5803-11. PubMed PMID: WOS:A1989AX60400006.
81. Shur KV, Maslov DA, Mikheecheva NE, Akimova NI, Bekker OB, Danilenko VN. The Intrinsic Antibiotic Resistance to beta-Lactams, Macrolides, and Fluoroquinolones of Mycobacteria Is Mediated by the whiB7 and tap Genes. Russ J Genet+. 2017;53(9):1006-15. PubMed PMID: WOS:000411516900007.
82. Champney WS, Chittum HS, Tober CL. A 50S ribosomal subunit precursor particle is a substrate for the ErmC methyltransferase in Staphylococcus aureus cells. Curr Microbiol. 2003;46(6):453-60. PubMed PMID: WOS:000182687800012.

83. Madsen CT, Jakobsen L, Buriankova K, Doucet-Populaire F, Pernodet JL, Douthwaite S. Methyltransferase Erm(37) slips on rRNA to confer atypical resistance in Mycobacterium tuberculosis. J Biol Chem. 2005;280(47):38942-7. PubMed PMID: WOS:000233362200007.

84. Murina V, Kasari M, Hauryliuk V, Atkinson GC. Antibiotic resistance ABCF proteins reset the peptidyl transferase centre of the ribosome to counter translational arrest. Nucleic Acids Res. 2018;46(7):3753-63. PubMed PMID: WOS:000431137900041.

85. Crowe-McAuliffe C, Graf M, Huter P, Takada H, Abdelshahid M, Novacek J, Murina V, Atkinson GC, Hauryliuk V, Wilson DN. Structural basis for antibiotic resistance mediated by the Bacillus subtilis ABCF ATPase VmlR. P Natl Acad Sci USA. 2018;115(36):8978-83. PubMed PMID: WOS:000443555000057.

86. Su WX, Kumar V, Ding YC, Ero R, Serra A, Lee BST, Wong ASW, Shi J, Sze SK, Yang L, Gao YG. Ribosome protection by antibiotic resistance ATP-binding cassette protein. P Natl Acad Sci USA. 2018;115(20):5157-62. PubMed PMID: WOS:000432120400049.

87. Sharkey LKR, O'Neill AJ. Antibiotic Resistance ABC-F Proteins: Bringing Target Protection into the Limelight. Acs Infect Dis. 2018;4(3):239-46. PubMed PMID: WOS:000427443200004.

88. Wilson DN. The ABC of Ribosome-Related Antibiotic Resistance. Mbio. 2016;7(3). PubMed PMID: WOS:000383440300037.

89. Chen B, Boel G, Hashem Y, Ning W, Fei JY, Wang C, Gonzalez RL, Hunt JF, Frank J. EttA regulates translation by binding the ribosomal E site and restricting ribosome-tRNA dynamics. Nat Struct Mol Biol. 2014;21(2):152-+. PubMed PMID: WOS:000331093600007.

90. Boel G, Smith PC, Ning W, Englander MT, Chen B, Hashem Y, Testa AJ, Fischer JJ, Wieden HJ, Frank J, Gonzalez RL, Hunt JF. The ABC-F protein EttA gates

ribosome entry into the translation elongation cycle. Nat Struct Mol Biol. 2014;21(2):143-+. PubMed PMID: WOS:000331093600006.

91. Hey T. Richard Feynman and computation. Contemp Phys. 1999;40(4):257-65. PubMed PMID: WOS:000081594300003.

92. Wollman AJM, Nudd R, Hedlund EG, Leake MC. From Animaculum to single molecules: 300 years of the light microscope. Open Biol. 2015;5(4). PubMed PMID: WOS:000355309700007.

93. Huxley HE. Citation Classic - Electron-Microscope Studies on the Structure of Natural and Synthetic Protein Filaments from Striated-Muscle. Cc/Life Sci. 1979(29):10-. PubMed PMID: WOS:A1979HZ36800001.

94. Scarff CA, Fuller MJG, Thompson RF, Iadaza MG. Variations on Negative Stain Electron Microscopy Methods: Tools for Tackling Challenging Systems. Jove-J Vis Exp. 2018(132). PubMed PMID: WOS:000426453400089.

95. Taylor KA, Glaeser RM. Electron-Diffraction of Frozen, Hydrated Protein Crystals. Science. 1974;186(4168):1036-7. PubMed PMID: WOS:A1974U879900018.

96. Haas DJ. The early history of cryo-cooling for macromolecular crystallography. Iucrj. 2020;7:148-57. PubMed PMID: WOS:000518799300004.

97. Adrian M, Dubochet J, Lepault J, Mcdowall AW. Cryo-Electron Microscopy of Viruses. Nature. 1984;308(5954):32-6. PubMed PMID: WOS:A1984SF11800038.

98. Jain T, Sheehan P, Crum J, Carragher B, Potter CS. Spotiton: A prototype for an integrated inkjet dispense and vitrification system for cryo-TEM. J Struct Biol. 2012;179(1):68-75. PubMed PMID: WOS:000305775400008.

99. Rubinstein JL, Guo H, Ripstein ZA, Haydaroglu A, Au A, Yip CM, Di Trani JM, Benlekbir S, Kwok T. Shake-it-off: a simple ultrasonic cryo-EM specimen-preparation device. Acta Crystallogr D. 2019;75:1063-70. PubMed PMID: WOS:000501134600003.

100. Ravelli RBG, Nijpels FJT, Henderikx RJM, Weissenberger G, Thewessem S, Gijsbers A, Beulen BWAMM, Lopez-Iglesias C, Peters PJ. Cryo-EM structures from

sub-nl volumes using pin-printing and jet vitrification. Nat Commun. 2020;11(1). PubMed PMID: WOS:000537134500010.

101. Henderson R, Unwin PNT. 3-Dimensional Model of Purple Membrane Obtained by Electron-Microscopy. Nature. 1975;257(5521):28-32. PubMed PMID: WOS:A1975AN78300016.

102. Chou FC, Sripakdeevong P, Dibrov SM, Hermann T, Das R. Correcting pervasive errors in RNA crystallography through enumerative structure prediction. Nat Methods. 2013;10(1):74-U105. PubMed PMID: WOS:000312810100041.

103. Brilot AF, Chen JZ, Cheng AC, Pan JH, Harrison SC, Potter CS, Carragher B, Henderson R, Grigorieff N. Beam-induced motion of vitrified specimen on holey carbon film. J Struct Biol. 2012;177(3):630-7. PubMed PMID: WOS:000301894900005.

104. Derosier DJ, Klug A. Reconstruction of 3 Dimensional Structures from Electron Micrographs. Nature. 1968;217(5124):130-&. PubMed PMID: WOS:A1968A441200006.

105. Frank J. Averaging of Low Exposure Electron-Micrographs of Non-Periodic Objects. Ultramicroscopy. 1975;1(2):159-62. PubMed PMID: WOS:A1975DT13700010.

106. Frank J, Goldfarb W, Kessel M, Eisenberg D, Baker TS. Reconstruction of Glutamine-Synthetase from Electron-Micrographs by Computer Averaging. Biophys J. 1978;21(3):A89-A. PubMed PMID: WOS:A1978ER10500442.

107. Penczek PA, Grassucci RA, Frank J. The Ribosome at Improved Resolution - New Techniques for Merging and Orientation Refinement in 3d Cryoelectron Microscopy of Biological Particles. Ultramicroscopy. 1994;53(3):251-70. PubMed PMID: WOS:A1994NA85900006.

108. Shaikh TR, Gao HX, Baxter WT, Asturias FJ, Boisset N, Leith A, Frank J. SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. Nat Protoc. 2008;3(12):1941-74. PubMed PMID: WOS:000265781700013.

109. vanHeel M, Harauz G, Orlova EV, Schmidt R, Schatz M. A new generation of the IMAGIC image processing system. J Struct Biol. 1996;116(1):17-24. PubMed PMID: WOS:A1996TX60600004.
110. Ludtke SJ, Baldwin PR, Chiu W. EMAN: Semiautomated software for high-resolution single-particle reconstructions. J Struct Biol. 1999;128(1):82-97. PubMed PMID: WOS:000084451600013.
111. Grigorieff N. FREALIGN: High-resolution refinement of single particle structures. J Struct Biol. 2007;157(1):117-25. PubMed PMID: WOS:000243391100013.
112. Scheres SHW. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol. 2012;180(3):519-30. PubMed PMID: WOS:000311471000014.
113. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods. 2017;14(3):290-+. PubMed PMID: WOS:000395661700027.
114. Mastronarde DN. Automated electron microscope tomography using robust prediction of specimen movements. J Struct Biol. 2005;152(1):36-51. PubMed PMID: WOS:000232572300004.
115. Carragher B, Kisseberth N, Kriegman D, Milligan RA, Potter CS, Pulokas J, Reilein A. Leginon: An automated system for acquisition of images from vitreous ice specimens. J Struct Biol. 2000;132(1):33-45. PubMed PMID: WOS:000166379300004.
116. Oikonomou CM, Jensen GJ. A new view into prokaryotic cell biology from electron cryotomography. Nat Rev Microbiol. 2016;14(4):205-20. PubMed PMID: WOS:000372318700009.
117. Oikonomou CM, Jensen GJ. Cellular Electron Cryotomography: Toward Structural Biology In Situ. Annual Review of Biochemistry, Vol 86. 2017;86:873-96. PubMed PMID: WOS:000407725800035.

118. Schertel A, Snaidero N, Han HM, Ruhwedel T, Laue M, Grabenbauer M, Mobius W. Cryo FIB-SEM: Volume imaging of cellular ultrastructure in native frozen specimens. J Struct Biol. 2013;184(2):355-60. PubMed PMID: WOS:000327171000029.
119. Zachs T, Schertel A, Medeiros J, Weiss GL, Hugener J, Matos J, Pilhofer M. Fully automated, sequential focused ion beam milling for cryo-electron tomography. Elife. 2020;9. PubMed PMID: WOS:000521157600001.
120. Jonathan Bouvette H-FL, Xiaochen Du, Ye Zhou, Andrew P. Sikkema, Juliana Da Fonseca Rezende E Mello, Bradley Klemm, Rick Huang, Roel M. Schaaper, Mario J. Borgnia, Alberto Bartesaghi. Beam image-shift accelerated data acquisition for near-atomic resolution single-particle cryo-electron tomography. bioRxiv. 2020.
121. Ousalem F, Singh S, Chesneau O, Hunt JF, Boel G. ABC-F proteins in mRNA translation and antibiotic resistance. Res Microbiol. 2019;170(8):435-47. PubMed PMID: WOS:000500919300015.
122. Aye ILMH, Paxton JW, Evseenko DA, Keelan JA. Expression, localisation and activity of ATP binding cassette (ABC) family of drug transporters in human amnion membranes. Placenta. 2007;28(8-9):868-77. PubMed PMID: WOS:000248720500015.
123. Sharkey LKR, Edwards TA, O'Neill AJ. ABC-F Proteins Mediate Antibiotic Resistance through Ribosomal Protection. Mbio. 2016;7(2). PubMed PMID: WOS:000377768700072.
124. Chakraburtty K, Triana-Alonso FJ. Yeast elongation factor 3: Structure and function. Biol Chem. 1998;379(7):831-40. PubMed PMID: WOS:000075141900008.
125. Kurata S, Nielsen KH, Mitchell SF, Lorsch JR, Kaji A, Kaji H. Ribosome recycling step in yeast cytoplasmic protein synthesis is catalyzed by eEF3 and ATP. P Natl Acad Sci USA. 2010;107(24):10854-9. PubMed PMID: WOS:000278807400017.
126. Dealdana CRV, Marton MJ, Hinnebusch AG. Gcn20, a Novel Atp Binding Cassette Protein, and Gcn1 Reside in a Complex That Mediates Activation of the Eif-2-Alpha Kinase Gcn2 in Amino Acid-Starved Cells. Embo J. 1995;14(13):3184-99. PubMed PMID: WOS:A1995RH80000022.

127. Dong JS, Lai R, Jennings JL, Link AJ, Hinnebusch AG. The novel ATP-binding cassette protein ARB1 is a shuttling factor that stimulates 40S and 60S ribosome biogenesis. Mol Cell Biol. 2005;25(22):9859-73. PubMed PMID: WOS:000233046700011.

128. Li ZH, Lee I, Moradi E, Hung NJ, Johnson AW, Marcotte EM. Rational Extension of the Ribosome Biogenesis Pathway Using Network-Guided Genetics. Plos Biol. 2009;7(10). PubMed PMID: WOS:000272031800003.

129. Kerr ID. Sequence analysis of twin ATP binding cassette proteins involved in translational control, antibiotic resistance, and ribonuclease L inhibition. Biochem Bioph Res Co. 2004;315(1):166-73. PubMed PMID: WOS:000188966400025.

130. Wilkens S. Structure and mechanism of ABC transporters. F1000prime reports. 2015;7:14. doi: 10.12703/P7-14. PubMed PMID: 25750732; PubMed Central PMCID: PMC4338842.

131. Davidson AL, Dassa E, Orelle C, Chen J. Structure, function, and evolution of bacterial ATP-binding cassette systems. Microbiology and molecular biology reviews : MMBR. 2008;72(2):317-64, table of contents. doi: 10.1128/MMBR.00031-07. PubMed PMID: 18535149; PubMed Central PMCID: PMC2415747.

132. Hofmann S, Januliene D, Mehdipour AR, Thomas C, Stefan E, Bruchert S, Kuhn BT, Geertsma ER, Hummer G, Tampe R, Moeller A. Conformation space of a heterodimeric ABC exporter under turnover conditions. Nature. 2019;571(7766):580-3. doi: 10.1038/s41586-019-1391-0. PubMed PMID: 31316210.

133. Smith PC, Karpowich N, Millen L, Moody JE, Rosen J, Thomas PJ, Hunt JF. ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. Mol Cell. 2002;10(1):139-49. doi: 10.1016/s1097-2765(02)00576-2. PubMed PMID: 12150914; PubMed Central PMCID: PMC3516284.

134. Procko E, Ferrin-O'Connell I, Ng SL, Gaudet R. Distinct structural and functional properties of the ATPase sites in an asymmetric ABC transporter. Mol Cell. 2006;24(1):51-62. doi: 10.1016/j.molcel.2006.07.034. PubMed PMID: 17018292.

135. Daniel J, Abraham L, Martin A, Pablo X, Reyes S. Rv2477c is an antibiotic-sensitive manganese-dependent ABC-F ATPase in Mycobacterium tuberculosis. Biochem Biophys Res Commun. 2018;495(1):35-40. doi: 10.1016/j.bbrc.2017.10.168. PubMed PMID: 29101040.

136. Faksri K, Tan JH, Disratthakit A, Xia E, Prammananan T, Suriyaphol P, Khor CC, Teo YY, Ong RT, Chaiprasert A. Whole-Genome Sequencing Analysis of Serially Isolated Multi-Drug and Extensively Drug Resistant Mycobacterium tuberculosis from Thai Patients. Plos One. 2016;11(8):e0160992. doi: 10.1371/journal.pone.0160992. PubMed PMID: 27518818; PubMed Central PMCID: PMC4982626.

137. Black PA, Warren RM, Louw GE, van Helden PD, Victor TC, Kana BD. Energy metabolism and drug efflux in Mycobacterium tuberculosis. Antimicrob Agents Chemother. 2014;58(5):2491-503. doi: 10.1128/AAC.02293-13. PubMed PMID: 24614376; PubMed Central PMCID: PMC3993223.

138. DeJesus MA, Gerrick ER, Xu W, Park SW, Long JE, Boutte CC, Rubin EJ, Schnappinger D, Ehrt S, Fortune SM, Sassetti CM, Ioerger TR. Comprehensive Essentiality Analysis of the Mycobacterium tuberculosis Genome via Saturating Transposon Mutagenesis. Mbio. 2017;8(1). doi: 10.1128/mBio.02133-16. PubMed PMID: 28096490; PubMed Central PMCID: PMC5241402.

139. Gupta AK, Katoch VM, Chauhan DS, Sharma R, Singh M, Venkatesan K, Sharma VD. Microarray analysis of efflux pump genes in multidrug-resistant Mycobacterium tuberculosis during stress induced by common anti-tuberculous drugs. Microbial drug resistance. 2010;16(1):21-8. doi: 10.1089/mdr.2009.0054. PubMed PMID: 20001742.

140. Philalay JS, Palermo CO, Hauge KA, Rustad TR, Cangelosi GA. Genes required for intrinsic multidrug resistance in Mycobacterium avium. Antimicrob Agents Chemother. 2004;48(9):3412-8. doi: 10.1128/AAC.48.9.3412-3418.2004. PubMed PMID: 15328105; PubMed Central PMCID: PMC514743.

141. Li X, Sun Q, Jiang C, Yang K, Hung LW, Zhang J, Sacchettini JC. Structure of Ribosomal Silencing Factor Bound to Mycobacterium tuberculosis Ribosome. Structure. 2015;23(12):2387. doi: 10.1016/j.str.2015.11.002. PubMed PMID: 29673486.
142. Selmer M, Dunham CM, Murphy FVt, Weixlbaumer A, Petry S, Kelley AC, Weir JR, Ramakrishnan V. Structure of the 70S ribosome complexed with mRNA and tRNA. Science. 2006;313(5795):1935-42. doi: 10.1126/science.1131127. PubMed PMID: 16959973.
143. Zhang K. Gctf: Real-time CTF determination and correction. J Struct Biol. 2016;193(1):1-12. PubMed PMID: WOS:000368317700001.
144. Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJH, Lindahl E, Scheres SHW. New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife. 2018;7. PubMed PMID: WOS:000450857100001.
145. Henderson R, Sali A, Baker ML, Carragher B, Devkota B, Downing KH, Egelman EH, Feng ZK, Frank J, Grigorieff N, Jiang W, Ludtke SJ, Medalia O, Penczek PA, Rosenthal PB, Rossmann MG, Schmid MF, Schroder GF, Steven AC, Stokes DL, Westbrook JD, Wriggers W, Yang HW, Young J, Berman HM, Chiu W, Kleywegt GJ, Lawson CL. Outcome of the First Electron Microscopy Validation Task Force Meeting. Structure. 2012;20(2):205-14. PubMed PMID: WOS:000300388000005.
146. Kucukelbir A, Sigworth FJ, Tagare HD. Quantifying the local resolution of cryo-EM density maps. Nat Methods. 2014;11(1):63-5. doi: 10.1038/nmeth.2727. PubMed PMID: 24213166; PubMed Central PMCID: PMC3903095.
147. Ramirez-Aportela E, Vilas JL, Glukhova A, Melero R, Conesa P, Martinez M, Maluenda D, Mota J, Jimenez A, Vargas J, Marabini R, Sexton PM, Carazo JM, Sorzano COS. Automatic local resolution-based sharpening of cryo-EM maps. Bioinformatics. 2020;36(3):765-72. doi: 10.1093/bioinformatics/btz671. PubMed PMID: 31504163.
148. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods in enzymology. 1997;276:307-26. PubMed PMID: 27754618.

149. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta crystallographica Section D, Biological crystallography. 2011;67(Pt 4):271-81. doi: 10.1107/S0907444910048675. PubMed PMID: 21460445; PubMed Central PMCID: PMC3069742.

150. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. Journal of applied crystallography. 2007;40(Pt 4):658-74. doi: 10.1107/S0021889807021206. PubMed PMID: 19461840; PubMed Central PMCID: PMC2483472.

151. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018;46(W1):W296-W303. doi: 10.1093/nar/gky427. PubMed PMID: 29788355; PubMed Central PMCID: PMC6030848.

152. Liebschner D, Afonine PV, Baker ML, Bunkoczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ, Moriarty NW, Oeffner RD, Poon BK, Prisant MG, Read RJ, Richardson JS, Richardson DC, Sammito MD, Sobolev OV, Stockwell DH, Terwilliger TC, Urzhumtsev AG, Videau LL, Williams CJ, Adams PD. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D. 2019;75:861-77. PubMed PMID: WOS:000497657100001.

153. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallographica Section D-Biological Crystallography. 2010;66:486-501. PubMed PMID: WOS:000275941300018.

154. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF chimera - A visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605-12. PubMed PMID: WOS:000223379100005.

155. Svidritskiy E, Demo G, Loveland AB, Xu C, Korostelev AA. Extensive ribosome and RF2 rearrangements during translation termination. Elife. 2019;8. doi:

10.7554/eLife.46850. PubMed PMID: 31513010; PubMed Central PMCID: PMC6742477.

156. Graf M, Huter P, Maracci C, Peterek M, Rodnina MV, Wilson DN. Visualization of translation termination intermediates trapped by the Apidaecin 137 peptide during RF3-mediated recycling of RF1. Nat Commun. 2018;9(1):3053. doi: 10.1038/s41467-018-05465-1. PubMed PMID: 30076302; PubMed Central PMCID: PMC6076264.

157. Dong Q, Ernst SE, Ostedgaard LS, Shah VS, Ver Heul AR, Welsh MJ, Randak CO. Mutating the Conserved Q-loop Glutamine 1291 Selectively Disrupts Adenylate Kinase-dependent Channel Gating of the ATP-binding Cassette (ABC) Adenylate Kinase Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) and Reduces Channel Function in Primary Human Airway Epithelia. J Biol Chem. 2015;290(22):14140-53. doi: 10.1074/jbc.M114.611616. PubMed PMID: 25887396; PubMed Central PMCID: PMC4447984.

158. Zolnerciks JK, Akkaya BG, Snippe M, Chiba P, Seelig A, Linton KJ. The Q loops of the human multidrug resistance transporter ABCB1 are necessary to couple drug binding to the ATP catalytic cycle. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2014;28(10):4335-46. doi: 10.1096/fj.13-245639. PubMed PMID: 25016028.

159. Qu L, Jiang Y, Cheng CY, Wu D, Meng B, Chen ZR, Zhu YP, Shaw N, Ouyang SY, Liu ZJ. Crystal Structure of ATP-Bound Human ABCF1 Demonstrates a Unique Conformation of ABC Proteins. Structure. 2018;26(9):1259-+. PubMed PMID: WOS:000443567100011.

160. Ernst R, Kueppers P, Klein CM, Schwarzmueller T, Kuchler K, Schmitt L. A mutation of the H-loop selectively affects rhodamine transport by the yeast multidrug ABC transporter Pdr5. Proc Natl Acad Sci U S A. 2008;105(13):5069-74. doi: 10.1073/pnas.0800191105. PubMed PMID: 18356296; PubMed Central PMCID: PMC2278221.

161. Ambudkar SV, Kim IW, Xia D, Sauna ZE. The A-loop, a novel conserved aromatic acid subdomain upstream of the Walker A motif in ABC transporters, is critical for ATP binding. FEBS letters. 2006;580(4):1049-55. doi: 10.1016/j.febslet.2005.12.051. PubMed PMID: 16412422.
162. Westfahl KM, Merten JA, Buchaklian AH, Klug CS. Functionally important ATP binding and hydrolysis sites in Escherichia coli MsbA. Biochemistry-Us. 2008;47(52):13878-86. doi: 10.1021/bi801745u. PubMed PMID: 19053284; PubMed Central PMCID: PMC2637178.
163. Whitford PC, Geggier P, Altman RB, Blanchard SC, Onuchic JN, Sanbonmatsu KY. Accommodation of aminoacyl-tRNA into the ribosome involves reversible excursions along multiple pathways. Rna. 2010;16(6):1196-204. doi: 10.1261/rna.2035410. PubMed PMID: 20427512; PubMed Central PMCID: PMC2874171.
164. Whitford PC, Onuchic JN, Sanbonmatsu KY. Connecting energy landscapes with experimental rates for aminoacyl-tRNA accommodation in the ribosome. Journal of the American Chemical Society. 2010;132(38):13170-1. doi: 10.1021/ja1061399. PubMed PMID: 20806913; PubMed Central PMCID: PMC2945706.
165. Cohen A, Mathiasen VD, Schon T, Wejse C. The global prevalence of latent tuberculosis: a systematic review and meta-analysis. Eur Respir J. 2019;54(3). doi: 10.1183/13993003.00655-2019. PubMed PMID: 31221810.
166. de Paus RA, van Meijgaarden KE, Prins C, Kamphorst MH, Arend SM, Ottenhoff THM, Joosten SA. Immunological characterization of latent tuberculosis infection in a low endemic country. Tuberculosis (Edinb). 2017;106:62-72. doi: 10.1016/j.tube.2017.07.001. PubMed PMID: 28802407.
167. Voskuil MI. Mycobacterium tuberculosis gene expression during environmental conditions associated with latency. Tuberculosis (Edinb). 2004;84(3-4):138-43. doi: 10.1016/j.tube.2003.12.008. PubMed PMID: 15207483.

168. Steitz TA. A structural understanding of the dynamic ribosome machine. Nature reviews Molecular cell biology. 2008;9(3):242-53. doi: 10.1038/nrm2352. PubMed PMID: 18292779.
169. Starosta AL, Lassak J, Jung K, Wilson DN. The bacterial translation stress response. Fems Microbiol Rev. 2014;38(6):1172-201. doi: 10.1111/1574-6976.12083. PubMed PMID: 25135187; PubMed Central PMCID: PMC4227928.
170. Polikanov YS, Blaha GM, Steitz TA. How Hibernation Factors RMF, HPF, and YfiA Turn Off Protein Synthesis. Science. 2012;336(6083):915-8. PubMed PMID: WOS:000304145600065.
171. Kato T, Yoshida H, Miyata T, Maki Y, Wada A, Namba K. Structure of the 100S ribosome in the hibernation stage revealed by electron cryomicroscopy. Structure. 2010;18(6):719-24. doi: 10.1016/j.str.2010.02.017. PubMed PMID: 20541509.
172. Khusainov I, Vicens Q, Ayupov R, Usachev K, Myasnikov A, Simonetti A, Validov S, Kieffer B, Yusupova G, Yusupov M, Hashem Y. Structures and dynamics of hibernating ribosomes from Staphylococcus aureus mediated by intermolecular interactions of HPF. Embo J. 2017;36(14):2073-87. doi: 10.15252/embj.201696105. PubMed PMID: 28645916; PubMed Central PMCID: PMC5510003.
173. Matzov D, Aibara S, Basu A, Zimmerman E, Bashan A, Yap MF, Amunts A, Yonath AE. The cryo-EM structure of hibernating 100S ribosome dimer from pathogenic Staphylococcus aureus. Nat Commun. 2017;8(1):723. doi: 10.1038/s41467-017-00753-8. PubMed PMID: 28959035; PubMed Central PMCID: PMC5620080.
174. Franken LE, Oostergetel GT, Pijning T, Puri P, Arkhipova V, Boekema EJ, Poolman B, Guskov A. A general mechanism of ribosome dimerization revealed by single-particle cryo-electron microscopy. Nat Commun. 2017;8(1):722. doi: 10.1038/s41467-017-00718-x. PubMed PMID: 28959045; PubMed Central PMCID: PMC5620043.
175. Beckert B, Abdelshahid M, Schafer H, Steinchen W, Arenz S, Berninghausen O, Beckmann R, Bange G, Turgay K, Wilson DN. Structure of the Bacillus subtilis

hibernating 100S ribosome reveals the basis for 70S dimerization. Embo J. 2017;36(14):2061-72. doi: 10.15252/embj.201696189. PubMed PMID: 28468753; PubMed Central PMCID: PMC5509997.

176. Flygaard RK, Boegholm N, Yusupov M, Jenner LB. Cryo-EM structure of the hibernating Thermus thermophilus 100S ribosome reveals a protein-mediated dimerization mechanism. Nat Commun. 2018;9(1):4179. doi: 10.1038/s41467-018-06724-x. PubMed PMID: 30301898; PubMed Central PMCID: PMC6177447.

177. Mishra S, Ahmed T, Tyagi A, Shi J, Bhushan S. Structures of Mycobacterium smegmatis 70S ribosomes in complex with HPF, tmRNA, and P-tRNA. Scientific reports. 2018;8(1):13587. doi: 10.1038/s41598-018-31850-3. PubMed PMID: 30206241; PubMed Central PMCID: PMC6133939.

178. Shukla AK, Westfield GH, Xiao K, Reis RI, Huang LY, Tripathi-Shukla P, Qian J, Li S, Blanc A, Oleskie AN, Dosey AM, Su M, Liang CR, Gu LL, Shan JM, Chen X, Hanna R, Choi M, Yao XJ, Klink BU, Kahsai AW, Sidhu SS, Koide S, Penczek PA, Kossiakoff AA, Woods VL, Jr., Kobilka BK, Skiniotis G, Lefkowitz RJ. Visualization of arrestin recruitment by a G-protein-coupled receptor. Nature. 2014;512(7513):218-22. doi: 10.1038/nature13430. PubMed PMID: 25043026; PubMed Central PMCID: PMC4134437.

179. Davies G, Boeree M, Hermann D, Hoelscher M. Accelerating the transition of new tuberculosis drug combinations from Phase II to Phase III trials: New technologies and innovative designs. PLoS medicine. 2019;16(7):e1002851. doi: 10.1371/journal.pmed.1002851. PubMed PMID: 31287813; PubMed Central PMCID: PMC6615592.

180. Pelchovich G, Zhuravlev A, Gophna U. Effect of ribosome-targeting antibiotics on streptomycin-resistant Mycobacterium mutants in the rpsL gene. International journal of antimicrobial agents. 2013;42(2):129-32. doi: 10.1016/j.ijantimicag.2013.04.001. PubMed PMID: 23664678.

181. Agyeman AA, Ofori-Asenso R. Efficacy and safety profile of linezolid in the treatment of multidrug-resistant (MDR) and extensively drug-resistant (XDR) tuberculosis: a systematic review and meta-analysis. Annals of clinical microbiology and antimicrobials. 2016;15(1):41. doi: 10.1186/s12941-016-0156-y. PubMed PMID: 27334498; PubMed Central PMCID: PMC4917997.

182. Koh WJ, Kang YR, Jeon K, Kwon OJ, Lyu J, Kim WS, Shim TS. Daily 300 mg dose of linezolid for multidrug-resistant and extensively drug-resistant tuberculosis: updated analysis of 51 patients. J Antimicrob Chemoth. 2012;67(6):1503-7. PubMed PMID: WOS:000304064400033.

183. Williams KN, Stover CK, Zhu T, Tasneen R, Tyagi S, Grosset JH, Nuermberger E. Promising antituberculosis activity of the oxazolidinone PNU-100480 relative to that of linezolid in a murine model. Antimicrob Agents Chemother. 2009;53(4):1314-9. doi: 10.1128/AAC.01182-08. PubMed PMID: 19075058; PubMed Central PMCID: PMC2663111.

184. Kanoh S, Rubin BK. Mechanisms of action and clinical application of macrolides as immunomodulatory medications. Clinical microbiology reviews. 2010;23(3):590-615. doi: 10.1128/CMR.00078-09. PubMed PMID: 20610825; PubMed Central PMCID: PMC2901655.

185. Buriankova K, Doucet-Populaire F, Dorson O, Gondran A, Ghnassia JC, Weiser J, Pernodet JL. Molecular basis of intrinsic macrolide resistance in the Mycobacterium tuberculosis complex. Antimicrob Agents Chemother. 2004;48(1):143-50. doi: 10.1128/aac.48.1.143-150.2004. PubMed PMID: 14693532; PubMed Central PMCID: PMC310192.

186. Luna-Herrera J, Reddy VM, Daneluzzi D, Gangadharam PR. Antituberculosis activity of clarithromycin. Antimicrob Agents Chemother. 1995;39(12):2692-5. doi: 10.1128/aac.39.12.2692. PubMed PMID: 8593004; PubMed Central PMCID: PMC163014.

187. Truffot-Pernot C, Lounis N, Grosset JH, Ji B. Clarithromycin is inactive against Mycobacterium tuberculosis. Antimicrob Agents Chemother. 1995;39(12):2827-8. doi: 10.1128/aac.39.12.2827. PubMed PMID: 8593032; PubMed Central PMCID: PMC163042.

188. Raleigh JW. Rifampin: clinical experience with a new anti-tuberculosis drug. Transactions of the American Clinical and Climatological Association. 1972;83:104-12. PubMed PMID: 5040107; PubMed Central PMCID: PMC2441274.

189. Hoagland DT, Liu J, Lee RB, Lee RE. New agents for the treatment of drug-resistant Mycobacterium tuberculosis. Advanced drug delivery reviews. 2016;102:55-72. doi: 10.1016/j.addr.2016.04.026. PubMed PMID: 27151308; PubMed Central PMCID: PMC4903924.

190. Swartz JR, Jewett MC, Woodrow KA. Cell-free protein synthesis with prokaryotic combined transcription-translation. Methods in molecular biology. 2004;267:169-82. doi: 10.1385/1-59259-774-2:169. PubMed PMID: 15269424.

191. Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng YF, Agard DA. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods. 2017;14(4):331-2. PubMed PMID: WOS:000397900500004.

192. Moriarty NW, Grosse-Kunstleve RW, Adams PD. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. Acta Crystallogr D. 2009;65:1074-80. PubMed PMID: WOS:000269845500007.

193. Dunkle JA, Xiong L, Mankin AS, Cate JH. Structures of the Escherichia coli ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action. Proc Natl Acad Sci U S A. 2010;107(40):17152-7. doi: 10.1073/pnas.1007988107. PubMed PMID: 20876128; PubMed Central PMCID: PMC2951456.

194. Hansen JL, Ippolito JA, Ban N, Nissen P, Moore PB, Steitz TA. The structures of four macrolide antibiotics bound to the large ribosomal subunit. Mol Cell. 2002;10(1):117-28. doi: 10.1016/s1097-2765(02)00570-1. PubMed PMID: 12150912.

195. Samir P, Browne CM, Rahul, Sun M, Shen BX, Li W, Frank J, Link AJ. Identification of Changing Ribosome Protein Compositions using Mass Spectrometry. Proteomics. 2018;18(20). PubMed PMID: WOS:000448086000002.

196. Hauser R, Pech M, Kijek J, Yamamoto H, Titz B, Naeve F, Tovchigrechko A, Yamamoto K, Szaflarski W, Takeuchi N, Stellberger T, Diefenbacher ME, Nierhaus KH, Uetz P. RsfA (YbeB) proteins are conserved ribosomal silencing factors. Plos Genet. 2012;8(7):e1002815. doi: 10.1371/journal.pgen.1002815. PubMed PMID: 22829778; PubMed Central PMCID: PMC3400551.

197. Brown A, Rathore S, Kimanius D, Aibara S, Bai XC, Rorbach J, Amunts A, Ramakrishnan V. Structures of the human mitochondrial ribosome in native states of assembly. Nat Struct Mol Biol. 2017;24(10):866-9. doi: 10.1038/nsmb.3464. PubMed PMID: 28892042; PubMed Central PMCID: PMC5633077.

198. Soufari H, Waltz F, Parrot C, Durrieu-Gaillard S, Bochler A, Kuhn L, Sissler M, Hashem Y. Structure of the mature kinetoplastids mitoribosome and insights into its large subunit biogenesis. P Natl Acad Sci USA. 2020;117(47):29851-61. PubMed PMID: WOS:000594737400007.

APPENDIX A

ASYMMETRIC CRYO-EM STRUCTURE OF THE CANONICAL ALLOLEVIVIRUS QB REVEALS A SINGLE MATURATION PROTEIN AND THE GENOMIC SSRNA IN SITU*

*Reprinted with permission from "Asymmetric cryo-EM structure of the canonical Allolevivirus Qb reveals a single maturation protein and the genomic ssRNA in situ" by Karl V. Gorzelnik, Zhicheng Cui, Catrina A. Reed, Joanita Jakana, Ry Young, Junjie Zhang. 2016. PNAS, 113, 11519-11524, Authors retain copyright.

Asymmetric cryo-EM structure of the canonical *Allolevivirus* Qβ reveals a single maturation protein and the genomic ssRNA in situ

Karl V. Gorzelnik[a,1], Zhicheng Cui[a,1], Catrina A. Reed[a,2], Joanita Jakana[b], Ry Young[a], and Junjie Zhang[a,3]

[a]Center for Phage Technology, Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843; and [b]National Center for Macromolecular Imaging, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030

Edited by David DeRosier, Brandeis University, Waltham, MA, and approved August 5, 2016 (received for review June 21, 2016)

Single-stranded (ss) RNA viruses infect all domains of life. To date, for most ssRNA virions, only the structures of the capsids and their associated protein components have been resolved to high resolution. Qβ, an ssRNA phage specific for the conjugative F-pilus, has a T = 3 icosahedral lattice of coat proteins assembled around its 4,217 nucleotides of genomic RNA (gRNA). In the mature virion, the maturation protein, A_2, binds to the gRNA and is required for adsorption to the F-pilus. Here, we report the cryo-electron microscopy (cryo-EM) structures of Qβ with and without symmetry applied. The icosahedral structure, at 3.7-Å resolution, resolves loops not previously seen in the published X-ray structure, whereas the asymmetric structure, at 7-Å resolution, reveals A_2 and the gRNA. A_2 contains a bundle of α-helices and replaces one dimer of coat proteins at a twofold axis. The helix bundle binds gRNA, causing denser packing of RNA in its proximity, which asymmetrically expands the surrounding coat protein shell to potentially facilitate RNA release during infection. We observe a fixed pattern of gRNA organization among all viral particles, with the major and minor grooves of RNA helices clearly visible. A single layer of RNA directly contacts every copy of the coat protein, with one-third of the interactions occurring at operator-like RNA hairpins. These RNA–coat interactions stabilize the tertiary structure of gRNA within the virion, which could further provide a roadmap for capsid assembly.

single-particle cryo-EM | *Allolevivirus* | ssRNA virus | genomic RNA | maturation protein

Single-stranded (ss) RNA viruses are an abundant type of virus and infect all domains of life (1–4). One of the best-studied ssRNA virus systems is the *Leviviridae*, which infects Gram-negative bacteria via a variety of retractile pili (5); extensive genetic and biochemical studies have been performed on two of these phages: MS2 and Qβ (5–11). All of the *Leviviridae* have the same core genome, spanning 3.4–4.3 kb, encoding the maturation protein, the coat protein, and a subunit of RNA-dependent RNA replicase (*SI Appendix*, Fig. S1) (10). The MS2-like phages, designated true leviviruses, have a fourth gene that encodes the lysis protein, whereas the Qβ-like phages, designated alloleviviruses, have the lysis function as an additional feature of the maturation protein (called A_2 in Qβ). Qβ also encodes a minor coat protein, called A_1, arising from occasional read-through of the stop codon of the major coat protein; it has been estimated that the A_1 protein replaces 3–10 copies of the major coat protein in the virion (11) and is required for infection (12). A_1 consists of a coat domain and a read-through domain separated by a flexible linker (13). Unlike most dsDNA phages, which use specialized protein machinery to pump their genomic DNA into a capsid preassembled around a protein scaffold (14–17), ssRNA viruses, including the *Leviviridae*, assemble their coat proteins around the genomic RNA (gRNA), presumably because the extremely small genome does not allow for genes to encode proteins that help package the genetic material. Therefore, ssRNA viruses require direct gRNA–coat protein interactions, some of which are specific, to self-assemble the virion (18). This raises interesting questions, including how the viral RNA is selectively encapsidated over host RNAs. Addressing these questions will lead to a better understanding of the physiology of ssRNA viruses, and potentially novel therapeutics against them (19).

In *Leviviridae*, both Qβ and MS2 have capsids with T = 3 morphology, meaning each coat protein monomer adopts one of three conformers, termed A, B, and C (20, 21). Within the viral capsid, conformers A and B form an asymmetric A/B dimer and two C conformers form a symmetric C/C dimer. The gRNAs of Qβ and MS2 form secondary structures both in vitro (22, 23) and in vivo (7). There is a specific RNA hairpin near the start of the replicase gene called the "operator," which has a high binding affinity for the coat protein (24, 25). Although both MS2 and Qβ operators form hairpins, the coat proteins for each phage are selectively attracted to their specific operator to the point at which, if purified RNAs from both phages are mixed with the coat protein of just one phage, only the cognate RNA is encapsidated (8). In MS2, the maturation protein binds to specific regions on the 5′ and 3′ of the gRNA during capsid assembly (26).

Neither the maturation protein nor the gRNA of *Leviviridae* was resolved in the previous high-resolution structural studies by X-ray crystallography, most likely because of the asymmetry of these two components in the stacking of symmetric particles within

Significance

Single-stranded (ss) RNA viruses have ribonucleic acid as their genetic material and infect animals, plants, and bacteria. Here we used cryo-electron microscopy to reveal the genomic RNA (gRNA) of the ssRNA virus Qβ. The asymmetric gRNA adopts a single dominant structure in all virions and binds the capsid of Qβ at each coat protein. At the same time, we determined the structure of the maturation protein, A_2, which functions both as the virion's "tail" and its lysis protein. We see the gRNA is more ordered when interacting with A_2. These results provide structural insights into gRNA packaging and host infection in ssRNA viruses.

Author contributions: J.Z. designed research; K.V.G., Z.C., J.J., and J.Z. performed research; C.A.R. and R.Y. contributed new reagents/analytic tools; K.V.G., Z.C., and J.Z. analyzed data; and K.V.G., Z.C., R.Y., and J.Z. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The data reported in this paper have been deposited in the EMDataBank database, www.emdatabank.org (accession nos. EMD-8253, EMD-8254, and EMD-8255). The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 5KIP).

See Commentary on page 11390.

[1]K.V.G. and Z.C. contributed equally to this work.

[2]Present address: Molecular and Cellular Medicine, College of Medicine, Texas A&M Health Science Center, College Station, TX 77843.

[3]To whom correspondence should be addressed. Email: junjiez@tamu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1609482113/-/DCSupplemental.

the crystals (20, 21). The first glimpse of RNA within the capsid came from a low-resolution structure of MS2 using single-particle cryo-EM (27). On the basis of another single-particle cryo-EM structure with icosahedral symmetry applied, it was proposed that MS2 has two concentric shells of RNA within the phage capsid (28). Subsequent work, using cryo-electron tomography and subtomogram averaging, yielded a 39-Å resolution symmetry-free density map of MS2 (29), suggesting the gRNA adopts the same conformation for each virus particle within the capsid and the maturation protein replaces a C/C dimer of the coat proteins. However, at this low resolution, it was not clear how the maturation protein, which contributes ∼1% of the molecular mass of the entire virion, interacts with the rest of the capsid or how the gRNA is organized inside the capsid shell.

In this study, we report the cryo-EM structures of the canonical *Allolevivirus* Qβ with and without symmetry applied, at 3.7- and 7-Å resolutions, respectively. Our structures reveal features never seen before for Qβ, such as the structure of A_2, symmetry deviation of the coat proteins, organization of gRNA, and the interactions between them. These results are discussed in terms of a model for viral assembly and gRNA release in Qβ.

Results

Icosahedral Structure of Phage Qβ. Ice-embedded Qβ phage particles were imaged using cryo-EM (*SI Appendix*, Fig. S2). A total of 712 image stacks were recorded on a direct detection camera in superresolution electron counting mode (30), yielding 51,815 high-quality phage particle images. Icosahedral symmetry was first applied to generate a structure of the coat protein shell at 3.7-Å resolution (*SI Appendix*, Fig. S3). Fig. 1 shows the icosahedral cryo-EM density map of Qβ with 180 copies of the 14-kDa coat protein. Each asymmetric unit within the capsid consists of three conformers: A, B, and C (Fig. 1*A*). Both types of dimers (A/B and C/C) are revealed. Protein α-helices and β-strands are clearly resolved throughout the cryo-EM density, with most of the bulky side-chains visible (Fig. 1 *B* and *C*). The cryo-EM structure is consistent with the previously solved crystal structure of the icosahedral Qβ phage capsid. PDB coordinates of the crystal structure fit into the cryo-EM density without the need for major adjustments (*SI Appendix*, Fig. S4). Two loops within the coat protein of note are the EF-loop, which protrudes into the interior of the capsid, and the FG-loop, which forms small pores at both the fivefold and threefold vertices of the capsid (*SI Appendix*, Fig. S5) (20). In our cryo-EM structure, the previously missing protein segment (residues 76–79) at the tip of the FG-loop of both conformers A and C is resolved (Fig. 1*D*), allowing us to build a more complete model for the capsid. The additional resolved residues are at the threefold vertices, with alternating FG-loops from conformers A and C. These residues form an ∼15-Å pore, in agreement with its predicted size (20). Residues 76–79 of conformer B are resolved both in our cryo-EM reconstruction and the previous crystal structure, mostly because of the location at the fivefold vertices, which allows smaller pores (∼8 Å in diameter) between neighboring subunits, restricting the movement of their FG-loops.

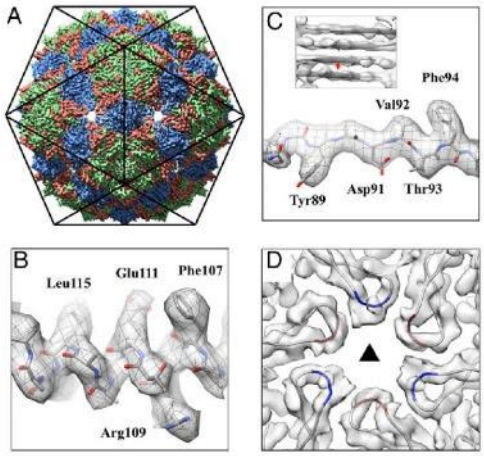


Fig. 1. Icosahedral reconstruction of the Qβ coat protein shell at 3.7-Å resolution. (*A*) The map is oriented with the twofold axis pointing out of the figure. Coat protein conformers A, B, and C are colored salmon, green, and blue, respectively. (*B*) The α-helix density of the coat protein showing α-helical pitches and bulky amino acid side-chains. (*C*) The bulky side-chains in the β-strand. (*Inset*) Separation of four β-strands in a β-sheet formed within coat protein dimers, with the red arrow indicating the location and viewing angle for the *C*. (*D*) The cryo-EM densities and models of the FG-loops, located at the threefold axis, with residues 76–79 for conformers A and C colored red and blue, respectively. The black triangle indicates the threefold axis.

Asymmetric Structure of Phage Qβ. To determine the structures of the A_2 protein and gRNA within the Qβ virion, we released the symmetry for our cryo-EM reconstruction. After 3D classification (*Experimental Procedures*), 12,975 particles were used to generate a symmetry-free reconstruction of Qβ at 7-Å resolution, with A_2 and the gRNA clearly defined. The resolution for the asymmetric reconstruction of the coat proteins and A_2 was further improved to 6.5 Å after masking out the RNA during map refinement (*SI Appendix*, Fig. S3 and S6). Fig. 2 and Movie S1 show the asymmetric cryo-EM structure of Qβ. The overall capsid stays similar to the icosahedral reconstruction of the coat proteins. The read-through domain of the A_1 coat protein is not visible in our asymmetric reconstruction, possibly because of the flexible connector to the coat domain and/or the heterogeneity in the locations of A_1 proteins among different virions. The most apparent difference in the capsid between our symmetric and asymmetric reconstructions is for A_2 (colored hot pink in Fig. 2). A_2 replaces one of the C/C dimers at a twofold axis (Fig. 2*A*), which leaves 178 copies of the coat protein in the virion. Much of the gRNA is clearly resolved, with major and minor grooves of the RNA helices discernible (*SI Appendix*, Fig. S7).

Structure of the Maturation Protein A_2. The Qβ maturation protein, A_2, consists of two structural regions positioned at an angle of 120° (Fig. 2). The top structural region, named the β-region, consists mostly of sheet-like densities from β-strands, whereas the bottom segment, termed the α-region, is mostly composed of rod-like densities from α-helices. Five α-helices were identified in the α-region and labeled α1 to α5, from the longest to shortest helix lengths (Fig. 2*B*), with α1 being ∼60 amino acids. Sequence-based secondary structure analysis predicted several α-helices in A_2 (*SI Appendix*, Fig. S8). Remarkably, one predicted α-helix (from residues 130–190) has a high confidence score and is consistent in length with α1. The α-region connects the gRNA through the capsid, the β-region, in contrast, is exposed on the surface of the phage with a tilt angle of 30° from the capsid shell (Fig. 2*C*). The dimers around A_2 are numbered clockwise from 1 to 10 (Fig. 2*D*). The top halves of helices α1 and α2 interact with two A/B dimers (labeled 1 and 10) of the coat proteins. The bottom halves of helices α1 and α2, along with α3–α5, insert into the gRNA. Although most of the A_2–gRNA interactions occur at the bottom of the α-helix bundle, one RNA stem-loop, labeled with an asterisk in Fig. 2, leans toward the top of the helix α5, forming an interaction between RNA with the middle of A_2 (seen as the black dashed oval in Fig. 2*B*). Another RNA

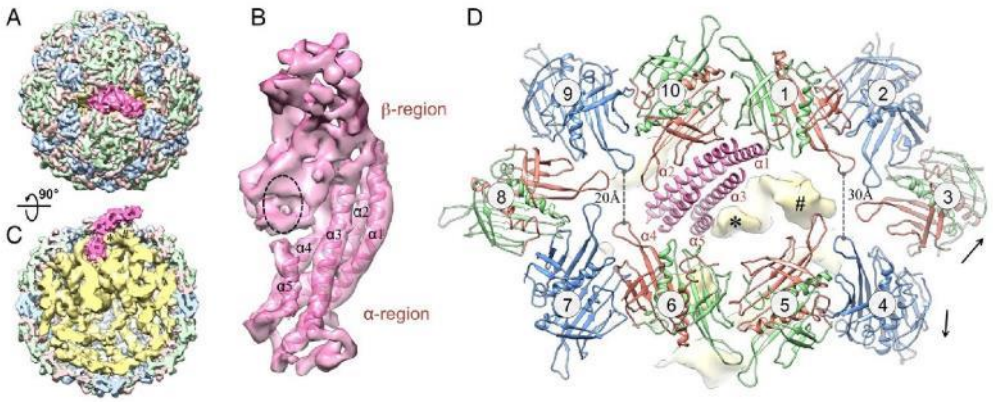


Fig. 2. Asymmetric structure of the Qβ virus at 7-Å resolution. (*A*) Top view of Qβ with coat protein conformers A, B, and C colored light red, light green, and light blue, respectively. The maturation protein A_2 is hot pink. (*B*) Segmented density of A_2 showing the top β-region and bottom α-region. Five α-helices were modeled into A_2. They are labeled from α1 to α5 according to their lengths, with the longest as α1. Black dashed oval indicates the location where one RNA stem-loop contacts. (*C*) Side view of Qβ rotated 90° from the top view, with A_2 pointing up. The virion is sliced down the middle to reveal the interaction of A_2 with the RNA (yellow). The RNA density is low-pass–filtered to 10-Å resolution. Densities of two RNA stem-loops, which are close to A_2, are labeled with star and pound signs. (*D*) Zoom-in top view around A_2 with the hot pink helix models shown for the α-helix bundle of α1 to α5. Ten coat protein dimers around A_2 are labeled from 1 to 10 clockwise and colored salmon, green, and blue for conformers A, B, and C, respectively. Black arrows label the disruption of the interface between dimers 3 and 4. The yellow densities of two RNA stem-loops are labeled with star and pound signs as in *C*.

stem-loop, labeled with a pound sign in Fig. 2, displaces the FG-loop on conformer A of dimer 5 (*SI Appendix*, Fig. S9 and Movie S2). The presence of A_2 and the gRNA forces the coat proteins to deviate from the perfect icosahedral arrangement seen in our symmetric reconstruction (*SI Appendix*, Fig. S9 and Movie S3), especially for the coat proteins around A_2. Notably, the deviations are larger for dimers 1–5 compared with dimers 6–10. The distance between the protein backbones of opposite FG-loops increases from 20 Å for dimers 6–10 to 30 Å for dimers 1–5. The break in symmetry around A_2 causes a fissure between dimers 3 and 4, disrupting the interactions that hold them together (Fig. 2*D*, *SI Appendix*, Fig. S9, and Movies S2 and S3), leaving more of an opening on the right toward which the β-region of A_2 is angled. These disrupted coat protein interactions may allow for an easier detachment of A_2, along with the gRNA, from the capsid upon infection. Furthermore, although A_2 interacts with neighboring coat proteins, it does not make up for all of the interactions that would have taken place with the two coat proteins for which it has substituted. As a result, the weakened coat protein interactions around A_2 may cause less obstruction during release of the gRNA from the capsid.

Organization of the gRNA. In the asymmetric structure, the gRNA exhibits well-defined secondary and tertiary structures within the ~30-Å-thick coat protein shell. Shown in Fig. 3 is a cutaway view of the asymmetric cryo-EM density map of Qβ oriented with A_2 at the "North Pole" and its β-region, pointing out of the figure. The RNA elements (yellow density in Fig. 3*A*) show clear RNA helices and junctions. Only one shell of RNA lies beneath the capsid, with RNA further inside the virion showing distinct asymmetric tertiary structures. Several long RNA helices are observed to traverse the interior of the capsid (Movie S1). One striking feature is a 200-Å-long RNA helix (density fitted with a red RNA helix model in Fig. 3*A*) spanning the middle of the capsid close to its equator. One end of this long helix touches the inner surface of an A/B dimer that is slightly above the equator (marked with a purple oval in Fig. 3 *A* and *B*); the other end of the long helix touches the inner surface of a C/C dimer on the equator at the opposite side (marked with an orange oval in Fig. 3 *A* and *C*). In our asymmetric reconstruction, ~60% of the RNA density is located above the equator, and ~40% below. South of the equator, the local resolution of the RNA is lower compared with the RNA in contact with the capsid or proximal to A_2 (*SI Appendix*, Fig. S6). This indicates that the RNA is more flexible in the interior of the southern hemisphere, possibly because of it being less packed in this region.

Interactions Between the Capsid and gRNA. The EF-loop of each coat protein thrusts inside the capsid, toward the RNA (Fig. 4*A* and *SI Appendix*, Fig. S5). Close inspection of the coat protein sequence of Qβ reveals that there are several positively charged residues at the tip of the EF-loop: residues 56–60 with the amino acid sequence SRNRK. These positively charged side-chains of Arginine and Lysine, at the tip of its EF-loop, may help anchor the coat proteins of Qβ to the negatively charged gRNA backbone during assembly. Moreover, it was previously proposed that, in many ssRNA viruses, the capsid proteins consist of highly basic semiflexible peptide arms, which provide nonspecific electrostatic interactions to control both the length and conformations of the gRNA (31).

We highlighted the tip of the EF-loops as spheres on our ribbon model of the coat proteins with colors salmon, green, and blue for conformers A, B, and C, respectively (Fig. 4*A*). Fig. 4*B* shows the density of gRNA and A_2 with the locations of EF-loops labeled as dots in corresponding colors. After defining A_2 as the North Pole, similar to that of Earth, we can unwrap the outmost spherical shell of the gRNA such that it would be like a map of the world. This is referred to as a Mercator projection, where the poles of the globe appear much larger in the projection than parts at the equator (Fig. 4*C*). What is readily seen is that each colored dot, representing EF-loops from the different-colored coat protein conformers, is immediately proximal to RNA, with most of the EF-loops appearing to be within the grooves of the RNA helices or at the tips of stem-loops. Interestingly, at some of the fivefold vertices, surrounded by the salmon ellipsoids (EF-loop of conformer A), the RNA also adopts a pseudo fivefold arrangement

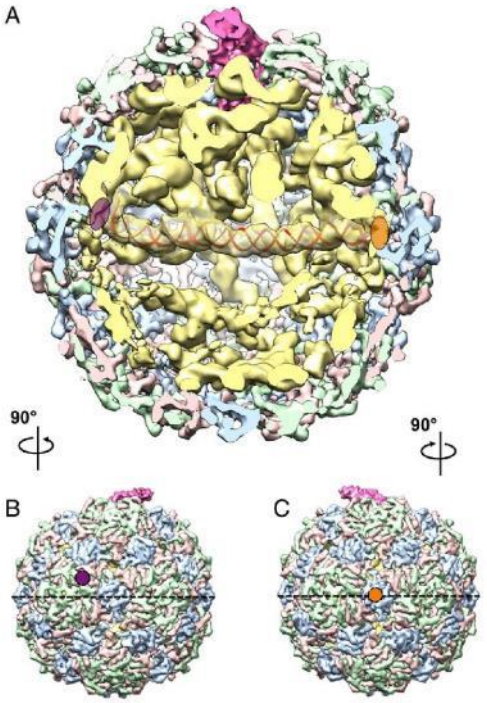


Fig. 3. gRNA organization inside the Qβ capsid. (*A*) Cutaway front view of Qβ with A_2 in the North Pole and pointing out of the paper. Proteins and RNA are colored as in Fig. 2. The density of a long RNA helix, ~200 Å in length, is docked with a red RNA duplex model and spans the two ends of the equator. Purple and orange ovals indicate the left and right of the helix, respectively. (*B*) Left outside view of the Qβ map, with the purple circle annotating the A/B dimer at the left end of the long helix in *A*. (*C*) Right outside view of the Qβ map, with the orange circle annotating the C/C dimer at the right end of the long helix in *A*. The black dashed lines, in *B* and *C*, label the equator.

(Fig. 4*C*), which provides evidence that the structure of the outmost shell of the RNA depends on steric contributions from the capsid inner surface topography (32).

The interactions between coat proteins and gRNA in Qβ were first seen through a crystal structure of the complex between the operator and a coat protein dimer (24). We can fit this crystal structure (PDB ID code 4L8H) as a template into our asymmetric density map to look for interactions that are similar to the one between operator and coat protein dimers. To our surprise, 31 of the 89 coat protein dimers interact with operator-like hairpins, suggesting they may share a similar mode of interaction with high binding affinity. The operator-like hairpins interact with both types of dimer, A/B and C/C (Fig. 5 *A* and *B*), with a preference of A/B dimers (23 of the 31 dimers). When visualized over the whole capsid in a Mercator projection, the operator-like hairpin-coat protein interactions are asymmetrically distributed within the capsid, mostly between longitudes 90° and 180° in the northern hemisphere, close to A_2 (Fig. 5 *C* and *D*). The coat protein dimers also interact with the gRNA in other locations, such as in the middle of RNA helices or on RNA junctions (*SI Appendix*, Fig. S10).

Discussion

Role of A_2 as the Maturation and Lysis Protein. The results presented here may have implications for the mechanism of host infection and of gRNA packaging in ssRNA phages. The Qβ gRNA is surrounded by a coat protein shell, whose symmetry is broken by the maturation protein, A_2. A_2 replaces a C/C dimer with its β-region tilting at an angle of ~30° from the virus surface, just barely rising above the capsid. As in all *Leviviridae*, the maturation protein A_2 is required for infection. The β-region of A_2, outside the capsid, may carry out the role of binding to the receptor, the F-pilus, whereas the α-region binds to the gRNA and hangs on until the RNA is inside the cell. In Qβ and Qβ-like phages, the portion of A_2 outside the capsid has the expanded role of causing lysis. At the time of lysis, it has been shown that A_2, either free in the cytosol or as a single protein on the virion, can both bind and inhibit MurA (33), a universally conserved enzyme that catalyzes the first committed step in cell wall biosynthesis. From our asymmetric structure of Qβ, the MurA-binding site of A_2 is likely to be exposed and around the β-region.

Structural Implication for gRNA Release. We see that A_2 coordinates denser packing of gRNA around it, locally expanding the capsid and causing a fissure between coat dimers (*SI Appendix*, Fig. S9 and Movie S3). Upon attachment to the F-pilus, the β-region of A_2 may have weakened interactions with the neighboring coat proteins, easing the detachment of A_2-bound gRNA. Upon infection, most of the RNA tertiary interactions, as well as the gRNA–coat interactions, need to be abolished for the gRNA to exit through the narrow slot (~60 × 40 Å), left on the capsid after A_2 is detached. As ssRNA viruses of *Leviviridae* do not have a pressurized capsid, the gRNA cannot be mechanically injected into the host from the capsid. The retractile character of the F-pilus may provide some or all of the driving force to pull the A_2-bound gRNA out of the capsid. Notably, the loose interactions of A_2 with the coat proteins and the disrupted interfaces between coat protein dimers around the exit may facilitate the gRNA release. The fissure between coat protein dimers 3 and 4 may further

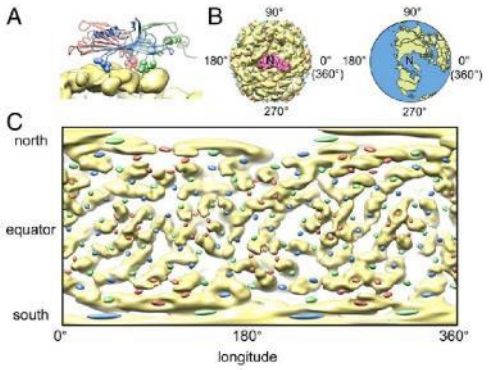


Fig. 4. EF-loops anchor the coat proteins onto the outmost shell of gRNA. (*A*) Interaction between the gRNA and an asymmetric unit of the coat protein with the residues at tip of the EF-loop shown as sphere models. The density of gRNA is in yellow, whereas the models of the coat proteins are colored salmon, green, and blue for conformers A, B, and C, respectively. (*B*, *Left*) A_2 (hot pink) and Qβ gRNA (yellow) are viewed with A_2 as the North Pole (labeled "N"). Tips of the EF-loop are presented as colored dots (with color conventions as in *A*) around the Qβ gRNA globe. (*Right*) Earth in the same orientation. (*C*) Mercator projection of the outmost shell of the Qβ gRNA (yellow) and the locations of EF-loops (salmon, green, and blue ellipsoids).

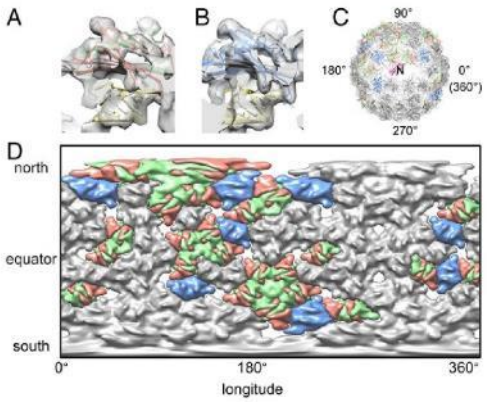


Fig. 5. Interaction between coat protein dimers and the operator-like hairpins of gRNA. (*A*) Operator-like hairpin interacting with an A/B dimer. (*B*) Operator-like hairpin interacting with a C/C dimer. (*C*) Model of the coat protein dimers interacting with operator-like hairpins are in salmon, green, and blue for their conformers A, B, and C, respectively. Dimers not interacting with operator-like hairpins are in gray. The α-helix bundle of A_2 is in hot pink and shown as the North Pole (labeled "N"). (*D*) Mercator projection of the capsid density with coat protein dimers colored as in *C*.

expand during gRNA release and come together when the capsid is emptied. The loose packing in the southern hemisphere of the capsid may enable the gRNA to sample different conformations to unfold during release through the narrow exit.

Structural Mechanism for Capsid Assembly. For Qβ and other ssRNA viruses, there is clearly an evolutionary mandate for efficient and accurate virion assembly (34). After the gRNA is replicated and coat proteins are produced, the rate-limiting step in virion production is encapsidation of the RNA (35). It was first proposed that the assembly pathway was initiated when the first coat protein dimer bound the operator hairpin, serving as a nucleation event for other coat protein dimers to "walk" down a "Hamiltonian path" binding neighboring RNAs until the virion is complete (36). To accommodate this perspective, it was later proposed that there are many sites within the gRNA, termed "packaging signals," that can interact with the coat protein (6). These packaging signals were thought to be related to the consensus sequence of the operator. In our reconstruction of Qβ, we see the outmost layer of the gRNA coming in contact with almost every coat protein, presenting 31 operator-like hairpins for the binding of coat protein dimers. Many of the gRNA secondary structures involve an extensive array of long-range tertiary interactions, which have been shown to play roles in gene regulation and capsid assembly (37). These tertiary interactions may contract the gRNA to a rigid structure in which the operator-like hairpins, which may be far apart in the genome, are located near each other. As seen in our structure, the operator-like hairpins are asymmetrically distributed on the outmost shell of the gRNA, possibly presenting a roadmap to sequentially recruit more coat protein dimers. The recruited coat protein dimers strengthen their interactions with neighboring coat protein dimers, which in turn stabilize the gRNA underneath. Such a mutual chaperoning of both RNA and coat protein conformations (38) may facilitate the fast and accurate assembly of these viruses in vivo.

Experimental Procedures

Sample and Grid Preparation. Phage Qβ was purified as previously described (9). The final protein concentration of Qβ was 5 mg/mL, ~10^{12} plaque-forming units. For imaging, 3 μL of the sample solution was first applied to a C-Flat 1.2/1.3 400 mesh holey carbon grid at 20 °C with 100% relative humidity and vitrified using a Vitrobot Mark III (FEI Company).

Cryo-EM Imaging. The thin-ice areas that showed clearly visible and monodispersed particles were imaged under a JEM3200FSC cryo-electron microscope with a field emission gun (JEOL) operated at 300 kV. Data were collected using the manual mode of SerialEM (39) on a Gatan K2 Summit direct detection camera (Gatan) in the superresolution electron counting mode. A nominal magnification of 30,000× was used, yielding a subpixel size of 0.6 Å. The beam intensity was adjusted to a dose rate of 7.2 e^-/pixel/s on the camera. A 50-frame movie stack was recorded for each picture, with 0.2 s per frame, for a total exposure time of 10 s. Each frame in the movie stack had an exposure dose of 1 $e^-/Å^2$ per frame. An in-column energy filter was used with a slit width of 29 eV.

Data Preprocessing. A total of 712 superresolution movie stacks were collected and first binned by 2 to yield a pixel size of 1.2 Å. Using Unblur (40), these stacks were aligned, filtered according to electron dose, and summed to generate two sets of sum images, with one set from frames 1–30 and the other set from frames 2–12, respectively.

These sum images were visually screened, and 582 of each set with strong power spectra were selected for further processing. Contrast transfer functions of the micrographs were estimated using CTFFind4 (41). Batchboxer in EMAN (42) was first used to automatically pick all the particles from these 582 sum images with a box size of 320 × 320 $pixel^2$, yielding 111,507 particles. The automatically picked particles were then shrunk by eight times to a pixel size of 9.6 Å and screened for high-contrast particles for three rounds of the reference-free 2D classification in Relion (43), with each round having 18, 16, and 25 iterations, respectively. The coordinates of the 53,999 screened particles were then imported into the e2boxer.py in EMAN2 (44) to manually screen for missed false positives, leaving only 51,815 particles.

Icosahedral Refinement for the Coat Proteins. The 51,815 clean particles were then refined in Relion, with the icosahedral symmetry applied. Initial refinement was run using the sum images from frame 1–30 (total dose of 31 $e^-/Å^2$ with dose filter on in Unblur). The starting map was generated from EMAN2 and low-pass-filtered to 60-Å resolution. The refinement first yielded a map at 4.0-Å resolution. This was followed by a continuing map refinement with sum images of only frame 2–12 (total dose of 11 $e^-/Å^2$ with dose filter on) and the local contrast transfer functions for each particle reestimated using Gctf (45). The final refined icosahedral map yielded a resolution of 3.7 Å.

Asymmetric Refinement for the Entire Virion. The same 51,815 clean particles were binned by 2 to a pixel size of 2.4 Å for the asymmetric reconstruction of the entire virion. The icosahedral map of the coat proteins was low-pass-filtered to 60 Å as an initial model. Unsupervised 3D classification in Relion was first performed with C1 symmetry, and 12,975 particles went into a class, which showed clear density of the gRNA inside the capsid. These 12,975 particles were further refined in Relion with C1 symmetry to yield a map at 7.0-Å resolution, where the protein secondary structures were clearly visible. To further improve the resolution of the coat proteins and A_2, a mask was applied to focus the refinement on the outer shell of the particle, which yielded a resolution of 6.5 Å. Further refinement of the asymmetric structure using all of the 51,815 particles did not improve the resolvability of the map, potentially because the 12,975 particles, selected from the 3D classification in Relion, have the best contrast and are most homogeneous in conformations.

Resolution Estimation. The overall resolutions of all these reconstructed maps were assessed using the gold-standard criterion of Fourier Shell Correlation, with a cutoff at 0.143, between two half maps from two independent half-sets of data (46). Local resolutions were estimated using Blocres (47).

Modeling, Map Segmentation, and Visualization. To build the model for the coat protein from the icosahedral reconstruction, the crystal structure (PDB ID code 1QBE) was used as the initial model. The missing loop in the crystal structure was built in COOT (48) and then further refined in the real-space refinement in Phenix (49). The refined model showed good model geometry from Molprobity (50) (*SI Appendix*, Table S1). The Fourier Shell Correlation between the icosahedral map and the refined PDB model agreed at 3.9-Å resolution at 0.5 threshold (*SI Appendix*, Fig. S3). This more complete model for the coat protein was then used to fit into the asymmetric map of the Qβ virion to guide the segmentation of the gRNA and A_2 protein with Segger (51) as a plugin in University of California, San Francisco (UCSF) Chimera (52).

The α-helices of the A_2 protein were identified in Helixhunter (53). Secondary structure prediction of A_2 was performed with I-TASSER (54). The 200-Å long RNA helix that spanned the equator of the gRNA was first built in UCSF Chimera with a straight A-form RNA helix with 70 GC base pairs and then refined in Molecular Dynamics Flexible Fitting to improve the fit to the density (55). All of the figures and movies were made in UCSF chimera and VMD (56). Unwrapping of the gRNA into Mercator projection was done in EMAN2 with the tool "e2unwrap3d.py" in two steps for the longitude and latitude, respectively.

ACKNOWLEDGMENTS. We thank Jeng-Yih Chang for performing the molecular dynamics flexible fitting of the long RNA helix model. We thank the Microscopy and Imaging Center at Texas A&M University for providing instrumentation for the initial screening of the phage particles; the Texas A&M High Performance Research Computing Center for providing the computational resources for the data processing; and the National Center for Macromolecular Imaging (NCMI) at Baylor College of Medicine for data collection. NCMI is supported by NIH Grants P41GM103832 and U24GM116787. J.Z. is supported by start-up funding from the Department of Biochemistry and Biophysics at Texas A&M University. R.Y. is supported by Public Health Service Grant GM27099. J.Z. and R.Y. acknowledge the support of the Center for Phage Technology, jointly sponsored by Texas AgriLife and Texas A&M University. K.V.G. is partly supported by the Welch Foundation Grant A-1863 (to J.Z.).

1. Koonin EV, Dolja VV, Krupovic M (2015) Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virology* 479-480:2–25.
2. Heil F, et al. (2004) Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science* 303(5663):1526–1529.
3. Bidet K, Garcia-Blanco MA (2014) Flaviviral RNAs: Weapons and targets in the war between virus and host. *Biochem J* 462(2):215–230.
4. Kolakofsky D (2015) A short biased history of RNA viruses. *RNA* 21(4):667–669.
5. Valentine RC, Strand M (1965) Complexes of F-Pili and RNA bacteriophage. *Science* 148(3669):511–513.
6. Rolfsson Ó, et al. (2016) Direct evidence for packaging signal-mediated assembly of bacteriophage MS2. *J Mol Biol* 428(2 Pt B):431–448.
7. Reed CA, Langlais C, Wang IN, Young R (2013) A_2 expression and assembly regulates lysis in Qβ infections. *Microbiology* 159(Pt 3):507–514.
8. Ling CM, Hung PP, Overby LR (1970) Independent assembly of Qβ and MS2 phages in doubly infected *Escherichia coli*. *Virology* 40(4):920–929.
9. Strauss JH, Jr, Sinsheimer RL (1963) Purification and properties of bacteriophage MS2 and of its ribonucleic acid. *J Mol Biol* 7:43–54.
10. Reed CA, Langlais C, Kuznetsov V, Young R (2012) Inhibitory mechanism of the Qβ lysis protein A_2. *Mol Microbiol* 86(4):836–844.
11. Weber K, Konigsberg W (1975) Proteins of the RNA phages. *RNA Phages*, ed Zinder ND (Cold Spring Harbor Lab Press, Cold Spring Harbor, New York), pp 51–84.
12. Hofstetter H, Monstein HJ, Weissmann C (1974) The readthrough protein A_1 is essential for the formation of viable Qβ particles. *Biochim Biophys Acta* 374(2):238–251.
13. Rumnieks J, Tars K (2011) Crystal structure of the read-through domain from bacteriophage Qβ A_1 protein. *Protein Sci* 20(10):1707–1712.
14. Lander GC, et al. (2006) The structure of an infectious P22 virion shows the signal for headful DNA packaging. *Science* 312(5781):1791–1795.
15. Jiang W, et al. (2006) Structure of epsilon15 bacteriophage reveals genome organization and DNA packaging/injection apparatus. *Nature* 439(7076):612–616.
16. Chen DH, et al. (2011) Structural basis for scaffolding-mediated assembly and maturation of a dsDNA virus. *Proc Natl Acad Sci USA* 108(4):1355–1360.
17. Yap ML, et al. (2016) Role of bacteriophage T4 baseplate in regulating assembly and infection. *Proc Natl Acad Sci USA* 113(10):2654–2659.
18. Stockley PG, et al. (2016) Bacteriophage MS2 genomic RNA encodes an assembly instruction manual for its capsid. *Bacteriophage* 6(1):e1157666.
19. McKnight KL, Heinz BA (2003) RNA as a target for developing antivirals. *Antivir Chem Chemother* 14(2):61–73.
20. Golmohammadi R, Fridborg K, Bundule M, Valegård K, Liljas L (1996) The crystal structure of bacteriophage Qβ at 3.5 Å resolution. *Structure* 4(5):543–554.
21. Golmohammadi R, Valegård K, Fridborg K, Liljas L (1993) The refined structure of bacteriophage MS2 at 2.8 A resolution. *J Mol Biol* 234(3):620–639.
22. Beekwilder J, Nieuwenhuizen R, Poot R, van Duin J (1996) Secondary structure model for the first three domains of Qβ RNA. Control of A-protein synthesis. *J Mol Biol* 256(1):8–19.
23. Beekwilder MJ, Nieuwenhuizen R, van Duin J (1995) Secondary structure model for the last two domains of single-stranded RNA phage Qβ. *J Mol Biol* 247(5):903–917.
24. Rumnieks J, Tars K (2014) Crystal structure of the bacteriophage Qβ coat protein in complex with the RNA operator of the replicase gene. *J Mol Biol* 426(5):1039–1049.
25. Valegård K, Murray JB, Stockley PG, Stonehouse NJ, Liljas L (1994) Crystal structure of an RNA bacteriophage coat protein-operator complex. *Nature* 371(6498):623–626.
26. Shiba T, Suzuki Y (1981) Localization of A protein in the RNA-A protein complex of RNA phage MS2. *Biochim Biophys Acta* 654(2):249–255.
27. Koning R, et al. (2003) Visualization by cryo-electron microscopy of genomic RNA that binds to the protein capsid inside bacteriophage MS2. *J Mol Biol* 332(2):415–422.
28. Toropova K, Basnak G, Twarock R, Stockley PG, Ranson NA (2008) The three-dimensional structure of genomic RNA in bacteriophage MS2: Implications for assembly. *J Mol Biol* 375(3):824–836.
29. Dent KC, et al. (2013) The asymmetric structure of an icosahedral virus bound to its receptor suggests a mechanism for genome release. *Structure* 21(7):1225–1234.
30. Li X, et al. (2013) Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods* 10(6):584–590.
31. Belyi VA, Muthukumar M (2006) Electrostatic origin of the genome packing in viruses. *Proc Natl Acad Sci USA* 103(46):17174–17178.
32. ElSawy KM, Caves LS, Twarock R (2011) On the origin of order in the genome organization of ssRNA viruses. *Biophys J* 101(4):774–780.
33. Bernhardt TG, Wang IN, Struck DK, Young R (2001) A protein antibiotic in the phage Qbeta virion: Diversity in lysis targets. *Science* 292(5525):2326–2329.
34. Dykeman EC, Stockley PG, Twarock R (2014) Solving a Levinthal's paradox for virus assembly identifies a unique antiviral strategy. *Proc Natl Acad Sci USA* 111(14):5361–5366.
35. Heldt FS, Frensing T, Reichl U (2012) Modeling the intracellular dynamics of influenza virus replication to understand the control of viral RNA synthesis. *J Virol* 86(15):7806–7817.
36. Dykeman EC, et al. (2011) Simple rules for efficient assembly predict the layout of a packaged viral RNA. *J Mol Biol* 408(3):399–407.
37. Klovins J, Berzins V, van Duin J (1998) A long-range interaction in Qβ RNA that bridges the thousand nucleotides between the M-site and the 3′ end is required for replication. *RNA* 4(8):948–957.
38. Rolfsson O, Toropova K, Ranson NA, Stockley PG (2010) Mutually-induced conformational switching of RNA and coat protein underpins efficient assembly of a viral capsid. *J Mol Biol* 401(2):309–322.
39. Mastronarde DN (2005) Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* 152(1):36–51.
40. Grant T, Grigorieff N (2015) Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *eLife* 4:e06980.
41. Rohou A, Grigorieff N (2015) CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* 192(2):216–221.
42. Ludtke SJ, Baldwin PR, Chiu W (1999) EMAN: Semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* 128(1):82–97.
43. Scheres SH (2012) RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 180(3):519–530.
44. Tang G, et al. (2007) EMAN2: An extensible image processing suite for electron microscopy. *J Struct Biol* 157(1):38–46.
45. Zhang K (2016) Gctf: Real-time CTF determination and correction. *J Struct Biol* 193(1):1–12.
46. Scheres SH, Chen S (2012) Prevention of overfitting in cryo-EM structure determination. *Nat Methods* 9(9):853–854.
47. Heymann JB, Belnap DM (2007) Bsoft: Image processing and molecular modeling for electron microscopy. *J Struct Biol* 157(1):3–18.
48. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
49. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
50. Chen VB, et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12–21.
51. Pintilie G, Zhang J, Chiu W, Gossard D (2009) Identifying Components in 3D Density Maps of Protein Nanomachines by Multi-scale Segmentation. IEEE NIH Life Sci Syst Appl Workshop 2009 2009:44–47.
52. Pettersen EF, et al. (2004) UCSF Chimera–a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
53. Jiang W, Baker ML, Ludtke SJ, Chiu W (2001) Bridging the information gap: Computational tools for intermediate resolution structure interpretation. *J Mol Biol* 308(5):1033–1044.
54. Yang J, et al. (2015) The I-TASSER suite: Protein structure and function prediction. *Nat Methods* 12(1):7–8.
55. Trabuco LG, Villa E, Schreiner E, Harrison CB, Schulten K (2009) Molecular dynamics flexible fitting: A practical guide to combine cryo-electron microscopy and X-ray crystallography. *Methods* 49(2):174–180.
56. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph* 14(1):27–38.

SI Appendix:

Asymmetric cryo-EM structure of the canonical Allolevivirus Qβ reveals a single maturation protein and the genomic ssRNA *in situ*

Karl V. Gorzelnik[a,1], Zhicheng Cui[a,1], Catrina A. Reed[a,2], Joanita Jakana[b], Ry Young[a], and Junjie Zhang[a,3]

[a] Center for Phage Technology, Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843

[b] National Center for Macromolecular Imaging, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030

Movie Legends

Movie S1. Asymmetric structure of Qβ. Coat proteins are in salmon (conformer A), green (conformer B) and blue (conformer C), respectively. A_2 is in hot pink. RNA is in yellow and low-pass filtered to 10Å resolution.

Movie S2. Symmetry breaking in the coat proteins around A_2.

Movie S3. The deviations of the coat protein dimers in the asymmetric structure from their locations in the icosahedral structure.

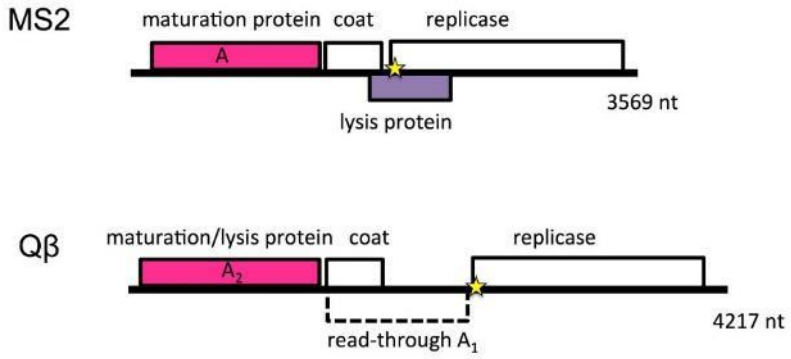


Figure S1. Genomic maps of *Leviviridae* MS2 (top; Levivirus) and Qβ (bottom; Allolevivirus). The maturation proteins are highlighted in hot pink. The lysis protein of MS2 is highlighted in purple. A_1 of Qβ is labeled as dashed box. Yellow stars indicate the operators near the start of their respective replicase genes.

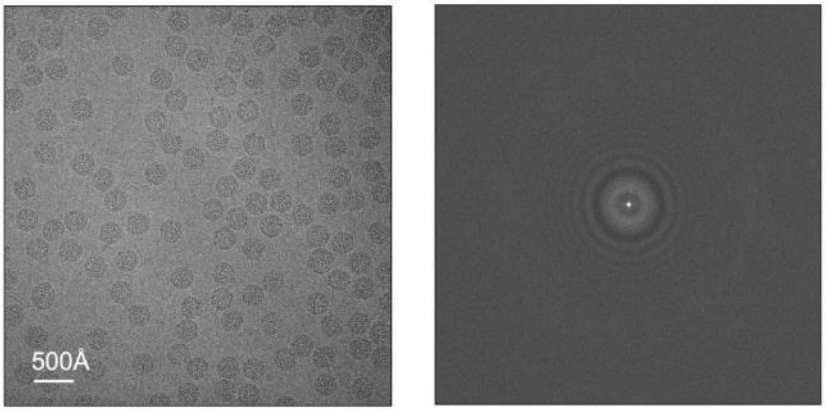


Figure S2. Drift-corrected micrograph of Qβ (left) and its power spectrum (right).

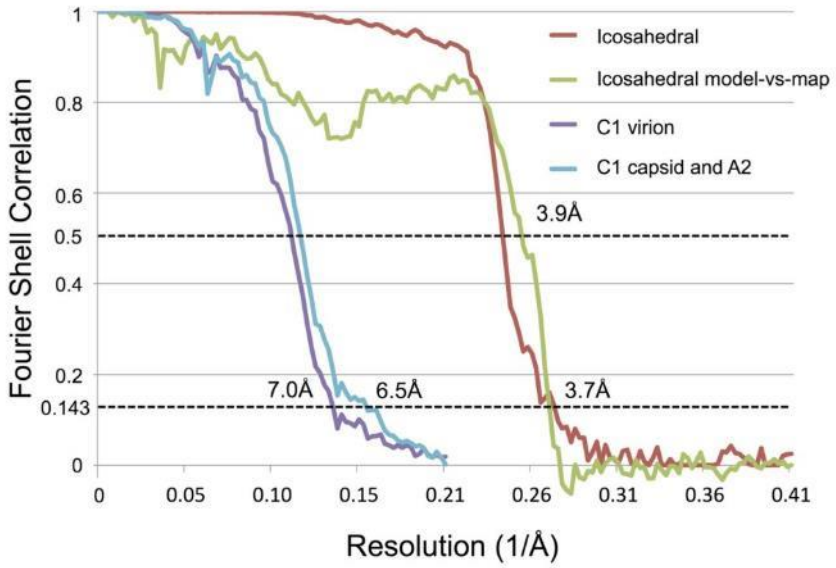


Figure S3. Gold-standard Fourier shell correlation (FSC) for the cryo-EM reconstructions of Qβ with and without symmetry and between the icosahedral map and the refined PDB model.

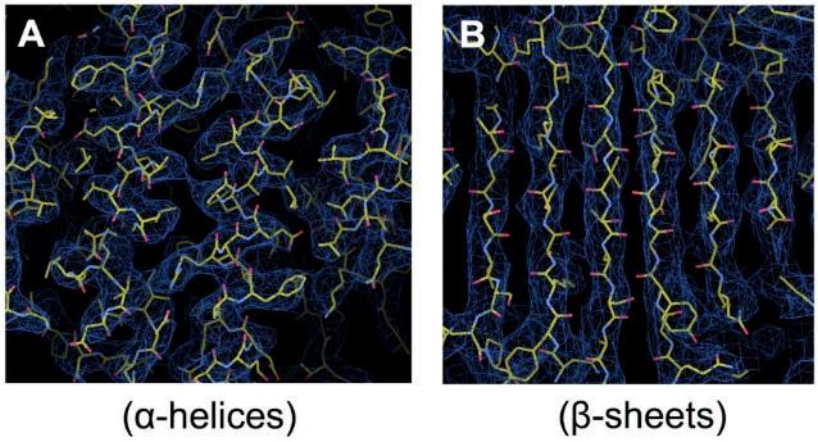


Figure S4. Fitting of the crystal structure (PDB ID 1QBE) into our icosahedral cryo-EM density map at 3.7Å resolution.

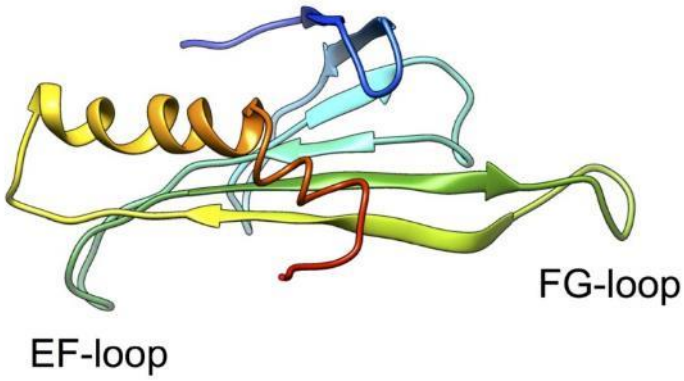


Figure S5. PDB model of Qβ coat protein showing the location of EF-loop and FG-loop. The protein backbone is colored blue to red from N-terminus to C-terminus.

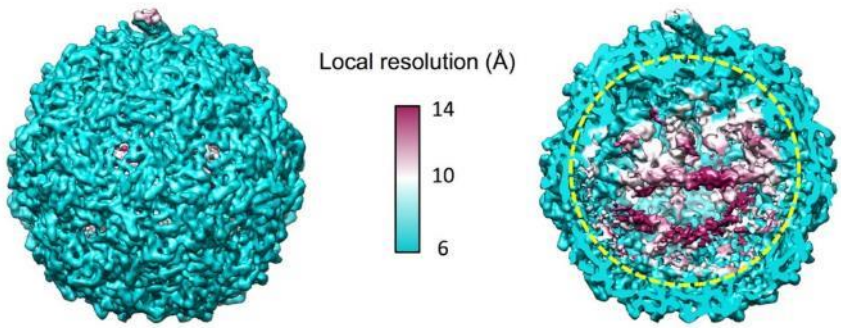


Figure S6. Local resolution of the asymmetric reconstruction of Qβ from outside (left) and cutaway (right) views. The yellow dashed circle encloses the gRNA densities.

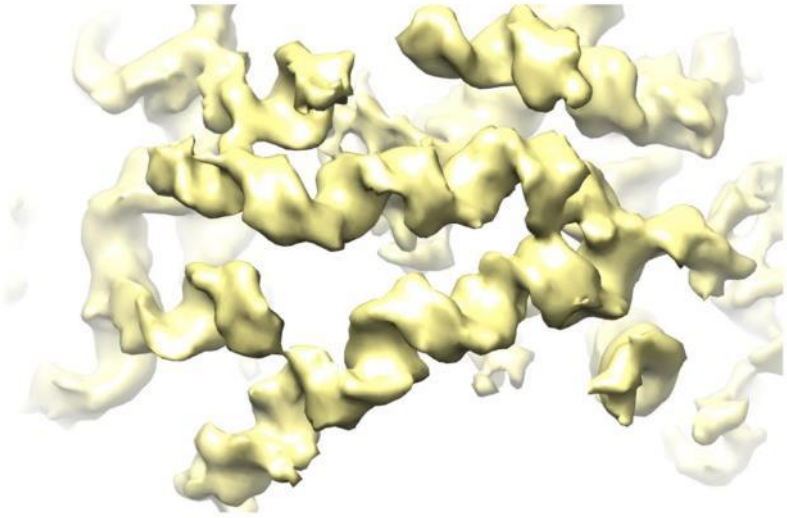


Figure S7. Segmented density map of genomic RNA showing major and minor grooves.

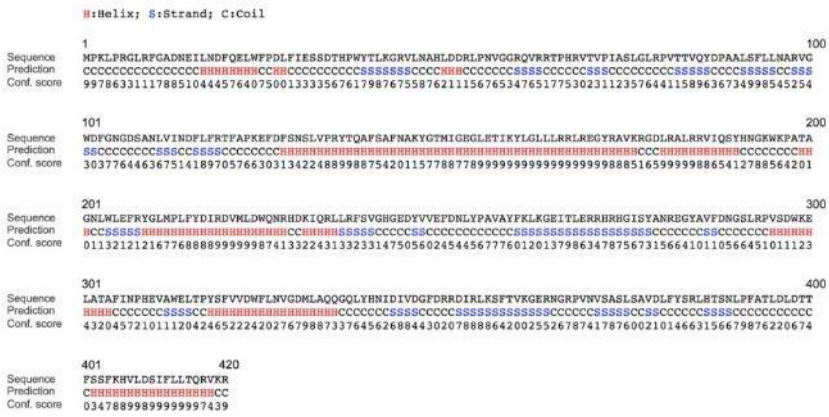


Figure S8. Secondary structure prediction for the A_2 protein from I-TASSER, with the first line as peptide sequence, second line as predicted secondary structures and the third line as the confidence score of the predicted secondary structures. The larger the confidence score, the higher the probability of the predicted secondary structures.

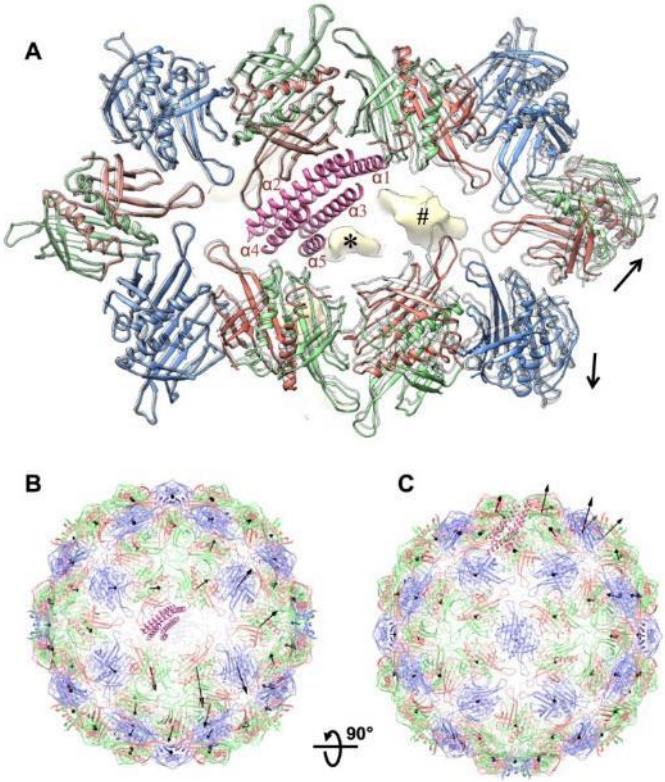


Figure S9. Deviation of the coat protein dimers' position in the asymmetric structure from their positions in the icosahedral reconstruction. (A) Zoom-in top view of the capsid proteins in the asymmetric structures (salmon, green and blue ribbon models) overlaid with their positions in the icosahedral structure (grey ribbon models). The α-region of A_2 is shown as a five-helix bundle (α1-5, hot pink ribbon models). Two RNA stem-loops (yellow density) are labeled with the star and pound symbols, respectively. Black arrows label the disruption of the interface between two dimers. (B) Zoom-out top view of asymmetric Qβ capsid with the black arrows indicating the direction and amount of deviation for each coat protein dimer from its position in the icosahedral structure. The length of each arrow equals 10 times the deviation between the corresponding dimer's centers-of-mass in the icosahedral and asymmetric capsids. (C) The side view with viewing angle rotated 90° from the top view.

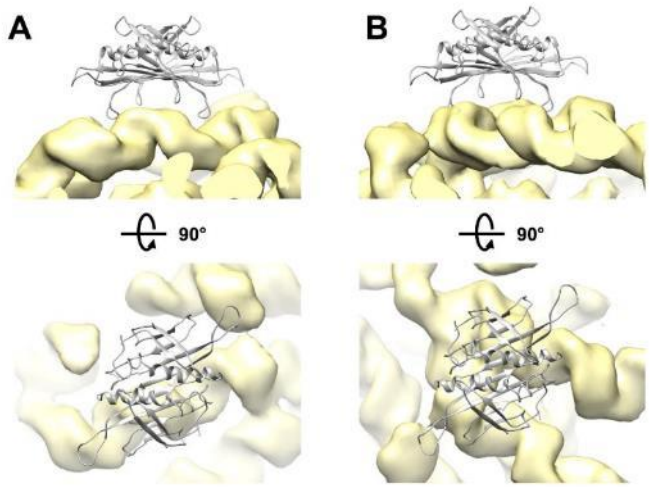


Figure S10. Examples of coat protein dimers (grey ribbon models) interacting with gRNA (yellow densities), which do not have an operator-like hairpin fold. (A) Side view and top view of one coat protein dimer interacting with the middle of an RNA helix. (B) Side view and top view of one coat protein dimer interacting with an RNA junction.

Table S1. Molprobity model statistics for the coat proteins built from the icosahedral map.

All-atom Clashscore	10.42
Poor rotamers	0.00%
Favored rotamers	98.06%
Ramachandran outliers	0.00%
Ramachandran favored	83.08%
MolProbity score	2.23
Cβ deviations >0.25Å	0.00%
Bad bonds:	0.00%
Bad angles:	0.07%

APPENDIX B

STRUCTURES OF QB VIRIONS, VIRUS-LIKE PARTICLES, AND THE QB–MURA COMPLEX REVEAL INTERNAL COAT PROTEINS AND THE MECHANISM OF HOST LYSIS*

Structures of Qβ virions, virus-like particles, and the Qβ–MurA complex reveal internal coat proteins and the mechanism of host lysis

Zhicheng Cui[a], Karl V. Gorzelnik[a], Jeng-Yih Chang[a], Carrie Langlais[a,1], Joanita Jakana[b], Ry Young[a], and Junjie Zhang[a,2]

[a]Department of Biochemistry and Biophysics, Center for Phage Technology, Texas A&M University, College Station, TX 77843; and [b]National Center for Macromolecular Imaging, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030

Edited by Joseph D. Puglisi, Stanford University School of Medicine, Stanford, CA, and approved September 25, 2017 (received for review May 2, 2017)

In single-stranded RNA bacteriophages (ssRNA phages) a single copy of the maturation protein binds the genomic RNA (gRNA) and is required for attachment of the phage to the host pilus. For the canonical *Allolevivirus* Qβ the maturation protein, A_2, has an additional role as the lysis protein, by its ability to bind and inhibit MurA, which is involved in peptidoglycan biosynthesis. Here, we determined structures of Qβ virions, virus-like particles, and the Qβ–MurA complex using single-particle cryoelectron microscopy, at 4.7-Å, 3.3-Å, and 6.1-Å resolutions, respectively. We identified the outer surface of the β-region in A_2 as the MurA-binding interface. Moreover, the pattern of MurA mutations that block Qβ lysis and the conformational changes of MurA that facilitate A_2 binding were found to be due to the intimate fit between A_2 and the region encompassing the closed catalytic cleft of substrate-liganded MurA. Additionally, by comparing the Qβ virion with Qβ virus-like particles that lack a maturation protein, we observed a structural rearrangement in the capsid coat proteins that is required to package the viral gRNA in its dominant conformation. Unexpectedly, we found a coat protein dimer sequestered in the interior of the virion. This coat protein dimer binds to the gRNA and interacts with the buried α-region of A_2, suggesting that it is sequestered during the early stage of capsid formation to promote the gRNA condensation required for genome packaging. These internalized coat proteins are the most asymmetrically arranged major capsid proteins yet observed in virus structures.

single-particle cryo-EM | ssRNA virus | genome packaging | maturation protein | lysis protein

Single-stranded RNA (ssRNA) viruses infect all domains of life and are important pathogens (1–4). Bacteriophages or phages, especially the canonical *Leviviridae* Qβ and MS2, have been model systems for studying ssRNA viral gene regulation (5–8), genome replication (9, 10), and encapsidation (11–14). They have a near-icosahedral capsid, encapsidating genomic RNAs (gRNAs) of ~3,500–4,300 nucleotides. Recent advances in single-particle cryoelectron microscopy (cryo-EM) have enabled characterization of asymmetric features in ssRNA phages. Of note, the gRNA has been found to adopt a single, well-folded, 3D structure within the virion (15–18).

The 4,217-nucleotide gRNA of Qβ has genes for the maturation protein (A_2), the coat protein, and the replicase (the viral subunit of the RNA-dependent RNA polymerase) (19–21). The gene for the Qβ coat protein contains a "leaky" stop codon that occasionally results in a readthrough by the host ribosome producing a minor coat protein A_1. A_1 is essential and consists of the full-length coat domain connected by a flexible linker to the readthrough domain, a 196-amino acid C-terminal extension (22). These four proteins have essential functions at the various stages of the Qβ life cycle, from genome replication and packaging to host lysis and the subsequent adsorption to the F pilus of a new host.

Unlike double-stranded DNA phages, which use specialized machinery to pump their genetic material into a preformed capsid, ssRNA phages assemble their capsid proteins around their gRNA. In Qβ, the maturation protein, A_2, along with 89 coat protein dimers, directly binds the gRNA, encapsidating it in a near-icosahedral (T = 3) capsid (15). Around 10 copies of A_1 are incorporated into the capsid, replacing monomers of the coat protein, but the positions of these substitutions and the basis for the requirement for A_1 are unknown (23, 24). The composition of ribonucleoprotein assembly intermediates of the related phage MS2 was estimated to be 1:15:1:1 for the replicase subunit, coat protein, maturation protein, and RNA, including both gRNA and ribosomal RNAs (25). This suggests that the maturation proteins of ssRNA phages bind before the vast majority of coat proteins, as there are ~180 coat proteins in the fully assembled virion. However, the detailed interactions between the components of Qβ were not clear, due to the lack of high-resolution structures of the Qβ virion.

One major difference between Qβ and other ssRNA phages, such as MS2, is that the maturation protein of Qβ, A_2, has an additional role. While MS2 dedicates a unique protein, L, for host lysis (26), Qβ uses the maturation protein A_2 to bind and inhibit MurA (27), the enzyme for the first committed step in bacterial cell wall biosynthesis (28). The inhibition of cell wall biosynthesis leads to lysis at the next cell division, releasing the progeny Qβ virions (29). Therefore, A_2, as well as the mature Qβ virion, has been regarded as a "protein antibiotic" (27). Although it has been shown that in vitro A_2 binds MurA in the presence of its first substrate, uridine 5′-diphosphate-*N*-acetylglucosamine (UDP-GlcNAc) (19), the basis

Significance

Host lysis and virion assembly are essential processes during the infection cycle of single-stranded RNA (ssRNA) viruses. Using single-particle cryoelectron microscopy, we visualized how the ssRNA virus, Qβ, uses its single-molecule "tail protein," A_2, to inhibit MurA, a bacterial enzyme essential for cell wall biosynthesis, leading to lysis of the host cell. We also revealed an extra coat protein dimer, which instead of being a part of the viral capsid, is sequestered within the virion, binding to an RNA hairpin from a five-way junction in the genomic RNA. The same five-way junction also presents hairpins to bind A_2 and other coat protein dimers in the capsid, potentially supporting a nucleation event for virion assembly.

Author contributions: Z.C. and J.Z. designed research; Z.C., K.V.G., J.-Y.C., J.J., and J.Z. performed research; C.L. and R.Y. contributed new reagents/analytic tools; Z.C., J.-Y.C., and J.Z. analyzed data; and Z.C., K.V.G., R.Y., and J.Z. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the EMDataBank (www.emdatabank.org) with accession IDs: EMD-8708 (Qβ virus-like particle), EMD-8709 (Qβ virion), EMD-8711 (Qβ virion bound with MurA), and EMD-8710 (local map including A2, the internalized coat protein dimer, and the interacting gRNA hairpins), and Protein Data Bank (https://www.wwpdb.org) with PDB IDs: 5VLY (capsid for the Qβ virus-like particle), 5VLZ (capsid for the Qβ virion), and 5VM7 (complex between A2 and MurA).

[1]Present address: AmpliPhi Biosciences Corporation, San Diego, CA 92130.

[2]To whom correspondence should be addressed. Email: junjiez@tamu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1707102114/-/DCSupplemental.

of how A_2 inhibits MurA was unknown, again largely due to the lack of structural information.

Here we present cryo-EM structures of Qβ both free and bound with MurA, to determine the structural details of bacterial cell lysis by Qβ. Moreover, by comparing the structures of the Qβ virion and noninfectious Qβ virus-like particles (VLPs), which are copurified with the Qβ virions, we reveal a structural rearrangement in the capsid coat proteins that is required for the assembly of viable phages. These results are discussed in terms of the pathway of Qβ assembly and the genetics of resistance to Qβ propagation.

Results

Qβ Virion at Near-Atomic Resolution. Using single-particle cryo-EM, 2,370 movie stacks of the Qβ phages were recorded on a direct detection camera in the superresolution electron counting mode (30), yielding 248,434 high-quality phage particles. Classification of these particles revealed that ~20% of them contain A_2 in the capsid and have a single, dominant gRNA conformation (*SI Appendix*, Fig. S1). These particles were then used to generate a symmetry-free reconstruction of the mature Qβ virion at an overall resolution of 4.7 Å, an improvement from the 7-Å resolution reconstruction in our previous work (15). The local resolution of the virion is not uniform, with the gRNA in the center of the virion having the lowest resolution, presumably due to the intrinsic flexibility of the RNA helices that are not directly bound to the capsid. The resolution for the asymmetric reconstruction of capsid components, including the coat proteins and A_2, was further improved to 4.4 Å after excluding the gRNA during the map refinement (*SI Appendix*, Fig. S2 *A* and *C*).

The cryo-EM density of the maturation/lysis protein, A_2, matches the recent crystal structure of purified A_2, in the absence of the capsid and gRNA (31). A_2 consists of an α-region and a β-region. It replaces a coat protein dimer at a twofold axis, with its α-region inserted into the capsid, bound to the gRNA, and its β-region exposed outside the virion. By labeling the coat protein dimers surrounding A_2 from 1 to 10 in a clockwise fashion, the β-region of A_2 is seen to lean toward the interface of neighboring coat protein dimers 3 and 4 (Fig. 1*A*). These results are consistent with our previous observations at a lower resolution (15). A cutaway side view (Fig. 1*B*) shows the components of the mature Qβ including A_2, the gRNA, and the coat proteins. Surprisingly, an extra protein density is observed inside the virion close to the bottom of A_2, around the α-region. This density does not resemble any RNA secondary structures and is not proximal to any coat proteins on the capsid. The identity of this density will be revealed in the present study.

The improved resolution of the capsid components in our cryo-EM reconstruction, from 6.5 Å in previous work (15) to 4.4 Å, allowed us to build the entire backbone model of the Qβ protein capsid into the cryo-EM density map and examine the detailed interactions between the A_2 and neighboring coat protein dimers. Four surrounding coat protein dimers (dimers 1, 5, 6, and 10 in Fig. 1 *A* and *C*) present loops to interact with A_2 (Fig. 1*C* and *SI Appendix*, Fig. S3), wherein the tightest interaction occurs from coat protein dimer 10. A zoomed-in view of the interface between coat protein dimer 10 and A_2 reveals that several amino acid residues from coat protein dimer 10 and A_2 carry opposite charges (Arg24, Glu45, and Arg47 on the coat protein and Arg281, Asp343, Asp396, and Asp398 from A_2, Fig. 1*D*), suggesting that salt bridges are involved in the structure.

Comparison of the Qβ Virion and Qβ Virus-Like Particles. The capsids of 70–80% of the particles are perfectly icosahedral and do not have A_2 incorporated into the surface of the capsid. Among these particles, the RNA density within the capsid does not adopt a dominant conformation (*SI Appendix*, Fig. S1). These particles are regarded as noninfectious VLPs. The electron density of the capsids for Qβ VLPs was resolved to 3.3-Å resolution with icosahedral symmetry applied (Fig. 2*A* and *SI Appendix*, Fig. S2*A*). Sidechains for the coat proteins in the Qβ VLP are clearly seen (Fig. 2*B*).

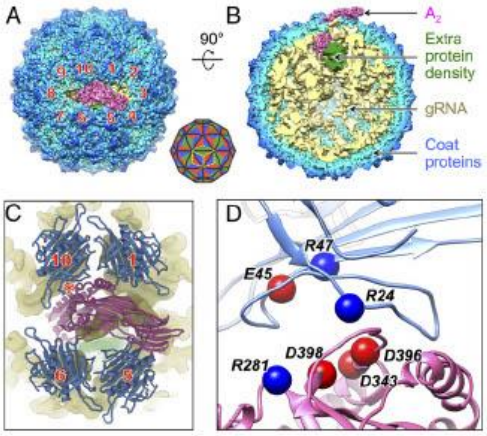


Fig. 1. The structure of the Qβ virion shows interactions between the maturation protein and the capsid coat proteins. (*A*) Near-atomic asymmetric structure of Qβ shows the maturation protein surrounded by 10 coat protein dimers, which are labeled clockwise 1–10. Diagram in the *Lower Right* corner shows the triangulation pattern for T = 3 and the placement of 3 coat protein dimers labeled by three yellow dashed diamond boxes. (*B*) A 90° turn and cutaway view of Qβ shows the coat proteins (radially colored from light blue to blue), the gRNA (yellow, low-pass filtered to 6-Å resolution), A_2 (hot pink), and the electron density of the previously uncharacterized proteins (green). (*C*) Close-up image of A_2 (hot pink) in proximity to the 4 coat protein dimers (blue) neighboring it; the red star labels the tightest interaction. The densities for the gRNA (yellow) and the previously uncharacterized proteins (green) are also shown. (*D*) A zoomed-in view of the tightest interaction between the maturation protein and coat protein dimer 10 with the clusters of charged residues labeled and shown as spheres (blue and red spheres for positively and negatively charged residues, respectively).

Compared with our previous cryo-EM structures (15), these improved resolutions for both the Qβ virion and the VLP allowed us to more accurately delineate the structural differences between these two types of capsids, which are particularly striking for coat dimers 2–6, with deviations in certain regions larger than 10 Å (*SI Appendix*, Fig. S4). When replacing the conformations of the coat proteins in the virion capsid with their corresponding conformations in the VLP (Fig. 2*C*), there are apparent steric clashes between the VLP capsid and the gRNA, particularly between a gRNA hairpin loop and one FG loop from coat protein dimer 5 (labeled by a star in Fig. 2*C*). A large structural rearrangement, expanding the capsid locally around these coat protein dimers, is necessary to avoid these collisions in the mature Qβ virions. The expansion also leaves a gap between the models of coat protein dimers 3 and 4. Interestingly, at this gap between the coat protein models, some extra density (colored red in Fig. 2*D*) is observed, protruding from the C termini of the coat proteins (*SI Appendix*, Fig. S5). The identity of this extra density is yet to be determined. One explanation is that one or a few coat proteins in dimers 3 and 4 are replaced by A_1 and this extra density belongs to part of the flexible linker of the readthrough domain in A_1.

MurA Binds A_2 in Mature Qβ Virions. During Qβ infection, A_2 inhibits MurA, resulting in host lysis. In addition, the mature Qβ virion has been shown to be a stoichiometric inhibitor of MurA, via its single molecule of A_2 (27). To reveal the interactions between A_2 and MurA, we collected 2,871 movie stacks of the Qβ–MurA complex on a direct detection camera in the superresolution electron counting mode using single-particle cryo-EM. Approximately 20% of total 130,013 particles correspond to

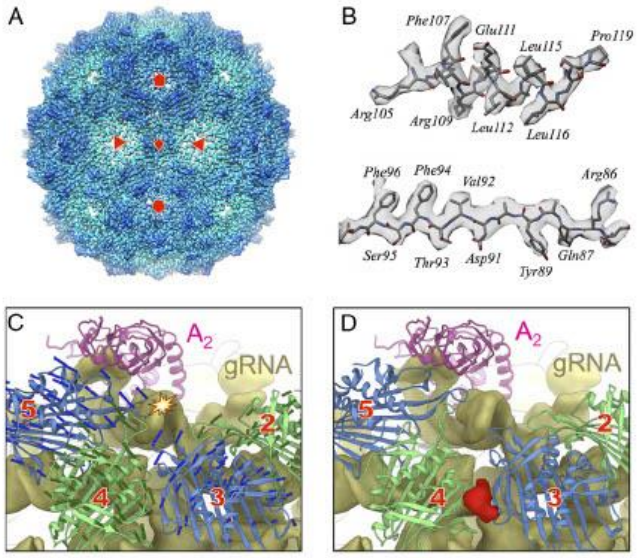


Fig. 2. A comparison between the capsids of Qβ virions and Qβ virus-like particles. (*A*) An icosahedral reconstruction of the Qβ VLPs at 3.3-Å resolution with the red diamond, triangles, and pentagons labeling the twofold, threefold, and fivefold axes, respectively. (*B*) Sidechains visualized on an α-helix (*Top*) and a β-strand (*Bottom*) in the coat proteins. (*C*) Coat protein dimers from the capsid of the Qβ virus-like particles are overlaid onto the gRNA (yellow density) and A_2 (pink model) of the Qβ virion. Only four coat protein dimers (green and blue alternating) are shown and numbered as in Fig. 1*A*. The arrows indicate the conformational difference between these four coat protein dimers in the Qβ virus-like particles and the Qβ virion. The star indicates a severe steric clash where the gRNA density collides with the EF loop of a coat protein in dimer 5 of the Qβ virus-like particles. (*D*) The same view as in *C* but with the four coat protein dimers in their conformations within a mature Qβ virion. An extra density (red) was observed at the interface between the models of coat protein dimers 3 and 4.

Qβ–MurA complexes. The resolution was determined to 6.1 Å for the Qβ–MurA complex and improved to 5.7 Å for the protein components by excluding gRNA during map refinement (*SI Appendix*, Fig. S2 *B* and *D*). The architecture of the Qβ–MurA complex is almost identical to that of the Qβ alone (Fig. 3 *A* and *B*), except for the additional density from MurA binding to A_2. The extra protein density sequestered inside the virion, seen in the structure of Qβ alone, is still observed. No significant conformational rearrangement of the gRNA is seen (*SI Appendix*, Fig. S6).

MurA consists of two globular domains: a catalytic domain and a C-terminal domain, resulting in a clamshell shape (32). The catalytic domain has an 11-amino acid loop (catalytic loop, residues 111–121), creating a "lid" over the active site, where the substrates UDP-GlcNAc and phosphoenolpyruvate (PEP) bind. By refining a model of the complex based on the crystal structures of MurA (32) and A_2 (31) into our cryo-EM density map, we are able to determine the overall interaction pattern (Fig. 3*C*). Both the C-terminal and the catalytic domains of MurA are seen to interact with the β-region of A_2, with the outer surface of A_2 binding at the concave side of MurA, where the catalytic loop is positioned. This interface between MurA and A_2, spanning ~170 $Å^2$, agrees with previous studies that mapped *murA* mutations blocking A_2 lysis in vivo (19). These MurA mutants are fully functional but cannot be inhibited by A_2. The positions of the *nt* (resistant to A-two) mutations, labeled as red stick models in Fig. 3*D*, are clearly capable of disrupting the interactions between A_2 and MurA (Fig. 3*D*). Of particular interest is Leu138 in MurA, the site of the first *nt* mutation (L138Q) (27). Leu138 is positioned directly at the interface where a three-stranded β-sheet (residues 106–110, 135–139, and 142–146) from MurA comes in contact with a five-stranded β-sheet (residues 38–45, 49–55, 91–103, 107–114, and 367–379) from A_2 (*SI Appendix*, Fig. S7). Mutating this residue from a hydrophobic leucine to a hydrophilic glutamine may prevent the proper association between the β-sheets from the two proteins. Further examination of the MurA–A_2 interface reveals that most of the interactions with MurA occur in the N-terminal β-sheet region of A_2 (between residues 30 and 120), as outlined in Fig. 3*E*.

To verify that the imaged interaction involving the N-terminal β-sheet region of A_2 with MurA had biological significance, we tested different truncations and deletions of A_2 for their effect on cell lysis caused by induction of plasmid-borne clones of A_2 (27) (*SI Appendix*, Fig. S9). As expected from the structure, deletions of N-terminal domains were uniformly nonlytic. The shortest N-terminal fragment that was capable of lysis comprised residues 1–180, which spans the entire interface observed by cryo-EM.

During catalysis, MurA undergoes a large conformational change between the unliganded (open) and liganded (closed) states (33). Binding of UDP-GlcNAc triggers the catalytic loop to close, before the association of the second substrate PEP (34, 35). Previous research has shown that the interactions between A_2 and MurA are conformation dependent, specifically favoring the conformation when MurA binds UDP-GlcNAc (19). To stabilize MurA in the UDP-GlcNAc bound state, we pretreated MurA with both UDP-GlcNAc and fosfomycin before incubating with Qβ. Fosfomycin covalently binds to Cys115 in the catalytic loop of MurA to lock it in the closed state, and the structure of the UDP-GlcNAc/fosfomycin bound state (PDB ID: 1UAE) is similar to that of the UDP-GlcNAc bound state (PDB ID: 3KQJ). In our cryo-EM map of Qβ–MurA complex, the density of UDP-GlcNAc is clearly visible in the active site of MurA (Fig. 3 *C* and *D* and *SI Appendix*, Fig. S8). Fosfomycin is not visualized, most likely due to its small mass and the limited resolution of the electron density map. In the unliganded open conformation of MurA, the catalytic loop flips out from the active site and may collide with the β-region of A_2 (Movie S1), which prevents their association. In fact, even a slight flipping out of the catalytic loop for the UDP-GlcNAc/PEP bound state could inhibit its binding to A_2 (19).

An Extra Coat Protein Dimer Is Sequestered Inside the Qβ Virion. In both the cryo-EM density maps of Qβ alone and the Qβ–MurA complex a strong protein density was observed inside the virion between the bottom of A_2 and the gRNA. To resolve this density better, we combined both datasets and subjected them to local 3D classification in the region including A_2, the extra protein density, and the interacting gRNA hairpins (*Experimental Procedures*). The

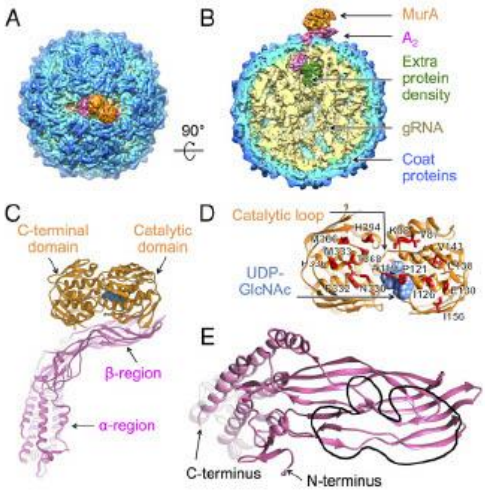


Fig. 3. The structure of Qβ bound with MurA reveals the structural mechanism of A_2-mediated host lysis. (*A*) The structure of Qβ bound to MurA colored as in Fig. 1 *A* and *B* with MurA colored orange. (*B*) A 90° turn and cutaway view of Qβ shows MurA bound to the maturation protein. (*C*) The ribbon model of A_2 bound to MurA with uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc) in the active site (cornflower blue). (*D*) The ribbon model of MurA viewed from the MurA–A_2 interface. The point mutations that render MurA resistant to A_2 are labeled and shown as red stick models. Locations of the catalytic loop and the UDP-GlcNAc are indicated by black arrows. (*E*) Ribbon model of A_2 as viewed from the MurA–A_2 interface. The region interacting with MurA, encompassing the *N*-terminal β-sheet region of residues 30–120, is outlined by a black lasso. The N and C termini are indicated by black arrows.

resultant map shows much better resolvability, with the improved extra density able to fit a coat protein dimer, which interacts with a gRNA hairpin in a mode that is similar to the interaction between a coat protein dimer and the RNA hairpin "operator" at the start of the replicase gene (36) (Fig. 4*A*). When compared at the local resolution (8 Å) in this region, the cross-correlation score between this density and the fitted crystal structure of the Qβ coat protein dimer, in complex with the operator RNA (PDB ID 4L8H), is 0.91.

An RNA domain that contains a five-way junction was observed close to A_2 and the extra coat protein dimer, with the five RNA helices around the junction labeled from 1 to 5 (Fig. 4 *B* and *C*). This local RNA domain connects back to other parts of the gRNA through helix 1. The hairpin loop of RNA helix 2 displaces the FG loop of one coat protein on coat protein dimer 5. RNA helices 3–5, each containing a hairpin loop, interact with A_2, the sequestered coat protein dimer and coat protein dimer 1, respectively. The sequestered coat protein dimer interacts with A_2 via its CD and DE loops, which clamp the bottom of the α-region of A_2 (Fig. 4*D*). Along with the interaction between coat protein dimer 1 and A_2, the interactions between these protein components may further compact the RNA five-way junction.

Discussion

Low Infectivity for Qβ May Be Due to Either VLP Formation or MurA Binding. Qβ, and many ssRNA phages, have low viability, as only a fraction of the purified particles are actually infectious (37, 38). In our cryo-EM data analysis, around 20% of the total number of particles contain A_2 and have a defined gRNA conformation inside the capsid (*SI Appendix*, Fig. S1). After further classification of the remaining dataset, we determined that the remaining particles are VLPs, inasmuch as they lack A_2, leading to a perfectly icosahedral capsid. The RNA in these VLPs could be viral RNA, which does not have A_2 bound; or host RNA, since it has been demonstrated that ssRNA viruses can encapsidate host RNAs (39). By testing coat protein binding to purified RNAs from Qβ and MS2, it was found that the Qβ coat protein has a higher affinity for both homologous and heterologous assembly than the MS2 coat protein (40). The higher affinity for the coat proteins of Qβ to bind RNA could be due to its positively charged interior EF loop (15), which may facilitate the nonspecific binding to the negatively charged RNA backbones. Indeed, the ability to encapsidate foreign RNA is the basis on which the Qβ coat protein is commercially used to protect heterologous RNAs (41). It is also possible that just as there is a sequestered coat protein dimer within the virion, the maturation protein may also bind to the Qβ gRNA in these VLPs but be internalized due to mispackaging. The low infectivity of Qβ could further be due to the cellular MurA binding A_2, wherein it prevents efficient assembly of the mature virion or occupies the interface of A_2 that is required for attachment to the host F pilus.

Potential Role of the Extra Coat Protein Dimer. The extra coat protein dimer sequestered in the interior of Qβ has not been seen in other ssRNA viruses. Its presence lends credence to the role of "packaging signals" (11) in virion morphogenesis. As the viral gRNA is replicated and translated, coat proteins are synthesized and form dimers to bind to packaging signals, mainly RNA hairpins, such as the operator, a hairpin at the start of the

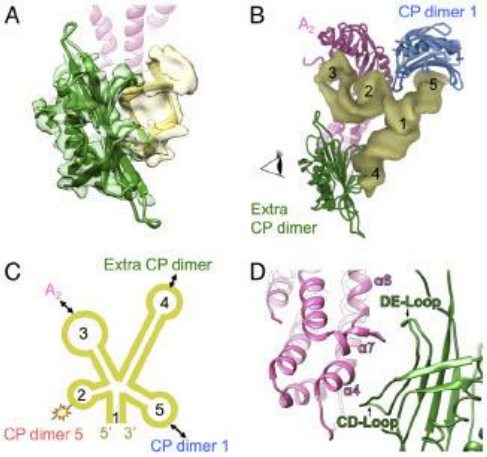


Fig. 4. A coat protein dimer is sequestered inside the capsid proximal to the bottom of A_2 but away from the rest of the capsid. (*A*) The additional protein density (transparent green) inside the virion fits the model of a coat protein dimer (green), which interacts with a gRNA hairpin (yellow ribbon model in transparent yellow electron density). The ribbon model of A_2 is colored pink and shown in the back. (*B*) A gRNA domain (yellow density), containing a five-way junction, interacts with A_2 (pink), the sequestered coat protein (CP) dimer (green) and CP dimer 1 (blue). The helices on this gRNA five-way junction are labeled 1–5 from the 5′ to the 3′ ends. (*C*) A scheme showing the secondary structures of the gRNA five-way junction domain with helices labeled from 1 to 5. Double-headed arrows indicate the interacting component of each hairpin on the corresponding helix. The star collision sign indicates where hairpin 2 displaces coat protein dimer 5 as in Fig. 2*C*. (*D*) Zoomed-in region as viewed by the eye cartoon in *B* showing the interactions between the extra coat protein dimer and the bottom of the α-region in A_2. The CD loop and DE loop from the sequestered CP dimer and the α-helices of α4, α6, and α7 from A_2 are labeled.

replicase gene that is known to bind coat proteins to repress translation of the replicase (42). The current resolutions of the Qβ RNA do not allow the assignment of the genome sequence solely based on its cryo-EM density. It needs to be determined whether the RNA sequence that the internalized coat protein dimer binds corresponds to the operator of the replicase gene or a sequence similar to the operator. It also cannot be ruled out that there may be a small fraction of Qβ particles that contain more coat protein dimers, internalized at different locations in the virion and bound to other gRNA hairpins facing inwards. The internalized coat protein dimer, observed in our Qβ virion, may participate in a nucleation event wherein the combination of the maturation protein, the sequestered coat protein dimer and the capsid coat protein dimers condense the gRNA locally around the RNA five-way junction, compacting the gRNA and enabling it to fit into the dimensions of the capsid.

Such an internalized coat protein dimer and its associated compact gRNA domain may also affect the genome release during viral infection. Helix 3 from this gRNA domain directly associates with A_2 (Fig. 4) and is likely the first portion of the RNA to exit the capsid during the genome release. The expansion of the capsid around A_2 (Fig. 2 and *SI Appendix*, Fig. S4) may facilitate the initial release of this ordered gRNA domain. It would be interesting to determine if the sequestered coat protein dimer remains inside the capsid or just simply diffuses away after the maturation protein and gRNA are removed from the virion. Alternatively, the sequestered coat protein dimer may recap the opening that is left after A_2 is removed from the capsid, leaving an empty but icosahedral VLP.

Structural Comparison of the Maturation Proteins from Two Model ssRNA Phages. A_2 from Qβ and the A protein from MS2 share high structural similarity within the α-region (*SI Appendix*, Fig. S10*A*). However, the connections between the α- and β-regions of these two proteins are very different. A_2 has a long α-helix (residues 128–177) from the α-region to the β-region, while the A protein has a zigzag helix–loop–helix (residues 117–169) motif to connect the two regions (*SI Appendix*, Fig. S10*B*). The different angles between the α- and β-regions for these two proteins lead to altered insertions into the capsid, at varying angles relative to the opening in the capsid (*SI Appendix*, Fig. S10*C*). When viewing A and A_2 in their corresponding virions, we see that the α-region of A_2 inserts at an orientation that is more perpendicular to the capsid opening (*SI Appendix*, Fig. S10*D*). Even though the coat proteins and α-region of the maturation proteins from MS2 and Qβ are structurally similar, the interactions between the α-region of the maturation protein and its surrounding coat proteins are different between MS2 and Qβ (*SI Appendix*, Fig. S10*C*), indicating these interactions are not conserved between the two model phages. The maturation protein may be loosely anchored in the capsid while being held in place by the neighboring coat proteins that wrap around it. The weak interactions between the maturation protein and the coat proteins could further facilitate the release of the maturation protein, along with the attached gRNA, from the capsid during infection.

Even though folded differently, the overall geometries of the β-regions in A and A_2 are similar, both presenting a flat outer surface that is angled ∼30° to the opening of the capsid (*SI Appendix*, Fig. S10*D*). Based on a low-resolution density map of MS2 bound to the F pilus obtained through cryoelectron tomography (43), the outer surface of the β-region of the MS2 A protein binds the pilus. Presumably, the similar outer surface of Qβ would also bind to the pilus. It has also been reported that treating Qβ with MurA abolishes its ability to infect *Escherichia coli* (29), suggesting that the F pilus and MurA bind the same or overlapping locations on the β-region of A_2. However, a detailed picture of the interaction between the F pilus and the maturation protein requires a high-resolution structure of the phage–pilus complex.

Experimental Procedures

Sample and Cryo-EM Grid Preparation. Qβ and MurA were purified as previously described with minor modifications (19, 44). After separation using a cesium chloride gradient, Qβ particles were further purified by a 200-mL hand-packed S500 gel-filtration column. The sample of MurA-bound Qβ was prepared by pretreating MurA with 1 mM UDP-GlcNAc and 1 mM fosfomycin on ice for 30 min, and then incubating with Qβ on ice for another 30 min. The molar ratio of MurA to Qβ was ∼5–1. Cryo-EM specimens were prepared by applying 3 μL of the sample solution to a C-Flat 1.2/1.3 or 2/1 400-mesh Holey Carbon Grid (Protochips) at 22 °C with 100% relative humidity and vitrified using a Vitrobot Mark III (FEI Company).

Cryo-EM Data Collection. Qβ alone and some of the Qβ–MurA complex images were recorded under a JEM3200FSC cryoelectron microscope with a field emission gun (JEOL) operated at 300 kV. Data were collected using SerialEM (45) on a K2 Summit direct detection camera (Gatan) in the superresolution electron counting mode (30). A nominal magnification of 30,000× was used, yielding a subpixel size of 0.608 Å. The beam intensity was adjusted to a dose rate of 7e^- per pixel per second on the camera. A 50-frame movie stack was recorded for each exposure with 0.2 s per frame for a total exposure time of 10 s. An in-column energy filter was used with a slit width of 29 eV.

Most of the Qβ–MurA complex images were obtained under an FEI Tecnai F20 cryoelectron microscope with a field emission gun (FEI Company) operated at 200 kV. Data were recorded using SerialEM on a K2 Summit direct detection camera (Gatan) in the superresolution electron counting mode. A nominal magnification of 29,000× was used, yielding a subpixel size of 0.625 Å. The beam intensity was adjusted to a dose rate of 8.7e^- per pixel per second on the camera. A 33-frame movie stack was recorded for each picture, with 0.2 s per frame for a total exposure time of 6.6 s.

Image Preprocessing. The collected superresolution movie stacks were first binned by 2, then aligned, summed, and filtered using Unblur (46). The defocus value of each summed micrograph was determined using CTFFIND4 (47). Micrographs with strong power spectrums were used to semi-automatically pick particles with e2boxer.py in EMAN2 (48). The 2D and 3D classifications and 3D refinement of the image data were performed in Relion (49).

The data processing for the Qβ alone sample was as follows (*SI Appendix*, Fig. S1): Particles were selected from 2,370 micrographs and subjected to 2D classifications in Relion to remove bad particles. After 2D classifications, 248,445 particles were chosen for three independent rounds of unsupervised 3D classifications, each for 30 interactions, requesting three, four, and five classes, respectively. Each round of the 3D classification produced a single class containing particles with defined densities for A_2 and the gRNA. These three single classes were combined to yield 76,843 particles without the duplication of particles. To improve the resolution for A_2, 50 iterations of a masked classification around A_2 with signal subtraction was performed (50). A total of 46,471 particles with good A_2 density were selected for the 3D refinement of the Qβ virion, giving a 4.7-Å resolution map of the Qβ virion. The resolution of the capsid for the Qβ virion was improved to 4.4 Å after a masked refinement excluding the gRNA density. The 3D classification of the remaining 171,602 particles revealed no A_2 density and the capsid of these particles was perfectly icosahedral. These particles were subjected to icosahedral refinement, giving a 3.3-Å resolution map of the Qβ VLP capsid.

The structure of the Qβ–MurA complex was determined by combining 1,117 and 1,754 images collected using the JEM3200FSC and the FEI Tecnai F20 microscopes, respectively. By carefully comparing the two maps calculated from icosahedral refinement in those two datasets, images collected on the JEM3200FSC microscope were downscaled to match the pixel size of those collected on the FEI Tecnai F20 microscope. After 2D and 3D classifications, 25,597 of the 130,013 total particles were used for the final reconstruction to yield a 5.7-Å map after using an outer shell mask, which exdudes the gRNA.

To achieve better map quality for the extra coat protein dimer density near the bottom of A_2, we combined all 5,241 images of Qβ alone and Qβ–MurA complex by the same procedure as we combined the two Qβ–MurA datasets. With a masked classification around A_2 and the extra coat protein dimer, 47,621 good particles were selected for the final refinement. A 6.2-Å map of A_2 and the coat protein dimer was obtained after masking.

Resolution Estimation. The overall resolutions of all these reconstructed maps were assessed using the gold-standard criterion of Fourier shell correlation,

with a cutoff at 0.143, between two half-maps from two independent half-sets of data (51). Local resolutions were estimated using Resmap (52).

Modeling and Visualization. To build the atomic model for the Qβ VLP capsid, our previous icosahedral Qβ structure (PDB ID: 5KIP) was used as the initial model and subsequently refined into the 3.3-Å icosahedral map of the Qβ VLP by the real-space refinement from PHENIX (53). To build the backbone model for the Qβ virion capsid, the crystal structure of A_2 (PDB ID: 5MNT) was used to replace one coat protein dimer on the Qβ VLP capsid from the previous step. The model for the internalized coat protein dimer was then added to the capsid of the Qβ virion by fitting into its corresponding density as a rigid body. The backbone of the Qβ virion capsid was then refined into the 4.4-Å asymmetric map using the real-space refinement with the secondary structure restraints in PHENIX. To model the interaction between A_2 and MurA, the density of the A_2–MurA complex was firstly segmented out from the 5.7-Å map of the Qβ–MurA complex using University of California San Francisco (UCSF) Chimera (54). The crystal structures of MurA (PDB ID: 1UAE) and A_2 were then rigidly fitted into the segmented density and further refined by the molecular dynamics flexible fitting method (55) followed by the real-space refinement with the secondary structure restraints in PHENIX. The model statistics are summarized in *SI Appendix*, Table S1.

The model of the *E. coli* MurA in its open conformation was built using homology modeling by SWISS-MODEL (56), based on the crystal structure of the *Enterobacter cloacae* MurA (PDB ID: 1EJD). All of the figures and movies were made using UCSF Chimera.

ACKNOWLEDGMENTS. We thank Karthik Chamakura for his expertise in ssRNA viruses and comments on the paper, and the Microscopy and Imaging Center at Texas A&M University and the National Center for Macromolecular Imaging (NCMI) at Baylor College of Medicine for the cryo-EM data collection. NCMI is supported by NIH Grants P41GM103832 and U24GM116787. We also thank the Texas A&M High Performance Research Computing Center for providing the computational resources for the data processing. J.Z. is supported by startup funding from the Department of Biochemistry and Biophysics at Texas A&M University and Welch Foundation Grant A-1863. R.Y. is supported by Public Health Service Grant GM27099. J.Z. and R.Y. acknowledge the support of the Center for Phage Technology jointly sponsored by Texas A&M AgriLife and Texas A&M University.

1. Koonin EV, Dolja VV, Krupovic M (2015) Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virology* 479–480:2–25.
2. Roossinck MJ (2003) Plant RNA virus evolution. *Curr Opin Microbiol* 6:406–409.
3. Kolakofsky D (2015) A short biased history of RNA viruses. *RNA* 21:667–669.
4. Barré-Sinoussi F, Ross AL, Delfraissy JF (2013) Past, present and future: 30 years of HIV research. *Nat Rev Microbiol* 11:877–883.
5. Kajitani M, Kato A, Wada A, Inokuchi Y, Ishihama A (1994) Regulation of the *Escherichia coli hfq* gene encoding the host factor for phage Q β. *J Bacteriol* 176:531–534.
6. Poot RA, Tsareva NV, Boni IV, van Duin J (1997) RNA folding kinetics regulates translation of phage MS2 maturation gene. *Proc Natl Acad Sci USA* 94:10110–10115.
7. de Smit MH, van Duin J (1993) Translational initiation at the coat-protein gene of phage MS2: Native upstream RNA relieves inhibition by local secondary structure. *Mol Microbiol* 9:1079–1088.
8. Klovins J, Berzins V, van Duin J (1998) A long-range interaction in Qbeta RNA that bridges the thousand nucleotides between the M-site and the 3′ end is required for replication. *RNA* 4:948–957.
9. Brown D, Gold L (1996) RNA replication by Q β replicase: A working model. *Proc Natl Acad Sci USA* 93:11558–11562.
10. Gytz H, et al. (2015) Structural basis for RNA-genome recognition during bacteriophage Qβ replication. *Nucleic Acids Res* 43:10893–10906.
11. Rolfsson Ó, et al. (2016) Direct evidence for packaging signal-mediated assembly of bacteriophage MS2. *J Mol Biol* 428:431–448.
12. Stockley PG, et al. (2013) Packaging signals in single-stranded RNA viruses: Nature's alternative to a purely electrostatic assembly mechanism. *J Biol Phys* 39:277–287.
13. Peabody DS (1997) Role of the coat protein-RNA interaction in the life cycle of bacteriophage MS2. *Mol Gen Genet* 254:358–364.
14. Horn WT, et al. (2006) Structural basis of RNA binding discrimination between bacteriophages Qbeta and MS2. *Structure* 14:487–495.
15. Gorzelnik KV, et al. (2016) Asymmetric cryo-EM structure of the canonical Allolevivirus Qβ reveals a single maturation protein and the genomic ssRNA in situ. *Proc Natl Acad Sci USA* 113:11519–11524.
16. Dai X, et al. (2017) In situ structures of the genome and genome-delivery apparatus in a single-stranded RNA virus. *Nature* 541:112–116.
17. Koning RI, et al. (2016) Asymmetric cryo-EM reconstruction of phage MS2 reveals genome structure in situ. *Nat Commun* 7:12524.
18. Zhong Q, et al. (2016) Genetic, structural, and phenotypic properties of MS2 coliphage with resistance to ClO_2 disinfection. *Environ Sci Technol* 50:13520–13528.
19. Reed CA, Langlais C, Kuznetsov V, Young R (2012) Inhibitory mechanism of the Qβ lysis protein A_2. *Mol Microbiol* 86:836–844.
20. Garwes D, Sillero A, Ochoa S (1969) Virus-specific proteins in Escherichia coli infected with phage Qb. *Biochim Biophys Acta* 186:166–172.
21. Blumenthal T, Carmichael GG (1979) RNA replication: Function and structure of Qbeta-replicase. *Annu Rev Biochem* 48:525–548.
22. Rumnieks J, Tars K (2011) Crystal structure of the read-through domain from bacteriophage Qβ A_1 protein. *Protein Sci* 20:1707–1712.
23. Weber K, Konigsberg W (1975) Proteins of the RNA phages. *Cold Spring Harb Monogr Arch* 5:51–84.
24. Hofstetter H, Monstein HJ, Weissmann C (1974) The readthrough protein A_1 is essential for the formation of viable Q β particles. *Biochim Biophys Acta* 374:238–251.
25. Richelson E, Nathans D (1967) Association of bacteriophage proteins and RNA in *E. coli* infected with MS2. *Biochem Biophys Res Commun* 29:842–849.
26. Kastelein RA, Remaut E, Fiers W, van Duin J (1982) Lysis gene expression of RNA phage MS2 depends on a frameshift during translation of the overlapping coat protein gene. *Nature* 295:35–41.
27. Bernhardt TG, Wang IN, Struck DK, Young R (2001) A protein antibiotic in the phage Qbeta virion: Diversity in lysis targets. *Science* 292:2326–2329.
28. Brown ED, Vivas EI, Walsh CT, Kolter R (1995) MurA (MurZ), the enzyme that catalyzes the first committed step in peptidoglycan biosynthesis, is essential in *Escherichia coli*. *J Bacteriol* 177:4194–4197.
29. Reed CA, Langlais C, Wang IN, Young R (2013) A_2(2) expression and assembly regulates lysis in Qβ infections. *Microbiology* 159:507–514.
30. Li X, et al. (2013) Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods* 10:584–590.
31. Rumnieks J, Tars K (2017) Crystal structure of the maturation protein from bacteriophage Qβ. *J Mol Biol* 429:688–696.
32. Skarzynski T, et al. (1996) Structure of UDP-N-acetylglucosamine enolpyruvyl transferase, an enzyme essential for the synthesis of bacterial peptidoglycan, complexed with substrate UDP-N-acetylglucosamine and the drug fosfomycin. *Structure* 4:1465–1474.
33. Schönbrunn E, Svergun DI, Amrhein N, Koch MH (1998) Studies on the conformational changes in the bacterial cell wall biosynthetic enzyme UDP-N-acetylglucosamine enolpyruvyl transferase (MurA). *Eur J Biochem* 253:406–412.
34. Zhu JY, et al. (2012) Functional consequence of covalent reaction of phosphoenolpyruvate with UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA). *J Biol Chem* 287:12657–12667.
35. Skarzynski T, Kim DH, Lees WJ, Walsh CT, Duncan K (1998) Stereochemical course of enzymatic enolpyruvyl transfer and catalytic conformation of the active site revealed by the crystal structure of the fluorinated analogue of the reaction tetrahedral intermediate bound to the active site of the C115A mutant of MurA. *Biochemistry* 37:2572–2577.
36. Rumnieks J, Tars K (2014) Crystal structure of the bacteriophage Qβ coat protein in complex with the RNA operator of the replicase gene. *J Mol Biol* 426:1039–1049.
37. Watanabe H, Watanabe M (1970) Comparative biology of five RNA phages, R23, f2, Q β, R34, and R40. *Can J Microbiol* 16:859–864.
38. Leipold B, Hofschneider PH (1975) Isolation of an infectious RNA-A-protein complex from the bacteriophage M12. *FEBS Lett* 55:50–52.
39. Routh A, Domitrovic T, Johnson JE (2012) Host RNAs, including transposons, are encapsidated by a eukaryotic single-stranded RNA virus. *Proc Natl Acad Sci USA* 109:1907–1912.
40. Ling CM, Hung PP, Overby LR (1969) Specificity in self-assembly of bacteriophages Q β and MS2. *Biochemistry* 8:4464–4469.
41. Bundy BC, Swartz JR, Chan W (2010) Encapsidation of heterologous entities into virus-like particles. US Patent 12/621,416.
42. Weber H (1976) The binding site for coat protein on bacteriophage Qbeta RNA. *Biochim Biophys Acta* 418:175–183.
43. Dent KC, et al. (2013) The asymmetric structure of an icosahedral virus bound to its receptor suggests a mechanism for genome release. *Structure* 21:1225–1234.
44. Strauss JH, Jr, Sinsheimer RL (1963) Purification and properties of bacteriophage MS2 and of its ribonucleic acid. *J Mol Biol* 7:43–54.
45. Mastronarde DN (2005) Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* 152:36–51.
46. Grant T, Grigorieff N (2015) Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *eLife* 4:e06980.
47. Rohou A, Grigorieff N (2015) CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* 192:216–221.
48. Tang G, et al. (2007) EMAN2: An extensible image processing suite for electron microscopy. *J Struct Biol* 157:38–46.
49. Scheres SH (2012) RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 180:519–530.
50. Bai XC, Rajendra E, Yang G, Shi Y, Scheres SH (2015) Sampling the conformational space of the catalytic subunit of human γ-secretase. *eLife* 4:e11182.
51. Scheres SH, Chen S (2012) Prevention of overfitting in cryo-EM structure determination. *Nat Methods* 9:853–854.
52. Kucukelbir A, Sigworth FJ, Tagare HD (2014) Quantifying the local resolution of cryo-EM density maps. *Nat Methods* 11:63–65.
53. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221.
54. Pettersen EF, et al. (2004) UCSF Chimera: A visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612.
55. Trabuco LG, Villa E, Mitra K, Frank J, Schulten K (2008) Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* 16:673–683.
56. Schwede T, Kopp J, Guex N, Peitsch MC (2003) SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* 31:3381–3385.

SI Appendix:

Structures of Qβ virions, virus-like particles, and the Qβ-MurA complex reveal internal coat proteins and the mechanism of host lysis

Zhicheng Cui [a], Karl V. Gorzelnik [a], Jeng-Yih Chang [a], Carrie Langlais [a,1], Joanita Jakana [b], Ry Young [a], Junjie Zhang [a,2]

[a] Department of Biochemistry and Biophysics, Center for Phage Technology, Texas A&M University, College Station, TX 77843

[b] National Center for Macromolecular Imaging, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030

Movie Legends

Movie S1. Substrate binding causes the closure of the catalytic loop in MurA, which allows MurA to bind to the outer surface of A_2. The model for the open conformation of the *E. coli* MurA with no substrate bound was built using homology modeling based on the crystal structure of *Enterobacter cloacae* MurA (PDB ID: 1EJD).

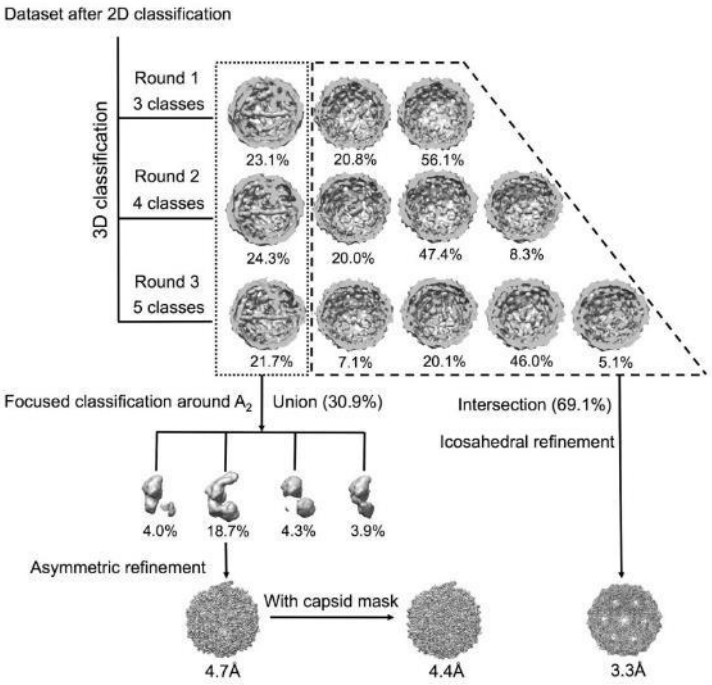


Figure S1. 3D classification of the Qβ image data. For the dataset screened after 2D classification, 30 iterations of the 3D classification was first performed for three independent rounds in Relion, requesting 3, 4 and 5 number of classes, respectively. For each round of the 3D classifications, one class showed defined densities of A_2 and the gRNA (outlined in a dotted rectangle). Particles within this class in each of the three rounds were combined as a union set with duplicate particles eliminated. This yielded about 30.9% of the total particles, which were subjected to another round of focused classification for 50 iterations around A_2. Finally, 18.7% of the particles showed good A_2 density, and particles in this class, corresponding to the Qβ virions, were asymmetrically refined to 4.7-Å resolution. The resolution was improved to 4.4Å after excluding the gRNA density during the final map refinement. The remaining particles in those classes that did not show a defined density for A_2 or RNA (outlined in a dashed trapezoid) turned out to have a perfect icosahedral capsid, corresponding to the virus-like particles (VLP). These VLPs from each round of the 3D classifications were combined as an intersection set, sharing only 69.1% of the total particles, and refined to 3.3-Å resolution with an icosahedral symmetry applied.

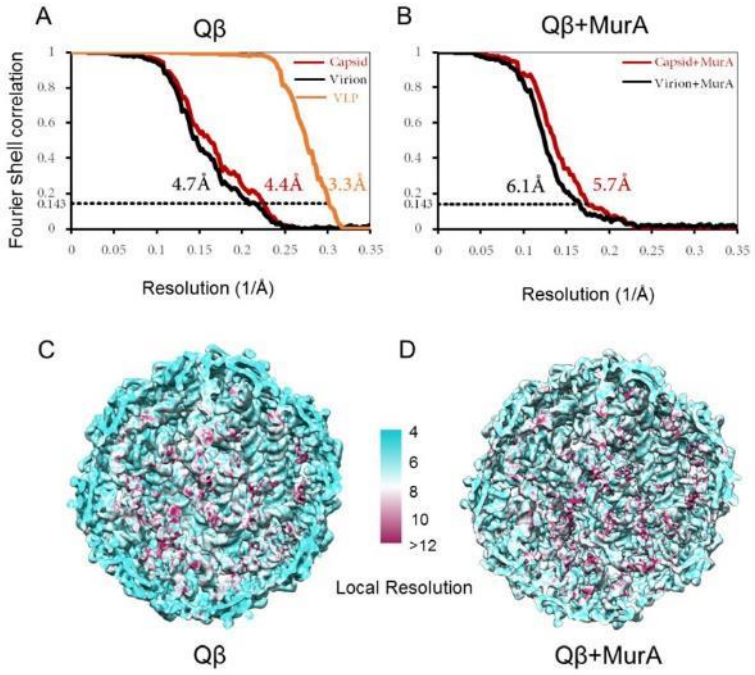


Figure S2. Resolution estimations. (A-B) Fourier shell correlation (FSC) of cryo-EM reconstructions of the Qβ virion, Qβ VLP and Qβ bound with MurA showing the overall resolutions based on the gold-standard criteria. The FSC curves for the complete virion, VLP and the virion capsid without the gRNA are in black, orange and red, respectively. (C-D) Cutaway view of each reconstruction showing the local resolution. The color scale indicates the local resolution. The map for Qβ alone is low-pass filtered to 6Å to better visualize the gRNA.

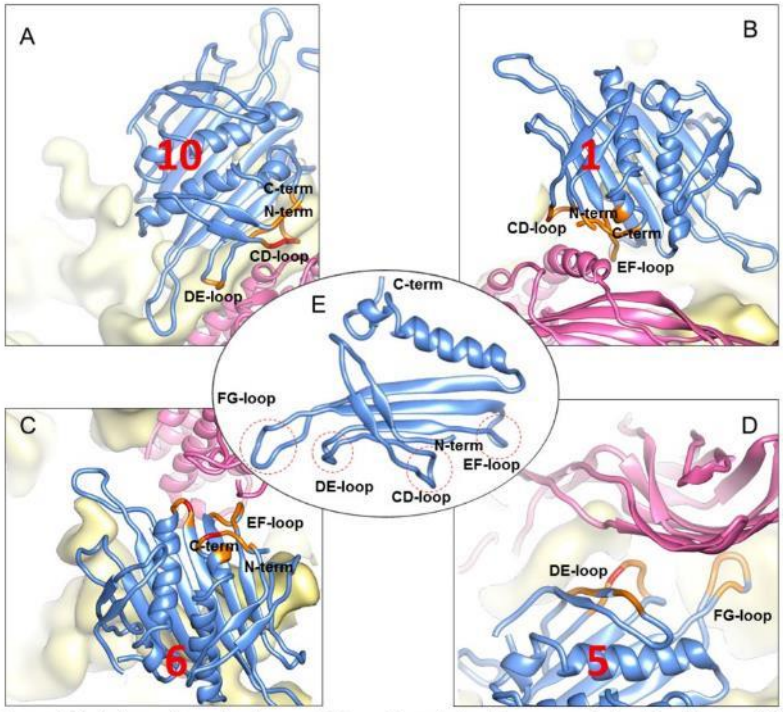


Figure S3. Interactions between Qβ coat protein dimers and A_2. Residues of the coat proteins within 5Å and 10Å Cα-Cα distances to A_2 are colored red and orange, respectively. The numbering for the coat protein dimers follows Figure 1. (A) For coat protein dimer 10, its N-terminus (residues 2-4), CD-loop (residues 23-35), DE-loop (residues 40-43) and C-terminus (residues 131-132) interact with the N-terminal end of α-helix 9 (residue 389-400), the N-terminal end of β-strand 12 (residues 340-344) and the C-terminal end of β-strand 11 (residues 280-283) of A_2. (B) For coat protein dimer 1, its N-terminus (residues 1-2), CD-loop (residues 26-32), EF-loop (residues 52-57) and C-terminus (residues 127-132) interact with α-helix 2 (residues 129-146) of A_2. (C) For coat protein dimer 6, its EF-loop (residues 54-57 and residues 59-60), the C-terminal of β-strand G (residues 97-101) and N-terminus (residues 123-130) interact with N-terminus (residues 3, residues 10-16) and the N-terminal end of α-helix 6 (residues 289-290 and 292-294) of A_2. (D) For coat protein dimer 5, its AB-loop (residues 16-18), β-strand B (residues 20-22), DE-loop (residues 36-42), and FG-loops (residue 75-76 and 79-82) interact with β-strand 5 (residues 51-53), the N-terminal end of β-strand 4 (residues 34-38), β-strand 9 (residues 114-116), β-strand 8 (residues 92-96), β-strand 13 (residues 372-379), and β-strand 12 (residues 355-358) of A_2. (E) A single copy of the coat protein with the loops and termini annotated.

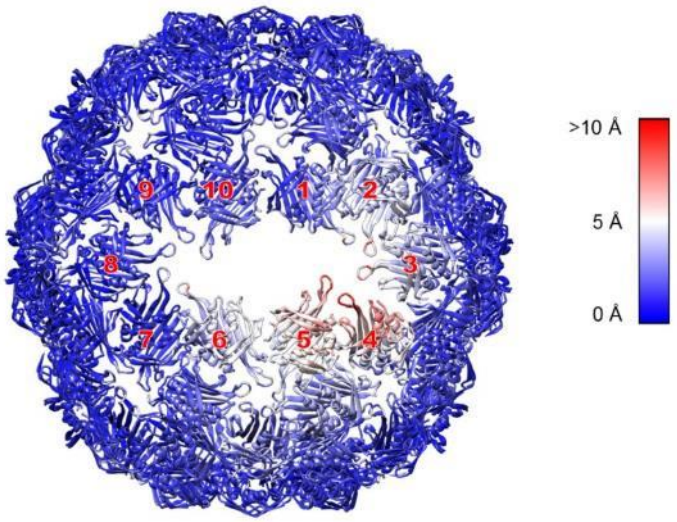


Figure S4. Conformational deviations of the corresponding Qβ coat proteins in the mature virion versus the virus-like particles (VLP). Coat protein dimers around the capsid opening are labeled as in Figure 1. The displacement of coat protein dimers in the mature virion compared to VLP is shown through a color scale. Coat protein dimers 2-6 are the most displaced in the mature virion compared to the virus-like particles, with some loop residues being displaced more than 10Å.

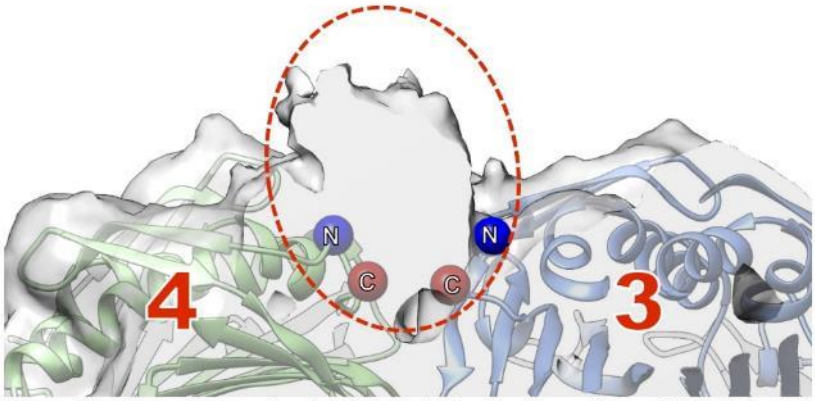


Figure S5. Extra electron density at the interface between the models of Qβ coat protein dimers 3 and 4. The N- and C-termini of the adjacent coat protein models at the interface are labeled as blue and red spheres, respectively. The extra density (from the unfiltered map of the Qβ virion) not occupied by the models of the coat proteins is marked by the red dashed oval.

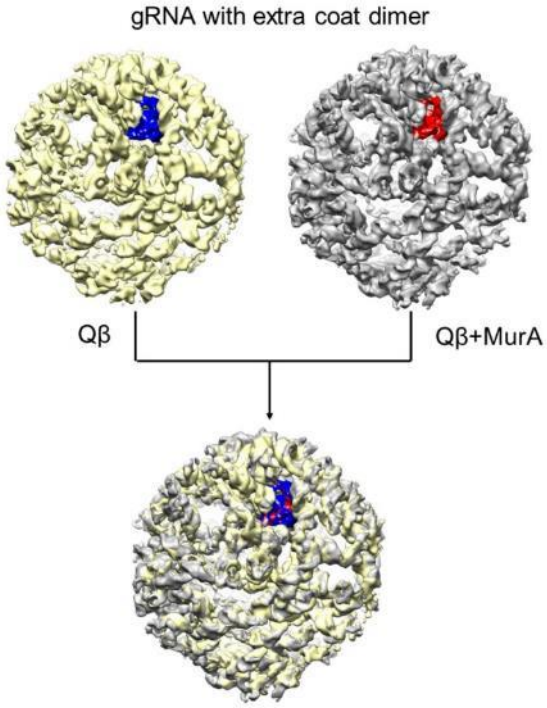


Figure S6. Structural comparison of the gRNA and the sequestered coat protein dimer between Qβ and Qβ+MurA. The protein capsid is removed (except for the sequestered coat protein dimer) and the gRNA is colored yellow and gray for Qβ alone and Qβ bound with MurA, respectively. The density for the sequestered coat protein dimer is colored blue and red in Qβ alone and Qβ bound with MurA, respectively. When the electron density maps are superimposed, there are no major differences between the two maps. The density from the Qβ alone map is low-pass filtered to 6-Å resolution.

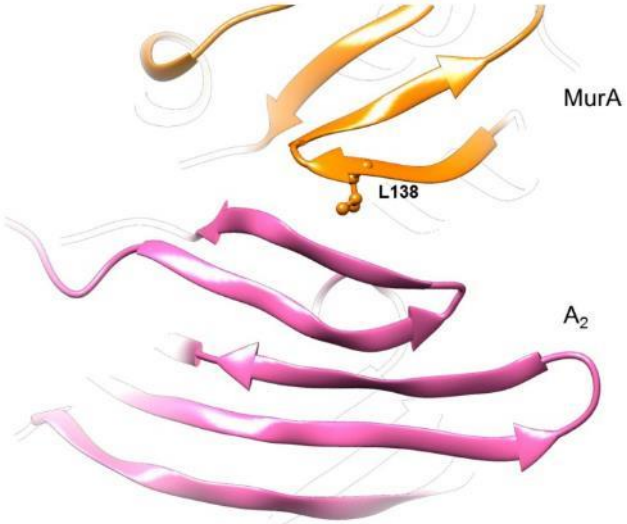


Figure S7. The interaction between the β-sheets from MurA and A_2. MurA is colored orange while A_2 is hot pink. Leu138, the residue with its mutation in MurA causing resistance to A_2, is labeled as a ball-and-stick model.

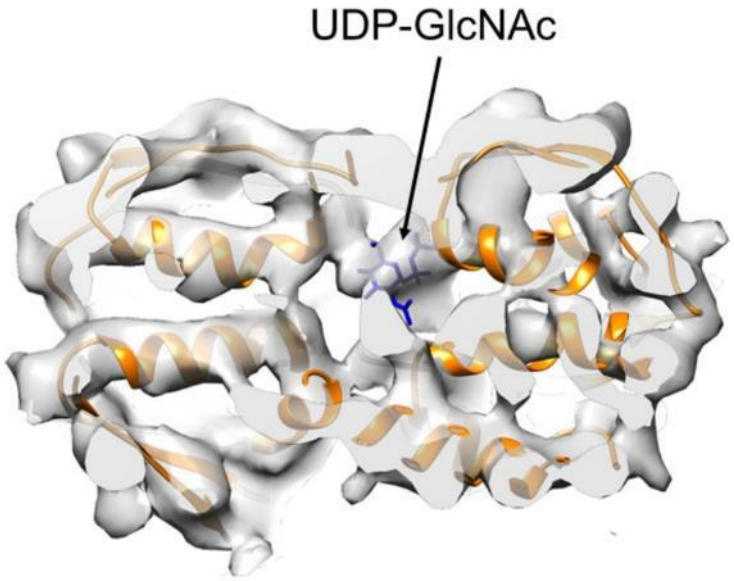


Figure S8. Slice view of the density for MurA from the MurA-Qβ complex. The models of MurA (orange) and its substrate UDP-GlcNAc (blue) fit well into the density (transparent gray).

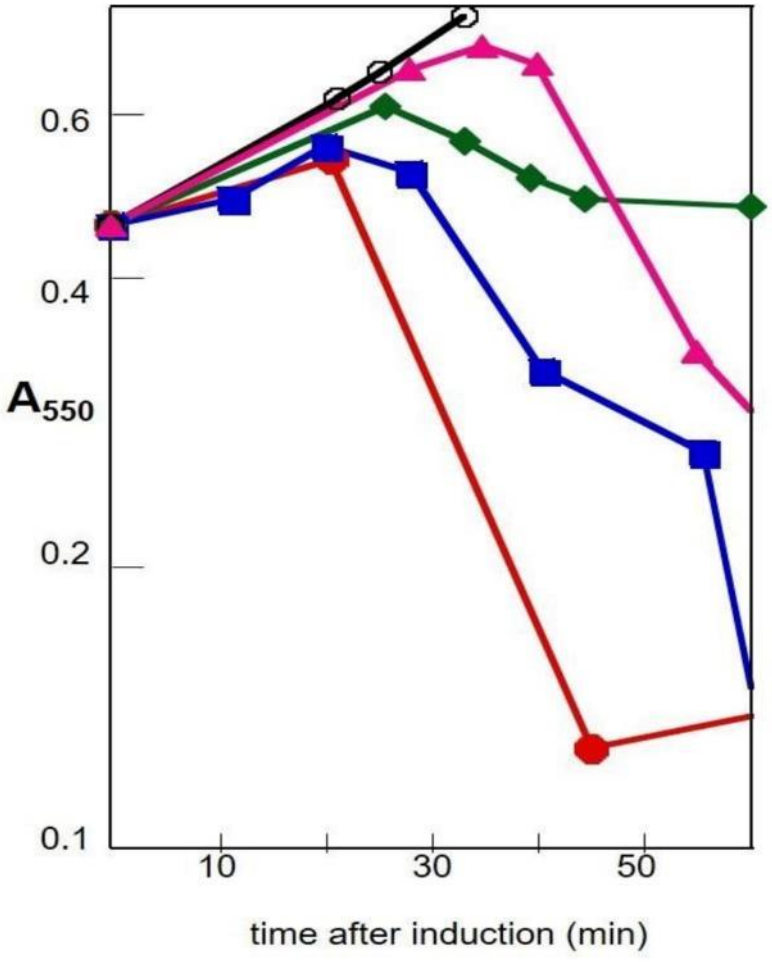


Figure S9. N-terminal fragments of A_2 can cause lysis of *E. coli*. Growth profiles of N-terminal fragments of A_2 as expressed in BL21(DE3) pLysS pET11a-(His)A_2. Induction of N-terminal fragments consisting of the first 190, 186 and 180 residues (red circles, blue squares, green diamonds respectively) but not 172 (black open circles) or fewer, cause lysis. Full length A_2 is shown (pink triangles). Culture turbidity indicating the cell lysis was monitored as a function of time.

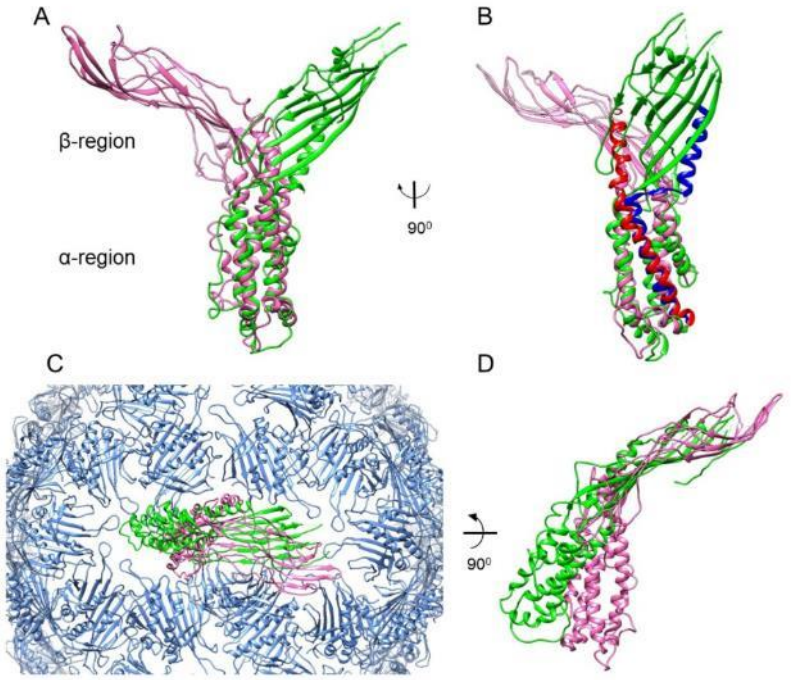


Figure S10. Structural comparison of the maturation proteins from Qβ and MS2. (A) The maturation protein, A protein, from MS2 (green, PDB ID: 5TC1) is aligned to the refined cryo-EM model of A_2 from Qβ (hot pink) based on their α-regions. (B) The long helix in A_2 and the helix-loop-helix in A that connect the α- and β-regions are colored red and blue, respectively. (C) The maturation proteins from MS2 and Qβ are aligned based on the their proximal coat protein dimers. For visualization only the capsid coat proteins from Qβ are shown (cornflower blue). (D) A 90° rotation with the capsid coat proteins removed shows the side view of the A and A_2 proteins, revealing that the α-regions of the two maturation proteins insert into their capsids at different angles.

Table S1. Statistics for the backbone model of the Qβ virion capsid, the atomic model of the Qβ VLP capsid and the pseudo-atomic model of the MurA-A_2 complex.

Model statistics	Backbone model of the Qβ virion capsid	Atomic model of the Qβ VLP capsid	Pseudo-atomic model of the MurA-A_2 complex
All-atom clash score	2.08	6.11	20.24
Ramachandran plot (%):			
outliers	0.00	0.00	0.24
allowed	6.12	7.94	7.2
favored	93.88	92.06	92.56
Rotamer outliers (%)	N/A	0.00	0.00
C-β deviations	N/A	0	0