Cryo-EM structure of the chloroplast ribosome

Tofayel Ahmed

2018

Tofayel Ahmed. (2018). Cryo-EM structure of the chloroplast ribosome. Doctoral thesis,
Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/75853>

<https://doi.org/10.32657/10356/75853>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**CRYO-EM STRUCTURE OF THE
CHLOROPLAST RIBOSOME**

TOFAYEL AHMED

SCHOOL OF BIOLOGICAL SCIENCES

2018

**CRYO-EM STRUCTURE OF THE
CHLOROPLAST RIBOSOME**

TOFAYEL AHMED

TOFAYEL AHMED

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**

2018

Acknowledgements

First of all, I would like to extend my heartfelt gratitude to my supervisor Asst. Prof. Shashi Bhushan for providing me with the opportunity to carry out my PhD research in his lab. His able guidance and patience during all these years have helped me to develop critical thinking skills. Without his constant encouragement and advice, this thesis would not have been possible today. I will, forever, remain indebted to him.

My sincere gratitude also goes to my Thesis Advisory Committee (TAC) members, Asst. Prof. Oliver Mueller-Cajar and Assoc. Prof. Surajit Bhattacharyya for their support and constructive suggestions during and after the TAC meetings.

My thanks go to Dr. Rahul Jaiswal for his uncountable help during experiments. The way he has spent his valuable time explaining details of many experiments to me is something I deeply appreciate. Special thanks to Satabdi for her valuable suggestions during some experiments. I also owe many thanks to Dawn and Zhan for helping me out with data processing and Dr. Eugene Chua for the numerous discussions on molecular modeling. My appreciation goes to all other members from both Asst. Prof. Shashi Bhushan and Asst. Prof. Sara Sandin's labs for providing me with a pleasant work environment. I am thankful to Arvind for being the source of much-needed moral support and providing technical help in writing this report. I also owe sincere thanks to Dev, Priya, Moumita and Manisha for their unconditional support, which kept me going during difficult times.

I owe special thanks to Dr. Andrew See Weng Wong and Dr. Jian Shi for their immense help and support during all phases of cryo-EM data collection. I can't imagine writing this thesis without them. I highly appreciate the help in computing from Mr. Rikky Wenang Purbojati and Mr. Darshan Chandrakant Redij. I also acknowledge the excellent computing facilities at National Supercomputing Center (NSCC), Singapore.

I gratefully acknowledge the financial support in the form of the student stipend and conference allowances provided by the School of Biological Sciences. I am also grateful to all the academic and technical staff members of our school who had relentlessly supported me during the period of this work.

I would like to express gratitude to my parents, Dr. Shamsuzzaman Ahmed and Mrs. Serina Khatun for encouraging me to pursue my dreams while being always at the receiving end during my setbacks. Finally, I would like to extend my love to my sister, Ela, for being a constant source of joy and happiness in my personal life.

Dedication

This thesis is dedicated to the memory of late Mr. Narayan Sanyal, whose literary work, *Viswasghatak* (The Traitor), was the earliest influence for me to pursue Science as a career.

Table of contents

Acknowledgements.....	i
Dedication.....	iii
Table of contents.....	iv
List of figures.....	vii
List of tables.....	ix
Abbreviations.....	x
Abstract.....	xiii
1. Introduction.....	1
1.1 Central dogma of Molecular Biology.....	2
1.2 The ribosome.....	3
1.3 Structural studies on the ribosome.....	3
1.4 Functionally important sites of the ribosome.....	4
1.4.1 Decoding center (DC).....	5
1.4.2 Peptidyl transferase center (PTC).....	6
1.4.3 The polypeptide exit tunnel and tunnel exit site.....	8
1.4.4 The mRNA entry and exit sites.....	9
1.4.5 Translation factor binding site.....	10
1.5 Mechanism of translation in bacteria.....	11
1.6 Hibernating ribosomes.....	12
1.7 Organellar ribosomes.....	14
1.8 Structure of the cp-ribosome.....	14
1.9 Chloroplastic translation and its regulation.....	15
1.10 Cryo-EM as the method of choice.....	17
1.11 Aim (outline of the thesis).....	18
2. Materials and Methods.....	19
2.1 Materials.....	20
2.1.1 Chemicals.....	21
2.1.2 Buffer compositions.....	21
2.1.3 Other instrumentation and materials.....	22
2.1.4 Computer tools and software.....	22

2.2	Methods.....	23
2.2.1	Biochemical method- Purification of cp-ribosomes from spinach leaves.....	24
2.2.2	Biophysical methods.....	25
2.2.2.1	History and introduction.....	25
2.2.2.2	The transmission electron microscope (TEM).....	26
2.2.2.3	The gun.....	27
2.2.2.4	The lens system.....	28
2.2.2.5	The detector.....	29
2.2.3	Principles of EM image formation.....	29
2.2.4	Contrast Transfer Function (CTF).....	30
2.2.5	Cryo-EM.....	31
2.2.6	Single particle analysis.....	31
2.2.6.1	Image processing.....	31
2.2.6.2	Particle picking.....	32
2.2.6.3	Estimation of CTF parameters and correction for their effects.....	33
2.2.6.4	2D clustering and formation of class averages.....	33
2.2.6.5	Calculation of the initial structures.....	34
2.2.6.6	Structure refinement and resolution.....	34
2.2.7	Latest developments in detector technology and image processing...	35
2.2.8	Limitations and future direction of the technology.....	36
2.2.9	Cryo-electron microscopy of cp-ribosome.....	37
2.2.10	Image processing of cp-ribosome Arctica-dataset.....	38
2.2.11	Image processing of cp-ribosome Krios-dataset.....	38
2.2.12	Model building and refinement of cp-ribosome LSU and SSU.....	39
2.2.13	Model validation.....	41
2.2.14	Figure generation.....	41
3.	Results.....	42
3.1	Structural insights into the large subunit of the cp-ribosome.....	45
3.2	Purification profile of the cp-ribosome.....	46
3.3	Ribosomal particles on cryo-EM grid.....	47
3.4	Image processing.....	47

3.5	2D classification.....	49
3.6	Initial 3D refinement and subsequent sorting by 3D classification	50
3.7	Estimation of resolution for the final LSU reconstruction.....	51
3.8	Overall architecture of the cp-ribosome LSU.....	53
3.9	Chloroplast-specific protein-extensions in the LSU.....	56
3.10	Localization and modeling of cL37 (PSRP5) and cL38 (PSRP6) in the LSU.....	59
3.11	Remodeling of rRNA structure in the LSU.....	62
3.12	Polypeptide exit tunnel and tunnel exit site in the LSU.....	67
4.1	Structural insights into the small subunit of the cp-ribosome.....	71
4.2	Image processing.....	72
4.3	Resolution estimation for the SSU, LSU and 70S ribosome from Krios-dataset.....	73
4.4	Model validation for the SSU, LSU and 70S ribosome from the Krios-dataset.....	75
4.5	Overall architecture of the cp-ribosome SSU.....	77
4.6	Localization and modeling of cS22 (PSRP2) at the SSU foot.....	81
4.7	Localization and modeling of bTHXc (PSRP4) at the SSU head....	83
4.8	Unique features of the mRNA entry and exit sites.....	85
4.9	An extended bS1c connects the mRNA entry and exit sites on the SSU.....	89
4.10	Mode of interaction of plastid pY with the ribosome.....	92
4.11	Intersubunit bridges in 70S cp-ribosome.....	96
5	Discussion.....	100
6	Conclusion.....	103
7	References.....	105
8	Appendix.....	123
9	Information regarding online access of cryo-EM maps and models	134
10	Author's publications related to this thesis during the Ph.D. period-Jan 2014 to Jan 2018.....	135

List of figures

Figure 1.1	The components of prokaryotic and eukaryotic ribosomes.....	3
Figure 1.2	The decoding center (DC) on the SSU.....	5
Figure 1.3	The peptidyl transferase center (PTC) and formation of peptide bond	7
Figure 1.4	The polypeptide exit tunnel in the LSU.....	8
Figure 1.5	The tunnel exit site located at the bottom of LSU.....	9
Figure 1.6	The mRNA entry and exit sites on the SSU.....	10
Figure 1.7	The GTPase-associated Center (GAC) of the ribosome.....	11
Figure 1.8	The steps in bacterial translation.....	12
Figure 1.9	The hibernation promoting factor (HPF).....	13
Figure 2.1	The components of a TEM.....	27
Figure 2.2	The typical workflow in cryo-EM based structural study.....	32
Figure 3.1	A ₂₈₀ profile of spinach cp-ribosome crude extract on a sucrose gradient.....	46
Figure 3.2	Distribution and contrast of particles on cryo-EM grid.....	47
Figure 3.3	Image processing flowchart for the Arctica-dataset.....	48
Figure 3.4	2D classification results.....	49
Figure 3.5	3D classification results.....	50
Figure 3.6	Overall and local resolution estimation and map quality.....	51
Figure 3.7	Overall architecture of the cp-ribosome LSU.....	53
Figure 3.8	Protein loss and gain in cp-ribosome compared to <i>E. coli</i>	54
Figure 3.9	Extensions in cp-ribosome LSU proteins.....	56
Figure 3.10	Comparison between <i>E. coli</i> ribosome LSU proteins and homologous cp-ribosome LSU proteins.....	59
Figure 3.11	The localization and modeling of cL37 (PSRP5).....	60
Figure 3.12	The localization and modeling of cL38 (PSRP6).....	61
Figure 3.13	The domain organization of the 23S rRNA in the cp-ribosome LSU	63
Figure 3.14	Remodeled rRNA elements in the cp-ribosome LSU.....	64
Figure 3.15	Protrusions in cp-ribosome 23S rRNA at K-turn locations of <i>H.</i> <i>marismortui</i>	67
Figure 3.16	The polypeptide exit tunnel and polypeptide exit site of cp-ribosome LSU.....	68
Figure 3.17	3D structure and multiple sequence alignment of protein L23.....	69

Figure 4.1	Image processing flowchart for the Krios-dataset.....	73
Figure 4.2	Resolution estimation of cp-ribosome maps derived from Krios-dataset.....	74
Figure 4.3	Local agreement between cp-ribosome maps and models.....	76
Figure 4.4	Overall architecture of the 70S cp-ribosome and SSU.....	78
Figure 4.5	Constituent proteins of cp-ribosome SSU.....	79
Figure 4.6	Localization and modeling of cS22.....	81
Figure 4.7	rRNA deletion in cp-ribosome SSU (16S rRNA).....	82
Figure 4.8	Localization and modeling of bTHXc.....	83
Figure 4.9	Multiple sequence alignment (MSA) and phylogenetic tree of selected bTHXc-homologs.....	84
Figure 4.10	The mRNA entry and exit sites.....	86
Figure 4.11	The mRNA exit site.....	88
Figure 4.12	Localization and modeling of bS1c.....	90
Figure 4.13	Localization and modeling of the factor, plastid pY.....	93
Figure 4.14	Interaction of plastid pY with the cp-ribosome SSU.....	94
Figure 4.15	Position of 16S rRNA bases involved in decoding and comparison between plastid pY and bacterial homolog.....	95
Figure 4.16	Intersubunit bridges in 70S cp-ribosome.....	97
Figure 4.17	The densities for intersubunit bridges in 70S cp-ribosome.....	99
Figure 5.1	Comparison between cp-ribosome SSU maps obtained in current study and by Bieri <i>et al</i>	102

List of tables

Table 3.1	Refinement and model statistics for the cp-ribosome LSU.....	52
Table 3.2	Constituent proteins of the cp-ribosome LSU.....	55
Table 3.3	Protein extensions in the spinach cp-ribosome LSU.....	57
Table 3.4	Data collection and refinement statistics.....	75
Table 3.5	Constituent elements of the cp-ribosome SSU and plastid pY	80
Table 3.6	Components of the intersubunit bridges in spinach 70S cp- ribosome.....	97

Abbreviations

2D	Two-dimensional
3D	Three-dimensional
3DEM	Three-dimensional electron microscopy
Å	Angström
Anti-SD	Anti-Shine-Dalgarno
ATP	Adenosine triphosphate
B	Bridge
βME	β-mercaptoethanol
CCD	Charge-coupled device
CMOS	Complementary metal-oxide semiconductor
Cp-ribosome	Chloroplast ribosome
CRW	Comparative ribonucleic acid website
Cryo-EM	Cryo-electron microscopy
CTD	C-terminal domain
CTF	Contrast transfer function
DC	Decoding center
DED	Direct electron detector
DNA	Deoxy ribonucleic acid
DQE	Detective quantum efficiency
<i>E. coli</i>	<i>Escherichia coli</i>
EF	Elongation factor
EM	Electron microscope
FEG	Field-emission gun
fMet-tRNA	N-formylmethionine transfer ribonucleic acid
FRET	Fluorescence resonance energy transfer
FRM2D	Two-dimensional fast rotational matching method
FSC	Fourier shell correlation
GAC	GTPase-associated center
GTP	Guanosine triphosphate
<i>H. marismortui</i>	<i>Haloarcula marismortui</i>
HPF	Hibernation promoting factor

IF	Initiation factor
IRES	Internal ribosome entry site
I-TASSER	Iterative threading assembly refinement
kV	Kilovolt
<i>L. lactis</i>	<i>Lactococcus lactis</i>
<i>L. monocytogenes</i>	<i>Listeria monocytogenes</i>
<i>L. paracasei</i>	<i>Lactobacillus paracasei</i>
LSU	Large subunit
MDa	Mega Dalton
Mitoribosome	Mitochondrial ribosome
ML	Maximum Likelihood
mRNA	Messenger ribonucleic acid
MSA	Multivariate statistical analysis
NADH	Nicotinamide adenine dinucleotide (reduced)
NMR	Nuclear magnetic resonance
NTD	N-terminal domain
OD	Optical density
PDB	Protein data bank
PRP	Plastid ribosomal protein
PSRP	Plastid-specific ribosomal protein
PTC	Peptidyl transferase center
RCSB	Research Collaboratory for Structural Bioinformatics
RCT	Random conical tilt
RELION	Regularised likelihood optimisation
RF	Release factor
RMF	Ribosome modulation factor
RNA	Ribonucleic acid
RPC	Resampling to polar coordinate method
RPM	Revolutions per minute
r-protein	Ribosomal protein
RRM	Ribonucleic acid recognition motif
rRNA	Ribosomal ribonucleic acid
S	Svedberg's unit

<i>S. aureus</i>	<i>Staphylococcus aureus</i>
SCF	Self-correction function method
SD	Shine-Dalgarno
SEM	Scanning electron microscope
SNR	Signal-to-noise ratio
SPA	Single particle analysis
SRL	Sarcin-Ricin-Loop
SRP	Signal recognition particle
SSU	Small subunit
<i>T. thermophilus</i>	<i>Thermus thermophilus</i>
TEM	Transmission electron microscope
tRNA	Transfer ribonucleic acid
UTR	Untranslated region

Abstract

Ribosomes are the universal machines which perform mRNA translation inside the cell. Chloroplasts harbor bacterial-like 70S ribosome but have acquired five plastid-specific ribosomal proteins (PSRPs) namely cL37, cL38, cS22, cS23 and bTHXc, extensions in the plastid ribosomal proteins (PRPs) and rRNA expansions and deletions. Here, a complete structure of the 70S spinach chloroplast ribosome (cp-ribosome) has been determined using cryo-electron microscopy and single particle analysis. The structures of the large (LSU) and small subunits (SSU) are resolved to 3.3 and 3.7 Å, respectively, demonstrating that the overall architecture of cp-ribosome is similar to the bacterial counterpart. The localization and modeling of the PSRPs, PRP-extensions and remodeled rRNAs suggest about the probable functions of these altered elements in the contexts of cp-ribosome structure and translation regulation inside chloroplasts. The mode of interaction of the hibernation factor, plastid pY, which is present at the subunit interface is also discussed.

1. Introduction

1.1 Central dogma of Molecular Biology

A cell is the smallest entity of life. It undergoes metabolism, homeostasis, growth, response, reproduction and many other properties of life. Reproduction helps in passing on the genetic information to the offspring. This genetic information makes the living beings who they are. Central to the conservation of the genetic information, lies the gene which carries this information and is therefore regarded as the unit of heredity. Genes are composed of deoxyribonucleic acids (DNA), a polymer of four nucleotide bases, namely Adenine (A), Thymine (T), Guanine (G) and Cytosine (C). DNA forms a double helix due to the presence of hydrogen bonds between the constituent nucleotide bases. These hydrogen bonds render high stability in their structure under different environmental conditions. DNA gets transcribed to generate ribonucleic acids (RNA), a polymer of Adenine (A), Uracil (U), Guanine (G) and Cytosine (C). Among many types of RNA polymers in the cell, messenger RNA (mRNA) is the one which is responsible for bearing a copy of the genetic information stored in DNA. mRNA itself gets translated to a protein molecule, a polymer of amino acids. Proteins are the major entities which perform all the functions necessary to sustain life. Transfer of information from DNA to RNA is done via a process known as transcription and involves a protein (enzyme) complex known as RNA polymerase, which acts along with other protein factors. The process to generate protein from mRNA, the so-called translation, is mediated by the ribosome, a macromolecular machine constituting many proteins and RNAs. The transfer of information from DNA to RNA and thereafter from RNA to protein is popularly known as the central dogma of molecular biology. Generally speaking, therefore, the flow of the genetic information follows a common path in all lifeforms on the earth. This points toward a shared origin for all the organisms dating back to ~ 3.5 billion years (4). Since the diversion of life from the Last Universal Ancestor (LUA) or Last Universal Common Ancestor (LUCA) (5,6), organisms have evolved to create three domains of life known as the Bacteria, Archaea and Eukarya. Reminiscent of their common origin, the primary sequence, structure and function of many macromolecular machines are similar, although, often the organisms themselves might have very different morphological traits.

1.2 The ribosome

In all living organisms, the translation of the message encoded in the mRNA to protein is mediated by a ribonucleoprotein complex known as the ribosome. Ribosomes are made up of two subunits, termed as the so-called large (LSU) and small (SSU) subunits. The size of a ribosomal subunit is traditionally measured using the sedimentation coefficient value and accordingly, the prokaryotic ribosomes are termed as the 70S while the eukaryotic ones are the 80S. The bacterial 70S ribosome, which is an example of both a prokaryotic and the most widely studied ribosome, consists of a 50S LSU and 30S SSU. The eukaryotic 80S ribosome, for e.g., the mammalian cellular ribosome, consists of a 60S LSU and 40S SSU. The increased mass of the eukaryotic ribosome is the result of increased protein and rRNA masses due to ribosomal protein (r-proteins) extensions, eukaryote-specific r-proteins and ribosomal RNA (rRNA) expansion segments. The molecular weights of the prokaryotic and eukaryotic ribosomes are around 2.5 and 4.5 MDa while the protein: rRNA ratios are 1:2 and 1:1, respectively. A general comparison between the compositions of prokaryotic and eukaryotic ribosomes are depicted in Figure 1.1.

Figure 1.1 The components of prokaryotic and eukaryotic ribosomes. Both are composed of two subunits. In comparison, eukaryotic ribosome is bigger than prokaryotic counterpart. One extra rRNA component (5.8S) is also present in eukaryotic ribosome LSU. The figure is taken from Pearson Education, Inc. 2008.

1.3 Structural studies on the ribosome

The early reports about the morphological structure of the ribosome were generated by using shadow casting with platinum and also studying negatively stained samples (7) (8) (9) (10) (11) (12) (13,14). Thereafter, the ribosomes were captured at various translational states as reported by cryo-EM and single particle analysis, detailing some of their structural features (15-18). However, due to low resolution, these structures

could not provide an atomic-level understanding of the ribosome and the translational process. Gradually, with the advent of X-ray crystallographic atomic models of the SSU and LSU from the laboratories of Yonath (19), Ramakrishnan (20) (*T. thermophilus* ribosomal SSU at 3.3 and 3 Å resolution, respectively) and Steitz (21,22) (*H. marismortui* ribosomal LSU at 2.4 Å resolution), an atomic-level understanding started to emerge about the ribosome. An atomic model for the complete 70S ribosome was still absent and the interpretation for a complete ribosome model relied on fitting the individual subunit models into the 5.5 Å map obtained from *T. thermophilus* by the Noller group (23). The complete 70S structure for the empty *E. coli* ribosome was solved at 3.5 Å in 2005 by Cate lab (24) and later a 2.8 Å structure of the *T. thermophilus* ribosome with bound mRNA and tRNA ligands was solved by Ramakrishnan lab (25) in 2006. What followed is the elucidation of the conformational dynamics of the bacterial ribosome at different translational states and in presence of antibiotics, translation factors, various RNAs (mRNA, tRNA and IRES) and protein translocation factors by cryo-EM and X-ray crystallography which is reviewed in details by various authors (26-30). Since then, high-resolution ribosome structures have been elucidated from multiple sources including model eukaryotes (31,32), mitochondria (33-35), pathogenic bacteria (36-39) and pathogenic parasitic eukaryotes (40-44), providing unparalleled insights into the conservation and diversity of ribosome structures.

It is noteworthy to mention that, apart from providing a detailed understanding of the translation process, these above-mentioned structural reports have increased the human knowledge base regarding RNA folding and RNA-protein interactions. Acknowledging the importance of the structural studies on the ribosome, three scientists, Ada Yonath, Thomas Steitz and Venkatraman Ramakrishnan were awarded the Nobel Prize in Chemistry in 2009.

1.4 Functionally important sites of the ribosome

The ribosome is a macromolecular machine specialized for synthesizing proteins inside the cell across all living organisms. Some of its important functional sites include the decoding center, peptidyl transferase center (PTC), polypeptide tunnel, tunnel exit site and translation factor binding site. The decoding center, located in the SSU, is responsible for recognizing the correct codon-anticodon pair between mRNA and

tRNA. Once the correct reading frame is established, a peptide bond is formed at the PTC located in the LSU, between the amino acids present at the A and P sites. The growing polypeptide chain then traverses along the polypeptide tunnel located in the LSU and emerges out of the ribosome at the tunnel exit site. This whole process of translation requires transient binding of many protein factors which bind at the translation factor binding site situated near the L7/L12 stalk of the LSU. A detailed discussion on the composition of these sites is presented below.

1.4.1 Decoding center (DC)

The DC is located at the interface of head, beak, platform and shoulder of the SSU (Figure 1.2A, B). It ensures translation of the correct reading frame, the precise mechanism of which remains unknown. Ramakrishnan and coworkers have done pioneering work and shed lights on the decoding mechanism (45). During protein synthesis, EF-Tu delivers an amino acid-charged tRNA to the aminoacyl tRNA-binding site (A-site) on the ribosome.

Figure 1.2 The decoding center (DC) on the SSU. (A) An overview of the SSU showing position of the DC as marked by a black box. Constituent elements are labelled as indicated. Two red arrows show the domain closure movement of the head and shoulder upon decoding. (B) Zoomed in view of the DC to show positions of important 16S rRNA bases G530, A1492 and A1493 near to the DC. (C) Codon-anticodon pairing to display the wobble position where non-Watson-Crick pairing is accepted. The figures are modified from (45).

If the anticodon present on the tRNA is recognized as the correct (cognate) pair for the codon on the mRNA, it leads to a conformational change in the SSU. This is known as the domain closure movement wherein the SSU head and shoulder rotate and move toward the LSU (Figure 1.2A). It is generally accepted that the Watson-Crick geometry between the two base pairs at the first two codon-anticodon (Figure 1.2C) is the reason behind domain closure as non-Watson-Crick base pairs are also capable of forming stable hydrogen bonds. Consequently, or simultaneously with the domain closure, three universally conserved bases in the 16S rRNA namely G530, A1492 and A1493 (Figure 1.2B) change their orientation in a way such that they stabilize the codon-anticodon interaction. Involvement of these three nucleotides has been confirmed by footprinting studies which demonstrated that they remained protected from chemical modifications upon tRNA binding to the A-site (46-48). Domain closure structural rearrangements of the SSU had earlier been achieved by the antibiotic paromomycin in the presence of near-cognate tRNAs (49) emphasizing that, Watson-Crick geometry rather than the count and position of hydrogen bonding are at the heart of the decoding process. Importance of this geometrical mimicry is corroborated by the results obtained by Demeshkina and coworkers who showed that a G.U base pair at the first or second codon-anticodon position can also adopt a G.C like geometry and promote erroneous incorporation of amino acids in the polypeptide chain (50,51). These results further suggested a rather static role of G530, A1492 and A1493 in decoding, contradicting earlier interpretations. A good place to start pondering over the intricacies of decoding would be the articles by Loveland *et al.* (52), Rodnina *et al.* (53) and Schrode *et al.* (54).

1.4.2 Peptidyl transferase center (PTC)

The PTC is the catalytic center for the peptide bond formation between amino acids and is located at the center of the LSU (Figure 1.3A). The PTC forming nucleotides, which are well conserved across three kingdoms of life, are from domain V of the 23S rRNA. The PTC wall is constituted by the A- and P-loops (Figure 1.3B) of the 23S rRNA and the floor of PTC faces the polypeptide exit tunnel at the bottom (Figure 1.3C).

Figure 1.3 The peptidyl transferase center (PTC) and formation of peptide bond. (A) The location of PTC in the LSU is indicated along with A, P and E-site tRNAs. (B) Secondary structure of rRNA from the PTC forming region. Numberings of the bases are based on *E. coli* ribosome. (C) PTC is connected to the polypeptide exit tunnel through which nascent polypeptide traverses while being synthesized at the PTC. (D) Schematic representation of the peptidyl transfer reaction between A and P-site tRNAs. The figures are taken and modified from (55), (56), (57) and (58).

In each cycle of polypeptide elongation, the α -amino group of aminoacyl-tRNA at the A-site performs a nucleophilic attack on the carbonyl carbon of the peptidyl-tRNA residing at the P-site (Figure 1.3D), producing a tetrahedral transition state product, which thereafter decomposes to produce peptidyl-tRNA with an extra amino acid at the A-site. This results in deacylated tRNA at the P-site. A comprehensive review of the related literature on the process has been undertaken by Simonovic *et al.* (55). It is noteworthy to mention here that the PTC is solely formed by RNA but two ribosomal proteins namely L27 (25,59,60) and L16 (61,62) have been implicated to play a role in stabilizing the CCA ends of A- and P-site tRNAs in prokaryotes and eukaryotes, respectively, facilitating peptidyl transfer reaction. Given the high importance of its function, it is not surprising that PTC is mostly conserved across the three kingdoms of life. However, some minute differences emerge after careful comparison (reviewed in (63)). Some of the antibiotics, as reported in bacteria (64), mediate their action by binding to the PTC, inhibiting translation.

1.4.3 The polypeptide exit tunnel and tunnel exit site

The polypeptide exit tunnel is a conduit through which the nascent polypeptide chain travels from the PTC to the solvent side of the LSU (Figure 1.4A). The existence of the exit tunnel was first proposed around 50 years back (65,66) and later shown structurally (67,68). Cryo-EM reconstructions and X-ray crystallographic studies have described the structure of the tunnel at increasing resolution (21,69-71). The length and width of the tunnel are about 80-100 and 10-20 Å, respectively (Figure 1.4A). 23S rRNA is the major contributing element in forming the tunnel, however, r-proteins L4, L22 (eukaryotic homolog is L17) and L23 are also responsible to form the tunnel wall and constriction (21,22,72) (Figure 1.4B).

Figure 1.4 The polypeptide exit tunnel in the LSU. (A) The exit tunnel connects the PTC to the solvent side of LSU. The exit tunnel is zoomed in to show the nascent polypeptide chain, which upon its release into the solvent side is processed by downstream factors to perform protein folding and targeting. (B) The composition of the tunnel wall in prokaryotic and eukaryotic ribosomes. The figures are modified from (61), (73).

Such composition of the tunnel is surprising as proteins are mainly responsible to form pores, tunnels and channels to transport molecules in and across the cell and cell membrane (74). The tunnel was initially thought to play a passive role in protein folding and biogenesis but recent studies have argued against this hypothesis. It has been shown (75,76) and reviewed (77,78) that protein starts to fold while in the tunnel.

The tunnel exit site (Figure 1.5) is the opening of the polypeptide tunnel toward the solvent side of the LSU. The rim of the opening is formed by four conserved

proteins, L22, L23, L24 and L29 and other organism-specific proteins. Therefore, unlike the polypeptide tunnel, the tunnel exit site is chiefly composed of r-proteins. It acts as a docking platform for many protein biogenesis factors which interact with the nascent polypeptide chain as it comes out of the tunnel (72).

Figure 1.5 The tunnel exit site located at the bottom of LSU. (A) The arrangement of the tunnel exit site forming proteins L22, L23, L24 and L29 surrounding the tunnel exit site are shown as colored surfaces. The exit site is indicated by a white arrow. (B) Interaction of Trigger factor (TF), shown in purple surface, at the tunnel exit site. TF and other protein translocation factors dock at the tunnel exit site to mediate their action. The figure is taken from (79).

1.4.4 The mRNA entry and exit sites

The genetic code is translated when the SSU binds the mRNA and a 70S ribosome is formed after binding of the LSU. The mRNA channel, which is located at the subunit interface side of the SSU and surrounded by the SSU head, beak, platform and shoulder, ideally positions the mRNA to ensure decoding fidelity. The mRNA has been visualized bound to the ribosome in many structures (25,80,81). Despite significant efforts (82-85), the movement of mRNA on the SSU during translation is still a matter of debate. The mRNA channel has two openings, the mRNA entry and exit sites (Figure 1.6), to reach the solvent side of the SSU. During translation, the mRNA wraps around the SSU in a way that its 3' and 5' ends are located at the mRNA entry and exit sites, respectively.

Figure 1.6 The mRNA entry and exit sites on the SSU. (A) Location of mRNA entry site on SSU is indicated. (B) Figure shown in (A) is rotated to indicate the location of mRNA exit site. In both (A) and (B), the constituent proteins are colored and labeled as indicated. 16S rRNA is displayed in grey atoms. *E. coli* ribosome high-resolution model (PDB ID 4YBB) (86) was used to prepare the figure.

The mRNA entry site (Figure 1.6A) is formed by three SSU r-proteins- S3, S4 and S5. Among these, S3 and S4 have been reported to be responsible for helicase activity in the ribosome and S5 is implicated in translation fidelity. The mRNA exit site (Figure 1.6B) is located near to the platform region of the SSU and is encircled by r-proteins S7, S11, S18 and S21 (absent in *T. thermophilus*) in the *E. coli* ribosome. Here, the 5'-untranslated region (UTR) of the mRNA interacts with the 3' end of 16S rRNA to form the so-called Shine-Dalgarno (SD):anti-SD complex (87) in bacteria. It has been proposed that a beta-hairpin loop of S7 prevents codon-anticodon interaction at the E-site (88,89). Also, the S7-S11 interaction has been implicated in the translational fidelity (90). S11 and S18 have been shown to interact with the SD-duplex and 5'-end of mRNA (83,84).

1.4.5 Translation factor binding site

During translation, many factors including some GTPases (for e.g., GTPases like IF2, EF-Tu, EF-G, RF3 and non-GTPases like RF1 and RF2 in bacteria) bind the ribosome transiently and mediate their function (81,91-96). These factors share a common binding site on the ribosome which is popularly known as the GTPase-associated center (GAC) (Figure 1.7). However, as clear from the examples of proteins that bind to this site, it can be considered as the canonical factor binding site on the ribosome. It is located close to the L7/L12 stalk base (H43 and H44) of the LSU and contacts r-proteins

and rRNAs from both the subunits. The Sarcin-Ricin-Loop (SRL) of the highly conserved H95 and r-protein L11 from the LSU, and the four rRNA helices- h5, h8, h14 and h15 and r-protein S12 in the SSU are the elements which contact the GAC.

Figure 1.7 The GTPase-associated Center (GAC) of the ribosome. (A) Binding mode of EF-Tu on the ribosome. Elements associated with the GAC are labeled as indicated. Unlabeled SSU and LSU proteins are in tan and sky blue colored ribbons, respectively. The figure is prepared using the model bearing PDB ID 4V5G (92). (B) Binding mode of EF-G on the ribosome displaying a common binding site (GAC) for the GTPases on the ribosome. Modified from (97).

1.5 Mechanism of translation in bacteria

Translation in bacteria is a complex process and involves multiple steps like initiation, elongation, termination and recycling (Figure 1.8). Initiation occurs when the SSU binds the mRNA, initiation factors (IF1, IF2 and IF3) and initiator tRNA (fMet-tRNA^{fMet}). A preinitiation complex is formed when fMet-tRNA^{fMet} at the P-site contacts the mRNA codon. Then, with the dissociation of the initiation factors and binding of the LSU, an initiation complex is established. To add the next and other amino acids thereafter during the elongation phase, aminoacyl-tRNA enters the A-site of ribosome to form a ternary complex between the tRNA, EF-Tu, GTP and SSU. Once correct codon-anticodon pair is recognized, GTP hydrolysis and subsequent accommodation of the aminoacyl end of the A-site tRNA occur into the PTC. Next, EF-Tu is released from the ribosome and a peptide bond is formed between the amino acids loaded onto the A- and P-site tRNAs. This also shifts the peptide onto the A-site tRNA and creates a deacylated tRNA at the P-site. Then, GTP hydrolysis by EF-G translocates the SSU on the mRNA to position the next codon into the empty A-site and consequently the deacylated tRNA at the P-site moves to the E-site. After this, EF-G falls off the ribosome and the next cycle of elongation begins. If during the process of elongation,

a stop codon is encountered, RF1 or RF2 starts the termination of translation by releasing the peptide. RF1 and RF2 are then released by RF3. Soon, ribosome enters the recycling phase where RRF and EF-G jointly act to dissociate the subunits of the ribosome. The bound mRNA and tRNA on the SSU are then removed by IF3 and this step concludes the translation cycle.

Figure 1.8 The steps in bacterial translation. The steps followed during bacterial translation- initiation, elongation, termination and recycling are shown in the order indicated by the arrows. The figure is taken from (98).

1.6 Hibernating ribosomes

During stressed conditions, cells can downregulate translation by factor-binding on the ribosome to render the complete ribosome intact but inactive. Such ribosomes are called hibernating ribosomes and structurally they are not so well-characterized to date. New knowledge is rapidly emerging from both single-particle and tomographic cryo-EM studies in bacteria which have captured such ribosomes in different states of rotation and monomeric/dimeric forms (99-104). As evident from these studies, four proteins namely ribosome modulation factor (RMF), YfiA (aka protein Y, pY, ribosome-associated inhibitor A or RaiA), hibernation promoting factor short variant (HPF^{short}) and hibernation promoting factor long variant (HPF^{long}) have been attributed in producing the hibernating state of ribosomes. Except for RMF, the other three proteins are related via homology as shown in Figure 1.9A, B.

Figure 1.9 The hibernation promoting factor (HPF). (A) Schematic representation of the domain organization among HPF homologs showing that N-terminal domain (sky blue) is conserved. (B) Phylogenetic tree built from primary sequences of HPF homologs. Please refer to Appendix Figure 8 for an enlarged view of (B). (A) and (B) are taken from (105). (C) Binding site of HPF N-terminal conserved domain at mRNA channel of 30S SSU that sterically clashes and excludes binding of A- and P-site tRNAs, as reported in *S. aureus* (106).

What appears from the studies in γ -proteobacteria *E. coli*, two different mechanisms could be followed to produce hibernating ribosomes. In first, the RMF and HPF^{short} bind sequentially to the ribosome to create a hibernating 100S ribosome (dimer of the 70S) (103). The second mechanism involves prevention of 100S formation and rather producing an inactive 70S due to binding of YfiA, a paralog of HPF^{short}. In plant plastids and most other bacteria except γ -proteobacteria, a homolog of HPF^{short} and YfiA is encoded that is almost double the size of HPF^{short} and is named HPF^{long}. HPF^{long} by itself can induce hibernation by mediating 100S ribosome formation as reported in *T. thermophilus*, *L. lactis*, *L. paracasei*, *L. monocytogenes* and *S. aureus* (103,105,107-110). The binding site of HPF^{long} on the SSU is shown in Figure 1.9C.

1.7 Organellar ribosomes

Origins of both mitochondria and chloroplast can be traced back to prokaryotic ancestors (111). They both contain their own genome and machineries to decode the genetic messages into mRNAs and thereafter synthesize proteins. It appears that the evolution of these two organelles have followed a different route which is reflected in the differential composition of their ribosomes. Cryo-EM structures were solved for mitoribosomes from yeast (33,34) and mammals (34,35,112) and to solve cp-ribosome structures, spinach (3,113-116) and green alga (117) sources were used. These studies reveal unique features of these ribosomes. In general, the cp-ribosome has maintained significant similarity with the bacterial ribosome while the mitoribosome has diverged substantially (118). Despite the similarity with the bacterial counterpart, novel features have evolved in the cp-ribosome. For example, cp-ribosomes have acquired five plastid-specific ribosomal-proteins (PSRPs) and extensions in many plastid ribosomal proteins (PRPs) while the rRNA content remains like bacteria (119,120). Due to this, the RNA: protein ratio is altered to 3:2 in the cp-ribosome, compared to the bacterial 2:1 (118). On the other hand, mitoribosomes possess significantly reduced rRNA and a lot many mitoribosome-specific proteins. The RNA: protein ratio in the mitoribosome is 1:2 (118). Mitoribosomes are larger in size possessing a diameter of ~ 320 Å compared to the cp- and bacterial ribosomes whose diameters are ~ 265 and ~ 260 Å, respectively. Despite the larger size, mitoribosomes are lighter due to its porous nature and sediment at 55S. Bacterial and cp-ribosomes sediment at 70S. Compared to the mitoribosome which synthesizes only membrane proteins (112,121), the cp-ribosome can synthesize both soluble and membrane proteins (119,122).

1.8 Structure of the cp-ribosome

The first report of the cp-ribosome structure originated from immunoelectron microscopy studies using negatively stained pea cp-ribosome (123). Then, in 2007, two reports were published using the spinach (124) and green algal (117) samples by cryo-EM at 9.4 and 15.5 Å resolution, respectively. These studies have enabled direct visualization of the cp-ribosome and proved that they are similar to the bacterial counterpart given the overall architecture. Recently, high-resolution cryo-EM structures of the spinach cp-ribosome have been solved by a number of labs (3,113,115,116). The results obtained from these studies, in general, agree well with each other and also support the ribosomal components identified by mass spectrometric

studies reported earlier (120,125). Due to the high-resolution obtained in these studies, it was possible for the first time to unambiguously identify the locations and build models for the PSRPs and PRP-extensions.

1.9 Chloroplastic translation and its regulation

As per the endosymbiotic theory (111), chloroplasts have a cyanobacterial origin. Over the course of evolution, the chloroplastic genes were transferred and integrated into the nuclear genome of the plant, leaving behind a reduced, double-stranded, circular genome (plastome) inside the organelle (126). Around 146 genes are harbored in spinach plastome among which ~98 are protein-coding (124,126), synthesizing ~5% (127) of the total proteins that function inside the chloroplast. As clear from this statistic, a vast majority of the proteins that are implicated in some functions inside the chloroplast are not encoded by the plastome. These proteins are encoded by the nuclear genome, and after getting synthesized by the 80S ribosomes, they are post-translationally imported into the chloroplast. The import of the so-called preprotein into the chloroplast stroma is made possible due to the presence of N-terminal signal/transit peptide which is thereafter chopped off, producing a shorter so-called mature protein (128). It is interesting to note that, some of the proteins possess dual-targeting signals for the chloroplast and mitochondria (129,130). Some of the protein complexes that are implicated in chloroplastic functions are constituted of subunits encoded by both the nuclear and chloroplastic genomes. A tight regulation between the nuclear and chloroplastic gene expression is therefore observed in plants, which ensures correct stoichiometry of the subunits in the protein complexes. A few examples of the protein complex whose subunits are encoded by both the genomes include complexes involved in photosynthesis and respiration like the photosystem -1 and -2, ATP synthase, RuBisCO and cytochrome *b6/f*. On the other hand, RNA polymerases, NADH dehydrogenases, r-proteins and factors involved in translation constitute examples of the proteins which are entirely encoded by the plastome. It is of interest that both soluble and membrane proteins are synthesized by the cp-ribosome, unlike the mitoribosome which is specialized to translate only membrane proteins.

Regulation of the chloroplastic gene expression happens primarily after transcription (131-134). During this stage, mRNA is either physically modified via editing, splicing and cleavage, or its stability is affected (131,135-138). The regulation

reaches its peak during translation, especially during initiation and elongation stages (134,139,140). This is opposed to what is observed in bacteria where due to the coupled nature of transcription and translation, the regulation of gene expression must occur at the transcriptional level. However, some translational regulation (translation attenuation) does happen in bacteria and is mediated by either masking of mRNA sequence by proteins or secondary structure formation in mRNA. In comparison, multiple chloroplastic mRNAs experience positive translation regulation in which important roles are played by *trans*-acting nuclear-encoded protein factors, as identified by the biochemical and genetic studies (139,141-146). These mRNAs are not translated in the absence of these factors. Thus, these *trans*-acting protein factors are able to mediate communication between the nucleus and chloroplast. Interestingly, these factors attach either to the 5'- or 3'-UTR of the chloroplastic mRNA, however, the former is more frequent than the latter. An important player in the translation regulation in plants is light, which mainly acts by changing the redox states of the components involved in the photosynthetic electron transport chain. This is only expected that light is a major regulatory element controlling translation, as photosynthesis is the main function of the chloroplast. One example of such light-induced regulation is the upregulation of the regulatory proteins that bind 5'-UTR of the *psbA* gene in response to light exposure (147). Apart from the *trans*-acting protein factors, *cis*-acting RNA elements located at the 5'-UTR of the chloroplastic mRNA also play a role in translation regulation. In bacteria, 5'-UTR mediates SD:anti-SD interactions with 16S rRNA and the positions of SD sequences on bacterial mRNAs are conserved (~7 bases upstream from translation start site) (87). Although anti-SD sequences in the chloroplastic 16S rRNA show conservation compared to bacteria (148), not all chloroplastic mRNAs contain SD-like sequences, nor are their positions conserved, which may even be situated as far as 100 nt upstream (140,149-151). This varied positioning and the fact that mutations in putative SD sequences were tolerated, moderately affecting or totally unaffacting translation, raise the question about the functional significance of the putative SD sequences (152-154). However, in some reports, SD sequences have been shown to play an essential role in translation (151,155-157). Therefore, the literature is diverse and controversial, and a complete picture about the role of SD:anti-SD in the chloroplastic translation is yet to emerge.

Another aspect of the chloroplastic translation system is the conservation of the translation factors compared to the bacterial system which is reminiscent of its prokaryotic origin (158).

1.10 Cryo-EM as the method of choice

The ribosome is a dynamic macromolecular machine. As shown for the first time by a Fluorescence resonance energy transfer (FRET) study (159), the SSU undergoes a rapid ratchet-like motion while it is situated on top of the LSU, even in the absence of EF-G and GTP which were initially considered as mandatory to drive such movements (160). This inherent motion in the ribosome creates a bottleneck for studying them using X-ray crystallography, as for successful crystal growth a minimal conformational flexibility is desirable. However, crystals were obtained for ribosomes from bacterial and human sources after employing significant efforts. No such standardized protocol is available for cp-ribosome crystallization. Another aspect is the large size of the ribosome which makes it beyond the reach of a Nuclear magnetic resonance (NMR) study. On the other hand, cryo-EM has proven itself as the technique where “bigger is better”. In cryo-EM, thousands of particle-images containing very low signal-to-noise ratio (SNR) are computationally aligned and thereafter averaged to produce sufficient high-contrast for 3D reconstruction. Relatively higher SNR obtained for bigger particles like ribosomes are better suited for cryo-EM reconstructions as compared to the lower SNR from smaller proteins. On top of that, adoption of maximum likelihood (ML) -based methods for 3D classification during the cryo-EM data processing as popularized by the software RELION (2,161) and FREALIGN (162,163), it has been possible to computationally sort conformational heterogeneity in a cryo-EM dataset. This is unlike a ribosome crystal which oftentimes provides information about only one functional state. Such sorting possibilities using computation also eliminates the necessity to begin with a highly pure sample. Therefore, considering the limitations of other techniques and benefits of using cryo-EM to solve cp-ribosome structure at high-resolution, cryo-EM and single particle analysis was adopted as the method of choice.

1.11 Aim (outline of the thesis)

The aim of the work is to reconstruct a high-resolution 3D model for the 70S cp-ribosome from spinach (*Spinacia oleracea*) using cryo-EM and single particle analysis. The model will shed light on the unique properties of the ribosome through analysis of the overall structures of the SSU and LSU along with the high-resolution details of functionally important sites like the polypeptide tunnel, mRNA entry and exit sites. Furthermore, localization and modeling of the PRP-extensions, PSRPs and rRNA expansions will provide a comprehensive picture of these ribosomes which previously eluded the scientific community. Chloroplastic translation is more complicated than the bacterial counterpart. When this work was initiated, very little was known about the chloroplastic ribosome and translation in structural terms. In light of the debate whether SD-antiSD interaction is necessary for the translation initiation inside the chloroplast, our data provide some insights into the matter. While this work was ongoing, two reports of the spinach cp-ribosome were published by Graf *et al.* (115) and Bieri *et al.* (114). A brief analysis of the data from their labs and the data obtained in the current study will be presented in the discussion section of the thesis.

2. Materials and Methods

2.1 Materials

2.1.1 Chemicals

All the chemicals used for the study were at least of analytical grade. Chemicals were purchased from the following companies:

Names	Suppliers
Tris hydrochloride	Promega
β -mercaptoethanol (β ME)	Sigma
Glycerol	Affymetrix
D-sorbitol	VWR
Magnesium Acetate	Sigma
Potassium Chloride	Sigma
Sucrose	Sigma
RNAse inhibitor	Promega
Protease inhibitor	Abcam
Triton X-100	Sigma

2.1.2 Buffer compositions

Names	Specifications
Buffer A	10 mM Tris HCl pH 7.5, 50 mM KCl, 10 mM Mg. Acetate, 7 mM β ME
Homogenization buffer	0.7 M Sorbitol in buffer A
Resuspension buffer	0.4 M Sorbitol in buffer A
Lysis buffer	2% (v/v) Triton X-100 in buffer A
Sucrose cushion buffer	0.75M or 1M Sucrose in buffer A, depending on experimental step
Buffer B	10% (v/v) Glycerol in buffer A
Sucrose gradient buffer	20 mM Tris HCl pH 7.5, 100 mM KCl, 10 mM Mg. Acetate, 10% or 40% (w/v) Sucrose, 7 mM β ME, 1 unit/mL RNasin (RNAase inhibitor)
Grid buffer/cp-ribosome suspension buffer	20 mM Tris HCl, 100 mM KCl, 10 mM Mg. Acetate, 7 mM β ME, 100 mM Sucrose, 1 unit/mL RNasin (RNAase inhibitor), 0.1% (w/v) Protease inhibitor

2.1.3 Other instrumentation and materials

Kitchen Blender	Philips
Nanodrop spectrophotometer	Thermofisher Scientific
Sucrose gradient maker	Biocomp Instruments
Ultracentrifuges	Beckman Coulter
Benchtop centrifuges	Thermofisher Scientific
Vortex genie2	Scientific Industries, Inc.
SW 41 Ti rotor	Beckman Coulter
TLA 110 rotor	Beckman Coulter
Cryo-TEM microscope	FEI Tecnai Arctica (200-kV), Nanyang Technological University, Singapore. FEI Titan Krios (300-kV), National University of Singapore, Singapore.
Cressington 208 high vacuum carbon coater	Cressington Scientific
Cheesecloth	Lazada Singapore
Miracloth	Calbiochem

2.1.4 Computer tools and software

Names	References
EMAN 2.1	Bell <i>et al.</i> , 2016 (1)
RELION 1.4	Scheres, S. H. (2)
ModeRNA server	Rother <i>et al.</i> , (164)
PDB to 3D Restraints server	http://rna.ucsc.edu/pdbrestraints/pdbtorestraints.html
PyMOL	Schrödinger, L.L.C. (165)
UCSF Chimera	Petersen <i>et al.</i> , (166)
COOT	Emsley <i>et al.</i> , 2010 (167)
MolProbity	Chen <i>et al.</i> , 2010 (168)
Phenix	Afonine <i>et al.</i> , (169)
I-TASSER server	Zhang, Y. (170)
Mapman	Kleywegt and Jones (171)

2.2 Methods

2.2.1 Biochemical method- Purification of cp-ribosomes from spinach leaves

To purify cp-ribosome, the protocol from Bartsch *et al.* was followed with some minor modifications. 6 kg of baby spinach (*Spinacia oleracea*) was bought from the supermarket on the day of purification. The leaves were deveined; damaged and infected leaves were discarded and washed thoroughly with distilled water. All steps thereafter were carried at 4° C. The leaves were homogenized in homogenization buffer (2 L per 1 kg of leaves) using a kitchen blender (2 shots of 10 sec each) and subsequently filtered through several layers of cheesecloth and one layer of miracloth. After the debris was discarded, the resultant filtrate was centrifuged at $1200 \times g$ for 15 min. The pelleted crude chloroplasts were washed and resuspended in resuspension buffer (50 mL per centrifuge bottle), pooled and centrifuged again at $1200 \times g$ for 15 min. The supernatant was discarded without disturbing the pellet. Pellet was completely resuspended in the lysis buffer and incubated on ice for 30 min. Most of the ribosomes in the cell remain bound to membranes and therefore, usage of a detergent is necessary to detach them from the membrane. The lysed suspension was centrifuged at $26000 \times g$ for 30 min to clarify and this step was repeated thrice to get rid of any membrane (which goes to the pellet fraction). Semi-crude ribosomes were pelleted at $86,000 \times g$ for 17 hr at 4° C by overlaying 15 mL of clear suspension on top of 8 mL of sucrose cushion buffer (1M sucrose in buffer A).

The greenish crude ribosomal pellet was resuspended in buffer B with gentle agitation and clarified at $26000 \times g$ for 15 min. This clarification step was repeated twice and the supernatant, containing crude ribosomes was snap frozen and stored at -80° C.

6 mL of 10% (w/v) sucrose gradient buffer was dispensed in a centrifugation tube. 40% (w/v) sucrose gradient buffer was dispensed slowly by immersing the needle till the bottom of the tube. Gradient was made on the gradient station using the preset program: Angle = 81.5, RPM = 17, Duration = 1.48. Tubes were balanced using 10% (w/v) sucrose gradient buffer (either add or remove) and 300 μ L buffer was removed carefully from the top. 200 μ L of crude ribosomes were loaded into each tube and centrifuged (SW 41 Ti rotor) at 30000 RPM for 4 hours at 4° C. After the centrifugation, ribosomes were fractionated on gradient station using following settings: Speed = 0.34, fraction size = 20 drops/ tube. The 70S fractions were pooled together and 2.5 mL of pooled fraction was loaded onto 1 mL of sucrose cushion buffer (+ 1 unit/mL RNasin)

and centrifuged (TLA 110 rotor) for 2.5 hours at 4 °C at 45000 RPM for pelleting the ribosomes. The ribosomal pellets from all the tubes were resuspended in 50 μ L (in each tube) of cp-ribosome suspension buffer or grid buffer. The concentration of the 70S ribosomes as determined by the absorbance measured at 260 nm was 59 OD/mL. The ribosomes were diluted to 25 OD/mL and snap frozen before storing at -80° C. Remaining crude ribosomes were also purified to homogeneity repeating the steps starting from sucrose gradient run.

2.2.2 Biophysical Methods

Cryo-electron microscopy (cryo-EM) and single particle analysis are the principal techniques used in this study. The technical details are provided below.

2.2.2.1 History and introduction

One of the most important advancements in modern science is the ability to see small objects using a microscope. A conventional light microscope typically has the resolution limit of 250 nm. In contrast, an electron microscope (EM), in principle, has the ability to produce images that are resolved below 1 Å (172). The EM comes in two flavors, the scanning electron microscope (SEM) and transmission electron microscope (TEM). In a SEM, the sample is coated with electron-reflecting metal from which electrons bounce off, once scanned with electrons. Therefore, the resulting 3D image contains information related to the surface of the sample. In contrast, in a TEM, the internal information can be captured as the bombarded electrons traverse the entire thickness of the sample. The first TEM was produced by Knoll and Ruska, in 1931 (173). However, imaging biological specimen by electrons did not produce desired results as electrons damaged the specimen due to energy deposition leading to bond breakage. Also, the high vacuum required to operate the microscope caused dehydration leading to the collapse of the biological sample. A workaround to this was found when negative stain was first introduced by Hall and coworkers (174) and later improved by other scientists (175,176). Negative stain is a dried solution of heavy metal salt which form a cast surrounding the specimen. The cast scatters electrons stronger than the specimen (thus producing enough contrast for visualization), is less affected by the radiation and prevents molecule from collapsing. A complete protocol to prepare negatively stained EM grids has been reported by Booth and coworkers (177). Comprehensive reviews about the procedure can be found elsewhere (178,179). The

first successful 3D model from negatively stained 2D projection images were produced by DeRosier and Klug for the tail of bacteriophage T₄ (180). This study kickstarted the field of three-dimensional electron microscopy (3DEM). However, negative staining was (and still is) inapplicable to capture finer details of the sample as it relied on the principle of stain-exclusion by the sample to produce contrast, and therefore, resolution was limited to about 20 Å (181). Despite the drawback, negative stain is still popular as it can provide information regarding conformational changes in a macromolecular complex (182,183), subunit stoichiometry/position (184,185) and complex stability (186). The first major breakthrough in obtaining high-resolution EM images was obtained when Taylor and Glaeser studied frozen-hydrated protein samples at cryogenic temperatures (187). This work popularized the use of cryogenic temperatures for EM imaging (cryo-EM). Later in 1988, Dubochet and coworkers developed the vitrification method for cryo-EM which has been in use till date (188). Technically speaking, in cryo-EM, biological specimen undergoes less radiation damage due to low temperature and also the native structure is preserved due to vitrification (189). In the following years, along with the developments in cryo-EM methodology, several works were published capturing biological specimen at better resolution which established cryo-EM as a potential technique for structural biology. Some of examples include the structures of bacteriorhodopsin (190), plant light-harvesting complex (191), acetylcholine receptor (192), icosahedral viruses (193,194), tubulin (195), amyloid fibril (196) and ribosomes (197). Among these examples, ribosomes were the only samples that did not form any helical symmetry, nor did they form 2D crystals and rather oriented randomly when vitrified on an EM grid. These works carried out during the early phases of cryo-EM developments are reviewed by Helen R. Saibil (189). In 2017, three scientists- Richard Henderson, Jacques Dubochet and Joachim Frank were awarded the Nobel Prize in Chemistry for their contribution in the cryo-EM field.

2.2.2.2 The transmission electron microscope (TEM)

The basic anatomy and working principle of a TEM are comparable to an optical microscope. However, compared to an optical microscope, the magnification and resolution obtained in an EM are greater, allowing it to see much smaller objects in finer details. At the top of the TEM, a source of illumination is placed which bombards electrons down the column kept under high vacuum. These electrons are focused

through a range of electromagnetic lens-system to finally produce a magnified image on a viewing screen or detector kept at the bottom of the microscope.

Figure 2.1 The components of a TEM. (A) A representative image of 120-kV TEM regularly used for viewing and screening negatively stained biological specimen. (B) A schematic view of basic layout of a TEM. Taken from Australian Microscopy and & Microanalysis Research Facility.

2.2.2.3 The gun

The gun (Figure 2.1) is the topmost component of a TEM which produces high-energy electron beam to obtain an image of the specimen. The gun comprises an electron-emitting source, a so-called Wehnelt cylinder and an anode. During the early phases of electron microscopy, thermionic emitters like the tungsten filament or lanthanum hexaboride (LaB_6) crystal were used as the electron source. Once a strong current is passed through the source, it gets superheated ($3000\text{ }^{\circ}\text{C}$) to release some electrons which are accelerated down the column of the microscope due to the electric potential

difference (80-300 kV) maintained between the top and bottom of the microscope. However, to obtain a good image, the electrons need to be collated into a coherent beam before they are accelerated down the column. Recently, an advancement has been achieved to obtain spatial and temporal coherence of the electrons by using a field-emission gun (FEG) as the electron source. In FEG, the tungsten filament is reduced to a small nanometer-sized tip which produces an electron beam with a smaller diameter and greater coherence. The greater coherency achieved with the FEGs has resulted in better signal-to-noise (SNR) ratio, thus improving the final resolution obtained from the images.

2.2.2.4 The lens system

An electromagnetic lens (Figure 2.1) is a coil of wire through which the current flows generating a magnetic field to push the electrons in a downward spiral. A series of electromagnetic lenses along with the deflectors, stigmators and apertures are used in a TEM to ensure that most of the electrons emanating from a point-source are converged and also that the parallel electron rays are focused on the focal plane. The two principal types of lenses in a TEM are the condenser lens and objective lens. The condenser lenses are situated on top of the specimen stage and gather electrons (forming the beam) to focus them on a specific area (the diameter of the beam) of the sample. The amount of current in the beam falling onto the sample can be controlled by the condenser lenses. But, the most important lenses in a TEM are the objective lenses which are located just around the specimen. These lenses generate the first magnified intermediate image (inverted) of the sample, the quality of which determines the resolution of the final image. In the back focal plane of the objective lens, a diffraction pattern is formed. An objective aperture is located at this place. Further down in the microscope, the intermediate and projector lenses are placed which further magnify the initial image that is formed by the objective lenses.

Apart from the lenses, the apertures (Figure 2.1) also play an important role in image formation. Apertures are different-sized holes in metal blocks, which are inserted at different places along the column. They limit the size of the beam that can pass through the column by filtering out the high-angle electrons. The apertures are named based on their association with the different lenses. The first aperture is situated at the top of the microscope and named as condenser aperture. It helps in condensing the beam to produce a coherent stream of electrons. The second aperture is located below the

sample and is known as the objective aperture which aids in the adjustment of the amplitude contrast of the generated image.

Other important components of the lens-system include the deflector and stigmator. A deflector is an electromagnetic component of a TEM which can change the direction of the electron beam with respect to the “optical axis”. Each EM lens is equipped with a pair of deflectors. A stigmator is another electromagnetic part of a TEM which corrects for the imperfections in the shape of the magnetic field used to focus the electrons.

2.2.2.5 The detector

During earlier times, electron micrographs were recorded either on photographic films or scintillator-based digital cameras such as charge-coupled device (CCD) or complementary metal-oxide semiconductor (CMOS) cameras (198). The usage of the films provided better detective quantum efficiency (DQE) and larger area of acquisition than the digital cameras but lacked automation while data collection, creating a bottleneck in high-throughput data screening. This problem with automation was solved with the advent of the digital cameras but the images obtained were noisy as the electrons were first converted to photons via interaction with a phosphor scintillator and thereafter recorded by the camera. However, in most of the cases, both the technologies could not produce images any better than 5-10 Å resolution. A surge in high-resolution structure determination with cryo-EM was observed once the direct electron detectors (DEDs) were developed which created a “resolution revolution” in the field (199). High DQE and faster frame-rates in DEDs have proven to be useful to solve structures for biological samples at near-atomic resolution (200). Examples of some of the DEDs which are in frequent use in the recent times are Falcon II and III (provided by FEI company), DE camera (Direct Electron) and K2 camera (Gatan, Inc.) (198).

2.2.3 Principles of EM image formation

Broadly speaking, there are three possibilities about what can happen to an electron when it passes through a thin object. If an electron does not encounter the sample, it would continue to move in the same original direction without losing any energy and finally reach the detector. This is the prevalent fate of an electron when used to image biological specimen and is called the transmitted electron. When an electron hits the

sample and is deflected from its original path but no energy transfer happens between the electron and the sample, then it is called as the elastic scattering event. It also happens that an electron encounters the sample which results in a change in the original directionality and deposition of some energy on the sample. This event might also accompany emission of secondary electrons and X-rays. This type of event is named as the inelastic scattering. The inelastically scattered electrons contribute to the noise in a TEM image. The image contrast generated in a TEM is dependent on the level of prevention of these inelastically scattered electrons to contribute to image formation. Another possibility involves a greater than 90° bent in the path of the electrons once they hit the sample and they are termed as the back-scattered electrons. These electrons are not used to form the image in a TEM, either. The elastically scattered electrons contain high-resolution information and are responsible for the generation of contrast in an image. Because biological specimens are best described as the weak-phase objects, this contrast is known as the phase contrast. The generation of contrast by electrons can be best understood if electrons are considered as waves. Waves can either positively or negatively interfere and therefore result in an increase or decrease in the amplitude of the resultant waves. A comprehensive review of image formation in TEM can be found elsewhere (201).

2.2.4 Contrast Transfer Function (CTF)

Biological specimens are weak-phase objects as they consist of atoms with low atomic numbers. Therefore, when imaged with an EM, these objects do not produce amplitude contrast. The visualization of the object is made possible by applying some underfocus (aka defocus) which generates phase contrast. But the transfer of information from the object to the detector is not straightforward. The two-dimensional (2D) images collected using an EM can be considered as the 2D projections of three-dimensional (3D) structures but with the addition of some modulation and noise. Therefore, to restore an ideal image, it is necessary to correct the image for the lost information and for this purpose, CTF correction is used at the early stage of image processing. The CTF describes the information transfer from the object to the detector as a function of spatial frequency. The amplitude of the transferred information decreases with the increasing spatial frequency. Because of the oscillatory nature of the CTF, it is necessary that images are collected at variable defocus values ($\sim 0.5\text{-}3.0\ \mu\text{m}$) to compensate for the absence of information transfer at the CTF zeroes. Accurate

estimation of the CTF parameters, especially defocus and astigmatism is of paramount importance for checking the quality of the micrographs and to obtain a good 3D reconstruction of the biological specimen.

2.2.5 Cryo-EM

In cryo-EM, the specimen is plunge-freezed into liquid nitrogen-cooled liquid ethane which has a temperature of around -170 °C. This forms amorphous/non-crystalline ice (aka vitreous ice) surrounding the specimen which helps it to maintain a native conformation and protects the specimen against beam-induced damage. To further prevent the biological specimen from beam-induced damage, low dose condition ($\sim 15\text{--}25\text{ e}^-/\text{\AA}^2$) is used for imaging. This imparts very low signal in the images but high-resolution information is preserved. With the advent of DEDs possessing fast frame-rates, an increasing tendency to collect data at a high dose and to discard the damaged frames later from the image stacks, is gaining popularity. This strategy thus results in higher contrast in the images which helps in particle picking and image alignment without killing the high-resolution information.

2.2.6 Single particle analysis

The field of cryo-EM has grown to include various techniques over the years such as single particle analysis, tomography and electron crystallography among which single particle analysis (SPA) has been the widely used variant. In SPA, many 2D projection images are collected of the identical copies of macromolecules or their assemblies, which are combined to produce a 3D reconstruction. Typically, in SPA, the first two steps of grid preparation and data collection are the quickest while the third step, image processing, is the longest. Image processing is again divided into multiple steps which are discussed below.

2.2.6.1 Image processing

A good quality micrograph is characterized by the isotropic Thon rings which extend to high-resolution in the 2D power spectra. Initial screening based on the quality of the Thon rings should be carried out to exclude bad quality micrographs containing drift and astigmatism. Described below are the steps followed thereafter to generate the 3D reconstruction.

2.2.6.2 Particle picking

The first step in single particle analysis is particle picking. In a typical high-resolution project, thousands of cryo-micrographs are collected in a 3-4 days session in an EM. Each of these micrographs contains hundreds of copies (aka particles) of the specimen, which are thereafter picked employing different computational algorithms that can broadly be divided into three categories: local statistical measures, cross-correlation and edge detection (202). Most of the software use different variations of these algorithms and operate in manual, semi-automatic or automatic modes. Manual selection of particles requires user supervision throughout the session of picking and is time-consuming. Semi-automatic picking is faster because after an initial round of manual picking based on the discretion of the user, the software automatically picks the remaining particles. Automatic picking requires the user to hand-pick a few particles and generate 2D classes based on the morphological features of the particles. The software, thereafter applies a template-matching algorithm to pick particles from the remaining micrographs, by itself, without requiring any further user intervention. However, owing to the noisy nature of the typical cryo-micrographs, fully-automated picking should be used with utmost care to avoid inclusion of too many false positives and empty boxes for downstream image processing.

Figure 2.2 The typical workflow in cryo-EM based structural study. For a molecule with unknown structure, an initial model is generated using negative stain images. This initial model is thereafter used to generate 2D projections to calculate particle orientation in cryo-EM images. Refinement of cryo-EM images using the optimal orientation parameters help to generate a high-resolution final structure. Taken from https://www.creative-biostructure.com/em-platform_66.htm.

2.2.6.3 Estimation of CTF parameters and correction for their effects

Image restoration by CTF correction is broadly done in three ways (203). 1. Filtering at first zero value- this results in truncation of the high-resolution signal post first zero. 2. Phase correction only- reversing (flipping) the phases of the CTF values which go below zero. This is equivalent to converting the negative contrast to positive contrast. No amplitude correction is done in this case. 3. Both phase and amplitude corrections- Wiener filtering method is applied here (203). A review on the relative performances of the various CTF determination programs is reported by Marabini *et al.*(204).

2.2.6.4 2D clustering and formation of class averages

Because of the low signal-to-noise (SNR) ratio in the cryo-EM images, while picking particles, lots of image artifacts, invalid particles, or empty background get included in the dataset. 2D classification sorts out these unwanted things and clusters the data into homogenous subsets known as the 2D classes, based on the morphological features of the particles. Thereafter, the members of the 2D classes are averaged to provide a high-contrast representative view of the classes. 2D classification and averaging begin with the alignment of particles to achieve the same relative orientations for the particles in the same class. Accuracy in image alignment is central to obtain high-resolution in the 3D reconstituted volume obtained by cryo-EM. For alignment, each particle is allowed three in-plane movements which include translations in x-y plane and rotation. Several computational algorithms are used to perform this alignment and clustering, which are broadly divided into supervised (aka model-based) and unsupervised (aka reference-free) mode of classification (205). In supervised mode, a reference is used to classify the particles, while the unsupervised mode operates completely based on the intrinsic features of the particles. Examples of supervised mode include Self-correction function method (SCF), Resampling to polar coordinate method (RPC), 2D fast rotational matching method (FRM2D), and Cross-common line method. The unsupervised mode is used in the Multivariate statistical analysis (MSA) and Maximum-likelihood (ML) methods. Supervised 2D classification comes with the risk of introducing reference-bias in the classification due to the low SNR of the particles. Therefore, unsupervised mode of 2D classification is preferred in most cases.

2.2.6.5 Calculation of the initial structures

Low-resolution negative-stain 3D reconstruction, X-ray model or EM-map of a homolog can serve as structural templates for the 3D structure calculation from the 2D cryo-EM images. When no reasonable structural templates are available, *ab initio* 3D structure determination is the way forward. Two *ab initio* methods that have gained popularity in the field are the random conical tilt (RCT) and common lines (aka the “Angular reconstitution”) (206,207).

In RCT, two images are collected for each particle by tilting the specimen stage by some definite angles (208,209). Typically, one image is collected at no tilt (0°) and the other at a very high tilt (usually around 60°). RCT data is mostly collected in negatively stained samples to avoid the pitfalls of electron-induced damage in cryo-specimen. Nonetheless, the reliability of RCT data is limited by artifact generation from staining and the incomplete sampling of a structure in Fourier space leading to distorted reconstructions (“missing cone” problem). The common lines method is based on the common-line projection theorem which states that two different 2D projections of the same 3D object will always have a 1D line projection in common (210,211). However, reliable detection of the common line is challenging due to low SNR of the cryo-EM images. Therefore, the multivariate statistical analysis is used to reduce noise while using this approach (210). However, both RCT and common lines methods are limited in producing angles for particle alignment which are far from the optimal and therefore, new advancements in this front are actively being pursued by the scientific community.

2.2.6.6 Structure refinement and resolution

After the initial structures are calculated, they are used to generate reprojections which in turn are used to modify the orientation parameters of single-particle projection images to achieve a better match. Thus, an updated and better 3D reconstruction is obtained which is thereafter used as a template. This whole process is iterated until a convergence is reached. To measure progress during the iterations, Fourier shell correlation (FSC) curve is used which provides information on the level of SNR as a function of the spatial frequency (212,213). The FSC curve is calculated by splitting the dataset into two halves, calculating two independent reconstructions, and then measuring the correlation coefficients between these two reconstructions within resolution shells extracted from Fourier transforms of the two volumes.

$$FSC_{1,2}(r) = \frac{\sum_{r_i \in r} F_1(r_i) \cdot F_2(r_i)^*}{\sqrt{\sum_{r_i \in r} |F_1(r_i)|^2 \cdot \sum_{r_i \in r} |F_2(r_i)|^2}}$$

r corresponding shell in Fourier space
F(r_i) complex “structure factor” at *r_i* in Fourier space
 $\sum_{r_i \in r}$ summation over all Fourier space voxels *r_i* in shell *r*

FSC curves can also be used to access the average resolution of the final 3D reconstruction. Although there are multiple opinions about what is the best strategy, three FSC cutoff values are used- 0.143 (214), 0.33 (214) and 0.5 (215). However, resolution values reported using FSC curve for cryo-EM maps are often misleading. Unlike X-ray crystallography where uniform resolution is obtained for the whole map, cryo-EM produces a varied resolution for the different parts of the same map. Therefore, coloring the 3D reconstruction according to the local resolution has become a common practice while reporting the FSC curve to indicate average resolution. The quality of the map can also be evaluated by checking for visibility of different resolvable features. These features include protein helix grooves, bulky side-chain densities and well-separated strands of beta sheets, to name a few.

2.2.7 Latest developments in detector technology and image processing

The above-mentioned steps are generally applicable for all EM-projects but the resolution of the final reconstruction was typically limited to 5-10 Å until even 3-4 years back. Bottlenecks to achieve higher resolution were mainly caused by: 1. the recording medium or the detector and 2. the image blurring, caused by the instability of the sample stage and motion induced by electron beam (216,217). Ideally, the detector should be able to read each incoming electron and would not add any noise to the resulting image. It should, therefore, possess DQE of 1, but this ideal condition is not even closely met in the best available detectors. In the past, photographic films and charge coupled device (CCD) detectors have been widely used for different reasons. The former possessed better DQE (~ 0.3) than the latter (~ 0.1) and offered a large field of view but came with the laborious process of film development and scanning (216).

The advent of automation with the CCDs in the market promoted its use. However, both detector varieties were eliminated soon after the advent of the direct electron detectors (DEDs) which possessed a better DQE ($\sim 0.4-0.8$) (200) and offered high throughput image collection. DEDs are commercially available from many vendors like Direct Electron (DE), FEI/Thermo Scientific (Falcon) and Gatan (K2). Another factor that contributed to the increased resolution in cryo-EM is the improvement in the image processing software to correct for the aberrations due to beam-induced motion of the particles. This is known as motion correction and popular programs to perform this include MotionCorr (218), Unblur (219) and MotionCor2 (220).

2.2.8 Limitations and future direction of the technology

The challenges in the cryo-EM field and their possible solutions are described in details in a number of review articles (217,221-223). Some of the challenges that users of the cryo-EM face are discussed here:

1. 3D reconstruction by cryo-EM is achieved by image averaging. Therefore, structural homogeneity is a requirement to obtain high-resolution. Typically, biological specimens display a range of structurally different conformations while in solution, thus posing a challenge for image alignment. It has now become possible to classify images into homogenous subsets after the advent of ML-based algorithms as implemented by the software packages FREALIGN (224), XMIPP (225) and RELION (161), direct detection cameras providing better quality images and motion-correction software. Nonetheless, too much of heterogeneity in the sample is undesirable to obtain a high-resolution 3D reconstruction. Glutaraldehyde cross-linking along with density gradient centrifugation have been used to bypass the heterogeneity issue but with limited applicability due to the non-specific nature of the reaction (223,226).
2. Introduction of structural heterogeneity during sample preparation is another problem. While some biological specimens are prone to denaturation when they encounter air-water interface while on the blotted EM grid, some assume preferred orientation at this interface, creating a bottleneck for a good 3D reconstruction. To circumvent these problems, technological advances to immobilize the particles on the specimen-support film of the EM-grid (223,227) are increasingly gaining popularity.

3. Cryo-EM has been successful in generating many high-resolution structures for large complexes like viruses and ribosomes, but its applicability is still limited for smaller proteins and complexes which do not produce enough contrast due to low SNR in the cryo-images. Phase plate and energy filter technology have proven to dramatically improve SNR, thus, holding the promise to be useful in such cases (222).
4. Applicability of cryo-EM to determine structures is limited due to the high maintenance cost of the facility. Another constraint is time. A high-resolution project entails collection of a large dataset consisting thousands of images. Processing of such dataset can take several months, only to yield a low-resolution blob at the end. So far, no estimate for the resolution of the final reconstruction can be obtained before the time-consuming data processing steps. Another limitation is the computational requirements. Because image files are huge, a powerful computing cluster is required to process data. This poses limitation when such facility is not readily available or available with limited accessibility. Centralized facilities such as Howard Hughes Medical Institute (HHMI) at the Janelia Farm campus in the USA, NeCEN in The Netherlands and eBIC at the Diamond Light Source, UK have made cryo-EM accessible to many researchers (181). Recently, FREALIGN-GPU (228) and RELION2 (229) has made GPU-based cryo-EM workflow popular among the scientific community as these software can obviate the use of large computing clusters.

2.2.9 Cryo-electron microscopy of cp-ribosome

Cp-ribosome particles were diluted to final concentration of 4 OD₂₆₀/ml (96 pmol/ml). 4 µl of this sample was applied to glow-discharged 2-nm carbon coated holey grids (Quantifoil R2/2) and incubated for 30 sec. FEI Vitrobot was used to blot the grids for 3 sec in 100% humidity at 4° C and thereafter grids were flash-frozen in liquid N₂-cooled liquid ethane. Two datasets were collected using 200-kV FEI Tecnai Arctica (henceforth mentioned as Arctica) and 300-kV FEI Titan Krios (henceforth mentioned as Krios) cryo-TEM. The Arctica was equipped with a back-thinned Falcon III direct electron detection device. Electron micrographs were collected automatically on Arctica by using FEI's automatic single particle acquisition software EPU at 0.2-2.5 µm underfocus values. 1,729 micrographs were recorded in a movie mode as a set of 7 frames (total dose of 26 electrons per Å²) at a calibrated magnification of 109,375

resulting in a pixel size of 1.28 Å on the object scale. The Krios was equipped with a back-thinned Falcon II direct electron detection device. Electron micrographs were collected automatically on Krios by using Leginon software (230) at 0.4-3.7 µm underfocus values. 3,161 micrographs were recorded in a movie mode as a set of 25 frames (total dose of 37.5 electrons per Å²) at a calibrated magnification of 133,333 which rendered a pixel size of 1.05 Å on the object scale.

2.2.10 Image processing of cp-ribosome Arctica-dataset

After data collection and screening for good micrographs, 1590 micrographs were selected for data processing. Whole-image drift correction was carried out using Motioncorr (218). e2boxer.py from EMAN 2.1 (1) was used to semi-automatically pick particles from motion-corrected micrographs and CTF parameters were estimated using CTFFIND3 (231). Data was processed with RELION 1.4 (2). In total, 338,305 particles were subjected to reference-free 2D classification to discard bad particles. 261,052 particles selected from the 2D classes were directly used for 3D refinement using a 60Å-filtered *E. coli* 70S map and thereafter 3D-classified without image alignment using a 50S mask. 202,766 particles from the 3D classes were taken and sorted according to the defocus values of the micrographs. Micrographs having defocus values higher than 1.5 µm were not included in subsequent refinements. 3D refinement of 174,949 particles with a mask around 50S subunit with defocus values less than 1.5 µm yielded a 50S density map at 3.47 Å resolution. ResMap (232) was used to calculate local resolution. Average resolution is calculated according to the gold-standard FSC= 0.143 criterion (214). The density map was corrected for the modulation transfer function (MTF) and sharpened using an automatically calculated B-factor (-104.3 Å²), before visualization.

2.2.11 Image processing of cp-ribosome Krios-dataset

A total of 3161 micrographs were collected and were subsequently evaluated for drift. After the evaluation, 1285 good micrographs were subjected to whole-image drift correction using Motioncorr (218). Contrast transfer function (CTF) parameters were determined using CTFFIND3 (231) and semi-automated particle picking was carried out using e2boxer.py from EMAN 2.1 (1). Data was processed using RELION 1.4 (2). At the outset, 187,946 particles were subjected to reference-free 2D classification applying 2X decimation. After sorting the particles into good and bad classes, 132,904

good particles were further subjected to 3D classification using 2X decimation guided by a 70S cp-ribosome map obtained previously (3). The dataset could be sorted into 22,737 (17%) poorly-aligned particles, 28,862 (22%) 50S and 81,305 (61%) 70S particles. 70S particles were used for high-resolution 3D refinement which yielded a 3.6 Å density map for the complete 70S cp-ribosome. Inclusion of movie frames for 3D refinement improved the resolution to 3.4 Å. Finally, Gaussian masks were used to perform 3D refinement on the ribosomal subunits which generated the final maps where the SSU and the LSU were resolved to 3.7 and 3.3 Å, respectively. Local resolution for the maps was calculated using ResMap (232) and gold standard FSC=0.143 criterion (214) was used for determination of the average resolution. The density maps were corrected for the modulation transfer functions (MTF) and sharpened using automatically calculated B-factors (-72.6 Å² for complete 70S ribosome, -97.9 Å² for 30S subunit and -89.6 Å² for 50S subunit), prior to visualization.

2.2.12 Model building and refinement of cp-ribosome LSU and SSU

The cp-ribosome LSU atomic model was initially built in the 3.5 Å Arctica map. This atomic model was then fitted in the 3.3 Å LSU map obtained from the Krios-dataset using real-space refinement. The SSU model was built in the 3.7 Å Krios SSU map. To localize and model the rRNAs and proteins in cp-ribosome maps, atomic coordinates of the *E. coli* ribosome (PDB ID: 4YBB) (86) were used as templates. A combination of homology and *de novo* modeling was used to build atomic coordinates for the complete 70S ribosome.

For the structural modeling of the rRNAs (23S, 5S, 4.8S and 16S), sequence information was obtained from the complete chloroplast genome sequence in NCBI (Accession no. AJ400848.1) and homology models were generated using the ModeRNA server (164). The models were thereafter rigid body fitted using the 'Fit in Map' functionality in UCSF Chimera (166) into the cp-ribosome maps. The model fitting was optimized by running *phenix.real_space_refine* (233) with the application of RNA base-pairing restraints generated from 'PDB to 3D Restraints' server (<http://rna.ucsc.edu/pdbrestraints/pdbtorestraints.html>) and RNA stacking restraints to maintain parallelity (default in phenix). These models were then manually rebuilt in COOT (167) using predicted secondary structure from the comparative RNA website (CRW) (56).

To model the ribosomal proteins, sequence information was obtained from UniProt (234) and GenBank (235) databases, and homology models were generated using the I-TASSER server (170). These proteins were initially rigid-body docked into the cp-ribosome maps and the fittings were optimized by running *phenix.real_space_refine* (233) using local rotamer fitting, morphing, simulated annealing and secondary structure restraints. Most of the chloro-specific protein extensions in the I-TASSER models did not fit the densities and were deleted from the models. These extensions were then manually rebuilt in density using a combination of ‘C-alpha baton mode’ and ‘add terminal residue’ functionalities in COOT (167). The protein extensions which assumed well-defined helical structures were *de novo* modeled using the tool ‘place helix here’ in COOT (167). In all stages of modeling in COOT (167), ‘Real Space Refine Zone’ and ‘Regularize Zone’ were used to locally fit model into density and maintain correct geometry. Because no suitable templates after extensive database search were found for the PSRPs cL37, cL38, cS22 and cS23, models for these were generated without specifying any template in the I-TASSER server (170) allowing the program to build the structures by performing multiple folding simulations from the identified templates in the Protein Data Bank (PDB). Homology models for plastid pY and bTHXc were generated using YfiA and bTHX models (PDB ID: 4Y4O) (236) as templates using the I-TASSER server (170). Placements of the PSRPs and plastid pY were achieved by applying a combination of guidance from previous literature (124) and based on the size and shape of the density.

At the final stage of modeling, the atomic coordinates for the SSU and LSU were merged and clashes at the subunit interface were removed using COOT (167). To reduce overfitting, poorly resolved regions of the maps were identified by generating model-map cross-correlation values using *phenix.map_model_cc* (237) and models in these regions were refined locally using *phenix.real_space_refine* (233) applying restraints (mentioned above) and tighter RMSD values before adding them back to the final model. Finally, the complete 70S ribosome structure was subjected to global refinement and minimization using *phenix.real_space_refine* (233) applying default parameters. Despite the high quality of the maps, register shifts could not be avoided at the poorly resolved flexible parts of the ribosome.

2.2.13 Model validation

The 70S cp-ribosome atomic coordinate was validated using the Molprobity server (168). The degree of model overfitting was evaluated by calculating FSC curves between the models and final cryo-EM maps. To generate the FSC curves, the model-derived maps (all frequencies until Nyquist included) were resampled onto the same grid as the final cryo-EM maps using UCSF Chimera (166) and FSC values were calculated using e2proc3d.py program in EMAN 2.1 (1). Maps were also colored to show local agreement between the models and maps using ‘vop localCorrelation’ command in UCSF Chimera (166).

2.2.14 Figure generation

UCSF Chimera (166) and PyMOL (165) were used to generate the figures in the manuscript. The plots showing local resolutions were generated using ResMap (232).

3. Results

High-resolution structure of the cp-ribosome from spinach

The work in this thesis is focused towards solving a high-resolution structure of the cp-ribosome from spinach, using cryo-EM and single particle analysis. To facilitate the understanding of structural interpretations, the result section is divided into two chapters.

Chapter 3 describes the cp-ribosome LSU in near-atomic details at the best resolved parts of the map. The main function of LSU is to harbor the catalytic peptidyl transferase center, to form peptide bonds. The LSU also harbors a tunnel through which the nascent polypeptide chain travels concomitant to its synthesis, which then emerges from the tunnel at a place known as the tunnel exit site. Here, in chapter 3, the similarities and unique features of the cp-ribosome LSU are presented in comparison with the *E. coli* ribosome LSU. It is observed that the structure of the chloroplastic 23S rRNA is altered at five places and most of these alterations are stabilized by the extensions in cp-ribosomal proteins. Chloroplastic 23S rRNA is also fragmented to produce 4.8S rRNA. Two bacterial ribosomal proteins L25 and L30 are absent from cp-ribosome LSU while two new proteins (PSRP5 and PSRP6, according to new nomenclature cL37 and cL38, respectively) have been gained although they structurally do not compensate for the missing proteins. The unique positioning of these PSRPs provide clues about their possible functions in the chloroplastic environment.

Chapter 4 describes the cp-ribosome SSU in near-atomic details at the best resolved parts of the map. SSU is responsible for the initiation of translation and therefore plays important role in the translational regulation. SSU harbors mRNA channel along with the mRNA entry and exit sites. In comparison with bacterial ribosomes, similarities and uniqueness of the cp-ribosome SSU are presented in this chapter. The locations and models of the 3 PSRPs (PSRP2, PSRP3 and PSRP4, according to new nomenclature cS22, cS23 and bTHXc, respectively) are presented which provide insights into the translation regulation inside the chloroplastic environment. The binding mode of the hibernation factor plastid pY, that co-purifies with our 70S ribosome, is discussed in near-atomic details. Finally, the presence of a novel intersubunit bridge compared to *E. coli* ribosome is discussed.

During the period of this thesis work, two cp-ribosome structures were published by Graf *et al.* (115) and Bieri *et al.* (114). While most parts of our structure and the analysis agree well with their results, some differences emerged after careful comparison. A brief analysis of such differences is presented in the discussion section.

3.1 Structural insights into the large subunit of the cp-ribosome

Towards solving the complete structure of the cp-ribosome LSU, a cryo-EM dataset was collected using an in-house (Nanyang Technological University, Singapore) Arctica microscope operating at 200 kV. This dataset could generate an average 3.5 Å resolution cryo-EM map for the LSU which enabled interpretation of the unique features of the subunit. However, the SSU could not be captured at high-resolution from this dataset. Probable causes behind obtaining low-resolution for the SSU include the rotation of the SSU while positioned on top of the LSU (so-called ratcheting) which results in conformational heterogeneity, use of a 200-kV microscope (Arctica) and also over-representation of the LSU on the cryo-grids.

3.2 Purification profile of the cp-ribosome

Cp-ribosomes were purified using sucrose density gradient centrifugation. The purification profile is shown in Figure 3.1. The fractions collected for cryo-EM studies are boxed in black dotted lines which represent intact 70S ribosome isolated from spinach chloroplast.

Figure 3.1 A_{280} profile of spinach cp-ribosome crude extract on a sucrose gradient. Crude ribosomes purified from the spinach chloroplast were layered on top of a 10-40% (w/v) sucrose gradient and spun at $111,000 \times g$ for 4 hr and thereafter gradients were fractionated from top to bottom. Modified from (3).

3.3 Ribosomal particles on cryo-EM grid

Purified cp-ribosomes were applied onto TEM copper grids and the particle distribution was checked collecting few test micrographs at a low and high defocus values. The particles showed optimal distribution covering the whole surface area of the micrographs. The image contrasts appeared to be optimal for particle picking and image alignment to obtain a reliable 3D reconstruction. Representative micrographs collected at two different defocus values are presented in Figure 3.2.

Figure 3.2 Distribution and contrast of particles on cryo-EM grid. To generate enough contrast for particle visualization, defocus was applied using the microscope lens system. Examples of such two images, collected at 0.3 (A) and 1.5 (B) μm defocus, are displayed to show optimal distribution and contrast of the cp-ribosome particles on copper grid.

3.4 Image processing

The image processing of the cryo-EM dataset collected using an in-house Arctica microscope is depicted in Figure 3.3 below. Following the figure, a detailed analysis of the results obtained at individual steps of the data processing is provided. It should be noted that particles were semi-automatically picked using the software EMAN 2.1 (1). The whole dataset thereafter was processed using RELION 1.4 (2).

Figure 3.3 Image processing flowchart for the Arctica dataset. From 1590 micrographs collected, a total of 338K particles were picked semi-automatically using e2boxer.py tool from EMAN 2.1 (1). Data was processed using RELION 1.4 (2). 3X binned particles were sorted using 2D classification step, discarding 77K particles. Thereafter, 3D refinement was used to find the correct angles for particle alignment. 3D classification using a 50S mask resulted in classes containing both low- (58K particles) and high-resolution (202K) ribosomal particles. After this, micrographs collected at higher defocus (>1.5 μm) were discarded and 3D refinement was performed with unbinned particles resulting in a density map resolved to 3.53 Å. Final 3D refinement applying a 50S mask improved resolution to 3.47 Å. Taken from (3).

3.5 2D classification

At the beginning of the data processing workflow, a 2D classification was carried out applying 3X binning (merging 3 pixels into 1) on 338,305 particles to sort out the so-called “junk” from the original ribosomal particles (Figure 3.4). Because low dose is used to image biological samples, particle binning helps in improving the contrast of the images which helps in better alignment. The “junk” originates from disintegrated particles, salt and water crystals, and other contaminations which produces contrast in a typical electron micrograph and are mistakenly picked by the particle picking software or a human user. 2D classification also helped in dividing the ribosomal particles with similar morphological features into distinct classes. Figure 3.4 shows the distribution of the particles after 2D classification.

Figure 3.4 2D classification results. Among 150 classes generated after 3X binning of the particles, only 50 are displayed. The red color boxes indicate the representative good classes that were taken for downstream data processing.

Counting particles in each class revealed that LSU is over-represented in the dataset. As the collected fractions from the purification profile graph were shoulder-free, it could be inferred that although the purification scheme resulted in intact 70S cp-ribosome, the mechanical force that ribosomes experience during the cryo-EM grid preparation, induce the subunit dissociation. As we will discuss later that, this creates a bottleneck in obtaining a high-resolution structure for the complete 70S ribosome and

the strategy to overcome this problem would also be discussed. The presence of different views of the ribosomal particles and the good contrast of the identifiable features of the secondary structural elements confirm that the dataset contains the potential for a reliable 3D reconstruction. Selection of particles from such classes provided with 261,052 particles which were then subjected to 3D refinement using RELION 1.4 (2).

3.6 Initial 3D refinement and subsequent sorting by 3D classification

After sorting out the bad particles from the dataset using 2D classification approach, a 3D refinement run was carried out using 3X decimated particles. This helped to find the correct angles for particle alignment and therefore was beneficial for a robust 3D sorting of the dataset in the next step.

Figure 3.5 3D classification results. The population of particles in each class is mentioned as percentages. Apart from class1 which represents LSU, the other classes contain the complete 70S cp-ribosome. Taken from (3).

Figure 3.5 shows the results from the 3D classification step. Class 1, representing just the LSU, turned out to be the most populated one, while the remaining classes showed a varied distribution of the 70S ribosomes. Because the initial goal of the work was to elucidate the high-resolution structure of the complete 70S ribosome, classes 2, 3, 4 and 5 were combined and refined. This provided a low resolution (> 4 Å) map for the complete 70S which was not suitable for atomic model building. Therefore, the focus was shifted to obtain a high-resolution structure for just the LSU by combining all five classes and applying a Gaussian mask on the LSU for further data processing. As shown earlier in Figure 3.3, application of a mask on the LSU resulted in a better sorting of data into classes based on the high- and low-resolution features.

3.7 Estimation of resolution for the final LSU reconstruction

The resolution for the cp-ribosome LSU cryo-EM map after applying masked refinement procedures was estimated to an average 3.5 Å, according to the FSC= 0.143 criterion (214) (Figure 3.6A). The local resolution calculated using the program ResMap (232) (Figure 3.6B) showed that the core of the ribosome is the best-resolved part of the map while the peripheral regions (mainly L1 and L7/L12 stalks) remained flexible and therefore could not be captured at high-resolution. The coloring of the map according to their local resolution revealed that the core is resolved to around 2.8 Å and the peripheral regions are at higher than 6 Å resolution. Consistent with the estimated resolution, densities for most of the rRNA bases and protein side chains were discernible throughout the map, except for the periphery (Figure 3.6C, D and E). The refinement and model validation statistics are presented in Table 3.1.

Figure 3.6 Overall and local resolution estimation and map quality. (A) Average resolution of the LSU cryo-EM map as estimated by the FSC= 0.143 criterion (214). The X-axis represents spatial frequency (fourier space) which is related to the resolution (real space) by the formula: spatial frequency= 1/resolution. The Y-axis represents the Fourier Shell Correlation values. (B) The local resolution estimation using ResMap (232) showing high-resolution at the core but low-resolution towards the periphery. (C), (D) and (E) Quality of the postprocessed LSU masked map displaying the features expected at 3.5 Å resolution. (C) A beta sheet from protein L19 is shown to display the separation of strands, (D) An alpha helix from protein L22 is shown to display the side chain densities. (E) Part of H88 (23S rRNA) showing clear separation of nucleotide bases in the density. (A) and (B) are taken from (3).

Table 3.1 Refinement and model statistics for the cp-ribosome LSU. The cp-ribosome LSU model was refined against the 3D reconstruction of LSU generated after the masked refinement. Taken from (3).

Data Collection	
Particles	338,305
Pixel Size (Å)	1.28
Defocus Range (µm)	0.2-2.5
Voltage (kV)	200
Electron Dose (e ⁻ Å ⁻²)	26

Model composition	
Non-hydrogen atoms	90,671
Protein residues	3,229
RNA bases	3,029

Refinement	
Resolution (Å)	3.47
Map sharpening B-factor (Å ²)	-104.3
CC _{map_model}	0.72

RMS deviation	
Bonds (Å)	0.01
Angles (°)	1.23

Validation (proteins)	
Molprobability score	2.65 (39 th percentile)
Clashscore, all atoms	25.65 (20 th percentile)
Poor rotamers (%)	0.51
Favoured rotamers (%)	97.15

Ramachandran plot	
Favoured (%)	78.87
Outliers (%)	0.35

Validation (RNA)	
Correct sugar puckers (%)	95.18
Good backbone conformations (%)	60.32

3.8 Overall architecture of the cp-ribosome LSU

Following the endosymbiotic theory, the origin of chloroplast can be traced back to cyanobacteria (111). Therefore, it would be ideal to compare the structure of the spinach cp-ribosome with the counterpart in cyanobacteria. But till date, detailed structural information on the ribosome and translation process is limited mainly to *Escherichia coli* and *Thermus thermophilus* ribosomes. Due to this reason, the cp-ribosome LSU map derived in this study was compared mostly with the *E. coli* ribosome with occasional comparison with the *T. thermophilus* ribosome, when deemed necessary. It was observed that the overall architecture of the cp-ribosome LSU is similar to its *E. coli* counterpart (Figure 3.7).

Figure 3.7 Overall architecture of the cp-ribosome LSU. The atomic model for the cp-ribosome LSU is presented from solvent-facing and subunit interface sides to indicate the location of its structural components. The proteins and RNAs are labelled as indicated. The two PSRPs, cL37 and cL38 are colored red. Modified from (3).

However, despite the similarity between the ribosomes, there are some differences which can be discerned after a careful analysis. The *E. coli* ribosome structure solved to the highest resolution till date is deposited in the RCSB PDB database with the ID 4YBB (86). This 4YBB ribosome model was downloaded and rigid-body fitted into the cp-ribosome LSU map which led to the identification and localization of all the bacterial-homologous proteins and the rRNA elements in the cp-ribosome LSU map. Apart from the two proteins L25 and L30 (Figures 3.7 and 3.8), all the bacterial-homologous proteins could be traced in cp-ribosome LSU.

Figure 3.8 Protein loss and gain in cp-ribosome compared to *E. coli*. bL25 and uL30 have been lost during evolution from cp-ribosome LSU while two PSRPs, cL37 and cL38 have been acquired. The cp-ribosome LSU map has been filtered and shown in transparency to improve the clarity. Fitted bacterial ribosome PDB model bears the ID 4YBB (86) in RCSB PDB. Modified from (3).

Interestingly, two new proteins namely cL37 (PSRP5) and cL38 (PSRP6) (Figures 3.7 and 3.8) have been acquired in the course of evolution. As clear from the Figure 3.8, these PSRPs do not compensate for the missing L25 and L30 but rather position themselves in new locations. Therefore, the total number of proteins remain the same between the *E. coli* and cp-ribosome LSU. However, most of the chloroplastic LSU proteins are longer in size due to the presence of extensions at their N- and C-termini (Figure 3.9). Further analysis to compare the rRNA elements revealed that the chloroplastic 23S rRNA is fragmented to give rise to a new species of rRNA namely 4.8S rRNA (Figure 3.7). Sequence comparison of the 23S rRNA between *E. coli* and chloroplast further revealed that deletions and insertions are spanned across the entire chloroplastic 23S rRNA sequence (Appendix Figure 1). Because the deletion lengths are often similar to the closely located insertion lengths (and vice-versa), they are mutually compensating which keeps the overall architecture of the 23S rRNA similar to *E. coli*. However, some of the peripheral regions in the 23S rRNA showed considerable structural differences and they were observed to be stabilized by the chloroplastic protein extensions.

Briefly, a total of 28 *E. coli* orthologous proteins, two PSRPs, 4.8S, 5S, and 23S rRNAs were built in the cp-ribosome LSU map (Figure 3.7). The ribosomal stalk regions, the so-called L1 and L7/L12 stalks were flexible and therefore the weak densities captured for these regions could not be used to build models. Therefore, the

final model of the cp-ribosome LSU lacks the protein models for uL1c, uL10c, uL11c and bL12c, all of which reside in the stalk regions.

Table 3.2 Constituent proteins of the cp-ribosome LSU. Protein residues and signal peptide positions are numbered according to the information present in UniProtKB/Swiss-Prot (234). UniProtKB/Swiss-Prot derives the signal peptide length from the sequencing study as reported in (119). The nomenclature of the ribosomal proteins is based on Ban *et al.* (238). Modified from (3).

Protein name	Chain ID	Uniprot ID	Mature protein (range, aa)	Built residues (range, aa)	Size of homologous <i>E. coli</i> proteins (aa)	Comments
uL1c		Q9LE95	73-352		234	Poor density, model could not be built
uL2c	E	P06509	1-272	23-270	273	
uL3c	F	P82191	85-305	86-297	209	
uL4c	G	O49937	51-293	52-261	201	
uL5c	H	P82192	39-258	56-230	179	
uL6c	I	P82193	39-220	41-214	177	
bL9c	J	P82180	42-196	43-95	149	
uL10c		P82139	53-232		165	Poor density, model could not be built
uL11c		P31164	67-224		142	Poor density, model could not be built
bL12c		P02398	57-189		121	Poor density, model could not be built
uL13c	L	P12629	48-250	100-246	142	
uL14c	M	P09596	1-121	1-121	123	
uL15c	N	P22798	62-271	93-254	144	
uL16c	O	P17353	1-135	2-135	136	
bL17c	P	P82194	11-126	11-126	127	
uL18c	Q	P82195	45-166	47-166	117	
bL19c	R	P82413	78-233	118-232	115	
bL20c	S	P28803	2-128	3-117	118	
bL21c	T	P24613	56-256	91-234	103	
uL22c	U	P09594	2-199	26-149	110	
uL23c	V	Q9LWB5	77-198	105-193	100	
uL24c	W	P27683	48-192	50-173	104	
bL27c	X	P82190	58-194	68-165	85	
bL28c	Y	P82245	72-148	72-145	78	
uL29c	Z	P82248	59-168	63-149	63	
bL31c	a	P82249	37-130	38-75	70	Poor density
bL32c	b	P28804	2-57	2-47	57	
bL33c	c	P28805	2-66	9-59	55	
bL34c	d	P82244	92-152	92-148	46	
bL35c	e	P23326	87-159	91-157	65	
bL36c	f	P12230	1-37	1-37	38	
cL37	g	P27684	63-142	94-142		PSRP variant 5 α , pI=10.01
cL38	h	P82411	48-116	48-93		pI= 9.87

3.9 Chloroplast-specific protein-extensions in the LSU

The total number of proteins are same (33 proteins) between the chloroplast and *E. coli* ribosome LSU but the chloroplastic proteins are usually longer as they contain extensions in their N- and C-termini (Figure 3.9). This is one of the reasons behind the altered protein: rRNA ratio in the cp-ribosome (2:3) compared to the *E. coli* ribosome which maintains a 1:3 ratio (119,120).

Figure 3.9 Extensions in cp-ribosome LSU proteins. Cp-ribosome LSU map is shown from the solvent-facing and subunit interface sides. Density for 23S rRNA, 5S rRNA and 4.8S rRNA are colored grey, goldenrod and sienna, respectively. The cp-ribosome proteins and their extensions are colored yellow and green, respectively. Position of cL37 and cL38 are indicated in red. Modified from (3).

There are 31 *E. coli*-orthologous proteins present in the cp-ribosome LSU among which 25 proteins are extended. The lengths of these extensions vary widely, ranging between 3 (as present in uL18c) to 97 (bL21c) amino acids (Table 3.3). Notable examples where the extension length exceeds 20 residues include L1 (+46 residues), L4 (+39), L5 (+41), L13 (+61), L15 (+55), L19 (+46), L21 (+97), L22 (+89), L23 (+35), L24 (+40), L27 (+53), L29 (+46), L31 (+23) and L32 (+23). It is interesting to note that, two chloroplastic proteins, bL21c and uL22c are doubled in size due to extensions. All these extensions are mostly located (Figure 3.9) in the peripheral solvent-exposed regions of the ribosome and are flexible. This flexibility resulted in poorly resolved density in some of these extensions although well-resolved density could be discernible for the extensions which formed secondary structural elements. Due to the absence of suitable homologs even after extensive database search using BLASTp, *de novo* modeling was applied to build models for 16 such protein-extensions (Figure 3.9) in the current study. Due to the poor quality of the density, some of the

extensions could not be built. Table 3.3 shows a comprehensive list of all the extensions along with the models that were built in this study.

Table 3.3 Protein extensions in the spinach cp-ribosome LSU. Length of the chloroplast-specific protein extensions are indicated along with the modeled lengths. Modified from (3).

Protein name	N-terminal extension length (aa)	N-terminal extension modeled (aa)	C-terminal extension length (aa)	C-terminal extension modeled (aa)
uL1c	42	0	4	0
uL2c	0	0	0	0
uL3c	4	3	7	0
uL4c	6	5	33	1
uL5c	15	0	26	0
uL6c	1	0	3	0
bL9c	7	6	0	0
uL10c	0	0	14	0
uL11c	5	0	12	0
bL12c	4	0	0	0
uL13c	52	0	9	5
uL14c	0	0	0	0
uL15c	18	0	37	20
uL16c	0	0	0	0
bL17c	0	0	0	0
uL18c	3	0	0	0
bL19c	46	6	0	0
bL20c	0	0	4	0
bL21c	67	32	30	8
uL22c	29	4	60	12
uL23c	28	0	7	2
uL24c	18	15	22	4
bL27c	0	0	53	24
bL28c	0	0	0	0
uL29c	3	0	43	24
bL31c	0	0	23	0
bL32c	0	0	23	13
bL33c	11	4	0	0
bL34c	0	0	10	8
bL35c	2	0	6	4
bL36c	0	0	0	0

Figure 3.10 Comparison between *E. coli* ribosome LSU proteins and homologous cp-ribosome LSU proteins. On the left, the chloroplastic proteins are shown in green along with their densities in transparent red mesh. Chloroplast-specific extensions are colored orange. On the right, chloroplastic proteins are overlaid with *E. coli* counterparts in grey (PDB ID 4YBB) (86) to show comparison. Modified from (3).

3.10 Localization and modeling of cL37 (PSRP5) and cL38 (PSRP6) in the LSU

Cp-ribosomes harbor 5 PSRPs, two in the LSU (cL37 and cL38) and three in the SSU (cS22, cS23 and bTHXc). After the assignment of all the *E. coli*-homologous proteins (with the exceptions of L1 and L7/L12 stalk proteins), protein-extensions and rRNAs into their respective densities, two densities remained unaccounted in the LSU map. The high-resolution map obtained in this study allowed unambiguous identification and modeling of these two densities as cL37 (Figure 3.11) and cL38 (Figure 3.12).

cL37 is a plastid-specific protein and is encoded as a 142 amino acids long protein by the nuclear genome. The mature protein after subtracting the 62 amino acids signal peptide is 80 amino acids long. The annotation for cL37 can be found at UniProtKB database bearing the identification number P27684. The mature protein contains a pI value of 10.72 as calculated using the ProtParam online tool (239). Consistent with its basic nature, cL37 was found to be buried in the rRNA environment which is negatively charged (Figure 3.11B). The structural modeling of cL37 was

carried out *de novo* by tracing the density and being guided by the placement of bulky side chains in the protein. The model revealed that cL37 contains a C-terminal small α -helix which is connected to a longer central α -helix via a linker (Figure 3.11). The observed density (Figure 3.11C) for the mature cL37 in the cp-ribosome LSU cryo-EM map encompassed amino acids 94-142. The density of the N-terminal 30 residues (amino acids 63-93) could not be discerned in the map most probably due to flexibility. Figure 3.11D displays the local resolution calculation for cL37 which shows that the structural modeling for the protein could be carried out with high confidence as most of the density was captured at high-resolution. At this location, cL37 mediates interaction with the rRNA helices H58 and H60 via the central and C-terminal α -helices, respectively. The reason behind the cp-ribosome acquiring cL37 at this location was further investigated by comparison with *E. coli* ribosome. This revealed that the structure of rRNA around cL37 is remodeled (Figure 3.14D in the later section) in the cp-ribosome and cL37 seems to stabilize these changes.

Figure 3.11 The localization and modeling of cL37 (PSRP5). (A) The cp-ribosome LSU model is displayed to point out the location of cL37. Central protuberance (CP) and L1 stalk are labelled. 23S rRNA is colored dark blue and all other ribosomal proteins except cL37-neighboring ones are colored sky blue. (B) The area surrounding cL37 is zoomed in from (A) to show cL37's environment. Neighboring protein and RNA elements are labelled as indicated. (C) cL37 model (red) displaying most of the side chains inside well-resolved density (grey). (D) ResMap (232) calculation of cL37 density to show local resolution in Å as indicated by the scale bar. (B) and (C) are modified from (3).

It should be noted here that the positioning of cL37 at this location disagrees with the localization of this protein near to the tRNA exit site by Sharma *et al.* (124). Analysis of the tRNA exit site in the current 3.5 Å map of cp-ribosome LSU revealed the presence of the C-terminal extension of uL15c. Therefore, it can be said that the erroneous localization of cL37 (mentioned as PSRP5 in the concerned paper (124)) by Sharma *et al.* is due to the low resolution of their reconstruction. Consistent with the findings in the current study, recently reported high-resolution cryo-EM structures of

the cp-ribosome by two more labs (114,115) have also placed and modeled cL37 at the correct position (near H58 and H60 of 23S rRNA) as mentioned above.

Coming to think of the functional significance of cL37, not much related literature could be found. A study on Arabidopsis has attributed the cause of impaired growth and reduced photosynthesis to the mutations in *psrp5* (gene for cL37) and therefore this protein is an essential PSRP (240). The same study also suggests that, cL37 might be involved in the assembly of the LSU and maintaining its stability (240). It seems that the close association of cL37 with the LSU rRNA is reminiscent of such possibilities. Concerning the speculation that PSRPs are involved in light-dependent regulation of translation (240), we feel that the RNA-buried nature of cL37 indicates its inaccessibility to interact with the protein factors necessary for the translational regulation in the chloroplastic environment.

Like all the PSRPs, cL38 (PSRP6) is also encoded by the nuclear genome. Synthesized as a 116-amino acids protein, the mature form of the protein is 69 amino acids long following the cleavage of 47-amino acids signal peptide as evident from the annotation in UniProtKB database (ID: P82411). Structural modeling of cL38 in cp-ribosome LSU map showed that the protein assumed an extended conformation (Figure 3.12).

Figure 3.12 The localization and modeling of cL38 (PSRP6). (A) The cp-ribosome LSU model is displayed to point out the location of cL38. Central protuberance (CP) and L1 stalk are labelled. 23S rRNA is colored dark blue and all other ribosomal proteins except cL38-neighboring ones are colored pink. (B) The area surrounding cL38 is zoomed in from (A) to show cL38's environment. Neighboring protein and RNA elements are labelled as indicated. (C) cL38 model (red) displaying most of the side chains inside well-resolved density (grey). (D) ResMap (232) calculation of cL38 density to show local resolution in Å as indicated by the scale bar. (B) and (C) are modified from (3).

Density (Figure 3.12C) for the residues 48-93 could be clearly discernible in the map while density for the 23 C-terminal residues could not be captured possibly due to flexibility. Figure 3.12D displays the local resolution calculation for cL38.

Mature cL38 possesses a pI value of 9.87 (239) suggesting a basic nature and consistently it interacts extensively with the rRNA. The N-terminus of cL38 which is more charged than the C-terminus interacts with the minor grooves of H89, H40 and H42. The two small and central α -helices are also in close association with the minor grooves of H40 and H42 (Figure 3.12B). These two α -helices are connected via a linker which mediates interaction with the 5S rRNA. The less-charged C-terminus interacts with the globular domain of bL21c and does not interact with rRNA (Figure 3.12B). Interestingly, in a previous report of the spinach cp-ribosome, authors could not identify cL38 (mentioned as PSRP6 in the concerned paper (124)) in their low-resolution map and therefore suggested that the cL38 might mediate loose association with the ribosome. The strong density for cL38 obtained in the current study suggests that the protein is an integral part of the 70S cp-ribosome. Interestingly, *psrp6* (gene for cL38) is dispensable for the cp-ribosome assembly and function, under normal growth conditions (240) and therefore, cL38 is a non-essential PSRP. However, it is probable that cL38 might be necessary when the plant faces unfavorable growth conditions. Of much interest, the modeling of cL38 revealed the presence of a surface-exposed proline-rich motif (residues 80-86). The fact that such proline-rich motifs have been reported (241) to mediate protein-protein interactions in many signaling cascades (241), hints to the possibility that cL38 might be involved in translational regulation via interaction with some protein factors.

3.11 Remodeling of rRNA structure in the LSU

While *E. coli* ribosome LSU contains two molecules of RNA namely 5S (120 nt) and 23S (2904 nt) rRNA, spinach cp-ribosomes contain three molecules of RNA namely 4.8S (106 nt), 5S (121 nt) and 23S (2810 nt) rRNA. Compared to the *E. coli* ribosome LSU which contains 3024 nucleotides of rRNA, the length of the spinach cp-ribosome LSU rRNA is 3037 nucleotides. Such conservation of sequence lengths compared to *E. coli* results from the compensation of rRNA deletion by the almost equal-sized rRNA insertion in the cp-ribosome LSU. This is evident from the global sequence alignment of 23S (Appendix Figure 1) and 5S rRNAs (Appendix Figure 2) between *E. coli* and cp-ribosome using EMBOSS needle tool (242) which reveals that they are 64.7% and 58.7% identical, respectively. Despite this conservation of sequence lengths, local

structures of the rRNA in cp-ribosome LSU are majorly altered at five different places (Figure 3.1.12). These alterations are mostly present in the 23S rRNA surrounding the positions 280-315, 939-968, 1502-1526 (chloroplast 23S rRNA numbering) and 129-153 (*E. coli* 23S rRNA numbering). The last alteration is observed in the 4.8S rRNA surrounding the position 1-28. It is important to note here that these rRNA structural alterations are exclusively located in the periphery of the ribosome and that, the peptidyl transferase center (PTC) forming core (domain V) and domain VI (Figure 3.13 and Appendix Figure 3) of the 23S rRNA remain conserved. Below is the description of all the remodeled rRNA segments.

Figure 3.13 The domain organization of the 23S rRNA in the cp-ribosome LSU. The six domains are colored and labelled as indicated. The central protuberance (CP) and L1 stalk (L1) are labelled. The color scheme is same as the 23S rRNA secondary structure diagram in Appendix Figure 3.

The first region of alteration is found in domain I (Figure 3.14A, B and Appendix Figure 4A) of 23S rRNA. This is the result of a combined 24-nucleotide long insertion (Appendix Figure 1, highlighted in yellow) that occurs between position 280-315, affecting H18 (uppercase H and lowercase h denote rRNA helices in the LSU and SSU, respectively; helix numbers are derived from the secondary structure diagram of rRNA as in http://rna.ucsc.edu/rnacenter/ribosome_images.html) and H19. The most prominent structural changes in this region are found in H16 which forms a variable loop and extends towards the solvent side of the ribosome (Figure 3.14B). The structures of H18 and H19 remain unaltered although a slight shift in the helical axis of H18 can be observed (Figure 3.14B).

Figure 3.14 Remodeled rRNA elements in the cp-ribosome LSU. (A) The LSU map is shown in grey surface and the rRNA insertion sites in 23S rRNA are highlighted in red. The numberings denote as follows: 1. 23S rRNA residues 280-315, 2. 938-969, 3. 1502-1526 and 4. 4.8S rRNA residues 25-36. (B) Region around 23S rRNA (in dark blue) residues 280-315 zoomed in to show comparison with *E. coli* ribosome model 4YBB (grey) (86). (C) Region around 938-969 zoomed in, (D) Region around 1502-1526 zoomed in, (E) rRNA deletion (H9) in chloroplastic 23S rRNA, compared with *E. coli*. (F) rRNA modifications in chloroplastic 4.8S rRNA. Cp-ribosome loses H98 in process of separating 4.8S from the original 23S rRNA. (B)-(F) are modified from (3).

The second region of alteration is found in domain II (Figure 3.14A, C and Appendix Figure 4B) of the 23S rRNA. This results from a combined 20-nucleotides long insertion between position 938-969 (Appendix Figure 1, highlighted in green) which creates a protrusion in H38. This protrusion extends towards the solvent side and

is localized close to a place where uL30 can be found in *E. coli* ribosome. In light of the fact that uL30 is absent in the cp-ribosome, it seems probable that the function of uL30 has been taken over by this rRNA protrusion. Of much interest, the bL27c C-terminal extension and cL38 (PSRP6) are located in a close neighborhood (Figure 3.14C). Further variability in this region arises from the shortening of H45 (Figure 3.14C and Appendix Figure 4B, black arrow) which is caused by a net 6 nucleotides deletion of the 23S rRNA sequence in the cp-ribosome.

The third region of alteration is located in domain III (Figure 3.14A, D and Appendix Figure 4C) of 23S rRNA which is caused by a combined 12 nucleotides insertion between 1502-1526 (Appendix Figure 1, highlighted in cyan). This results in a protrusion in H58 which is extended towards the solvent-exposed side. This insertion causes changes in the structure of the rRNA locally by causing shortening of H58 which is located towards uL2c. In *E. coli* ribosome, this region is also characterized by the presence of H63 which is truncated by 26 nucleotides in the cp-ribosome (Figure 3.14D and Appendix Figure 4D).

The above-discussed regions of remodeled rRNA arise due to rRNA insertion events. A region where structural changes occur due to rRNA deletion is situated in domain I (Figure 3.14E and Appendix Figure 4E) of the 23S rRNA. H9 has been omitted due to a combined 14 nucleotides deletion between position 129-153 (*E. coli* numbering) (Appendix Figure 1, highlighted in pink) in the chloroplastic 23S rRNA. Interestingly, the C-terminal extension of uL29c is positioned in nearby region suggesting that function of H9 might have been replaced by this protein extension.

The last region of major alterations in the rRNA structure originates due to a deletion and an insertion in the 4.8S rRNA (Figure 3.14A, F and Appendix Figure 4F). The deletion is comprised of 16 nucleotides at the 5' end of the 4.8S rRNA (Appendix Figure 5, highlighted in yellow) and results in the loss of H98 from the chloroplastic 23S rRNA. The insertion is comprised of 9 nucleotides at position 28 (Appendix Figure 5, highlighted in green) of the 4.8S rRNA (equivalent to the domain VI of *E. coli* 23S rRNA) creating a bulge between H100 and H101. This bulge resides in a close proximity of the NTE of bL19c.

Sequence analysis reveals that chloroplastic 4.8S rRNA is equivalent to the 3' tail of the *E. coli* ribosome 23S rRNA (Appendix Figure 5). It seems probable that an

ancestor of the 23S rRNA had given rise to 4.8S rRNA via cleavage of bonds to facilitate rRNA remodeling by losing H98 at this site, during evolution. Thereafter, 4.8S rRNA had co-existed alongside the new 23S rRNA molecule. This also results in shortening of the chloroplastic 23S rRNA by 94 nucleotides compared to its *E. coli* counterpart. It seems that the loss of H98 must be significant in the functioning of the cp-ribosome but no conclusion could be drawn from the current structural data.

Among the above-discussed rRNA remodeling sites, H38 and H58 are of special importance as these two helices have earlier been reported to contain Kink-turns (K-turns) as reported in the *H. marismortui* ribosome (243) (Figure 3.15). A signature of the K-turns is the presence of helix-turn-helix structure which introduces flexibility in the rRNA structure (244). K-turns can form nucleation points during ribosome assembly process and can act as the rRNA recognition motifs for ribosomal proteins. The introduction of protrusions near to these K-turn sites (as reported in the *H. marismortui* ribosome) is an important feature of the chloroplastic rRNA, the functional significance of which remains unknown (Figure 3.15).

Figure 3.15 Protrusions in cp-ribosome 23S rRNA at K-turn locations of *H. marismortui*. 23S rRNA models are overlaid from *H. marismortui* (orange) (21) and cp-ribosome to highlight presence of protrusions (blue boxes) in cp-ribosome near to sites where K-turn structures (orange circles) are present in *H. marismortui* rRNA. *E. coli* 23S rRNA model (grey) (86) is also overlaid for comparison. Insertions in cp-ribosome sequence are marked in red. Modified from (3).

3.12 Polypeptide exit tunnel and tunnel exit site in the LSU

The LSU harbors an about 100 Å long “polypeptide exit tunnel” (Figure 3.16A, B) via which the nascent polypeptide chain traverses from its origin at the peptidyl transferase center (PTC). The tunnel opens in the “tunnel exit site” (Figure 3.16C, D) to the solvent side of the ribosome where it interacts with a variety of protein biogenesis factors required for protein maturation, folding and targeting. Therefore, the tunnel exit site functions as the docking platform for these factors. The polypeptide exit tunnel in *E. coli* is mainly composed of 23S rRNA along with the proteins uL4, uL22 and uL23 while the tunnel exit site is formed by the proteins uL22, uL23, uL24 and uL29. Cp-

ribosome LSU model was overlaid with *E. coli* counterpart for comparison which revealed that they both share a common overall architecture. However, some differences are prominently visible which are discussed below.

Figure 3.16 The polypeptide exit tunnel and polypeptide exit site of cp-ribosome LSU. (A) The densities for the exit tunnel lining proteins are displayed and labelled. The simulated path of the polypeptide chain is modeled based on a previous publication (245). (B) The exit tunnel lining chloroplastic proteins are overlaid with *E. coli* counterparts from PDB model 4YBB (86) to show comparison. (C) The densities for the tunnel exit site lining proteins are displayed and labelled. (D) Tunnel exit site lining proteins are compared with *E. coli* counterpart (86). Modified from (3).

The most prominent difference is the absence of a conserved beta-hairpin loop in the protein uL23c which in *E. coli* ribosome extends into the polypeptide exit tunnel (Figures 3.16 and 3.17A). Structural changes in uL23c further involve a disordered N-terminus (Figure 3.17A) which extends towards a site where H9 is located in the *E. coli* ribosome. In comparison with *E. coli* where it is extended, the C-terminus of uL23c adopts a small helical turn (Figure 3.17A). Such structural changes in uL23c should be

studied with much interest as it has earlier been reported to harbor a universal docking site for various protein biogenesis factors like the signal recognition particle (SRP), trigger factor (TF) (246,247) and SecA (248), etc. Experiments showed that truncation of this beta-hairpin loop did not impair cell growth under normal growth conditions but it lowered the affinity of SRP binding to the ribosome-nascent chain complex (RNC). Consistently, it had been suggested in bacteria that the nascent chain interacts with this loop inducing a conformational change in uL23 at the tunnel exit site which in turn increases the affinity of SRP towards the RNC (249). As evident, uL23 is of central importance to the bacterial translation system.

Figure 3.17 3D structure and multiple sequence alignment of protein L23. (A) Structural comparison between L23 from chloroplast and *E. coli* (86) ribosomes to display deletion of β -hairpin loop in cp-ribosome. (B) Multiple sequence alignment of L23 primary sequences across different bacterial groups with spinach to display conservation of the β -hairpin loop in bacteria. Modified from (3).

Chloroplastic SRP (cpSRP) system is unique as the equivalent 4.8S RNA component of *E. coli* SRP is replaced with a protein called cpSRP43 and unlike the *E. coli* SRP which functions only co-translationally, cpSRP has been proposed to function co-translationally as well as post-translationally (250). Due to the above-mentioned structural variations in the uL23c structure, it is tempting to speculate that the mode of SRP recruitment in the chloroplastic environment might be different than the bacterial system but further experiments are required to establish this hypothesis.

Apart from uL23c, chloroplast-specific features are visible in the protein uL24c (Figure 3.16). Density for 15 among 18 amino acids of the uL24c NTE could be traced in the map which winds its way into the polypeptide tunnel lumen and seems to be strategically located to interact with the chloroplastic nascent polypeptide chain (Figure 3.16). Structural variation also arises due to the presence of 24 amino acids CTE of uL29c which runs parallel to the NTE of uL23c (Figures 3.16). In uL29c CTE, the residues spanning 130-141 assume a helical conformation while extended conformation is observed for the residues 142-149. In summary, it can be said that the visible structural variations (compared to *E. coli*) in the three (uL23c, uL24c and uL29c) proteins at the tunnel exit site might suggest a chloroplast-specific fashion of protein biogenesis factor recruitment on the cp-ribosome.

4.1 Structural insights into the small subunit of the cp-ribosome

As stated in the introduction section of the previous chapter, the dataset collected using the Arctica microscope could not be used to generate a high-resolution cryo-EM reconstruction for the SSU. To this end, another dataset was collected using the FEI Titan Krios microscope operated at 300 kV (National University of Singapore, Singapore). This dataset was of superior quality and yielded a 3.7 Å resolution map for the SSU which enabled the modeling of the complete SSU. The resolution for the LSU improved too (3.3 Å) but no new information could be discerned compared to the earlier interpretation of the 3.5 Å Arctica-map of the LSU. It is well-accepted in the scientific community that obtaining a high-resolution cryo-EM map for the SSU is difficult due to the conformational heterogeneity, as mentioned earlier. The fact that the LSU was over-represented on the cryo-grids, further worsened the situation. To mitigate this, 300-kV Krios microscope was used as electrons accelerated at 300 kV are more coherent than those at 200 kV (Arctica). Consequently, images recorded using the Krios microscope are less noisy compared to the images recorded using the Arctica microscope, which results in a better 3D reconstruction.

4.2 Image processing

The steps taken during image processing of the Krios-dataset are depicted in Figure 4.1 below. Like the procedure followed before, during the processing of Arctica-dataset (Section 3.4), the particles were picked semi-automatically in EMAN 2.1 (1) and thereafter processed by RELION 1.4 (2). As the data quality was better compared to the previous Arctica-dataset (Section 3.4), the computational time spent to process this data was much lesser than before. Also, the number of particles needed to generate the high-resolution structure at the final refinement step is reduced to almost half (Figure 4.1) than what was needed previously (Section 3.4). Compared to 174K particles in final refinement step of the Arctica-dataset, only 81K particles were used for the Krios-dataset.

Figure 4.1 Image processing flowchart for the Krios-dataset. 3161 micrographs were collected, 1285 turned out to be of suitable quality with minimal drift and astigmatism. From these, 188K particles were semi-automatically picked using e2boxer.py of EMAN 2.1. Data processing was done using RELION 1.4. 2X-binned particles were sorted by 2D classification providing 133K good particles which were subjected to 3D classification using a cp-ribosome 70S map as reference. 3D classification sorted out the bad and 50S particles from the dataset. Thereafter, 81K good 70S particles were refined using unbinned particles, producing a 3.66 Å-resolved 70S map which yielded a 3.36 Å-resolved map after inclusion of movie frames in subsequent 3D refinement runs. Application of Gaussian masks on the ribosomal subunits rendered maps of the LSU and SSU at 3.3 and 3.7 Å, respectively. Taken from (113).

4.3 Resolution estimation for the SSU, LSU and 70S ribosome from Krios-dataset

The cryo-EM reconstruction obtained for the SSU from the Krios-dataset was resolved to an average 3.7 Å as reported by the FSC=0.143 criterion (Figure 4.2A, B and C) (214). As expected, local resolution calculation using ResMap (232) showed that the core of the SSU was better resolved (3- 3.5 Å) than the periphery (> 5 Å) (Figure 4.2D, E and F). Components identified in the low-resolution regions consisted of the protein cS22 at the SSU foot, the beak and part of the head region, and the platform region of the SSU. The reasons behind obtaining low-resolution in these regions might be the high flexibility of cS22 structure and head swiveling motion of the SSU. A few examples are shown in Figure 4.2G and H which show the well-resolved features in the SSU map.

Figure 4.2 Resolution estimation of cp-ribosome maps derived from Krios-dataset. (A), (B) and (C) Overall resolution calculation for masked 70S, 50S and 30S maps according to FSC=0.143 and FSC=0.5 criterions. The former criterion is used to report FSC between cryo-EM half maps generated at the last stage of refinement showing maps are not overfitted. The later criterion is used to calculate map-model FSC to show that model is not overfitted. (D), (E) and (F) Local resolution calculation for masked 70S, 50S and 30S maps using ResMap (232). (G) and (H) Quality of 30S map as displayed by the side chain densities for some amino acids in an α -helix from uS3c, separation of strands in β -sheet (uS3c) and nucleotide bases (around C389 in 16S rRNA). (A)-(F) taken from (113).

4.4 Model validation for the SSU, LSU and 70S ribosome from the Krios-dataset

The atomic models of the SSU, LSU and 70S cp-ribosome were validated using the Molprobity server (168). The model validation and refinement statistics are listed in Table 4.1. Model overfitting was estimated by calculating FSC curves (Figure 4.2A, B and C) between the maps and derived models. Good agreement between these FSCs and the FSCs calculated between the cryo-EM half maps (Figure 4.2A, B and C) shows that the generated models do not contain any alarming level of overfitting and are therefore useful for a reliable interpretation. The models were further validated by generating colored maps based on the local agreement between map and model as shown in Figure 4.3.

Table 4.1. Data collection and refinement statistics. Taken from (113).

Data Collection	70S ribosome	30S subunit	50S subunit
Particles	81305	81305	81305
Pixel Size (Å)	1.05	1.05	1.05
Defocus Range (μM)	0.4-3.7	0.4-3.7	0.4-3.7
Voltage (kV)	300	300	300
Electron Dose (e ⁻ Å ⁻²)	37.5	37.5	37.5

Model composition			
Non-hydrogen atoms	146,970	54,641	92,329
Protein residues	6,167	2,859	3308
RNA bases	4,496	1478	3018

Refinement			
Resolution (Å)	3.36	3.73	3.30
Map sharpening B-factor (Å ²)	-72.6	-97.9	-89.6
CC _{map-model} (whole unit cell)	0.775	0.778	0.713

RMS deviation			
Bonds (Å)	0.01	0.02	0.01
Angles (°)	1.31	1.54	1.17

Validation (proteins)			
Molprobity score	2.59	2.86	2.16
Clashscore, all atoms	8.58	13.01	6.15
Poor rotamers (%)	1.26	1.54	1.02
Favoured rotamers (%)	94.78	94.95	94.63

Ramachandran plot			
Favored (%)	85.97	85.76	86.17
Outliers (%)	0.85	1.61	0.18

Validation (RNA)			
Correct sugar puckers (%)	99.78	99.80	99.77
Good backbone conformations (%)	77.48	78.56	76.95

Figure 4.3 Local agreement between cp-ribosome maps and models. The masked 70S, 50S and 30S maps are colored based on local correlation values calculated with respect to fitted models. The good correlation values denote good model quality without overfitting. Taken from (113).

4.5 Overall architecture of the cp-ribosome SSU

The overall architecture of the cp-ribosome SSU (Figure 4.4) is comparable to its *E. coli* counterpart. The cp-ribosome SSU contains one molecule of 16S rRNA which is 1491 nucleotides in length. Compared to the *E. coli* 16S rRNA which is 1542 nucleotides long, the chloroplastic 16S rRNA is reduced in length due to rRNA deletion at three locations. The cp-ribosome SSU also contains all 21 *E. coli*-orthologous proteins and additionally contains three PSRPs (cS22, cS23 and bTHXc) (Figures 4.4 and 4.5). Half of these proteins are encoded by the plastome while the remaining half are encoded by the nuclear genome. The twelve proteins encoded by the plastome include uS2c, uS3c, uS4c, uS7c, uS8c, uS11c, uS12c, uS14c, uS15c, bS16c, bS18c and uS19c. The remaining twelve proteins encoded by the nuclear genome include bS1c, uS5c, bS6c, uS9c, uS10c, uS13c, uS17c, bS20c, bS21c, cS22, cS23, bTHXc and are post-translationally imported inside the chloroplast.

As discussed in detail in the earlier chapter on the LSU, the high-resolution *E. coli* model bearing PDB ID 4YBB (86) was overlaid to localize and model the SSU proteins and RNA components. After localizing all *E. coli*-homologous components, clear masses of densities remained unaccounted at four places which were assigned to the three PSRPs (cS22, cS23 and bTHXc) and factor plastid pY (Figure 4.4). In summary, 21 PRPs and some PRP-extensions, 3 PSRPs and 16S rRNA formed the chloroplastic ribosome SSU which agrees with the mass spectrometry results reported by Yamaguchi and coworkers (120). Presence of the factor pY is consistent with earlier reported cp-ribosome purification from spinach (124). A list of all the components in the SSU is presented in Table 4.2.

Figure 4.4 Overall architecture of the 70S cp-ribosome and SSU. (A) The 70S cp-ribosome map is displayed with the major structural landmarks labelled. The PSRPs are colored red and the factor plastid pY located at the intersubunit space is colored blue. Other components are colored as indicated. (B) The masked refined map for the SSU showing the position of all the bacterial-homologous proteins and the three PSRPs of SSU (cS22, cS23 and bTHXc). (C) Atomic model for the density presented in (B). Taken from (113).

Figure 4.5 Constituent proteins of cp-ribosome SSU. On the left, isolated densities for the cp-ribosome SSU proteins are in blue colored meshes; fitted models are in dark red. On the right, the chloroplastic protein models are overlaid with corresponding homologous *E. coli* proteins (grey) for comparison. Density for cS22 is filtered to 8 Å for better visualization of the connecting density between the two RNA recognition motif (RRM) domains. Plastid-specific protein extensions are colored green. Taken from (113).

Table 4.2. Constituent elements of the cp-ribosome SSU and plastid pY. Mature protein lengths are taken from UniProtKB/Swiss-Prot (234). Annotation in UniProtKB/Swiss-Prot is based on the sequencing results from (120). The nomenclature of the ribosomal proteins is based on Ban *et al.* (238). Modified from (113).

Protein/ RNA name	Old Name	Chain ID	Uniprot ID (GenBank ID)	Mature protein (range, aa)	Built residues (range, aa)	Size of homologous <i>E. coli</i> proteins (aa)	Comments
bS1c	RPS1	8	P29344	370	49-74, 249-377	557	Poor density. N, C- terminal extensions, OB domain rigid body fitted
uS2c	RPS2	b	P08242	236	5-231	241	
uS3c	RPS3	c	P09595	217	3-215	233	
uS4c	RPS4	d	P13788	200	2-197	206	
uS5c	RPS5	e	Q9ST69	253	138-308	167	
bS6c	RPS6	f	P82403	146	97-207	131	
uS7c	RPS7	g	P82129	154	4-152	179	
uS8c	RPS8	h	P09597	133	1-134	130	
uS9c	RPS9	i	P82278	157	63-195	130	
uS10c	RPS10	j	P82162	122	98-195	103	
uS11c	RPS11	k	P06506	137	21-138	129	
uS12c	RPS12	l	P62128	122	1-123	124	
uS13c	RPS13	m	P82163	126	32-141	118	
uS14c	RPS14	n	P06507	99	2-100	101	
uS15c	RPS15	o	Q9M3I4	90	21-82	89	
bS16c	RPS16	p	P28807	88	1-80	82	
uS17c	RPS17	q	P82137	108	58-135	84	
bS18c	RPS18	r	Q9M3K7	101	21-84	75	
uS19c	RPS19	s	P06508	91	7-84	92	
bS20c	RPS20	t	P82130	108	76-180	87	
bS21c	RPS21	u	P82024	137	95-138	71	
bTHXc	PSRP4	x	P47910	47			
cS22	PSRP2	v	P82277	198	71-260		RRM domains rigid body fitted. Helical linker added thereafter
cS23	PSRP3	w	P82412	121	96-179		Core fold rigid body fitted
plastid pY	PSRP1	y	P19954	236	70-177	113	
16S rRNA		a	AJ400848. 1	1491 nucleotides	1-1488		

4.6 Localization and modeling of cS22 (PSRP2) at the SSU foot

In the cryo-EM reconstruction of the SSU at an average 3.7 Å, a highly fragmented density was observed in the foot region (Figure 4.6). Gaussian filtering (8 Å) was applied to visualize low-resolution features of this density, revealing that the density is bilobular. This density was assigned to cS22 (PSRP2), since the generated model for cS22 using the I-TASSER server (170) could be completely accommodated inside the density. This location of cS22 at the SSU foot is consistent with the report of the cp-ribosome by Sharma *et al.* (124).

Figure 4.6 Localization and modeling of cS22. (A) At the left, SSU model fitted in 3.7 Å SSU map. On the right, SSU foot is zoomed in to show rigid body fitting of cS22 model into filtered density. (B) I-TASSER (170) generated model of cS22. (C) Charge distribution on cS22 surface, displaying the interaction between negatively charged rRNA and positively charged patch of cS22. Taken from (113).

Figure 4.6B displays the I-TASSER (170) server-generated model which contains two globular RNA recognition motif (RRM) domains. These two domains were first independently rigid-body fitted in the two lobes of the filtered cS22 density. Subsequently, the interdomain linker was placed in this density. The filtered map of the SSU was further analyzed to figure out the stabilization of cS22 at this location. Although not conclusive from the low-resolution cS22 density obtained in the current study, it is possible that the flipped-out bases in the hairpin loop of h6 and h10 stack with the aromatic amino acids of cS22. Further stabilization might have originated from the electrostatic attraction between the negatively charged rRNA and the basic patches of cS22 (Figure 4.6C). It is interesting to note that, cS22 is the only

candidate among the five PSRPs of the cp-ribosome which contains RRM domains (251). The RRM domains are ubiquitous in eukaryotes and possess an $\alpha\beta$ -sandwich structure which folds in $\beta 1\alpha 1\beta 2\beta 3\alpha 2\beta 4$ manner. They can bind a range of RNA, DNA and protein sequences. RRM domains also demonstrate high plasticity and therefore render the RRM-domain containing proteins with specificity and high RNA-binding affinity (252). Analysis of the filtered SSU map reveals that the two RRM domains of cS22 do not interact with each other and therefore, cS22 seems to adopt an open conformation while bound to the ribosome (Figure 4.6A).

Figure 4.7 rRNA deletion in cp-ribosome SSU (16S rRNA). On the left, 16S rRNA from cp- (cyan) and *E. coli* (grey) (86) ribosomes are overlaid to show comparison. On the right, the SSU foot is zoomed in for better visualization of the rRNA helices (h6, h10 and h17) which are affected by deletion. Taken from (113).

In comparison with the *E. coli* counterpart, chloroplastic 16S rRNA contains three deletions which are in h6, h10 and h17 (Figure 4.7, Appendix Figures 6 and 7). Presence of cS22 at a nearby location seems to structurally compensate for the loss of h6 and h10 and is responsible for an increased protein mass at the cp-ribosome SSU foot (Figure 4.6A). However, the importance of such structural remodeling remains unknown. Interestingly, under standard greenhouse condition, knockdown of *psrp2* (gene for cS22) does not affect ribosome function and plant growth, as reported by Tiller and coworkers (240). Thus, cS22 seems to be a dispensable PSRP under green-house conditions and therefore, its function remains unknown (240).

4.7 Localization and modeling of bTHXc (PSRP4) at the SSU head

The density for bTHXc (PSRP4) could be located at the SSU head region (Figure 4.8). This protein is highly basic as it possesses a pI value of 11.14 and therefore is surrounded by rRNA elements. This localization of bTHXc is consistent with the position of the homolog bTHX in *T. thermophilus* ribosome (236) and also agrees with the findings of Sharma *et al.* (124) in the earlier reported low-resolution model of the spinach cp-ribosome.

Figure 4.8 Localization and modeling of bTHXc. (A) Cp-ribosome SSU model showing the localization of bTHXc (red) at the SSU head. 16S rRNA is colored sky blue. Ribosomal proteins except bTHXc-neighboring ones are colored golden. (B) Zoomed in view of SSU head from (A) to show rRNA buried location of bTHXc. (C) *T. thermophilus* (grey) ribosome model (83) overlaid to show that bTHXc is extended compared to bTHX. Taken from (113).

A PSI-BLASTp search using the full-length bTHXc sequence against the Non-redundant protein database reveals a wide distribution of bTHXc-homologs across multiple organisms like cyanobacteria, gamma proteobacteria and Cytophaga-Flavobacterium-Bacteroides (CFB) group of bacteria and plants. Figure 4.9 shows the multiple sequence alignment and phylogenetic tree of such distribution. Curiously, similar search against the model organism *E. coli* does not reveal any homologs which is corroborated by the structural study on *E. coli* ribosome (86).

Figure 4.9 Multiple sequence alignment (MSA) and phylogenetic tree of selected bTHXc-homologs. (A) MSA generated from selected PSI-BLASTp hits against the bTHXc preprotein sequence, using default options in MultAlin server (253). (B) Distribution of bTHXc-homologs across different organisms is shown by a phylogenetic tree generated using the MSA as produced by the MultAlin server (253).

bTHXc is synthesized by the nuclear genome as 101 amino acids long protein which after the signal peptide removal, generates 47 amino acids long mature protein inside the chloroplast. bTHX in *T. thermophilus* ribosome is only 27 amino acids long and the increased length in the mature bTHXc is due to the chloroplast-specific CTE (Figure 4.8). However, bTHXc and bTHX share a common fold as evident from the superimposition of the two models as displayed in Figure 4.8C. This fold extensively interacts and stabilizes 16S rRNA helices h41 and h42. The CTE of bTHXc further mediates hydrophobic contacts with uS13c. Interestingly, uS13c is the only protein which is situated in a close vicinity of bTHXc and that uS13c protein is a major component of two intersubunit bridges B1a and B1b in the 70S cp-ribosome. It seems that bTHXc is strategically located to provide stabilization to the bridge B1b and thus helps to maintain an intact 70S ribosome under conditions of thermal fluctuations that are routinely experienced by plants. This hypothesis makes sense given the fact that, homologs for bTHXc are only found in ribosomes of the *Thermus* bacteria who experience extreme temperatures in their habitats. The importance of bTHXc is further corroborated by the fact that knockdown of *psrp4* (gene for bTHXc) in *Arabidopsis* causes impaired growth by reducing photosynthesis (240) and therefore, bTHXc is an essential PSRP.

4.8 Unique features of the mRNA entry and exit sites

On the intersubunit side of the SSU, mRNA binds to a narrow constriction in the neck region between the head and body of the SSU, which is known as the mRNA channel. In the intact 70S ribosome, this channel is connected to two openings on the solvent-exposed side of the SSU: (a) the mRNA entry site, situated close to the beak region and (b) the mRNA exit site, located near the platform region (Figure 4.10).

Figure 4.10 The mRNA entry and exit sites. (A) Comparative location of mRNA entry and exit sites on the SSU from the solvent exposed side. Density for SSU is shown in transparent grey. Fitted SSU model is in golden. (B) mRNA entry site zoomed in to show the constituent proteins and protein extensions. (C) Overall architecture of mRNA entry sites from bacteria (PDB ID 4YBB) (86), cp- and mitochondrial ribosomes (PDB ID 5AJ3) (34) showing source-specific features. (D) Zoomed-in representation of mRNA entry sites from (C). The entry sites are marked with an asterisk and neighboring proteins are labelled. Taken from (113).

The mRNA entry site is formed by the proteins uS3c, uS4c and uS5c which is consistent with the bacterial ribosome. The facts that proteins S3 and S4 possess RNA-helicase activity and S5 is involved in maintaining translation fidelity, suggest that these three proteins play important role in facilitating translation (254). Compared to the *E. coli* counterpart, uS5c has acquired an 86-amino acid long NTE while the uS3c and uS4c remain mostly conserved. Therefore, a richer protein mass content and the resulting constriction is evident for the mRNA entry site of the cp-

ribosome in comparison with *E. coli* (Figure 4.10C, D). However, a major part of the uS5c NTE remained indiscernible in the current map which suggests that it is flexible in nature.

The mRNA exit site (Figure 4.11A) which is located near to the platform region is an important site on the SSU as it harbors the so-called Shine-Dalgarno (SD):anti-SD interaction between the mRNA and 16S rRNA in bacterial ribosomes (Figure 4.11B). SD:anti-SD interaction helps in the correct positioning of the mRNA reading frame for translation. In other words, it ensures that the start codon AUG is placed at the ribosomal P-site. Unlike the bacterial ribosomes where the SD sequences are invariably found around -7 to -4 nucleotides of the start codon (255), only two-thirds of the chloroplastic mRNAs contain a SD sequence and their positions are variable (156). It is therefore not certain whether SD:anti-SD interactions are necessary for translation of the chloroplastic mRNAs. Whatever may be the case, a solid mass of density at the mRNA exit site was observed in the cp-ribosome SSU map. The density was surface-exposed and positioned between a cleft formed by the neck (solvent side) and platform of the SSU. At this point of the model building, all other PRPs, PRP extensions, PSRPs and rRNAs were already assigned to their respective densities and therefore this new density was assigned to cS23 (PSRP3) (Figure 4.11A). cS23 was the last among the three PSRPs to be located and modeled in the cp-ribosome SSU. Although the cS23 density was sufficiently big to accommodate the core fold of cS23, the poor resolution obtained for this density allowed only rigid-body docking of the model and restricted any molecular interpretation of the contacts of cS23 with other neighboring elements of the ribosome. It is imperative to mention here that, apart from the structural data reported in this current study, no biochemical data have been generated to validate the claim that this density arises from cS23. However, assigning the density to cS23 seems logical as mass spectrometry studies had reported the presence of only five PSRPs (PSRP1 is now known as plastid pY which binds the ribosome at hibernating state) and that, apart from plastid pY, no other protein factor has been reported so far to co-purify with the 70S cp-ribosome (124,256).

Figure 4.11 The mRNA exit site. (A) The constituent proteins and rRNA of the mRNA exit site in cp-ribosome are displayed. Presence of cS23 at this location is unique feature of cp-ribosome. (B) SD:antiSD interaction at the mRNA exit site is displayed overlaying bacterial ribosome model (PDB ID 4V6F) (grey) (82) with cp-ribosome to show closeness of cS23 to SD:antiSD site. (C) Shifted positions of bS21c (cyan) and 3' end of cp-ribosome 16S rRNA (pink) in comparison with *E. coli* (grey) ribosome SSU (PDB ID: 5AFI) (257). (D) Overall architecture of mRNA exit sites from bacteria (PDB ID 4YBB) (86), cp- and mitochondrial ribosomes (PDB ID 5AJ3) (34) showing source-specific features. (E) Zoomed-in representation of mRNA exit sites from (D). The exit sites are marked with asterisks and neighboring proteins are labelled. Taken from (113).

Apart from the presence of cS23, there are other visible changes at the cp-ribosome mRNA exit site, compared to the *E. coli* ribosome. A detailed study of these structural alteration at this site is important to understand the translation process in the chloroplast since there is an increasing amount of literature claiming that translation initiation is far more complicated than anticipated (134,140,258). As evident from the discussion on cS23, the mRNA exit site in the cp-ribosome is protein-rich compared to *E. coli* (Figure 4.11D, E). In the cp-ribosome, the mRNA exit site is formed by bS1c, uS7c, uS11c, bS21c, cS23 and 16S rRNA helix h28 (Figure 4.11A). Compared to its *E. coli* counterpart, bS21c is longer by 66 amino acids due to NTE/CTE, but density for the extension is not visible due to flexibility. Unique features of the cp-ribosome mRNA exit site include repositioning of the bS21c C-terminal helix ranging from amino acids 123-136 which mediates extensive interaction with the 16S rRNA 3' end (Figure 4.11C). Thus, it appears that bS21c locks and stabilizes the 16S rRNA 3' end and therefore does not allow 16S rRNA to participate in SD:anti-SD interaction.

Considering the function of cS23, it is interesting to note that knockdown Arabidopsis mutants of *psrp3* (gene for cS23) resulted in defective chloroplast translation and therefore caused impaired growth and reduced photosynthesis, thus making cS23 an essential PSRP (240). Because cS23 is positioned at the SSU platform, it is possible that cS23 acts as a scaffold for the recruitment of trans-acting protein factors to mediate efficient translation of certain mRNAs, especially which lack or possess weak SD sequences at the canonical position. cS23 has also been suggested to have a moon-lighting function in gene expression and signaling pathways in plants (240).

4.9 An extended bS1c connects the mRNA entry and exit sites on the SSU

Consistent with bacterial ribosomes, bS1c is the largest protein in the cp-ribosome. It can bind mRNA while associated to the ribosome or in free-form (259,260). The bacterial homolog S1 is 557 amino acids long and contains six Oligonucleotide/oligosaccharide-binding-fold (OB) domains. In comparison, the truncated spinach bS1c (Figure 4.12) contains only 3 OB-domains spanning a sequence length of 411 amino acids.

Figure 4.12 Localization and modeling of bS1c. bS1c assumes an extended structure on the cp-ribosome SSU and connects the mRNA entry and exit sites. Taken from (113).

S1 plays an essential role in translation initiation as it catches mRNA and facilitates its placement in the mRNA channel, but the precise molecular mechanism remains unclear (261-265). In bacteria, S1 has been reported to interact with the SSU in an on-and-off manner. Bacterial S1 is particularly essential for the translation of leaderless or weak SD-containing mRNAs (262,264,266). A diverse array of literature describe the mode of interaction between the bacterial S1 and SSU, according to which, the S1 NTE binds the SSU while the long and flexible C-terminus binds the 5' UTR of mRNA due to its high affinity toward the pyrimidine sequences (267,268). It has been further reported that a binary complex between S1 and the SSU (259,264) is recruited to the 5' UTR of mRNA, which then begins translation. Unlike bacterial S1 which can direct the 30S complex for translation initiation, chloroplastic bS1c needs help from the gene-specific nuclear-encoded translation factors to initiate translation (268). Three such mediator proteins, which are also light-regulated, have been reported in the *Chlamydomonas reinhardtii* chloroplast (269).

Despite the biochemical evidence behind the importance of S1 in translation initiation, the structure remains unresolved till date. In the cp-ribosome SSU map, partial density for bS1c could be located after filtering the map to low-resolution (8 Å) (Figure 4.12). In this density, one of the three OB-domains of bS1c and the chloroplast-specific N- and C-terminal extensions were rigid-body fitted. From the density, it is evident that bS1c reaches to both the mRNA entry and exit sites on the cp-ribosome SSU. This finding agrees with bacterial S1 which is reported to have an elongated structure on the bacterial ribosome SSU (270,271). Bacterial S1 is highly flexible on the SSU and therefore, it is very difficult to capture S1 density on the bacterial ribosome. Because capturing partial bS1c density was possible in the current study, it seems that the interaction between bS1c with cp-ribosome SSU is stronger than its bacterial counterpart. A report shows that most of the *C. reinhardtii* S1 (CreS1) remains associated with the SSU during the entire period of translation (259). Therefore, it is tempting to speculate that, in spinach, bS1c remains associated with the ribosome at all times. Whatever might be the case, the low-resolution bS1c density obtained in the current study provides insights into bS1c interaction with the SSU. uS2c forms the major interaction surface, while cS23, uS5c and uS8c are the other contributing partners for bS1c-stabilization on the ribosome (Figure 4.12). Overall, the N- and C-terminal extensions and one of the three OB-domains are responsible for ribosome binding. It is possible that the other two OB-domains for which the density could not be discerned due to flexibility, are involved in the mRNA binding. It is also well-known that OB-domains are functionally diverse as they can bind to a range of substrates like DNA, RNA and proteins (272,273). This could mean that OB-domains are also capable of forming the launching pads for translation factors, to facilitate translation initiation by mediating protein-protein and protein-mRNA interactions. Presence of cS23 in the close vicinity could mean that cS23 and bS1c might act together in this respect, although experiments would be required to validate this hypothesis. Interestingly, bacterial S1 has been reported to possess RNA helicase activity (263,274-277). It is, therefore, possible that chloroplastic bS1c might also possess a similar activity which might help in the destabilization of the secondary structural elements of the 5' UTR in the chloroplastic mRNAs to facilitate translation initiation.

4.10 Mode of interaction of plastid pY with the ribosome

The factor plastid pY was co-purified with the ribosome preparation in the current study (Figure 4.13A, B). This is consistent with the earlier purifications of the spinach 70S cp-ribosome (124,256). The density for plastid pY is located at the subunit interface overlapping the A- and P-sites of the mRNA channel and the binding site is contributed solely by the SSU (Figure 4.13A, B and C). This agrees with the binding mode of plastid pY-homolog, YfiA to the *T. thermophilus* 70S ribosome (236). Therefore, due to steric clash, plastid pY precludes binding of mRNA and tRNAs and leads to a translationally inactive hibernating state of the chloroplast ribosome.

Interestingly, mainly due to the close association of plastid pY with ribosome, plastid pY was earlier considered as a *bona fide* chloroplast ribosomal protein named as PSRP1 (120). But, a biochemical study revealed that under both *in vitro* and *in vivo* conditions, plastid pY interacts with the *E. coli* 70S ribosome at the same location (mRNA channel) and stabilizes the ribosome from dissociation by inducing a conformational change within several intersubunit bridges (256). This indeed suggests that plastid pY is a functional homolog of the bacterial cold shock protein pY (also known as YfiA, RaiA and SpotY). However, inside the chloroplast, it is possible that the role of plastid pY might not be limited to just promoting ribosome hibernation during stressed conditions, as anti-miscoding activity and alternative functions like promoting post-stress survival under normal growth condition have been proposed for plastid pY-homologs in *E. coli* (278) and *Synechocystis* (279), respectively.

Figure 4.13 Localization and modeling of the factor, plastid pY. (A) Position of plastid pY at the intersubunit space is indicated with a black box. The subunits of ribosome are labelled. SSU proteins are colored golden rod, 16S rRNA is colored sky blue and LSU rRNAs (4.8S, 5S and 23S) are in green. Plastid pY is colored rosy brown. (B) Enlarged view of plastid pY's environment to show the factor's exclusive association with the SSU. The orientation of the constituent α -helices and β -sheets are indicated. (C) Binding of plastid pY at mRNA channel excludes binding of A- and P-site tRNAs and therefore halts translation. Models for the tRNAs and mRNA are derived from an *E. coli* ribosome bearing PDB ID 5AFI (257). (D) Local resolution of plastid pY density as derived using ResMap (232). Modified from (113).

Plastid pY, a 302 amino acids long protein, is encoded by the nuclear-genome. Entry for this protein, bearing the ID P19954, can be found at UniProtKB (234). The first 66 residues constitute the signal peptide and therefore the mature protein inside the chloroplast is 236 amino acids long. The homology model generated using the I-TASSER server (170) reveals a two-domain architecture for plastid pY. The structure of the NTD is conserved (252,280) and folds in a $\beta 1\alpha 1\beta 2\beta 3\alpha 2\beta 4$ fashion (Figure 4.13B). In the current study, a well-resolved density for the NTD could be captured. The local resolution of plastid pY density was calculated using ResMap (232) and, is shown in Figure 4.13D. The density of the CTD could not be captured possibly due to flexibility. Therefore, the structure and function of CTD remain elusive although the interactions between the NTD and SSU could be characterized revealing insights about the hibernation state of the cp-ribosome.

Figure 4.14 Interaction of plastid pY with the cp-ribosome SSU. Plastid pY rRNA residues that participate in binding SSU are displayed in ball-and-stick and labelled in green. 16S rRNA is in sky blue; only those nucleotide bases which directly interact with plastid pY are labelled (red). Interaction with plastid pY is also mediated by sugar-phosphate backbone of 16S rRNA but labels are hidden for clarity. Taken from (113).

The mode of plastid pY interaction with the SSU was analyzed in detail. Helix $\alpha 1$ of plastid pY positions itself in front of the LSU rRNA H69 and the helix $\alpha 2$ is situated above the mRNA channel, while the four beta strands are oriented toward the SSU head (Figure 4.13A, B). As evident from the fitted model, the interaction of $\alpha 2$ with the SSU is more extensive than $\alpha 1$ (Figure 4.13B). While only two positively charged residues of $\alpha 1$ (Lys31 and Lys38) interact with the negatively charged 16S rRNA backbone, multiple polar and charged residues in $\alpha 2$ (Ser89, Asp92, Lys101, Arg103, Lys104, Lys106 and Asp107) mediate interaction with the rRNA backbone (Figure 4.14). Interaction with the rRNA backbone is further strengthened by the beta strand residues Lys10, Gln14, Arg49, Glu67 and Arg79 (Figure 4.14). It is of interest that the 16S rRNA bases G641, A738, G875, C1003, A1287, A1288 and G1454 mediate interactions with both acidic and basic residues of plastid pY while G478 (G530 in *E. coli*) is stabilized by the stacking interaction partner Pro127 (Figure 4.14).

However, chloroplastic G478 (G530 in *E. coli*) and C1003 (C1054) maintain similar orientations compared to their bacterial counterparts. A small displacement in the position of the base G478 is observed (Figure 4.15A) in the cp-ribosome structure, which is caused probably due to the complete head-swiveling motion of the SSU.

Another interesting feature of the plastid pY structure is a longer loop between beta strands $\beta 2$ and $\beta 3$, which gets stabilized by interacting with h18 and h34 (Figure 4.15B). Figure 4.15C shows the alignment between bacterial and plastid pY's primary sequences. It is evident that the origin of this loop is due to sequence insertion in plastid pY, but why such modification became necessary during the evolution of the chloroplast, remains elusive. In all, the presence of plastid pY at the mRNA channel of the SSU solves two purposes during unfavorable conditions: (a) it prevents the binding of mRNA and tRNA ligands (A- and P sites) and, (b) it locks the ribosome in the so-called hibernating state to render it translationally inactive. When the environment becomes conducive to translation, plastid pY might get kicked out of the ribosome by elongation factor G (EF-G) and ribosome recycling factor (RRF) but the mechanism has remained unknown.

4.11 Intersubunit bridges in the 70S cp-ribosome

The intersubunit bridges (Figure 4.16) were thereafter studied in the 70S cp-ribosome which revealed their conserved nature compared to bacterial ribosomes. The major types of interactions are RNA-RNA, followed by RNA-protein and this is consistent with the *E. coli* ribosome (Figure 4.16).

Figure 4.16 Intersubunit bridges in the 70S cp-ribosome. (A) Subunit interface view of the ribosomal subunits to show the rRNA and protein residues which contribute in intersubunit bridge formation. Brown and pink spheres represent contributing SSU and LSU elements, respectively. 16S rRNA and the SSU proteins (left) are colored in sky blue and golden rod, respectively. Plastid pY is shown in rosy brown. 23S rRNA and LSU proteins (right) are colored green and golden rod, respectively. Taken from (113).

Table 4.3 lists all the intersubunit bridges which are formed in the cp-ribosome. The densities for the bridges are displayed in Figure 4.17. It is of interest that, bridge B7b is not observed in the 70S cp-ribosome structure. This bridge is conserved in bacterial ribosomes. Mutation studies have shown that B7b is important in maintaining fidelity of translation initiation in bacteria (282).

Table 4.3. Components of the intersubunit bridges in spinach 70S cp-ribosome. Taken from (113).

Bridge	SSU component	LSU component	Type
B1a	uS13c: Ile111, Glu112, Arg113	23S: H38 (A-site finger): 896-900	R-P
B1b	uS13c: Ile37	uL5c: Arg163, Thr164	P-P
	uS13c: Glu96, Val97	uL5c: Arg165, Asp166	P-P
	uS13c: Arg113	bL31c: C-terminus*	P-P
B1c	uS14c: Asp39, Lys40	bL31c: C-terminus*	P-P
	uS19c: Pro9, Pro42	bL31c: C-terminus*	P-P
	16S: h42 (A1259, G1260)	bL31c: C-terminus*	R-P

B2a/d	16S: h44 (U1444)	23S: H69 (A1926)	R-R
	16S: h44 (C1445)	23S: H69 (C1934)	R-R
	16S: h44 (C1356, A1357)	23S: H69 (A1932, A1933)	R-R
	16S: h45 (G1466)	23S: H69 (A1933)	R-R
	16S: h24 (739)	23S: H69 (G1935, G1936)	R-R
B2b	16S: h24 (G732)	23S: H68 (C1846, C1847)	R-R
B2c	16S: h27 (C848, A849)	23S: G1841, C1842	R-R
B3	16S: h44 (C1369, U1370, G1371, G1372)	uL14c: Pro48, Asn49, Arg54	R-P
	16S: h44 (G1431, A1432, C1433, G1366, A1367, G1368)	23S: H71 (C1961, G1962, G1963, C1972, G1973, A1974)	R-R
B4	uS15c: Ile33, Thr37, Leu40, Glu41, Gln50, Leu53, Arg54, Leu57	23S: H34 (U725, G726, A727)	R-P
	16S: h20 (G710)	23S: H34 (A729)	R-R
B5	16S: h44 (U1376, G1377, C1378, A1422, G1423, G1424)	23S: H62 (C1722, G1723, A1725, A1737, G1738, G1739, G1740)	R-R
B6	16S: h44 (C1380, G1381)	bL19c: Arg228	R-P
B7a	16S: h23 (A650)	23S: H68 (G1856, A1857, A1858, C1908, C1909)	R-R
B7b	Not observed in our cryo-EM reconstruction		
B7c	bS6c: Lys161**, Ala162**	uL2c: Pro132, Leu133, Ala162, Lys163	P-P
B8	16S: h14 (A315, C316, G317)	bL19c: Lys159, Arg160	R-P
	16S: h14 (G308, A309)	uL14c: Arg97	R-P

Abbreviations: R-R=RNA-RNA, R-P=RNA-protein, P-P=Protein-protein

The absence of bridge B7b is compensated by the presence of a novel bridging interaction between the chloroplast-specific loop of bS6c (Lys161 and Ala162) and uL2c (Pro132, Leu133, Ala162 and Lys163) (Figure 4.17M, N). This novel bridge is named as B7c (Figure 4.17M, N). The absence of B7c in bacteria is due to truncation in the chloroplast-specific loop of bS6c.

Figure 4.17 The densities for intersubunit bridges in 70S cp-ribosome. (A)-(L) Conserved intersubunit bridges between cp- and bacterial ribosome are displayed to show constituent proteins and rRNAs. Densities (grey surface) are filtered for clarity. (M) Density for the novel intersubunit bridge (B7c) in 70S cp-ribosome. The density is 8 Å-filtered, neighboring proteins and rRNA hidden, for better visualization. (N) Bridge B7c forming proteins of the cp-ribosome and their counterparts in *E. coli* (PDB ID: 4YBB) (86) are overlaid to show absence of bridge-forming loop in bacterial S6 protein. Taken from (113).

5. Discussion

The aim of this study was to elucidate a high-resolution structure of the cp-ribosome from spinach. The first dataset collected using the 200-kV Arctica microscope had provided a 3.5 Å structure of the LSU which was subsequently used for atomic model building. It is to be noted that, cryo-EM provides a map which demonstrates variable resolution throughout its volume. Therefore, the interpretation of the map was performed after determining the local resolution using a computational program. Using the Arctica-dataset, localization and modeling of cL37 and cL38 (LSU-residing PSRPs), chloroplast-specific extensions in PRPs and remodeled rRNA could be possible. A number of insights could be obtained from the LSU structure. For example, rRNA-buried nature of cL37 is indicative of its role in stabilizing extended rRNA around the protein. This also seems true for the PRP-extensions which are mostly found near to rRNA remodeled sites. Surface-exposed nature of cL38 points toward its role in translation regulation. Furthermore, the modifications at the polypeptide tunnel and tunnel exit site hint toward the existence of a separate pathway for the chloroplastic signal recognition particle (cpSRP) recruitment on the ribosome. The cpSRP contains two proteins, cpSRP54 and cpSRP43. This is unlike its bacterial counterpart which is a ribonucleoprotein containing a protein (SRP54) and 4.5S rRNA. The rRNA is known to facilitate the interaction of SRP54 with its receptor (283). It remains to be seen what is the function of SRP43 in cpSRP-stabilization on the cp-ribosome. To elucidate this, efforts can be employed to recombinantly purify the complex of cpSRP54 and cpSRP43 (cpSRP) and *in vitro* reconstitute with cp-ribosome before collecting images using cryo-EM.

The SSU map obtained from the Arctica-dataset was of insufficient quality to build a reliable SSU model. Therefore, a 300-kV Krios microscope was used to collect data. This data turned out to be of superior quality and the SSU could be resolved to 3.7 Å which enabled an atomic model to be built. Subsequently, localization and modeling of cS22, cS23 and bTHXc (SSU-residing PSRPs), bS1c and plastid pY were possible. It is to be remembered that the density obtained for cS22, cS23 and bS1c are not well resolved due to flexibility. To interpret their mode of interaction on the ribosome, therefore, it would be necessary to devise strategies for stabilization of these proteins on the ribosome using biochemical experiments.

During the period of this work, two cryo-EM structures of the cp-ribosome were published by Graf *et al.* (115) and Bieri *et al.* (114). Unfortunately, the SSU model

from the study carried out by Graf *et al.* (115) is not available in the PDB, so a meaningful comparison could not be drawn. Findings related to LSU from the current study are in good agreement with their structure. Compared to Bieri *et al.*'s report (114), where the LSU and SSU are reported at 3.2 and 3.6 Å, respectively, two differences stand out. First, the density at the mRNA exit site that has been assigned to cS23 in the current study, is absent from their map. Second, their map also did not capture the density at the SSU foot where the second RRM domain of cS22 has been accommodated in the current study. Therefore, unlike their study, in the current report, both the RRM domains of cS22 could be fitted in the protein density found at the SSU foot. Figure 5.1 shows the comparison between the 3.7 Å-resolved density obtained for the SSU in the current study and that of Bieri *et al.*'s (114) which is resolved to 3.6 Å. The comparison clearly demonstrates the presence of extra densities at the mRNA exit site and SSU foot, in the current study.

Figure 5.1 Comparison between cp-ribosome SSU maps obtained in current study and by Bieri *et al.* (A) Both maps are overlaid to show presence of an extra density at SSU foot in the current study. (B) The above-mentioned maps are overlaid again to display an extra density at the mRNA exit site in the current study. The maps obtained in current study and by Bieri *et al.* (114) are represented in grey and yellow, respectively. The fitted SSU model is taken from the current study. The densities are filtered and shown as transparent for clarity.

6. Conclusion

A high-resolution complete 70S structure of the cp-ribosome from spinach is presented in this study. This is the first step towards elucidating details of the intricacies of cp-ribosome and the chloroplastic translation in structural terms. It has long been known that the cp-ribosome contains PRP-extensions and PSRPs along with rRNA insertions and deletions. However, the precise location and structure along with the possible implication of these altered elements in the chloroplastic translation system remained elusive so far. This issue is now partly solved as this study describes the structure and interaction of many such elements in near-atomic details providing insights into the plausible function of these elements in the chloroplastic environment. Apart from that, the current study advances the understanding of the translational regulation in the chloroplastic environment, especially regarding the translational initiation. The major parts of the results agree well with the contemporary studies carried out in other labs. This serves as an additional level of validation of the results obtained in this study. Some of the PRP-extensions and parts of some PSRPs could not be modeled due to flexibility, posing limitations in the complete understanding of the cp-ribosome structure. In future, improvements in the detector-technology and image-processing software would usher in new advances in the cryo-EM field, allowing it to capture the flexible regions at a higher resolution. Once achieved, a better picture of the cp-ribosome and chloroplastic translation would emerge.

7. References

1. Bell, J.M., Chen, M., Baldwin, P.R. and Ludtke, S.J. (2016) High resolution single particle refinement in EMAN2.1. *Methods*, **100**, 25-34.
2. Scheres, S.H. (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol*, **180**, 519-530.
3. Ahmed, T., Yin, Z. and Bhushan, S. (2016) Cryo-EM structure of the large subunit of the spinach chloroplast ribosome. *Sci Rep*, **6**, 35793.
4. Djokic, T., Van Kranendonk, M.J., Campbell, K.A., Walter, M.R. and Ward, C.R. (2017) Earliest signs of life on land preserved in ca. 3.5 Ga hot spring deposits. *Nature Communications*, **8**, 15263.
5. Glansdorff, N., Xu, Y. and Labedan, B. (2008) The last universal common ancestor: emergence, constitution and genetic legacy of an elusive forerunner. *Biol Direct*, **3**, 29.
6. Fitch, W.M. and Upper, K. (1987) The phylogeny of tRNA sequences provides evidence for ambiguity reduction in the origin of the genetic code. *Cold Spring Harb Symp Quant Biol*, **52**, 759-767.
7. Hall, C.E. and Slayter, H.S. (1959) Electron microscopy of ribonucleoprotein particles from E. Coli. *Journal of Molecular Biology*, **1**, 329-IN321.
8. Vasiliev, V.D. (1974) Morphology of the ribosomal 30S subparticle according to electron microscopic data. *Acta Biol Med Ger*, **33**, 779-793.
9. Lake, J.A. (1976) Ribosome structure determined by electron microscopy of Escherichia coli small subunits, large subunits and monomeric ribosomes. *J Mol Biol*, **105**, 131-139.
10. Huxley, H.E. and Zubay, G. (1960) Electron microscope observations on the structure of microsomal particles from Escherichia coli. *Journal of Molecular Biology*, **2**, 10-IN18.
11. Kastner, B., Stoffler-Meilicke, M. and Stoffler, G. (1981) Arrangement of the subunits in the ribosome of Escherichia coli: demonstration by immunoelectron microscopy. *Proc Natl Acad Sci U S A*, **78**, 6652-6656.
12. Nonomura, Y., Blobel, G. and Sabatini, D. (1971) Structure of liver ribosomes studied by negative staining. *J Mol Biol*, **60**, 303-323.
13. Vasiliev, V.D., Selivanova, O.M. and Ryazantsev, S.N. (1983) Structure of the Escherichia coli 50 S ribosomal subunit. *J Mol Biol*, **171**, 561-569.
14. Vasiliev, V.D. (1971) Electron microscopy study of 70 S ribosomes of Escherichia coli. *FEBS Lett*, **14**, 203-205.
15. Stark, H., Orlova, E.V., Rinke-Appel, J., Junke, N., Mueller, F., Rodnina, M., Wintermeyer, W., Brimacombe, R. and van Heel, M. (1997) Arrangement of tRNAs in pre- and posttranslocational ribosomes revealed by electron cryomicroscopy. *Cell*, **88**, 19-28.
16. Stark, H., Rodnina, M.V., Rinke-Appel, J., Brimacombe, R., Wintermeyer, W. and van Heel, M. (1997) Visualization of elongation factor Tu on the Escherichia coli ribosome. *Nature*, **389**, 403-406.
17. Frank, J., Verschoor, A., Li, Y., Zhu, J., Lata, R.K., Radermacher, M., Penczek, P., Grassucci, R., Agrawal, R.K. and Srivastava, S. (1995) A model of the translational apparatus based on a three-dimensional reconstruction of the Escherichia coli ribosome. *Biochem Cell Biol*, **73**, 757-765.
18. Agrawal, R.K., Penczek, P., Grassucci, R.A. and Frank, J. (1998) Visualization of elongation factor G on the Escherichia coli 70S ribosome: the mechanism of translocation. *Proc Natl Acad Sci U S A*, **95**, 6134-6138.
19. Schlutzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F. *et al.* (2000) Structure of

- functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell*, **102**, 615-623.
20. Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Jr., Morgan-Warren, R.J., Carter, A.P., Vonnrhein, C., Hartsch, T. and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit. *Nature*, **407**, 327-339.
 21. Ban, N., Nissen, P., Hansen, J., Moore, P.B. and Steitz, T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science*, **289**, 905-920.
 22. Nissen, P., Hansen, J., Ban, N., Moore, P.B. and Steitz, T.A. (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science*, **289**, 920-930.
 23. Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H. and Noller, H.F. (2001) Crystal structure of the ribosome at 5.5 Å resolution. *Science*, **292**, 883-896.
 24. Schuwirth, B.S., Borovinskaya, M.A., Hau, C.W., Zhang, W., Vila-Sanjurjo, A., Holton, J.M. and Cate, J.H. (2005) Structures of the bacterial ribosome at 3.5 Å resolution. *Science*, **310**, 827-834.
 25. Selmer, M., Dunham, C.M., Murphy, F.V.t., Weixlbaumer, A., Petry, S., Kelley, A.C., Weir, J.R. and Ramakrishnan, V. (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. *Science*, **313**, 1935-1942.
 26. Schmeing, T.M. and Ramakrishnan, V. (2009) What recent ribosome structures have revealed about the mechanism of translation. *Nature*, **461**, 1234-1242.
 27. Steitz, T.A. (2008) A structural understanding of the dynamic ribosome machine. *Nat Rev Mol Cell Biol*, **9**, 242-253.
 28. Yusupov, M. (2014) In Ito, K. (ed.), *Regulatory Nascent Polypeptides*. Springer Japan, Tokyo, pp. 23-43.
 29. Wilson, D.N. (2009) The A-Z of bacterial translation inhibitors. *Crit Rev Biochem Mol Biol*, **44**, 393-433.
 30. Innis, C.A., Blaha, G., Bulkley, D. and Steitz, T.A. (2011) In Rodnina, M. V., Wintermeyer, W. and Green, R. (eds.), *Ribosomes: Structure, Function, and Dynamics*. Springer Vienna, Vienna, pp. 31-43.
 31. Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G. and Yusupov, M. (2011) The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science*, **334**, 1524-1529.
 32. Khatter, H., Myasnikov, A.G., Natchiar, S.K. and Klaholz, B.P. (2015) Structure of the human 80S ribosome. *Nature*, **520**, 640-645.
 33. Desai, N., Brown, A., Amunts, A. and Ramakrishnan, V. (2017) The structure of the yeast mitochondrial ribosome. *Science*, **355**, 528-531.
 34. Greber, B.J., Bieri, P., Leibundgut, M., Leitner, A., Aebersold, R., Boehringer, D. and Ban, N. (2015) Ribosome. The complete structure of the 55S mammalian mitochondrial ribosome. *Science*, **348**, 303-308.
 35. Amunts, A., Brown, A., Toots, J., Scheres, S.H.W. and Ramakrishnan, V. (2015) Ribosome. The structure of the human mitochondrial ribosome. *Science*, **348**, 95-98.
 36. Khusainov, I., Vicens, Q., Bochler, A., Grosse, F., Myasnikov, A., Menetret, J.F., Chicher, J., Marzi, S., Romby, P., Yusupova, G. *et al.* (2016) Structure of the 70S ribosome from human pathogen *Staphylococcus aureus*. *Nucleic Acids Res*, **44**, 10491-10504.

37. Yang, K., Chang, J.Y., Cui, Z., Li, X., Meng, R., Duan, L., Thongchol, J., Jakana, J., Huwe, C.M., Sacchettini, J.C. *et al.* (2017) Structural insights into species-specific features of the ribosome from the human pathogen *Mycobacterium tuberculosis*. *Nucleic Acids Res*, **45**, 10884-10894.
38. Eyal, Z., Matzov, D., Krupkin, M., Wekselman, I., Paukner, S., Zimmerman, E., Rozenberg, H., Bashan, A. and Yonath, A. (2015) Structural insights into species-specific features of the ribosome from the pathogen *Staphylococcus aureus*. *Proc Natl Acad Sci U S A*, **112**, E5805-5814.
39. Sohmen, D., Chiba, S., Shimokawa-Chiba, N., Innis, C.A., Berninghausen, O., Beckmann, R., Ito, K. and Wilson, D.N. (2015) Structure of the *Bacillus subtilis* 70S ribosome reveals the basis for species-specific stalling. *Nat Commun*, **6**, 6941.
40. Zhang, X., Lai, M., Chang, W., Yu, I., Ding, K., Mrazek, J., Ng, H.L., Yang, O.O., Maslov, D.A. and Zhou, Z.H. (2016) Structures and stabilization of kinetoplastid-specific split rRNAs revealed by comparing leishmanial and human ribosomes. *Nat Commun*, **7**, 13223.
41. Hashem, Y., des Georges, A., Fu, J., Buss, S.N., Jossinet, F., Jobe, A., Zhang, Q., Liao, H.Y., Grassucci, R.A., Bajaj, C. *et al.* (2013) High-resolution cryo-electron microscopy structure of the *Trypanosoma brucei* ribosome. *Nature*, **494**, 385-389.
42. Liu, Z., Gutierrez-Vargas, C., Wei, J., Grassucci, R.A., Ramesh, M., Espina, N., Sun, M., Tutuncuoglu, B., Madison-Antenucci, S., Woolford, J.L., Jr. *et al.* (2016) Structure and assembly model for the *Trypanosoma cruzi* 60S ribosomal subunit. *Proc Natl Acad Sci U S A*, **113**, 12174-12179.
43. Shalev-Benami, M., Zhang, Y., Matzov, D., Halfon, Y., Zackay, A., Rozenberg, H., Zimmerman, E., Bashan, A., Jaffe, C.L., Yonath, A. *et al.* (2016) 2.8-A Cryo-EM Structure of the Large Ribosomal Subunit from the Eukaryotic Parasite *Leishmania*. *Cell Rep*, **16**, 288-294.
44. Wong, W., Bai, X.C., Brown, A., Fernandez, I.S., Hanssen, E., Condrón, M., Tan, Y.H., Baum, J. and Scheres, S.H. (2014) Cryo-EM structure of the *Plasmodium falciparum* 80S ribosome bound to the anti-protozoan drug emetine. *Elife*, **3**.
45. Ogle, J.M., Carter, A.P. and Ramakrishnan, V. (2003) Insights into the decoding mechanism from recent ribosome structures. *Trends Biochem Sci*, **28**, 259-266.
46. Moazed, D. and Noller, H.F. (1990) Binding of tRNA to the ribosomal A and P sites protects two distinct sets of nucleotides in 16 S rRNA. *J Mol Biol*, **211**, 135-145.
47. Moazed, D. and Noller, H.F. (1986) Transfer RNA shields specific nucleotides in 16S ribosomal RNA from attack by chemical probes. *Cell*, **47**, 985-994.
48. Noller, H.F. and Chaires, J.B. (1972) Functional modification of 16S ribosomal RNA by kethoxal. *Proc Natl Acad Sci U S A*, **69**, 3115-3118.
49. Ogle, J.M., Brodersen, D.E., Clemons, W.M., Jr., Tarry, M.J., Carter, A.P. and Ramakrishnan, V. (2001) Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science*, **292**, 897-902.
50. Demeshkina, N., Jenner, L., Westhof, E., Yusupov, M. and Yusupova, G. (2012) A new understanding of the decoding principle on the ribosome. *Nature*, **484**, 256-259.
51. Demeshkina, N., Jenner, L., Westhof, E., Yusupov, M. and Yusupova, G. (2013) New structural insights into the decoding mechanism: translation

- infidelity via a G.U pair with Watson-Crick geometry. *FEBS Lett*, **587**, 1848-1857.
52. Loveland, A.B., Demo, G., Grigorieff, N. and Korostelev, A.A. (2017) Ensemble cryo-EM elucidates the mechanism of translation fidelity. *Nature*, **546**, 113-117.
 53. Rodnina, M.V., Fischer, N., Maracci, C. and Stark, H. (2017) Ribosome dynamics during decoding. *Philos Trans R Soc Lond B Biol Sci*, **372**.
 54. Schrode, P., Huter, P., Clementi, N. and Erlacher, M. (2017) Atomic mutagenesis at the ribosomal decoding site. *RNA Biol*, **14**, 104-112.
 55. Simonovic, M. and Steitz, T.A. (2009) A structural view on the mechanism of the ribosome-catalyzed peptide bond formation. *Biochim Biophys Acta*, **1789**, 612-623.
 56. Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D'Souza, L.M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., Muller, K.M. *et al.* (2002) The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics*, **3**, 2.
 57. Yamamoto, H., Qin, Y., Achenbach, J., Li, C., Kijek, J., Spahn, C.M. and Nierhaus, K.H. (2014) EF-G and EF4: translocation and back-translocation on the bacterial ribosome. *Nat Rev Microbiol*, **12**, 89-100.
 58. Hentschel, J., Burnside, C., Mignot, I., Leibundgut, M., Boehringer, D. and Ban, N. (2017) The Complete Structure of the Mycobacterium smegmatis 70S Ribosome. *Cell Rep*, **20**, 149-160.
 59. Voorhees, R.M., Weixlbaumer, A., Loakes, D., Kelley, A.C. and Ramakrishnan, V. (2009) Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. *Nat Struct Mol Biol*, **16**, 528-533.
 60. Jenner, L., Demeshkina, N., Yusupova, G. and Yusupov, M. (2010) Structural rearrangements of the ribosome at the tRNA proofreading step. *Nat Struct Mol Biol*, **17**, 1072-1078.
 61. Bhushan, S., Meyer, H., Starosta, A.L., Becker, T., Mielke, T., Berninghausen, O., Sattler, M., Wilson, D.N. and Beckmann, R. (2010) Structural basis for translational stalling by human cytomegalovirus and fungal arginine attenuator peptide. *Mol Cell*, **40**, 138-146.
 62. Armache, J.P., Jarasch, A., Anger, A.M., Villa, E., Becker, T., Bhushan, S., Jossinet, F., Habeck, M., Dindar, G., Franckenberg, S. *et al.* (2010) 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*, **107**, 19754-19759.
 63. Wilson, D.N. and Doudna, J.H. (2012) The structure and function of the eukaryotic ribosome. *Cold Spring Harb Perspect Biol*, **4**.
 64. Auerbach-Nevo, T., Baram, D., Bashan, A., Belousoff, M., Breiner, E., Davidovich, C., Camicata, G., Eyal, Z., Halfon, Y., Krupkin, M. *et al.* (2016) Ribosomal Antibiotics: Contemporary Challenges. *Antibiotics (Basel)*, **5**, 24.
 65. Malkin, L.I. and Rich, A. (1967) Partial resistance of nascent polypeptide chains to proteolytic digestion due to ribosomal shielding. *J Mol Biol*, **26**, 329-346.
 66. Blobel, G. and Sabatini, D.D. (1970) Controlled proteolysis of nascent polypeptides in rat liver cell fractions. I. Location of the polypeptides within ribosomes. *J Cell Biol*, **45**, 130-145.

67. Milligan, R.A. and Unwin, P.N. (1986) Location of exit channel for nascent protein in 80S ribosome. *Nature*, **319**, 693-695.
68. Yonath, A., Leonard, K.R. and Wittmann, H.G. (1987) A tunnel in the large ribosomal subunit revealed by three-dimensional image reconstruction. *Science*, **236**, 813-816.
69. Frank, J., Zhu, J., Penczek, P., Li, Y., Srivastava, S., Verschoor, A., Radermacher, M., Grassucci, R., Lata, R.K. and Agrawal, R.K. (1995) A model of protein synthesis based on cryo-electron microscopy of the E. coli ribosome. *Nature*, **376**, 441-444.
70. Becker, T., Bhushan, S., Jarasch, A., Armache, J.P., Funes, S., Jossinet, F., Gumbart, J., Mielke, T., Berninghausen, O., Schulten, K. *et al.* (2009) Structure of monomeric yeast and mammalian Sec61 complexes interacting with the translating ribosome. *Science*, **326**, 1369-1373.
71. Beckmann, R., Bubeck, D., Grassucci, R., Penczek, P., Verschoor, A., Blobel, G. and Frank, J. (1997) Alignment of conduits for the nascent polypeptide chain in the ribosome-Sec61 complex. *Science*, **278**, 2123-2126.
72. Kramer, G., Boehringer, D., Ban, N. and Bukau, B. (2009) The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nat Struct Mol Biol*, **16**, 589-597.
73. Balchin, D., Hayer-Hartl, M. and Hartl, F.U. (2016) In vivo aspects of protein folding and quality control. *Science*, **353**, aac4354.
74. Bogdanov, A.A., Sumbatyan, N.V., Shishkina, A.V., Karpenko, V.V. and Korshunova, G.A. (2010) Ribosomal tunnel and translation regulation. *Biochemistry (Mosc)*, **75**, 1501-1516.
75. Lu, J. and Deutsch, C. (2005) Folding zones inside the ribosomal exit tunnel. *Nat Struct Mol Biol*, **12**, 1123-1129.
76. Bhushan, S., Gartmann, M., Halic, M., Armache, J.P., Jarasch, A., Mielke, T., Berninghausen, O., Wilson, D.N. and Beckmann, R. (2010) alpha-Helical nascent polypeptide chains visualized within distinct regions of the ribosomal exit tunnel. *Nat Struct Mol Biol*, **17**, 313-317.
77. Wilson, D.N., Bhushan, S., Becker, T. and Beckmann, R. (2011) In Rodnina, M. V., Wintermeyer, W. and Green, R. (eds.), *Ribosomes: Structure, Function, and Dynamics*. Springer Vienna, Vienna, pp. 393-404.
78. Wilson, D.N. and Beckmann, R. (2011) The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling. *Curr Opin Struct Biol*, **21**, 274-282.
79. Wilson, D.N. and Nierhaus, K.H. (2005) Ribosomal proteins in the spotlight. *Crit Rev Biochem Mol Biol*, **40**, 243-267.
80. Weixlbaumer, A., Murphy, F.V.t., Dziergowska, A., Malkiewicz, A., Vendeix, F.A., Agris, P.F. and Ramakrishnan, V. (2007) Mechanism for expanding the decoding capacity of transfer RNAs by modification of uridines. *Nat Struct Mol Biol*, **14**, 498-502.
81. Weixlbaumer, A., Petry, S., Dunham, C.M., Selmer, M., Kelley, A.C. and Ramakrishnan, V. (2007) Crystal structure of the ribosome recycling factor bound to the ribosome. *Nat Struct Mol Biol*, **14**, 733-737.
82. Jenner, L.B., Demeshkina, N., Yusupova, G. and Yusupov, M. (2010) Structural aspects of messenger RNA reading frame maintenance by the ribosome. *Nat Struct Mol Biol*, **17**, 555-560.

83. Yusupova, G., Jenner, L., Rees, B., Moras, D. and Yusupov, M. (2006) Structural basis for messenger RNA movement on the ribosome. *Nature*, **444**, 391-394.
84. Yusupova, G.Z., Yusupov, M.M., Cate, J.H. and Noller, H.F. (2001) The path of messenger RNA through the ribosome. *Cell*, **106**, 233-241.
85. Jenner, L., Rees, B., Yusupov, M. and Yusupova, G. (2007) Messenger RNA conformations in the ribosomal E site revealed by X-ray crystallography. *EMBO Rep*, **8**, 846-850.
86. Noeske, J., Wasserman, M.R., Terry, D.S., Altman, R.B., Blanchard, S.C. and Cate, J.H. (2015) High-resolution structure of the Escherichia coli ribosome. *Nat Struct Mol Biol*, **22**, 336-341.
87. Shine, J. and Dalgarno, L. (1974) 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*, **71**, 1342-1346.
88. Hosaka, H., Nakagawa, A., Tanaka, I., Harada, N., Sano, K., Kimura, M., Yao, M. and Wakatsuki, S. (1997) Ribosomal protein S7: a new RNA-binding motif with structural similarities to a DNA architectural factor. *Structure*, **5**, 1199-1208.
89. Carter, A.P., Clemons, W.M., Brodersen, D.E., Morgan-Warren, R.J., Wimberly, B.T. and Ramakrishnan, V. (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature*, **407**, 340-348.
90. Robert, F. and Brakier-Gingras, L. (2003) A functional interaction between ribosomal proteins S7 and S11 within the bacterial ribosome. *J Biol Chem*, **278**, 44913-44920.
91. Nilsson, J. and Nissen, P. (2005) Elongation factors on the ribosome. *Curr Opin Struct Biol*, **15**, 349-354.
92. Schmeing, T.M., Voorhees, R.M., Kelley, A.C., Gao, Y.G., Murphy, F.V.t., Weir, J.R. and Ramakrishnan, V. (2009) The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science*, **326**, 688-694.
93. Gao, Y.G., Selmer, M., Dunham, C.M., Weixlbaumer, A., Kelley, A.C. and Ramakrishnan, V. (2009) The structure of the ribosome with elongation factor G trapped in the posttranslocational state. *Science*, **326**, 694-699.
94. Rawat, U., Gao, H., Zavialov, A., Gursky, R., Ehrenberg, M. and Frank, J. (2006) Interactions of the release factor RF1 with the ribosome as revealed by cryo-EM. *J Mol Biol*, **357**, 1144-1153.
95. Rawat, U.B., Zavialov, A.V., Sengupta, J., Valle, M., Grassucci, R.A., Linde, J., Vestergaard, B., Ehrenberg, M. and Frank, J. (2003) A cryo-electron microscopic study of ribosome-bound termination factor RF2. *Nature*, **421**, 87-90.
96. Gao, H., Zhou, Z., Rawat, U., Huang, C., Bouakaz, L., Wang, C., Cheng, Z., Liu, Y., Zavialov, A., Gursky, R. *et al.* (2007) RF3 induces ribosomal conformational changes responsible for dissociation of class I release factors. *Cell*, **129**, 929-941.
97. Li, W., Liu, Z., Koripella, R.K., Langlois, R., Sanyal, S. and Frank, J. (2015) Activation of GTP hydrolysis in mRNA-tRNA translocation by elongation factor G. *Sci Adv*, **1**.
98. Keiler, K.C. (2015) Mechanisms of ribosome rescue in bacteria. *Nat Rev Microbiol*, **13**, 285-297.

99. Ortiz, J.O., Brandt, F., Matias, V.R., Sennels, L., Rappsilber, J., Scheres, S.H., Eibauer, M., Hartl, F.U. and Baumeister, W. (2010) Structure of hibernating ribosomes studied by cryoelectron tomography in vitro and in situ. *J Cell Biol*, **190**, 613-621.
100. Kato, T., Yoshida, H., Miyata, T., Maki, Y., Wada, A. and Namba, K. (2010) Structure of the 100S ribosome in the hibernation stage revealed by electron cryomicroscopy. *Structure*, **18**, 719-724.
101. Yoshida, H., Maki, Y., Kato, H., Fujisawa, H., Izutsu, K., Wada, C. and Wada, A. (2002) The Ribosome Modulation Factor (RMF) Binding Site on the 100S Ribosome of *Escherichia coli*. *The Journal of Biochemistry*, **132**, 983-989.
102. Polikanov, Y.S., Blaha, G.M. and Steitz, T.A. (2012) How hibernation factors RMF, HPF, and YfiA turn off protein synthesis. *Science*, **336**, 915-918.
103. Franken, L.E., Oostergetel, G.T., Pijning, T., Puri, P., Arkhipova, V., Boekema, E.J., Poolman, B. and Guskov, A. (2017) A general mechanism of ribosome dimerization revealed by single-particle cryo-electron microscopy. *Nat Commun*, **8**, 722.
104. Matzov, D., Aibara, S., Basu, A., Zimmerman, E., Bashan, A., Yap, M.F., Amunts, A. and Yonath, A.E. (2017) The cryo-EM structure of hibernating 100S ribosome dimer from pathogenic *Staphylococcus aureus*. *Nat Commun*, **8**, 723.
105. Ueta, M., Ohniwa, R.L., Yoshida, H., Maki, Y., Wada, C. and Wada, A. (2008) Role of HPF (hibernation promoting factor) in translational activity in *Escherichia coli*. *J Biochem*, **143**, 425-433.
106. Khusainov, I., Vicens, Q., Ayupov, R., Usachev, K., Myasnikov, A., Simonetti, A., Validov, S., Kieffer, B., Yusupova, G., Yusupov, M. *et al.* (2017) Structures and dynamics of hibernating ribosomes from *Staphylococcus aureus* mediated by intermolecular interactions of HPF. *EMBO J*, **36**, 2073-2087.
107. Starosta, A.L., Lassak, J., Jung, K. and Wilson, D.N. (2014) The bacterial translation stress response. *FEMS Microbiol Rev*, **38**, 1172-1201.
108. Puri, P., Eckhardt, T.H., Franken, L.E., Fusetti, F., Stuart, M.C., Boekema, E.J., Kuipers, O.P., Kok, J. and Poolman, B. (2014) *Lactococcus lactis* YfiA is necessary and sufficient for ribosome dimerization. *Mol Microbiol*, **91**, 394-407.
109. Kline, B.C., McKay, S.L., Tang, W.W. and Portnoy, D.A. (2015) The *Listeria monocytogenes* hibernation-promoting factor is required for the formation of 100S ribosomes, optimal fitness, and pathogenesis. *J Bacteriol*, **197**, 581-591.
110. Ueta, M., Wada, C., Daifuku, T., Sako, Y., Bessho, Y., Kitamura, A., Ohniwa, R.L., Morikawa, K., Yoshida, H., Kato, T. *et al.* (2013) Conservation of two distinct types of 100S ribosome in bacteria. *Genes Cells*, **18**, 554-574.
111. Margulis, L. (1970) *Origin of eukaryotic cells: evidence and research implications for a theory of the origin and evolution of microbial, plant, and animal cells on the Precambrian earth*. Yale University Press New Haven.
112. Greber, B.J. and Ban, N. (2016) Structure and Function of the Mitochondrial Ribosome. *Annu Rev Biochem*, **85**, 103-132.
113. Ahmed, T., Shi, J. and Bhushan, S. (2017) Unique localization of the plastid-specific ribosomal proteins in the chloroplast ribosome small subunit provides mechanistic insights into the chloroplastic translation. *Nucleic Acids Res*, **45**, 8581-8595.

114. Bieri, P., Leibundgut, M., Saurer, M., Boehringer, D. and Ban, N. (2017) The complete structure of the chloroplast 70S ribosome in complex with translation factor pY. *EMBO J*, **36**, 475-486.
115. Graf, M., Arenz, S., Huter, P., Donhofer, A., Novacek, J. and Wilson, D.N. (2017) Cryo-EM structure of the spinach chloroplast ribosome reveals the location of plastid-specific ribosomal proteins and extensions. *Nucleic Acids Res*, **45**, 2887-2896.
116. Forsberg, B.O., Aibara, S., Kimanius, D., Paul, B., Lindahl, E. and Amunts, A. (2017) Cryo-EM reconstruction of the chlororibosome to 3.2 Å resolution within 24 h. *IUCrJ*, **4**, 723-727.
117. Manuell, A.L., Quispe, J. and Mayfield, S.P. (2007) Structure of the chloroplast ribosome: novel domains for translation regulation. *PLoS biology*, **5**, e209.
118. Agrawal, R.K., Sharma, M.R., Yassin, A., Lahiri, I. and Spremulli, i.L. (2011) In Rodnina, M. V., Wintermeyer, W. and Green, R. (eds.), *Ribosomes: Structure, Function, and Dynamics*. Springer Vienna, Vienna, pp. 83-96.
119. Yamaguchi, K. and Subramanian, A.R. (2000) The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast). *J Biol Chem*, **275**, 28466-28482.
120. Yamaguchi, K., von Knoblauch, K. and Subramanian, A.R. (2000) The plastid ribosomal proteins. Identification of all the proteins in the 30 S subunit of an organelle ribosome (chloroplast). *J Biol Chem*, **275**, 28455-28465.
121. Ott, M. and Herrmann, J.M. (2010) Co-translational membrane insertion of mitochondrially encoded proteins. *Biochim Biophys Acta*, **1803**, 767-775.
122. Harris, E.H., Boynton, J.E. and Gillham, N.W. (1994) Chloroplast ribosomes and protein synthesis. *Microbiol Rev*, **58**, 700-754.
123. Trempe, M.R. and Glitz, D.G. (1981) Chloroplast ribosome structure. Electron microscopy of ribosomal subunits and localization of N6,N6-dimethyladenosine by immunoelectronmicroscopy. *J Biol Chem*, **256**, 11873-11879.
124. Sharma, M.R., Wilson, D.N., Datta, P.P., Barat, C., Schlutzen, F., Fucini, P. and Agrawal, R.K. (2007) Cryo-EM study of the spinach chloroplast ribosome reveals the structural and functional roles of plastid-specific ribosomal proteins. *Proc Natl Acad Sci U S A*, **104**, 19315-19320.
125. Yamaguchi, K. and Subramanian, A.R. (2003) Proteomic identification of all plastid-specific ribosomal proteins in higher plant chloroplast 30S ribosomal subunit. PSRP-2 (U1A-type domains), PSRP-3alpha/beta (ycf65 homologue) and PSRP-4 (Thx homologue). *European Journal of Biochemistry*, **270**, 190-205.
126. Schmitz-Linneweber, C., Maier, R.M., Alcaraz, J.P., Cottet, A., Herrmann, R.G. and Mache, R. (2001) The plastid chromosome of spinach (*Spinacia oleracea*): complete nucleotide sequence and gene organization. *Plant Mol Biol*, **45**, 307-315.
127. Koussevitzky, S., Nott, A., Mockler, T.C., Hong, F., Sachetto-Martins, G., Surpin, M., Lim, J., Mittler, R. and Chory, J. (2007) Multiple Signals from Damaged Chloroplasts Converge on a Common Pathway to Regulate Nuclear Gene Expression. *Science*.
128. Jarvis, P. and Robinson, C. (2004) Mechanisms of protein import and routing in chloroplasts. *Curr Biol*, **14**, R1064-1077.

129. Carrie, C. and Small, I. (2013) A reevaluation of dual-targeting of proteins to mitochondria and chloroplasts. *Biochim Biophys Acta*, **1833**, 253-259.
130. Peeters, N. and Small, I. (2001) Dual targeting to mitochondria and chloroplasts. *Biochim Biophys Acta*, **1541**, 54-63.
131. Monde, R.A., Schuster, G. and Stern, D.B. (2000) Processing and degradation of chloroplast mRNA. *Biochimie*, **82**, 573-582.
132. Nakamura, T., Schuster, G., Sugiura, M. and Sugita, M. (2004) Chloroplast RNA-binding and pentatricopeptide repeat proteins. *Biochem Soc Trans*, **32**, 571-574.
133. Rochaix, J.-D. (2006) In Wise, R. R. and Hooper, J. K. (eds.), *The Structure and Function of Plastids*. Springer Netherlands, Dordrecht, pp. 145-165.
134. Marin-Navarro, J., Manuell, A.L., Wu, J. and S, P.M. (2007) Chloroplast translation regulation. *Photosynth Res*, **94**, 359-374.
135. Schuster, G. and Bock, R. (2001) In Aro, E.-M. and Andersson, B. (eds.), *Regulation of Photosynthesis*. Springer Netherlands, Dordrecht, pp. 121-136.
136. Asakura, Y. and Barkan, A. (2006) Arabidopsis orthologs of maize chloroplast splicing factors promote splicing of orthologous and species-specific group II introns. *Plant Physiol*, **142**, 1656-1663.
137. Hayes, M.L. and Hanson, M.R. (2007) Identification of a sequence motif critical for editing of a tobacco chloroplast transcript. *RNA*, **13**, 281-288.
138. Yukawa, M., Kuroda, H. and Sugiura, M. (2007) A new in vitro translation system for non-radioactive assay from tobacco chloroplasts: effect of pre-mRNA processing on translation in vitro. *Plant J*, **49**, 367-376.
139. Zerges, W. (2000) Translation in chloroplasts. *Biochimie*, **82**, 583-601.
140. Sugiura, M., Hirose, T. and Sugita, M. (1998) Evolution and mechanism of translation in chloroplasts. *Annu Rev Genet*, **32**, 437-459.
141. Wostrikoff, K., Choquet, Y., Wollman, F.-A. and Girard-Bascou, J. (2001) TCA1, a Single Nuclear-Encoded Translational Activator Specific for *petA* mRNA in *Chlamydomonas reinhardtii* Chloroplast. *Genetics*, **159**, 119-132.
142. Yohn, C.B., Cohen, A., Rosch, C., Kuchka, M.R. and Mayfield, S.P. (1998) Translation of the chloroplast *psbA* mRNA requires the nuclear-encoded poly(A)-binding protein, RB47. *J Cell Biol*, **142**, 435-442.
143. Stampacchia, O., Girard-Bascou, J., Zanasco, J.L., Zerges, W., Bennoun, P. and Rochaix, J.D. (1997) A nuclear-encoded function essential for translation of the chloroplast *psaB* mRNA in *chlamydomonas*. *Plant Cell*, **9**, 773-782.
144. Girard-Bascou, J., Pierre, Y. and Drapier, D. (1992) A nuclear mutation affects the synthesis of the chloroplast *psbA* gene production *Chlamydomonas reinhardtii*. *Curr Genet*, **22**, 47-52.
145. Rochaix, J.D., Kuchka, M., Mayfield, S., Schirmer-Rahire, M., Girard-Bascou, J. and Bennoun, P. (1989) Nuclear and chloroplast mutations affect the synthesis or stability of the chloroplast *psbC* gene product in *Chlamydomonas reinhardtii*. *EMBO J*, **8**, 1013-1021.
146. Rattanachaikunsopon, P., Rosch, C. and Kuchka, M.R. (1999) Cloning and characterization of the nuclear AC115 gene of *Chlamydomonas reinhardtii*. *Plant Mol Biol*, **39**, 1-10.
147. Kim, J. (1997) Protein Disulfide Isomerase as a Regulator of Chloroplast Translational Activation. *Science*, **278**, 1954-1957.

148. Maidak, B.L., Cole, J.R., Lilburn, T.G., Parker, C.T., Jr., Saxman, P.R., Farris, R.J., Garrity, G.M., Olsen, G.J., Schmidt, T.M. and Tiedje, J.M. (2001) The RDP-II (Ribosomal Database Project). *Nucleic Acids Res*, **29**, 173-174.
149. Bonham-Smith, P.C. and Bourque, D.P. (1989) Translation of chloroplast-encoded mRNA: potential initiation and termination signals. *Nucleic Acids Res*, **17**, 2057-2080.
150. Ruf, M. and Kössel, H. (1988) Occurrence and spacing of ribosome recognition sites in mRNAs of chloroplasts from higher plants. *FEBS Letters*, **240**, 41-44.
151. Scharff, L.B., Ehrnthal, M., Janowski, M., Childs, L.H., Hasse, C., Gremmels, J., Ruf, S., Zoschke, R. and Bock, R. (2017) Shine-Dalgarno Sequences Play an Essential Role in the Translation of Plastid mRNAs in Tobacco. *Plant Cell*, **29**, 3085-3101.
152. Zhang, D.C.F.M. (1998) Shine-Dalgarno-like sequences are not required for translation of chloroplast mRNAs in *Chlamydomonas reinhardtii* chloroplasts or in *Escherichia coli*. *Molecular and General Genetics MGG*, **257**, 271-282.
153. Nickelsen, J., Fleischmann, M., Boudreau, E., Rahire, M. and Rochaix, J.D. (1999) Identification of cis-acting RNA leader elements required for chloroplast psbD gene expression in *Chlamydomonas*. *Plant Cell*, **11**, 957-970.
154. Sakamoto, W., Chen, X., Kindle, K.L. and Stern, D.B. (1994) Function of the *Chlamydomonas reinhardtii* petD 5' untranslated region in regulating the accumulation of subunit IV of the cytochrome b6/f complex. *Plant J*, **6**, 503-512.
155. Hirose, T., Kusumegi, T. and Sugiura, M. (1998) Translation of tobacco chloroplast rps14 mRNA depends on a Shine-Dalgarno-like sequence in the 5'-untranslated region but not on internal RNA editing in the coding region. *FEBS Lett*, **430**, 257-260.
156. Hirose, T. and Sugiura, M. (2004) Functional Shine-Dalgarno-like sequences for translational initiation of chloroplast mRNAs. *Plant & cell physiology*, **45**, 114-117.
157. Kuroda, H., Suzuki, H., Kusumegi, T., Hirose, T., Yukawa, Y. and Sugiura, M. (2007) Translation of psbC mRNAs starts from the downstream GUG, not the upstream AUG, and requires the extended Shine-Dalgarno sequence in tobacco chloroplasts. *Plant & cell physiology*, **48**, 1374-1378.
158. Manuell, A., Beligni, M.V., Yamaguchi, K. and Mayfield, S.P. (2004) Regulation of chloroplast translation: interactions of RNA elements, RNA-binding proteins and the plastid ribosome. *Biochem Soc Trans*, **32**, 601-605.
159. Cornish, P.V., Ermolenko, D.N., Noller, H.F. and Ha, T. (2008) Spontaneous intersubunit rotation in single ribosomes. *Mol Cell*, **30**, 578-588.
160. Frank, J. and Agrawal, R.K. (2000) A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature*, **406**, 318-322.
161. Scheres, S.H. (2012) A Bayesian view on cryo-EM structure determination. *J Mol Biol*, **415**, 406-418.
162. Lyumkis, D., Brilot, A.F., Theobald, D.L. and Grigorieff, N. (2013) Likelihood-based classification of cryo-EM images using FREALIGN. *J Struct Biol*, **183**, 377-388.
163. Grigorieff, N. (2007) FREALIGN: high-resolution refinement of single particle structures. *J Struct Biol*, **157**, 117-125.

164. Rother, M., Milanowska, K., Puton, T., Jeleniewicz, J., Rother, K. and Bujnicki, J.M. (2011) ModeRNA server: an online tool for modeling RNA 3D structures. *Bioinformatics*, **27**, 2441-2442.
165. Schrödinger, L.L.C. (2010) Py-MOL, The PyMOL Molecular Graphics System, V1.8.
166. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, **25**, 1605-1612.
167. Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*, **66**, 486-501.
168. Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*, **66**, 12-21.
169. Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H. and Adams, P.D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr*, **68**, 352-367.
170. Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, **9**, 40.
171. Kleywegt, G.J. and Jones, T.A. (1996) xdlMAPMAN and xdlDATAMAN - programs for reformatting, analysis and manipulation of biomacromolecular electron-density maps and reflection data sets. *Acta Crystallogr D Biol Crystallogr*, **52**, 826-828.
172. Smith, D.J. (2008) Ultimate resolution in the electron microscope? *Materials Today*, **11**, 30-38.
173. Lambert, L. and Mulvey, T. (1996) In Hawkes, P. W. (ed.), *Advances in Imaging and Electron Physics*. Elsevier, Vol. 95, pp. 2-62.
174. Hall, C.E., Jakus, M.A. and Schmitt, F.O. (1945) The Structure of Certain Muscle Fibrils as Revealed by the Use of Electron Stains. *Journal of Applied Physics*, **16**, 459-465.
175. Brenner, S. and Horne, R.W. (1959) A negative staining method for high resolution electron microscopy of viruses. *Biochim Biophys Acta*, **34**, 103-110.
176. Huxley, H.E. and Zubay, G. (1961) Preferential staining of nucleic acid-containing structures for electron microscopy. *J Biophys Biochem Cytol*, **11**, 273-296.
177. Booth, D.S., Avila-Sakar, A. and Cheng, Y. (2011) Visualizing proteins and macromolecular complexes by negative stain EM: from grid preparation to image acquisition. *J Vis Exp*, 3227.
178. Ohi, M., Li, Y., Cheng, Y. and Walz, T. (2004) Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy. *Biol Proced Online*, **6**, 23-34.
179. De Carlo, S. and Harris, J.R. (2011) Negative staining and cryo-negative staining of macromolecules and viruses for TEM. *Micron*, **42**, 117-131.
180. De Rosier, D.J. and Klug, A. (1968) Reconstruction of three dimensional structures from electron micrographs. *Nature*, **217**, 130-134.
181. Thompson, R.F., Walker, M., Siebert, C.A., Muench, S.P. and Ranson, N.A. (2016) An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods*, **100**, 3-15.

182. Hesketh, E.L., Parker-Manuel, R.P., Chaban, Y., Satti, R., Coverley, D., Orlova, E.V. and Chong, J.P. (2015) DNA induces conformational changes in a recombinant human minichromosome maintenance complex. *J Biol Chem*, **290**, 7973-7979.
183. Roberts, A.J., Numata, N., Walker, M.L., Kato, Y.S., Malkova, B., Kon, T., Ohkura, R., Arisaka, F., Knight, P.J., Sutoh, K. *et al.* (2009) AAA+ Ring and linker swing mechanism in the dynein motor. *Cell*, **136**, 485-495.
184. Rasmussen, S.G.F., DeVree, B.T., Zou, Y., Kruse, A.C., Chung, K.Y., Kobilka, T.S., Thian, F.S., Chae, P.S., Pardon, E., Calinski, D. *et al.* (2011) Crystal structure of the β_2 adrenergic receptor-Gs protein complex. *Nature*, **477**, 549.
185. Low, H.H., Gubellini, F., Rivera-Calzada, A., Braun, N., Connery, S., Dujancourt, A., Lu, F., Redzej, A., Fronzes, R., Orlova, E.V. *et al.* (2014) Structure of a type IV secretion system. *Nature*, **508**, 550-553.
186. Wu, S., Avila-Sakar, A., Kim, J., Booth, D.S., Greenberg, C.H., Rossi, A., Liao, M., Li, X., Alian, A., Griner, S.L. *et al.* (2012) Fabs enable single particle cryoEM studies of small proteins. *Structure*, **20**, 582-592.
187. Taylor, K.A. and Glaeser, R.M. (1974) Electron diffraction of frozen, hydrated protein crystals. *Science*, **186**, 1036-1037.
188. Dubochet, J., Adrian, M., Chang, J.-J., Homo, J.-C., Lepault, J., McDowell, A.W. and Schultz, P. (2009) Cryo-electron microscopy of vitrified specimens. *Quarterly Reviews of Biophysics*, **21**, 129.
189. Saibil, H.R. (2000) Macromolecular structure determination by cryo-electron microscopy. *Acta Crystallogr D Biol Crystallogr*, **56**, 1215-1222.
190. Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E. and Downing, K.H. (1990) Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J Mol Biol*, **213**, 899-929.
191. Kuhlbrandt, W., Wang, D.N. and Fujiyoshi, Y. (1994) Atomic model of plant light-harvesting complex by electron crystallography. *Nature*, **367**, 614-621.
192. Unwin, N. (1995) Acetylcholine receptor channel imaged in the open state. *Nature*, **373**, 37-43.
193. Bottcher, B., Wynne, S.A. and Crowther, R.A. (1997) Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. *Nature*, **386**, 88-91.
194. Conway, J.F., Cheng, N., Zlotnick, A., Wingfield, P.T., Stahl, S.J. and Steven, A.C. (1997) Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy. *Nature*, **386**, 91-94.
195. Nogales, E., Wolf, S.G. and Downing, K.H. (1998) Structure of the alpha beta tubulin dimer by electron crystallography. *Nature*, **391**, 199-203.
196. Serpell, L.C. and Smith, J.M. (2000) Direct visualisation of the beta-sheet structure of synthetic Alzheimer's amyloid. *J Mol Biol*, **299**, 225-231.
197. Frank, J., Penczek, P., Agrawal, R.K., Grassucci, R.A. and Heagle, A.B. (2000), *Methods in Enzymology*. Academic Press, Vol. 317, pp. 276-291.
198. Wu, S., Armache, J.P. and Cheng, Y. (2016) Single-particle cryo-EM data acquisition by using direct electron detection camera. *Microscopy (Oxf)*, **65**, 35-41.
199. Kuhlbrandt, W. (2014) Biochemistry. The resolution revolution. *Science*, **343**, 1443-1444.

200. Ruskin, R.S., Yu, Z. and Grigorieff, N. (2013) Quantitative characterization of electron detectors for transmission electron microscopy. *J Struct Biol*, **184**, 385-393.
201. Orlova, E.V. and Saibil, H.R. (2011) Structural analysis of macromolecular assemblies by electron microscopy. *Chem Rev*, **111**, 7710-7748.
202. Wong, H.C., Chen, J., Mouche, F., Rouiller, I. and Bern, M. (2004) Model-based particle picking for cryo-electron microscopy. *J Struct Biol*, **145**, 157-167.
203. Jiang, L., Liu, Z., Georgieva, D., Kuil, M.E. and Abrahams, J.P. (2010) A novel approximation method of CTF amplitude correction for 3D single particle reconstruction. *Ultramicroscopy*, **110**, 350-358.
204. Marabini, R., Carragher, B., Chen, S., Chen, J., Cheng, A., Downing, K.H., Frank, J., Grassucci, R.A., Bernard Heymann, J., Jiang, W. *et al.* (2015) CTF Challenge: Result summary. *J Struct Biol*, **190**, 348-359.
205. Jonic, S. (2016) Cryo-electron Microscopy Analysis of Structurally Heterogeneous Macromolecular Complexes. *Comput Struct Biotechnol J*, **14**, 385-390.
206. Cheng, Y., Grigorieff, N., Penczek, P.A. and Walz, T. (2015) A primer to single-particle cryo-electron microscopy. *Cell*, **161**, 438-449.
207. Frank, J. (1996), *Three-Dimensional Electron Microscopy of Macromolecular Assemblies*. Academic Press, Burlington, pp. 182-246.
208. Radermacher, M., Wagenknecht, T., Verschoor, A. and Frank, J. (1986) A new 3-D reconstruction scheme applied to the 50S ribosomal subunit of *E. coli*. *J Microsc*, **141**, RP1-2.
209. Radermacher, M., Wagenknecht, T., Verschoor, A. and Frank, J. (1987) Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *J Microsc*, **146**, 113-136.
210. van Heel, M., Gowen, B., Matadeen, R., Orlova, E.V., Finn, R., Pape, T., Cohen, D., Stark, H., Schmidt, R., Schatz, M. *et al.* (2000) Single-particle electron cryo-microscopy: towards atomic resolution. *Q Rev Biophys*, **33**, 307-369.
211. Van Heel, M. (1987) Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy*, **21**, 111-123.
212. Penczek, P.A. (2010) Resolution measures in molecular electron microscopy. *Methods Enzymol*, **482**, 73-100.
213. van Heel, M. and Harauz, G. (1986) Resolution criteria for three dimensional reconstruction. *Optik*, **73**, 119-122.
214. Rosenthal, P.B. and Henderson, R. (2003) Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol*, **333**, 721-745.
215. van Heel, M. and Schatz, M. (2005) Fourier shell correlation threshold criteria. *J Struct Biol*, **151**, 250-262.
216. Bai, X.C., McMullan, G. and Scheres, S.H. (2015) How cryo-EM is revolutionizing structural biology. *Trends Biochem Sci*, **40**, 49-57.
217. Elmlund, D., Le, S.N. and Elmlund, H. (2017) High-resolution cryo-EM: the nuts and bolts. *Curr Opin Struct Biol*, **46**, 1-6.
218. Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A. and Cheng, Y. (2013) Electron counting and beam-induced

- motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods*, **10**, 584-590.
219. Grant, T. and Grigorieff, N. (2015) Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *Elife*, **4**, e06980.
 220. Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y. and Agard, D.A. (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods*, **14**, 331-332.
 221. Frank, J. (2017) Advances in the field of single-particle cryo-electron microscopy over the last decade. *Nat Protoc*, **12**, 209-212.
 222. Murata, K. and Wolf, M. (2018) Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *Biochim Biophys Acta*, **1862**, 324-334.
 223. Glaeser, R.M. (2016) How good can cryo-EM become? *Nat Methods*, **13**, 28-32.
 224. Grigorieff, N. (2016) In Crowther, R. A. (ed.), *Methods in Enzymology*. Academic Press, Vol. 579, pp. 191-226.
 225. Scheres, S.H., Nunez-Ramirez, R., Sorzano, C.O., Carazo, J.M. and Marabini, R. (2008) Image processing for electron microscopy single-particle analysis using XMIPP. *Nat Protoc*, **3**, 977-990.
 226. Kastner, B., Fischer, N., Golas, M.M., Sander, B., Dube, P., Boehringer, D., Hartmuth, K., Deckert, J., Hauer, F., Wolf, E. *et al.* (2008) GraFix: sample preparation for single-particle electron cryomicroscopy. *Nat Methods*, **5**, 53-55.
 227. Cheng, Y. (2015) Single-Particle Cryo-EM at Crystallographic Resolution. *Cell*, **161**, 450-457.
 228. Li, X., Grigorieff, N. and Cheng, Y. (2010) GPU-enabled FREALIGN: accelerating single particle 3D reconstruction and refinement in Fourier space on graphics processors. *J Struct Biol*, **172**, 407-412.
 229. Kimanius, D., Forsberg, B.O., Scheres, S.H. and Lindahl, E. (2016) Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *Elife*, **5**, e18722.
 230. Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C.S. and Carragher, B. (2005) Automated molecular microscopy: the new Leginon system. *J Struct Biol*, **151**, 41-60.
 231. Mindell, J.A. and Grigorieff, N. (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol*, **142**, 334-347.
 232. Kucukelbir, A., Sigworth, F.J. and Tagare, H.D. (2014) Quantifying the local resolution of cryo-EM density maps. *Nat Methods*, **11**, 63-65.
 233. Afonine, P.V., Headd, J.J., Terwilliger, T.C., and Adams, P.D. (2013) New tool: phenix.real_space_refine. *Computational Crystallography Newsletter*, **4**, 43-44.
 234. Apweiler, R., Bairoch, A., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M. *et al.* (2004) UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res*, **32**, D115-119.
 235. Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. and Sayers, E.W. (2013) GenBank. *Nucleic Acids Res*, **41**, D36-42.
 236. Polikanov, Y.S., Melnikov, S.V., Söll, D. and Steitz, T.A. (2015) Structural insights into the role of rRNA modifications in protein synthesis and ribosome assembly. *Nature Structural & Molecular Biology*, **22**, 342.

237. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*, **66**, 213-221.
238. Ban, N., Beckmann, R., Cate, J.H., Dinman, J.D., Dragon, F., Ellis, S.R., Lafontaine, D.L., Lindahl, L., Liljas, A., Lipton, J.M. *et al.* (2014) A new system for naming ribosomal proteins. *Curr Opin Struct Biol*, **24**, 165-169.
239. Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D. and Hochstrasser, D.F. (1999) Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol*, **112**, 531-552.
240. Tiller, N., Weingartner, M., Thiele, W., Maximova, E., Schottler, M.A. and Bock, R. (2012) The plastid-specific ribosomal proteins of *Arabidopsis thaliana* can be divided into non-essential proteins and genuine ribosomal proteins. *Plant J*, **69**, 302-316.
241. Yamaguchi, K. and Subramanian, A.R. (2003) Proteomic identification of all plastid-specific ribosomal proteins in higher plant chloroplast 30S ribosomal subunit. *Eur J Biochem*, **270**, 190-205.
242. Rice, P., Longden, I. and Bleasby, A. (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet*, **16**, 276-277.
243. Klein, D.J., Schmeing, T.M., Moore, P.B. and Steitz, T.A. (2001) The kink-turn: a new RNA secondary structure motif. *EMBO J*, **20**, 4214-4221.
244. Huang, L. and Lilley, D.M.J. (2016) The Kink Turn, a Key Architectural Element in RNA Structure. *J Mol Biol*, **428**, 790-801.
245. Bhushan, S., Hoffmann, T., Seidelt, B., Frauenfeld, J., Mielke, T., Berninghausen, O., Wilson, D.N. and Beckmann, R. (2011) SecM-stalled ribosomes adopt an altered geometry at the peptidyl transferase center. *PLoS biology*, **9**, e1000581.
246. Bornemann, T., Holtkamp, W. and Wintermeyer, W. (2014) Interplay between trigger factor and other protein biogenesis factors on the ribosome. *Nat Commun*, **5**, 4180.
247. Wegrzyn, R.D., Hofmann, D., Merz, F., Nikolay, R., Rauch, T., Graf, C. and Deuerling, E. (2006) A conserved motif is prerequisite for the interaction of NAC with ribosomal protein L23 and nascent chains. *J Biol Chem*, **281**, 2847-2857.
248. Singh, R., Kraft, C., Jaiswal, R., Sejwal, K., Kasaragod, V.B., Kuper, J., Burger, J., Mielke, T., Lührink, J. and Bhushan, S. (2014) Cryo-electron microscopic structure of SecA protein bound to the 70S ribosome. *J Biol Chem*, **289**, 7190-7199.
249. Bornemann, T., Jockel, J., Rodnina, M.V. and Wintermeyer, W. (2008) Signal sequence-independent membrane targeting of ribosomes containing short nascent peptides within the exit tunnel. *Nat Struct Mol Biol*, **15**, 494-499.
250. Eichacker, L.A. and Henry, R. (2001) Function of a chloroplast SRP in thylakoid protein export. *Biochim Biophys Acta*, **1541**, 120-134.
251. Xu, T., Lee, K., Gu, L., Kim, J.I. and Kang, H. (2013) Functional characterization of a plastid-specific ribosomal protein PSRP2 in *Arabidopsis thaliana* under abiotic stress conditions. *Plant Physiol Biochem*, **73**, 405-411.
252. Maris, C., Dominguez, C. and Allain, F.H. (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J*, **272**, 2118-2131.

253. Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research*, **16**, 10881-10890.
254. Kurkcuoglu, O., Doruker, P., Sen, T.Z., Kloczkowski, A. and Jernigan, R.L. (2008) The ribosome structure controls and directs mRNA entry, translocation and exit dynamics. *Phys Biol*, **5**, 046005.
255. Godbey, W.T. (2014), *An Introduction to Biotechnology*. Woodhead Publishing, pp. 65-105.
256. Sharma, M.R., Donhofer, A., Barat, C., Marquez, V., Datta, P.P., Fucini, P., Wilson, D.N. and Agrawal, R.K. (2010) PSRP1 is not a ribosomal protein, but a ribosome-binding factor that is recycled by the ribosome-recycling factor (RRF) and elongation factor G (EF-G). *J Biol Chem*, **285**, 4006-4014.
257. Fischer, N., Neumann, P., Konevega, A.L., Bock, L.V., Ficner, R., Rodnina, M.V. and Stark, H. (2015) Structure of the E. coli ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. *Nature*, **520**, 567-570.
258. Danon, A. (1997) Translational regulation in the chloroplast. *Plant Physiol*, **115**, 1293-1298.
259. Merendino, L., Falcioratore, A. and Rochaix, J.D. (2003) Expression and RNA binding properties of the chloroplast ribosomal protein S1 from *Chlamydomonas reinhardtii*. *Plant Mol Biol*, **53**, 371-382.
260. Franzetti, B., Carol, P. and Mache, R. (1992) Characterization and RNA-binding properties of a chloroplast S1-like ribosomal protein. *J Biol Chem*, **267**, 19075-19081.
261. Sorensen, M.A., Fricke, J. and Pedersen, S. (1998) Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in *Escherichia coli* in vivo. *J Mol Biol*, **280**, 561-569.
262. Tzareva, N.V., Makhno, V.I. and Boni, I.V. (1994) Ribosome-messenger recognition in the absence of the Shine-Dalgarno interactions. *FEBS Lett*, **337**, 189-194.
263. Kolb, A., Hermoso, J.M., Thomas, J.O. and Szer, W. (1977) Nucleic acid helix-unwinding properties of ribosomal protein S1 and the role of S1 in mRNA binding to ribosomes. *Proc Natl Acad Sci U S A*, **74**, 2379-2383.
264. Boni, I.V., Isaeva, D.M., Musychenko, M.L. and Tzareva, N.V. (1991) Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. *Nucleic Acids Res*, **19**, 155-162.
265. Farwell, M.A., Roberts, M.W. and Rabinowitz, J.C. (1992) The effect of ribosomal protein S1 from *Escherichia coli* and *Micrococcus luteus* on protein synthesis in vitro by *E. coli* and *Bacillus subtilis*. *Mol Microbiol*, **6**, 3375-3383.
266. Komarova, A.V., Tchufistova, L.S., Dreyfus, M. and Boni, I.V. (2005) AU-rich sequences within 5' untranslated leaders enhance translation and stabilize mRNA in *Escherichia coli*. *J Bacteriol*, **187**, 1344-1349.
267. Byrgazov, K., Manoharadas, S., Kaberdina, A.C., Vesper, O. and Moll, I. (2012) Direct interaction of the N-terminal domain of ribosomal protein S1 with protein S2 in *Escherichia coli*. *PLoS One*, **7**, e32702.
268. Shteiman-Kotler, A. and Schuster, G. (2000) RNA-binding characteristics of the chloroplast S1-like ribosomal protein CS1. *Nucleic Acids Res*, **28**, 3310-3315.
269. Somanchi, A. and Mayfield, S.P. (1999) Nuclear-chloroplast signalling. *Curr Opin Plant Biol*, **2**, 404-409.

270. Sengupta, J., Agrawal, R.K. and Frank, J. (2001) Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. *Proc Natl Acad Sci U S A*, **98**, 11991-11996.
271. Lauber, M.A., Rappsilber, J. and Reilly, J.P. (2012) Dynamics of ribosomal protein S1 on a bacterial ribosome with cross-linking and mass spectrometry. *Mol Cell Proteomics*, **11**, 1965-1976.
272. Flynn, R.L. and Zou, L. (2010) Oligonucleotide/oligosaccharide-binding fold proteins: a growing family of genome guardians. *Crit Rev Biochem Mol Biol*, **45**, 266-275.
273. Arcus, V. (2002) OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Curr Opin Struct Biol*, **12**, 794-801.
274. Rajkowitsch, L. and Schroeder, R. (2007) Dissecting RNA chaperone activity. *RNA*, **13**, 2053-2060.
275. Qu, X., Lancaster, L., Noller, H.F., Bustamante, C. and Tinoco, I., Jr. (2012) Ribosomal protein S1 unwinds double-stranded RNA in multiple steps. *Proc Natl Acad Sci U S A*, **109**, 14458-14463.
276. Thomas, J.O., Kolb, A. and Szer, W. (1978) Structure of single-stranded nucleic acids in the presence of ribosomal protein S1. *J Mol Biol*, **123**, 163-176.
277. Bear, D.G., Ng, R., Van Derveer, D., Johnson, N.P., Thomas, G., Schleich, T. and Noller, H.F. (1976) Alteration of polynucleotide secondary structure by ribosomal protein S1. *Proc Natl Acad Sci U S A*, **73**, 1824-1828.
278. Agafonov, D.E. and Spirin, A.S. (2004) The ribosome-associated inhibitor A reduces translation errors. *Biochemical and Biophysical Research Communications*, **320**, 354-358.
279. Galmozzi, C.V., Florencio, F.J. and Muro-Pastor, M.I. (2016) The Cyanobacterial Ribosomal-Associated Protein LrtA Is Involved in Post-Stress Survival in *Synechocystis* sp. PCC 6803. *PLoS One*, **11**, e0159346.
280. Birney, E., Kumar, S. and Krainer, A.R. (1993) Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res*, **21**, 5803-5816.
281. Laursen, B.S., Sorensen, H.P., Mortensen, K.K. and Sperling-Petersen, H.U. (2005) Initiation of protein synthesis in bacteria. *Microbiol Mol Biol Rev*, **69**, 101-123.
282. Sun, Q., Vila-Sanjurjo, A. and O'Connor, M. (2011) Mutations in the intersubunit bridge regions of 16S rRNA affect decoding and subunit-subunit interactions on the 70S ribosome. *Nucleic Acids Res*, **39**, 3321-3330.
283. Bradshaw, N. and Walter, P. (2007) The signal recognition particle (SRP) RNA links conformational changes in the SRP to protein targeting. *Mol Biol Cell*, **18**, 2728-2734.

8. APPENDIX

Pairwise global sequence alignment between 23S rRNAs from spinach cp- and *E. coli* (257) ribosomes

```
#####
#
# Aligned_sequences: 2
# 1: spinach23S
# 2: ecoli23S
# Matrix: EDNAFULL
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 3120
# Identity:      2018/3120 (64.7%)
# Similarity:    2018/3120 (64.7%)
# Gaps:          530/3120 (17.0%)
# Score: 6663.5
#
#
#####

spinach23S      1 --UCAAACGAGGAAAGGCUU--ACGGUGGAUACCUAGGCACCCAGAGAC      46
      || ||.|||.|| ||. | ||||| |||.|||.|||.|||.|||.||
ecoli23S        1 GGUU-AAGCGACUAA--GCGUACACGGUGGAUGCCCUGGCAGUCAGAGGC      47

spinach23S      47 GAGGAAGGGCGUAUUAUCGACGAAAUGCUUCGGGGAGUUGAAA-----      90
      ||.|||||.|||.|||.|||.|||.|||.|||.|||.|||.|||.||
ecoli23S        48 GAUGAAGGACGUGCUAUCUGCGAUAAAGCGUCGGUAAGGUGAUUGAACC      97

spinach23S      91 ---AUAAGCAGAGAUCCGGAGAUUCCCGAAUAGGUCAACC-----UUU      130
      |||| | |||.|||||.|||||.|||.|||| | |||
ecoli23S        98 GUUAUAA-----CCGGCGAUUUCGAAUGGGGAAACC CAGUGUGUUU      139

spinach23S     131 CGA---ACU---UCUGCUGAAUCCAUGGGCAGG--CAAGAGAC-AACCU      170
      ||| ||| |...||||| ||| |||.|||.|||.|||.|||.|||.||
ecoli23S     140 CGACACACUAUCAUUAACUGAAUCCAU----AGGUUAAUGAGGCGAACC      185

spinach23S     171 GGCGAACUGAAACAUCUUAUGUAGCCAGAGGAAAAGAAAGCAAAGCGAUU      220
      ||.||||| |||.|||.|||.|||.|||.|||.|||.|||.|||.|||.||
ecoli23S     186 GGGGAACUGAAACAUCUAAAGUACCCCGAGGAAAAGAAUCAACCGAGAUU      235

spinach23S     221 CCCGUAGUAGCGGCGAGCGAAAUGGGAGCAGCCUAAACCGUGAAAACGGG      270
      |||.||||| |||.|||.|||.|||.|||.|||.|||.|||.|||.|||.||
ecoli23S     236 CCCCAGUAGCGGCGAGCGAACGGGGAGCAGCCAGAGCCUG-AAUCAGU      284

spinach23S     271 GU-UGUGGGAGAGCAAUACAAGCGUC GUGCUGCUAGGCGAAUCAGUGGA-      318
      || |||||.|| |.|||||| | |||
ecoli23S     285 GUGUGUGUUAG-----UGGAAGCGUC-----UGGAA      310

spinach23S     319 --GUGCG-GAACCCUAGAUGGUGAAAAGUCCAGUAGCCGAAAAGCAU-CACU      364
      |.||| || |||.|||||.|||.|||.|||.|||.||| || |||
ecoli23S     311 AGGCGCGCGA----UACAGGGUGACAGCCCCGUACACAAAA--AUGCAC-      353

spinach23S     365 AGCUUAUGCUCUGACC---CGAGUAGCAUGGGGCACGUGGAUCCCGUG      410
      |||||.|||.|||.|||||.|||.|||||.|||.|||.|||.|||.|||.||
ecoli23S     354 -----AUGCUGUGAGCUGAUGAGUAGGGCGGGACACGUGGUAUCCUGUC      398

spinach23S     411 UGAAUCAGCAAGGACCACCUUGCAAGGCUAAAUACUCCUGGGUGACCGAU      460
      |||||.|||.|||.|||.|||.|||||.|||.|||||.|||.|||||.|||.||
ecoli23S     399 UGAAUAGGGGGGACCAUCCUCCAAGGCUAAAUACUCCUGACUGACCGAU      448

spinach23S     461 AGCGAAGUAGUACCGUGAGGGAAGGGUGAAAAGAACCC---CCAUCGGGG      507
      ||.|||||.||||| |||.|||.|||||.|||.|||.|||.|||.|||.|||.||
ecoli23S     449 AGUGAACCAUACCGUGAGGGAAGGCGAAAAGAACCCCGGCGA--GGGG      496

spinach23S     508 AGUGAAAUAGAACAUGAAACCGUAAGCUCUCAAGCAGUGGGAGGGGGACC      557
      |||||.|||.|||.|||||.|||.|||.|||.|||.|||||.|||.|||||.||
ecoli23S     497 AGUGAAAAGAACCUGAAACCGUGUACGUACAAGCAGUGGGAG-----      539

spinach23S     558 AGACCC-----UGACCGCGUGCCUGUUGAAGAAUGAGCCGGCGACU      598
      .||. | ||||.|||||.|||.|||.|||.|||.|||.|||.|||.|||
```

ecoli23S	540	-CACGCUUAGGCGUGUGACUGCGUACCUUUUGUAUAAUGGGUCAGCGACU	588
spinach23S	599	CAUAGGCAGUGGCUUGGUUAAAGGAACCCACCGAGCCGUAGCGAAAGCG	648
ecoli23S	589	UAUAUUCUGUAGCAAGGUUAAACCGAA--UAGGGGAGCCGAAGGGAAACCG	636
spinach23S	649	AGUCUUCAUAGGGC----AAUUGUCACUGCUUAUGGACCCGAACCUGGGU	694
ecoli23S	637	AGUCUUAACUGGGCGUUAAGUUG---CAGGGUUAUAGACCCGAAACCCGGU	683
spinach23S	695	GAUCUAUCCAUGACCAGGAUGAAGCUUGGGUGAAACUAAGUGGAGGUCCG	744
ecoli23S	684	GAUCUAGCCAUGGGCAGGUUGAAGGUUGGGUAAACACUAACUGGAGGACCG	733
spinach23S	745	AACCGACUGAUGUUGAAGAAUCAGCGGAUGAGUUGUGGUUAGGGGUGAAA	794
ecoli23S	734	AACCGACUAAUNNNGAAAAUAGCGGAUGACUUGUGGCUGGGGUGAAA	783
spinach23S	795	UGCCACUCGAACCCAGAGCUAGCUGGUUCUCCCCGAAAUAGC-GUUGAG--	841
ecoli23S	784	GGCCAAUCAAAACCGGGAGAUAGCUGGUUCUCCCCGAAA-GCUAUUUAGGU	832
spinach23S	842	-GCGCAGCAGUUGACUGGACAUCU--AGGGGUAAAGCACUGUUUCGG--U	886
ecoli23S	833	AGCGC--CUCGUGAAU--UCAUCUCCGGGGUAGAGCACUGUUUCGGCAA	878
spinach23S	887	GCGGGCCGCGAGAGCGGUACCAAAUCGAGGCAACUCUGAAUAC--UAGA	934
ecoli23S	879	GGGGGUCACGA---CUUACCAACCCGAUGCAAACUGCGAAUACCGGAGA	924
spinach23S	935	UAUGCCUCCAAAUAACAGGGGUCAGGUCGGCGAGUGAGACGAUGGGGG	984
ecoli23S	925	-AUG-----UUAUCACGGG-----AGACACACGGCGGGNG	953
spinach23S	985	AUAAGCUUCAUCGUCGAGAGGGAAACAGCCCGGAUACCAGCUAAGGCC	1034
ecoli23S	954	CUAAGCUGCCGUCGUGAAGAGGGAAACAACCCAGACCGCCAGCUAAGGUCC	1003
spinach23S	1035	CUAAAUGACCGCUCAGUGAUAAAGGAGGUAGGGGUGCAGAGACAGCCAGG	1084
ecoli23S	1004	CAAAGUCAUGGUUAAAGUGGGAAACGAUGUGGGAAGGCCAGACAGCCAGG	1053
spinach23S	1085	AGGUUUGCCUAGAAGCAGCCACCCUUGAAAGAGUGCGUAAUAGCUCACUG	1134
ecoli23S	1054	AUGUUGGCUUAGAAGCAGCCAUCAUUUAAAGAAAGCGUAAUAGCUCACUG	1103
spinach23S	1135	AUCGAG-CGCUCUUGCGCCGAAGAUG-AACGGGGCUAAGCGGUCUGCCGA	1182
ecoli23S	1104	GUCGAGUCG-GCCUGCGCGGAAGAUGUAACGGGGCUAAACCAUGCACCGA	1152
spinach23S	1183	AGCUG-----UGGGAUGUAAAAAACAUCGGUAGGGGAGC	1217
ecoli23S	1153	AGCUGCGGCAGCGACGCUUAUGCGUUGU-----UGGGUAGGGGAGC	1193
spinach23S	1218	GUUCCGUGUUGAG----GGAGAAACGCGUGC-GUGAGCCGCGUUGGACGA	1261
ecoli23S	1194	GUUC--UGUAAGCCUGCG---AAGGUGUGCUGUGAGGCAUGCUGGAGGU	1237
spinach23S	1262	AGCGGAAGCGAGAAUGUCGGCUUGAGUAACGCAAAACAUUGGUGAGAAUCC	1311
ecoli23S	1238	AUCAGAAGUGCGAAUGCUGACAUAAGUAACGAUAAAGCGGGUGAAAAGCC	1287
spinach23S	1312	AAUGC---CCCGAAAACCUAAGGGUUCUCCGCAAGGUUCGUCCACGGAG	1358
ecoli23S	1288	--CGCUCGCCGGAAGACC-AAGGGUUCUGUCCAACGUUAAUCGGGGCAG	1334
spinach23S	1359	GGUGAGUCAGGGCCUAAGAUCAGGCCGAAAGGCGUAGUCGAUGGACAACA	1408
ecoli23S	1335	GGUGAGUCGACCCCUAAGGCGAGGCCGAAAGGCGUAGUCGAUGGGAAACA	1384
spinach23S	1409	GGUGAAUAUUCUGUACUACCCCUUGUUGGUCC--CGA---GGGACGGA	1452
ecoli23S	1385	GGUUAUAUUCUGUA-----CUUGGUGUACUGCGAAGGGGGACGGA	1428
spinach23S	1453	GGAGGCUAGGUUAGCCGAAAGAUGGUUAU--CGGUUCAAGGACGCAAGGU	1500

ecoli23S	1429	GAAGGCUAUGUUGGCCGGGCGACGGUUGUCCCGUUUAAGCGUGUAGGCU	1478
spinach23S	1501	GACCCUGUUUUUCAGGGUAAGAAGGGGUAGAGAAAAU---GCCUCGAGCC	1547
ecoli23S	1479	G-----GUUUUCCAGGCAAA-----UCCGGA AAAUCAAGGCU-GAGGC	1515
spinach23S	1548	AAUGUUCGAGUACACAGGCGCUACGGCGCUGAAGUAACCGAUGCCAUACUC	1597
ecoli23S	1516	---GU--GAUGACGAGGCACUACGGUGCUGAAGCAACAAUGCCCGUCUU	1560
spinach23S	1598	CCAGGAAAAG-CUCGAACGACCUUCAACAAAAGGGUACC-----UGUA	1639
ecoli23S	1561	CCAGGAAAAGCCUCUA---AGCAUCA-----GGUAACAUCAAUCGUA	1600
spinach23S	1640	CCCGAAACCGACACAGGUAGGU-AGGUAGAGAAUACCUAGGGGCGCGAGA	1688
ecoli23S	1601	CCCCAAACCGACACNGGU-GGUCAGGUAGAGAAUACCAAGGCGCUUGAGA	1649
spinach23S	1689	CAACUCUCUCUAAGGAACUCGGCAAAAUAGCCCCGUAACUUCGGGAGAAG	1738
ecoli23S	1650	GAACUCGGGUGAAGGAACUAGGC AAAAUGGUGCCGUAACUUCGGGAGAAG	1699
spinach23S	1739	G-----GGU--GCCCCUCACAAAGGG-----GGUC	1762
ecoli23S	1700	GCACGCUGAU AUGUAGGUGAGGUCCUCGCGGAUGGAGCUGAAAUACAGUC	1749
spinach23S	1763	GAAGUGACCAGGCCCGG--GCGACUGUUUACCAAAAACACAGGUCUCCGC	1810
ecoli23S	1750	GAAGAUACCAG--CUGGCGCAACUGUUUAUUA AAAACACAGCAGUCUGC	1797
spinach23S	1811	AAAGUCGUAAGACCAUGUAUGGGGGCUGACGCCUGCCCAGUGCCGGAAGG	1860
ecoli23S	1798	AAACACGAAAGUGGACGUAUACGGUGUGACGCCUNCCCGGUGCCGGGAGG	1847
spinach23S	1861	UCAAGGAAGUUGGUGACCUGAUGACAGGGG--AGC-----CG--GC----	1897
ecoli23S	1848	UUAA-----UUGAU-----GGG GUUAGCGCAAGCGAAGCUCUU	1880
spinach23S	1898	GACCGAAGCCCCGGUGAACGGCGGCCGUAACUAUAACGGUCCUAAGGUAG	1947
ecoli23S	1881	GAUCGAAGCCCCGGUA AACGGCGGCCGNAACNANAACGGUCCUAAGGUAG	1930
spinach23S	1948	CGAAAUUCCUUGUCGGGUAAGUUCGACCCGCACGAAAGGCGUAACGAUC	1997
ecoli23S	1931	CGAAANUCCUUGUCGGGUAAGUUCGACNUGCACGAAUGGCGUAAUGA--	1978
spinach23S	1998	UGGGCA--CUGUCU---CGGAGAGAGGCUCGGUGAAAUAGACAUGUCUG	2041
ecoli23S	1979	UGGCCAGGCUGUCUCCACCCGAGA---CUCAGUGAAAUUGAACUCGCUG	2024
spinach23S	2042	UGAAGAU GCGGACUACCUGCACCUGGACAGAAAGACCCUAUGAAGCUUUA	2091
ecoli23S	2025	UGNAGAUGCAGUGUACCCGCGGCAAGACGGAAGACCCCGUNAACCUUUA	2074
spinach23S	2092	CUGUCCCCUGGGAUUG-GCUUUGGGCUUUUCCUGCGCAGCUUAGGUGGAA	2140
ecoli23S	2075	CUAUAGCUUGACACUGAACAUUGAGC-CUUGAUGUGUAGGAUAGGUGGGA	2123
spinach23S	2141	GGCGAAGAAGGCCCCCUUCCGGGGGGGCC-----CGAGCCAUCAGU	2181
ecoli23S	2124	GGCUUUGAAG-----UGUGGACGCCAGUCUGCAUGGAGCCGACCUU	2164
spinach23S	2182	GAGAUACCACUCUGGAAGAGCUAGA-AUUCUAACCUUGUGUCAGGACC--	2228
ecoli23S	2165	GAAAUACCACCCUUUAA-UGUUUGAUGUUCUAACGUU-----GACCCG	2206
spinach23S	2229	---UACGG--GCCAAGGGACAUUCUCAGGUAGACAGUUUCUAUGGGGCG	2272
ecoli23S	2207	UAAUCCGGGUUGC---GGACAGUGUCUGGUGGGUAGUUUGACUGNGGCG	2252
spinach23S	2273	UAGGCCUCCCAAAAGGUAAACGGAGGCGUGCAAAGGUU-----UCCUCGG	2316
ecoli23S	2253	GUCUCCUCCUAAAGAGUAACGGAGGAGCACGAAGGUUGGCUAAUCCU--G	2300
spinach23S	2317	GCCGGAC---GGAGAUUGGCCCUCCGAGUGCAAAGGCAGAAAGGGAGCUUG	2362

ecoli23S	2301	GUCGGACAUCAGGAGGUU-----AGUGCAAUGGCAUAAGCCAGCUUG	2342
spinach23S	2363	ACUGCAAGACCCACCCGU-----CGAGCAGGGACGAAAGUCGGCCUUA	2405
ecoli23S	2343	ACUGCGAG-----CGUGACGGCGCGAGCAGGUGCGAAAGCAGGUCAUA	2385
spinach23S	2406	GUGAUCCGACGGUGCCGAGUGGAAGGGCCGUCGCUAACGGAUAAAAGUU	2455
ecoli23S	2386	GUGAUCCGGUGGUUCUGAAUGGAAGGGCCAUCGCUAACGGAUAAAAGGU	2435
spinach23S	2456	ACUCUAGGGAUAACAGGCUGAUCUCCCCAAGAGUUCACAUCGACGGGAA	2505
ecoli23S	2436	ACUCCGNGGANAACAGGCNGAUACCGCCAAGAGUUCAUUAUCGACGGCGG	2485
spinach23S	2506	GGUUUGGCACCUCGAUGUCGGCUCUUCGCCACCUGGGGCGUAGUAUGUU	2555
ecoli23S	2486	UGUUUGGCANCUCGNNUGCGGCUCAUCAUCCUGGGGCGUAGUAGGUC	2535
spinach23S	2556	CCAAGGGUUGGGCUGUUCGCCCAUUAAGCGGUACGUGAGCUGGGUUCAG	2605
ecoli23S	2536	CCAAGGGUAUGGCNGUUCGCCAUUUAAGUGGUACGCGAGCNGGGUUUAG	2585
spinach23S	2606	AACGUCGUGAGACAGUUCGGUCCAUAUCCGGUGUGGGCGUAGAGCAUUG	2655
ecoli23S	2586	AACGUCGUGAGACAGNNCGGUCCUAUCUGCCGUGGGGCGUGGAGAACUG	2635
spinach23S	2656	AGAGGACCUUCCCUAGUACGAGAGGACCGGGAAGGACGCACCUCUGGUG	2705
ecoli23S	2636	AGGGGGGCGUCUCCUAGUACGAGAGGACCGGAGUGGACGCAUCACUGGUG	2685
spinach23S	2706	UACCAGUUAUCGUGCCACGGUAAACGCUG---GGUAGCCAAGUGCGGAG	2752
ecoli23S	2686	UUCGGGUUGUCAUGCCAAUGG---CACUGCCCGGUAGCUAAAUGCGGAA	2731
spinach23S	2753	CGGAUAACUGCUGAAAGCAUCUAAGUAGUAAGCCCACCCCAAGAUGAGUG	2802
ecoli23S	2732	GAGAUAAUGUGCUGAAAGCAUCUAAGCACGAAACUUGCCCCGAGAUAGUU	2781
spinach23S	2803	CUCUCCUA-----	2810
ecoli23S	2782	CUC-CCUGACCCUUUAAGGGUCCUGAAGGAACGUUGAAGACGACGACGUU	2830
spinach23S	2811	-----	2810
ecoli23S	2831	GAUAGGCCGGGUGUGUAAGCGCAGCGAUGCGUUGAGCUAACCGGUACUAA	2880
spinach23S	2811	-----	2810
ecoli23S	2881	UGAACCGUGAGGCUUAACCU	2900

#-----
#-----

Appendix Figure 1

Pairwise global sequence alignment between 23S rRNAs from spinach cp- and *E. coli* (86) ribosomes

```
#=====
#
# Aligned_sequences: 2
# 1: spinach5S
# 2: ecoli5S
# Matrix: EDNAFULL
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 138
# Identity:      81/138 (58.7%)
# Similarity:    81/138 (58.7%)
# Gaps:          35/138 (25.4%)
# Score: 168.0
#
#=====

spinach5S      1 UAUU--CUGGUGUCC-UAG-GCGUAGAGGAACCAC---ACCAAUCCAUCC
43
      |  |||.|.|.|| ||| |||  |.||..||||  |||  |||.|
ecoli5S        1 ---UGCCUGGCGGCCCGUAGCGCG--GUGGUCCCACCUGACC---CCAUGC
42

spinach5S      44 CGAACUUGGUGGUUAAACUC--UACUGCGGUGACGAUACUGUAGGGGAGG
91
      |||||..|.|.|.|.|.|.  ||  |||.|.|.|.|.|.  |||...
ecoli5S        43 CGAACUCAGAAGUGAAACGCCGUA--GCGCCGAUGGUAGUGU--GGGGUC
88

spinach5S      92 UCC---UGC GGAAAAAUAG----CUCGACGCCAGG-AU      121
      |||  ||||  |.|.|||  ||  ||||| ||
ecoli5S        89 UCCCCAUGCG--AGAGUAGGGAACU----GCCAGGCAU      120

#-----
#-----
```

Appendix Figure 2

Secondary structure of cp-ribosome 23S rRNA

Appendix Figure 3 (Taken from Bieri et al. (114))

Comparison between secondary structures of the 23S rRNA from cp- and *E. coli* ribosomes

Appendix Figure 4 (Taken from Sharma et al. (124))

Pairwise global sequence alignment between 23S rRNAs from spinach cp- and *E. coli* (86) ribosomes

```
#=====
#
# Aligned_sequences: 2
# 1: spinach4.8S
# 2: ecoli23S
# Matrix: EDNAFULL
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 131
# Identity:      68/131 (51.9%)
# Similarity:    68/131 (51.9%)
# Gaps:          41/131 (31.3%)
# Score: 143.0
#
#=====

spinach4.8S      1  -----AG-----AGAAGGUCACGGCG-AGACGA-GCCGUUUAUC
32
                        ||      .||||| .|||. | ||||| |.||||
ecoli23S         1  CUGACCCUUUAAGGGUCCUGAAGG-AACGUUGAAGACGACGACGUU----
45

spinach4.8S      33  AUUACGAUAG---GUGUCAAGUGGAAGUGCAGUGAUGUAUGCAGCUGA--
77
                        ||||| |.||  |||.|||.||||.||||..|..|||. |
ecoli23S         46  -----GAUAGGCCGGGU---GUGUAAGCGCAGCGAUGCGUUGAGCUAACC
87

spinach4.8S      78  GGCAUCCUAACAGACCCACAGACUUGAAC--      106
                        || |.|||||...|||...|||.||| |||
ecoli23S         88  GG--UACUAAUGAACCGUGAGGCUU-AACCU      115

#-----
#-----
```

Appendix Figure 5

Comparison between secondary structures of the 16S rRNA from cp- and *E. coli* ribosomes

Appendix Figure 6 (Taken from Sharma et al. (124))

Secondary structure of cp-ribosome 16S rRNA

Appendix Figure 7 (Taken from Bieri et al. (114))

Enlarged view of Figure 1.9 (B)

Appendix Figure 8

9. Information regarding online access of cryo-EM maps and models

1. 200-kV Arctica-dataset:

The 3.5 Å cryo-EM density map of the spinach cp-ribosome LSU has been deposited to the Electron Microscopy Data Bank (EMDB) with accession code EMD-9572, the coordinates of the atomic structure have been deposited to the Protein Data Bank under accession code 5H1S.

2. 300-kV Krios-dataset:

The 3.4 Å 70S, 3.7 Å SSU and 3.3 Å LSU cryo-EM density maps of the spinach cp-ribosome have been deposited to the EMDB with accession codes EMD-6709, EMD-6710 and EMD-6711, respectively. The coordinates of the atomic structures of the 70S ribosome, SSU and LSU have been deposited to the Protein Data Bank as 5X8P, 5X8R and 5X8T, respectively.

10. Author's publications related to this thesis during the Ph.D. period- Jan 2014 to Jan 2018

Research articles:

- [1] Ahmed, T., Shi, J. and Bhushan, S. (2017) Unique localization of the plastid-specific ribosomal proteins in the chloroplast ribosome small subunit provides mechanistic insights into the chloroplastic translation. *Nucleic Acids Res*, **45**, 8581-8595.
- [2] Ahmed, T., Yin, Z. and Bhushan, S. (2016) Cryo-EM structure of the large subunit of the spinach chloroplast ribosome. *Sci Rep*, **6**, 35793.

Talk presentation:

9th Annual Retreat, Konstanz Research School Chemical Biology (KoRS-CB), Konstanz, Germany. Aug 23-25, 2017.

Poster presentation:

6th International Caesar Conference on “Overcoming barriers, atomic resolution and beyond: advances in Molecular Electron Microscopy”, Bonn, Germany. Sep 12-15, 2017.