Cryo-EM structure of the Mycobacterium smegmatis ribosome

Mishra, Satabdi

2018

Mishra, S. (2018). Cryo-EM structure of the Mycobacterium smegmatis ribosome. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/81569>

<https://doi.org/10.32657/10220/47514>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**
SINGAPORE

**CRYO-EM STRUCTURE OF THE
MYCOBACTERIUM SMEGMATIS RIBOSOME**

SATABDI MISHRA

SCHOOL OF BIOLOGICAL SCIENCES

2019

**CRYO-EM STRUCTURE OF THE
MYCOBACTERIUM SMEGMATIS RIBOSOME**

SATABDI MISHRA

SCHOOL OF BIOLOGICAL SCIENCES

**A thesis submitted to the Nanyang Technological
University in partial fulfilment of the requirement
for the degree of Doctor of Philosophy**

2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

11.01.2019

Date

Satabdi Mishra

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

16/1/2019
.....
Date

.....
Dr. Shashi Bhushan

Authorship Attribution Statement

This thesis contains material from one paper published in the following peer-reviewed journal where I was the first author.

Chapter 3 is published as Mishra, S., T. Ahmed, A. Tyagi, J. Shi and S. Bhushan (2018). "Structures of Mycobacterium smegmatis 70S ribosomes in complex with HPF, tmRNA, and P-tRNA." Scientific Reports 8(1): 13587.

The contributions of the co-authors are as follows:

- Prof S. Bhushan conceived and supervised the project.
- I prepared the manuscript drafts. The manuscript was revised by A/Prof S. Bhushan, and all the authors contributed in the final version of manuscript.
- The sample preparation for microscopy was done by me and A. Tyagi.
- Cryo-EM data collection was done by Prof S. Bhushan and J. Shi in the Cryo-EM Facilities at Nanyang Technological University, Singapore, and National University of Singapore, Singapore.
- I analyzed the data with the help of T. Ahmed.
- All the figures and tables were prepared by me.

11.01.2019
.....
Date

.....
Satabdi Mishra

Acknowledgements

My first and foremost gratitude goes to my supervisor, Prof. Shashi Bhushan for giving me this opportunity and the project towards completion of my dream to become a doctorate. Under his guidance, I got a chance to work on several projects, and learnt a myriad of molecular biology techniques. He has been patient with me through my projects and his supervision has lead me to the end of this PhD journey.

I am thankful to Nanyang Technological University, Singapore, for this opportunity and the research scholarship which has aided me through this journey.

My Thesis Advisory Committee members, Prof. Gerhard Gruber, Prof. Gao Yonggui, and Prof. Oliver Mueller Cajar, whose comments and suggestions have greatly helped in the path and direction of this work.

My thanks go to my lab members, Rahul, my mentor and friend, who has helped me with many of my projects and guided me diligently through years; Tofayel, my colleague who helped me though the protein modeling; Anu, who has helped me in ribosome purification; and Gaurav, Dawn, Zhan, and Toan who have been a part of all the lab fun though these years.

I am thankful to my friends and colleagues at NTU, Barbara, my constant source of positivity through these years; Eugene, a helpful friend and colleague whenever in need; Harvey, Oviya, Harshya, and Vinod, who have been the constant source of entertainment while work. I also thank our cryo-EM manager, Andrew for his help in cryo-EM techniques.

A special thanks to my best friends, Namrata and Nidhi, for being my support, friends who have supported me through the worst days of PhD and celebrated the best results.

My family, who are the blood and breath of me. My Mom who cheers me up before I share my low-feelings. My Dad, who has always taught me that there is nothing in this world out of my reach. And my dear brother, for all his love and charm.

Lastly, I would like to thank my partner, my husband Akash, who has been a trusting partner I can always depend on. This is a joyful moment for him as well as for me.

TABLE OF CONTENTS

Statement of originality	i
Supervisor Declaration Statement	ii
Authorship Attribution Statement	iii
Acknowledgements	iv
Table of contents	v
List of Figures	ix
List of Tables	xii
Abbreviation	xiii
Abstract	xvi
1 Introduction	1
1.1 Scope of Introduction	2
1.2 Mycobacteria	2
1.2.1 Model organisms to study TB	4
1.3 Central dogma of Life	4
1.4 Ribosomes	5
1.5 Structure of Ribosome	6
1.6 Structure of the prokaryotic ribosome	7
1.6.1 Interaction of mRNA and tRNAs with the ribosome	8
1.6.2 Decoding Center	13
1.6.3 Peptidyl Transferase Center	14
1.6.4 Polypeptide exit tunnel	15
1.6.5 L1 stalk on the LSU	17
1.6.6 GTPase-Associated Center (GAC)	18
1.6.7 Inter-subunit bridges between the 50S and 30S	18
1.6.8 Factors that influence translation	20
1.7 Overview of the translation process	20
1.7.1 Initiation	21
1.7.2 Elongation	22
1.7.3 Termination	22
1.7.4 Recycling	23
1.8 Hibernating Ribosome	23

1.8.1	Mechanism of formation of hibernating ribosome dimers (100S ribosome)	24
1.8.2	Physiological significance of hibernating ribosome	26
1.9	Trans-translating Ribosome	27
1.9.1	Structure of tmRNA	28
1.9.2	The mechanism of trans-translation	30
1.9.3	Physiological significance of trans-translation	31
1.10	Mycobacterial Ribosome	31
1.11	Cryo-EM	32
1.12	Aim of this dissertation	36
2	Materials and Methods	37
2.1	Materials	38
2.1.1	Buffers, Media and Salts	38
2.1.2	Microbiology strains	38
2.1.3	Buffer Composition	39
2.1.4	Instruments and accessories	39
2.1.5	Computer tools and softwares	40
2.2	Methods	41
2.2.1	Purification of 70S ribosomes from <i>M. smegmatis</i>	41
2.2.2	Ribosome protein identification by SDS PAGE and Mass Spectrometry	42
2.2.3	Purification of 70S ribosomes from <i>M. smegmatis</i> growing in different growth phases	42
2.2.4	Cryo-Electron Microscopy	43
2.2.5	Cryo-EM Single Particle Analysis (Data Processing)	51
2.2.6	Current capabilities and limitations of cryo-EM and Single Particle Analysis	56
2.2.7	Cryo-EM imaging of <i>M. smegmatis</i> ribosome	59
2.2.8	Model building and refinement	60
2.2.9	Model validation	63
2.2.10	Figure generation	63

3	Results	64
	Cryo-EM structures of <i>M. smegmatis</i> ribosomes in different functional states	65
3.1	Purification profile of <i>M. smegmatis</i> ribosomes in log and stationary phase	66
3.2	Micrograph inspection	66
3.3	Image Processing	69
3.3.1	2D classification	71
3.3.2	3D classification and 3D refinement	71
3.3.3	Resolution estimation and Resmap calculation for the 70S P/P state maps	72
3.3.4	Model Validation	74
3.3.5	Resolution estimation for the Hibernating and Trans-translating maps	78
3.4	Structure of <i>M. smegmatis</i> 70S ribosome in P/P state	78
3.4.1	<i>M. smegmatis</i> small ribosomal subunit	79
3.4.2	<i>M. smegmatis</i> large ribosomal subunit	88
3.4.3	Species specific alteration in the 23S rRNA of the LSU in the <i>M. smegmatis</i> ribosome	93
3.4.4	Species specific alteration in r-proteins clustered near the functional sites of <i>M. smegmatis</i> ribosome	100
3.5	Structure of <i>M. smegmatis</i> 70S ribosome in the Hibernating state	104
3.5.1	Hibernation Promoting Factor	104
3.5.2	Structure of mRNA entry and exit sites of <i>M. smegmatis</i> 70S ribosome	117
3.6	Trans-Translating state of <i>M. smegmatis</i> 70S ribosome	119
3.6.1	Unique mechanism of tmRNA during trans-translation	123
4	Discussion	125
5	Conclusion	130
6	References	132
7	Appendix	155

8	Online access of cryo-EM maps and models	162
9	Author's publications related to this thesis during the Ph.D. period- Aug 2014 to Aug 2018	163

List of Figures

Figure 1.1	The structure of bacterial ribosome	7
Figure 1.2	Structures of prokaryotic mRNA and tRNA	10
Figure 1.3	The codon-anticodon base pairing in a cognate, near-cognate and non-cognate tRNA	11
Figure 1.4	Interactions of A, P, and E - tRNAs with the ribosome	12
Figure 1.5	Decoding center of the ribosome	13
Figure 1.6	Peptidyl transferase center	15
Figure 1.7	Polypeptide exit tunnel	16
Figure 1.8	L1 stalk movement	17
Figure 1.9	The location of the inter-subunit bridges in the ribosome	19
Figure 1.10	An overview of the bacterial translation	21
Figure 1.11	Schematic representation of Ribosome during and after translation	23
Figure 1.12	Mechanism of hibernation and formation of 100S ribosome dimers in <i>E. coli</i>	24
Figure 1.13	Mode of ribosome dimer formation in bacteria	26
Figure 1.14	Secondary structure of <i>T. thermophilus</i> tmRNA	29
Figure 1.15	The mechanism of trans-translation	30
Figure 1.16	Principle of cryo-electron microscopy	33
Figure 2.1	The process of vitrification	44
Figure 2.2	An example of micrograph and respective thon rings	45
Figure 2.3	Behaviour of the electrons in an electron microscope	46
Figure 2.4	The electron microscope	47
Figure 2.5	Computed power spectrum for different types of CTF during imaging in cryo-EM	51
Figure 2.6	A workflow of single particle analysis	52
Figure 2.7	Flowchart of image processing for <i>M. smegmatis</i> ribosome	53
Figure 3.1	Purification profile of <i>Ms</i> ribosomes in log and stationary phase	66
Figure 3.2	FFT showing motion correction	68
Figure 3.3	Micrographs recorded at different defocuses	68

Figure 3.4	Cryo-EM data processing pipeline for the 70S-log dataset	70
Figure 3.5	2D classification of 70S-log dataset	71
Figure 3.6	3D classification of 70S-log dataset	72
Figure 3.7	Estimation of resolution and validation against model-overfitting	73
Figure 3.8	Quality of the P/P state (3.4 Å) and hibernating state (4.1 Å) maps	74
Figure 3.9	Local agreement between the cryo-EM maps and models	77
Figure 3.10	Resolution estimation for the trans-translation and hibernation states 70S ribosome maps	78
Figure 3.11	Unique structural features of the <i>Mycobacterium smegmatis</i> 70S ribosome	79
Figure 3.12	Architecture of the <i>Ms</i> SSU	80
Figure 3.13	Comparison of <i>Ms</i> SSU r-proteins with the <i>Ec</i> counterparts	83
Figure 3.14	<i>Ms</i> r-protein uS5	84
Figure 3.15	r-protein bS16 interaction with uS4 in <i>Ms</i> 30S	85
Figure 3.16	r-protein uS17 interactions with 16S rRNA in <i>Ms</i> 30S	86
Figure 3.17	Comparison of <i>Ms</i> mRNA exit site with <i>Ec</i>	87
Figure 3.18	<i>Ms</i> 16S rRNA secondary structure comparison with <i>Ec</i>	88
Figure 3.19	Architecture of the <i>Ms</i> LSU	89
Figure 3.20	Comparison of <i>Ms</i> LSU r-proteins with the <i>Ec</i> counterparts	91-92
Figure 3.21	<i>Ms</i> bL25 with its two functional domains	93
Figure 3.22	<i>Ms</i> 16S rRNA showing H15-H16a	94
Figure 3.23	Depiction of H15-16a with <i>Ms</i> bL9	95
Figure 3.24	<i>Ms</i> 16S rRNA showing the H31a insertion	96
Figure 3.25	<i>Ms</i> 16S rRNA showing the H54a insertion	96
Figure 3.26	Species-specific modifications near H54a	97
Figure 3.27	H54a movement comparison between P/P state, hibernating state and <i>Mtb</i> 50S	99
Figure 3.28	Location of the <i>Ms</i> bS22 in its rRNA environment	101
Figure 3.29	A comparison of the bS22 with its eukaryotic counterparts	102
Figure 3.30	Location of <i>Ms</i> bL37 in the rRNA pocket and comparison with other ribosomal proteins	103

Figure 3.31	Structure of the hibernating state of <i>Ms</i> 70S ribosome	104
Figure 3.32	Sequence alignment of HPF from <i>Ms</i> , <i>Sa</i> , and <i>Tth</i>	105
Figure 3.33	Location of <i>Ms</i> HPF	106
Figure 3.34	Interaction of HPF with the tRNAs and the 16S rRNA residues	107
Figure 3.35	Ribosome purification under high and low salt conditions for <i>Ms</i> cultures harvested at different growth periods	108
Figure 3.36	Comparison of S1 structure in association with the ribosome as captured in different studies	110
Figure 3.37	Depiction of the four OB domains of <i>Ms</i> bS1 interacting with uS2	111
Figure 3.38	Analysis of the extra density near the mRNA exit site in the hibernating state map	112
Figure 3.39	bS1 density in the P/P state 70S ribosome	112
Figure 3.40	SDS PAGE and Mass spectrometry analysis of bS1 protein from stationary phase 70S ribosomes	113
Figure 3.41	Comparison between <i>Ms</i> HPF, <i>Sa</i> HPF and <i>Ll</i> HPF	115
Figure 3.42	Architecture of mRNA entry and exit sites of the <i>Ms</i> 70S ribosome	119
Figure 3.43	Structure of the trans-translating state of <i>Ms</i> 70S ribosome	120
Figure 3.44	Secondary structure diagram of <i>Ms</i> tmRNA	121
Figure 3.45	Architecture of tmRNA and SmpB when bound to the 70S	122
Figure 3.46	Comparison of the structure of <i>Ms</i> tmRNA with <i>Tth</i> tmRNA	123
Figure 4.1	Two new r-proteins of mycobacterial ribosome	127

List of Tables

Table 1.1	List of the inter-subunit bridges in bacteria	19
Table 3.1	Model refinement statistics for the <i>Ms</i> 50S and 30S ribosome subunits and the complete 70S ribosome in P/P state	75
Table 3.2	Composition of P/P state SSU model	81
Table 3.3	Composition of P/P state LSU model	90

Abbreviations

2D	Two-dimensional
3D	Three-dimensional
A	Adenine
Å	Angström
Anti-SD	Anti-Shine-Dalgarno
BCG vaccine	Bacillus Calmette-Guerin vaccine
C	Cytosine
CCD	Charge-coupled device
CMOS	Complementary metal-oxide semiconductor
Cryo-EM	Cryo Electron Microscopy
CTD	C-terminal Domain
CTE	C-terminal Extension
CTF	Constant Transfer Function
DC	Decoding Centre
DED	Direct Electron Detector
DNA	Deoxy ribonucleic acid
DQE	Detective quantum efficiency
<i>Ec</i>	<i>Escherichia coli</i>
EFG	Elongation Factor G
EF-Tu	Elongation Factor ‘temperature control’
EM	Electron Microscope
EMDB	Electron Microscopy Data Bank
FEG	Field Emission Gun
FSC	Fourier shell correlation
G	Guanine
GAC	GTPase Associated Center
GTP	Guanosine triphosphate
H	Helices
HIV	human immunodeficiency virus
HPF	Hibernation Promoting Factor
IC	Initiation Complex

IF	Initiation Factor
I-TASSER	Iterative threading assembly refinement
kDa	kilo Dalton
kV	kilo Volt
LB broth	Luria-Bertani broth
<i>Ll</i>	<i>Lactococcus lactis</i>
LSU	Large Subunit
MDa	Mega Dalton
MLD	mRNA like Domain
mRNA	messenger RNA
<i>Ms</i>	<i>Mycobacterium smegmatis</i>
<i>Mtb</i>	<i>Mycobacterium tuberculosis</i>
nt	nucleotides
OD	Optical density
PDB	Protein Data Bank
PIC	Pre-Initiation Complex
PK	Pseudoknots
PSF	Point Spread Function
PSRP	Plastid-specific ribosomal protein
PTC	Polypeptide Transfer Centre
RELION	Regularised likelihood optimisation
RF	Release Factor
RMF	Ribosome Modulation Factor
RNA	Ribonucleic acid
RNA pol	RNA polymerase
r-protein	ribosomal protein
RPM	Revolutions per minute
RRF	Ribosome Release Factor
rRNA	ribosomal RNA
Rsfs	Ribosome Silencing factor during Starvation
MLD	mRNA like Domain
NMR	Nuclear Magnetic Resonance
NTD	N-terminal Domain
NTE	N-terminal Extension

S	Svedberg's unit
<i>Sa</i>	<i>Staphylococcus aureus</i>
SD	Shine Dalgarno
SEM	Scanning electron microscope
SmpB	Small Protein B
SNR	Signal to Noise Ratio
SPA	Single particle analysis
SRA proteins	Stationary phase induced Ribosome Associated proteins
SRL	Sarcin Resin Loop
<i>ssrA</i>	Small Stable RNA
SSU	Small Subunit
T	Thymine
TB	Tuberculosis
TEM	Transmission electron microscope
TLD	tRNA like Domain
tmRNA	transfer-messenger RNA
tRNA	transfer RNA
<i>Th</i>	<i>Thermus thermophilus</i>
U	Uracil
UTR	Untranslated Region
WHO	World Health Organisation

Abstract

Ribosomes are the universal protein synthesis machineries in all living cells. These macromolecular structures can range from 2.5 (prokaryotes) to 4.2 MDa (eukaryotes) in size, undergo various different and precise movements to synthesis the nascent polypeptide chain. Ribosomes in the cells can exist in various functional states. Thus, to understand the process of translation completely, we must learn about these functional states as well. Although, many X-ray and high-resolution structures are available for ribosomes from different bacteria, a high-resolution structure of ribosome from *Mycobacteria* was lacking at the time this dissertation work started. Using cryo-electron microscopy and single-particle reconstruction, I have determined the structure of *Mycobacterium smegmatis* ribosome at its P/P state, characterized by the tRNA bound at the P-site to the resolution of 3.4 Å. The high-resolution structure has presented various insights about the species-specific rRNA and r-protein extensions. Along with the P/P state, two different functional states of *M. smegmatis*, the hibernating state (at 4.1 Å) and trans-translating state (12.5 Å) of 70S ribosomes were also investigated in this work. A possible theory for mycobacterial ribosome resuscitation from inactive ribosomal state is also discussed. The comparison of the three functional states of the mycobacterial ribosome will help in understanding the structure, function, and evolution of *Mycobacteria*. These studies will also help to understand the translation process in mycobacteria which could aid in discovery of novel targets against the pathogenic mycobacterial strains, a step towards the cure for Tuberculosis.

Chapter 1

INTRODUCTION

1 Introduction

1.1 Scope of Introduction

The introduction in this thesis aims to provide a summarized overview of various topics. The chapter starts with an overview of Mycobacteria and why the pathogenic strains from this genus have been extensively studied. It is followed by an overall understanding about the ribosomes, ribosome structures, and an overview of the translation in the bacteria. Next, the chapter explains two different ribosome functional states observed in the bacteria, the hibernating state and the trans-translating state, along with their physiological importance. This follows with an introduction about the mycobacterial ribosomes, and the significance of its study. And finally, an introduction about the technique cryo-electron microscopy, which is the primary technique used for the structural determination of the mycobacterial ribosome in this dissertation. The introduction is followed by the aim of this dissertation, summarizing the main objectives of this work.

1.2 Mycobacteria

Mycobacteria belongs to the phylum Actinobacteria, known for their high morphological diversity [1]. The genus *Mycobacterium* included many pathogenic strains, most commonly known two strains, *Mycobacterium tuberculosis* (*Mtb*), the pathogen causing tuberculosis (TB), and *Mycobacterium leprae*, which causes leprosy [2].

The DNA of *Mtb* has been discovered in Egyptian mummies from 2000 B.C. [3] and TB was described by Hippocrates as early as 400 B.C. [4]. Historical texts identify the disease as “consumption,” “wasting away,” “king’s evil,” “lupus vulgaris,” “the white plague” or “phthisis” based on its clinical manifestations [4, 5].

TB epidemiology

TB has now existed for a millennium and yet remains one of the major global health issue. TB caused one in four deaths during industrial revolution [6]. About 10 million new cases of TB are registered in the world annually; 30 % of these can be found in India and China and 80 % in the 20-25 highest-burden countries, preferentially in Africa, South America, and Asia [7]. Nearly two

million persons die from TB each year, and it is estimated that one third of the world's population is infected with *Mtb* [8-10]. Most individuals who are exposed to the pathogen, however, do not develop disease. It is believed that these individuals (up to 50 % of those exposed) clear the infection through a robust innate immune response. It is difficult to firmly establish this number, however, as these individuals do not show signs of disease and have no immunological memory against the pathogen, giving a negative result in diagnostic tests based on the presence of memory T cells. On the other hand, clearance through an adaptive immune response or the establishment of a latent TB infection, which occurs in the remaining 50 % of individuals, will leave primed T cells behind, whose response can be detected. This means that it is difficult to distinguish latent TB infection from a resolved infection using a test based on immunological memory. Furthermore, only 5 % of latently infected individuals go on to develop active TB within five years, while the remaining 95 % contain their latent TB throughout their lifetime; only developing active TB if immunocompromised, e.g. through simultaneous HIV-infection, treatment with immunosuppressive drugs, old age, or through re-infection [10, 11]. The current vaccine, Bacillus Calmette-Guerin (BCG) vaccine provides immunity against TB and leprosy in young children, but is not 100% efficient against pulmonary TB in adults [12]. Current antibiotics developed against TB target the actively replicating bacteria, but they are not efficient against the latent stage TB. A latent TB is the absence of clinical TB symptoms, yet showing a positive reaction to the TB test, which is the reaction to the purified protein derivative (PPD) skin test [13]. The first-line drugs against TB include rifamycin, isoniazid, ethambutol and pyrazinamide and needs to be administered for the period of six months. Due to its long length of treatment or reduced patient compliance or antibiotic inefficiency, genetic antibiotic resistance has been developed in TB. Today, latent TB and TB reactivation are one of the major public health concerns, which requires new and novel antimicrobial strategies against the TB pathogens.

Majority of the antibiotics against bacteria target the ribosome of the cells [14, 15]. Ribosomes are the protein synthesis machineries of the cell [16, 17]. In addition to this, the cell viability also greatly depends on the translation efficiency of the cell. Therefore, it is essential to determine the structure of

mycobacterial ribosome and understand the translation mechanism of *Mycobacteria*.

1.2.1 Model organisms to study Tuberculosis

The organism *Mtb* is a slow-growing human pathogen, and thus is extremely challenging and labour-intensive to study. It is classed as a biosafety contaminant level 3 organism, which further raises its limitation to be studied in a research laboratory. Since it has to be handled in a dedicated facility within biological containment cabinets, the physical constraints of this containment affect the sample sizes and type of experiments that can be carried out with *Mtb* in the average laboratory. Furthermore, *Mtb* grows very slowly doubling approximately once every 24 hours, which lengthens the time of the most experiments considerably.

Model organisms used for the study of *Mtb* overcome such limitations. A non-virulent strain from the genus *Mycobacterium*, *Mycobacterium smegmatis* (*Ms*), is a saprophytic fast-growing *Mycobacterium* and possess no health and safety risk under a research facility. It offers a convenient proxy for the study of *Mtb* physiology. However, *Ms* is genetically distinct from *Mtb*. This caused a significant debate in the tuberculosis community regarding its value as a model organism [18][19]. Regardless, if used carefully, *Ms* can provide important information on specific aspects of *Mtb* physiology and therefore the most preferred model in the study against TB [20].

1.3 Central dogma of Life

Life as we know it continues to exist based on its ability to replicate and pass on its genetic information to the next progeny. The genetic information of an organism is stored in deoxyribonucleic acid (DNA), which gets duplicated and get transferred to the next generation, making the DNA carrier of genetic information in all organisms [21]. The DNA exists in a double-stranded helix conformation composed of nucleotides. Each of the nucleotide has one of the four bases (adenine (A), thymine (T), guanine (G), or cytosine (C)), a sugar called deoxyribose, and a phosphate group. The nucleotides are joined together into a chain by covalent bonds between the sugar of one nucleotide and the phosphate of the next nucleotide, which results in an alternating sugar-phosphate backbone.

The nitrogen bases of the two strands are stabilized by the hydrogen bonds, following the base pairing rules (A-T, G-C) to make the double-stranded DNA. The genetic information in the DNA must be 'read' for the synthesis of proteins. The RNA polymerases (RNA pol) use the double stranded DNA as a template to generate RNA in the process called transcription. The RNA is used in the next step, called translation, to synthesize proteins. The two interdependent steps, transcription and translation constitute the 'central dogma of life'. There are three different types of RNA polymerases (RNA pol) which synthesize three different types of RNAs. The messenger RNA (mRNA), which guides the polypeptide sequence synthesized by the ribosomes, is synthesized by the RNA pol II. Every three nucleotides in the mRNA sequence codes for an amino-acid. The RNA pol III synthesizes transfer RNA (tRNA). tRNA functions by carrying the specific amino-acid encoded by the mRNA and incorporates it into the growing polypeptide chain in the ribosome. The RNA pol I carries out the transcription of the ribosomal RNA (rRNA) which forms the catalytic component of the macromolecular structure responsible for translation, the ribosome.

1.4 Ribosomes

Ribosomes were first viewed in the tissue section under the electron microscope. They appeared as dense granules and were called as the "ribonucleoprotein particles of the microsome fraction" [22-24] and were later termed as "ribosome". Ribosomes are the translating machineries of the cells, which translate the information in the mRNA into functional proteins [16, 17]. These macromolecules have been the interest of the structure biologists for a very long time because of their unique precision in conversion of the genetic information stored in DNA into functional proteins. The importance of ribosome study came to limelight in the scientific world when the three pioneers in the field of ribosome structure, Thomas A. Steitz, Venkatraman Ramakrishnan, and Ada E. Yonath were awarded Nobel Prize in chemistry in the year 2009 for their contribution in the study of structure and function of Ribosome.

Over the period of evolution, the ribosomes have also evolved with various unique species-specific characteristics, maintaining the underlying function of translation of mRNA to synthesize proteins. Ribosomes being the only mechanism of the cell to synthesize proteins, are also vital for the cell survival

and cell duplication. This fact is exploited in the pathogenic bacteria and one of the major reason why half of the antibiotics in use today target ribosomes [25, 26]. Thus, a complete understanding of these macromolecules along with their unique species-specific features is an absolute necessity from pathogenic bacteria, in the process of discovering novel targets for antibiotics.

1.5 Structure of Ribosome

The ribosomes are made up of ribosomal RNA (rRNA) which contribute two-third of the ribosome mass and the rest one-third mass is made up of ribosomal proteins (r-proteins). In all living cells, the ribosomes consist of two parts, a large subunit (LSU) and a small subunit (SSU) (Figure 1.1). Both the subunits of ribosome are responsible for carrying out unique functions during translation. In archaea and eubacteria, the small subunit (30S) and the large subunit (50S), together form the complete 70S ribosome complex, while in eukaryotes these counterparts are much larger in size and are the 40S, 60S, and 80S, respectively. Looking at the conserved functions in general, the SSU is responsible for the decoding of the mRNA sequences, and the LSU catalyzes the peptide bond formation and also hosts the tunnel through which the nascent polypeptide exits the ribosome. In bacteria, the LSU is composed of two rRNA molecules, the 23S rRNA (~2300 nucleotides (nt)) and the 5S rRNA (~120 nt), and the SSU has one, the 16S rRNA (~1500 nt). The r-proteins (~55 r-proteins) in bacteria range from approximately 4-60 kDa in size. The rRNAs and the r-proteins together make the entire ribosome mass about 2.5 MDa in bacteria.

A complete understanding about the structure of ribosomes was obtained when the first atomic structure of the of the 30S subunit from *Thermus thermophilus* was reported at 3.0 [27], and at 3.3 Å [28], and the 50S subunit from *Haloarcula marismortui* was reported at 2.4 Å [29]. This was followed by the structure of the complete 70S ribosome from *Thermus thermophilus* at 5.5 Å [30]. Along with these structures, the structure of the large subunit from *Deinococcus radiodurans* [31], the structure of 70S ribosome from *Thermus thermophilus* [32, 33], and the structure of 70S ribosome from *E. coli* [34] are considered as the foundation ribosome structures in the field of study of ribosomes and translation. After several structural reports on ribosomes from higher eukaryotes, it was also understood how the evolution in the eukaryotic

ribosomes have prevented them against the antibiotics which are capable of targeting and creating severe malfunctions in the bacterial ribosome (reviewed in [35]).

Figure 1.1 The structure of bacterial ribosome. (A) The top view of the 70S ribosome bound with A-, P-, and E-tRNAs (Pink, Green, and Yellow, respectively). The mRNA is shown in black. (B) The side view of the 30S, showing the head, beak, body, and spur. (C) The subunit interface view of the 50S, with the functional sites labelled, along with the structure and location of the L7/L12 arm. Adapted from [36].

1.6 Structure of the prokaryotic ribosome

The structure of ribosome from the prokaryotes has been studied extensively using X-ray crystallography and cryo-EM. The entire understanding of the structure and function of the ribosome can be divided into three important

aspects, (i) key players of translation, the mRNA and the tRNAs, and how they interact with the ribosome in a fast and precise manner during translation, (ii) study of the functional domains, the Decoding Center (DC) on the 30S, the Peptidyl Transferase Center (PTC), the L1 stalk, and the GTPase-Associated Center (GAC) on the 50S, and (iii) study of the inter-subunit bridges between the 50S and 30S, these inter-subunit bridges are important to keep the 70S ribosome intact and guide the various steps of translation in the ribosome. Apart from these parameters, which includes the ribosome, the mRNA, and the tRNAs, there are several external factors which influence translation in the bacteria. These translational factors and their location in the ribosome are discussed first in the upcoming sections, before we move to the mechanism of translation in ribosome and the various functional states of the ribosome present in the cell.

1.6.1 Interaction of mRNA and tRNAs with the ribosome

mRNA: The mRNA carries the genetic information that is transformed into the functional polypeptides during translation in the ribosome. The genetic information of the mRNA exists in the form of codons, comprising of three nucleotides. With four nucleotides (A, U, G, and C; U: uracil), there are 64 different triplet combinations possible. Out of these 64 possibilities, 61 codons specify individual amino-acids. Thus, the genetic code is degenerative and few of the amino-acids are coded by more than one codon. Most often, the process of translation starts with the start codon, AUG, which codes for methionine. There are three codons (UGA, UAA, and UAG) which do not code for any amino-acid, and code the termination signal in the mRNA, causing the translation to stop when the ribosomes encounter any of these codons. Therefore, the ribosomes translate the region between the start codon and the termination codon, and this stretch of the mRNA is termed as the *reading frame*.

The mRNA (~30 nucleotides) enfolds around the ‘neck’ or ‘beak’ region of the 30S (shown in black/grey color in figure 1.1). Interestingly, the entire mRNA molecule is not translated by the ribosome. The sequence near the 5’ end of the mRNA is a stretch of untranslating sequence, known as the untranslated region (UTR) or the leader sequence (Figure 1.2 A). The leader sequence plays an important role in the assembly of the translation initiation complex. It contains the ribosome binding site, which in bacteria is also known as the Shine-Dalgarno

sequence (SD sequence) [37]. A similar stretch of sequence is also observed in the vertebrates and called as the Kozak box [38]. The SD sequence gets stabilized at the cleft between the 'head' and 'beak' of the 30S [30]. The SD sequence near the mRNA exit site interacts with the 3' end of the 16S rRNA (which harbors the anti-SD sequence). The SD-antiSD interaction between the mRNA and 16S rRNA is crucial for the correct positioning of the start codon of the mRNA at the P-site in the SSU (discussed on the next page).

tRNAs: To understand translation, it is important to understand the interaction of the ribosome with its substrates, the tRNAs. The tRNAs serve as the physical link between the nucleotide sequence of the mRNA which is translated into the amino-acid sequence of the nascent polypeptide synthesized. This single strand of the tRNA, which is about 75-95 nucleotides long, adapts a L-shaped structure with various structural loops (Figure 1.2 B and C). The tRNA anticodon loop interacts with the small ribosomal subunit, while the CCA tail (acceptor stem) interacts with the large ribosomal subunit.

Figure 1.2 Structures of prokaryotic mRNA and tRNA. (A) The depiction of a prokaryotic mRNA, showing the arrangement of the start codon, the Shine-Dalgarno sequence located upstream of the protein-coding sequence, and the stop codon. The mRNA also ends with a stretch of non-coding sequence at its 3' end. (B) A secondary structure of the prokaryotic tRNA, showing the cloverleaf like structure of the tRNA. (C) A model of the tRNA with different loops of tRNA highlighted in different colors as labelled. Figures modified from [39-41].

There are three different types of tRNAs and they have their unique binding sites in the ribosome. The A-site on the ribosome binds with the incoming aminoacyl-tRNA or A-tRNA; the P-site stabilizes the peptidyl-tRNA or P-tRNA which is attached to the growing nascent polypeptide chain, and the E-site (exit-site) where the deacylated P-tRNA or E-tRNA localizes and after the formation of the peptide bond and before its ejection from the ribosome (Figure 1.1). The 3' ends of the three tRNAs extend towards the 50S, with the A- and P-tRNA located in close proximity of the PTC, while the 3' end of the E-tRNA is almost 50 Å away from the PTC of the 50S [30, 36].

Interactions at the A-site. The A-tRNA anticodon loop interacts with the DC of the 30S. The universally conserved 16S rRNA bases (G530, A1492, and

A1493, *E. coli* numbering) in the DC interact extensively with the first two codon-anticodon base pairs [42] (Figure 1.4). The interaction at A-site is important for the discrimination between a cognate tRNA (having all the three Watson-crick base pairs matching between the mRNA codon and the tRNA anticodon), a near cognate tRNA (with one mismatch out of the three codons), and non-cognate tRNA (with two or three mismatches). A representation of the matching of cognate, near-cognate, and non-cognate tRNA with the mRNA codons is given in the figure 1.3.

Figure 1.3 The codon-anticodon base pairing in a cognate, near-cognate and non-cognate tRNA. Cognate tRNA (yellow), with two watson-crick base pairing between the codon-anticodon and one mismatch at the wobble position (third) position is accepted. A near cognate with two acceptable codon-anticodon base pairing is rejected because of the mis-match at the first position. A non-cognate (purple) with no codon-anticodon matches. Figure adapted from [43].

These interactions are responsible for the binding of a correct aminoacyl-tRNA, complement to the mRNA sequence at the P-site, so the correct amino-acid could be added to the growing nascent polypeptide chain. The r-protein S12 interacts with the backbone on the 30S side. On the 50S side, the A-tRNA interacts with the r-protein L16 and the Helix 38 (H-helix of 23S rRNA, h-helix of 16S rRNA) of the 23S rRNA. The H38 is thus known as the A-site finger because of its interaction with the A-site tRNA.

Interactions at the P-site. Several 16S rRNA bases and the 30S r-proteins S9 and S13 interact extensively with the anticodon stem-loop of P-tRNA (Figure 1.4 B) [32, 33, 44]. U1498 of 16S rRNA interacts with the P-site codon, and is the first position of interaction between the P-tRNA with the 16S rRNA. The 16S rRNA base U1498 (*E. coli* numbering) interacts with the P-site codon first, while

its phosphate group interacts with the G926 (*E. coli* numbering) of 16S rRNA. the P-tRNA in addition to this also interacts with G966, A1229, C1230, G1338, and A1339 (*E. coli* numbering) and stacks with C1400 (*E. coli* numbering) of the 16S rRNA. Mutation on these 16S rRNA bases leads to severe decrease (up to 20-fold) in the translation activity [45, 46]. On the 50S side, the elbow of the P-tRNA locates very closely to the r-protein L5 [47]. The P-tRNA also interacts with the H69 in the 50S.

Interactions at the E-site. The E-tRNA on the 30S side interacts with the r-protein S7 and S11 (Figure 1.4 C). The E-tRNA is loosely regulated by the 16S rRNA but interacts exclusively with the 23S rRNA on the 50S side. The 23S rRNA bases G693 and A695 (*E. coli* numbering) interact first with the E-tRNA. The acceptor arm of E-tRNA shares extensive interactions with the 23S rRNA on the 50S subunit side via the bases G2421, A2422, C2394, C1892, and C1893, *E. coli* numbering). The acceptor arm of E-tRNA also interacts with the r-proteins L28 and L33. In addition to this, the E-tRNA also interacts with the L1 stalk extensively.

Figure 1.4 Interactions of A, P, and E - tRNAs with the ribosome. (A) The P-site tRNA interaction with the 16S rRNA (yellow) and with the r-proteins L5, L16, L27, S9, and S13 (orange). (B) The A-site tRNA interaction with the 16S rRNA (yellow) and with the r-proteins L16, L27, S12, and S13 (orange). (C) The E-site tRNA interaction with the 16S rRNA (yellow) and with the r-proteins L1, L28, L33, S7, and S11 (orange). The specific bases are colored in red and mRNA is shown in black. Figure modified from [48].

1.6.2 Decoding Center

The DC is located at the neck region of the 30S, just above the mRNA channel, where the A-site of the ribosome is located (Figure 1.1 B). As the name suggests, the decoding of the mRNA takes place at this site, and it ensures that the translation is carried out with the correct reading frame. 16S rRNA helices, h44 and h34 form the DC, with the universally conserved nucleotides A1492, A1493, and G530 (*E. coli* numbering) at this site which are essential for the recognition of cognate (correct) codon-anticodon interaction (Figure 1.5 B) [49-52]. The DC not only oversees the translocation of the mRNA and tRNA, it also controls the fidelity of the codon-anticodon interaction at this site [16]. The interactions between the mRNA, tRNA and ribosome at the DC is well conserved through evolution (reviewed in [43]).

Figure 1.5 Decoding center of the ribosome. (A) The decoding centre at the 30S. Interaction of the A-site tRNA with the DC. (B) interaction of a cognate tRNA (orange) with the universally conserved residues A1492, A1493, and G530 of 16S rRNA (highlighted in red) at the DC. (C) Schematic representation of the binding of ternary complex at the DC. Figures modified from [43, 53, 54].

During translation, the decoding (a part of elongation) begins with an empty A-site and a paired tRNA at the P-site. Decoding of the mRNA is followed by the binding of the correct ternary complex (aminoacyl-tRNA/EF-Tu/GTP) at the A-site of the ribosome. If the aminoacyl tRNA's anti-codon matches with the mRNA codon at the A-site, the aminoacyl tRNA binds to the A-site, which is catalysed by GTP hydrolysis by EF-Tu (Elongation Factor Tu) (Figure 1.5 C) [53]. EF-Tu then dissociates from the ribosome and the aminoacyl end of the A-

tRNA then moves to the PTC (accommodation) for peptide bond formation. This movement of the tRNA from the A-site to P-site is assisted by the rotation of the SSU. The SSU near the DC stays in an open conformation until the binding of the cognate tRNA takes place at the A-site, which in turn causes the shoulder and head of the SSU to rotate upwards towards the subunit centre [43], facilitating the accommodation of the tRNA at the P-site. This unique movement of the SSU is also known as the 30S domain closure. The three conserved residues, A1492, A1493, and G530 (*E. coli* numbering), change their stoichiometry at the DC to accommodate the 30S domain closure and stabilize the codon-anticodon interaction [43, 49]. Though believed that the domain closer is specific with the cognate recognition of the tRNA, it can be achieved in the presence of the antibiotic paromomycin with near-cognate tRNAs.

1.6.3 Peptidyl Transferase Center

The peptidyl transferase center is the site of the catalytic reaction where the peptide bond is formed between the amino-acids in the growing polypeptide chain [55]. It is located in the 50S and is formed by the 23S rRNA residues U2506, G2583, U2584, and U2585 (*E. coli* numbering) (Figure 1.1 C). The new peptide bond is formed between the α -amino group of the aminoacyl-tRNA and the ester carbon of the peptidyl-tRNA (Figure 1.6) [36, 56-58]. As a result of the formation of peptide bond, the tRNA is left deacylated at the P-site (Figure 1.6 C), and is translocated to the E-site before being ejected from the ribosome in the next step of translation. The CCA ends of both the A- and P-tRNAs exist near the close vicinity of the PTC, along with domain V of the 23S rRNA. The interaction of the 23S rRNA at the PTC with the A- and P-tRNAs is universally conserved (reviewed in [59]). Although, no r-protein directly interferes with the catalytic peptide bond formation, two r-proteins L27 and L16 are present very closely to the vicinity of the PTC [59]. The N-terminus of the r-protein L27 is located close to the CCA ends of the A- and P-tRNAs and is believed to interact with the tRNAs [33, 43, 60]. And the L16 interacts with the acceptor arm of the A-tRNA [44, 61].

Figure 1.6 Peptidyl transferase center (A) Location of the PTC in the 50S, model is shown with the three tRNAs; A-tRNA (red), P-tRNA (green), and E-tRNA (yellow) (B) Nucleophilic attack during the catalytic reaction of peptide bond formation. (C) Chemical reaction that follows during the peptide bond formation. Figures are adapted from [36, 62].

1.6.4 Polypeptide exit tunnel

The growing nascent polypeptide chain exits the ribosome through the polypeptide exit tunnel located in the LSU. This universally conserved tunnel is approximately 80 Å in length and 10-20 Å in diameter. Unlike the functional domains, DC and PTC, the tunnel is composed of both the r-proteins and rRNAs. The 23S rRNA constitutes along with L4, L22, L17 (only in eukaryotes), and L23 to form the wall of this tunnel.

Figure 1.7 Polypeptide exit tunnel. (A) Bacterial polypeptide exit tunnel, which is about 20 Å wide and 100 Å long, is trapped by the various rRNA bases and r-proteins. The nucleotide bases shown in orange constitute the PTC. The nucleotides in pink (U2609, A752, and A751) and red (A2062, A2503, and A2058) are essential for ribosome stalling. The nucleotides in cyan (U2586 and U1782) are responsible to connect the nucleotides U2585 and U2609. The white arrow indicates the possible pathway of the nascent polypeptide. (B) A transverse section of the polypeptide exit tunnel revealing the rRNA, P-tRNA, and r-protein components. Figures modified from [63, 64].

The tunnel was initially believed to act as a transport channel for the newly synthesized proteins, but it has also shown to be involved in protein folding and translation regulation. Overall the tunnel possesses an electronegative potential, which could cause the stalling or slow down translation by interacting with the polypeptides with long stretches of charged residues (like arginine or lysine). Translation stalling is carried out by the eukaryotic specific r-proteins in the tunnel, L4 and L17, while in prokaryotes it is carried out by the L22 [65]. The translation stalling is another method for regulation of translation, or efficient targeting of the membrane proteins (Figure 1.7 A).

The nascent polypeptide emerges from the ribosome at the polypeptide exit site. In bacteria, the exit site is flanked by three major r-proteins, L22, L23, L24, and L29 [65, 66]. The exit site is also the site of interaction of various post-translation enzymes or chaperones. To list a few examples, the trigger factor (a bacterial protein chaperone) interacts with L23 and L29 at the exit site to assist with the proper folding of the emerging protein, an enzyme called Deformylase binds to the emerging N-terminal of the protein in bacteria to cleave the formyl

group bound to the first methionine residue, and Zuotin (in eukaryotes) interacts with L32 (eukaryotic specific protein at the peptide exit site) and assists in the co-translational folding of the new polypeptide.

1.6.5 L1 stalk on the LSU

The L1 stalk (of the LSU) consists of the conserved 23S rRNA helices H76-78 and the r-protein L1 [29-31, 67-69]. It is a highly dynamic stalk is located near to the E-site on the LSU. Various structural analysis (both by X-ray and cryo-EM) reveal that the L1 stalk can exist in an 'open' and 'closed' conformation [32, 33, 70, 71]. The 'open' conformation happens when the ribosome is at the classical A/A state (with the A-tRNA at the A-site, P-tRNA at the P-site, and/or empty E-site), whereas during the 'closed' conformation, the ribosome exist in its rotated state containing the A/P and P/E hybrid tRNAs (Figure 1.8). In its closed conformation, the L1 stalks contacts the P/E hybrid tRNA. The displacement from the 'open' to 'closed' conformation would be about 30° (when viewed from the inter-subunit interface) (reviewed in [44, 72]). The dynamic nature of the L1 stalk is coupled with the inter-subunit rotation of the ribosome subunits and movement of tRNAs [73]. Thus, it is assumed to be involved in assisting the tRNA translocation and also ejection of the tRNA from the E-site.

Figure 1.8 L1 stalk movement. The movement of L1 stalk between 'open conformation' (A) and 'closed conformation' (B). The E-site is empty in the open conformation, whereas in the closed conformation, the L1 stalk interacts with the P/E tRNA. L1 protein – magenta, L1 stalk – shades of brown, 16S rRNA – cyan, 23S rRNA – light brown, 5S

rRNA – light purple, 50S r-proteins – pink, 30S r-proteins – dark blue. Figure is modified from [74].

1.6.6 GTPase-Associated Center (GAC)

The GTPase-Associated Center is where the GTPase translation factors bind to the ribosome (Figure 1.1 C, labelled as ‘GTPase factor binding site’). The GTPase factors include the initiation factor 2 (IF2), Elongation Factor ‘temperature unstable’ (EF-Tu), EF-G, release factor 3 (RF3), LepA (also known as EF4) and BipA. The GAC also includes the sarcin-resin loop (SRL) of the 23S rRNA and the r-proteins L10, L11, and the L7/L12 stalk. The GAC has been verified by various biochemical analysis [75-79]. Its location is visualised in cryo-EM structures [80-86] and X-ray structures [29, 30, 87-94] of ribosomes bound with these GTPase factors. The SRL of Helix 95 (nucleotides 2653-2667, *E. coli* numbering) is primarily responsible for the elongation cycle in the ribosome translation [95]. During decoding, the distortion of the aminoacyl-tRNA due to the codon-anticodon recognition near the DC signals the EF-Tu of the GAC, which is located about 80 Å away, to hydrolyse the binding of the aminoacyl-tRNA at the A-site [96, 97].

1.6.7 Inter-subunit bridges between the 50S and 30S

Inter-subunit bridges are the connections shared between the 50S and the 30S in the 70S ribosome complex (Figure 1.9). In bacteria, there are 12 highly conserved bridges (table 1.1) [30]. There could be more species-specific inter-subunit bridges, based on insertions in their rRNAs and r-proteins. These inter-subunit bridges maintain the ribosome assembly in its 70S form, and any disruption in these is indispensable in ribosome functioning. The accurate formation of these bridges requires the formation of a functional 70S initiation complex. During elongation, several of these bridges undergo reconfiguration to accommodate the inter-subunit rotations for the translocation of the mRNA and tRNAs.

Figure 1.9 The location of the inter-subunit bridges in the ribosome. The inte-subunit bridges for 50S (A) and 30S (B) as viewed from the subunit interaface. The bridges are coloured in magenta and orange. The other components are coloured as: 23S rRNA (grey), 50S r-proteins (brown), 16S rRNA (cyan), 30S r-rproteins (dark blue), P-tRNA (red), and A-tRNA (green). Figure adapted from [48].

Bridge	Type	30S subunit		50S subunit	
		16S rRNA helix or S protein	RNA or protein positions	23S rRNA helix or L protein	RNA or protein positions
B1a	Prot-RNA	S13	92–94	H38(L)	886–888
B1b	Prot-Prot	S13	NH ₂ -term	L5	134–153
B2a	RNA-RNA	H44(m)	1408–1410, 1494–1495	H69(Lm)	1913–1914, 1918
B2b	RNA-RNA	H24(m,LM)	784–785,794	H67(m), H69(M)	1836–1837, 1922
	RNA-RNA	H45(LM,Lm)	1516–1519	H71(M), H69(B)	1919–1920, 1932
B2c	RNA-RNA	H24(Bm)	770–771	H67(B)	1832–1833
	RNA-RNA	H27(Bm)	900–1	H67(B)	1832–1833
B3	RNA-RNA	H44(m)	1484–1486	H71(m)	1947–1948, 1960–1961
B4	RNA-RNA	H20(m)	763–764	H34(Lm)	717–718
	Prot-RNA	S15	40–44, COOH-term	H34(LB,LM)	713, 717
B5	RNA-RNA	H44(m)	1418–1419	H64(m)	1768–1769
	RNA-Prot	H44(B)	1420–1422	L14	44–49
	RNA-RNA	H44(B)	1474–1476	H62(Bm)	1689–1690
	RNA-RNA	H44(B)	1474–1476	H64(m)	1989
B6	RNA-RNA	H44(m)	1429–1430, 1474–1476	H62(m)	1689–1690, 1702–1705
B7a	RNA-prot	H44(B)	1431	L19	(Hm24e:R44)
	RNA-RNA	H23(L,m)	698,702	H68(m)	1848–1849, 1896
B7b	RNA-Prot	H23(M,m)	712–713	L2	162–164, 172–174, 177–178
	RNA-Prot	H24(M,m)	773–776	L2	177–178, 198–202
B8	RNA-Prot	H14(LM)	345–347	L14	116–119

Table 1.1 List of the inter-subunit bridges in bacteria. The bridges are numbered as B1a, B1b, etc., as shown in the figure 1.9. The rRNA helices numbers are for *E. coli*. Abbreviations are: M- major groove, m- minor groove, L- loop, B- backbone, Lm- minor groove on the side of loop, LB- to the loop backbone. Table adapted from [30].

The inter-subunit bridges are of three different kinds. The RNA-RNA bridges (B2a, B2b, B2c, and B3) which are located in the core of the inter-subunit interface. The RNA-protein bridges and the protein-protein bridges (B1b, B1a, B7b, and B8) found in the peripheral regions of the ribosome. The inter-subunit bridges including the rRNAs are conserved through all the three kingdoms [98]. It is believed that the inter-subunit bridges located at the core are formed during the 70S initiation complex rapidly followed by the inter-subunit bridges at the periphery of the ribosome [99].

1.6.8 Factors that influence translation

There are several translation factors that influence and bind transiently to the ribosome through various steps during translation (Table 1.1). Translation initiation in the prokaryotes is regulated by three initiation factors (IF1, IF2, and IF3) (reviewed in [100]), whereas, in the eukaryotes there are ten initiation factors. Out of the three initiation factors, IF3 is assumed to bind to the 30S strongly, preventing the dissociation of the 30S from 50S. It is also known to help in the selection and binding of the fMet-tRNA, by destabilizing the binding of the other tRNAs at the P-site [62, 101]. A crystal structure of IF3 bound to 30S showed that IF3 binds at the A-site of the 30S and prevents the binding of tRNA at the A-site. By doing so, it might be involved in the translation regulation [102]. During the elongation of the polypeptide chain, the ribosomes interacts with the GTPases EF-Tu and EF-G. The EF-Tu is responsible for bringing the aminoacyl tRNA to the A-site. The correct recognition of the mRNA codon by the aminoacyl tRNA anti-codon stimulates the binding of this aminoacyl tRNA at the A-site, hydrolysed by EF-Tu [103]. The EF-G however, acts after the peptidyl transferase reaction is complete and help in the movement of the tRNAs from the A to P and P to E sites. The release factors (RF1, RF2, and RF3) recognize the stop codon in the mRNA and enable the termination of translation. Ribosome recycling factor (RRF) and EF-G assist in ribosome recycling and preparing ribosomes for another round of translation.

1.7 Overview of the translation process

Translation is a multi-step complex process where the triple codons of mRNA sequences are translated into amino-acids to form functional proteins.

Translation is a highly regulated mechanism which takes place in the ribosomes, with the help of several factors. Bacterial translation is simpler than eukaryotic translation, and can be sub-divided into four major steps, as initiation, elongation, termination, and recycling (Figure 1.10).

Figure 1.10 An overview of the bacterial translation. The translation system in bacteria showing the translation initiation, elongation and translocation along with the recycling of the ribosomes. Various factors which form an integral part of the translation are also shown and labelled. Figure adapted from [36].

1.7.1 Initiation

Translation initiation starts in a step-by-step manner with 30S subunit and ends with the formation of a 70S initiation complex (the 70S IC). First the mRNA positions itself over the 30S with the help of the complementary binding of its SD sequence to the anti-SD sequence in 16S rRNA [37]. The correct positioning of the mRNA in the 30S ensures that the first start codon AUG is positioned at the P-site, which leads to the formation of the 30S pre-initiation complex (30S PIC). Translation factors, IF1, IF2, and IF3 act collectively and recruit and position the first aminoacylated tRNA (aa-tRNA) (bound with the formyl derivative of methionine, fMet) into the P-site, leading to the 30S initiation complex (30S IC). The IF1 and IF3 bind to the A- and E-site, respectively, leaving and directing the aa-tRNA fMet to the P-site [102, 104, 105]. The

dissociation of the IFs and binding of the 50S to the 30S results in the formation of the 70S IC, ready for the next step elongation.

1.7.2 Elongation

The elongation step starts with the 70S IC bound with the aa-tRNA fMet at the P-site. Elongation takes place in an iterative cycle, and is divided into three steps, decoding, tRNA accommodation and peptide bond formation, and translocation (Figure 1.10). The Decoding step is carried out at the DC where the cognate tRNA is delivered to the A-site via the ternary complex. The ternary complex first binds at the A/T site of the ribosome, located far from the PTC. At the A/T site, the tRNA adopts a different conformation than the usual L-shaped conformation at the tRNA binding sites. The modified conformation at the A/T sites abets in the codon-anticodon pairing at the DC of the 30S. If there is a mismatch with the codon at the A-site, the ternary complex is released from the A/T site and replaced with another ternary complex. Once the codon-anticodon matches, the aminoacyl-tRNA binds at the A-site, a hydrolysis reaction carried out by EF-Tu. The EF-Tu then dissociates from the ribosome. This is followed by the peptide bond formation, and the nascent polypeptide chain is transferred from the peptidyl-tRNA at the P-site to the aminoacyl-tRNA at the A-site. The next step is translocation, where the tRNAs translocate from the P to E-site and A to P-site along with the movement of the mRNA by one codon (three nucleotides), which is catalyzed by the EF-G. Completion of the translocation results with an empty A-site for the incoming next aa-tRNA. This elongation cycle continues with addition of an amino-acid after each cycle and release of the poly-peptide chain through the polypeptide exit tunnel.

1.7.3 Termination

Translation is terminated when the elongation cycle reaches any of the stop codons (UAA, UGA, or UAG) on the mRNA sequence. The termination codon is recognized by either RF1 (recognizes UAA and UAG) or RF2 (recognizes UAA and UGA) (reviewed in [106]). The RF1/RF2 binds at the A-site of the 30S and mimics the A-tRNA and breaks the ester linkage of the polypeptide chain, facilitating the release of the nascent polypeptide chain through the polypeptide exit tunnel. RF3-GTPase plays an important role in the

dissociation of RF1/RF2 after the release of the nascent polypeptide chain from the ribosome, completing the termination of translation [106].

1.7.4 Recycling

After the release of the polypeptide chain from the ribosome and termination of translation, the ribosomes prepare themselves for new round of translation. The recycling is carried out by the ribosome recycling factor (RRF), which works with EF-G, to separate the two subunits, along with the dissociation of the mRNA and the deacylated tRNA (reviewed in [107]).

A schematic representation of the ribosome during translation and post translation is given in the figure 1.11. It shows how the complex retains its structure in a folded state integrated with mRNA, tRNAs, and nascent polypeptide chain.

Figure 1.11 Schematic representation of ribosome during and after translation. The 70S complex formed during translation with mRNA, tRNAs, nascent polypeptide chain emerging from the 50S (left). The dissociation of the ribosomal complex separating showing the individual components of the ribosome translation complex. Figure is adapted from The Science Visualization Lab.

1.8 Hibernating Ribosome

The growth of bacteria can be influenced by several stress factors, like temperature, aeration, nutrient availability, etc. Under such conditions, ribosomes downregulate their activity and play an important role in adaptation of the bacterial cells to such adverse environmental conditions. Under normal growth condition, bacterial ribosomes may exist in 30S or 50S state or in 70S

complex. When the bacteria enter the stationary phase, there is an increase in the stationary phase induced ribosome associated proteins (SRA proteins). Examples of the SRA protein family are hibernation promoting factor (HPF), ribosome modulation factor (RMF), YifA and plastid pY factor. These SRA proteins bind reversibly to the SSU in the 70S ribosome under unfavourable growth conditions arresting the ribosomes in an inactive state, known as the hibernation state. These translationally inactive hibernating ribosomes have been reported to exist in a dimer form of the 70S, known as the 100S, formed by two 70S ribosomes transiently connected at the 30S subunits. The structure of 100S ribosome dimers have been reported from *E. coli*, *B. subtilis*, *T. thermophilus*, *L. lactis* and *S. aureus* [108-113].

1.8.1 Mechanism of formation of hibernating ribosome dimers (100S ribosome)

The mechanism of formation of the 100S ribosome dimers differs among different species. In *E. coli*, the both RMF and HPF^{short} are required to bind with the 30S of the 70S ribosome to facilitate the dimerization (Figure 1.12). The HPF binds at the mRNA channel at the A- and P-tRNA binding sites in the 30S, and the RMF binds near the mRNA exit site, thereby inhibiting translation [111, 114, 115]. The binding of the RMF and HPF to the 30S causes conformational changes, resulting in the formation of the inactive 100S ribosome dimers [112].

Figure 1.12 Mechanism of hibernation and formation of 100S ribosome dimers in *E. coli*. The figure illustrates the general mechanism of the formation of 100S dimer in bacteria during ribosome hibernation in the stationary phase cells. The mechanism is

based on the binding of two stationary phase associated proteins, HPF and RMF. Figure is adapted from [112].

HPF in bacteria can exist in two forms, short HPF (HPF^{short}) with only one domain, the N-terminal domain (NTD) and long HPF (HPF^{long}) with an additional C-terminal domain (CTD). Interestingly, the species that carry the gene encoding for the HPF^{long}, do not have the gene for RMF. In the recent mechanism unearthed in the formation of ribosome dimers, the CTD of HPF^{long} has substituted the function of RMF. HPF^{short} (and NTD of HPF^{long}) is capable of inactivating 70S ribosome by binding to the 30S at the mRNA channel. Recent ribosome dimer structures from *S. aureus* and *L. lactis*, possessing the HPF^{long} showed that the mechanism of dimerization differs from the species with HPF^{short} [108, 109]. The CTD of HPF^{long} binds to the protein uS2 on 70S and induce 100S formation by homo-dimerizing with the CTD of second HPF^{long} from the neighbouring 70S [108, 109]. The HPF^{long} mediated 100S dimers have been reported in *T. thermophilus*, *S. aureus*, *L. lactis*, and *B. subtilis*, [108, 109, 111, 116, 117].

The binding of the SRA protein, HPF, results in ribosome inactivation. Binding of HPF at the A- and P-tRNA binding sites protects the tRNA binding sites from various environmental factors which might be released during stationary phase and also against several ribosomes targeting antibiotics. Also, the 100S ribosome structure is more rigid and compact compared to the 70S, and thus is more protected from the stress factors. In *Pseudomonas aeruginosa*, HPF is required for ribosome preservation and during nutrient starvation. The ribosome preservation is needed for the resuscitation of *P. aeruginosa* cells from the dormancy state [120].

Thus, the study of the hibernating state from different species might assist in the further understanding of the functioning of ribosomes. Such studies from a pathogenic microorganism like *Mtb* might help in studying how mycobacterial species are able to stay dormant for a prolonged period.

1.9 Trans-translating ribosome

From the translation steps, it is well understood that an encounter with the stop codon is essential for the translation termination. The bacteria, unlike eukaryotes, lack a proof reading mechanism for mRNA for the onset of translation. This sometimes results in the translation of the mRNA which lacks a stop codon or contains a defected stop codon. The ribosome under such circumstances is stalled at this position, resulting in the decrease of the translation efficiency, as the ribosomes cannot be recycled for the next cycle of translation. It sometimes also leads in the production of an incomplete polypeptide which might be toxic for the cell. Fortunately, in bacteria, this issue is controlled by the trans-translation mechanism.

In trans-translation, a transfer messenger RNA (tmRNA) binds to the ribosome, mimicking both the mRNA and tRNA. The tmRNA is encoded by the gene *ssrA* (Small Stable RNA). The functioning of tmRNA in the trans-translation requires a small protein B (SmpB), and together they rescue the ribosome from the stalled state and resume translation. They also add a small degradation tag to the incomplete polypeptide chain which is then subjected to proteolysis in the cell. This efficient translation control system also fosters the degradation of the defected mRNA, preventing further ribosome stalling.

1.9.1 Structure of tmRNA

tmRNA is a highly structured bifunctional RNA molecule which structurally and functionally mimics both the tRNA and the mRNA during trans-translation. The secondary structure of *T. thermophilus* tmRNA is represented in the figure 1.14. The size of the tmRNA can range from 230 - 400 nucleotides, depending on the species. The 3' and 5' ends of this RNA molecule fold in a tRNA like domain (TLD), which forms the acceptor arm similar to the tRNA, but it lacks the anticodon-stem loop. The tmRNA also mimics the mRNA with its stretch of RNA sequence known as the mRNA like domain (MLD). The size of tmRNA is about six times the size of a tRNA. The TLD and MLD constitute the functional elements of the tmRNA. This large RNA molecule is stabilized by various pseudoknots (PKs) and helices (H) in the tmRNA, which constitute the structural elements of the tmRNA. There are four pseudoknots in tmRNA (PK 1-4) and two helices (H2 and H5). The exact function of the PKs is still unknown, although it has been suggested that the PK2 and PK4 are responsible for the correct folding of the tmRNA during trans-translation [121]. The PK1 precedes the MLD, and could be assumed to be responsible for the correct positioning of the tmRNA's TLD and MLD at the 30S [122, 123]. The two helices interact with the 30S r-proteins on the solvent side and facilitate the various bends and movements in tmRNA during trans-translation [124]. The long and somewhat disrupted helix 2 (H2) connects the TLD with the pseudoknots, and the helix 5 (H5) together with the MLD carry the termination codon needed for the termination of the translation of the MLD of tmRNA.

Figure 1.14 Secondary structure of *T. thermophilus* tmRNA. The secondary structure shows the various structural components of the tmRNA, the pseudoknots PK1-4 and Helices H2 and H5, and the functional elements, the tRNA like domain (TLD) and mRNA like domain (MLD), colored and indicated. Figure is modified from [125].

Trans-translation required the protein SmpB, which is an integral part in the functioning of the tmRNA. The SmpB protein and the *ssrA* gene are conserved across prokaryotes. In species like *E. coli* [126], *B. subtilis* [127], *Salmonella enterica* [128], and *Yersinia pseudotuberculosis* [129], the *ssrA* is not an essential gene under normal growth condition albeit, the *ssrA* influences the cell growth and recovery of bacteria from stressed conditions. The *ssrA* deletion strains in *E. coli*, showed reduced motility, were more sensitive to antibiotics, showed slower recovery rate from carbon starvation [130, 131], and in *B. subtilis*, showed growth defect in high and low temperatures [127]. The *ssrA* gene has been predicted to be an essential gene in few pathogenic species like *Neisseria gonorrhoeae*, *Mycoplasma pneumoniae*, *Mycoplasma genitalium* and *Haemophilus influenza* [132]. Crystal structures of the tmRNA-TLD bound with SmpB verified that the SmpB binds to the D-loop of the TLD in the complex in an open conformation [133]. Thus, the SmpB together with the TLD of the tmRNA mimic the entire tRNA molecule. SmpB compensates for the D stem and

the anticodon loop of tRNA in tmRNA. It also provides the codon-anticodon base pairing in the trans-translation, and thus is an indispensable element in the functioning of tmRNA. SmpB is believed to interact with the tmRNA outside the TLD, and these interactions might help in the correct positioning of the open reading frame of the MLD [134-138].

1.9.2 The mechanism of trans-translation

The binding of tmRNA-SmpB complex to the ribosome is quite captivating (Figure 1.15). Trans-translation initiates with a stalled ribosome with an empty A-site. The aminoacyl tmRNA-SmpB complex binds to the ribosome at the A/T site with the help of EF-Tu. This is followed by the codon-anticodon pairing and hydrolysis by EF-Tu, resulting in the binding of the aa-tmRNA-SmpB complex at the A-site [124]. At the A-site, tmRNA-SmpB complex mimics the function of the A-tRNA and accepts the truncated polypeptide from the P-site, followed by the peptide bond formation and translocation to the P-site. Interestingly, the translocation of the tmRNA-SmpB complex is analogous to tRNA translocation from the A- to P-site, even though the tmRNA is almost six times the size of a tRNA [139, 140]. At this point, the deacylated P-tRNA moves to the E-site, and exits the ribosome thereafter. The MLD which is already in position, then gets translated into the polypeptide chain, which encodes a degradation tag. The MLD also carries a stop codon at the end of the MLD sequence which helps in release of the truncated tagged polypeptide and the dissociation of the tmRNA from the ribosome, resuming translation. The truncated tagged polypeptide is then subjected to proteolysis in the cell [139, 141].

Figure 1.15 The mechanism of trans-translation. A schematic representation of the mechanism of trans-translation process in bacteria. The tmRNA binds to the stalled ribosome with the help of EF-Tu and carries out the process of translation in the same

manner as the tRNA. Trans-translation helps in the degradation of the defective mRNA and proteolysis of the tagged nascent polypeptide chain. Figure adapted [142].

1.9.3 Physiological significance of trans-translation

As mentioned before, *ssrA* is not essential in many bacteria, although it might affect the cell's response under stress or other environmental factors. In *E. coli* and *B. subtilis*, the *ssrA* deletion strains show slower physiological response to stress [127, 130, 131]. During the antibiotic challenge, there was an eight-fold increase in the tmRNA level in *Thermotoga maritime* [143]. It is interesting to notice that the deletion of the *ssrA* gene affects the virulence in few pathogens like, *Y. pseudotuberculosis* [129] and *S. enterica* [128]. Studies showed that in the absence of *ssrA* and SmpB, *Y. pseudotuberculosis* did not have any lethal effect on the mouse model [129]. Therefore, the study of tmRNA-smpB complex from Mycobacteria may shed some light about their virulence and help in discovery of new drug targets against these pathogens.

1.10 Mycobacterial ribosome

The first report of the mycobacterial ribosome was the structure of *M. smegmatis* 70S ribosome at 12 Å [144]. Even though the resolution of the structure was low, several distinct features unseen from bacterial ribosome were recognizable. The 50S showed a long extension emerging near the L1 stalk, and was termed as steeple. They reported various rRNA expansions in the ribosome periphery and suggested the presence of extra domains around the r-proteins bL17, bL25, uS3, uS5, and uS16. Apart from these features, the overall ribosome shared the conserved bacterial 70S core architecture. Since no ribosome modeling was done on this structure, there was no molecular interpretations of the ribosome at this stage.

The second report was the structure of the *Mtb* LSU bound with the ribosomal silencing factor during starvation or stationary phase (RsfS) at a resolution of 8.5 Å [145]. This structure discussed about the interaction of RsfS with protein L14 in the 50S ribosomal subunit of *Mtb*.

While this dissertation work was undergoing, two structures of *M. smegmatis* 70S ribosome and one structure of *M. tuberculosis* 70S ribosome in their P/P state at near atomic resolution were published [146-148]. These

structures showed the various rRNA and r-protein extensions present in the Mycobacteria ribosome. The most prominent feature of Mycobacteria ribosome is the highly flexible rRNA extension near the L1 stalk, the H54a, which is about 100 nucleotides in length. The exact function of the helix is still unknown. There are also two new r-proteins (one in each 30S and 50S) in the mycobacterial ribosome near the functional centers, DC and PTC, respectively. A detailed analysis of these three structures presented an overview how the species-specific features of mycobacterial ribosomes have brought a new understanding of the Mycobacteria and how can this information be exploited in structure based drug discovery.

1.11 Cryo-EM

Structural biology focuses on the architectural information of the biological complexes. Atomic or near atomic resolutions (broadly defined as the resolution from 1.5 Å to 4.5 Å) information of the macromolecular complexes becomes a necessity in understanding the molecular mechanism and in developing drugs against specific targets for pathogenic microorganisms, and assists in gaining insights about these macromolecular complexes. Until recently, Nuclear Magnetic Resonance (NMR) and X-ray crystallography were the two pioneered techniques widely used to study the molecular structures of cellular complexes. Even though these two techniques have contributed in innumerable structures over the last 90 years, there are still some bottlenecks in these techniques. While NMR is limited to the study of the structures above 50 kDa, X-ray requires a high amount of highly purified sample which can form crystals. The process of crystallization becomes even more difficult for the large multi-protein complexes.

An alternative to this, and a third technique in structural biology is the cryo-electron microscopy (cryo-EM) that has emerged since 1980s and has become a powerful tool to visualize large macromolecular complexes. Such macromolecular complexes are either too large or too flexible for crystallographic methods, and thus can be better studied with Cryo-EM in combination with single-particle reconstruction. Cryo-EM has three most important gains over NMR and X-ray crystallography, i) it does not require crystals, ii) it needs very low amount of sample and has a fairly easy sample

preparation procedure, and iii) it allows certain level of heterogeneity in the sample purification.

Cryo-EM as the name suggest is a microscopic technique which uses electrons as a source to image the frozen specimen and these electron microscopic images of the specimen are used in their 3D structure determination. In a more simplified manner, a small amount of the sample is applied to an EM grid and is plunge-frozen. These frozen grids are transferred to the transmission electron microscope for data collection. Using several processing advancements in the area of single particle analysis, the 2D projections of these images are processed to yield the 3D information of the biological sample (Figure 1.16) [149]. These 2D projections are computationally processed into a 3D volume of the sample at near atomic resolution. The term ‘cryo’ states that the specimen is viewed in a cryogenic state. A cryogenic state prevents the specimen from the radiation damage by electrons.

Figure 1.16 Principle of cryo-electron microscopy. The arrows symbolize the electrons hitting the specimen (blue) embedded in the vitreous ice. The 2D projections of the specimen is shown in the image below, is an example of the images obtained from cryo-electron microscope. Figure is adapted from [150].

The single particle analysis (SPA) aims to solve the reconstruction problem: how to obtain a good 3D representation of the sample molecule from the experimental 2D projection data. The technique takes advantage of a large collection of EM projections of structurally identical particles, laying in random orientations. The SPA technique involves two main stages: determination of the projection image parameters, and 3D reconstruction of the desired macromolecule. In the first stage, parameters such as the defocus value of the

images collected from the microscope must be determined for later image correction. A faithful reconstruction of the original macromolecule requires image correction of the experimental data for the aberrations of the electron microscope imaging system (discussed in chapter 2). Individual particle images are isolated and their respective projection angles are determined for alignment of the data and reconstruction of the final density map.

The SPA technique implicitly assumes that multiple projection images originate from structurally identical particles. However, in cryo-EM, it is difficult to find biochemical means of ensuring complete structural homogeneity among free, isolated particles. Heterogeneity exists in two main forms: compositional and conformational. Regarding the former, variation in ligand occupancies can result in complexes of differing composition [151]. Regarding the latter, the particle in its native state can thermodynamically assume a large range of conformational states [152]. Separation of conformers is essential because structural details are diminished if data from structurally different particles are merged together. Iterative classification techniques (discussed Chapter 2) have been developed to separate different conformers into smaller but more homogeneous subsets, thus allowing an inventory of molecules in different functional states to be reconstructed from a single dataset. Originally considered a major obstacle for cryo-EM, structural heterogeneity can now be seen as an advantage of the SPA method because an entire spectrum of functional states can be captured and investigated in a single experiment [153]. Impressive studies with large datasets have successfully produced inventories for studying the dynamics of the tRNA with the 70S ribosome [154] or the biogenesis of the 30S small ribosomal subunit [155].

Cryo-EM can provide structural information for a wide range of biological specimens, from cells to macromolecular complexes, which includes cells, viruses, intact tissues, and protein molecules. The sub-disciplines of cryo-EM, which include Cryo-electron tomography, electron crystallography, and single-particle cryo-electron microscopy have been extensively used to analyze and interpret various biological structures. These methods, either used individually or in combinations, have furnished structural information complemented with X-ray crystallography and NMR spectrometry [156]. Cryo-EM is a constantly evolving method. And because of the increased study and rise in the cryo-EM

structures and publications, Nature Methods named cryo-EM as its ‘method of the year’ for 2015. Advances in cryo-EM and its influence in science was recognized with the 2017 Nobel prize in chemistry, awarded to Jacques Dubochet, Joachim Frank and Richard Henderson for their contribution in developing the cryo-EM. From the belief that electron microscope cannot generate structures below 20 Å, cryo-EM journey has led to today’s resolution of *E. coli* β -galactosidase better than 1.8 Å [157], at the time of writing this dissertation. Achieving near atomic resolution with cryo-EM in the recent years has also enabled in *de novo* model building of molecules which has no homologous information before. The resolution revolution, which is the great leap forward in the resolution has goaded the structure biologist to push the limits with cryo-EM to study the macromolecular structures which were once near impossible to study with X-ray crystallography.

Cryo-EM is undoubtedly the most preferred technique used for the study of ribosomes, as can be seen from the abundant increase in the EMDB submissions of ribosome structures in the past years. The ribosomes are the large protein-RNA macromolecular complexes which can range from about 2.5 MDa in prokaryotes to about 3.5 - 4.0 MDa in eukaryotes, and can be a challenging sample to be crystallized. Ribosome as a molecule is always in a constant motion. It undergoes various states of movement during translation. This results in the presence of different states of ribosome coexisting in the same sample, which makes the formation of ribosomes crystals very difficult. This heterogeneity of ribosomes is not a limiting factor in cryo-EM, because of the classification of images done during the processing of a cryo-EM dataset. And sometimes, this heterogeneity draws a better understanding of translation process by separating the transient stages of ribosomes into different ribosomal structures. With recent advances in cryo-EM, it is becoming possible to image and separate different translational intermediates of ribosomes. The highest resolution of ribosome from cryo-EM was obtained for the structure of the large subunit of *Trypanosoma cruzi*, which is a pathogen of Chagas disease, resolved to 2.5 Å [158].

A review of the mathematical concepts underlying TEM and the SPA methods in cryo-EM and the details of the evolution of the cryo-EM as an imaging technique were beyond the scope of this dissertation. These concepts are well described by the pioneers in the field of cryo-EM and few articles to mention

will be [159], [160], [161], [162], [163], [164], [165], and [166]. A concise and practical review of each step is detailed in the chapter 2, which give an overview of the cryo electron microscopy and data processing by single particle analysis.

1.12 AIM OF THIS DISSERTATION

The aim for my doctoral dissertation is the determination of a high-resolution structure of *Mycobacterium smegmatis* 70S ribosome with single-particle cryo-electron microscopy. Along with this, I also aim to do a structural analysis of the *M. smegmatis* ribosome, and understand the various species-specific features of mycobacterial ribosome. Such a study will provide us with a better understanding of the ribosome structure and translation mechanism in Mycobacteria.

This dissertation also focuses on the different ribosomal functional states found in the Mycobacteria under different growth phases. The study of these different functional states was possible because of a large dataset collected for the *M. smegmatis* ribosomes, and with meticulous data processing via single particle analysis cryo-EM. An evaluation of the different functional states from *M. smegmatis*, which shares high genetic similarity with *M. tuberculosis*, can help in a better understanding about the pathogenic mycobacterial ribosome.

Chapter 2

MATERIALS AND METHODS

2 Materials and Methods

2.1 Materials

2.1.1 Buffers, Media and Salts

All the chemicals used for in this study were of analytical grade. A list of the chemicals used with their company names are listed as following:

β -Mercaptoethanol	Sigma
2% Bis, 30% Acrylamide	Bio-Rad
Acetic Acid	Merck
Coomassie brilliant blue, G-250	Applichem
Glycine	1st Base Biochemicals
Glycerol	Affymetrix
HEPES	Sigma
Kanamycin Sulfate	Gibco
LB agar	BD Biosciences
LB broth	BD Biosciences
Magnesium Acetate	Merck
Methanol	Merck
Potassium Chloride	Merck
Protease inhibitor	Roche
RNAse inhibitor	Promega
Sodium dodecyl sulfate (SDS)	Hoefer
Sucrose	Sigma
Tetramethylethylenediamine (TEMED)	Biorad
Tris base	Promega
Tween 80	Bio-Rad

2.1.2 Microbiology strains

<i>M. smegmatis</i> strain mc ² 155 (kan ⁺)	Prof. Gerhard Gruber's Lab (Nanyang Technological University, Singapore)
--	---

2.1.3 Buffer Composition

Ribosome purification buffer composition

<i>Buffer</i>	<i>Composition</i>
Extraction Buffer	40 mM Hepes pH 7.5, 500 mM Potassium acetate, 25 mM Magnesium acetate, 250 mM Sucrose and 5 mM β -mercaptoethanol
High-salt sucrose cushion	40 mM Hepes pH 7.5, 500 mM Potassium acetate, 10 mM Magnesium acetate, 750 mM Sucrose and 5 mM β -mercaptoethanol
Grid buffer	20 mM Hepes pH 7.5, 50 mM Potassium acetate, 10 mM Magnesium acetate, 5 mM β -mercaptoethanol, 0.1% PI pill/ml and 1 unit/ml RNASin
Sucrose gradient buffer	20 mM Hepes pH 7.5, 50 mM Potassium acetate, 10 mM Magnesium acetate, 5 mM β -mercaptoethanol, 0.1% PI pill/ml, 1 unit/ml RNASin and 10/40% Sucrose
Low-salt sucrose cushion	20 mM Hepes pH 7.5, 50 mM Potassium acetate, 10 mM Magnesium acetate, 750 mM Sucrose and 5 mM β -mercaptoethanol

2.1.4 Instruments and accessories

Cryo-TEM microscope (200 kV)

FEI Tecnai Arctica
Nanyang Technological
University, Singapore.

Cryo-TEM microscope (300 kV)

FEI Titan Krios
National University of
Singapore, Singapore.

Bio-photometer Plus	Eppendorf
Gel Doc EZ System	Bio-Rad
Fraction Collector	BioComp
Ultracentrifuges	Beckman Coulter
Sonicator	Proscientific
Microfluidizer	Microfluidics
Vitrobot	FEI
Glow-discharger	Harrick
MALDI-TOF/TOF	School of Biological Sciences, Nanyang Technological University, Singapore.
Syringe, needles and accessories	BD Biosciences

2.1.5 Computer tools and softwares

COOT - model building for proteins and RNAs [167]

EMAN 2.1 - SPA data processing [168]

I-TASSER - to generate homology models for the proteins [169]

ModeRNA server - to generate homology models for the RNAs [170]

MolProbity - to validate the model [171]

PDB to 3D Restraints server - for model refinement in *phenix.refine*
(<http://rna.ucsc.edu/pdbrestraints/pdbtorestraints.html>)

phenix.refine - for model refinement [172]

Phyre - to generate homology models for the proteins [173]

RELION 1.4 - SPA data processing [174]

Resmap - to calculate the local resolution of the cryo-EM maps [175]

UCSF Chimera - for structure analysis, estimating map-model
overfitting, and making figures [176]

2.2 Methods

2.2.1 Purification of 70S ribosomes from *M. smegmatis*

70S ribosomes of *M. smegmatis* were purified from *M. smegmatis* strain mc²155. 70S ribosomes were prepared under two different growth phases, log phase (24 h growth) and stationary phase (48 h growth). Cells were grown in LB medium supplemented with 0.05% (v/v) Tween 80 and 0.2% glycerol (v/v). Cell pellet was collected by centrifugation at 8,000 g for 20 min using the Beckman ultracentrifuge (rotor - Beckman JA 25.5).

Cell lysis: Cell pellet was resuspended in extraction buffer and sonicated at 30% amplitude (10 sec ON and 20 sec OFF) for 5 min. The lysed cells were further broken down using microfluidizer. The broken cell suspension was centrifuged at 15,000 g for 45 min using the Beckman ultracentrifuge to remove cell debris and/or unbroken cells.

Crude Ribosome Preparation: The supernatant from the above centrifuge was overlaid on a high-salt sucrose cushion (sucrose cushion: supernatant = 1: 3) and centrifuged at 100,000 g for 2.5 h using the Beckman ultracentrifuge (rotor - Beckman TI 50). The pellet obtained after centrifugation is considered as the crude ribosome. The pellet was resuspended in grid buffer over ice for about 2 h. The concentration of the ribosome was calculated by measuring the absorbance of RNA in the spectrophotometer at 260 nm. The crude ribosomes are stored by flash freezing them in liquid nitrogen at -80°C.

70s Ribosome preparations from crude ribosome: The crude ribosomes were further purified over sucrose gradient of 10-40%. The gradient was made by first adding 6 ml of 10% sucrose gradient buffer into the centrifuge tube. 40% sucrose gradient buffer was dispensed slowly by immersing from a needle at the bottom of the tube. Linear gradients of 10-40% sucrose was made with Biocomp gradient station (angle = 81.5°, rpm = 17, duration = 1.48 min). After the gradients were made, crude ribosomes were loaded on each tube at concentration of 10-12 OD₂₆₀. The gradient fractions were centrifuged at 111,132 g for 4 h using the Beckman ultracentrifuge (rotor - Beckman SW 41). After centrifugation, the ribosomes were fractionated on the BioComp gradient station at a speed of 0.34 cm/sec and fraction size of 20 drops/tube. The 70S fractions were pooled together (about 3.5 ml) and overlaid on low-salt sucrose cushion and centrifuged at

100,000 g for 2.5 h using the Beckman benchtop ultracentrifuge (rotor - Beckman TLA 110). The ribosome pellet obtained after centrifugation was resuspended in grid buffer over ice for about 2 h. Concentration of the ribosome is calculated as 1OD₂₆₀ equals to 23 pmols. Small aliquots were made and snap frozen in liquid nitrogen and stored in -80°C until further use.

2.2.2 Ribosome protein identification by SDS PAGE and Mass Spectrometry

Purified ribosomes were analyzed on the 12% SDS denaturing PAGE for protein identification. 15 pmol of ribosomes were denatured with SDS loading buffer prior to electrophoresis. The predicted band of protein based on its migration pattern was cut and analyzed by mass spectrometry for protein identification.

The protein identification was carried out by MALDI-TOF/TOF, in the facility located at the Nanyang Technological University, Singapore. The peptide mass fingerprinting results were used in the Swiss-Prot database using the Mascot search engine to identify the protein.

2.2.3 Purification of 70S ribosomes from *M. smegmatis* growing in different growth phases

To study the ribosome behavior during cell growth, *M. smegmatis* cells were collected from the different growth phase cultures, log phase (24 h), where the OD_{260nm} was increasing; stationary phase (48 h), where the OD_{260nm} remained constant; and extended stationary phase (66 h), which was collected 18 h after the stable OD_{260nm} at the stationary phase to check if there have been any further variations in the cell growth. The ribosomes were purified using the protocol stated in 2.2.1 with changes in the extraction and high-salt buffers. The ribosomes were purified under low-salt (50 mM potassium acetate) and high-salt (500 mM potassium acetate) conditions. The rest of the buffer composition remained the same. The purification profile after sucrose gradient purification gives an understanding about the different ribosomal complexes found under these conditions.

2.2.4 Cryo-Electron Microscopy

The technique used for the analysis of the *Mycobacterium smegmatis* 70S ribosome in this thesis is Cryo-Electron Microscopy. The process of cryo-EM comprises three major parts; first is sample preparation, followed by data collection using cryo-EM, and data processing of the micrographs collected. This section as gives a detailed description of every step of cryo-EM.

2.2.4.1 Cryo-EM Sample Preparation

The sample preparation is a very critical step in cryo-EM. These samples are subjected to high vacuum; thus, the sample needs to be prepared in such a way that the structure of the initially hydrated molecules of the specimen is preserved when the samples are positioned in vacuum. In addition to this, the specimen must be protected from the radiation damage caused by the exposure to electron beam. For TEM, samples can be prepared by two different techniques, negative staining and vitrification. Both the sample preparation procedures are explained in the following section.

Negative staining

The method of negative staining was developed by Brenner and Horne [177], where a heavy metal salt (e.g. uranyl acetate, uranyl formate) is used to reduce the radiation damage and protect the molecules, and as the name suggests, stains the background giving a better contrast to the sample molecules when observed under microscope. Under this technique the structural information is limited to the shape of the molecule. The samples under this technique suffer deformation from the air drying and thus are preserved to a limited extent. Therefore, the negative stained samples are limited in resolution. Though limited in the resolution information, it gives an idea about the heterogeneity of the sample, concentration required for cryo-EM, and most importantly gives an idea about the formation of protein aggregates in studies of higher macromolecular complexes. This fast and easy sample preparation technique has thus made it a widely used method for screening the suitability of the sample for high-resolution cryo-EM.

Vitrification

In the technique of vitrification the sample is embedded in amorphous ice (vitreous ice), which enables the molecules to be fully hydrated despite the

requirement for vacuum conditions inside the EM, preserving the molecule in a biologically relevant conformation. The vitrification technique was pioneered by Jacques Dubochet nearly 30 years ago [178] and is continually developing over the years. Amorphous ice, an essential element in vitrification, is formed when water is cooled extremely fast (cooling rate $\sim 10^5$ K/s) [178] preventing the formation of ice crystals which could damage the sample. This is achieved by plunge freezing the grid (containing the sample) into a cryogen (often liquid ethane or liquid ethane-propane mixtures, -181°C) cooled by liquid nitrogen (-196°C). Once the samples are vitrified, they have to be kept at liquid nitrogen temperature to avoid devitrification and ice crystal formation. A schematic representation of the sample preparation via vitrification is given in the figure 2.1. Before sample application, sometimes, a thin layer of amorphous carbon is placed onto the EM grid [179]. The amorphous carbon assists in the determination of the contrast transfer function, to be discussed later in sections 2.2.4.3 and 2.2.5.3. The white Thon rings, shown in the figure 2.2, in the power spectrum originate from the scattering of the amorphous carbon film that was laid onto the EM grid during sample preparation.

Figure 2.1 The process of vitrification. The sample is first applied to the grid held by tweezers. The sample is blotted (to remove excess sample) by filter paper and then plunged frozen into a cryogen precooled at the liquid nitrogen temperature. The specimen/biomolecules are embedded in a thin amorphous ice film and are then images in an electron microscope. Figure adapted from [180].

Figure 2.2 *An example of micrograph and respective thon rings. (A) micrograph collected at the magnification $62000\times$ and defocus $-1\mu\text{m}$. (B) The power spectrum density (PSD) of the micrograph showing the thon rings. As can be seen, the thon rings are not perfect circular format due to astigmatism (discussed in the section 2.2.4.3). The scale bar is 0.5nm^{-1} . Figure adapted from [181].*

Vitrification of sample has many advantages over negative staining; it protects sample from radiation damage, samples do not suffer from deformation, and the contrast observed is related to the biological molecule and not from the contrasting agent (like heavy metal atoms in case of negative staining). Although, this is also the cause of low contrast observed in cryo-EM images of vitrified samples, which can cause problems in the subsequent image processing.

During imaging under EM, the electron beam can eject electrons (called ‘secondary electrons’) from the sample. This can lead to a positive charge build up which can distort the recorded image through an ‘effect termed charging’. Charging can be reduced by pre-irradiating the carbon film that coats the EM grids before sample application. In this dissertation work, the holey carbon coated grids were used. As the name indicates, these grids were coated with holey carbon.

2.2.4.2 The Electron Microscope

An electron microscope, as the name indicates, is a microscope where the electrons are used as the source of imaging. Electrons can behave in many different ways when they interact with the specimen/sample (Figure 2.3). These different electron interactions and the resulting signals can provide complementary information about the specimen, and different methods have been developed to study and take advantage of this. Two principle techniques in

electron microscopy are the scanning electron microscopy (SEM), and the transmission electron microscopy (TEM) (Figure 2.4). In a broad definition, SEM analyses the signals that are emitted from the surface of the specimen, whereas, TEM analyses the signals that are transmitted through the specimen. In this dissertation work, I have used TEM as the principal electron microscope technique, which is explained in detail in the upcoming sections. A schematic representation of a typical SEM and TEM is given in the figure 2.4.

Figure 2.3 Behaviour of the electrons in an electron microscope. When a high-energy beam interacts with a thin specimen, various different types of signals are generated, which are detected by different types of electron microscopes. The directions shown for each signal might not represent the exact physical direction of the signal, rather gives an idea about where the signal can be detected. Figure redrawn from [182].

The TEM consists of an electron source, a lens system (condenser lens, objective lens, and projector lens), an imaging system (this can be a viewing screen, a photographic film, a charge-coupled device, or a direct detector device), and a specimen loading system (Figure 2.3).

Figure 2.4 The electron microscope. A schematic diagram of the Transmission Electron Microscope (TEM) (left) and Scanning Electron Microscope (SEM) (right) and the comparison of their components. Figure adapted from [183].

Electron Source

Electron source is the apparatus responsible for generating and emitting the electrons in the TEM. The different electron sources available are, a heated tungsten filament, a lanthanum hexaboride (LaB₆) crystal, and tungsten field emission gun (FEG) [184]. The field emission gun (FEG) is the most advanced electron source which comes with a small size tip (10-25 nm radius), and emits coherent near monochromatic, high current density electrons. These thermally emitted electrons are extracted by a potential gradient and accelerated through voltages ranging from 80 - 300 kV. The acceleration voltage of the electrons is determined by the potential difference set up between the tip and a grounded anode (Figure 2.4). The best acceleration voltage to be used depends on the nature of the specimen. Increasing voltages will permeate electrons with greater penetrating power but reduce the number of scattering events that will occur between the electron beam and the specimen. Structural information about the specimen is encoded in the distribution and phases of these scattered electrons, and thus, these scattering events carry the resolution information of the specimen.

The lens system

The accelerated electrons generated by the electron source needs to stay focused into a coherent beam while travelling through the vacuum before hitting

the sample. The coherency and acceleration of the electrons are maintained in the column of the microscope through a series of electro-magnetic lenses, called condensers. In an optical microscope, the lenses are made up of glass, while a lens in the electron microscope is a magnetic field, produced by running current through the coils of a conducting wire.

The accelerated electrons (from the electron beam) are directed through these electromagnetic lens systems to interact with the specimen and form an image. The condenser lenses maintain the parallel nature of the electrons. The objective lens, immersing or after the specimen, focuses the parallel electron beam. An objective aperture is placed in the back focal plane of the objective lens to remove electrons scattered at high angles and increase the image contrast. The electron beam leaving the specimen is first focused by the objective lens to produce a first image (20-50×), which is subsequently magnified (about 1000×) by a set of projection lenses, before the electrons reach the imaging system.

The image detection system

In cryo-EM, the choice of the image detectors plays a very important role in achieving the desired final resolution. To evaluate the performance of the detector one parameter is often used, the detective quantum efficiency (DQE) [185]. The DQE is defined as the ratio of the signal to noise ratio (SNR) output squared to the SNR input squared [185].

$$DQE = (S/N)^2_{out} / (S/N)^2_{in}$$

It means that the detector determines how much the signal in the detector degrades compared to the original signal in the image, which is measured in terms of the signal-to noise (S/N) ratio in the output compared to that of the input [186]. This parameter determines the percentage of the incident electrons counted as a function of different spatial frequencies and shows how well a given detector can detect an electron event [187], since detection and estimation of these electron events is the underlying fundamental of the TEM. A perfect detector would have a DQE of 1, which means that every electron event is been detected and recorded. In cryo-EM three types of detectors are used, photographic emulsion film, charge-coupled-devices (CCD), and the direct electron detectors (DED).

Photographic emulsion film consists of silver halide grains suspended in a gel matrix. Electrons striking the film cause ionization of the halide grains, transforming them into metallic silver. Photographic films have a higher DQE value than the CCDs at high spatial frequencies. However, the great disadvantage is they require the dark room development, and digitization before assessment of the data, and image processing can take place.

In scintillator-CCDs, an indirect method of electron detection is used. In this system, incident electrons first collide with a scintillator, where the electron signal are converted into a photon signal. The photons emitted are transported by a dense bundle of fibre optics to a CCD array consisting of potential wells where the light signal is converted into an electrical signal. The electrical charge is accumulated in the wells before being read out line by line and recorded as a digital micrograph [188]. The signal conversion process results in a loss of resolution information; thus, the DQE of the CCD is lower than the photographic films.

The introduction of the DED, a completely digital method of image detectors, has revolutionized the electron microscopy imaging techniques. The DED uses complementary metal-oxide-semiconductor (CMOS) sensor, which is exposed directly to the electron beam, omitting the need for a scintillator fibre optics system. Thus, the DED sensor has the ability to detect and map each electron strike event, instead of integrating an accumulated charge [189]. Overall, the DED device offers convenience of automation, immediate access to the data, and has a high DQE performance across all spatial frequencies. Additionally, because of the increased readout speed on the CMOS chip, the exposure of each image can be broken up into dose-fractionated frames (sub-frames).

The ability of splitting the image into sub-frames (also known as ‘movie-mode’) has enabled correction of beam-induced motion and drift that occurs while acquiring images in TEM. The motion vectors are determined globally and for individual particles, and then the frames are aligned computationally [190, 191]. Using a high electron dose is unsuitable for the biological samples, as it might lead to radiation damage over time [192]. However, a higher dose improves the contrast of the images acquired, which is required for the identification of the particles and determination of their orientations. Now, with the help of DED and the ability to split the images into sub-frames, the images can be acquired at

higher electron dose, and during structure determination, the final frames of the images with the radiation damages can be discarded [190]. With constant improvements in the detector system in cryo-EM, the DEDs have played an important role in the ‘resolution revolution’.

2.2.4.3 Image formation

The basic underlying principle in TEM is the detecting the interaction of the electrons with the specimen. At the subatomic level, matter is mostly empty, and thus the majority of the electrons in the focused beam pass through the sample unscattered [193]. Interactions that occur between the incident electrons with the specimen result in two principal forms of scattering: elastic and inelastic (Figure 2.3).

In elastic scattering, the path of an incident electron passing close to an atomic nucleus is changed due to its attraction to the positive potential of the nucleus. The elastically scattered electrons exit the specimen with a change in phase from the incident phase but without change in kinetic energy [194]. Structural information about the particles is encoded within these elastically scattered electrons.

In inelastic scattering, kinetic energy is transferred from the incident electron to the atomic orbital electrons of the specimen [195]. The resulting inelastically scattered electrons leave the specimen at random angles with an overall loss of energy. Inelastic scattering ionizes the electron cloud, acts as a source of incoherent signal, and induces free radicals that scatter and break chemical bonds in the specimen [192]. This is the major source of the radiation damage to the specimen as well as a source of background noise in the electron micrograph. Inelastic scattering is diffused and thus provides little structural information about the specimen.

Elastically scattered electrons contain meaningful structural information about the specimen [196]. It is the interference of the phase-shifted elastically scattered electrons with the unscattered electrons that generates phase contrast, which is the main contribution to the image contrast in the final micrograph. The pattern of phase shifts in the back focal plane can be changed by defocusing the microscope with the objective lens [197].

In the electron microscope, the electromagnetic lenses experiences many of the similar defects as the optical lenses in the light microscope. In the microscope, the final image generated is convoluted by a function, called the point spread function (PSF), which depends on the aberrations of the imaging system. The defects in the TEM are the spherical aberrations, chromatic aberrations, and axial astigmatism [198]. The Fourier transform of the PSF is the contrast transfer function (CTF), which include the spherical aberration, axial astigmatism, and defocus value [182, 199, 200]. The signature of the CTF can be seen in the absolute-squared Fourier transform of the micrograph, called the computed power spectrum [197]. Therefore it is important to correct the collected images for their CTF of the electron microscope used for imaging (reviewed in [196]). Figure 2.5 shows the Fourier transform of the various types of CTF.

Figure 2.5 *Computed power spectrum for different types of CTF during imaging in cryo-EM. The CTF can be seen in the pattern of Thon rings computed from the power spectra of each micrograph image. (A) An example of a good micrograph, with round, full Thon rings. (B) Patterns produced by continuous drift of the specimen stage. (C) Aberrations of the microscope, such as axial astigmatism gives rise to elliptical Thon rings. (D) An under-focused micrograph shows small and closely spaced Thon rings. (E) Patterns produced by jump drift of the specimen stage.*

2.2.5 Cryo-EM Single Particle Analysis (Data Processing)

Cryo-EM results in a set of electron micrographs. Each single-particle image is approximately a 2D projection of the 3D electron density map of the molecule. On the 2D micrograph, the single particles appear as different

projection of the biological molecules. The individual projection images (commonly called as particles) of the molecules need to be selected for the subsequent image processing procedure to achieve a 3D reconstruction of the molecule. This entire process is known as Single Particle Analysis (SPA).

There are several SPA tools available for processing single-particle image data, including IMAGIC [201], SPIDER [154, 202], XMIPP [203], EMAN [168], FREALIGN [204], and RELION [174]. Out of these tools, EMAN and RELION are briefly introduced next, which are the major computational tools used in this work. SPA includes a number of steps, micrograph sorting, motion correction of the micrographs, particle selection from the micrographs, sorting and averaging of the particles into good and bad classes (known as 2D classification in RELION), 3D reconstruction and assignment of initial projection angles (3D classification in RELION), 3D refinement, which is an iterative process of angular refinement. SPA can be a challenging process, because of the low contrast and low signal-to-noise ratio (SNR). These issues are constantly being overcome with the advancement and development of new direct electron detectors and better cameras. Figure 2.6 depicts a workflow, summarising the steps involved in SPA and Figure 2.7 gives a flowchart of the data processing for this thesis.

Figure 2.6 A workflow of single particle analysis. The sample preparation is first done using vitrification. The vitrified sample is imaged in TEM. The particles are selected

from micrographs. The particles are then subjected to various rounds of 2D class averaging and 3D classification and refinement. The refined 3D volume maps are the used for model building and structure analysis. Figure adapted from [205].

Figure 2.7 Flowchart of image processing for *M. smegmatis* ribosome. A flowchart depicting the various steps of image processing for *Ms* 70S ribosome. The various steps involved are explained in detail in the sub-sections of 2.2.5.

2.2.5.1 Micrograph Sorting and Motion Correction

SPA starts with the sorting of the micrographs into good and bad micrographs by viewing the entire dataset. The good micrographs are then corrected for the aberrations caused due to beam-induced motion. This is called as motion correction. This step contributes to the better resolution in the cryo-EM SPA. The motion correction is carried out by the program Motioncorr [190].

2.2.5.2 Particle picking

The next step is individual particle selection from the electron micrographs. For a high-resolution 3D reconstruction, hundreds of thousands of particles are required. Particle are picked in a manual mode, semi-automated mode, or by using automated particle-picking programs. Manual particle picking has been a huge bottleneck in SPA taking weeks of repetitive work. Since then,

the development of automated particle picking program has been a great advancement in the area of single particle reconstruction. With the developments in machine learning and automations, the automated particle picking algorithms are improving very fast [206-208]. In this dissertation work, a semi-automated mode of particle picking by EMAN 2.1 was carried out. In the semi-automated mode, the first few particles are selected by the user, and then the software uses these selected particles as a reference to select the particles from the remaining micrographs. The micrographs are manually screened for the quality of the particles picked. Once done, the coordinates of the particles are imported into the software RELION 1.4, which was used for further processing.

2.2.5.3 CTF estimation

After the micrographs are particle picked, the next step is the CTF correction. As discussed before, this correction accounts for the defects from the defocus values and the spherical aberration constant of the lens system on the acquired images from the microscope. The effects of CTF have been extensively studied in the interest of image analysis in SPA, and have been explained in detail in [197], [209], and [210]. The CTF is determined by adjustable parameters, such as defocus, astigmatism and the acceleration voltage, as well as fixed properties of the electron microscope, such as lens aberrations and beam coherence [182]. The CTF oscillates around zero contrast, modulating the amplitude and also reversing the phase for some frequency intervals. Moreover, the CTF has an envelope, which leads to an attenuation of signal transfer with increasing spatial frequency [182]. A CTF curve is the plot against resolution in inverse angstroms. The limits of CTF can be avoided by collecting images at various defocus values. Dependent on the defocus value, features in the image at certain spatial frequency points have zero contrast, and in some spatial frequency range have reverse contrast. CTF correction can be performed either on all particle images before 3D reconstruction, or separately on each reconstruction obtained from a set of micrographs with the same defocus settings, before the reconstructions are combined. CTF correction can be done through various methods, including: (1) phase flipping in the spatial frequency regions that have reverse contrast, and (2) multiplying the Fourier transform of the image with a Wiener filter, which corrects for both phase and amplitude and suppresses noise near CTF zeroes.

Currently, CTF correction is most commonly done as an integral part of the 3D reconstruction and the CTF correction has been integrated into the RELION software. In the current datasets, the Wiener filter method using the CTFFIND 3 [211] was used on the acquired micrographs before particle classification.

2.2.5.4 2D class averaging

After selecting the particles, and CTF correction of the micrographs, the particles were sorted based on their shape and structural features. The assumption of this approach is that the signal to noise has a random distribution and the features of the particles have identical signals. The classification starts with random seeds, so the entire data set is divided into a number of classes (specified by the user, and usually kept 100 or more classes for a big dataset), which are randomly averaged. Next, all particles are classified to the generated classes and are subsequently aligned [212]. Thus, individual particles showing the same orientation and conformation are aligned to increase the SNR of the individual classes [213]. This procedure operates in an iterative fashion: the generated output of the previous iteration is used as a model for the next iteration. This 2D classification is useful to clean the data set from non-particle candidates, e.g. ice contamination or background selected as particles. It is also useful to remove particles, which depict only low-resolution information due to blurring arising from drift or beam-induced charging. The cleaned particle classes are further separated into homologous particles by 3D classification.

2.2.5.5 3D classification and refinement

After the 2D classification is completed, particles from the good class averages are saved and again classified into different classes to obtain 3D reconstructions of different conformations, based on a reference or initial model [214]. To start this step, initial model/reference is required which is used as a template for the alignment of the particles in the 3D classification. A reference can be generated from a homolog structure or 3D volume and by low pass filtering it. If such a reference is not available, then the initial model can be generated from the selected particles of 2D classes with data processing software like EMAN. In our dataset, a filtered *E. coli* SecA bound 70S ribosome map [215] was used as a reference in 3D classification. The RELION software package implements classification of single-particle cryo-EM data in 3D based on

Maximum Likelihood [174]. The 3D classification step can be repeated several times applying different parameters (no. of classes, no. of iterations, addition of movie frames, etc.) to generate the 3D reconstructions of different or best conformation.

The individual 3D classes are then refined to a higher resolution by the auto-refinement procedure. In RELION, this automated procedure employs the gold-standard Fourier Shell Correlation (FSC) calculations for the resolution estimation to avoid over-fitting [216]. The FSC is a function of consistency. It measures over shells of equal spatial frequencies in Fourier space and the normalized cross-correlation coefficients between the Fourier transform of two 3D volumes [217]. The two 3D volumes that are compared in the FSC calculations are obtained from splitting the dataset into two halves and refining them separately ('gold standard').

The ribosomes exist in a dynamic state in the cell, where the two subunits show different movements during translation. This dynamic state affects the resolution of the entire 70S ribosome. To overcome this issue, the two subunits are refined by applying a mask over the individual subunits and refining them separately ('focused mask refinement') [218]. Applying the mask defines the region outside as noise and only refines the volume inside. A focused mask refinement was applied to both the subunits, LSU and SSU, to yield two maps of the respective subunits. Additionally, the density is sharpened according to an automatically estimated B-factor [219]. The B-factor or temperature factor describes the attenuation of structure factors by thermal mobility of the molecule atoms.

2.2.6 Current capabilities and limitations of cryo-EM and Single Particle Analysis

The cryo-EM and single-particle reconstruction technique have seen great improvements over the years. The current trends in technical advances include (i) automation in data collection and processing, and (ii) hardware improvement, especially the use of direct detector devices.

'Automation' in cryo-EM enables the elimination of repetitive user interactions to process, translate or move datasets between image-processing modules using an intuitive mode of user learning [220]. Automated EM data

collection has enabled fast collection of large datasets on the level of tens of gigabytes to terabytes. It has been made possible by combining the digital recorders (CCD and DDD) with the computer-controlled operation of the lens system and stage movements. Automated data collection systems include Legion [221], JADAS [222], and EPU (FEI). In the current work, EPU is used for automated data collection. The use of DED has greatly increased the quality of images (see sections 2.2.4.1 and 2.2.4.2), thus improving the resolution of reconstructions and expanding the range of amenable molecule size. Theoretically, it is possible to solve the structure of a 100 kDa molecule close to 3 Å resolution using only 10,000 particles, as predicted by Richard Henderson [161]. And today, DED combined with careful computation is reaching this limit for asymmetrical complexes. For example, a 70S ribosome structure was solved to 6.5 Å resolution with using only 15,000 tilt-pair particles collected on a Falcon I DED (FEI), using movie frame alignment to compensate for the beam-induced drift in micrographs (Bai et al., 2013). An 80S eukaryotic ribosome complex was solved to 3.4 Å resolution with only 37,000 particles after classification [223]. Automation has also reduced the time required for the data processing. The structure of chloroplast ribosome was resolved to 3.2 Å in a matter of 24 hours [224].

DED has not only pushed the resolution limits, it has enabled in studying the structures of smaller molecules. The small molecules (considered as less than 200 kDa in size) have low contrast in the micrograph, and thereby making the particle selection difficult and also the aligning of these particles in the subsequent data processing steps. Few years back, the general notion about such small molecules was, they cannot be better than 20 Å resolution [225]. But with recent advancements in cryo-EM hardware, which included the field emission gun (see section 2.2.4.2), use of phase plate [226, 227], and DED, the image contrast for small molecules has greatly enhanced, and thereby pushing the resolution limits for small molecules. For example, the γ -secretase (~200 kDa) was solved to 3.4 Å using cryo-EM images collected on Gatan K2 summit DDD [228]. Recently a group pushed resolution limits by solving the structure of isocitrate dehydrogenase (93 kDa), cancer targets lactate dehydrogenase (145 kDa), and glutamate dehydrogenase (334 kDa) to a resolution of 3.8, 2.8, and 1.8 Å, respectively [229].

Limitations with cryo-EM and SPA

Even though there has been a surge in the resolution race, factors like sample preparation, camera detection efficiency, specimen stabilization under the beam, better electron optics, in-focus phase contrast, computational means for structure validity checks, wider access to high-end microscopes, and better training need to be considered for an improved data collection. Apart from these, there are several SPA limitations one needs to consider.

i) Use of cryo-EM frames. Preparation of perfect EM grids, algorithms for frame alignment, and dose weighting aim to reduce the Beam induced movements in the cryo-EM frames. Yet, some of the beam induced motions are not analysed by any method like 3D motion correction.

ii) Finer aberrations correction. Microscope aberrations which are not corrected by hardware are corrected by software. Many attempts have been done to correct for spherical aberrations, magnification anisotropy, or local defocus changes, but their use is not widespread, probably indicating that still a better match into the processing workflow is required. Also, the introduction of phase plates as a way to avoid defocusing poses additional challenges, since focus determination in these conditions becomes further difficult.

iii) Particle selection. Microscopy sample often consist of certain degree of heterogeneity. In SPA, selection of particles is the most important step and could be extremely complex when dealing with thousands, or even millions of particles. Current trend is to use automatic algorithms to select on the micrographs all those subimages that “might” be a particle. In the best case, these algorithms have a false positive range between 80 and 95%, meaning that between 5 and 20% of the data does not correspond to the particle of interest but to something else. Better algorithms are needed to decrease the false positive rate as well as to analyse the particle candidates and discard those that are not true particles. With the levels of noise in EM, particle pruning is still an open issue in the field.

iv) Flexibility and heterogeneity of macromolecules. Particle flexibility and heterogeneity is at the same time the blessing and the curse of EM. Flexibility helps to reveal the dynamics of the macromolecule under study, whereas, only homogeneous sets of particles can be reconstructed to atomic resolution. This becomes challenging where conformational changes being addressed are not

correctly modelled as a discrete collection of classes, but they explore a continuous of states, although with different probabilities. Indeed, the very successful Maximum Likelihood approach explicitly considers that the set of macromolecular views can be classified into an unknown number of distinct classes. It is not only that the determination of the number of classes may be a tedious and subjective task, but it is that the modeling does not conform to the reality behind, creating instabilities in the classification/reconstruction process. This issue has been explored in some works [230, 231], but these new approaches are not in general use in cryo-EM, again possibly indicating that they have to further evolve and mature before having a significant impact in the field. Still, this issue is crucial to the continuation of the success of cryo-EM, with the scarcity of works in this area only indicating how difficult it is.

Therefore, the EM has certainly seen a boost in the resolution, processing capabilities and is not competing well with the results of X-ray diffraction and NMR. These open problems are continuing to be worked upon and thus limitations with cryo-EM certainly looks hazy in the future, and no doubt soon to achieve a near atomic resolution for any specimen is possible without the worry of size or symmetry of the molecule.

2.2.7 Cryo-EM imaging of *M. smegmatis* ribosome

As mentioned in the section 2.2.1, *M. smegmatis* ribosomes were prepared under two different conditions, from cells growing in log (24 h), and stationary phase (48 h).

Imaging of 70S log phase ribosome. The log-phase ribosomes were imaged on a FEI Titan Krios electron microscope located at the National University of Singapore (NUS), Singapore, operating at 300 kV. The images were recorded for the period of 48 h on a FEI Falcon II direct electron detector at a calibrated magnification of 126,000 X yielding a pixel size of 1.11 Å. The defocus values ranged from 0.5 to 3.5 µm for this dataset. The images were recorded with 23 movie frames with a combined dose of 35 electrons per Å². This dataset resulted in two different 70S states. The first state was resolved to the resolution of 3.4 Å and was identified as the P/P state 70S ribosome because of the presence of a strong density of P-site tRNA in the non-rotated 70S state.

Focused mask refinement for LSU and SSU resulted in two cryo-EM maps at an average resolution of 3.2 and 3.7 Å, respectively. These two refined maps were used for model building of the entire 70S ribosome of *M. smegmatis*. This dataset also produced a minor class (0.6% particles) refined to an average 12.5 Å resolution (third round of 3D classification with fine angular sampling), which represents the trans-translating state of 70S ribosome with tmRNA bound at the P-site.

Imaging of 70S stationary phase ribosome. The stationary phase ribosomes were imaged with FEI Tecnai Arctica electron microscope located at the Nanyang Technological University (NTU), Singapore, operating at 200 kV. The images for this dataset were also recorded over 48 h on a FEI Falcon II direct electron detector with a calibrated magnification of 71,000 X, yielding a pixel size of 1.9 Å. The defocus range for this dataset ranged from 0.5 to 3.5 µm. The images were recorded using 7 movie frames with a combined dose of 26 electrons per Å². This dataset yielded a ribosome map resolved to an average resolution of 4.1 Å.

2.2.8 Model building and refinement

Cryo-EM single particle analysis results in a 3D volume of the structure of interest. Model-building is a process in structure determination, in which coordinates of each atom are obtained by interpreting these 3D volumes/electron density maps. To interpret the structural features, a model (of the macromolecule) needs to be built into this density. The maps obtained from the 70S-log dataset, 70S map at 3.4 Å, 50S map at 3.2 Å, and 30S map at 3.7 Å, were used to build the model for the *M. smegmatis* 70S ribosome. For systems where there are structural homologues available, structure determination from cryo-EM maps typically starts with docking of the detectable atomic models from either X-ray crystallography or NMR to build a pseudo-atomic model. For the model building of the *M. smegmatis* 70S ribosome, *E. coli* 70S ribosome model was used as a starting template. When there were no structural homologues available, *de novo* model-building was done, and ultimately the complete 70S model for the *M. smegmatis* was built.

2.2.8.1 Model building in P/P state of *M. smegmatis* ribosome

The model for the 70S P/P state ribosome were first built into 50S and 30S maps resulting in the 50S and 30S models, which were later combined and refined into the 70S map to result in a complete 70S ribosome model for P/P state of *M. smegmatis* ribosome.

Model building for the r-proteins. The *M. smegmatis* r-protein sequences were obtained from UniProt [232]. The homology models for the *M. smegmatis* r-proteins were built with I-TASSER [169] with *E. coli* r-proteins (PDB ID: 4YBB) [233] as template for almost all r-proteins, except for modeling of bL25, bS22, and bL37. *M. smegmatis* bL25 is more than double the size of *E. coli* bL25. To model this protein, *T. thermophilus* bL25 was used as a template (PDB ID: 4V5D) [44]. Two Mycobacteria-specific proteins, bS22 and bL37, were modeled from the recently published 70S ribosome structure of *M. smegmatis* [147]. The homology models thus obtained were then rigid body fitted into the density maps obtained for LSU and SSU in UCSF Chimera [176]. To model the highly flexible parts of the map, original maps (without post-processing) were used. Each protein was extensively built in COOT [167]. COOT is an interactive 3D modeling program which is used to build and validate the macromolecular structures. The modeling program COOT, is a widely used modeling software used to build models into the electron density map. While modeling for the r-proteins, Ramachandran restraints were applied throughout. The *M. smegmatis* specific extensions were manually built into the density by using the tool ‘add terminal residue’ in COOT. During modeling in COOT, the model was refined with both the ‘Regularize Zone’ and ‘Real Space Refine Zone’ to obtain optimal fit into the EM density and maintain the correct protein geometry.

Model building for the r-RNAs. Ribosomal RNA sequences, 23S, 16S and 5S rRNA were obtained from GenBank [234]. *M. smegmatis* rRNA sequence alignment and model building with *E. coli* as template was done in ModeRNA server [170]. The rRNA homology models were first rigid body fitted into the density of the map in UCSF Chimera and then extensively rebuilt into the density with COOT. The rRNA extensions adopting a perfect RNA helix were built by placing a RNA helix model and modifying it according to the *M. smegmatis* rRNA sequence, while other parts were built by addition of the terminal residue.

Throughout model building, the tool ‘Regularize Zone’ was applied for optimal fit into the EM density.

The rRNA and r-protein models were combined to yield the complete LSU and SSU models. To further improve the fitting of LSU and SSU, both the models were refined with *phenix.refine* [172]. The refinement was performed by applying RNA base pairing restraints (generated by ‘PDB to 3D Restraints’ server) for rRNAs and secondary structure restraints for proteins with *phenix.real_space_refine*.

The models obtained for LSU and SSU were then combined and rigid body docked into the 70S map (at 3.4 Å from log phase dataset). An *E. coli* derived P-tRNA^{fMet} (PDB ID: 5AFI) [235] was fitted onto the density of P-tRNA in the 70S map. The clashes in the final model were removed in COOT. The complete 70S model was further refined with *phenix.real_space_refine* to improve the fit into the 70S map. This 70S model (refined into the 70S map at the resolution 3.4 Å) served both as a template and reference for modeling in the two other maps at 4.1 Å and 12.5 Å.

2.2.8.2 Model building in the hibernating state 70S map at 4.1 Å.

The 70S model obtained for the P/P state ribosome was used as a 70S model for hibernating state. HPF from *S. aureus* (*Sa*) (PDB ID: 5ND8) [119] was used as the template to generate homology model for *Ms* HPF with I-TASSER and rigid body docked into the density of HPF at the inter-subunit space. An *Ec* derived E-tRNA^{fMet} (PDB ID: 5AFI) [235] was fitted into the density of E-tRNA in the 70S map. *Ms* bS1 protein sequence was obtained from UniProt and a homology model was generated with Phyre2 software [173]. The best Phyre2 result was the model containing 216 residues of S1 protein from *Streptococcus pneumoniae* (PDB ID: 3GO5). *Ms* bS1 was built with unassigned UNK residues as the backbone and rigid body docked into the density observed for bS1 protein near the mRNA exit site. The clashes in the 70S model were removed in coot and the fitting of the complete 70S model into hibernating state 70S ribosome map was further refined with *phenix.real_space_refine* by applying base-pairing and secondary structure restraints for RNA and proteins, respectively.

2.2.8.3 Model building in the trans-translating state 70S map at 12.5 Å.

M. smegmatis tmRNA sequence was obtained from tmRNA database [236]. *T. thermophilus* tmRNA (PDB ID: 3IYR) [139] was used as template for tmRNA model building with ModeRNA server. *M. smegmatis* tmRNA was rigid body docked into the 12.5 Å map of the trans-translating state of ribosome. *T. thermophilus* SmpB protein (PDB ID: 3IYR) [139] was used as a template to build homology model for *M. smegmatis* SmpB protein in I-TASSER, and rigid body docked into the density for SmpB protein. An *E. coli* derived A-tRNA^{Met} (PDB ID: 5AFI) [235] was rigid body docked onto the density of A-tRNA in the 70S map.

2.2.9 Model validation

Model validation for P/P state 70S ribosome model was done with MolProbity server [171]. To validate against model over-fitting, FSC curve between the final reconstructed cryo-EM maps and the maps generated from refined atomic coordinates were computed with resolution calculated according to the FSC = 0.5 criterion. The observed FSC curves and computed FSC curves display good agreement and thus showed the absence of any model-overfitting. To show the local agreement between map and models, maps were colored using ‘*vop localCorrelation*’ in UCSF Chimera [176].

2.2.10 Figure generation

The figures for the ribosome structures were prepared in UCSF Chimera [176]. The secondary structure diagram of the tmRNA was prepared with Adobe illustrator. The local resolution plots were generated with Resmap [175].

Chapter 3

RESULTS

3 Cryo-EM structures of *M. smegmatis* ribosomes in different functional states.

The work in this thesis focuses on the structural analysis of different functional states of ribosomes in *M. smegmatis* using cryo-EM and single particle analysis. Ribosomes were purified from two different growth phases, log (24 h), and stationary (48 h) phase. The purified ribosomes from log phase (70S-log) were imaged using FEI Krios (located at National University of Singapore, Singapore) and the stationary phase ribosomes (70S-stat) were imaged in FEI Arctica (located at the Nanyang Technological University, Singapore).

Two functional states of ribosomes were identified and analyzed from the 70S-log dataset. The first state is the 70S ribosome bound with P-tRNA in P/P state. The P/P state 70S map was refined to an average resolution of 3.4 Å. This high-resolution map enabled in building a near complete 70S ribosome model of *M. smegmatis* for most of the rRNAs and r-proteins. The most interesting feature of the mycobacterial ribosome is an approximately 100 nucleotide long insertion in 23S rRNA, known as H54a, near the L1 stalk. The second state from the 70S-log dataset is the trans-translation state of 70S ribosome. The trans-translation state of ribosome is a rescue mechanism by bacteria to recover ribosomes from its stalled state, which can arise from the translation of defected mRNAs. This class was distinguishable due to the presence of the large crown shaped tmRNA wrapped around the 70S ribosome. Because of limited number of particles in this class, the trans-translating 70S map could be refined to a resolution of 12.5 Å.

From the 70S-stat dataset, we identified the 70S ribosome bound with Hibernation Promoting Factor (HPF). And this functional state of ribosome is thus called as the hibernating state. The hibernating state 70S map was refined to an average resolution of 4.1 Å. The most unique feature of this state is the presence of HPF at the inter-subunit space preventing the binding of A- and P-tRNAs and the mRNA, thus arresting ribosomes in an inactive state. Apart from this, we also observed a partial density for the bS1 protein near the mRNA exit site.

During the course of this dissertation work, two structures of *M. smegmatis* ribosomes [147, 148] and one structure of *M. tuberculosis* ribosomes [146] were published. Our structures aligned well with these mycobacterial ribosome

structures. A comparison with these structures and several unique features are presented in this and the discussion sections.

For the ease of discussion, the organisms mentioned frequently are abbreviated as; *E. coli* (*Ec*), *M. smegmatis* (*Ms*), *M. tuberculosis* (*Mtb*), *S. aureus* (*Sa*), *T. thermophilus* (*Tth*), and *L. lactis* (*Ll*).

3.1 Purification profile of *M. smegmatis* ribosomes in log and stationary phase

Crude ribosomes prepared from the log-phase and stationary phase *Ms* cultures were purified by sucrose density gradient (10-40%) centrifugation. After centrifugation, the ribosomes were fractionated (from 10% to 40%) and further concentrated over sucrose cushion. Figure 3.1 represent the ribosome purification profile of log-phase (A) and stationary-phase (B). The 70S-fractions used for cryo-EM imaging are boxed in dotted grey. In the log phase, the peaks after 70S denote the polysomes population. Presence of polysomes indicate that the ribosomes in the cell are actively translating.

Figure 3.1 Purification profile of *Ms* ribosomes in log and stationary phase. Crude ribosomes prepared from *Ms* culture in log phase (A), and stationary phase (B) are further purified by sucrose gradient centrifugation (10-40%). The samples were fractionated (from 10% to 40%) and the 70S fractions, boxed in dotted grey, are further concentrated to be imaged by cryo-EM.

3.2 Micrograph inspection

The purified ribosomes were transferred to the TEM copper grids for sample preparation and imaged with EM under cryogenic conditions. The micrographs were collected over a wide range of defocus values and corrected for the beam induced motion with the program Motioncorr [190]. Figure 3.2 shows an example of FFT before (A) and after (B) motion correction. After

motion correction, the missing information (possibly caused by drift) were recovered, as can be seen from the complete rings in the FFT. The motion corrected images were manually inspected and the micrographs with distorted FFT were deleted. Figure 3.3 shows examples of micrographs at high and low defocus values. Micrographs at different defocuses yielded in varied contrast of the particles. Though challenging to pick particles from the lower defocus micrographs, it is essential to include these micrographs as these retain the high-resolution information.

Figure 3.2 FFT showing motion correction. (A) The FFT of the raw micrograph. (B) The FFT after motion correction. As can be seen from the image, the thon rings are denser and complete in the FFT after motion correction.

Figure 3.3 Micrographs recorded at different defoci. Micrographs collected at the defocus value $0.7\ \mu\text{m}$ (A) and $2.05\ \mu\text{m}$ (B). The micrograph collected at the lower defocus range has lower contrast compared to the micrograph at higher defocus.

3.3 Image Processing

After motion correction and discarding the bad micrographs, particles were picked from the micrographs in a semi-automated mode using the software EMAN 2.1 [168]. For 70S-log dataset, 392,000 particles were picked from a total of 3503 good micrographs. And for the 70S-stat dataset, 114,282 particles were picked from a total of 748 good micrographs. The particle's coordinates were then imported into RELION 1.4 [174]. The next stages of the data processing were done using the software RELION 1.4. In the following sections, a detailed analysis of the data processing is presented for the 70S-log dataset (FEI Krios). A similar pipeline was also applied for the data processing of the 70S-stat dataset (FEI Arctica). Figure 3.4 presents an overall pipeline of the data processing for the 70S-log dataset.

Figure 3.4 Cryo-EM data processing pipeline for the 70S-log dataset. Out of 5231 collected micrographs, 3503 micrographs were selected as good micrographs showing minimal drift and no astigmatism. 392K particles were picked semi-automatically using EMAN 2.1 [168]. Images were processed further in RELION 1.4 [174]. 2X-binned particles were sorted with 2D classification. 330K good particles were selected and subjected to several rounds of 3D classification with a filtered *Ec* SecA bound 70S ribosome map as a reference. In the 3rd round of 3D classification, fine angular sampling was applied and one 3D class (with 1313 particles, 0.6% particles) was identified as the trans-translation state of ribosome. The final resolution obtained for this class after unbinned 3D refinement was 12.5 Å. The rest of the un-binned particles were 3D-refined generating a 70S map at 3.62 Å resolution. Movie frames were then included in the 3D

refinement, which yielded in a 3.41 Å resolved map for the 70S ribosome. Focused mask refinement for LSU and SSU resulted in two cryo-EM maps at an average resolution of 3.2 and 3.7 Å, respectively. Figure is adapted from [237].

3.3.1 2D classification

Particles were first sorted into good or bad particles, by classifying them into different groups or classes, known as 2D classification. The 2D classification was done with 300 classes for the 70S-log dataset (Figure 3.5). A higher number of classes was applied here to sort the particles efficiently, which helped in the elimination of most of the bad particles/classes. Out of approximately 392,000 particles, the good classes from 2D classification resulted in a total of 330,000 particles.

Figure 3.5 2D classification of 70S-log dataset. The particles are classified into 300 classes. The good classes are easily distinguishable in the top five rows, showing distinct structural features and better contrast for the ribosome, while the bad classes neither show any contrast nor any distinguishable features.

3.3.2 3D classification and 3D refinement

The good particles selected from the 2D classification were further classified into six classes during 3D classification (Figure 3.6). Out of these six classes, the first five good classes (about 270,000 particles) were further classified into different classes applying different parameters (Figure 3.4). The particles from the five good classes were used for 3D refinement yielding a 70S ribosome map at 3.62 Å. Addition of the movie frames improved the resolution

to 3.41 Å for the 70S ribosome map. At this stage, focused masked refinement is also done on the two subunits which resulted in two maps at 3.2 Å and 3.7 Å for LSU and SSU, respectively.

Figure 3.6 3D classification of 70S-log dataset. The first five classes (from left) were selected for the further 3D classification. The 3D classes obtained were viewed in chimera as 3D volumes before deciding to proceed with the specific class or not. The sixth class was discarded as the 3D volume did not look good (shown in Figure 3.4).

During the third round of 3D classification, one class (approximately 1000 particles) was uniquely distinguishable from the rest of the 3D classes because of the presence of a large crown shaped tmRNA wrapped around the head of the 30S (Figure 3.4). Because of the limited number of particles, this 70S state could only be resolved to 12.5 Å.

For the 70S-stat dataset, out of 114,282 particles selected, 91,303 particles resulted in one major class during 3D classification. During 3D refinement with the movie frames, this 70S state was resolved to a 4.1 Å map of the hibernating state ribosome with additional densities observed at the inter-subunit space and near the mRNA exit site (when compared with the P/P state 70S ribosome).

3.3.3 Resolution estimation and Resmap calculation for the 70S P/P state maps

The three P/P state maps, 30S, 50S, and 70S were used for the model building of the *Ms* 70S ribosome. The resolutions of the maps were calculated using the gold standard Fourier Shell Correlation (FSC) = 0.143 criterion [219]. The local resolution of the map was calculated using Resmap [175]. Figure 3.7 shows the Resmap calculation for the 30S, 50S, and 70S maps for the P/P state 70S *Ms* ribosome. Generally, the resolution at the core of the ribosomes (showed in spiced view in figure 3.7 A-C, right side) is higher than the ribosome periphery. The regions shown in colors yellow and red are at lower resolution (5-6 Å), as these regions are highly flexible in the ribosome. To model into such flexible

regions, filtered maps were used (whenever needed), thus enabling in the modeling of the complete 70S *Ms* ribosome.

The 70S model of the *Ms* ribosome at the P/P state was used as a template to model the 4.1 Å hibernating state and the 12.5 Å trans-translating state.

Figure 3.7 Estimation of resolution and validation against model-overfitting. (A-C) Local resolution estimation with ResMap [175] for the 50S and 30S subunits, and 70S ribosome. Both solvent side views (left) and sliced views (right) are shown. The core regions across all three maps are well-resolved (sliced view) compared to the peripheral regions. (D-F) Average resolution of the final refined cryo-EM maps and validation against model-overfitting. Red: Fourier Shell Correlation (FSC) curves for 50S, 30S and 70S maps with estimated average resolutions of 3.3, 3.7, and 3.4 Å, respectively, calculated according to the FSC = 0.143 criterion [219]. Black: FSC curve computed between the final cryo-EM maps and maps generated from refined atomic coordinates with resolution calculated according to the FSC = 0.5 criterion. Figure is adapted from [237].

The quality of the maps obtained for the P/P state were good for modeling of the most of the ribosome. A comparison of the map quality between the 3.4 Å map (P/P state) and 4.1 Å map (hibernating state) is shown in the figure 3.8.

Figure 3.8 Quality of the P/P state (3.4 Å) and hibernating state (4.1 Å) maps. (A and D) A part of 23S rRNA (Helix 11) showing the quality of map and fitting of nucleotides. (B and E) Density and fitting of a small part of 23S rRNA (C2271-A2305) extracted in chimera. (C and F) Density and fitting of LSU protein bL28 extracted in chimera. A-C, P/P state (3.4 Å) map; D-F hibernating state (4.1 Å) map. Figure is adapted from [237].

3.3.4 Model Validation

The model validation for the P/P state 70S ribosome model was done with the server MolProbity [171]. Table 3.1 details the model validation statistics for the P/P state models.

Field	SSU	LSU	70S Ribosome
Data Collection			
Particles	391,837	391,837	391,837
Pixel Size (Å)	1.11	1.11	1.11
Defocus Range (µm)	0.3 – 4.3	0.3 – 4.3	0.3 – 4.3
Voltage (kV)	300	300	300
Electron Dose (e ⁻ Å ⁻²)	35	35	35
Model Composition			
Non-hydrogen atoms	52,674	97,108	148,419
Protein residues	2,406	3,687	6,093
RNA bases	1,506	3,207	4,713
Model Refinement			
Resolution (Å)	3.28	3.7	3.4
Map sharpening B-factor (Å ²)	-134.86	-85.95	-84.62
CC _{map_model}	0.789	0.775	0.804
RMS deviation			
Bonds (Å)	0.02	0	0.01
Angles (°)	1.04	0.71	1.03
Validation (proteins)			
Molprobity score	2.22 (63 rd percentile)	2.64 (39 th percentile)	2.57 (43 rd percentile)
Clashscore, all atoms	16.04 (45 th percentile)	16.47 (43 rd percentile)	18.45 (36 th percentile)
Poor rotamers (%)	0.90	3.4	2.39
Favoured rotamers (%)	94.05	90.42	91.89
Ramachandran plot			
Favoured (%)	91.08	90.91	90.98
Outliers (%)	0.42	1.13	0.85
Validation (RNA)			
Correct sugar puckers (%)	98.8	99.35	99.77
Good backbone conformations (%)	74.04	75.05	74.72
Bond Outliners (%)	0	0	0
Angle Outliners (%)	0.02	0.01	0.01

Table 3.1 Model refinement statistics for the Ms 50S and 30S ribosome subunits and the complete 70S ribosome in P/P state. Table is adapted from [237].

In cryo-EM, it is essential to validate model over-fitting. This is achieved by calculation the FSC curve between the final reconstructed cryo-EM maps and the maps generated from the refined atomic coordinates, computed with resolutions calculated according to the FSC = 0.5 criterion. For the P/P state maps, the observed FSC curves and computed FSC curves display a good

agreement and thus showed the absence of any model-overfitting (Figure 3.7). The local agreement between the map and model can be shown by 'vop localCorrelation' estimated in UCSF Chimera [176], and shown in the figure 3.9 for the P/P state maps. The pixel size for both the maps were also reassessed and the best fitting was achieved at pixel size 1.11 Å for P/P state 70S map and translating state 70S map and 1.9 Å for the hibernating state 70S map.

Figure 3.9 Local agreement between the cryo-EM maps and models. The 50S subunit (A), 30S subunit (B) and 70S (C) cryo-EM maps from P/P state are colored based on the local correlation values between the cryo-EM maps and respective atomic models (calculated using UCSF Chimera [176]) to indicate the presence of minimal overfitting in the final models. Figure is adapted from [237].

3.3.5 Resolution Estimation for the Hibernating and Trans-translating maps

The resolution for the two maps of the trans-translation state and hibernating state were also calculated using the gold standard Fourier Shell Correlation (FSC) = 0.143 criterion [219], and the FSC curves are shown in the Figure 3.10.

Figure 3.10 Resolution estimation for the trans-translation and hibernation states 70S ribosome maps. FSC curves for (A) trans-translation and (B) hibernation state 70S ribosome map as indicated, using FSC= 0.143 criterion [219]. Figure is modified from [237].

In the following sections, the detailed structural analysis is presented for the *M. smegmatis* 70S ribosomes in P/P state, trans-translating state and hibernating state.

3.4 Structure of *M. smegmatis* 70S ribosome in P/P state

In the field of ribosome structures, the ribosomes from *Ec* and *Tth* have been well characterized, and often compared with, when the ribosome from a new organism is analyzed. *Ms* ribosome shares the conserved ribosomal architecture with *Ec* ribosome. The 70S ribosome model of *Ec* (PDB ID: 4YBB) [233] which was solved at high-resolution, fitted well in our map and was used as the template for modeling of the r-proteins and rRNAs of *Ms* 70S. Homology models were obtained with ITASSER [169] and ModeRNA [170] servers and were subsequently used for modeling and fitting with COOT [167] and UCSF Chimera [176] (as described in the section 2.2.8).

Sequence comparison with *Ec* showed that several of the r-proteins and rRNAs of *Ms* showed prominent extension. The major extensions observed in the P/P state are highlighted in the figure 3.11.

Figure 3.11 Unique structural features of the *Mycobacterium smegmatis* 70S ribosome. (A) Cryo-EM map of *Ms* 70S ribosome at 3.4 Å resolution in the non-rotated P/P state. Density is shown as surface and colored as- rRNA: grey, r-proteins: blue and P-tRNA: light green. *Ms*-specific rRNA and r-protein extensions (in comparison to *Ec*, PDB: 4YBB) [233] are shown in deep-red and deep-green, respectively. (B) LSU model of *Ms* 70S from the solvent-side. *Ms*-specific protein extension of bL25 and repositioning of bL9 (in green) are shown in the zoomed view (blue color) with superimposition of respective *Ec*-homologs in grey. Mycobacteria-specific protein bL37 is colored in yellow. (C) SSU model of *Ms* 70S from the solvent-side. *Ms*-specific protein extensions for uS5, uS16, uS17 and bS18 (in green) are shown in zoomed view (blue color) with superimposition of respective *Ec*-homologs in grey. Mycobacteria-specific protein bS22 is colored in yellow. Figure is adapted from [237].

3.4.1 *M. smegmatis* small ribosomal subunit

***Ms* SSU r-proteins.** *Ms* SSU has 21 r-proteins which are bS1, uS2, uS3, uS4, uS5, bS6, uS7, uS8, uS9, uS10, uS11, uS12, uS13, uS14, uS15, uS16, uS17, bS18, uS19, bS20, and bS22. However, in contrast to *Ec* SSU, protein bS21 is missing in *Ms* and one additional protein, bS22, is present near the decoding center keeping the overall number of r-proteins in *Ms* SSU similar to *Ec*. In P/P state 70S map, we could identify almost all the SSU r-proteins, except bS1.

Although, partial density for the N-terminal domain of bS1 was observed when the P/P state map was filtered to 10 Å (discussed in the section 3.5). The core structure of *Ms* SSU remains similar to *Ec*. The overall architecture of the *Ms* SSU showing the rRNA and r-proteins is represented in the figure 3.12.

Figure 3.12 Architecture of the *Ms* SSU. The overall architecture of the *Ms* SSU shown in the form the density of the r-proteins and rRNA (**A**) and protein and rRNA models (**B**) as seen from the solvent view (left) and interface view (right). The proteins are colored and labelled as shown in the figure. H-head, Pt-platform, B-beak/neck, F-foot. The densities of the r-proteins and 16S rRNA are rendered at the threshold value of 0.07 and radius 2.0 in UCSF chimera [176], except for the density of the uS2 (threshold: 0.04, radius 3.0).

Sequence comparison showed that seven r-proteins (uS2, uS3, uS5, uS9, uS16, uS17, and bS18) of *Ms* SSU are longer by 10 or more residues when compared to their *Ec* homologs because of the presence of the N- and/or C-terminal extensions (Table 3.2).

Protein/ RNA name	Chain ID in PDB ID:	Uniprot ID (GenBank ID)	Complete size	Modeled	Homologous <i>Ec</i> size	comments
16S rRNA	a	X52922.1	1528	9-1518	1541	
uS2	b	A0QVB8	277	1-228	241	
uS3	c	A0QSD7	275	1-210	233	
uS4	d	A0QSL7	201	2-201	206	
uS5	e	A0QSG6	214	17-214	167	
bS6	f	A0A0D6J3X3	96	1-96	135	
uS7	g	A0QS97	156	1-156	179	
uS8	h	A0QSG3	132	3-132	130	
uS9	i	A0QSP9	150	25-150	130	
uS10	j	A0QSD0	101	5-101	103	
uS11	k	A0QSL6	138	22-138	129	
uS12	l	A0QS96	124	2-123	124	
uS13	m	A0QSL5	124	2-117	118	
uS14	n	A0QSG2	61	2-61	101	
uS15	o	A0QVQ3	89	2-88	89	
uS16	p	A0QV37	156	2-114	52	
uS17	q	A0QSE0	98	6-97	84	
bS18	r	A0R7F7	84	16-79	75	
uS19	s	A0QSD5	93	6-83	92	
bS20	t	A0R102	86	3-86	87	
bS22	u	A0QR10	33	2-33	-	
P-site tRNA ^{fMet}	v	CP011124.1	77	1-77	77	Ec P- tRNA ^{fMet}

Table 3.2 Composition of P/P state SSU model. The r-protein IDs are from Uniprot and the rRNA IDs are from Genbank. The chain IDs are according to the PDB ID: 5ZEU. Table is adapted from [237].

Figure 3.13 details the individual SSU r-proteins modeled in the P/P state of 70S ribosome map in comparison to the *Ec* SSU r-proteins. In comparison to *Ec* SSU r-proteins, *Ms* uS9, uS17, and bS18 are extended at their N-terminal by 20, 14, and 10 residues, respectively. Similarly, uS2, uS3, and uS16 are extended at their C-terminal by 50, 42, and 74 residues, respectively. *Ms* uS5 is extended

at both N-terminal (33 residues) and C-terminal (22 residues). We could localize the extensions for uS5, uS16, and uS17 in our map while the other extensions are not visible possibly due to the flexibility of these regions. These extensions are discussed in detail in the following sections.

Figure 3.13 Comparison of Ms SSU r-proteins with the Ec counterparts. On the right, Ms SSU r-proteins are colored in deep red and shown with their respective density as light grey surface; On the right, Ms SSU r-proteins (deep red) are super positioned with the respective Ec r-proteins (grey), showing the mycobacteria specific extensions (deep green).

r-protein uS5. *Ms* uS5 is extended both at its N- and C-terminal (in comparison to *Ec*). The C-terminal extension (CTE) of *Ms* uS5 forms an extra α -helix which is localized on the SSU surface (Figure 3.11, Figure 3.13, and Figure 3.14). This helix might impart additional stability to the mycobacterial SSU surface. The density for N-terminal extension (NTE, 33 residues) is not visible clearly in the P/P 70S state. However, it could be partially (residues 17-33) localized in the hibernating state 70S map. uS5 encircles the mRNA entry site and plays an important role in subunit assembly and maintaining translation fidelity [238]. Considering the position of uS5 at the mRNA entry site and its known critical function in subunit assembly and translation fidelity, it seems probable that these Mycobacteria-specific N- and C-terminal extensions in uS5 are required for regulation of translation initiation in the Mycobacteria-specific environment. uS5 in the context of mRNA entry site is discussed in detail in the section 3.5.2.

Figure 3.14 *Ms r-protein uS5.* (A) Model of uS5 modeled in the hibernating state map. (B) A comparison of the model of *Ms* uS5 (deep red) and *Ec* uS5 (grey) showing the additional CTE helix and NTE loop. (C) Local resolution estimation with Resmap [175] for uS5 from the P/P state with the resolution shown in Å in the scale bar.

r-protein uS16. *Ms* uS16 is extended by 74 residues at the C-terminal in contrast to *Ec*. The density for the first 30 residues (out of 74 residues) of CTE in uS16 could be localized in the P/P state 70S map. This region of CTE is extended at the beginning and ends with a short helical conformation further away from the core of uS16 where it reaches towards uS4 to interact with it (Figure 3.15). The Mycobacteria-specific interaction between uS16 and uS4 is different from *Ec*, where these two proteins do not directly interact to each other. uS4 is essential for the binding of uS16 to SSU. Both uS4 and uS16 proteins are assumed to be

indispensable for the 30S ribosome assembly [239, 240]. Also, the uS4 together with uS17 and bS20, are considered as the early ribosome assembly proteins [239]. And this Mycobacteria-specific interaction between uS16 and uS4 might provide additional stability during SSU assembly and maturation in these organisms.

Figure 3.15 *r*-protein uS16 interaction with uS4 in *Ms* 30S. (A) The CTE of uS16 (pink) extends towards the mRNA channel and interacts with the protein uS4 (orange). (B) Local resolution estimation with Resmap [175] for uS16 from the P/P state with the resolution shown in Å in the scale bar.

***r*-protein uS17.** *Ms* uS17 is extended at its N-terminal. In the P/P state, the NTE (modeled residues: Gly 6 to Gly 18) is extended to interact with the helix 9 (h9) of 16S rRNA (Figure 3.16 A). In Mycobacteria, the h9 has 18 nucleotides insertion (compared to *Ec*), which seems to be stabilized by the interaction with *Ms* uS17. As mentioned before, the uS17 belongs to the early ribosome assembly proteins, and thus needs to interact extensively with the 16S rRNA during the assembly of the 30S subunit. Thus, the modification of the uS17 N-terminal in Mycobacteria could be a result to accommodate and stabilize the 16S rRNA h9 alteration.

Figure 3.16 r-protein uS17 interactions with 16S rRNA in *Ms* 30S. (A) The NTE of uS17 (green) extends to interact with the 16S rRNA helix 9 (h9-deep yellow) stabilizing it at this region in the *Ms* 30S. (B) Local resolution estimation with Resmap [175] for uS17 from the P/P state with the resolution shown in Å in the scale bar.

r-protein bS21. r-protein bS21, a constituent part of mRNA exit site of *Ec* SSU, is missing from *Ms* genome. Consistently, we did not observe any density for bS21 at this location in our map when compared with *Ec* ribosome (Figure 3.17). The density for bS21 was also not observed in the *Ms* and *Mtb* ribosome structures published recently [146-148]. bS21 is a late ribosome binding protein. A viable Δ bS21 *Ec* strain shows that it can be considered as a non-essential protein for ribosome translation [241, 242]. In *B. subtilis*, the Δ bS21 strains display slow growth phenotype than the wild type under normal laboratory conditions [243]. In the genus *Mycobacterium*, a homolog of bS21 is present in *Mycobacterium abscessus*, an early mycobacterial species, and in *Mtb*. However, *Mtb* bS21 protein was not localized in the recently published 70S ribosome structure of *Mtb* [146].

Figure 3.17 Comparison of *Ms* mRNA exit site with *Ec*. Models of the *Ms* (light blue) and *Ec* ribosomes (grey, PDB ID: 4YBB) [233] are super-positioned to show the overall conserved architecture of the mRNA exit site. Protein bS21 (highlighted in deep-grey) is absent in *Ms*. The LSU model is hidden for clarity. Figure is modified from [237].

Ms SSU r-proteins bS6, uS7, and uS14 are significantly shorter in length by more than 20 residues than their *Ec* homologs because of the truncation at N- or C-terminal (Table 3.2). However, these deletions are not substituted by any alterations in their surrounding environment.

***Ms* SSU rRNA.** The *Ms* SSU has one rRNA molecule, 16S rRNA which is 1528 nucleotides (nt) long (*Ec* 16S rRNA is 1541 nt). The *Ms* 16S rRNA is fairly conserved when compared with *Ec* 16S rRNA except for helix 9 and 17. Structure of *Ms* h9 is altered by an 18 nt insertion (185 G – 202 G) when compared to *Ec* (Figure 3.18). The *Ms* uS17 NTE seems to stabilize this altered h9 in *Mycobacteria*. *Ms* h17 has 21 nucleotides deletion (compared to *Ec*) (Figure 3.18). This deletion is not substituted by any insertion by other rRNA helices or r-proteins. Apart from this, several small insertions and deletions are present throughout the RNA structure, but the overall topology of the rRNA remains similar to *Ec*.

Figure 3.18 *Ms* 16S rRNA secondary structure comparison with *Ec*. Secondary structure of *Ms* 16S rRNA with rRNA extension and deletion compared with *Ec*. Extension/deletion are shown in a zoomed view in the boxes with respective helix numbers.

3.4.2 *M. smegmatis* large ribosomal subunit

***Ms* LSU r-proteins.** *Ms* LSU has 34 r-proteins (uL1, uL2, uL3, uL4, uL5, uL6, bL7/bL12, bL9, uL10, uL11, uL13, uL14, uL15, uL16, bL17, uL18, bL19, bL20, bL21, uL22, uL23, uL24, bL25, bL27, bL28, uL29, uL30, bL31, bL32, bL33, bL34, bL35, bL36, and bL37) and two rRNA molecules, 23S rRNA (3120 nt) and 5S rRNA (118 nt) (Figure 3.19). Compared to *Ec*, *Ms* LSU has one additional protein bL37 (Figure 3.11). We were able to localize and model 32 r-proteins of LSU except for the stalk proteins uL1 and bL7/bL12. In comparison to *Ec* LSU, six out of 32 r-proteins showed significant extensions (above 10 residues) in *Ms* LSU. uL4, bL17, bL20, uL22 and bL25 are extended at the C-terminal by 12, 65, 11, 35, and 121 residues, respectively (Figure 3.20). uL18 is extended at the N-terminal by 11 residues (Table 3.3). Out of these, only density for bL25 extension is clearly visible in our map. The overall architecture of the *Ms* LSU is represented in the figure 3.19.

Figure 3.19 Architecture of the Ms LSU. The overall architecture of the Ms LSU shown in the form of the density of the r-proteins and rRNA (**A**) and protein and rRNA models (**B**) as seen from the interface view (left) and solvent view (right). The proteins are colored and labelled as shown in the figure. CP-central protuberance, L1-S- L1 stalk, and L7/12-S- L7/L12 stalk. The densities of the r-proteins and 16S rRNA are rendered at the threshold value of 0.07 and radius 2.0 in UCSF chimera [176].

Protein/ RNA name	Chain ID in LSU	Uniprot ID (GenBank ID)	Complete size	Modeled	Homolo- gous <i>Ec</i> size	Com- ments
23S rRNA	A	NR_076104.1	3120	2-3119	2897	
5S rRNA	B	NR_075650.1	118	1-117	120	
uL1	-	A0QS46	235	-		not visible
uL2	C	A0QSD4	278	3-275	273	
uL3	D	A0QSD1	217	2-215	209	
uL4	E	A0QSD2	215	3-209	201	
uL5	F	A0QSG1	187	6-186	179	
uL6	G	A0QSG4	179	2-177	177	
bL7/bL12	-			-		not visible
bL9	H	A0R7F6	151	1-151	149	
uL10	I	A0QS62	175	1-126	165	
uL11	J	A0QS45	142	10-142	142	
uL13	K	A0QSP8	147	1-147	147	
uL14	L	A0QSF9	122	1-121	123	
uL15	M	A0QSG8	147	3-147	144	
uL16	N	A0QSD8	138	1-134	136	
bL17	O	A0QSL9	199	2-118	127	
uL18	P	A0QSG5	127	2-127	117	
bL19	Q	A0QV42	113	1-113	115	
bL20	R	A0QYU6	129	2-125	118	
bL21	S	A0R151	103	2-103	103	
uL22	T	A0QSD6	153	6-119	110	
uL23	U	A0QSD3	100	4-97	100	
uL24	V	A0QSG0	105	1-105	104	
bL25	W	A0R3D2	215	6-193	94	
bL27	X	A0R150	88	5-86	85	
bL28	Y	A0QV03	64	2-64	78	
uL29	Z	A0QSD9	77	5-57	63	
uL30	1	A0QSG7	61	2-61	59	
bL31	2	A0R215	82	1-66	70	
bL32	3	A0R3I9	57	2-55	57	
bL33	4	A0QS39	55	6-55	55	
bL34	5	A0R7K0	47	3-47	46	
bL35	6	A0QYU7	64	2-64	65	
bL36	7	A0QSL4	37	1-37	38	
bL37	8	A0QTP4	24	2-24	-	

Table 3.3 Composition of P/P state LSU model. The r-protein IDs are from Uniprot and the rRNA IDs are from Genbank. The chain IDs are according to the PDB ID: 5ZET. Table is adapted from [237].

Figure 3.20 Comparison of *Ms* LSU r-proteins with the *Ec* counterparts. On the right, *Ms* LSU r-proteins are colored in deep red and shown with their respective density as light grey surface; On the right, *Ms* LSU r-proteins (deep red) are super positioned with the respective *Ec* r-proteins (grey), showing the mycobacteria specific extensions (deep green).

***r*-protein bL25.** *Ms* bL25 is more than double in size (215 residues) compared to its *Ec* homolog (94 residues) because of its long C-terminal extension. This extension folds into a C-terminal domain (CTD) (Figures 3.11 and 3.21) similar to other bacteria (e.g. *B. subtilis* and *Tth*) known to have a longer version of bL25. Model of *Ms* bL25-CTD was built using *Tth* bL25 (PDB ID: 4V5D) as a template which fitted well in our map [44]. *Ms* bL25 with two functional domains (NTD and CTD) belongs to the catabolic-controlled (CTC) protein family. The NTD interacts with the loop E of 5S rRNA and binds to the ribosome to form a part of the central protuberance [244]. The CTD is sandwiched between uL16 and H38 (Figure 3.21). The structure of *Ms* bL25 CTD is similar to *Tth* bL25 CTD. In *Tth*

the bL25-CTD assists in tRNA proofreading [245]. While CTC proteins are known to interact extensively with 5S rRNA, bL25 deletion reduces cell growth by several folds [246]. Figure 3.21 displays the two domains of *Ms* bL25, where the NTD is seen interacting with the 5S rRNA and the CTD with H38 and uL16.

Figure 3.21 *Ms* bL25 with its two functional domains. The *Ms* bL25 has two functional domains, the NTD interacts with the 5S rRNA, and the CTD is sandwiched between the uL16 (green) and H38 of 23S rRNA.

3.4.3 Species specific alteration in the 23S rRNA of the LSU in the *M. smegmatis* ribosome

The total length of rRNA in *Ms* LSU is 3238 nt and is significantly larger than *Ec* LSU rRNA (3024 nt). The total length of *Ms* 23S rRNA is 3120 nt and is 216 nt longer than *Ec* 23S rRNA. Whereas *Ms* 5S rRNA is just 2 nt shorter than its *Ec* counterpart. Sequence comparison revealed that *Ms* 23S rRNA and 5S rRNA are 65.6% and 58.5% identical to the *Ec* 23s rRNA and 5S rRNA, respectively. The *Ms* 23S rRNA can be divided into six domains similar to *Ec* 23S rRNA based on its function in ribosome translation. Compared to *Ec*, there are three major insertions in domain I, II and III which have altered the topology of the 23S rRNA structure in mycobacterial ribosome when compared to *Ec*. These insertions are H15 and H16a at the L1 stalk base (domain I), H31a at the base of the central protuberance (domain II), and H54a located near the mRNA exit site and L1 stalk (domain III). Other than these three major alterations in the rRNA topology of *Ms*, there are various small insertions/deletions throughout the RNA structure in a manner that insertions are balanced by deletions and vice versa, maintaining the overall topology of rRNA similar to *Ec* in those regions.

H15 and H16a: The first rRNA extension exists in domain I at the L1 stalk base in the H15 (A272 - U318). H15 is seen in the *Tth* 23S rRNA and *Ms* H15 is extended by 36 nt in comparison to *Tth* H15. *Ms* H16 has 21 nt insertion in comparison to *Ec* H16 which give rise to a new helix H16a. *Ms* H15 and H16a are two complementary RNA helices exhibiting a RNA kissing loop interaction in the mycobacterial ribosome. This unique rRNA complementary interaction has also been reported in the recent mycobacterial ribosome structure [147] (Figure 3.22). The density observed for these two helices is weak in our map and thus the filtered maps were used to build the models in this region.

Figure 3.22 *Ms* 16S rRNA showing H15-H16a. (A) The H15-H16a kissing loop interaction seen in *Ms* 16S rRNA, shown in the secondary structure of H15-H16a (left) and in the model of H15-H16a (right). (B) The H15-H16a kissing loop interaction with the ribosome density. The P/P state map is filtered to 6 Å to show the density in this region. The rest of the LSU is hidden in the background for clarity.

The altered topology of the H15-H16a is complemented with a different orientation of bL9 near the L1 stalk. *Ms* bL9 shows a flipped orientation compared to *Ec* and gets accommodated near the *Ms* H15-H16a as shown in the figure 3.23. The long helix of *Ec* bL9 (Lys 41 - Leu 75) is kinked in *Ms* into two regions (Gly 42 - Val 58; Glu 63 - Leu 72) connected by a linker (Ile 59 - Ile 62). This broken helix in *Ms* assists in the flipped orientation of the *Ms* bL9 at its new position in the *Ms* LSU. *Ms* bL9 long helix (Gly 42 - Val 58) is kinked in comparison to *Ec*, and the rest of the protein flips on this kink (Figure 3.23). This

flipped orientation will also help in the accommodation of the H54a, the rRNA insertion in the domain III of the *Ms* 23S rRNA. RNA kissing loop interactions provide unusual RNA stability in RNA molecules (as seen in human immunodeficiency virus, HIV). Such interactions have been exploited as potential drug targets. Such Mycobacteria-specific interactions will therefore, be more interesting to investigate in pathogenic mycobacterial species.

Figure 3.23 Depiction of H15-16a with *Ms* bL9. (A) LSU r-protein bL9 in *Ms* is flipped inwards to accommodate the H54a extension (also shown in figure 3.26). The H5-H16a shown in different shades of blue, Ec bL9 (PDB ID: 4YBB) [233] in grey, *Ms* bL9 in dark-red and rest of the LSU in light-blue. The rest of the ribosome is hidden for clarity. (B) Model of Ec bL9 (grey) and *Ms* bL9 (dark-red) showing the location of the kinked helix which assists in the flipping of the *Ms* bL9 in *Ms* 70S ribosome.

H31a: At the base of central protuberance, a 23 nt long insertion at position U748 in *Ms* H31 results in a new helix H31a (Figure 3.24 A). This insertion extends from the base of the central protuberance to interact with bL27 (Figure 3.24 B). *Ms* bL27 has an extended N-terminal tail which reaches near the inter-subunit space to interact with P- and A-tRNA. The flexible N-terminal extension of bL27 interacts with the peptidyl transferase centre (PTC) and is responsible for the correct positioning of tRNA substrates. Mutations or deletions at this N-terminal tail severely affects the functioning of the PTC [60]. The overall architecture of rRNA and r-proteins is well conserved in this region except for this insertion. This insertion is conserved across Mycobacteria and is 4 nt longer in *Mtb* [147].

Figure 3.24 Ms 16S rRNA showing the H31a insertion. (A) The secondary structure of the Ms 23S rRNA compared with the Ec 23S rRNA, showing the H31a insertion (highlighted in red). (B) The interaction of the H31a (deep blue) with the r-protein bL27 (dark-red). The bases A758 and G759 interact with the residues R85 and P86. The 23S rRNA is shown in light blue, and the rest of the 70S model is hidden for clarity.

H54a: H54a is a Mycobacteria-specific 113 nt long insertion in H54 at position G1532 which emerges at the solvent side of LSU near the L1 stalk, in the domain III of the 23S rRNA (Figure 3.25). This is unequivocally the most prominent feature of mycobacterial ribosomes (Figure 3.11 A). This helix was first reported as a structure named steeple [144] and is recently renamed as Helix 54a [147] following the new nomenclature for ribosome [247]. Species-specific unusual long r-RNA extensions have been observed before, such as in yeast 80S ribosomes (ES6S, 200 nt long), and in *Plasmodium falciparum* (ES6S and ES10S) [248, 249]. Such extensions could impart species-specific functions in ribosome translation or regulation. The helix H54a is conserved among Mycobacteria species and has also been reported in the recent *Mtb* 70S ribosome structure [146].

Figure 3.25 Ms 16S rRNA showing the H54a insertion. The secondary structure of the Ms 23S rRNA compared with the Ec 23S rRNA, showing the H54a insertion (highlighted in red).

H54a is highly flexible and the observed density was fragmented near the tip of the helix in the P/P state 70S ribosome and could only be resolved to the resolution of 5 to 6 Å. Its close proximity to the mRNA exit site and L1 stalk indicates its possible function in translation initiation and/or regulation. Because the mRNA exit site is the site of translation initiation where the protein bS1 binds the ribosome during translation initiation state and the L1 stalk is responsible for subunit rotation during translation, the H54a located at this site might also have related functions. r-proteins bL9, bS6, and uS11 are also present in the close proximity of H54a. *Ms* bL9 shows a different orientation compared to its *Ec* counterpart. *Ms* bL9 is flipped towards to H54a to accommodate the rRNA extension H54a. The SSU protein bS6 shares surface contact with H54a at the position Arg17, creating a small new inter-subunit interaction (Figure 3.26).

Figure 3.26 Species-specific modifications near H54a. (A) *Ms* bL9 exists in a flipped orientation around H15-H16a, in a manner to accommodate the flexible nature of H54a in the *Ms* LSU. (B) SSU r-protein bS6 shares surface contact with H54a via Arg17 in the *Ms* P/P state 70S ribosome.

To further understand the flexibility and movement of H54a, the H54a was compared between the P/P state and hibernating state 70S ribosomes along with the recently published *Mtb* 50S [146]. The H54a moves 16.4 Å towards the mRNA exit site in the hibernating state to interact with the protein *Ms* bS1. This was observed when the P/P state and hibernating state 70S models were overlaid. Interaction of H54a with the bS1 near the mRNA exit site is a direct indication that H54a might influence the translation initiation in Mycobacteria. This interaction might be responsible for a stable 70S initiation complex formation during translation initiation. Surprisingly, the overall rotation of the 50S and 30S remains unchanged, except for the movement of the H54a. This interaction

stabilizes the hibernating ribosome in the 70S state and might promote ribosome resuscitation from hibernation. Also, the interaction of 23S rRNA with bS1 has never been seen before, and might be a species-specific interaction observed in mycobacterial ribosome occurring via H54a. In *Mtb* 50S, the H54a gets localized in the inter-subunit space and predicted to prevent the formation of the 70S ribosomal initiation complex [146]. The prevention of the formation of 70S by H54a in mycobacterial ribosome might be an accessory to protect the 50S in the cell against unfavourable cellular environments. It might impart additional stability for the pathogenic mycobacterial ribosome to survive under dormant conditions for a prolonged period. During the event of translation, the H54a has to move about 58.2 Å out towards the mRNA exit site, allowing the successful formation of the 70S initiation complex [146]. A comparison of the movement of H54a from the 50S complex with the *Ms* P/P state and movement of H54a towards the mRNA exit site to interact with the bS1 protein is shown in the figure 3.27.

Figure 3.27 H54a movement comparison between P/P state, hibernating state and Mtb 50S. The Ms P/P state, Ms hibernating state, and Mtb 50S (PDB ID: 5V7Q) [146] maps and models were superimposed and the movement of the H54a was then compared. Between Ms P/P state and Mtb 50S, the entire helix H54a displaces about 58.2 Å (32°) to form a stable 50S in the cell. Between the Ms P/P state and Ms hibernating state, the helix movement is restricted towards the tip and displaces about 16.4 Å (15°) to form a stable hibernating 70S state. The mRNA exit site is denoted by (*). The movement of the

H54a is shown by the dotted red arrow. The H54a of the Ms hibernating state, Ms P/P state, and Mtb 50S are shown in deep red, deep grey, and deep green, respectively.

Interestingly, the H54a movement observed in the hibernating state could have been driven from the presence and interaction with bS1 protein near the mRNA exit site (Figure 3.27). The density for the protein bS1 was not visible in the high-resolution map of the P/P state of 70S, but was seen in the hibernating state 70S map (discussed in the section 3.5). Also, the interaction of bS1 with 23S rRNA is not observed before in the bacterial ribosome before, and is only visible in Mycobacteria due to the long 113 nt extension, H54a. In the hibernating state 70S ribosome, H54a moves about 16.4 Å (about 15° angular tilt at A1564 – G1605) towards the mRNA exit site to interact with the protein bS1 observed in the hibernating state 70S.

The localization of the H54a near the tRNA ejection site and mRNA exit site in the P/P state suggest a possible function of Mycobacteria-specific H54a rRNA in translation elongation (involved in tRNA ejection).

3.4.4 Species-specific alteration in r-proteins clustered near the functional sites of *M. smegmatis* ribosome

In the cryo-EM map of *Ms*, additional densities for two new proteins were observed near the decoding center (DC) in SSU and peptidyl transferase center (PTC) in LSU, not present in *Ec*, *Tth*, and several other bacterial ribosomes. The DC and PTC of the bacterial ribosome are highly conserved in their rRNA and r-protein composition. Therefore, any alterations in these or near these regions might cause species-specific differences in translation. Structural information of such alterations from Mycobacteria could be exploited for the discovery of novel drug targets. In the *Ms* 70S ribosome map, both these densities could fit a small protein with a helical geometry.

The first density was observed near the DC where a small protein, bS22, was fitted [146-148]. bS22 is a small, 33 amino acid long protein located in a rRNA pocket formed by h27, h44, and h45 of 16S rRNA and H70 of 23S rRNA (Figure 3.28). Because of the presence of arginines and lysines, the bS22 is highly basic (pI: 12.9) and shares numerous surface contacts with the surrounding rRNAs.

Figure 3.28 Location of the *Ms* bS22 in its rRNA environment. *bS22* is located near the DC and is surrounded by the 16S rRNA helices h27, h44, and h45 and the 23S rRNA helix H70. The rest of the ribosome model is hidden for clarity.

This protein is conserved through Mycobacteria and has also been located in the *Mtb* ribosome [146]. The sequence of *Ms* bS22 is similar to the *Mtb* bS22. In *Mtb*, the sequence of this protein is listed under unknown in the Tuberculist (TubercuList: Rv0500B). In the WHO's tropical disease research (TDR) database, this unknown *Mtb* protein is listed as a potential drug target [250, 251]. In most bacterial ribosome (eg. *Ec* and *Tth*), this pocket formed by the rRNAs h27, h44 and h45 is conserved and is empty. However, in *Ms* 70S ribosome it is occupied by the protein bS22. Interestingly, the location of protein bS22 in *Ms* coincides with the eukaryotic ribosomal protein eL41 and mitochondrial ribosomal protein mS38. Figure 3.29 gives a comparison of the protein location of bS22 in Mycobacteria, *Ec*, and the eukaryotic counterparts [252-255]. In eukaryotic ribosome, protein eL41 plays an important role in the formation of the inter-subunit bridge eB14 [256]. Although the protein eL41 is a non-essential ribosomal protein, its deletion causes changes in the peptidyl transferase activity affecting the translation fidelity [256-259]. The interaction of bS22 with H70 might also give rise to a species-specific inter-subunit bridge in Mycobacteria and may impart additional stability to the 70S complex.

Thus, the presence of bS22 near the decoding centre, and its location in the inter-subunit interface close to the mRNA channel, hints that bS22 might be involved in the translation initiation, ribosome biogenesis or might be involved in the species-specific translation disparities.

Figure 3.29 A comparison of the bS22 with its eukaryotic counterparts. The location of the protein *Ms* bS22 and *Mtb* bS22 compared with the eukaryotic eL41 and mitochondrial mS38 proteins. In *Ec*, the space is empty as no protein/rRNA is located at this position. The *Ms* rRNA is shown in yellow and the helices h27, h44, h45 are shown in deep yellow. The *Ms* bS22 is shown in red; *Mtb* bS22 in purple (PDB ID: 5V93) [146]; *H. sapiens* eL41 in deep blue (PDB ID: 4UG0) [254]; *H. sapiens* mS38 in light blue (PDB ID: 3J9M) [253]; *S. cerevisiae* eL41 in deep green (PDB ID: 5MRC) [255]; *S. cerevisiae* mS38 in light green (PDB ID: 4V88) [260].

The second density observed was near the PTC, and was assigned to the helical protein bL37 [146-148]. Protein bL37 is 24 residues long, and resides in the pocket formed by the rRNA helices H39-H40 (of domain II of 23S rRNA) and H72-H89 (of domain V in 23S rRNA) (Figure 3.30). There are no homologous protein residing at this location from bacterial or mammalian ribosome. Although, the position of bL37 in the map surrounded by the rRNA helices coincides with the location of the chloroplastic r-protein PSRP6 (Figure 3.30) [261]. The bL37 basic protein (pI: 12.19) is stabilized by the negatively charged rRNA surroundings. The highly conserved PTC and the location of tRNA binding loops (A and P loops) are in close vicinity of this protein bL37. It

also resides closely to interact with the D-loop of 5S rRNA on one side. The 5S rRNA is uniquely placed in the ribosome and is proposed to link all the functional centres of the ribosome [262]. Therefore, the interaction of bL37 with H38, H89 and the loop D of 5S rRNA might contribute in the stabilization of the ribosome functional centres at this junction. The location of bL37 also raises further interests because of its interaction with H89, which is an essential component of the PTC. Although the exact function of bL37 still remains to be elucidated, its critical location near these functional domains hints towards its involvement in the stabilization of the ribosome architecture.

Figure 3.30 Location of *Ms* bL37 in the rRNA pocket and comparison with other ribosomal proteins. (A) *Ms* bL37 is located near the PTC in the LSU. It is sandwiched between the helices H39-H40 and H72-H89, shown in dark blue. The 5S rRNA and the rest of 23S rRNA are shown in green and light blue, respectively. The rest of the ribosome model is hidden for clarity. (B) A comparison of the rRNA model (near the region of bL37) between *Ms* and *Ec*. (C) The location of the chloroplastic PSRP6 (yellow) (PDB ID: 5MMM) [261] sharing the same location as the *Ms* bL37.

3.5 Structure of *M. smegmatis* 70S ribosome in the Hibernating state

Hibernation state 70S map was obtained from the stationary phase (48 h) purified ribosomes, and the ribosomes were imaged on the FEI Arctica Cryo-Electron Microscope. The 70S EM density map was refined to an average resolution of 4.1 Å (Figure 3.10). When the map was compared with the P/P state 70S map, additional densities were observed at the inter-subunit space for HPF and near mRNA exit site for the protein bS1 (Figure 3.31). As this map was bound to HPF, it was called as the hibernating state of the ribosome [112]. The overall structure of the hibernating state 70S is very similar to the P/P state 70S except for the two distinct additional densities (for HPF and bS1) and the observed movement in the H54a towards the mRNA exit site. The P/P state 70S model was fitted into the hibernating state 70S and refined further (with the help of *phenix.refine*) to obtain the 70S model of the hibernating state. The HPF and bS1 were modeled into the additional densities at the inter-subunit space and near the mRNA exit site, respectively.

Figure 3.31 Structure of the hibernating state of *Ms* 70S ribosome. Cryo-EM map of the hibernating state of 70S ribosome at 4.1 Å resolution (A) along with the model (B). LSU r-proteins are colored in light-blue, LSU rRNAs in steel-blue, SSU r-proteins in light-yellow, SSU rRNAs in golden-yellow, H54a in deep red, bS1 in deep-blue, E-tRNA in pink, and HPF in green. Figure is modified from [237].

3.5.1 Hibernation Promoting Factor

Ms HPF is a 230 residues long protein belonging to the HPF^{long} protein family. It shares 43.5% and 40.2% sequence similarity with the *Sa* HPF^{long} and

Tth HPF^{long}, respectively (Figure 3.32). The homology model for the conserved NTD of *Ms* HPF (residues 31 – 130) was built using *Tth* HPF as a template (PDB ID: 4V8G) [111]. The homology model was rigid body docked onto the observed density of HPF at the inter-subunit space.

Figure 3.32 Sequence alignment of HPF from *Ms*, *Sa*, and *Tth*. The sequence alignment of HPF from *M. smegmatis*, *S. aureus*, and *T. thermophilus*. The sequence similarity between *Ms* and *Sa* is 43.5%, and between *Ms* and *Tth* is 40.2%. Figure is adapted from [237].

The *Ms* hibernating state 70S bound with HPF exists in an un-rotated state of ribosome similar to *Sa* HPF-70S (PDB ID: 5ND8) and *Tth* HPF-70S (PDB ID: 4V8G) [111, 119]. The NTD of *Ms* HPF adopts the conserved β_1 - α_1 - β_2 - β_3 - β_4 - α_2 arrangement and is located between the SSU head and the body, in the pocket formed by the 16S rRNA helices on the mRNA channel at the inter-subunit space (Figure 3.33 A). Since the density of HPF is moderately resolved here (Figure 3.33 B), sequence analysis and comparison with the known HPF models was done to understand the mode of interaction between HPF and SSU.

Figure 3.33 Location of *Ms* HPF (A) A model of the NTD of *Ms* HPF with density in grey translucent surface. HPF binds to the mRNA channel at the inter-subunit space, surrounded by 16S r-RNA and E-tRNA, adopting the conserved β_1 - α_1 - β_2 - β_3 - β_4 - α_2 arrangement. **(B)** Local resolution calculation for HPF density using resmap [175] in the hibernating state 70S map showing the local resolution in Å as indicated in the scale bar. Figure is modified from [237].

The stacking interaction between β -sheets of NTD *Ms* HPF and 16S rRNA helices (Lys36/Arg38 and U947, and Glu100 and G948) are conserved in the head of SSU (Figure 3.34 B). HPF binds at the mRNA channel and block the binding of the A- and P-tRNA at the SSU in the inter-subunit space. Therefore, it transitionally inactivates the ribosomes by physically occupying the A-tRNA, P-tRNA and mRNA binding sites on the SSU (Figure 3.34 A) and thereby blocking the binding of A-tRNA, P-tRNA, and mRNA to the 70S. The binding of HPF to the ribosome is irreversible and is generally stress induced and the HPF is known to dissociate from the ribosome once the cells encounter favourable conditions. The conserved residues (Lys 53, Arg 56, and Arg 59) of helix α_2 overlap with the position of phosphate backbones of A- and P- tRNA at the body of SSU (Figure 3.34 B). HPF also blocks certain antibiotics like hygromycin B, tetracycline, and edeine targeting A- and P- tRNA binding sites on the 70S in the similar manner [263]. Thus, an inactive ribosome under stressed condition also gets protected by the external agents like antibiotics. This could be another reason that the HPF bound inactive ribosomes are also observed in the log phase cells of some pathogenic microorganisms. Since most of the antibiotics that target ribosomes, generally target the translationally active ribosome, they might have reduced impact on these HPF-bound translationally inactive ribosomes, thus imparting another drug-resistant mechanism by the pathogenic microorganism against antibiotics.

Figure 3.34 Interaction of HPF with the tRNAs and the 16S rRNA residues (A) HPF occupies the A- and P-tRNA binding sites in SSU. Models of mRNA (black), A-tRNA (salmon), and P-tRNA (blue) from *Ec* (PDB: 5AFI) [235] are super-positioned with HPF to show steric clashes. **(B)** The binding mode of the NTD of *Ms* HPF^{long} at the mRNA channel as predicted by the superimposition of the *Sa* HPF^{long} model to the *Ms* hibernating state map [119]. Top box: Conserved stacking interactions between Lys 36/Arg 38 with U947 and Glu 100 with G948. Bottom Box: Lys 53, Arg 56 and Arg 59 interact with U1479, C1480 and G1481 and obstruct the binding of A- and P-tRNAs. Figure is modified from [237].

The density for the first 30 amino acids of HPF is not visible in our map and could be because of the flexibility and low resolution at this region. In addition to this, the entire CTD of the HPF is also not visible in the map. From the recent ribosome dimer structures, also known as 100S dimers, of *S. aureus* and *L. lactis*, it is shown that the CTD of HPF^{long} is involved in the ribosome dimerization [108, 109]. Surprisingly, no 100S dimers have been reported for *Mycobacteria* even during hyoxic state, which indicate that the *Ms* ribosomes do not undergo 100S dimerization [116].

To further confirm this, the ribosomes were purified from the log (24 h), stationary (48 h) and extended stationary phase (66 h). They were purified under two different salt conditions (potassium acetate), low salt (50 mM) and high salt (500 mM) and analysed by sucrose density gradient centrifugation. The reason to include the low salt condition of purification is to avoid any dissociation of ribosomes which might result from high salt concentration in the buffer. Under all these six conditions, the ribosomes continued to exist as a stable 70S complex without any detectable formation of 100S dimer, even during hibernation as shown in the figure 3.35.

Figure 3.35 Ribosome purification under high and low salt conditions for *Ms* cultures harvested at different growth periods. Ribosomes purified from *Ms* culture harvested at

different growth periods of 24, 48, and 66 h, as indicated. For each growth period, ribosomes were purified under low (50 mM Potassium acetate) and high salt (500 mM Potassium acetate) conditions. Crude ribosomes were then analysed by linear sucrose gradient (10 - 40%) to observe ribosomal profile under these conditions. Neither any of the ribosomal subunits (30S and 50S) nor the 100S ribosomal dimers could be detected under these conditions even after extended stationary phase (66 h growth) while 70S seems to be stably present during stationary phases. Figure is adapted from [237].

The absence of 100S dimers from *M. smegmatis* raises the possibility that the CTD of HPF in Mycobacteria might be involved in some other function which yet remains to be elucidated. Interestingly, a distinct density of bS1 protein interacting with uS2 and H54a was observed in the HPF bound 70S map. This additional density was too large for the CTD of HPF but was large enough to accommodate the four functional domains of the bS1 protein. The assigning of this additional density to bS1 protein is consistent with the earlier reports on localization of bS1 at this position across various ribosome structures known so far (Figure 3.36) [264-267].

Figure 3.36 Comparison of S1 structure in association with the ribosome as captured in different studies. (A) *Ms* bS1 N-terminal α -helix (blue) in P/P map (filtered to 10 Å), along with H54a (deep red) and uS2 (yellow). (B) *Ms* bS1 (deep-blue) with its four OB domains in hibernating map along with H54a (deep red) and uS2 protein (yellow). (C) Cryo-EM density for bS1 in RNA polymerase-bound *Ec* SSU (PDB ID: 6AWB) [265]. Model for S1 is shown in dark green and uS2 in yellow. (D) Cryo-EM density of bS1 in *Ec* SSU (PDB ID: 4TOI) [266]. (E) Cryo-EM bS1 density on the *Ec* ribosome bound to translocon SecY (PDB ID: 3J46) [267]. (F) Cryo-EM density of bS1c in the chloroplast SSU (PDB ID: 5X8R) [264]. Figure is adapted from [237].

Ms bS1 is the largest and the most acidic protein in the *M. smegmatis* ribosome comprising of 449 residues. The ribosomal bS1 protein belongs to the oligonucleotide/oligosaccharide binding (OB)-fold superfamily proteins. The sequence analysis of *Ms* bS1 predicted the presence of four OB domains in contrast to six OB domains in *Ec* bS1 [268]. The homology model for *Ms* bS1 was built with Phyre2 server [173] as explained in the section 2.2.8.2. We were able to fit all the four predicted OB domains in the additional density observed near the mRNA exit site in the map (Figure 3.37 A). The first OB domain and the N-terminal α -helix interacting with uS2 could be fitted into the density

unambiguously while the three remaining OB domains were manually fitted into the remaining fragmented densities. The first OB domain of bS1 is the first site of interaction of bS1 with the SSU during translation initiation. The N-terminal α -helix of bS1 interacts with uS2 by salt bridges and is conserved among bacteria [266, 269]. The first OB domain of *Ms* bS1 also shares surface contact with the 16S rRNA via h40 in this map (Figure 3.37 B). These conserved and strong interactions contribute towards the well resolved density for the first OB domain of bS1 NTD in the hibernating state 70S map. This also agrees with the assumption that NTD of bS1 is the anchoring domain to the ribosome [270].

Figure 3.37 Depiction of the four OB domains of *Ms* bS1 interacting with uS2. (A) Density for bS1 is in transparent grey, the four OB domains of bS1 are labelled and highlighted in different shades of blue as indicated. Densities for the OB domains and uS2 protein are shown at the different thresholds for clarity. (B) At a lower threshold, the *Ms* bS1 is seen to share surface contact (highlighted by dotted arrow) with 16S rRNA (h40), shown in yellow.

To further confirm the assigning of the additional density to *Ms* bS1, we compared the fitting of *Ms* bS1 N-terminal and *Ms* HPF-CTD into the density. This density is much bigger to accommodate the CTD of HPF whereas the NTD of bS1 fits well in this density (Figure 3.38).

Figure 3.38 Analysis of the extra density near the mRNA exit site in the hibernating state map. The models of *Ms* HPF CTD (**A**) and *Ms* bS1 OB-I (**B**) were fitted into the extra density observed near the mRNA exit site of the hibernating state 70S map. The *Ms* bS1 NTD fitted well into this density compared to *Ms* HPF CTD. The *Ms* HPF CTD is shown in the same orientation as found in the *S. aureus* and *L. lactis* 100 dimers [108, 109]. Figure is adapted from [237].

Interestingly, a partial density of bS1 protein was also noticeable in the P/P state 70S ribosome when filtered to 10 Å (Figure 3.39). The N-terminal α -helix of bS1 could be fitted into this filtered density and first OB domain only appears at very low threshold while rest of the bS1 is not visible. This suggests that protein bS1 could be present during both the P/P and the hibernating states in *Ms* as opposed to the general proposal of an on- and -off mechanism of bS1 during translation initiation and elongation [271, 272].

Figure 3.39 bS1 density in the P/P state 70S ribosome. bS1 density in P/P state 70S map when filtered to 10 Å and shown at 0.01 threshold in UCSF Chimera [176]. The density of N-terminal α -helix appears at this resolution. A fragmented density is observed for the first OB domain. Figure is adapted from [237].

The presence of bS1 protein in the purified ribosomes from the log phase and stationary phase was confirmed by SDS PAGE analysis and also identified by Mass spectrometry analysis (Figure 3.40). In the recent report of bS1 from *Mtb*, it was reported to interact with tmRNA to rescue stalled mycobacterial ribosomes [273]. Although, this interaction was not observed in our trans-translation 70S map possibly due to the its low-resolution.

Figure 3.40 SDS PAGE and Mass spectrometry analysis of bS1 protein from stationary phase 70S ribosomes. (A) 15 pmol of 70S ribosomes from *Ec* (as control), *Ms* 70S log-phase ribosome and *Ms* 70S stationary-phase ribosome were analysed on 12% SDS-PAGE with coomassie blue staining. The bS1 protein is highlighted. *Ms* 70S (stationary phase) bS1 protein band (boxed) is further analysed by Mass Spectrometry (B-C) The sequences colored in bold red are the matched peptides from Peptide mass fingerprinting. The sequence coverage obtained was 75%. With the highest hit for bS1 protein from *Mycobacterium smegmatis*. Figure is adapted from [237].

It is plausible to assume that the interaction between *Ms* bS1 and H54a in the hibernating state 70S makes it visible in our map. As shown in the figure 3.27, the H54a moves about 16.4 Å towards the mRNA exit site. It interacts with the protein bS1 at this location. This stable interaction of H54a-bS1 protein at the mRNA exit site and tRNA ejection site, suggest a possible function for the *Mycobacteria* specific H54a rRNA insertion both in translation elongation (involved in tRNA ejection) and initiation or resuscitation from the hibernation state of ribosome (stabilizing bS1).

The protein bS1 has not been observed in any hibernating state ribosomes or in ribosome dimers of *E. coli*, *T. thermophilus*, *S. aureus* and *L. lactis*. Contrary to this, there is a clear well-ordered density of bS1 protein visible in the *Ms* hibernating state 70S map. To understand the function of the presence of bS1 in the hibernating state ribosome, we super-positioned the ribosome dimers of *S. aureus* (PDB ID: 5NG8) and *L. lactis* (PDB ID: 5MYJ) (as both *Sa* and *Ll* have HPF^{long} similar to *Ms* HPF^{long}). The super-positioning revealed that the first OB domain of *Ms* bS1 in the HPF-70S map sterically overlaps with the CTD of *Sa* and *Ll* HPF^{long} in their ribosome dimers [108, 109]. The CTDs of both HPFs in *Sa* and *Ll* (interacting with the uS2) are responsible for the dimerization, and are localized at the same position as the first OB domain of bS1 in our HPF-70S map (Figure 3.41).

between *Ms* (*M. smegmatis*), *Mtb* (*M. tuberculosis*), *Sa* (*S. aureus*) and *Ll* (*L. lactis*). One out of five conserved residues involved in HPF dimer interaction is mutated in *Ms* and *Mtb* (shown in red dotted box) [109]. Figure is adapted from [237].

The fact that bS1 is located at the same position as that of the CTD domain of HPF, it might explain the absence of any 100S dimer formation in Mycobacteria as seen in *S. aureus* and *L. lactis*, even though *Ms* possess HPF^{long}, which is known to be responsible for 100S formation in *Ll* and *Sa*. This strongly hints towards a direct role of *Ms* bS1 in ribosome resuscitation from the inactive/hibernating state. Such a phenomenon is definitely advantageous for the recently evolved slow-growing Mycobacteria.

The formation of 100S dimer in bacteria is a way to keep ribosomes stable and prevent subunit dissociation and rRNA degradation. 100S dimers have not been reported in Mycobacteria [116]. Also, in the biochemical analysis of ribosomes purified from different growth phases, no 100S dimers were observed (Figure 3.35). Thus, not forming 100S dimers could be disadvantageous for Mycobacteria, unless there is some other mechanism employed for the stabilization of 70S ribosomes during hibernation state in Mycobacteria. Therefore, it seems plausible that the presence of *Ms* bS1 could provide stability to the HPF bound hibernating 70S ribosome in Mycobacteria without forming the 100S dimers. bS1 was also seen to interact with the H54a near the mRNA exit site (Figure 3.27). Because of the interaction between H54a and bS1 in the hibernating state 70S ribosome, it is plausible to think a possible function of H54a in translation initiation or resuscitation from the inactive hibernating state of ribosome, via stabilizing bS1 protein at this location. In addition, this interaction may play a role in the translation of the leaderless mRNA transcripts, which are in abundance in Mycobacteria (nearly one fourth of mRNAs in Mycobacteria are leaderless). Thus, this unique rRNA extension might be involved in translation in Mycobacteria.

The results in the next section (3.5.2) has been interpreted from the *Ms* P/P state 70S ribosome (at 3.4 Å) and the *Ms* hibernating state 70S ribosome (at 4.1 Å).

3.5.2 Structure of mRNA entry and exit sites of *M. smegmatis* 70S

ribosome

The function of ribosome is translation of mRNAs into nascent polypeptide chains. The mRNA encapsulates the neck of the SSU and is exposed to the LSU to interact with the tRNAs for translation. mRNA binding and recognition during translation initiation is generally facilitated by the mRNA entry and exit sites on the SSU [266, 269, 274]. mRNA is the driving force of the translation, and therefore the correct positioning of mRNA in the ribosome is absolutely crucial. It is believed that in bacteria the translation is initiated by the two most important factors, i) bS1 protein, and ii) the Shine Dalgarno (SD) sequence of the mRNA. The bS1 protein is involved in the translation initiation which acts as a RNA chaperone and unfolds the secondary structure elements of the mRNA which might obscure the start signal embed in the start codon [270, 275-277]. It is believed to function in an on-off fashion with the ribosome, and is considered to be one of the most flexible protein of the ribosome complex, and is thus arduous to get localized with ribosome [271, 272]. It has been the most tedious protein to be localized in the structures of ribosome. Although, in the *Ms* hibernating state, bS1 protein was localized near the mRNA exit site, showing how the highly conserved ribosomal protein bS1 could show such diversity across bacteria. The second important factor is the SD sequence, which is the sequence upstream of the start codon in the mRNA. The correct positioning of the mRNA in the ribosome is stabilized by the interaction of the Shine Dalgarno (SD) sequence of mRNA to the anti-SD sequence in the 3' end of the 16S rRNA, otherwise known as the 3' tail of 16S rRNA [274, 278]. Interestingly, in *Mycobacteria*, one fourth of the total mRNAs, so-called the leaderless mRNAs are devoid of SD sequences [279]. Therefore, *Mycobacteria*-specific remodeling of the mRNA recognition and binding sites might be required for the efficient translation initiation of these classes of leaderless mRNA transcripts.

mRNA entry site. The mRNA entry site is one of the most conserved sites of bacterial ribosome and is located at the neck of the SSU. It is composed of three proteins, uS3, uS4, and uS5. *Ms* mRNA entry site is architecturally similar to *Ec*, except for the protein uS5. *Ms* uS5 is extended at both the ends (in comparison to its *Ec* counterpart). The CTE forms an additional helix over the surface of the

SSU, and might contribute towards additional structural stability on the surface of SSU during translation initiation. The NTE extends towards the mRNA entry site. This extension might interact with the leaderless mRNA transcripts and provide the correct positioning on the first codon, as the protein uS5 is responsible in translation fidelity [238]. Extended NTE in uS5 could also help in directing the mRNA to the ribosome. An overview of the mRNA entry site is shown in the figure 3.42 B.

mRNA exit site. The mRNA exit site is the location of the SD-antiSD interaction, and thus is very crucial for the translation initiation. mRNA exit site in the bacteria is composed of four major proteins, uS7, uS11, bS18, and bS21 (Figure 3.17 and Figure 3.42 C). The *Ms* mRNA exit site superimposes well with *Ec*, except for the missing bS21 (Figure 3.17) and the presence of 10 residue NTE in the protein *Ms* bS18 (compared to *Ec* bS18). The absence of the bS21 is not substituted by either rRNA or r-protein extension at this location. And also, we observed distinct density for the protein bS1 in hibernating state and a partial density in the P/P state 70S maps. H54a is highly flexible and is located very near to the mRNA exit site. In the hibernating state 70S, the tip of the helix also moves further towards the exit site to interact with the bS1 protein. Thus, it can be assumed that the H54a-bS1 interaction might have species-specific functions in translation initiation. This means that in Mycobacteria both the 16S and 23S rRNAs are involved in the translation initiation. And as this interaction seems to be more stable in the hibernating state ribosome, it can also suggest that the H54a-bS1 interaction might function in the ribosome resuscitation in Mycobacteria.

Figure 3.42 Architecture of mRNA entry and exit sites of the *Ms* 70S ribosome. (A) A surface view of the SSU in hibernating state of 70S shown from the solvent-side. Structural landmarks of SSU and the proteins around the mRNA entry and exit sites are labelled as indicated. (B) mRNA entry site in *Ms* is formed by three r-proteins, uS3, uS4 and uS5. The uS5 NTE extends towards the mRNA entry site (marked with star) and CTE forms an α -helix on SSU surface. Density of uS5 NTE is shown in grey mesh. (C) mRNA exit site (marked by a dotted circle) in *Ms* is formed by two r-proteins uS7 and uS11, and 16S rRNA. bS1 protein is present near to the mRNA exit site. The N-terminal α -helix of *Ms* bS1 interacts with uS2 and one of the OB domain (domain III) with H54a. bS1 is colored in deep blue, uS2 in green, uS3 in pink, uS4 in orange, and uS5 in grey colors in A-C. Figure is modified from [237].

3.6 Trans-Translating state of *M. smegmatis* 70S ribosome

Trans-translation state of ribosome is a state of rescue mechanism present in bacteria to rescue ribosomes from their stalled state. Ribosome stalling in bacteria can arise because of the translation of a truncated mRNA or mRNA lacking a stop codon. Trans-translation is driven by a RNA molecule known as transfer-messenger RNA or tmRNA and a small protein B (SmpB). These two molecules together carry out the function of trans-translation in bacteria to rescue the ribosome from its stalled state.

In the 70S-log dataset, the trans-translation state 70S map was obtained as one of the 3D classes during the third round of 3D classification (Figure 3.4). The identification of this class was possible because of a big dataset of 391,000

particles used, which allowed in the classification of this functional state. And was distinguishable because of the presence of this structured tmRNA wrapped around the 70S ribosome. The presence of the trans-translation state in the log phase of *Ms* could also arise from the abundance of leaderless mRNA (about one fourth of total mRNA) in *Mycobacteria* [279]. The presence of trans-translation state is essential in few bacteria [128, 129, 132], thus their presence in log phase could be essential for the sustainability of *Mycobacteria*. Because of the limited number of particles (~1300), this class could only be refined to an average resolution of 12.5 Å (Figure 3.10). The docking and comparison with the known trans-translation states of 70S models bound with tmRNA [139, 140, 280, 281] confirmed that the *Ms* trans-translation state represents the resumed state with tmRNA-smpB complex at the P-site and a tRNA occupying the A-site (Figure 3.43).

Figure 3.43 Structure of the trans-translating state of *Ms* 70S ribosome. Cryo-EM map of the tmRNA-SmpB complex-bound 70S ribosome at 12.5 Å resolution. Density is shown as surface. tmRNA is shown in purple, SmpB in green and A-tRNA in red (left). LSU in transparent blue showing the tmRNA-TLD and SmpB occupying the P-site and tRNA occupying A-site (right). Figure is modified from [237].

The *Ms* tmRNA is 369 nt long RNA molecule and shares 50.7% sequence identity with *Tth* (349 nt) and 51.5% with *Ec* (363 nt). The secondary structure of tmRNA is presented in the figure 3.44.

Figure 3.44 Secondary structure diagram of *Ms* tmRNA. Secondary structure of tmRNA from *Ms*, highlighting the structural domains TLD, MLD, pseudoknots and helices. MLD coding for protein degradation tag (ADSNQRDYALAA) is underlined with resume codon shown in red and stop codon in black. The different domains are colored and labelled. The dots denote counts after every 10 nucleotides. Watson-crick base pairs are shown in lines (-), G $\circ$ U base pairs with rings ($\circ$), and non-standard base pairs with dots ($\bullet$). Figure is adapted from [237].

The homology models for *Ms* tmRNA and SmpB were built using *Tth* tmRNA and SmpB (PDB ID: 3IYR) as templates [139]. tmRNA is a highly structured RNA molecule. It consists of four pseudoknots (PKs 1-4), two helices (H2 and H5), one tRNA like domain (TLD), and one mRNA like domain (MLD) (Figure 3.45 A). The structural integrity of tmRNA is maintained by the pseudoknots and helices which constitute the structural elements of tmRNA, while the tmRNA functions through its functional elements TLD and MLD [124, 139, 281]. In the trans-translation state, the tmRNA encircles the SSU of the

ribosome. The four pseudoknots are arranged in an arc-like structure extending from the beak of the SSU to the mRNA entry site. The two helices (H2 and H5) assist in the accurate positioning of the tmRNA on to the SSU of the ribosome. The 3' and 5' ends of the tmRNA form the TLD [282] and together with the small protein B (SmpB) mimic the tRNA structure (Figure 3.45 B). SmpB, which is the conserved protein in prokaryotes, is an essential component in the trans-translation [280]. SmpB stabilizes the binding of the tmRNA at the A- and P-tRNA binding sites. The TLD of the tmRNA and the SmpB protein mimic the tRNA. While the TLD of the tmRNA accounts for the acceptor stem (T-arm) of the tRNA, the SmpB protein compensates for the D stem and the anticodon loop of tRNA in tmRNA (Figure 3.45 B). SmpB provides the codon-anticodon pairing in the trans-translation state of ribosome, and is an indispensable element in the functioning of tmRNA. *Ms* SmpB is a highly basic (pI: 10.27) 18.2 kDa protein. It shares 45.7% and 36.4% sequence identity with *Tth* SmpB and *Ec* SmpB, respectively. Being highly basic, it gets stabilized by the surrounding RNA molecules.

Figure 3.45 Architecture of tmRNA and SmpB when bound to the 70S. (A) TLD is shown in purple, pseudoknots (PK) 1-4 are in deep green, pink, brown, and yellow, respectively. Helix 2 (H2) is in magenta, helix 5 (H5) in light blue, MLD in grey (stop codon in black), and SmpB in green. These colors are consistent with the secondary structure of the *Ms* tmRNA in figure 3.44. (B) Comparison of the structure of a typical tRNA with tmRNA-TLD and SmpB complex showing how tmRNA (TLD) and SmpB complex mimics a tRNA. Figure is modified from [237].

The MLD of the tmRNA is the most crucial part in the trans-translation of ribosome (shown in grey color in the figures 3.44 and 3.45 A). It mimics the mRNA during trans-translation process. The MLD positions itself in the mRNA

channel and the MLD sequence codons gets translated into the stalled polypeptide chain.

The MLD encodes for a degradation tag which can vary in its length (8 - 35 residues) depending on species. The degradation tag has a conserved 3' end in most bacteria except for mycoplasma [283]. In *Ms*, this degradation tag, ADSNQRDYALAA, is 12 residues long (*Ec* - ANDENYALAA (10 residues), *Tth* - ANTNYALAA (9 residues)) and carries a small insertion in the MLD in comparison to *Ec* and *Tth* (Figure 3.46).

Figure 3.46 Comparison of the structure of *Ms* tmRNA with *Tth* tmRNA. The *Ms* tmRNA (purple) is superimposed over the *Tth* tmRNA (grey) (PDB: 3IYR) and sequence comparison of the degradation tags encoded by MLD. An insertion of 9 nucleotides is boxed in MLD of *Ms* [139]. Figure is modified from [237].

3.6.1 Unique mechanism of tmRNA during trans-translation.

tmRNA-SmpB complex binding to the ribosome is quite intriguing. In a stalled ribosome, the A-site tRNA is unoccupied. The tmRNA-SmpB complex first binds at the A-site in the SSU with the help of elongation factor Tu (EF-Tu) in a similar fashion as the tRNA binds to A-site (assisted by EF-Tu) [124]. At the A-site, tmRNA acts like a tRNA and accepts the truncated polypeptide from the P-site and the TLD is translocated to the P-site. Interestingly, the translocation of the tmRNA-SmpB complex is analogous to tRNA translocation from the A- to P-site, even though the tmRNA is almost six times the size of a tRNA [139, 140]. At this point, the deacylated P-tRNA moves to the E-site. The MLD is placed in

the ribosome and the translation of the degradation tag encoded by the MLD thus begins. Another unique property of the MLD is, it encodes a stop codon at the end of the degradation tag, which helps in release of the new hybrid protein (Figure 3.45 A). These truncated tagged proteins are then subjected to proteolysis in the cell. The tmRNA and SmpB are then released from the ribosome and the translation resumes [139, 141].

The current model of *Ms* 70S with tmRNA suggest that the truncated mRNAs are subjected to the trans-translation state for rescue making this state of ribosome essential for bacteria to recover from ribosome stalling and degradation of incomplete or truncated polypeptide chains which might be deleterious to bacteria [132]. Cryo-EM and single particle analysis assisted in classifying this unique class of ribosome during data processing and enabled us to study the trans-translation state of 70S ribosome in *M. smegmatis*.

Chapter 4

DISCUSSION

Discussion

In the present work, I have analyzed three different functional states of *M. smegmatis* ribosomes. A high-resolution structure of the P/P state of *Ms* helped in identifying various species-specific insertions and deletions in the ribosomal proteins and rRNAs of Mycobacteria. The three different functional states were compared and analyzed with the recently published structures of ribosomes from *Ms* and *Mtb* to identify and study the various unique features of the mycobacterial ribosome [146-148].

Mycobacterial ribosome shares a conserved ribosome core structure for both of the large subunit and small subunit. In the SSU, several r-proteins have terminal extensions (in comparison to *E. coli*). Out of those extensions which were visible in our maps, they were located on the solvent exposed surface. The CTE of uS5 forms an additional helix on the neck region of the SSU surface (Figure 3.14). The CTE of the uS16 protein which is located near the foot region of the SSU, extends towards the neck and folds into a small helical structure to interact with the protein uS4 (Figure 3.15). The NTE of uS17 extends towards the foot region of the *Ms* SSU to interact and stabilize the h9 insertion in 16S rRNA (Figure 3.16). And lastly, a small NTE of bS18 (about 10 residues) also exists on the beak region of the SSU surface. It is hard to not notice that all these extensions are arranged in a manner to cage the SSU surface and provide additional stability (Figure 3.11 C). In the ribosome, the SSU is highly flexible compared to the LSU. Such Mycobacteria-specific protein extensions in the SSU might provide additional stability to the mycobacterial SSU in the cell. And whether this characteristic feature is responsible for the long dormant stability of mycobacterial species, is still an unanswered question.

Two new proteins were identified in the mycobacteria close to the functional domains of the ribosome. These proteins have not been seen before in any of the bacterial ribosome structures known so far. The first protein is a SSU r-protein bS22, which lies close to the decoding center of the SSU, and might assist in the binding of the tRNAs at the mRNA channel (Figure 4.1). bS22 shares similar location as the eukaryotic LSU r-protein eL41 and eukaryotic mitochondrial SSU r-protein mS38. The second protein is a LSU r-protein, bL37 located close to the Peptidyl Transferase Center (Figure 4.1). The location coincides with the

chloroplastic protein PSRP6, though the exact role of this protein in the mycobacterial ribosome translation is still unknown.

Figure 4.1 Two new r-proteins of mycobacterial ribosome. *bS22* and *bL37* are the two new proteins identified in the mycobacterial ribosome. *bS22* exists near the decoding centre in the *Ms* SSU, whereas *bL37* is present near the peptidyl transferase center.

Mycobacterial ribosome has many rRNA extensions, which are extended significantly in the LSU and have changed the surface topology of the mycobacterial LSU. H15-H16a rRNA extension of the domain I of 23S rRNA takes part in the kissing loop interaction. Kissing loop interactions of RNA are known to impart great structural stability to the RNA complexes. The presence of such an interaction in mycobacterial ribosome again hint towards the structural differences contributing to the stability of the mycobacterial ribosome complex.

The most prominent feature and longest rRNA insertion of mycobacterial ribosome is the ~100 nucleotide long extension in the domain III of the 23S rRNA, the Helix 54a (H54a). Emerging near the tRNA exit site and the L1 stalk, it is the most flexible insertion of the r-RNA in mycobacterial ribosome. Studying its flexibility from the current states, the P/P 70S and hibernating state 70S, and the 50S from *Mtb* [146], we know that it is capable of moving at least 80 Å laterally. In the LSU, it can translocate towards the inter-subunit space to prevent the binding of the 30S (inhibiting the formation of 70S), and this way protects the 50S degradation in the cell [146]. In the current study's functional states, we observed that the flexibility and the unique location of H54a allows it to interact with the bS1 protein. This interaction can lead to the formation of a stable initiation complex during translation initiation. As this state of 70S was also found with the Hibernation Promoting Factor, the interaction of H54a-bS1

complex might help in ribosome resuscitation from this functionally inactive 70S ribosome of the stationary phase.

Apart from the P/P state, which is one of the most abundant ribosomal state found in bacteria, we studied the hibernating 70S state ribosome of *Mycobacteria* bound with HPF. In *Ms*, HPF exists in its long form, with two functional domains; the NTD domain binds to the mRNA channel in the inter-subunit space and inactivates the 70S complex by physically blocking the binding of A- and P-site tRNA. Though, in this state, it does not interfere with the binding of E-tRNA (Figure 3.33 and Figure 3.34). HPF^{long} is known to form ribosome dimerization in bacteria, as seen in *S. aureus* and *L. lactis* [108, 109]. It seems that *Mycobacteria* do not favor the formation of 100S ribosome dimers. Even under different growth periods the mycobacterial ribosomes do not dimerize (Figure 3.35). The CTD of HPF^{long} responsible for the formation of ribosome dimers was not visible in the hibernating state 70S ribosome, and instead the density of the NTD of bS1 was present at the same location (Figure 3.41). Thus, the CTD of *Ms* HPF might be functionally irrelevant as *Mycobacteria* ribosome do not form 100S dimers. And the CTD of *Ms* HPF might be involved in some other unknown function.

Using a large dataset and with the help single particle analysis data processing, we were able to separate a unique functional state of ribosome, the trans-translation 70S state, recognized by its crown shaped tmRNA bound to the 70S ribosome. Separation of this class which consist of about 1000 particles out of the ~390,000 particles demonstrates the robustness in cryo-EM data processing. The trans-translation is a mechanism in bacteria which rescues the stalled ribosomes during translation. The trans-translation is an essential mechanism for the survival of the pathogenic organisms like *Neisseria gonorrhoeae*, *Mycoplasma pneumoniae*, *Mycoplasma genitalium* and *Haemophilus influenza* [132], although it is still unclear if trans-translation is essential for the survival of the pathogenic *Mycobacteria*. The overall structure of tmRNA and protein SmpB is similar to *Ec* and *Tth*, except for a small insertion in the MLD region. The trans-translation map of *Ms* is in its resumed state with tmRNA-SmpB complex bound at the P-site and tRNA at the A-site. Because of the low resolution of this map (12.5 Å), further structural interpretations were not

possible. Though it is worth mentioning that this is the first-time trans-translation state of mycobacterial ribosome has been visualized.

Study of these functional states show that most of the unique species-specific features of mycobacterial ribosome contribute towards the stable ribosome structure, even when the subunits stay in their dissociated states. Such extensive species-specific features thus might help the bacteria to survive in its dormant state for a prolonged period and could be a mechanism of Mycobacteria's continued adoption of drug resistance.

Chapter 5

CONCLUSION

Conclusion

In this dissertation, an attempt was made to understand the structure of mycobacterial ribosome purified from *M. smegmatis*. In the process, three different functional states of ribosomes were studied. With the help of a high-resolution structure of ribosome, and the presence of different functional states, various species-specific insights were discovered in the study. The structure of mycobacterial ribosome has unique insertions in its rRNA and r-proteins compared to other bacterial ribosomes, which in a certain manner are responsible for imparting additional stability to the mycobacterial ribosomes. In this study, we have discovered a new interaction between 23S rRNA extension H54a with the translation initiation r-protein bS1. The presence of this interaction in an inactive ribosome state suggests a possible new mechanism of the bS1 and H54a in ribosome resuscitation in mycobacterial ribosome.

Apart from this, we saw how robust the single-particle analysis can be in cryo-EM, which helped us in separating (or classifying) the trans-translation state of mycobacterial ribosome (consisting of only ~1000 particles) from the total dataset. The presence and isolation of this state of ribosome in the Mycobacteria could arise from the abundance of leaderless mRNA in Mycobacteria, which could lead to stalled ribosomes and the need for trans-translation in Mycobacteria.

Studying different functional states from the pathogenic microorganisms helps in developing a better understanding about the mycobacterial translation system and discovery of novel targets for antibiotics. As the non-pathogenic mycobacterial strain *M. smegmatis* shares high genetic similarity with *M. tuberculosis* and is easy to work in laboratory, it was chosen as a model to study the structure of mycobacterial ribosome.

In future, the ribosome study could be directed with different antibiotics targeting the mycobacterial ribosome, to have an understanding about their binding sites, and observe if there are any differences from other pathogenic bacterial ribosome. As the rRNA insertions are prominent in Mycobacteria, they could also serve as potential drug targets against Tuberculosis.

Chapter 6

REFERENCES

1. Servin, J.A., et al., *Evidence excluding the root of the tree of life from the actinobacteria*. Mol Biol Evol, 2008. **25**(1): p. 1-4.
2. Henriques, A.O. and C.P. Moran, Jr., *Structure and assembly of the bacterial endospore coat*. Methods, 2000. **20**(1): p. 95-110.
3. Zink, A.R., et al., *Characterization of Mycobacterium tuberculosis complex DNAs from Egyptian mummies by spoligotyping*. J Clin Microbiol, 2003. **41**(1): p. 359-67.
4. Ducati, R.G., et al., *The resumption of consumption -- a review on tuberculosis*. Mem Inst Oswaldo Cruz, 2006. **101**(7): p. 697-714.
5. Donoghue, H.D., *Human tuberculosis – an ancient disease, as elucidated by ancient microbial biomolecules*. Microbes and Infection, 2009. **11**(14): p. 1156-1162.
6. Donoghue, H.D., *Human tuberculosis--an ancient disease, as elucidated by ancient microbial biomolecules*. Microbes Infect, 2009. **11**(14-15): p. 1156-62.
7. Dye, C. and B.G. Williams, *The population dynamics and control of tuberculosis*. Science, 2010. **328**(5980): p. 856-61.
8. Snowden, F.M., *Emerging and reemerging diseases: a historical perspective*. Immunol Rev, 2008. **225**: p. 9-26.
9. Navin, T.R., S.J. McNabb, and J.T. Crawford, *The continued threat of tuberculosis*. Emerg Infect Dis, 2002. **8**(11): p. 1187.
10. Dheda, K., et al., *The immunology of tuberculosis: from bench to bedside*. Respiriology, 2010. **15**(3): p. 433-50.
11. Rustad, T.R., et al., *Hypoxia: a window into Mycobacterium tuberculosis latency*. Cell Microbiol, 2009. **11**(8): p. 1151-9.
12. Singh, S., I. Saraav, and S. Sharma, *Immunogenic potential of latency associated antigens against Mycobacterium tuberculosis*. Vaccine, 2014. **32**(6): p. 712-6.
13. Barry, C.E., 3rd, et al., *The spectrum of latent tuberculosis: rethinking the biology and intervention strategies*. Nat Rev Microbiol, 2009. **7**(12): p. 845-55.
14. Wilson, D.N., et al., *Species-specific antibiotic-ribosome interactions: implications for drug development*. Biol Chem, 2005. **386**(12): p. 1239-52.

15. Wilson, D.N., *Ribosome-targeting antibiotics and mechanisms of bacterial resistance*. Nat Rev Microbiol, 2014. **12**(1): p. 35-48.
16. Green, R. and H.F. Noller, *Ribosomes and translation*. Annu Rev Biochem, 1997. **66**: p. 679-716.
17. Wilson, D.N. and K.H. Nierhaus, *The ribosome through the looking glass*. Angew Chem Int Ed Engl, 2003. **42**(30): p. 3464-86.
18. Reyrat, J.-M. and D. Kahn, *Mycobacterium smegmatis: an absurd model for tuberculosis?* Trends in microbiology, 2001. **9**(10): p. 472-473.
19. Tyagi, J.S. and D. Sharma, *Mycobacterium smegmatis and tuberculosis*. Trends in microbiology, 2002. **10**(2): p. 68-69.
20. Shiloh, M.U. and P.A.D. Champion, *To catch a killer. What can mycobacterial models teach us about Mycobacterium tuberculosis pathogenesis?* Current opinion in microbiology, 2010. **13**(1): p. 86-92.
21. Berg, J.M., Tymoczko, J.L., Stryer, L., *Biochemistry*, ed. 6. 2007, New York, W. H. Freeman and Company.
22. Luria, S.E., M. Delbruck, and T.F. Anderson, *Electron Microscope Studies of Bacterial Viruses*. J Bacteriol, 1943. **46**(1): p. 57-77.
23. Palade, G.E., *A study of fixation for electron microscopy*. J Exp Med, 1952. **95**(3): p. 285-98.
24. Palade, G.E., *A small particulate component of the cytoplasm*. J Biophys Biochem Cytol, 1955. **1**(1): p. 59-68.
25. Myasnikov, A.G., et al., *Structure-function insights reveal the human ribosome as a cancer target for antibiotics*. Nat Commun, 2016. **7**: p. 12856.
26. Tenson, T. and A. Mankin, *Antibiotics and the ribosome*. Mol Microbiol, 2006. **59**(6): p. 1664-77.
27. Wimberly, B.T., et al., *Structure of the 30S ribosomal subunit*. Nature, 2000. **407**(6802): p. 327-39.
28. Schlueder, F., et al., *Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution*. Cell, 2000. **102**(5): p. 615-23.
29. Ban, N., et al., *The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution*. Science, 2000. **289**(5481): p. 905-20.
30. Yusupov, M.M., et al., *Crystal structure of the ribosome at 5.5 Å resolution*. Science, 2001. **292**(5518): p. 883-96.

31. Harms, J., et al., *High resolution structure of the large ribosomal subunit from a mesophilic eubacterium*. Cell, 2001. **107**(5): p. 679-88.
32. Korostelev, A., et al., *Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements*. Cell, 2006. **126**(6): p. 1065-77.
33. Selmer, M., et al., *Structure of the 70S ribosome complexed with mRNA and tRNA*. Science, 2006. **313**(5795): p. 1935-42.
34. Schuwirth, B.S., et al., *Structures of the Bacterial Ribosome at 3.5 Å Resolution*. Science, 2005. **310**(5749): p. 827.
35. Wilson, D.N., *The A-Z of bacterial translation inhibitors*. Crit Rev Biochem Mol Biol, 2009. **44**(6): p. 393-433.
36. Schmeing, T.M. and V. Ramakrishnan, *What recent ribosome structures have revealed about the mechanism of translation*. Nature, 2009. **461**(7268): p. 1234-42.
37. Shine, J. and L. Dalgarno, *The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites*. Proc Natl Acad Sci U S A, 1974. **71**(4): p. 1342-6.
38. Kozak, M., *An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs*. Nucleic Acids Res, 1987. **15**(20): p. 8125-48.
39. Pierce, B.A., W.H. Freeman, and Company, *Genetics : a conceptual approach*. 2017.
40. Alberts, B., et al., *Molecular Biology of the Cell*. 4 ed. 2005.
41. Shi, H. and P.B. Moore, *The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited*. RNA, 2000. **6**(8): p. 1091-1105.
42. Ogle, J.M., et al., *Recognition of cognate transfer RNA by the 30S ribosomal subunit*. Science, 2001. **292**(5518): p. 897-902.
43. Ogle, J.M., A.P. Carter, and V. Ramakrishnan, *Insights into the decoding mechanism from recent ribosome structures*. Trends Biochem Sci, 2003. **28**(5): p. 259-66.
44. Voorhees, R.M., et al., *Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome*. Nat Struct Mol Biol, 2009. **16**(5): p. 528-33.

45. Abdi, N.M. and K. Fredrick, *Contribution of 16S rRNA nucleotides forming the 30S subunit A and P sites to translation in Escherichia coli*. RNA, 2005. **11**(11): p. 1624-1632.
46. Qin, D., N.M. Abdi, and K. Fredrick, *Characterization of 16S rRNA mutations that decrease the fidelity of translation initiation*. RNA, 2007. **13**(12): p. 2348-2355.
47. Ali, I.K., et al., *Deletion of a Conserved, Central Ribosomal Intersubunit RNA Bridge*. Molecular Cell, 2006. **23**(6): p. 865-874.
48. Liu, Q. and K. Fredrick, *Contribution of intersubunit bridges to the energy barrier of ribosomal translocation*. Nucleic Acids Res, 2013. **41**(1): p. 565-74.
49. Demeshkina, N., et al., *A new understanding of the decoding principle on the ribosome*. Nature, 2012. **484**(7393): p. 256-9.
50. Moazed, D. and H.F. Noller, *Transfer RNA shields specific nucleotides in 16S ribosomal RNA from attack by chemical probes*. Cell, 1986. **47**(6): p. 985-94.
51. Yoshizawa, S., D. Fourmy, and J.D. Puglisi, *Recognition of the codon-anticodon helix by ribosomal RNA*. Science, 1999. **285**(5434): p. 1722-5.
52. Powers, T. and H.F. Noller, *Selective perturbation of G530 of 16 S rRNA by translational miscoding agents and a streptomycin-dependence mutation in protein S12*. J Mol Biol, 1994. **235**(1): p. 156-72.
53. Shimizu, Y., *Biochemical aspects of bacterial strategies for handling the incomplete translation processes*. Front Microbiol, 2014. **5**: p. 170.
54. Miller, M.R., et al., *The role of SmpB and the ribosomal decoding center in licensing tmRNA entry into stalled ribosomes*. RNA, 2011. **17**(9): p. 1727-36.
55. Pape, T., W. Wintermeyer, and M.V. Rodnina, *Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the E. coli ribosome*. EMBO J, 1998. **17**(24): p. 7490-7.
56. Kruger, K., et al., *Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena*. Cell, 1982. **31**(1): p. 147-57.

57. Zaug, A.J., P.J. Grabowski, and T.R. Cech, *Autocatalytic cyclization of an excised intervening sequence RNA is a cleavage-ligation reaction*. Nature, 1983. **301**(5901): p. 578-83.
58. Nissen, P., et al., *The structural basis of ribosome activity in peptide bond synthesis*. Science, 2000. **289**(5481): p. 920-30.
59. Simonovic, M. and T.A. Steitz, *A structural view on the mechanism of the ribosome-catalyzed peptide bond formation*. Biochim Biophys Acta, 2009. **1789**(9-10): p. 612-23.
60. Maguire, B.A., et al., *A protein component at the heart of an RNA machine: the importance of protein l27 for the function of the bacterial ribosome*. Mol Cell, 2005. **20**(3): p. 427-35.
61. Armache, J.P., et al., *Localization of eukaryote-specific ribosomal proteins in a 5.5-A cryo-EM map of the 80S eukaryotic ribosome*. Proc Natl Acad Sci U S A, 2010. **107**(46): p. 19754-9.
62. Ramakrishnan, V., *Ribosome structure and the mechanism of translation*. Cell, 2002. **108**(4): p. 557-72.
63. Wilson, D.N. and R. Beckmann, *The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling*. Curr Opin Struct Biol, 2011. **21**(2): p. 274-82.
64. Martinez, A.K., et al., *Crucial elements that maintain the interactions between the regulatory TnaC peptide and the ribosome exit tunnel responsible for Trp inhibition of ribosome function*. Nucleic Acids Res, 2012. **40**(5): p. 2247-57.
65. Ishida, H. and S. Hayward, *Path of nascent polypeptide in exit tunnel revealed by molecular dynamics simulation of ribosome*. Biophys J, 2008. **95**(12): p. 5962-73.
66. Voss, N.R., et al., *The geometry of the ribosomal polypeptide exit tunnel*. J Mol Biol, 2006. **360**(4): p. 893-906.
67. Agrawal, R.K., et al., *Visualization of tRNA movements on the Escherichia coli 70S ribosome during the elongation cycle*. J Cell Biol, 2000. **150**(3): p. 447-60.
68. Malhotra, A., et al., *Escherichia coli 70 S ribosome at 15 A resolution by cryo-electron microscopy: localization of fMet-tRNA^{fMet} and fitting of L1 protein*. J Mol Biol, 1998. **280**(1): p. 103-16.

69. Nikulin, A., et al., *Structure of the L1 protuberance in the ribosome*. Nat Struct Biol, 2003. **10**(2): p. 104-8.
70. Valle, M., et al., *Locking and unlocking of ribosomal motions*. Cell, 2003. **114**(1): p. 123-34.
71. Schuwirth, B.S., et al., *Structures of the bacterial ribosome at 3.5 Å resolution*. Science, 2005. **310**(5749): p. 827-34.
72. Achenbach, J. and K.H. Nierhaus, *Translocation at work*. Nat Struct Mol Biol, 2013. **20**(9): p. 1019-22.
73. Agirrezabala, X., et al., *Structural characterization of mRNA-tRNA translocation intermediates*. Proc Natl Acad Sci U S A, 2012. **109**(16): p. 6094-9.
74. Mohan, S. and H.F. Noller, *Recurring RNA structural motifs underlie the mechanics of L1 stalk movement*. Nat Commun, 2017. **8**: p. 14285.
75. Savelsbergh, A., et al., *An elongation factor G-induced ribosome rearrangement precedes tRNA-mRNA translocation*. Mol Cell, 2003. **11**(6): p. 1517-23.
76. Ryan, P.C., M. Lu, and D.E. Draper, *Recognition of the highly conserved GTPase center of 23 S ribosomal RNA by ribosomal protein L11 and the antibiotic thiostrepton*. J Mol Biol, 1991. **221**(4): p. 1257-68.
77. Ryan, P.C. and D.E. Draper, *Detection of a key tertiary interaction in the highly conserved GTPase center of large subunit ribosomal RNA*. Proc Natl Acad Sci U S A, 1991. **88**(14): p. 6308-12.
78. Rosendahl, G. and S. Douthwaite, *Ribosomal proteins L11 and L10.(L12)4 and the antibiotic thiostrepton interact with overlapping regions of the 23 S rRNA backbone in the ribosomal GTPase centre*. J Mol Biol, 1993. **234**(4): p. 1013-20.
79. Rodnina, M.V., et al., *GTPases mechanisms and functions of translation factors on the ribosome*. Biol Chem, 2000. **381**(5-6): p. 377-87.
80. Agrawal, R.K., et al., *EF-G-dependent GTP hydrolysis induces translocation accompanied by large conformational changes in the 70S ribosome*. Nat Struct Biol, 1999. **6**(7): p. 643-7.
81. Allen, G.S., et al., *The cryo-EM structure of a translation initiation complex from Escherichia coli*. Cell, 2005. **121**(5): p. 703-12.

82. Gabashvili, I.S., et al., *Solution structure of the E. coli 70S ribosome at 11.5 Å resolution*. Cell, 2000. **100**(5): p. 537-49.
83. Gao, H., et al., *RF3 induces ribosomal conformational changes responsible for dissociation of class I release factors*. Cell, 2007. **129**(5): p. 929-41.
84. Stark, H., et al., *Ribosome interactions of aminoacyl-tRNA and elongation factor Tu in the codon-recognition complex*. Nat Struct Biol, 2002. **9**(11): p. 849-54.
85. Villa, E., et al., *Ribosome-induced changes in elongation factor Tu conformation control GTP hydrolysis*. Proc Natl Acad Sci U S A, 2009. **106**(4): p. 1063-8.
86. Kumar, V., et al., *Structure of BipA in GTP form bound to the ratcheted ribosome*. Proc Natl Acad Sci U S A, 2015. **112**(35): p. 10944-9.
87. Ban, N., et al., *Placement of protein and RNA structures into a 5 Å-resolution map of the 50S ribosomal subunit*. Nature, 1999. **400**(6747): p. 841-7.
88. Gao, Y.G., et al., *The structure of the ribosome with elongation factor G trapped in the posttranslocational state*. Science, 2009. **326**(5953): p. 694-9.
89. Pulk, A. and J.H. Cate, *Control of ribosomal subunit rotation by elongation factor G*. Science, 2013. **340**(6140): p. 1235970.
90. Schmeing, T.M., et al., *The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA*. Science, 2009. **326**(5953): p. 688-694.
91. Tourigny, D.S., et al., *Elongation factor G bound to the ribosome in an intermediate state of translocation*. Science, 2013. **340**(6140): p. 1235490.
92. Voorhees, R.M., et al., *The mechanism for activation of GTP hydrolysis on the ribosome*. Science, 2010. **330**(6005): p. 835-838.
93. Zhou, J., et al., *Crystal structures of EF-G-ribosome complexes trapped in intermediate states of translocation*. Science, 2013. **340**(6140): p. 1236086.
94. Zhou, J., et al., *Crystal structure of release factor RF3 trapped in the GTP state on a rotated conformation of the ribosome*. RNA, 2012. **18**(2): p. 230-40.

95. Chan, Y.-L., C.C. Correll, and I.G. Wool, *The Location and the Significance of a Cross-link Between the Sarcin/Ricin Domain of Ribosomal RNA and the Elongation Factor-G*. Journal of Molecular Biology, 2004. **337**(2): p. 263-272.
96. Schmeing, T.M., et al., *The Crystal Structure of the Ribosome Bound to EF-Tu and Aminoacyl-tRNA*. Science, 2009. **326**(5953): p. 688.
97. Villa, E., et al., *Ribosome-induced changes in elongation factor Tu conformation control GTP hydrolysis*. Proceedings of the National Academy of Sciences, 2009. **106**(4): p. 1063.
98. Mears, J.A., et al., *Modeling a minimal ribosome based on comparative sequence analysis*. J Mol Biol, 2002. **321**(2): p. 215-34.
99. Hennelly, S.P., et al., *A time-resolved investigation of ribosomal subunit association*. J Mol Biol, 2005. **346**(5): p. 1243-58.
100. Gualerzi, C.O. and C.L. Pon, *Initiation of mRNA translation in prokaryotes*. Biochemistry, 1990. **29**(25): p. 5881-9.
101. Hartz, D., et al., *Domains of initiator tRNA and initiation codon crucial for initiator tRNA selection by Escherichia coli IF3*. Genes Dev, 1990. **4**(10): p. 1790-800.
102. Carter, A.P., et al., *Crystal structure of an initiation factor bound to the 30S ribosomal subunit*. Science, 2001. **291**(5503): p. 498-501.
103. Stark, H., et al., *Visualization of elongation factor Tu on the Escherichia coli ribosome*. Nature, 1997. **389**(6649): p. 403-6.
104. Dallas, A. and H.F. Noller, *Interaction of translation initiation factor 3 with the 30S ribosomal subunit*. Mol Cell, 2001. **8**(4): p. 855-64.
105. Moazed, D., et al., *Specific protection of 16 S rRNA by translational initiation factors*. J Mol Biol, 1995. **248**(2): p. 207-10.
106. Kisselev, L.L. and R.H. Buckingham, *Translational termination comes of age*. Trends Biochem Sci, 2000. **25**(11): p. 561-6.
107. Janosi, L., et al., *Ribosome recycling by ribosome recycling factor (RRF)-an important but overlooked step of protein biosynthesis*. Adv Biophys, 1996. **32**: p. 121-201.
108. Matzov, D., et al., *The cryo-EM structure of hibernating 100S ribosome dimer from pathogenic Staphylococcus aureus*. Nat Commun, 2017. **8**(1): p. 723.

109. Franken, L.E., et al., *A general mechanism of ribosome dimerization revealed by single-particle cryo-electron microscopy*. Nat Commun, 2017. **8**(1): p. 722.
110. Sato, A., et al., *Solution structure of the E. coli ribosome hibernation promoting factor HPF: Implications for the relationship between structure and function*. Biochem Biophys Res Commun, 2009. **389**(4): p. 580-5.
111. Polikanov, Y.S., G.M. Blaha, and T.A. Steitz, *How hibernation factors RMF, HPF, and YfiA turn off protein synthesis*. Science, 2012. **336**(6083): p. 915-8.
112. Kato, T., et al., *Structure of the 100S ribosome in the hibernation stage revealed by electron cryomicroscopy*. Structure, 2010. **18**(6): p. 719-24.
113. Beckert, B., et al., *Structure of the Bacillus subtilis hibernating 100S ribosome reveals the basis for 70S dimerization*. EMBO J, 2017. **36**(14): p. 2061-2072.
114. Ueta, M., et al., *Role of HPF (hibernation promoting factor) in translational activity in Escherichia coli*. J Biochem, 2008. **143**(3): p. 425-33.
115. Wada, A., et al., *Ribosome modulation factor: stationary growth phase-specific inhibitor of ribosome functions from Escherichia coli*. Biochem Biophys Res Commun, 1995. **214**(2): p. 410-7.
116. Trauner, A., et al., *The dormancy regulator DosR controls ribosome stability in hypoxic mycobacteria*. J Biol Chem, 2012. **287**(28): p. 24053-63.
117. Yoshida, H. and A. Wada, *The 100S ribosome: ribosomal hibernation induced by stress*. Wiley Interdiscip Rev RNA, 2014. **5**(5): p. 723-32.
118. Ortiz, J.O., et al., *Structure of hibernating ribosomes studied by cryoelectron tomography in vitro and in situ*. J Cell Biol, 2010. **190**(4): p. 613-21.
119. Khusainov, I., et al., *Structures and dynamics of hibernating ribosomes from Staphylococcus aureus mediated by intermolecular interactions of HPF*. EMBO J, 2017. **36**(14): p. 2073-2087.

120. Akiyama, T., et al., *Resuscitation of Pseudomonas aeruginosa from dormancy requires hibernation promoting factor (PA4463) for ribosome preservation*. Proc Natl Acad Sci U S A, 2017. **114**(12): p. 3204-3209.
121. Wower, I.K., C. Zwieb, and J. Wower, *Contributions of pseudoknots and protein SmpB to the structure and function of tmRNA in trans-translation*. J Biol Chem, 2004. **279**(52): p. 54202-9.
122. Lee, S., et al., *Determinants on tmRNA for initiating efficient and precise trans-translation: some mutations upstream of the tag-encoding sequence of Escherichia coli tmRNA shift the initiation point of trans-translation in vitro*. RNA, 2001. **7**(7): p. 999-1012.
123. Williams, K.P., K.A. Martindale, and D.P. Bartel, *Resuming translation on tmRNA: a unique mode of determining a reading frame*. EMBO J, 1999. **18**(19): p. 5423-33.
124. Ramrath, D.J., et al., *The complex of tmRNA-SmpB and EF-G on translocating ribosomes*. Nature, 2012. **485**(7399): p. 526-9.
125. Giudice, E., K. Mace, and R. Gillet, *Trans-translation exposed: understanding the structures and functions of tmRNA-SmpB*. Front Microbiol, 2014. **5**: p. 113.
126. Oh, B.K. and D. Apirion, *10Sa RNA, a small stable RNA of Escherichia coli, is functional*. Mol Gen Genet, 1991. **229**(1): p. 52-6.
127. Muto, A., et al., *Requirement of transfer-messenger RNA for the growth of Bacillus subtilis under stresses*. Genes Cells, 2000. **5**(8): p. 627-35.
128. Julio, S.M., D.M. Heithoff, and M.J. Mahan, *ssrA (tmRNA) plays a role in Salmonella enterica serovar Typhimurium pathogenesis*. J Bacteriol, 2000. **182**(6): p. 1558-63.
129. Okan, N.A., J.B. Bliska, and A.W. Karzai, *A Role for the SmpB-SsrA system in Yersinia pseudotuberculosis pathogenesis*. PLoS Pathog, 2006. **2**(1): p. e6.
130. Abo, T., et al., *SsrA-mediated protein tagging in the presence of miscoding drugs and its physiological role in Escherichia coli*. Genes Cells, 2002. **7**(7): p. 629-38.
131. Komine, Y., et al., *A tRNA-like structure is present in 10Sa RNA, a small stable RNA from Escherichia coli*. Proc Natl Acad Sci U S A, 1994. **91**(20): p. 9223-7.

132. Personne, Y. and T. Parish, *Mycobacterium tuberculosis possesses an unusual tmRNA rescue system*. Tuberculosis (Edinb), 2014. **94**(1): p. 34-42.
133. Gutmann, S., et al., *Crystal structure of the transfer-RNA domain of transfer-messenger RNA in complex with SmpB*. Nature, 2003. **424**(6949): p. 699-703.
134. Wower, I.K., C. Zwieb, and J. Wower, *Transfer-messenger RNA unfolds as it transits the ribosome*. RNA, 2005. **11**(5): p. 668-73.
135. Watts, T., et al., *SmpB contributes to reading frame selection in the translation of transfer-messenger RNA*. J Mol Biol, 2009. **391**(2): p. 275-81.
136. Metzinger, L., M. Hallier, and B. Felden, *The highest affinity binding site of small protein B on transfer messenger RNA is outside the tRNA domain*. RNA, 2008. **14**(9): p. 1761-72.
137. Metzinger, L., M. Hallier, and B. Felden, *Independent binding sites of small protein B onto transfer-messenger RNA during trans-translation*. Nucleic Acids Res, 2005. **33**(8): p. 2384-94.
138. Konno, T., et al., *A functional interaction of SmpB with tmRNA for determination of the resuming point of trans-translation*. RNA, 2007. **13**(10): p. 1723-31.
139. Weis, F., et al., *tmRNA-SmpB: a journey to the centre of the bacterial ribosome*. EMBO J, 2010. **29**(22): p. 3810-8.
140. Valle, M., et al., *Visualizing tmRNA entry into a stalled ribosome*. Science, 2003. **300**(5616): p. 127-30.
141. Fu, J., et al., *Visualizing the transfer-messenger RNA as the ribosome resumes translation*. EMBO J, 2010. **29**(22): p. 3819-25.
142. Keiler, K.C., *Mechanisms of ribosome rescue in bacteria*. Nat Rev Microbiol, 2015. **13**(5): p. 285-97.
143. Montero, C.I., et al., *Colocation of genes encoding a tRNA-mRNA hybrid and a putative signaling peptide on complementary strands in the genome of the hyperthermophilic bacterium Thermotoga maritima*. J Bacteriol, 2006. **188**(19): p. 6802-7.

144. Shasmal, M. and J. Sengupta, *Structural diversity in bacterial ribosomes: mycobacterial 70S ribosome structure reveals novel features*. PLoS One, 2012. **7**(2): p. e31742.
145. Li, X., et al., *Structure of Ribosomal Silencing Factor Bound to Mycobacterium tuberculosis Ribosome*. Structure, 2015. **23**(10): p. 1858-65.
146. Yang, K., et al., *Structural insights into species-specific features of the ribosome from the human pathogen Mycobacterium tuberculosis*. Nucleic Acids Res, 2017.
147. Hentschel, J., et al., *The Complete Structure of the Mycobacterium smegmatis 70S Ribosome*. Cell Rep, 2017. **20**(1): p. 149-160.
148. Li, Z., et al., *Cryo-EM structure of Mycobacterium smegmatis ribosome reveals two unidentified ribosomal proteins close to the functional centers*. Protein Cell, 2017.
149. De Rosier, D.J. and A. Klug, *Reconstruction of three dimensional structures from electron micrographs*. Nature, 1968. **217**(5124): p. 130-4.
150. Frank, J., *Single-particle imaging of macromolecules by cryo-electron microscopy*. Annu Rev Biophys Biomol Struct, 2002. **31**: p. 303-19.
151. Gao, H., et al., *Dynamics of EF-G interaction with the ribosome explored by classification of a heterogeneous cryo-EM dataset*. J Struct Biol, 2004. **147**(3): p. 283-90.
152. Frank, J. and R.L. Gonzalez, Jr., *Structure and dynamics of a processive Brownian motor: the translating ribosome*. Annu Rev Biochem, 2010. **79**: p. 381-412.
153. Frank, J., *Story in a sample-the potential (and limitations) of cryo-electron microscopy applied to molecular machines*. Biopolymers, 2013. **99**(11): p. 832-6.
154. Frank, J., B. Shimkin, and H. Dowse, *Spider—A modular software system for electron image processing*. Ultramicroscopy, 1981. **6**(4): p. 343-357.
155. Mulder, A.M., et al., *Visualizing ribosome biogenesis: parallel assembly pathways for the 30S subunit*. Science, 2010. **330**(6004): p. 673-7.
156. Milne, J.L.S., et al., *Cryo-electron microscopy: A primer for the non-microscopist*. The FEBS journal, 2013. **280**(1): p. 28-45.

157. Bartesaghi, A., et al., *Atomic Resolution Cryo-EM Structure of b-Galactosidase*. Structure, 2018. **26**(6): p. 848-856.e3.
158. Liu, Z., et al., *Determination of the ribosome structure to a resolution of 2.5 Å by single-particle cryo-EM*. Protein Sci, 2017. **26**(1): p. 82-92.
159. Brilot, A.F., et al., *Beam-Induced Motion of Vitrified Specimen on Holey Carbon Film*. Journal of structural biology, 2012. **177**(3): p. 630-637.
160. Henderson, R., *Image contrast in high-resolution electron microscopy of biological macromolecules: TMV in ice*. Ultramicroscopy, 1992. **46**(1): p. 1-18.
161. Henderson, R., *The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules*. Q Rev Biophys, 1995. **28**(2): p. 171-93.
162. Frank, J., *Chapter 2 - Electron Microscopy of Macromolecular Assemblies*, in *Three-Dimensional Electron Microscopy of Macromolecular Assemblies*, J. Frank, Editor. 1996, Academic Press: Burlington. p. 12-53.
163. van Heel, M., et al., *Single-particle electron cryo-microscopy: towards atomic resolution*. Q Rev Biophys, 2000. **33**(4): p. 307-69.
164. Nogales, E., *The development of cryo-EM into a mainstream structural biology technique*. Nature Methods, 2015. **13**: p. 24.
165. Glaeser, R.M., *How good can cryo-EM become?* Nature Methods, 2015. **13**: p. 28.
166. Sigworth, F.J., *Principles of cryo-EM single-particle image processing*. Microscopy (Oxf), 2016. **65**(1): p. 57-67.
167. Emsley, P., et al., *Features and development of Coot*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
168. Bell, J.M., et al., *High resolution single particle refinement in EMAN2.1*. Methods, 2016. **100**: p. 25-34.
169. Roy, A., A. Kucukural, and Y. Zhang, *I-TASSER: a unified platform for automated protein structure and function prediction*. Nat Protoc, 2010. **5**(4): p. 725-38.
170. Rother, M., et al., *ModeRNA server: an online tool for modeling RNA 3D structures*. Bioinformatics, 2011. **27**(17): p. 2441-2.

171. Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 12-21.
172. Afonine, P.V., et al., *Towards automated crystallographic structure refinement with phenix.refine*. Acta Crystallogr D Biol Crystallogr, 2012. **68**(Pt 4): p. 352-67.
173. Kelley, L.A., et al., *The Phyre2 web portal for protein modeling, prediction and analysis*. Nat Protoc, 2015. **10**(6): p. 845-58.
174. Scheres, S.H., *RELION: implementation of a Bayesian approach to cryo-EM structure determination*. J Struct Biol, 2012. **180**(3): p. 519-30.
175. Kucukelbir, A., F.J. Sigworth, and H.D. Tagare, *Quantifying the local resolution of cryo-EM density maps*. Nat Methods, 2014. **11**(1): p. 63-5.
176. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis*. J Comput Chem, 2004. **25**(13): p. 1605-12.
177. Brenner, S. and R.W. Horne, *A negative staining method for high resolution electron microscopy of viruses*. Biochim Biophys Acta, 1959. **34**: p. 103-10.
178. Dubochet, J., et al., *Cryo-electron microscopy of vitrified specimens*. Q Rev Biophys, 1988. **21**(2): p. 129-228.
179. Grassucci, R.A., D.J. Taylor, and J. Frank, *Preparation of macromolecular complexes for cryo-electron microscopy*. Nat Protoc, 2007. **2**(12): p. 3239-46.
180. Murata, K. and M. Wolf, *Cryo-electron microscopy for structural analysis of dynamic biological macromolecules*. Biochim Biophys Acta, 2018. **1862**(2): p. 324-334.
181. Vulović, M., et al., *Precise and unbiased estimation of astigmatism and defocus in transmission electron microscopy*. Ultramicroscopy, 2012. **116**: p. 115-134.
182. Williams, D.B. and C.B. Carter, *Transmission Electron Microscopy - 2009*. 2009.
183. Mescher, A.L., *Chapter 1. Histology & Its Methods of Study*, in *Junqueira's Basic Histology, 13e*. 2013, The McGraw-Hill Companies: New York, NY.

184. Spence, J.C.H., *High-resolution electron microscopy*. 2013, Oxford University Press: Oxford University Press.
185. Vulovic, M., et al., *A toolkit for the characterization of CCD cameras for transmission electron microscopy*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 97-109.
186. McMullan, G., A.R. Faruqi, and R. Henderson, *Chapter One - Direct Electron Detectors*, in *Methods in Enzymology*, R.A. Crowther, Editor. 2016, Academic Press. p. 1-17.
187. Booth, C.R., J. Jakana, and W. Chiu, *Assessing the capabilities of a 4kx4k CCD camera for electron cryo-microscopy at 300kV*. J Struct Biol, 2006. **156**(3): p. 556-63.
188. Brink, J. and W. Chiu, *Applications of a slow-scan CCD camera in protein electron crystallography*. J Struct Biol, 1994. **113**(1): p. 23-34.
189. Milazzo, A.C., et al., *Initial evaluation of a direct detection device detector for single particle cryo-electron microscopy*. J Struct Biol, 2011. **176**(3): p. 404-8.
190. Li, X., et al., *Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM*. Nat Methods, 2013. **10**(6): p. 584-90.
191. Bai, X.C., et al., *Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles*. Elife, 2013. **2**: p. e00461.
192. Baker, L.A. and J.L. Rubinstein, *Chapter Fifteen - Radiation Damage in Electron Cryomicroscopy*, in *Methods in Enzymology*, G.J. Jensen, Editor. 2010, Academic Press. p. 371-388.
193. L. Reimer, H.K., *Transmission Electron Microscopy: Physics of Image Formation* 2008: Springer.
194. Williams, D.B. and C.B. Carter, *Elastic Scattering*, in *Transmission Electron Microscopy: A Textbook for Materials Science*, D.B. Williams and C.B. Carter, Editors. 2009, Springer US: Boston, MA. p. 39-51.
195. Williams, D.B. and C.B. Carter, *Inelastic Scattering and Beam Damage*, in *Transmission Electron Microscopy: A Textbook for Materials Science*, D.B. Williams and C.B. Carter, Editors. 2009, Springer US: Boston, MA. p. 53-71.

196. Wade, R.H., *A brief look at imaging and contrast transfer*. Ultramicroscopy, 1992. **46**(1): p. 145-156.
197. Frank, J., *Three-Dimensional Electron Microscopy of Macromolecular Assemblies: Visualization of Biological Molecules in Their Native State*. 2006: Oxford University Press.
198. Penczek, P.A., *Image restoration in cryo-electron microscopy*. Methods Enzymol, 2010. **482**: p. 35-72.
199. H. Wade, R. and J. Frank, *Electron microscope transfer functions for partially coherent axial illumination and chromatic defocus spread*. Optik. Vol. 49. 1977.
200. Zhu, J., et al., *Three-Dimensional Reconstruction with Contrast Transfer Function Correction from Energy-Filtered Cryoelectron Micrographs: Procedure and Application to the 70S Escherichia coli Ribosome*. Journal of Structural Biology, 1997. **118**(3): p. 197-219.
201. van Heel, M., et al., *A new generation of the IMAGIC image processing system*. J Struct Biol, 1996. **116**(1): p. 17-24.
202. Shaikh, T.R., et al., *SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs*. Nat Protoc, 2008. **3**(12): p. 1941-74.
203. Sorzano, C.O.S., et al., *XMIPP: a new generation of an open-source image processing package for electron microscopy*. Journal of Structural Biology, 2004. **148**(2): p. 194-204.
204. Grigorieff, N., *FREALIGN: High-resolution refinement of single particle structures*. Journal of Structural Biology, 2007. **157**(1): p. 117-125.
205. Chung, J.-H. and H. Min Kim, *The Nobel Prize in Chemistry 2017: High-Resolution Cryo-Electron Microscopy*. Vol. 47. 2017. 218-222.
206. Arbeláez, P., et al., *Experimental evaluation of support vector machine-based and correlation-based approaches to automatic particle selection*. Journal of Structural Biology, 2011. **175**(3): p. 319-328.
207. Zhao, J., M.A. Brubaker, and J.L. Rubinstein, *TMacS: A hybrid template matching and classification system for partially-automated particle selection*. Journal of Structural Biology, 2013. **181**(3): p. 234-242.

208. Langlois, R., J. Pallesen, and J. Frank, *Reference-free particle selection enhanced with semi-supervised machine learning for cryo-electron microscopy*. Journal of Structural Biology, 2011. **175**(3): p. 353-361.
209. Moody, M.F., *Chapter 1 - Overview*, in *Structural Biology Using Electrons and X-rays*, M.F. Moody, Editor. 2011, Academic Press: Boston. p. 1-8.
210. Glaeser, R.M. and R.J. Hall, *Reaching the information limit in cryo-EM of biological macromolecules: experimental aspects*. Biophys J, 2011. **100**(10): p. 2331-7.
211. Mindell, J.A. and N. Grigorieff, *Accurate determination of local defocus and specimen tilt in electron microscopy*. J Struct Biol, 2003. **142**(3): p. 334-47.
212. Yang, Z., et al., *Iterative stable alignment and clustering of 2D transmission electron microscope images*. Structure, 2012. **20**(2): p. 237-47.
213. van Heel, M. and M. Stoffler-Meilicke, *Characteristic views of E. coli and B. stearothermophilus 30S ribosomal subunits in the electron microscope*. EMBO J, 1985. **4**(9): p. 2389-95.
214. Scheres, S.H., et al., *Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization*. Nat Methods, 2007. **4**(1): p. 27-9.
215. Singh, R., et al., *Cryo-electron microscopic structure of SecA protein bound to the 70S ribosome*. J Biol Chem, 2014. **289**(10): p. 7190-9.
216. Scheres, S.H. and S. Chen, *Prevention of overfitting in cryo-EM structure determination*. Nat Methods, 2012. **9**(9): p. 853-4.
217. van Heel, M. and M. Schatz, *Fourier shell correlation threshold criteria*. J Struct Biol, 2005. **151**(3): p. 250-62.
218. Chen, S., et al., *High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy*. Ultramicroscopy, 2013. **135**: p. 24-35.
219. Rosenthal, P.B. and R. Henderson, *Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy*. J Mol Biol, 2003. **333**(4): p. 721-45.

220. Lyumkis, D., et al., *Chapter Fourteen - Automation in Single-Particle Electron Microscopy: Connecting the Pieces*, in *Methods in Enzymology*, G.J. Jensen, Editor. 2010, Academic Press. p. 291-338.
221. Suloway, C., et al., *Automated molecular microscopy: The new Leginon system*. Journal of Structural Biology, 2005. **151**(1): p. 41-60.
222. Zhang, W., J.A. Dunkle, and J.H. Cate, *Structures of the ribosome in intermediate states of ratcheting*. Science, 2009. **325**(5943): p. 1014-7.
223. Voorhees, R.M., et al., *Structure of the mammalian ribosome-Sec61 complex to 3.4 Å resolution*. Cell, 2014. **157**(7): p. 1632-43.
224. Forsberg, B.O., et al., *Cryo-EM reconstruction of the chlororibosome to 3.2 Å resolution within 24 h*. IUCrJ, 2017. **4**(Pt 6): p. 723-727.
225. Wu, S., et al., *Fabs enable single particle cryoEM studies of small proteins*. Structure, 2012. **20**(4): p. 582-92.
226. Chua, E.Y.D. and S. Sandin, *Advances in phase plate cryo-EM imaging of DNA and nucleosomes*. Nucleus, 2017. **8**(3): p. 275-278.
227. Murata, K., et al., *Zernike Phase Contrast Cryo-Electron Microscopy and Tomography for Structure Determination at Nanometer and Subnanometer Resolutions*. Structure, 2010. **18**(8): p. 903-912.
228. Bai, X.C., et al., *An atomic structure of human gamma-secretase*. Nature, 2015. **525**(7568): p. 212-217.
229. Merk, A., et al., *Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery*. Cell, 2016. **165**(7): p. 1698-1707.
230. Dashti, A., et al., *Trajectories of the ribosome as a Brownian nanomachine*. Proc Natl Acad Sci U S A, 2014. **111**(49): p. 17492-7.
231. Jin, Q., et al., *Iterative elastic 3D-to-2D alignment method using normal modes for studying structural dynamics of large macromolecular complexes*. Structure, 2014. **22**(3): p. 496-506.
232. Apweiler, R., et al., *UniProt: the Universal Protein knowledgebase*. Nucleic Acids Res, 2004. **32**(Database issue): p. D115-9.
233. Noeske, J., et al., *High-resolution structure of the Escherichia coli ribosome*. Nat Struct Mol Biol, 2015. **22**(4): p. 336-41.
234. Benson, D.A., et al., *GenBank*. Nucleic Acids Res, 2013. **41**(Database issue): p. D36-42.

235. Fischer, N., et al., *Structure of the E. coli ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM*. Nature, 2015. **520**(7548): p. 567-70.
236. Hudson, C.M. and K.P. Williams, *The tmRNA website*. Nucleic Acids Res, 2015. **43**(Database issue): p. D138-40.
237. Mishra, S., et al., *Structures of Mycobacterium smegmatis 70S ribosomes in complex with HPF, tmRNA, and P-tRNA*. Scientific Reports, 2018. **8**(1): p. 13587.
238. Kamath, D., S.T. Gregory, and M. O'Connor, *The Loop 2 Region of Ribosomal Protein uS5 Influences Spectinomycin Sensitivity, Translational Fidelity, and Ribosome Biogenesis*. Antimicrob Agents Chemother, 2017. **61**(2).
239. Abeysirigunawardena, S.C., et al., *Evolution of protein-coupled RNA dynamics during hierarchical assembly of ribosomal complexes*. Nat Commun, 2017. **8**(1): p. 492.
240. Held, W.A. and M. Nomura, *Escherichia coli 30 S ribosomal proteins uniquely required for assembly*. J Biol Chem, 1975. **250**(8): p. 3179-84.
241. Nikolay, R., et al., *Ribosomal Proteins: Role in Ribosomal Functions*. eLS. John Wiley & Sons Ltd, 2015.
242. Davis, J.H. and J.R. Williamson, *Structure and dynamics of bacterial ribosome biogenesis*. Philos Trans R Soc Lond B Biol Sci, 2017. **372**(1716).
243. Akanuma, G., et al., *Inactivation of ribosomal protein genes in Bacillus subtilis reveals importance of each ribosomal protein for cell proliferation and cell differentiation*. J Bacteriol, 2012. **194**(22): p. 6282-91.
244. Lu, M. and T.A. Steitz, *Structure of Escherichia coli ribosomal protein L25 complexed with a 5S rRNA fragment at 1.8-Å resolution*. Proc Natl Acad Sci U S A, 2000. **97**(5): p. 2023-8.
245. Jenner, L., et al., *Structural rearrangements of the ribosome at the tRNA proofreading step*. Nat Struct Mol Biol, 2010. **17**(9): p. 1072-8.
246. Gongadze, G.M., et al., *Bacterial 5S rRNA-binding proteins of the CTC family*. Biochemistry (Mosc), 2008. **73**(13): p. 1405-17.

247. Ban, N., et al., *A new system for naming ribosomal proteins*. Curr Opin Struct Biol, 2014. **24**: p. 165-9.
248. Jenner, L., et al., *Crystal structure of the 80S yeast ribosome*. Curr Opin Struct Biol, 2012. **22**(6): p. 759-67.
249. Sun, M., et al., *Dynamical features of the Plasmodium falciparum ribosome during translation*. Nucleic Acids Res, 2015. **43**(21): p. 10515-24.
250. Agüero, F., et al., *Genomic-scale prioritization of drug targets: the TDR Targets database*. Nature reviews. Drug discovery, 2008. **7**(11): p. 900-907.
251. Lew, J.M., et al., *TubercuList – 10 years after*. Tuberculosis, 2011. **91**(1): p. 1-7.
252. Ben-Shem, A., et al., *The structure of the eukaryotic ribosome at 3.0 Å resolution*. Science, 2011. **334**(6062): p. 1524-9.
253. Amunts, A., et al., *Ribosome. The structure of the human mitochondrial ribosome*. Science, 2015. **348**(6230): p. 95-98.
254. Khatter, H., et al., *Structure of the human 80S ribosome*. Nature, 2015. **520**(7549): p. 640-5.
255. Desai, N., et al., *The structure of the yeast mitochondrial ribosome*. Science, 2017. **355**(6324): p. 528-531.
256. Kisly, I., et al., *The functional role of eL19 and eB12 intersubunit bridge in the eukaryotic ribosome*. Journal of molecular biology, 2016. **428**(10 Pt B): p. 2203-2216.
257. Yu, X. and J.R. Warner, *Expression of a micro-protein*. J Biol Chem, 2001. **276**(36): p. 33821-5.
258. Meskauskas, A., et al., *Decreased peptidyltransferase activity correlates with increased programmed -1 ribosomal frameshifting and viral maintenance defects in the yeast Saccharomyces cerevisiae*. RNA, 2003. **9**(8): p. 982-992.
259. Dresios, J., et al., *A dispensable yeast ribosomal protein optimizes peptidyltransferase activity and affects translocation*. J Biol Chem, 2003. **278**(5): p. 3314-22.
260. Ben-Shem, A., et al., *The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution*. Science, 2011. **334**(6062): p. 1524.

261. Bieri, P., et al., *The complete structure of the chloroplast 70S ribosome in complex with translation factor pY*. EMBO J, 2017. **36**(4): p. 475-486.
262. Dinman, J.D., *5S rRNA: Structure and Function from Head to Toe*. Int J Biomed Sci, 2005. **1**(1): p. 2-7.
263. Usachev, K.S., et al., *NMR assignments of the N-terminal domain of Staphylococcus aureus hibernation promoting factor (SaHPF)*. Biomol NMR Assign, 2017.
264. Ahmed, T., J. Shi, and S. Bhushan, *Unique localization of the plastid-specific ribosomal proteins in the chloroplast ribosome small subunit provides mechanistic insights into the chloroplastic translation*. Nucleic Acids Res, 2017. **45**(14): p. 8581-8595.
265. Demo, G., et al., *Structure of RNA polymerase bound to ribosomal 30S subunit*. Elife, 2017. **6**.
266. Byrgazov, K., et al., *Structural basis for the interaction of protein S1 with the Escherichia coli ribosome*. Nucleic Acids Res, 2015. **43**(1): p. 661-73.
267. Park, E., et al., *Structure of the SecY channel during initiation of protein translocation*. Nature, 2014. **506**(7486): p. 102-6.
268. Salah, P., et al., *Probing the relationship between Gram-negative and Gram-positive S1 proteins by sequence analysis*. Nucleic Acids Res, 2009. **37**(16): p. 5578-88.
269. Byrgazov, K., et al., *Direct interaction of the N-terminal domain of ribosomal protein S1 with protein S2 in Escherichia coli*. PLoS One, 2012. **7**(3): p. e32702.
270. Duval, M., et al., *Escherichia coli ribosomal protein S1 unfolds structured mRNAs onto the ribosome for active translation initiation*. PLoS Biol, 2013. **11**(12): p. e1001731.
271. Delvillani, F., et al., *S1 ribosomal protein and the interplay between translation and mRNA decay*. Nucleic Acids Res, 2011. **39**(17): p. 7702-15.
272. Subramanian, A.R. and J. van Duin, *Exchange of individual ribosomal proteins between ribosomes as studied by heavy isotope-transfer experiments*. Mol Gen Genet, 1977. **158**(1): p. 1-9.

273. Fan, Y., et al., *Structural basis for ribosome protein S1 interaction with RNA in trans-translation of Mycobacterium tuberculosis*. Biochem Biophys Res Commun, 2017. **487**(2): p. 268-273.
274. Kurkcuoglu, O., et al., *The ribosome structure controls and directs mRNA entry, translocation and exit dynamics*. Phys Biol, 2008. **5**(4): p. 046005.
275. Vellano, R.L. and J.C. Rabinowitz, *The influence of ribosome-binding-site elements on translational efficiency in Bacillus subtilis and Escherichia coli in vivo*. Mol Microbiol, 1992. **6**(9): p. 1105-14.
276. Prabhakaran, R., S. Chithambaram, and X. Xia, *Escherichia coli and Staphylococcus phages: effect of translation initiation efficiency on differential codon adaptation mediated by virulent and temperate lifestyles*. J Gen Virol, 2015. **96**(Pt 5): p. 1169-79.
277. Abolbaghaei, A., J.R. Silke, and X. Xia, *How Changes in Anti-SD Sequences Would Affect SD Sequences in Escherichia coli and Bacillus subtilis*. G3 (Bethesda), 2017. **7**(5): p. 1607-1615.
278. Steitz, J.A. and K. Jakes, *How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in Escherichia coli*. Proc Natl Acad Sci U S A, 1975. **72**(12): p. 4734-8.
279. Shell, S.S., et al., *Leaderless Transcripts and Small Proteins Are Common Features of the Mycobacterial Translational Landscape*. PLoS Genet, 2015. **11**(11): p. e1005641.
280. Kaur, S., et al., *Cryo-EM visualization of transfer messenger RNA with two SmpBs in a stalled ribosome*. Proc Natl Acad Sci U S A, 2006. **103**(44): p. 16484-9.
281. Fu, J., et al., *tmRNA on its way through the ribosome: two steps of resume, and what next?* RNA Biol, 2011. **8**(4): p. 586-90.
282. Okada, T., et al., *Contribution of the second OB fold of ribosomal protein S1 from Escherichia coli to the recognition of TmRNA*. Biosci Biotechnol Biochem, 2004. **68**(11): p. 2319-25.
283. Gur, E. and R.T. Sauer, *Evolution of the ssrA degradation tag in Mycoplasma: specificity switch to a different protease*. Proc Natl Acad Sci U S A, 2008. **105**(42): p. 16113-8.

7. APPENDIX

Appendix figure 1. Secondary structure of M. smegmatis 16S rRNA

Figure adapted from [147]

Appendix figure 2. Secondary structure of *M. smegmatis* 23S rRNA and 5S rRNA

Figure adapted from [147]

B

Figure adapted from [147]

Appendix Figure 3. Alignments of the H54a region, actinobacterial 23S rRNA genes

Figure adapted from [147]

Composition of Hibernating state state LSU model. The r-protein IDs are from Uniprot and the rRNA IDs are from Genbank. The chain IDs are according to the PDB ID: 5ZEP.

Protein/RNA name	Chain ID in 50S	Uniprot ID (GenBank ID)	Complete size	Modeled	Homologous Ec size	comments
23S rRNA	A	NR_076104.1	3120	2-3119	2897	
5S rRNA	B	NR_075650.1	118	1-117	120	
uL1	-	A0QS46	235	-		not visible
uL2	C	A0QSD4	278	3-275	273	
uL3	D	A0QSD1	217	2-215	209	
uL4	E	A0QSD2	215	3-209	201	
uL5	F	A0QSG1	187	6-186	179	
uL6	G	A0QSG4	179	2-177	177	
bL7/bL12	-			-		not visible
bL9	H	A0R7F6	151	1-151	149	
uL10	I	A0QS62	175	1-126	165	
uL11	J	A0QS45	142	10-142	142	
uL13	K	A0QSP8	147	1-147	147	
uL14	L	A0QSF9	122	1-121	123	
uL15	M	A0QSG8	147	3-147	144	
uL16	N	A0QSD8	138	1-134	136	
bL17	O	A0QSL9	199	2-118	127	
uL18	P	A0QSG5	127	2-127	117	
bL19	Q	A0QV42	113	1-113	115	
bL20	R	A0QYU6	129	2-125	118	
bL21	S	A0R151	103	2-103	103	
uL22	T	A0QSD6	153	6-119	110	
uL23	U	A0QSD3	100	4-97	100	
uL24	V	A0QSG0	105	1-105	104	
bL25	W	A0R3D2	215	6-193	94	
bL27	X	A0R150	88	5-86	85	
bL28	Y	A0QV03	64	2-64	78	
uL29	Z	A0QSD9	77	5-57	63	
uL30	v	A0QSG7	61	2-61	59	
bL31	y	A0R215	82	1-66	70	
bL32	z	A0R3I9	57	2-55	57	
bL33	1	A0QS39	55	6-55	55	
bL34	2	A0R7K0	47	3-47	46	
bL35	3	A0QYU7	64	2-64	65	
bL36	4	A0QSL4	37	1-37	38	
bL37	5	A0QTP4	24	2-24	-	

Composition of Hibernating state state SSU model. The r-protein IDs are from Uniprot and the rRNA IDs are from Genbank. The chain IDs are according to the PDB ID: 5ZEP.

Protein/RNA name	Chain ID in SSU	Uniprot ID (GenBank ID)	Complete size	Modeled	Homologous <i>Ec</i> size	comments
16S rRNA	a	X52922.1	1528	9-1518	1541	
uS2	b	A0QVB8	277	1-228	241	
uS3	c	A0QSD7	275	1-210	233	
uS4	d	A0QSL7	201	2-201	206	
uS5	e	A0QSG6	214	17-214	167	
bS6	f	A0A0D6J3X3	96	1-96	135	
uS7	g	A0QS97	156	1-156	179	
uS8	h	A0QSG3	132	3-132	130	
uS9	i	A0QSP9	150	25-150	130	
uS10	j	A0QSD0	101	5-101	103	
uS11	k	A0QSL6	138	22-138	129	
uS12	l	A0QS96	124	2-123	124	
uS13	m	A0QSL5	124	2-117	118	
uS14	n	A0QSG2	61	2-61	101	
uS15	o	A0QVQ3	89	2-88	89	
uS16	p	A0QV37	156	2-114	52	
uS17	q	A0QSE0	98	6-97	84	
bS18	r	A0R7F7	84	16-79	75	
uS19	s	A0QSD5	93	6-83	92	
bS20	t	A0R102	86	3-86	87	
bS22	u	A0QR10	33	2-33	-	
E-site tRNA ^{fMet}	w	CP011124.1	77	1-77	77	<i>Ec</i> E-tRNA ^{fMet}
HPF	x	A0QTK6	230	31-130	95	Rigid body fitted
bS1	0	A0QYY6	479	223 residues	556	Rigid body fitted

Composition of Trans-translating state model. The *r*-protein IDs are from Uniprot and the *r*RNA IDs are from Genbank. The chain IDs are according to the PDB ID: 5ZEY.

Protein/RNA name	Chain ID in SSU	Uniprot ID (GenBank ID)	Complete size	Modeled	Homologous <i>Ec</i> size	Comments
tmRNA	A	HG522090.1	369	1-369	363	Rigid body fitted
A-site tRNA ^{fMet}	B	CP011124.1	77	1-77	77	<i>Ec</i> A-site tRNA ^{fMet} , Rigid body fitted
SmpB	C	A0QU63	161	1-161	160	Rigid body fitted

8. Online access of cryo-EM maps and models

1. 70S-log dataset (Krios - 300 kV):

The 3.4 Å 70S, 3.2 Å 50S and 3.7 Å 30S cryo-EM density maps of the *M. smegmatis* 70S ribosome at P/P state have been deposited to the Electron Microscopy Data Bank (EMDB) with accession code EMD - 6054, EMD - 6941, and EMD - 6943 and the coordinates of the atomic structures are deposited to the Protein Data Bank (PDB) under accession code PDB - 5ZEB, PDB - 5ZET, and PDB - 5ZEU, respectively.

The 12.5 Å cryo-EM density map of the *M. smegmatis* 70S ribosome at the trans-translation state has been deposited to the EMDB with accession code EMD - 6937 and the coordinates of the atomic structures are deposited to the PDB under accession code PDB - 5ZEY.

2. 70S-stat dataset (Arctica - 200 kV):

The 4.1 Å cryo-EM density map of the *M. smegmatis* 70S ribosome at the hibernating state has been deposited to the EMDB with accession code EMD - 6922 and the coordinates of the atomic structures are deposited to the PDB under accession code PDB - 5ZEP.

**9. Author's publications related to this thesis during the Ph.D. period-
Aug 2014 to Aug 2018**

Research articles:

Mishra, S., T. Ahmed, A. Tyagi, J. Shi and S. Bhushan (2018). "Structures of Mycobacterium smegmatis 70S ribosomes in complex with HPF, tmRNA, and P-tRNA." Scientific Reports 8(1): 13587.

Talk presentation:

10th Annual Retreat, Konstanz Research School Chemical Biology (KoRS-CB), Konstanz, Germany. Aug 14-17, 2018.

Poster presentation:

10th Annual Retreat, Konstanz Research School Chemical Biology (KoRS-CB), Konstanz, Germany. Aug 14-17, 2018.