Poorly understood aspects of ribosome biology: An untapped resource for novel antibiotics

Dushyant Narendra Jahagirdar

Department of Anatomy and Cell Biology
McGill University, Montreal
Quebec, Canada

May 2022

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Table of Contents

Abstract iii
Résumé iv
Acknowledgments v
Original Contributions to Knowledge vii
List of Figures ix
List of Tables xi
List of Abbreviations xii
CHAPTER 1: Introduction and Literature Review 1

1.1. Overview of bacterial ribosomes 1

1.2. Role of magnesium for ribosome structural integrity and function 3

1.3 Bacterial ribosome biogenesis 4

1.4 Overview of canonical translation mechanism in bacteria 13

1.5 Built-in mechanisms of Ribosome Fidelity 17

1.6 Antibiotics mainly target the protein synthesis aspect of the bacterial ribosome 19

1.7 Cryo-EM, ribosomes, and air-water interface 20

1.8 Thesis Objective 24

CHAPTER 2: Alternative Conformations and Motions Adopted by 30S Ribosomal Subunits Visualized by Cryo-Electron Microscopy 26

2.1 Abstract 27

2.2 Introduction 28

2.3 Materials and Methods 31

2.4 Results 38

2.5 Discussion 52

2.6 Supplemental Material 57

2.7 References 63

Connecting Text 67
CHAPTER 3: Discovering the Role of YjeQ in Catalysis of Small Ribosome Subunit Assembly, Translational Fidelity, and Ribosome Rescue 68

3.1 Abstract 69

3.2 Introduction 70

3.3 Materials and Methods 73

3.4 Results 82

3.5 Discussion 96

3.6 Supplemental Material 101

3.7 References 106

Connecting Text 110
CHAPTER 4: Structural basis of sequestration of the Anti-Shine-Dalgarno sequence in the Bacteroidetes ribosome 111

4.1 Abstract 112

4.2 Introduction 113

4.3 Material and Methods 116

4.4 Results 128

4.5 Discussion 150

4.6 Supplementary Data 156

4.7 References 157

CHAPTER 5: Discussion and Conclusion 165
Bibliography 174

Abstract

The discovery of antibiotics in the late 1920s was a major medical breakthrough. However, the global rise of antibiotic resistance and the lack of new antibiotics threaten life-saving medical procedures and our ability to treat infectious diseases. Accordingly, we need more potent antibiotics that have novel action mechanisms. Many antibiotics used in the clinic target the bacterial ribosome since protein synthesis is central to all biological processes, and this mechanism is well understood. However, other aspects of ribosome biology remain poorly understood and constitute untapped targets for novel antibiotics. My thesis aims to understand these unexplored aspects of ribosome biology. While we are familiar with the 30S structure from X-ray crystallography, biochemical studies have hinted at an alternative state of the 30S. Accordingly, I solved the Escherichia coli small ribosome subunit (30S) structure by cryo-electron microscopy (cryo-EM). I found that in solution, the 30S subunit displays dynamic movements contrary to the static X-ray crystallography structure we have become familiar with over the last two decades. Translational fidelity is another poorly understood aspect of the ribosome that could be exploited as an antibiotic target. YjeQ GTPase is among many ribosome assembly factors that bind to the 30S. However, this interaction is unique since YjeQ induces conformational changes on the 30S, typically observed during messenger RNA decoding. Combining cryo-EM with biochemical assays, I show how yjeQ deletion/mutation affects the stability of critical 30S regions and reduces the translational accuracy of the ribosome. My final thesis chapter shows how translation initiation mechanisms are not conserved through our bacterial kingdom. Canonical translation initiation requires the recognition of the Shine-Dalgarno sequence (SD) on the messenger RNA (mRNA) by the anti-Shine-Dalgarno sequence (ASD) on the ribosome. However, in the case of Bacteroidetes, recognition of these two components is inconsequential, yet these bacteria efficiently initiate the translation mechanism. Our high-resolution cryo-EM structure demonstrates that Bacteroidetes sequester ASD and therefore do not recognize SD and propose a mechanism for translation autoregulation during ribosome biogenesis. Overall, my thesis lays the groundwork for leveraging uncharacterized aspects of ribosome biology to develop new antibiotic targets.

Résumé

La découverte des antibiotiques à la fin des années 1920 a constitué une avancée médicale majeure. Cependant, l'augmentation mondiale de la résistance aux antibiotiques et le manque de nouveaux antibiotiques menacent des procédures médicales vitales et notre capacité à traiter les maladies infectieuses. Par conséquent, nous avons besoin d'antibiotiques plus puissants et dotés de nouveaux mécanismes d'action. De nombreux antibiotiques utilisés en clinique ciblent le ribosome bactérien, car la synthèse des protéines est au cœur de tous les processus biologiques, et ce mécanisme est bien compris. Cependant, d'autres aspects de la biologie du ribosome restent mal compris et constituent des cibles inexploitées pour de nouveaux antibiotiques. Ma thèse vise à comprendre ces aspects inexplorés de la biologie du ribosome. Bien que la structure du 30S nous soit familière grâce à la cristallographie aux rayons X, des études biochimiques ont laissé entrevoir un autre état du 30S. J'ai donc résolu la structure de la petite sous-unité du ribosome (30S) d'Escherichia coli par cryo-microscopie électronique (cryo-EM). J'ai découvert qu'en solution, la sous-unité 30S présente des mouvements dynamiques, contrairement à la structure statique de la cristallographie aux rayons X que nous connaissons depuis deux décennies. La fidélité translationnelle est un autre aspect mal connu du ribosome qui pourrait être exploité comme cible des antibiotiques. La GTPase YjeQ fait partie des nombreux facteurs d'assemblage du ribosome qui se lient au 30S. Cependant, cette interaction est unique car YjeQ induit des changements de conformation sur le 30S, typiquement observés lors du décodage des ARN messagers. En combinant la cryo-EM avec des essais biochimiques, je montre comment la délétion/mutation de YjeQ affecte la stabilité des régions critiques du 30S et réduit la précision de la traduction du ribosome. Le dernier chapitre de ma thèse montre que les mécanismes d'initiation de la traduction ne sont pas conservés dans notre règne bactérien. L'initiation canonique de la traduction nécessite la reconnaissance de la séquence Shine-Dalgarno (SD) sur l'ARN messager (ARNm) par la séquence anti-Shine-Dalgarno (ASD) sur le ribosome. Cependant, dans le cas des Bacteroidetes, la reconnaissance de ces deux composants est sans conséquence, pourtant ces bactéries initient efficacement le mécanisme de traduction. Notre structure cryo-EM à haute résolution démontre que les Bacteroidetes séquestrent l'ASD et ne reconnaissent donc pas la SD et propose un mécanisme d'autorégulation de la traduction pendant la biogenèse du ribosome. Dans l'ensemble, ma thèse jette les bases de l'exploitation des aspects non caractérisés de la biologie des ribosomes pour développer de nouvelles cibles antibiotiques.

Acknowledgments

First and foremost, I want to thank my supervisor Dr. Joaquin Ortega for accepting me into his laboratory, creating countless opportunities for me to learn and succeed, and always being available for advice and encouragement. I cannot thank you enough for going above and beyond to help me with so many things that have transpired over the last six years, especially for supporting me in reuniting with my wife during the uncertainties of the COVID-19 pandemic. I am grateful for every minute you invested in me because it has played a significant role in my personal growth.

I want to thank the present, and the past members of the Ortega laboratory, each one of you has made this journey incredible. I want to thank Dr. Aida Razi for taking me under her wing during my initial years in the laboratory. Studying for ANAT-690D and journal clubs wouldn't be fun without you. To Armando Palacios, I have had the best time in this laboratory in your company, and I want to thank you for always being there to help and stimulating discussions.

I am thankful to my advisory committee members, Dr. Alfredo Ribeiro-da-Silva, Dr. Martin Schmeing, Dr. Javier Vargas, and Dr. Khanh Huy Bui, for their guidance and valuable advice. I also want to thank Dr. Alba Guarné for unlimited advice, access to her laboratory, and overall support throughout these years. I want to thank Dr. Anusha Naganathan and Dr. Gloria Culver for their collaborative expertise with Miller's Fidelity Assays. I also want to thank Dr. Kurt Fredrick and his extended team for their invaluable collaboration with the *Flavobacterium Johnsoniae* ribosome project.

Special thanks to all the staff at the Facility for Electron Microscopy Research (FEMR) for providing me with high-quality data throughout my time at McGill. Specifically, I want to thank Dr. Kaustuv Basu for training me and working with diligence to collect the highest quality cryo-EM data. I also want to thank Dr. S. Kelly Sears and Dr. Eric Boucher for their non-stop support within the workplace.

I also want to thank *Centre de Recherche en Biologie Structurale* (CRBS) for the studentship award and several opportunities to showcase my research. I am also grateful to the McGill Anatomy and Cell Biology Department for several fellowships. Thanks to Ms. Lisa Kush at the Biochemistry and Biomedical Sciences Department, McMaster University, and Ms. Joelle Denomy-Hasilo, ACB, McGill University, for their help with admissions and transition from transition McMaster to McGill and the overall administrative process.

None of this would have happened without the unconditional support from my family and friends. Words cannot describe how thankful I am to my mother, Supriya Jahagirdar, who never stopped believing in me. Without her support and determination, I would have never reached here. I cannot thank her enough for holding our family together through the tough times after my father passed away so that I could focus on my career. I also want to thank my father, Dr. Narendra Jahagirdar, for instilling the fighting spirit in me. I have missed you a lot, and I hope I have made you proud. I also want to thank my younger brother, Shrinivas Jahagirdar, for stepping up and supporting our family when I couldn't be there. To my aunt, Anagha Tongaonkar, and uncle, Rahul Phadnis, without your unconditional support, Smruti and I could not have reunited in this pandemic if it wasn't for you. I also want to thank my friends Atharva Gajendragadkar and Dr. C Jason Baby; I am more thankful for our friendship than you'll ever know. I could have never imagined having two of my best friends on different continents and yet being easily reachable whenever I needed you both. To my friends Ajinkya Ghagre and Chaitra Shetty, you were my home away from home. Thank you so much for your companionship over the last five years.

Finally, to my wife, Smruti, I am eternally grateful to you for being with me through thick and thin. Thank you for believing in me, bearing my frustrations and stress, and constantly encouraging me. I could not have done this without your support and patience.

Original Contributions to Knowledge

Chapter 2

Until the 'resolution-revolution' in cryo-EM, high-resolution structural information was solely contributed through X-ray crystallography. Consequently, only conformational states of macromolecules that could be stabilized in a crystal lattice were accessible. In the case of the bacterial 30S ribosomal subunit, it was reported five decades ago that 30S subunits in solution readily interconvert between 'active' and 'inactive' conformations. The structure of the 'active' 30S subunit was described by crystallography in the year 2000, but the structure of the inactive form remained unsolved. Using cryo-EM, we obtained a high-resolution structure of the 'inactive' conformation which presents a distinct conformation in the functional center of the 30S ribosomal subunit. This study will advance our understanding of how the 'active-inactive' transition of the 30S governs the accessibility of several protein factors involved in translation initiation and ribosome biogenesis. Finally, many researchers use cryo-EM to study ribosome biogenesis which typically involves analyzing a heterogeneous mixture of 30S ribosome subunits that may or may not resemble the mature 30S. Therefore, cryo-EM structures obtained in this study will serve as a comparative reference to draw meaningful conclusions for ribosome biology studies aimed at understanding the assembly and maturation of the 30S ribosome subunit.

Chapter 3

YjeQ is a biochemically and structurally well-characterized GTPase protein that plays a role in assembling the 30S ribosome subunit. How YjeQ catalyzes ribosome assembly and promotes the maturation of the decoding region remains unknown. Using cryo-EM, we uncover the mechanism through which YjeQ ensures that the 30S assembly follows the canonical pathway for folding

different 30S domains. We also reveal the mechanism YjeQ uses to ensure that the mature 30S ribosome subunits can accurately perform the decoding step during protein synthesis. Finally, our high-resolution structural analysis of the YjeQ bound to the 30S reveals a potential role for YjeQ in rescuing ribosomes that may abruptly stall during protein synthesis.

Chapter 4

The Bacteroidetes are a large phylum of bacteria that are significant clinical pathogens and remain understudied. Gene expression mechanisms like translation initiation in these bacteria are distinct compared to commonly used model organisms such as *Escherichia coli* and *Bacillus subtills*. Typically, translation initiation machinery within most bacteria requires recognition of the Shine-Dalgarno (SD) on messenger RNA by the anti-Shine Dalgarno sequence (ASD) on the ribosome for accurate start codon selection. However, Bacteroidetes can successfully and efficiently initiate translation even without the need for recognition of SD by ASD. This is because these bacteria lack SD sequence in their genes but oddly have retained the ASD sequence on their ribosomes. Reporter gene studies in representative species have shown that introduced SDs fail to stimulate translation, indicating that these ribosomes cannot recognize SD in vivo. How the SD-ASD pairing is prevented in these organisms remained a puzzle. Using cryo-EM, we obtained a 2.8 Å resolution structure of 70S ribosomes from *Flavobacterium johnsoniae*, representative of the Bacteroidetes. This high-resolution structure revealed the structural basis of ASD occlusion in the Bacteroidetes, explaining why these ribosomes are blind to the SD sequence. Our structural analysis also suggests that Bacteroidetes uses the same ASD occlusion mechanism to regulate ribosome biogenesis which is a metabolically expensive process. Additionally, our cryo-EM structure also uncovers a novel ribosomal protein which is a part of the 50S ribosome subunit and named it bL38.

List of Figures

CHAPTER 1

Figure 1. 1 Overview of the Escherichia coli ribosome and its subunits. 2
Figure 1. 2 Co-transcriptional ribosome assembly and rRNA processing. 5
Figure 1. 3 Nomura assembly for 30S assembly. 8
Figure 1. 4 Overview of the Escherichia coli ribosome. 12
Figure 1. 5 Overview of canonical prokaryotic translation cycle. 16
Figure 1. 6 30S assembly intermediates accumulated in the Era-depleted strain. 21
Figure 1. 7 Air-Water Interface in the context of cryo-EM. 23

CHAPTER 2

Figure 2. 1 Cryo-EM structure of the 30S-inactivated-high-Mg2+ particle. 39
Figure 2. 2 Molecular model of the 30S-inactivated-high-Mg2+ particle. 41
Figure 2. 3 Cryo-EM structure of the 30S-inactivated-high-Mg2+ particle in grids containing a continuous layer of carbon. 43
Figure 2. 4 Cryo-EM structure of the 30S-activated-high-Mg2+ particle. 45
Figure 2. 5 Cryo-EM structure of 30S-inactivated-low-Mg2+ particle. 49
Figure 2. 6 Multibody refinement analysis of the cryo-EM structures. 51

Supplemental Figure S2. 1 Resolution analysis of the cryo-EM maps obtained under high Mg2+ concentrations. 57
Supplemental Figure S2. 2 Resolution analysis of the cryo-EM maps obtained under low Mg2+ concentrations. 58

CHAPTER 3

Figure 3. 1 Various 16S rRNA helices remain unfolded in the 30SΔyjeQ particles. 84
Figure 3. 2 Cryo-EM structure and analysis of YjeQ double mutant bound to wildtype mature 86
Figure 3. 3 Cryo-EM structure of YjeQ double mutant treated the 30S in the 'active' state. 88

Figure 3. 4 (A) Effect of deletion of yjeQ on frameshifting errors and stop-codon readthrough in vivo. 91
Figure 3. 5 Cryo-EM characterization of YjeQ N-terminal extension. 95
Figure 3. 6 Proposed model for the role of YjeQ in the maturation of the 30S subunit......97

Figure S3. 1 Image processing workflow for the 30S+YjeQ double mutant dataset. 101
Figure S3. 2 YjeQ double mutant in presence of GMP-PNP can dissociate 70S ribosomes under associating conditions.....102
Figure S3. 3 Resolution analysis and Fourier Shell Correlation (FSC) of the 30S+YjeQ double mutant complex......103
Figure S3. 4 Resolution analysis and Fourier Shell Correlation (FSC) of the 'active' and 'inactive' 30S. 104
Figure S3. 5 Comparison between interaction of YjeQ double mutant and wildtype GTPase domain residues with 16S rRNA. 105

CHAPTER 4

Figure 4. 1 Biochemical evidence that E. coli and F. johnsoniae ribosomes differ in ASD function. 129
Figure 4. 2 Structure of the Flavobacterium johnsoniae ribosome. 131
Figure 4. 3 Comparison of structural differences in the 16S rRNA between **F. johnsoniae** *and* **E. coli.**.....134
Figure 4. 4 Homolog of bS22 in the 30S subunit of F. johnsoniae ribosome. 135
Figure 4.5 Structural Basis for the Sequestration of the ASD sequence within the F. johnsoniae 30S subunit......137
Figure 4. 6 Unique features of bS21 and bS18 of the Bacteroidetes......139
Figure 4. 7 Comparison of structural differences in the 23S rRNA between the 50S subunit in F. johnsoniae and E. coli. 141
Figure 4. 8 A novel ribosomal protein: bL38. 143
Figure 4. 9 Sequences near the start codon of rpsU in various Flavobacteriia.....145
Figure 4. 10 A potential mechanism of bS21 autoregulation......152

List of Tables

CHAPTER 2

Supplemental Table S2. 1 Data Deposition 60

Supplemental Table S2. 2 Cryo-EM data acquisition, processing and map and model statistics. 61

CHAPTER 4

Table 4. 1 Occurrences of strong SD sequences in the Bacteroidetes......146

List of Abbreviations

16S	16S ribosomal ribonucleic acid
17S	17S ribosomal ribonucleic acid/ precursor 16S
30S	bacterial small ribosome subunit
30S PIC	30S Pre-Initiation Complex
30S IC	30S Initiation Complex
50S	bacterial large ribosome subunit
70S	bacterial/prokaryoic ribosome
70SIC	70S Initiation Complex
A	adenine
A-site	A-site aminoacyl site
A260	A260 Light absorbance at 260 nanometres (wavelength)
ASD	Anti-Shine Dalgarno Sequence
ASD2	Anti-Shine Dalgarno Sequence 2
AWI	Air Water Interface
C	Cytosine
CP	Central Protuberance
CTD	C-terminal domain
CTF	Contrast Trasnfer Function
DNA	Deoxyribonucleic acid
DSSR	Dissecting the Spatial Structure of RNA program
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EF	Elongation factor (bacteria)
EM	Electron microscopy
EMDB	Electron Microscopy Data Bank
FPLC	Fast protein liquid chromatography
FRT	Flippase recognition target
FSC	Fourier shell correlation
g	gravitational force
G	Guanine
GC	Guanine and cytosine
GDP	Guanosine diphosphate
GMPPNP	Guanosine 5′-[βγ-imido]triphosphate trisodium salt hydrate
GTP	Guanosine triphosphate
GTPase	Guanosine triphosphate hydrolase
HEPES	N-2-Hydroxyethylpiperazine-N'-2-Ethanesulfonic Acid

IC	Initiation Complex
IF	Initiation Factor
IF1	Initiation Factor 1
IF2	Initiation Factor 2
IF3	Initiation Factor 3
IPC	Isoelectric Point Calculator
IPC2	Isoelectric Point Calculator 2.0
IPTG	Isopropyl ß-D-1-thiogalactopyranoside
kDa	kilo Dalton
KH	domain RNA-binding K homology domain
LB	Luria-Bertani broth viii
MOPS	3-(N-morpholino) propanesulfonic acid
MDa	megadalton
Mg^{2+}	Magnesium ion
MS	Mass spectrometry
mRNA	messenger RNA
MSD	mock Shine Dalgarno
MSD2	mock Shine Dalgarno 2
MST	Microscale Thermophoresis
NCBI	National Center for Biotechnology Information
NEB	New England Biology
Ni^{2+}	Nickel ion
NMR	Nuclear Magnetic Resonance
P1	P1 promoter
PC	Pulse-Chase
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PIC	Pre-Initiation Complex
PTC	peptidyl-transferase center
RF	Release Factor
RF1	Release Factor 1
RF2	Release Factor 2
RF3	Release Factor 3
r-protein	ribosomal protein
rRNA	ribosomal ribonucleic acid
RNA	ribonucleic acid
RRF	Ribosome rescue factor
S	Svedberg unit of measurement (sedimentation rate)

SD	Shine-Dalgarno sequence
SD2	Shine-Dalgarno sequence 2
SDS	sodium dodecyl sulfate
SEC	Size Exclusion Chromatography
SHAPE	selective 2′-hydroxyl acylation analyzed by primer extension
SRL	sarcin-ricin-loop
T	Thiamin
TEV	Tobacco Etch Virus
tRNA	Translation Initiation Region
TIR	Transfer RNA
U	Uracil
UTR	untranslated region
WT	Wild Type
Zn^{2+}	Zinc ion

CHAPTER 1: Introduction and Literature Review

1.1. Overview of bacterial ribosomes

Ribosomes are complex cellular machines responsible for one of the essential processes in all living organisms - translating genetic code into functional proteins. The bacterial 2.5MDa ribosome (70S) consists of two subunits, the small (30S) and the large (50S) subunits **(Figure 1.1)**. While the 30S is responsible for accurately decoding the genetic information encoded in the messenger RNA (mRNA), the 50S functional core containing the peptidyl transferase center catalyzes peptide bond synthesis of amino acids during translation (Ban et al., 2000) (Wimberly et al., 2000). The 30S comprises a single ribosomal RNA (rRNA) molecule (16S) composed of 1542 nucleotides and approximately 21 ribosome proteins (r-proteins) (Wimberly et al., 2000). The 16S rRNA can be further divided into four domains: the body (5' end of 16S rRNA), the platform (central domain), the head (3' major domain), and the 3' minor domain formed by helices 44 and 45 (Noller and Woese, 1981). Apart from including the central decoding region for accurate translation, the 16S rRNA also contains the anti-Shine-Dalgarno (ASD) sequence essential for mRNA recognition during the translation initiation (Shine and Dalgarno, 1974). The distribution of r-proteins is asymmetric, with most of the proteins located on the head, periphery, and solvent sides (back) of the 30S.

The 50S comprises two rRNA molecules, 5S and 23S, composed of 115 and 2904 nucleotides, respectively, and ~34 r-proteins. The 23S rRNA houses the peptidyl transferase center (PTC) responsible for catalyzing the peptide bond synthesis. The 23S rRNA is divided into six secondary structure domains, and 5S rRNA forms the seventh domain. The landmark feature of the 50S subunit includes the L1 stalk, the central protuberance, and the L7/L12 stalk. The 50S also consists of the polypeptide exit tunnel, which is mainly formed by rRNA and a few r-proteins that

contribute to the geometry of this tunnel (Ban et al., 2000) (Nissen et al., 2000) (Kudva et al., 2018). Unlike the 30S, the distribution of the r-proteins on the 50S is nearly uniform, with most proteins located on the solvent side, leaving the interface accessible for the inter-subunit bridge formation (Ban et al., 2000).

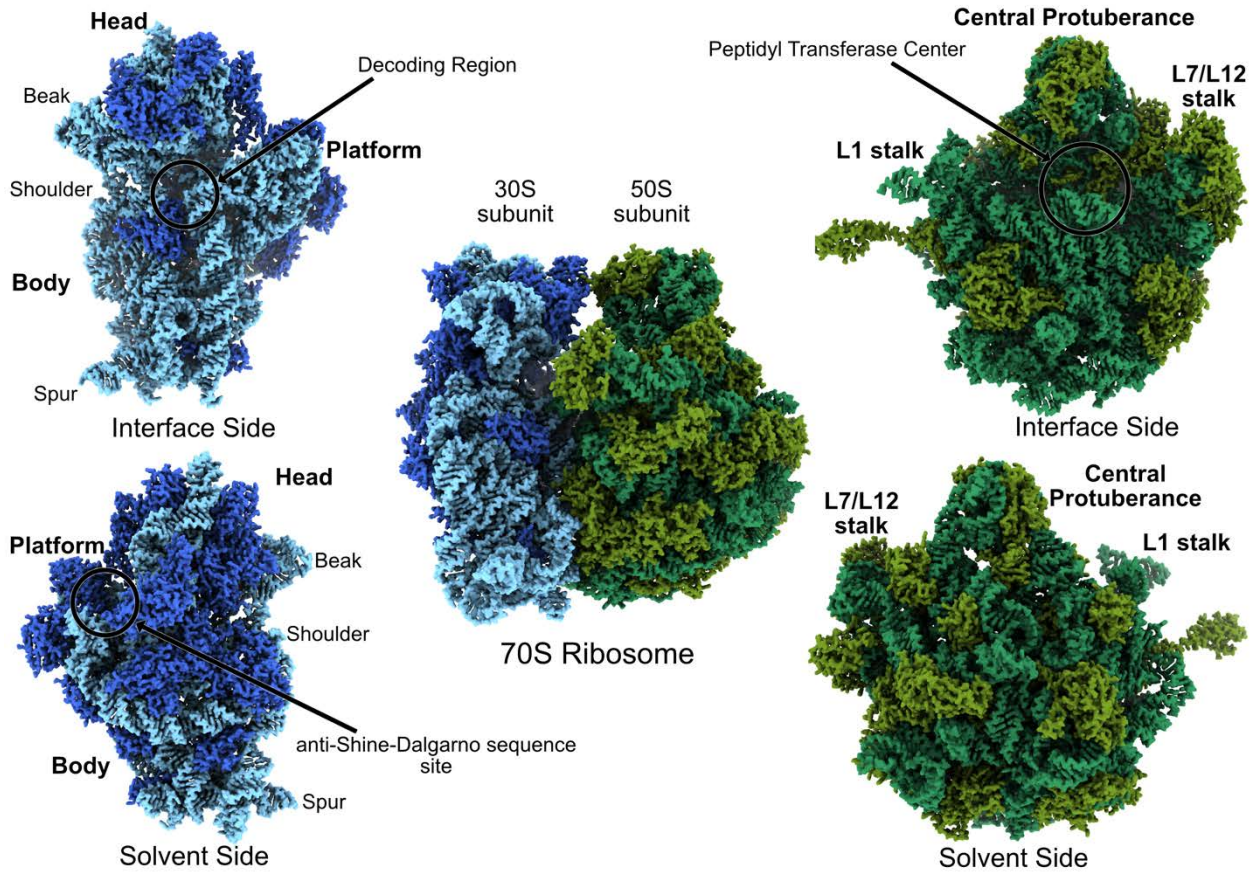


Figure 1. 1 Overview of the Escherichia coli ribosome and its subunits.

The 70S ribosome is positioned in the center of this figure. The 30S (small subunit) is colored in shades of blue (16S rRNA = light sky blue; ribosome proteins = royal blue), while the 50S (large subunit) is colored in shades of green (23S rRNA = dark sea green; ribosome proteins = olive drab). The left side of the figure displays the interface (top) and solvent (bottom) views of the 30S. The decoding region is encircled in bold font in the interface view and landmark domains Body, Platform, Head. The site for the anti-Shine-Dalgarno sequence is encircled on the solvent side. The right side of the figure displays the interface (top) and solvent (bottom) views of the 50S. The peptidyl transferase center is encircled on the interface side. Landmark domains L1 Stalk, Central Protuberance, L7/L12 Stalk are labeled in bold font. Volumes from the PDB 4V4Q were created using the script e2pdb2mrc.py from EMAN2.

1.2. Role of magnesium for ribosome structural integrity and function

Metal ions play an essential role as cofactors in various biological processes and are critical for the stability and functioning of proteins and nucleic acids (Sissi and Palumbo, 2009) (McCall et al., 2000) (Gesteland, 1966). Accordingly, magnesium (Mg^{2+}) is a multivalent cation abundant in all living cells (Wacker, 1969) and plays an essential role in replication, transcription, and translation. In the context of the ribosomes, each *E. coli* 70S contains about 170 Mg^{2+} ions (Schuwirth et al., 2005), which highlights the essentiality of Mg^{2+} for structural stability. In an *in vitro* Mg^{2+} deficient environment (~1mM Mg^{2+}), the 70S ribosomes tend to dissociate into individual subunits (Gesteland, 1966). On the contrary, in an *in vitro* Mg^{2+} abundant environment (10mM Mg^{2+}), the 70S ribosomes remain intact. Similarly, in an *in vivo* Mg^{2+} deficient environment, *E. coli* cells tend to lose ribosomes due to degradation, thus affecting protein synthesis and growth (Mccarthy, 1962). This study also showed that growth in *E. coli* cells could be revived by supplementing growth conditions with Mg^{2+} abundance. Such a dramatic effect on ribosomes due to the loss or gain of Mg^{2+} is due to its small ionic radius (0.6 Å), high charge density, and ability to coordinate octahedral geometry (Petrov et al., 2012) (Nierhaus, 2014) (Rozov et al., 2019). These properties allow Mg^{2+} to efficiently polarize water molecules and coordinate the interaction of r-proteins and rRNA.

The abundance and scarcity of Mg^{2+} also affect activities and conformational changes of each subunit. In the case of the 30S, levels of Mg^{2+} can reversibly dictate the interconversion of 30S into ‘active’ and ‘inactive’ states due to conformational changes in the decoding region (Zamir et al., 1971) (Moazed et al., 1986). During translation initiation, Mg^{2+} ions stabilize mRNA-tRNA interaction at the A-site and assist in positioning the codons at A and P sites, thus preventing slippage (Konevega et al., 2004) (Selmer et al., 2006). In the case of the 50S, the PTC is exclusively

dependent on the abundance of Mg^{2+} ions for structural integrity and its ability to catalyze the peptide bond synthesis (Miskin et al., 1970) (Klein et al., 2004). Such high dependence on Mg^{2+} ions is mainly due to the absence of ribosome proteins at the functional core of the 50S ribosome subunit. In ribosome biogenesis, Mg^{2+} plays an essential role in neutralizing negative charges on the phosphate backbone in nascent rRNA synthesized during the ribosome biogenesis (Pontes et al., 2015). High levels of Mg^{2+} suppress defects related to the formation of 70S ribosomes that arise due to the absence of r-proteins uL1, uL23, bL36, bL34, and bS6 (Akanuma et al., 2014) (Akanuma et al., 2018).

Several other metal ions also play a role in the structural integrity and function of the ribosome. In *vitro*, the most efficient protein synthesis has been observed in the combination of Mg^{2+}, K^{+}, NH_4^{+}, and polyamines like spermidine (Rozov et al., 2019), (Belinite et al., 2021). While Mg^{2+} is the most abundant divalent cation, the potassium ions (K^{+}) are predominant monovalent cations inside the cells (Nierhaus, 2014). A recent study from Rozov et al. demonstrated that K^{+} is distributed throughout the 70S ribosomes and coordinates several rRNA interactions in the 30S decoding region and the 50S PTC. K^{+} has also been found to facilitate mRNA-tRNA interactions within loops of several r-proteins. Therefore, K+ plays a vital role in the overall functionality of ribosomes.

In the context of unexplored aspects of ribosome biogenesis, this section will mainly focus on assembling the 30S ribosome subunit.

1.3 Bacterial ribosome biogenesis

Bacterial ribosome biogenesis is a complex process involving a series of simultaneous events (1) transcription, processing, folding, and modification of rRNA; (2) translation, modification, and folding and simultaneous binding of approximately ~54 r-proteins **(Figure 1.2 A)** (Shajani et al.,

2011) (Davis and Williamson, 2017). This process is so rapid and efficient that the ribosome subunits assemble in ~2 minutes, with each bacterium synthesizing up to 100,000 ribosomes/hour (Davis et al., 2016) (Chen et al., 2012). As a result, such a well-orchestrated process consumes up to 40% of energy in a rapidly growing *E. coli* (Maguire, 2009).

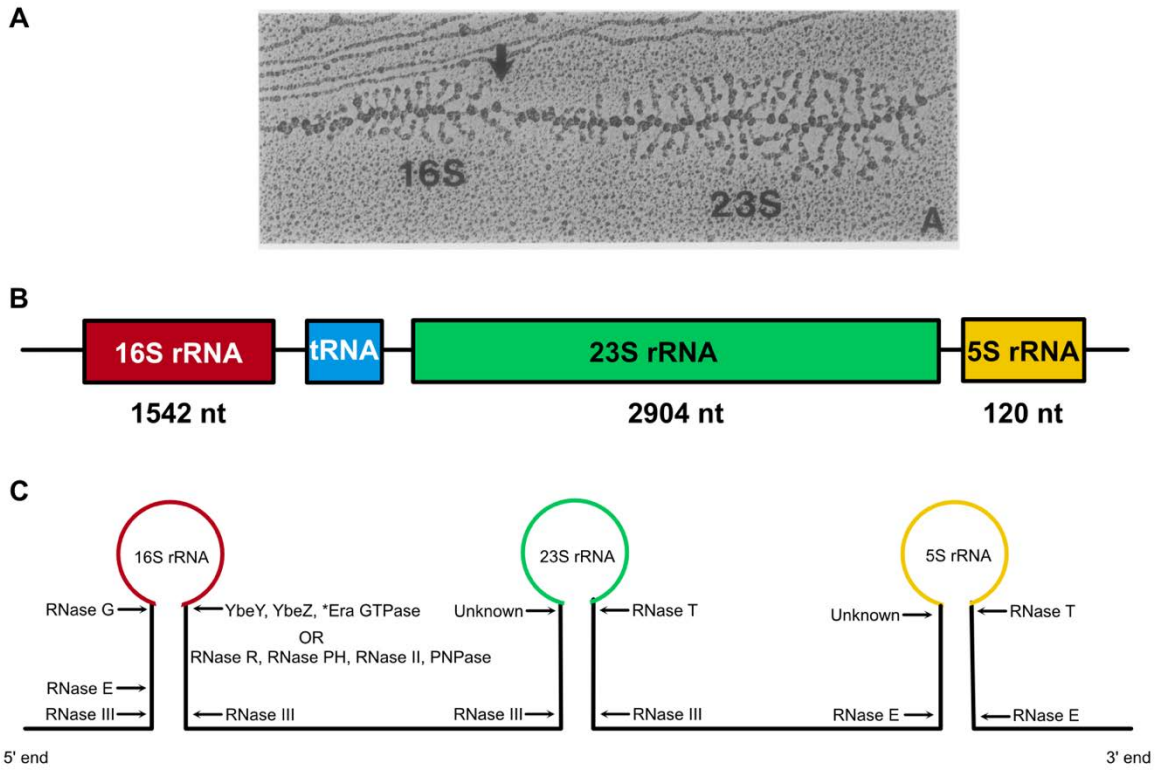


Figure 1. 2 Co-transcriptional ribosome assembly and rRNA processing.

(A) Electron micrograph showing Christmas tree morphology of the co-transcriptional ribosome assembly in E. coli. The arrow indicates the RNase III cleavage site that separates the 16S rRNA and 23S rRNA. This figure was adapted from (Gotta et al., 1991) with permission from the Journal of Bacteriology. (B) Cartoon depicting a typical E. coli rRNA operon with genes encoding the 16S rRNA, tRNA, 23S rRNA, and 5S rRNA. (C) Cartoon depicting primary rRNA transcript marked with labeled arrows for endo- and exo-ribonucleases cleavage site.

1.3.1 rRNA processing

In vivo, bacterial ribosome biogenesis begins with a single rRNA operon transcription. Each of the seven rRNA operons (*rrnA, rrnB, rrnC, rrnD, rrnE, rrnG, and rrnH*) in *E. coli* contains genes for 16S, 23S, and 5S rRNA and one to two genes encoding tRNA **(Figure 1.2 B)** (Hillebrand et al.,

2005) (Ginsburg and Steitz, 1975). Transcription by RNA Polymerases proceeds in the 5' to 3' direction coupled with rRNA processing by several endoribonucleases and exoribonucleases **(Figure 1.2 C)**. The primary transcript is initially cleaved by the double-strand-specific endonuclease – RNase III and results in the release of the precursor-16S rRNA (17S rRNA), precursor-23S rRNA, and precursor-5S rRNA (9S rRNA) (Redko et al., 2008) (Christiansen, 1988).

The 17S rRNA contains +115 and +33 nucleotides on the 5′ and 3′ ends, respectively, compared to the mature 16S rRNA. The 49 nucleotides on the 5' end are first cleaved by RNase E. In comparison, the remaining 66 nucleotides are cleaved by RNase G. Processing of the additional 33 nucleotides by the 3' end by RNases remains under-characterized because of (1) well-characterized cleavage activity by exonucleases RNase R, RNase II, PNPase, and RNase PH, and (2) poor characterization of YbeY endonuclease (Sulthana and Deutscher, 2013) (Davies et al., 2010) (Jacob et al., 2013). Biochemical studies have suggested that the complex of YbeY, r-protein uS11, ribosome assembly factor GTPase Era, and YbeZ, an RNA helicase, may provide an alternative cleavage pathway (Vercruysse et al., 2016) (Andrews and Patrick, 2022).

The additional 3 to 7 nucleotides on the precursor 23S rRNA, how the 5' end is removed is still unknown, while the extra 7 to 9 nucleotides on the 3'end are removed by exonuclease RNase T (Li et al., 1999). In the case of the 9S rRNA, RNase E cleaves the overhang 84 and 42 nucleotides at the 5' and 3' ends. The final stage of maturation on the 3' end is led by RNase T; however, further cleavage on the 5' end for getting the final mature 5S rRNA remains uncharacterized (Misra and Apirion, 1979).

1.3.2 16S rRNA folding and r-protein binding

Classical studies by the Nomura group in the late 1960s provided a framework for how 30S ribosome subunits can be assembled *in vitro* (Traub and Nomura, 1969) (Nomura and Traub, 1968) (Nomura et al., 1969). These *in vitro* reconstitution experiments determined the hierarchy and cooperativity of r-protein binding to the 16S rRNA to assemble a functional 30S **(Figure 1.3A)**. Accordingly, the Nomura Assembly Map categorizes 30S r-proteins into three categories: primary r-proteins (uS4, uS7, uS8, uS15, uS17, and bS20) that bind the directly to the rRNA, followed by secondary r-proteins (uS5, bS6, uS9, uS11, uS12, uS13, bS16, bS18, and uS19) that depend on at least one primary r-protein bound to 16S rRNA, and finally tertiary r-proteins (uS2, uS3, uS10, uS14, and bS21) that require pre-binding of at least one primary and secondary r-protein. r-protein bS1 is excluded from the Nomura Assembly Map due to its weak and transient association with the 30S ribosome subunit (Wilson and Nierhaus, 2005). Since rRNA transcription proceeds in the 5' to 3' direction, the primary r-proteins interact with the 5' domain (body) of the rRNA, secondary r-proteins with the central domain (platform), and tertiary r-proteins with the 3'domain (head) (Sykes and Williamson, 2009).

Recent studies using techniques like the time-resolved hydroxyl radical footprinting (Adilakshmi et al., 2008), pulse-chase quantitative mass spectrometry (PC/QMS) (Talkington et al., 2005) (Bunner et al., 2010), and fluorescence resonance energy transfer (Duss et al., 2019) (Rodgers et al., 2019) have provided additional kinetic information regarding how the binding of r-protein promotes the folding of rRNA, that was missing from the Nomura assembly maps. These kinetic studies also show that ribosome assembly can proceed to maturation via multiple parallel pathways. Accordingly, the work from the Woodson lab demonstrated that major domains of the 16S rRNA (i.e., 5' body domain, central platform domain, and 3' head domain) could start folding

independently and simultaneously, resulting in a diverse population of assembling 30S particles. Similarly, the PC/QMS study from the Williamson lab demonstrated that r-proteins do not have an absolute dependence on each other and may bind in diverse orders. Multiple parallel pathways thus ensure flexibility, making ribosome assembly robust and efficient (Thurlow et al., 2016).

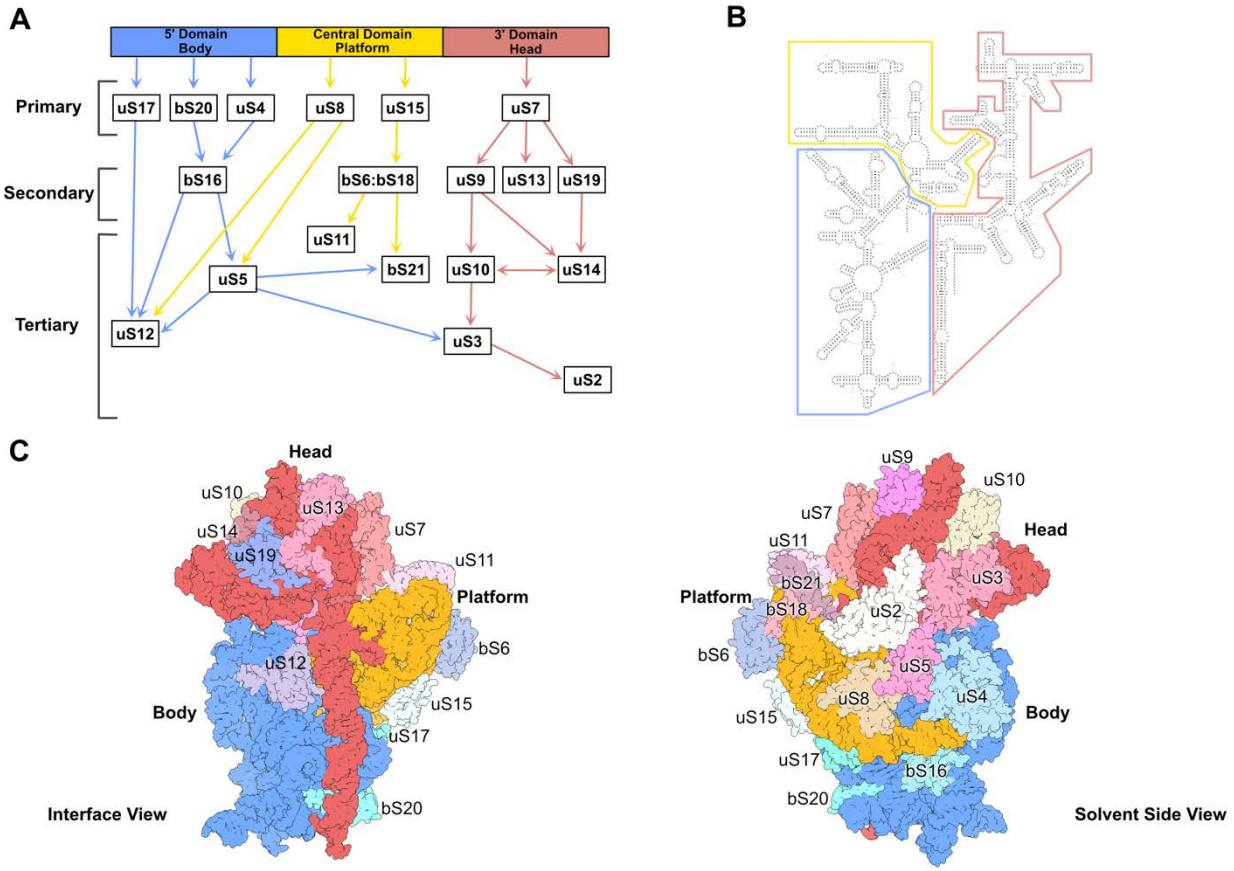


Figure 1. 3 Nomura assembly for 30S assembly.

(A) The arrows in this figure show thermodynamic dependencies of 30S ribosome subunit r-proteins binding to the 16S rRNA during in vitro ribosome assembly. The rRNA depicted as the segmented bar is colored according to the major domains of the 30S ribosome subunit: 5' body domain (blue), central domain (protein), and 3' Head domain. The top row represents 'primary' r-proteins that can directly bind to the 16S rRNA. The middle row represents 'secondary' r-proteins that depend on primary r-proteins to bind to the 16S rRNA. The bottom row represents 'tertiary' r-proteins that rely on primary and secondary r-proteins to bind to the 16S rRNA. This figure was adapted from (Shajani et al., 2011) with permission from the Annual Review of Biochemistry. (B) Secondary structure of the 16S rRNA depicting major domains color-coded same as panels A. Secondary structure (#URS00000ABFE9_562) represents the 30S from PDB ID: 3J9Z; this was downloaded from RNA Central Database. (C) The cartoon of the 30S molecular

model (PDB: 4V4Q) here shows the 16S rRNA colored according to the significant domains depicted in panel A; r-proteins are labeled next to their binding sites and follow the nomenclature from (Ban et al., 2014).

1.3.3 Assembly factors catalyze and orchestrate ribosome assembly

In vitro reconstitution experiments by Nomura and colleagues have shown that the 30S can be assembled with purified rRNA and r-proteins. However, these reaction conditions are far from physiological conditions with high salts and high temperatures (Nomura et al., 1984). An additional problem with the ribosome assembly *in vitro* is that rRNA tends to start folding autonomously and may misfold as a part of this process, thus getting kinetically trapped (Connolly and Culver, 2009). As a result, ribosome assembly efficiency *in vitro* is less efficient than *in vivo* (Chen and Williamson, 2013). This difference has been attributed to the existence of specialized proteins known as ribosome assembly factors (RAF) (Wilson and Nierhaus, 2007) (Shajani et al., 2011) (Woodson, 2011). These are classified as rRNA processing and modification enzymes (RNases, methylase, acetylases, and pseudouridinylase); enzymes that unwind and fold rRNAs (RNA helicases); and maturation factors that play a role during the final steps of ribosome subunit assembly (may or may not be GTPases). In the case of assembling 30S ribosomes, the final stages are focused on the appropriate maturation decoding region where mRNA is decoded during translation. Bacteria have specific assembly factors that ensure the correct folding of 16S rRNA helices 44 and 45, leading to functional subunits. Accordingly, these factors are RbfA, RimP, KsgA (RsmA), and GTPase YjeQ (RsgA).

<u>Ribosome assembly factor – GTPase YjeQ:</u>

GTPases are universally conserved proteins that play a diverse role in signal transduction, translation, cytoskeleton formation, intracellular transport, and biogenesis of ribosomes in both prokaryotes and eukaryotes (Karbstein, 2007) (Britton, 2009). Extensive biophysical and

biochemical studies over the last few years have provided detailed insights into the role of GTPases in the ribosome biogenesis (Karbstein, 2007) (Maiti et al., 2021). Accordingly, these proteins can (1) act as rRNA chaperons to promote conformational changes in assembling ribosomes to facilitate rRNA folding; (2) play a role in preventing premature binding of r-proteins thus preventing ribosome assembly from falling into kinetic traps; (3) facilitate recruitment of r-proteins or assembly factors; and (4) prevent premature association of ribosome subunits helping to avoid entry of faulty ribosomes into the translation cycle. One such biochemically well-characterized GTPase that plays a role in the assembly and maturation of the 30S ribosome subunit is YjeQ.

YjeQ (RsgA–Ribosomal small subunit GTPase A) is a 39 kDa protein broadly conserved in all bacterial species. Homologs of YjeQ have not been discovered in archaea and exceptionally have one homolog in eukaryote in the plant species *Arabidopsis thaliana* (Janowski et al., 2018). X-ray crystallography studies showed that the YjeQ N-terminal domain comprises an uS1-like OB (oligonucleotide/oligosaccharide binding)-fold domain, followed by a circularly permuted GTPase domain, and a C-terminal zinc-finger domain **(Figure 1.4A)** (Daigle et al., 2002) (Shin et al., 2004).

OB-fold Domain: Composed of five anti-parallel β-sheets that form a β-barrel and the presence of such a uS1 like OB-fold domain is a common feature of several RNA binding proteins such as translation factors. Interestingly, the YjeQ OB-fold domain has structural similarities with the translation initiation factor-1 (IF1) and binds to the 30S A-site like IF-1 (Carter et al., 2001). YjeQ in a GTP or GMP-PNP state can dissociate 70S ribosome into individual subunits and may perform this activity because of the YjeQ OB-fold domain's ability to sense an empty A-site on the 30S. Mutation in the OB-fold domain also affects the ability of YjeQ to hydrolyze GTP. Through its

β1/β2-hairpin loop in the OB-fold field, YjeQ probes the final turn on helix 44 and induces conformational changes in the decoding region observed during translation (Razi et al., 2017b) (Lopez-Alonso et al., 2017a). These observations have led to the hypothesis that YjeQ acts as a quality control protein to probe and check the assembly of the 30S before the 30S are released in a pool of actively translating ribosomes.

GTPase Domain: YjeQ belongs to the translation factor class (TRAFAC) of GTPases having five motifs G1-G5 (Shin et al., 2004). These motifs are not ordered sequentially as G1-G2-G3-G4-G5 but are present in the G4–G5–G1–G2–G3, thus making the GTPase domain circularly permuted. Within the TRAFAC GTPases, YjeQ belongs to the sub-family of Hydrophobic Amino acid Substituted GTPases (HAS-GTPase) since the glutamine, which is typically the catalytic residue for GTPases, is replaced by valine in YjeQ. Therefore, YjeQ has an alternative GTP hydrolysis mechanism where it triggers GTP hydrolysis through a histidine residue in the GTPase domain. Further, this catalytic histidine is in switch 1 rather than switch 2, which is the case for TRAFAC GTPases like translation elongation factor-Tu (Schmeing et al., 2009). YjeQ probes several critical 16S rRNA residues that are key for interacting mRNA and tRNA residues during translation through its GTPase domain. Intrinsically YjeQ has a low GTPase activity but stimulates >100 fold in the presence of the 30S ribosome subunit.

Zinc-Finger Domain: This C-terminal portion of YjeQ comprises a 3_{10}- helix and a loop containing three cysteines and a histidine residue coordinating a zinc ion required for the overall structural stability of the YjeQ. Additionally, two α-helices may probe the overall mature conformation of the 30S. YjeQ Zinc-finger domain is not essential for the binding of YjeQ but is required for efficient GTP hydrolysis. This domain is critical for releasing another ribosome assembly factor, RbfA, from the 30S ribosome subunit.

Although not essential, *yjeQ* deletion results in slow growth, and accumulation of immature 30S ribosome subunits ($30S_{\Delta yjeQ}$) with 17S rRNA, affecting virulence and overall fitness (Jomaa et al., 2011b) (Jeganathan et al., 2015) (Leong et al., 2013) (Campbell et al., 2005) (Himeno et al., 2004). Early cryo-EM analysis of the $30S_{\Delta yjeQ}$ particles revealed structural defects in the decoding region of the 30S **(Figure 1.4C)**. Quantitative mass spectrometry analysis revealed that only tertiary binding proteins uS2 and bS21 were severely depleted on the purified $30S_{\Delta yjeQ}$ particles and the lysate analyzed from the *yjeQ* deletion strain. Pulse-chase labeling experiments showed that the accumulated $30S_{\Delta yjeQ}$ particles are not dead-end products of the ribosome assembly process because these particles are eventually incorporated into the 70S ribosome (Thurlow et al., 2016).

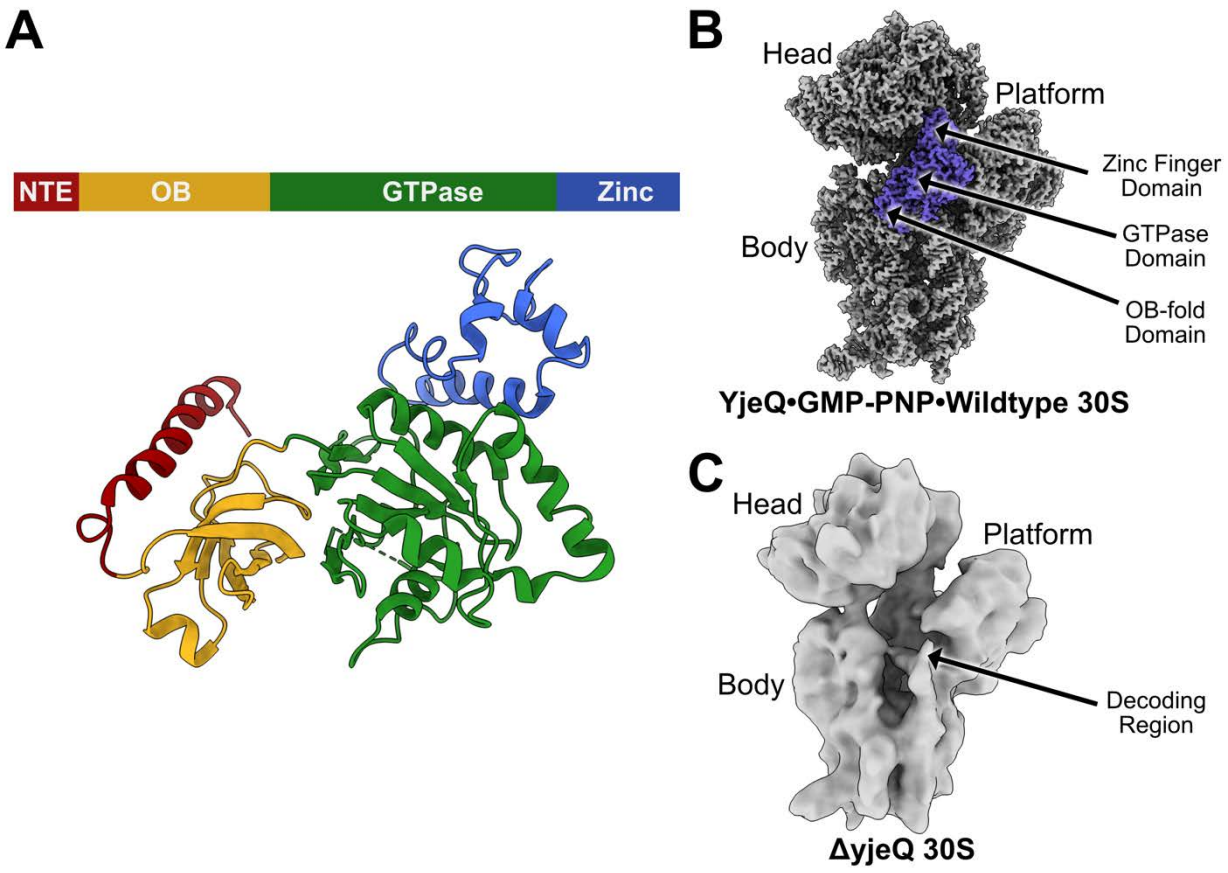


Figure 1. 4 Overview of the Escherichia coli ribosome.

(A) Cartoon depicting different domains of the ribosome assembly factor YjeQ GTPase (PDB: 5UZ4). (B) Volume showing the binding site of YjeQ bound to mature 30S ribosome subunit in the presence of GMP-

PNP; this volume was created using PDB 7NAR by using the script e2pdb2mrc.py from EMAN2. (C) 11.6 Å cryo-EM map (EMD-1774) of the immature 30S ribosomal subunits purified from ΔyjeQ E. coli strain.

1.4 Overview of canonical translation mechanism in bacteria

The translation cycle is divided into four significant steps: initiation, elongation, termination, and recycling **(Figure 1.5C)**.

1.4.1 Translation Initiation:

Successful translation initiation begins with the correct positioning of the mRNA start codon (generally AUG) in the P site (Julian et al., 2011). This is facilitated by the recognition of the Shine-Dalgarno sequence (SD) present in the 5' untranslated region (UTR) of mRNA by the anti-Shine-Dalgarno sequence (ASD) on the 3' end of the 16S rRNA (Shine and Dalgarno, 1974). Recruitment of an initiator N-formyl-methionine-tRNA (fMet-tRNA^{fMet}) in the presence of all three Initiation Factors (IF1, IF2, IF3) results in the formation of a 30S pre-initiation complex (30S PIC) (Hussain et al., 2016). Assembly of 30SIC results in structural rearrangements on the interface side of the 30S allowing association with the 50S subunit. This association triggers GTPase activity of IF2 and results in dissociation of all three initiation factors leaving the fMet-tRNA^{fMet} bound to the start codon of the P-site and formation of inter-subunit bridges between both subunits, thus creating a 70S initiation complex (70SIC) that are elongation competent (Milon and Rodnina, 2012).

1.4.2 Translation Elongation:

Three steps govern polypeptide synthesis during translation elongation: (A) decoding, (B) peptide bond formation, and (C) translocation **(Figure 1.5)**.

During the decoding step, Elongation Factor-Tu (EF-Tu) in a GTP bound-state carrying an aminoacyl-tRNA (aa-tRNA) binds to the 70SIC with an empty A-site. As the EF-Tu·GTP·aa-tRNA complex is delivered to the A-site, the tRNA anticodon attempts to form a Watson-Crick

base pairing with the codon on the mRNA (Schmeing et al., 2009) (Loveland et al., 2020). The 30S decoding region (helix 44) monitors the geometry of the codon-anticodon complex. In the case of correct Watson-Crick base-pairing, ribosome triggers conformational changes in the EF-Tu·GTP·aa-tRNA complex, which results in hydrolysis of GTP to GDP. This hydrolysis leads to the release of EF-Tu·GDP and leaves the tRNA in the A-site (Schuette et al., 2009).

(2) Peptide bond synthesis occurs in the peptidyl transferase center (PTC) of the 23S rRNA. Accordingly, nucleophilic attack by the α-amino group of the aa-tRNA in the A-site on the carbon of the ester group of the P-site tRNA creates a new peptide bond (Schmeing et al., 2005) (Polikanov et al., 2014). However, the reaction mechanism of this chemical reaction is not entirely understood since both studies propose different mechanisms.

(3) As a result of the peptide bond synthesis, the anti-stem loop (ASL) of the tRNAs that contain the anti-codon sequence remains in the A and P sites on the 30S ribosome subunit. The acceptor stems of these tRNAs now shift to the P and E sites, respectively, resulting in tRNAs in the A/P and P/E hybrid sites (Loveland et al., 2017). These movements are coupled with 30S head rotation and simultaneous contact by the L1 stalk with the tRNA in the P/E state. As a result, ribosome subunits rachet (move relatively to each other) until the L7/L12 stalk recruits Elongation Factor-G (EF-G) in a GTP-state. EF-G·GTP binding to the ribosome stabilizes the ratcheting motion. GTP hydrolysis leads to a conformational change in EF-G, causing it to adopt an elongated state and catalyze the movement of mRNA to shift by one codon (Carbone et al., 2021) (Li et al., 2015).

1.4.3 Translation Termination and Recycling

Translation elongation continues until the ribosome encounters one of the stop codons (UAA, UAG, or UGA) on the A-site (Oparina et al., 2005). Class I Release Factors (RF1 and RF2) recognize these stop codons and trigger conformational changes in the decoding center bases

A1492, A1493, and G530, and the nucleotides from the 23S rRNA involved in inter-subunit bridge formation between the two subunits (Fu et al., 2019). These movements within the ribosome promote conformational changes in the RFs. They now allow placement of the GGQ motif on the RF into the PTC, which triggers hydrolysis of the ester bond between the nascent polypeptide chain and P-site tRNA (Zavialov et al., 2002). This leads to the release of nascent polypeptide followed by the dissociation of RF1 and RF2 triggered by the GTP hydrolysis of RF3, leaving the ribosome with mRNA and deacylated tRNA in the P site (Korostelev et al., 2010) (Zaher and Green, 2011).

To recycle these ribosomes for another round of translation, bacteria recruit proteins known as the ribosome recycling factor (RRF) to split ribosome subunits with help from EF-G·GTP. GTP hydrolysis from the EF-G induces reverse rotation of the 30S and catalyzes tRNA release and dissociation of both subunits (Klimova et al., 2019).

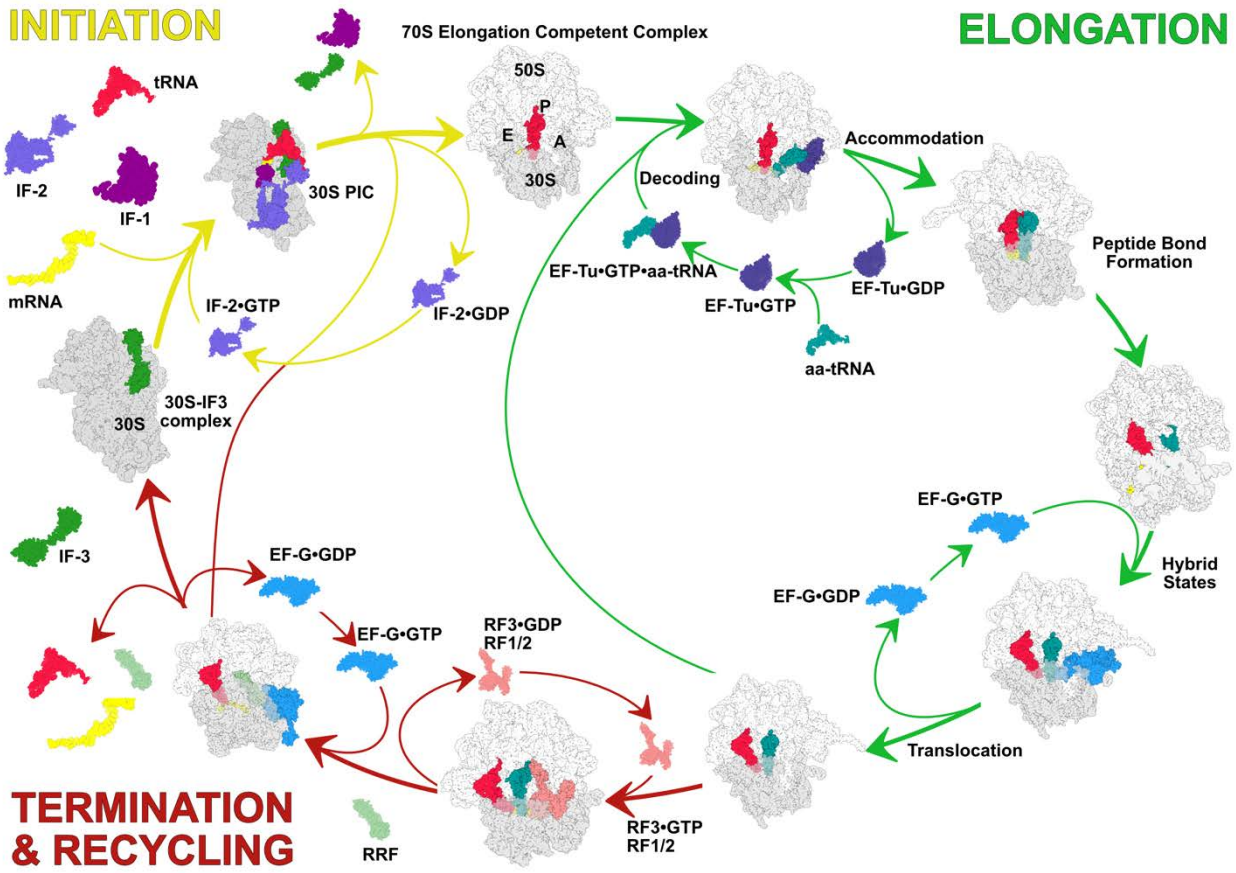


Figure 1. 5 Overview of canonical prokaryotic translation cycle.

The typical translation cycle is divided into Initiation, Elongation, Termination & Recycling. The initiation stages are colored in Yellow and start with Initiation Factor-3 (IF3) binds to the 30S ribosome. This facilitates the binding of messenger RNA (mRNA), Initiation Factor-1 (IF1), and Initiation Factor-2 (IF2) and an initiator tRNA for making the 30S Pre-Initiation Complex (30SPIC). Association of the 50S to the 30SPIC leads to IF-2 GTP hydrolysis mediated conformational change releasing the initiation factors and formation of a 70S elongation competent complex. The Elongation states are colored in Green. Elongation Factor-Tu (EF-Tu) in the GTP state is recruited to the 70S ribosome subunit to present the aminoacyl-transfer messenger RNA (tRNA) to the 30S A site, where it undergoes decoding and proofreading. Cognate codon-anticodon leads to the accommodation of the two tRNAs in the PTC for peptide bond synthesis. Successful peptide bond synthesis rotates the subunits in the hybrid state, a substrate for Elongation Factor-G (EF-G). GTP bound EF-G helps the tRNA and mRNA translocate on the next codon and moves the deacetylated tRNA into the hybrid binding site P/E. The elongation cycle continues until the ribosome encounters stop codons and mark the translation termination and recycling state colored in Red. Encountering the stop codons allows binding of the Release Factors-1 or 2 (RF1 or RF2) that releases the nascent polypeptide chain through the polypeptide exit tunnel in the 50S. The GTP hydrolysis activity of RF3 releases RF-1 or RF2. Finally, the Ribosome Recycling Factor (RRF) and EF-G recycle the ribosome subunits for making the components available to the translation initiation machinery. This figure was adapted from (Sohmen et al., 2009) with permission from Cell. Following molecular models with PDB ID were used to recreate the individual components of this figure in UCSF Chimera X: 5LMV, 5AFI, 4V6A, 7SSN, 3J9Z, 7N31, 4V7P, 6UCQ.

1.5 Built-in mechanisms of Ribosome Fidelity

Ribosomes decode genetic information with high accuracy. The frequency of error during the translation is one codon per 1,000 to 10,000 codons (Ogle and Ramakrishnan, 2005). However, accurate protein synthesis is not solely based on recognizing the codon on the mRNA by the correct anti-codon on the tRNA (cognate interaction). This is because the free energy difference (ΔG) between the cognate and the non-cognate base-pairing is insignificant (≤3 kcal/mol) (Thompson and Stone, 1977). Therefore, if only Watson-Crick base-pairing (a single-step mechanism) were the only criteria for accurate decoding, the translation error would be one codon in every few hundred codons. To achieve this high fidelity ribosomes have innate mechanisms contributing to high translational fidelity, a two-step reaction: (1) initial decoding and (2) proofreading (Mohler and Ibba, 2017).

1.5.1 Initial decoding

The first step involves L7/L12 stalk mediated recruitment of the EF-Tu·GTP·aa-tRNA to the empty A-site of the 30S ribosome subunit within the 70S ribosome (Wieden et al., 2001). At this stage, the aa-tRNA interacts with the ribosome without needing a codon-anticodon interaction (Kothe et al., 2004). While the aa-tRNA can bind to the ribosome A-site without the need for EF-Tu·GTP, this binding is slow. The EF-Tu·GTP mediated approach allows the rapid presentation of the aa-tRNA to the ribosome. Since the tRNAs are not lined up in the order of the mRNA codon sequence, this allows quick selection and rejection of the tRNAs, which happens in the following codon recognition step. Placement of the anti-codon near the mRNA codon at the A-site triggers a conformational change in the decoding region. Accordingly, A1492 (16S rRNA), which typically stacks with A1913 (23S rRNA) as a part of the inter-subunit bridge formation, now unstacks and flips out of the helix 44 (Kipper et al., 2009). Similarly, 16S rRNA residue A1493 which forms a

pseudo-Watson-Crick base pair with A1408 in the same helix, also flips out of the helix 44 (Gregory et al., 2005). The aa-tRNA at the A-site is then approached by 16S rRNA residue G530 (helix 18) in the shoulder region of the 30S, which then monitors the geometry of the codon-anticodon interaction as follows: Via hydrogen bonding, A1493 and G530 monitor the first position, while A1492 and G530 inspect the second position of the codon (Loveland et al., 2017). Since the third position does not need accurate base pairing, mismatch in that position is tolerated (Murphy and Ramakrishnan, 2004). In case of cognate interaction at the first two positions, the 30S shoulder region transitions into a closed state which locks the tRNA and places the EF-Tu·GTP·aa-tRNA near the sarcin-ricin-loop (SRL) in the 50S subunit Helix 95 (Diaconu et al., 2005).

1.5.2 Proofreading

Placement of the EF-Tu·GTP·aa-tRNA near the sarcin-ricin-loop (SRL) causes the catalytic Histidine residue in the EF-Tu GTP domain to approach A2662 (H95 nucleotide), which then triggers the GTP hydrolysis. However, in the case of near-cognate interaction, the decoding residues A1492, A1493 and G530 fail to establish a hydrogen bond network with the codon-anticodon helix. As a result, the 30S shoulder region fails to transition from an open to a closed state, thus allowing the near-cognate tRNA to be released (Korostelev, 2022). Recent studies have shown that exit of the whole EF-Tu·GTP·aa-tRNA complex is unnecessary as EF-Tu can hydrolyze GTP while remaining bound to the 30S shoulder region (Morse et al., 2020). According to this model, EF-Tu can undergo several GTP hydrolysis cycles until the tRNA capable of forming cognate interaction is captured and presented to the ribosome, which triggers domain closure and GTP hydrolysis mediated dissociation the EF-Tu·GDP.

1.6 Antibiotics mainly target the protein synthesis aspect of the bacterial ribosome

Bacterial ribosomes have been a primary target for antibiotics, with many of the existing ones mainly targeting the translation elongation step (Wilson, 2014) (Lin et al., 2018). Accordingly, antibiotics that target the 30S are localized around the decoding region and target diverse aspects like (1) Preventing binding of aa-tRNA in the A-site (Tetracyclines); (2) Stabilizing the aa-tRNA in the A-site, thus preventing translocation and induce miscoding (Negamycin, HygromycinB); (3) Induce flipping of universally conserved helix 44 nucleotides A1492 and A1493 in the decoding region thus promoting miscoding by accepting of non-cognate and near-cognate tRNA (Streptomycin, Aminoglycosides); (4) Interact with 16S rRNA on the P and E site to stall the mRNA thus preventing translocation of mRNA and tRNA and may also affect initiation mechanism (Kasugamycin, Pactamycin, Edeine, tetrapeptide antibiotic GE81112, and amicoumacin A). Several antibiotics specifically target the 30S ribosome subunit to interfere with the translation mechanism. Similarly, several antibiotics exist that interfere with the translation elongation mechanism by (1) Generally acting on the PTC on the 23S rRNA and preventing the peptide bond synthesis by clashing with the CCA motif on the 3'end of the aa-tRNA, thus preventing its accommodation (Chloramphenicol, Linezolid); (2) Some antibiotics target translation by blocking the polypeptide exit tunnel (Macrolide, Ketolides, etc.).

While we have a plethora of antibiotics that target the ribosome, antibiotic resistance has been on the rise and poses a severe threat to our healthcare system. Accordingly, resistance to aminoglycoside, tetracyclines, Macrolides, Lincosamides, Streptogramins B, and Ketolides have been reported. However, the availability of antibiotics that can effectively intervene in translation initiation is limited. Similarly, no clinically approved antibiotics specifically target bacterial ribosome biogenesis.

1.7 Cryo-EM, ribosomes, and air-water interface

High-resolution structures from X-ray crystallography have facilitated our understanding of ribosome structure, provided us with a glimpse of protein synthesis, and helped us understand the mechanism of action of several antibiotics. However, the choice of studying ribosomes using X-ray crystallography comes with severe limitations: (1) ribosomes are extremely dynamic molecules, therefore, obtaining crystals that can diffract at high resolution is difficult; (2) only certain states of the ribosome may crystallize, and the ones that do are packed within constraints of a crystal. Therefore, it is impossible to capture an ensemble of ribosome intermediates, limiting our understanding of ribosome function; (3) crystallization requires large concentrations of any given molecule; therefore, obtaining a large yield of ribosomes may be tedious.

Recent and rapid hardware and software advancements in single-particle cryo-electron microscopy (cryo-EM) have made determining the structures of biological molecules at near-atomic resolution routine. These advancements have allowed us to capture conformational and compositional states of ribosomes never visualized before in translation, ribosome assembly, ribosome hibernation, transcription-translation coupling, and several other biological processes. However, like X-ray crystallography, cryo-EM has several limitations: sample preparation specifically, cryo-EM grid preparation. To highlight why cryo-EM sample preparation is critical and needs to be flawless, let's consider studying ribosome biogenesis using cryo-EM. Ribosome biologists typically remove the component of interest and study its effect on ribosome assembly to understand how ribosome subunits are assembled. This is generally achieved by using genetic or chemical approaches. The genetic approach works by either deleting or depleting ribosome proteins or ribosome assembly factors (depending on their essentiality), which disrupts the normal assembly pathway, leading to the accumulation of immature ribosome subunits **(Figure 1.6)**.

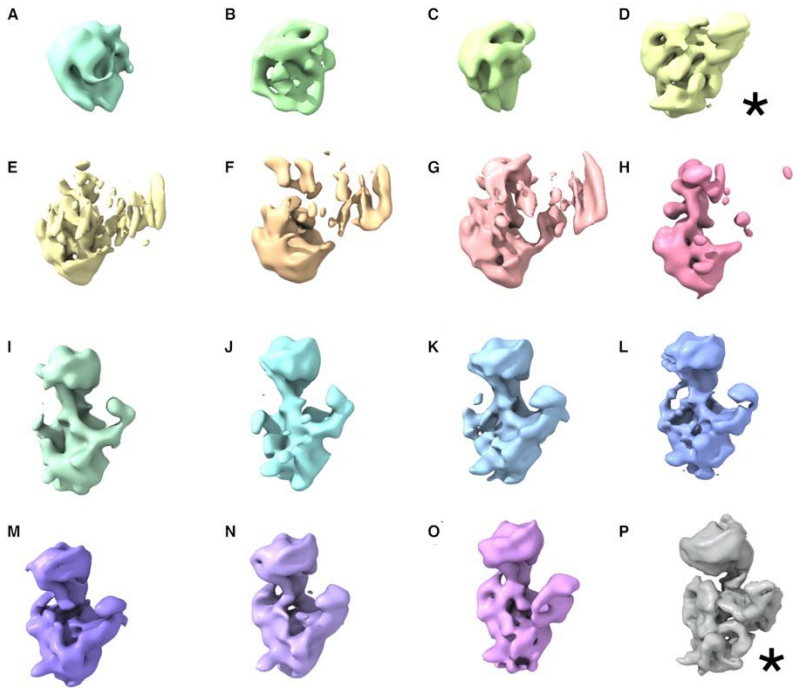


Figure 1. 6 30S assembly intermediates accumulated in the Era-depleted strain.

Ensemble an immature 30S ribosome subunit purified from E. coli after depleting the essential protein GTPase Era, a ribosome assembly factor that plays a role in maturing the platform region of the 30S ribosome subunit. Volumes in alphabetical order represent different stages of maturation, with the initial ones being extremely immature to the later ones resembling the overall 30S morphology. This figure was adapted from (Razi et al., 2019) with permission from Nucleic Acid Research.

Similarly, in the chemical approaches, small inhibitor molecules, if available, can be used to temporarily halt ribosome assembly and study its effects to understand ribosome assembly. These accumulated ribosome assembly intermediates can then be visualized and analyzed by cryo-EM and usually result in a collection of a heterogeneous pool of ribosome particles with a broad range of maturity (structural similarity or difference to a functional ribosome subunit). Since some ribosome assembly intermediates have incomplete protein complement or unfolded domains, it raises the question of whether the particles that do not resemble mature ribosomes are bona fide or artifacts of the technique. Therefore, this section aims to provide an overview of the sample vitrification process (ideally glass-like transparent ice film formation) used in cryo-EM, existing problems with the vitrification, and solutions developed by the field.

While there have been recent technological advancements in cryo-EM grid freezing, the essential aspect of sample vitrification in liquid cryogen near its melting point, typically using ethane (or a mixture of ethane and propane), remains the same as designed by Jacques Dubochet (Dubochet et al., 1988). A traditional and commonly used workflow is as follows: (1) Approximately ~3 μL of purified sample is applied to a cryo-EM grid treated with appropriate glow discharge for to reduce hydrophobicity on the surface of the grid; (2) Excess of the sample is then blotted by filter paper to produce buffer film of ideally 100 nm thickness; (3) The grid is immediately plunged into liquid ethane for vitrification of the thin buffer film on the cryo-EM grid and can be stored in liquid nitrogen until imaging. One of several limitations within this workflow is that once the excess sample is blotted, 90% of protein or macromolecules within the sample tend to migrate to the air-water interface (AWI) within milliseconds and can instantly denature (D'Imprima et al., 2019). This denaturation is triggered by the unfolding of motifs or domains with the macromolecules that contact the AWI leading to partial disruption of the quaternary structure and particle aggregation. An effective way to prevent this catastrophe is to coat the cryo-EM grid with a support layer which will cause the particles to adsorb on the support layer and prevent their interaction with the AWI. Of the several approaches that exist, using a carbon evaporator has been the most common way to coat cryo-EM grids with a thin continuous carbon layer (3 to 4 nm) (Russo and Passmore, 2014). While this method introduces noise in the cryo-EM images, it remains a go-to method for AWI problems. It also avoids preferential orientation issues, which can cause a poor 3D reconstruction of 2D cryo-EM images. Graphene as a support layer for cryo-EM grids has recently gained popularity (Sader et al., 2013). These grids may help avoid particle migration to the AWI, have minimum noise contribution in cryo-EM images, and can be functionalized to immobilize particles with different sizes and biochemical properties. Similarly, other options include detergents in the

sample buffer or affinity-tagged grids, but the efficacy of these options in the context of ribosomes remains to be tested.

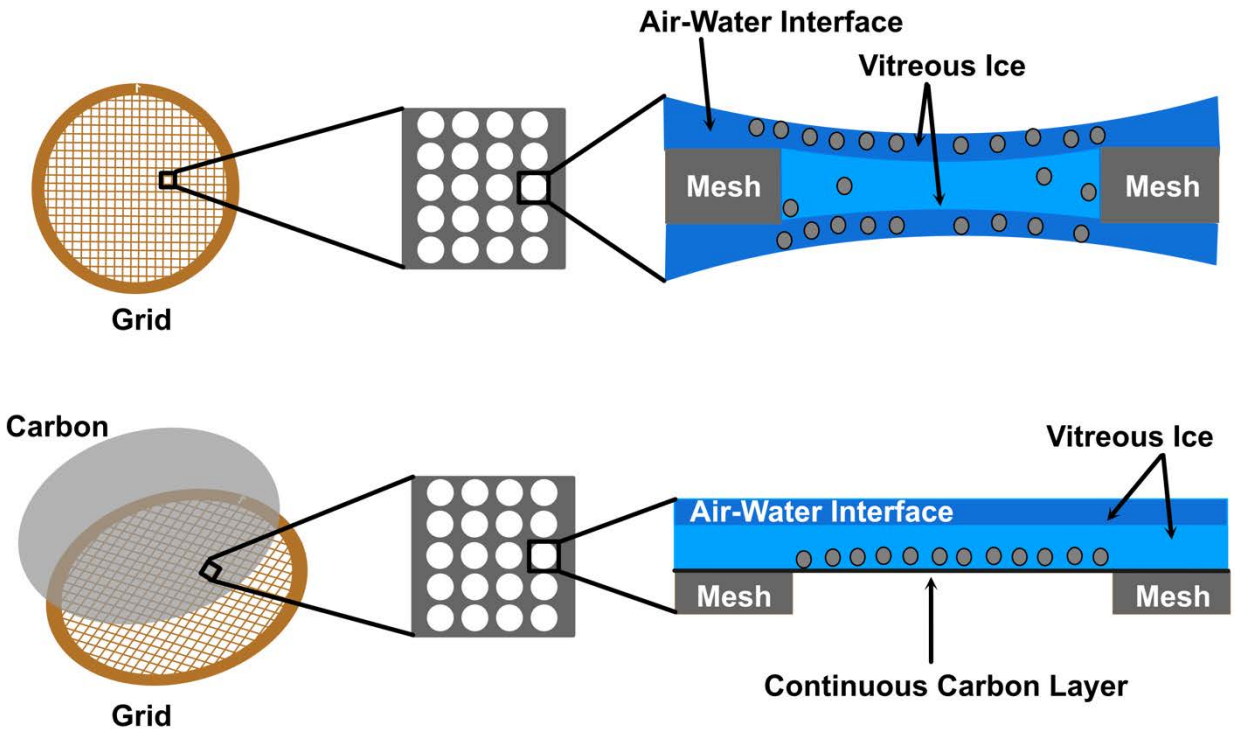


Figure 1. 7 Air-Water Interface in the context of cryo-EM.

The top panel shows the distribution of particles in a cryo-EM grid frozen without any support layer. Most particles migrate to the air-water interface, with very few remaining in the buffer. The bottom panel shows the effect of having a support layer, for example, a thin continuous carbon layer on the cryo-EM grid. In this case, Particles tend to adsorb on the support later, preventing migration to the air-water interface and helping avoid unfolding, denaturation, and aggregation.

1.8 Thesis Objective

This literature review highlights that decades of biophysical and biochemical studies have provided unprecedented details on the structure and function of the bacterial ribosome. However, there are several shortcomings in our understanding of ribosome biology:

- While X-ray crystallography provided us with high-resolution structures of the bacterial ribosome and its subunits, a significant limitation of these structures was that they limited our understanding of conformational dynamics. This is mainly because the ribosome is a highly dynamic molecule, and not all ribosome conformations can be crystallized. A great example in this context is the structure of the 30S ribosome subunit. Biochemical studies dating back more than five decades have shown that the 30S exists in active and inactive states. Accordingly, the crystal structure of *Thermus* Thermophilus 30S from Venki Ramakrishnan's group described the 'active' state. However, no structural information on the 'inactive' state was described.
- Since protein synthesis is an essential and central biological process, the bulk of ribosome biology studies are focused on gaining detailed insights into the translation process. In contrast, other aspects such as ribosome biogenesis remain inadequately explored. For example, ribosome assembly factor YjeQ has been structurally well characterized. However, ambiguity over its role related to the 30S ribosome subunit remains.
- Our current understanding of prokaryotic ribosome biology mainly comes from studies carried out using model organisms such as *E. coli* or *B. subtilis*. Therefore, we understand the translation initiation mechanisms only from a subset of prokaryotic ribosome systems. Contrastingly, bacteria belonging to the Bacteroidetes phylum seem to have alternative translation initiation mechanisms since Bacteroidetes genes lack the Shine-Dalgarno (SD) sequences even though their ribosome contains the anti-Shine-Dalgarno (ASD) sequence.

Typically, these two components are essential for successful translation initiation. However, Bacteroidetes ribosome fails to initiate translation even when an mRNA with SD sequence is presented to these ribosomes suggesting that Bacteroidetes ribosome are blind to SD sequence. Therefore, the overall objective of this thesis was to explore poorly understood aspects of ribosome biology that may constitute untapped targets for novel antibiotics. Accordingly, we seek to understand:

(1) How different is the inactive state of the 30S ribosome subunit compared with the active state?

(2) What role does YjeQ play in the context of mature and immature 30S ribosome subunits?

(3) How is SD-ASD pairing prevented in Bacteroidetes ribosomes?

CHAPTER 2: Alternative Conformations and Motions Adopted by 30S Ribosomal Subunits Visualized by Cryo-Electron Microscopy

Dushyant Jahagirdar, Vikash Jha, Kaustuv Basu, Josue Gomez-Blanco, Javier Vargas and Joaquin Ortega*

Department of Anatomy & Cell Biology, McGill University, Montreal, Quebec H3A 0C7, Canada. Centre for Structural Biology, McGill University, 3649 Promenade Sir William Osler, Montreal, Quebec H3G 0B1, Canada.

Published

RNA 2020 Dec;26(12):2017-2030

Contribution of Authors

D.J purified ribosomes, prepared samples for cryo-EM, and performed cryo-EM image analysis. D.J and J.O. collected cryo-EM data with strategies developed by K.B. J.G.B and J.V developed and performed isotropic map normalization method. D.J. and V.J built molecular models. J.O designed and supervised the study. J.O. and D.J wrote the manuscript, and all authors contributed to manuscript editing.

2.1 Abstract

It is only after recent advances in cryo-electron microscopy that it is now possible to describe at high-resolution structures of large macromolecules that do not crystalize. Purified 30S subunits interconvert between an “active” and “inactive” conformations. The active conformation was described by crystallography in the early 2000s, but the structure of the inactive form at high resolution remains unsolved. Here we used cryo-electron microscopy to obtain the structure of the inactive conformation of the 30S subunit to 3.6 Å resolution and study its motions. In the inactive conformation, alternative base pairing of three nucleotides causes the region of helix 44, forming the decoding center to adopt an unlatched conformation and the 3’ end of the 16S rRNA positions similarly to the mRNA during translation. Incubation of inactive 30S subunits at 42 ºC reverts these structural changes. The air-water interface to which ribosome subunits are exposed during sample preparation also peel off some ribosomal proteins. Extended exposures to low magnesium concentrations make the ribosomal particles more susceptible to the air-water interface causing the unfolding of large rRNA structural domains. Overall, this study provides new insights about the conformational space explored by the 30S ribosomal subunit when the ribosomal particles are free in solution.

2.2 Introduction

Ribosomes in bacteria undergo constant conformational changes that are essential for the translation process. These include from small-scale base flipping events at the decoding center to much large-scale motions of ribosomal subunit domains induced by mRNA, tRNA and translation factor binding (Frank, 2017). Similarly, the ribosome assembly process involves constant conformational changes, as the rRNA folds and ribosomal proteins are incorporated into the assembling particle (Sashital et al., 2014) (Razi et al., 2017a) (Mulder et al., 2010). Until recently, high-resolution structural information about these conformational states was solely contributed through X-ray crystallography. Consequently, only those states that could be stabilized in a crystal lattice were accessible providing only a reduced breath of the conformational heterogeneity existing in these processes and potentially masking important details about local and global conformational dynamics.

Today, structural biology is in the midst of a "resolution revolution" (Kuhlbrandt, 2014). Due to continuous advances, cryo-electron microscopy (cryo-EM) can now routinely contribute high-resolution models of large macromolecular machines with dynamic composition and conformations that have remained impervious to crystallization (Cheng, 2015, Cheng et al., 2017, Nogales and Scheres, 2015). In the context of the ribosome, these cryo-EM models are illuminating new relevant transition steps in the protein translation (Hussain et al., 2016) and ribosome assembly processes (Razi et al., 2019) (Ni et al., 2016) (Seffouh et al., 2019) (Nikolay et al., 2018). Mainly, in the study of the ribosome assembly process, structural biologists have widely used genetic approaches to trigger the accumulation of assembly intermediates (Razi et al., 2017a, Stokes and Brown, 2015). The essence of this approach consists of creating single deletion or depletion strains for one of the assembly factors to disable or slow down the ribosome biogenesis

process (Daigle and Brown, 2004). Invariably, these methods produce a heterogeneous mixture of immature ribosomal particles from which is not possible to grow crystals. However, structural characterization of these assembly intermediates using cryo-EM combined with image classification approaches has shown to be a powerful approach to identify the role of protein factors in assisting specific steps in the ribosome assembly process (Jomaa et al., 2011a) (Guo et al., 2013) (Leong et al., 2013) (Razi et al., 2019).

Inferring the role of assembly factors from the structural deficiencies observed in the assembling particles accumulating in the null or depleted cells requires the existence of a common standard or reference structure that is used in the comparative analysis. Previous publications have typically used the mature 30S subunit structure obtained by crystallographic approaches as the reference structure (Wimberly et al., 2000). However, the mature 30S subunit not constrained on a crystal lattice exists in more than one conformation. Five decades ago, Elson and colleagues (Zamir et al., 1969, Zamir et al., 1971) reported that purified 30S subunits readily interconvert between "active" and "inactive" conformations. Later Noller's group determined using chemical probing that transition between both states involves structural changes in the neck and decoding center regions of the 16S rRNA (Moazed et al., 1986). This conformational variability of the mature 30S subunit is not restricted to the bacterial ribosome. The eukaryotic 40S subunit also seems to sample multiple conformations (Swiatkowska et al., 2012). More recently, it was found using RNA SHAPE (selective 2-hydroxyl acylation analyzed by primer extension) that in exponentially growing *Escherichia coli* cells, 16S rRNA mainly adopts the inactive conformation in free mature 30S subunits and the active conformation in translating 70S ribosomes (McGinnis et al., 2015, McGinnis and Weeks, 2014). The reactivity patterns on the 30S subunit associated with the 50S subunits are entirely consistent with the RNA secondary structure exhibited by the 30S subunit in

the crystal structure, suggesting that both approaches describe the same structure. However, the high-resolution structure of the inactive conformation observed in the free 30S subunits has never been obtained by X-ray crystallography or cryo-EM.

To visualize the 3D structure of the inactive conformation of the 30S subunit, and potentially other alternative conformations the mature 30S subunits adopt outside the constraints of a crystal lattice, we exposed purified 30S subunits to buffer conditions that recall those in ribosome purification approaches (Jomaa et al., 2011a) (Daigle and Brown, 2004). Cryo-EM revealed that the mature 30S subunits in solution adopt a variety of conformations. Magnesium concentration in the purification buffers had a large effect. Exposure to low magnesium concentrations switched the decoding center of the 30S subunit to a drastically different conformation from that observed in the crystal structure. Incubation of the purified 30S subunits at 42 ºC induced the decoding center to switch back to the canonical conformation. Our experiments also showed how the air-water interface to which the 30S subunits are exposed during the vitrification process in cryo-EM could also induce structural variability by causing the loss of r-proteins. Extended exposure to low magnesium concentration increased the ribosomal particles' susceptibility to the air-water interface, causing complete unfolding of the head domain and partial unfolding of the platform domain. Studying the motion of these ribosomal particles revealed that the most predominant motions exhibited by the head domain in the free 30S subunits are similar those observed in the 70S ribosome during protein translation.

2.3 Materials and Methods

Cell strains

We used the *Escherichia coli* K-12 (BW25113) strain from the Keio collection (Baba et al., 2006) to produce the purified 30S subunits used in this study.

Purification of 30S ribosomal particles

To obtain purified 30S subunits, we first produced a 30 mL saturated overnight culture of *E. coli* K-12 strain (BW25113) in LB media and this culture was used to inoculate a 3 L culture of LB media. Cells were grown at 37 °C with shaking at 225 rpm in an Innova 44 incubator shaker (New Brunswick) until they reached an OD_{600} of 0.6. Harvesting of the cells was done by centrifugation at 3,700 *g* for 15 min and the obtained pellets were chilled to 4 °C and all the subsequent steps were done at this temperature. These pellets were resuspended in 7 mL of buffer A (20 mM Tris-HCl at pH 7.5, 10 mM magnesium acetate, 100 mM NH_4Cl, 0.5 mM EDTA, 3 mM 2-mercaptoethanol, and a protease inhibitor mixture (cOmplete Protease Inhibitor Mixture Tablets; Roche) and DNaseI (Roche). Cell lysis was done by passing the cell suspension through a French pressure cell at 1,400 kg/cm^2 three consecutive times, and the cell debris was cleared by spinning the lysate at 59,000*g* for 30 min. The supernatant was layered over a sucrose cushion of equal volume composed of 37.6% sucrose in buffer B (20 mM Tris-HCl pH 7.5, 10 mM magnesium acetate, 500 mM NH_4Cl, 0.5 mM EDTA, and 3 mM 2-mecaptoethanol), and then spun down for 4.5 hours at 321,000 *g*. The pellet containing the ribosomal particles was resuspended in buffer C containing 10 mM Tris-HCl pH 7.5, 10 mM magnesium acetate, 500 mM NH_4Cl, 0.5 mM EDTA and 3 mM 2-mecaptoethanol and then spun for 16 hours at 100,000 *g*. This produced a pellet containing washed ribosomal particles that was resuspended in buffer F (10 mM Tris-HCl, pH 7.5, 1.1 mM magnesium acetate, 60 mM NH_4Cl, 0.5 mM EDTA, and 2 mM 2-mercaptoethanol). To

separate the fraction containing the 30S subunits, approximately 120 A_{260} units of resuspended crude ribosomes were then applied to 34 mL of 10–30% (wt/vol) sucrose gradients prepared with buffer F. The gradient was centrifuged for 16 hours at 40,000 *g* on a Beckman Coulter SW32 Ti rotor and fractionated using a Brandel fractionator apparatus and an AKTA Prime FPLC system (GE Healthcare). Fractions containing the 30S subunits were selected and pooled together based on the UV absorbance at A_{254}. Subsequently, they were spin down for another 4.5 hours at 321,000*g* on a Beckman 70Ti rotor. The pellet was resuspended differently according to the final conditions we intended to study the 30S subunit. The 30S subunits under 'low Mg^{2+} conditions' (30SInactivated-low-Mg^{2+}) were resuspended back in buffer F, whereas the 30S subunits under 'high Mg^{2+} conditions' (30SInactivated-high-Mg^{2+} and 30SInactivated-Carbon-high-Mg^{2+}) and were resuspended in buffer E (10 mM Tris-HCl, pH 7.5, 10 mM magnesium acetate, 60 mM NH_4Cl and 3 mM 2-mecaptoethanol). 'Activated' 30S subunits (30SActivated-high-Mg^{2+}) were resuspended in modified buffer E with a different concentration of magnesium acetate (20 mM) and stored at -80 °C. To perform the activation, these 30S subunits were heated at 42 °C for 20 minutes and then diluted by mixing equal volumes of the sample with buffer E not containing any magnesium acetate.

Cryo-electron microscopy

Sample vitrification was performed using a Vitrobot (Thermo Fisher Scientific Inc.). For all samples, a volume of 3.6 µL of the diluted sample was applied to holey carbon grids (c-flat CF-2/2-2C-T) previously washed in chloroform for two hours and treated with glow discharged in air at 5 mA for 15 seconds right before the sample was applied. For the 30SInactivated-Carbon-high-Mg^{2+} and 30S-Inactivated- Carbon-low-Mg^{2+} samples, we used holey carbon grids containing an extra layer of continuous thin carbon (5-10nm). The concentration of ribosomal particles in the

solution applied to the grid varied between 200-300 nM depending on the sample. In the Vitrobot, each grid was blotted once for 3 seconds and with a blot force +1 before they were plunged into liquid ethane. The Vitrobot was set at 25 ºC and 100% relative humidity.

Most of the data acquisition was performed using EPU software at FEMR-McGill using a Titan Krios microscope at 300 kV equipped with a Falcon II direct electron detector (Thermo Fisher Scientific Inc.). Movies for the 30S-Inactivated-high-Mg^{2+}, 30S-Inactivated-Carbon-high-Mg^{2+}, 30S-Activated-high-Mg^{2+} and 30S-Inactivated-low-Mg^{2+} datasets were collected with a total dose of 50, 52, 52 and 50 $e^-/Å^2$, respectively. All datasets except for the 30S-Inactivated-Carbon-high-Mg^{2+} particles were collected as movies with seven frames acquired in 1 second exposure at a magnification of 75,000x, producing images with a calibrated pixel size of 1.073 Å. Movies for the 30S-Inactivated-Carbon-high-Mg^{2+} particle were collected in the same manner but with 30 frames. The nominal defocus range used to collect all these datasets was between -1.25 to -2.75 μm. The dataset for the 30S-Inactivated- Carbon-low-Mg^{2+} sample was also collected in the same Titan Krios microscope but using SerialEM software (Schorb et al., 2019) and a Gatan K3 direct electron detector equipped with a Quantum LS imaging filter. Images were collected as 36-frames movies using 3.6 second exposures in counting mode using a total dose per movie of 60 e^-/A^2 at a nominal magnification of 81,000x corresponding to a calibrated pixel size of 1.09 Å. Defocus ranged for this dataset was -1 to -2.25 μm.

Image processing

Collected movies were corrected for beam-induced motion using RELION's implementation of the MotionCor 2 algorithm (Zivanov et al., 2018) (Zheng et al., 2017). Correction for the Contrast Transfer Function (CTF) was done using the Gctf program (Zhang, 2016). Subsequent processing steps were done using RELION-3 (Zivanov et al., 2018). Particle images were selected and

extracted from the micrographs using auto-picking and subsequently subjected to one or more cycles of reference-free 2D classification to remove false positive and damaged particles. This process produced clean datasets for the 30S-Inactivated-high-Mg^{2+}, 30SInactivated-Carbon-high-Mg^{2+}, 30SActivated-high-Mg^{2+}, 30SInactivated-low-Mg^{2+} and 30S-Inactivated- Carbon-low-Mg^{2+} samples comprised of 565,255, 334,903, 407,623, 658,065 and 264,418 particles, respectively. These datasets were used for 3D classifications to separate the different conformational subpopulations existing in each sample. The initial 3D reference used for these classifications was either a 60 Å low pass filtered map of the mature 30S subunit created from 4V4Q.pdb (Schuwirth et al., 2005) using the Xmipp program (de la Rosa-Trevin et al., 2013) or the intermediate cryo-EM maps obtained during classification. No mask was used for the 3D classifications. All 2D classification and 3D classification jobs were performed using particle images binned by 4. In each dataset, maps obtained in the classification steps were visually inspected in Chimera (Pettersen et al., 2004) and particles assigned to maps representing the same conformation were pooled together and used for standard 3D auto-refinement. A soft-mask was applied in all 3D auto-refinements. These masks were created with 'relion_mask_create' command extending the binary mask by four pixels and creating a soft-edge with a width of four pixels. The initial threshold for binarization of the mask varied depending on the structure. As initial map for the refinement procedures we used either a 60 Å low pass filtered map of the mature 30S subunit created from 4V4Q.pdb (Schuwirth et al., 2005) using the Xmipp program (de la Rosa-Trevin et al., 2013) or the cryo-EM maps obtained during classification after they were re-scaled back to full-size. Suitable classes were subsequently subjected to Bayesian polishing to correct for per-particle beam-induced motions, followed by CTF refinement for per-particle defocus, per-micrograph astigmatism, beam tilt, optical aberrations and a final round of 3D auto-refinement

(Zivanov et al., 2019) (Zivanov et al., 2020). Sharpening of the final cryo-EM maps and the local analysis was done with RELION (Zivanov et al., 2018).

Isotropic map normalization

The cryo-EM maps for 30S-Inactivated-high-Mg^{2+} class A and 30S-Activated-high-Mg^{2+} were slightly affected by directional resolution anisotropy caused by the presence of preferential specimen orientations. In these cases, the underrepresented macromolecular views result on low map amplitudes along these directions causing a stretching in the reconstructed volume. In this work, we compensate the map stretching caused by preferential specimen orientations by an isotropic amplitude map normalization in Fourier space. In this approach, we normalize map amplitudes at a particular resolution along all possible map directions for resolutions higher than 9-10 Å. At these resolution ranges, respective Fourier components encode the information of secondary structures presented in the map. We hypothesized that these secondary structures should be oriented approximately random wise. Then, the amplitudes of map Fourier components should be approximately isotropic for resolutions higher than 9-10 Å. Our map restoration method follows these steps: (1) The Fourier transform of the cryo-EM map is obtained; (2) Starting from 9-10 Å, a shell of the map Fourier transform is extracted; (3) The map amplitudes within the extracted shell and the respective q=75% percentile value is obtained; (4) Map amplitudes at the extracted shell are modified so their value is set to q; (5) A new shell at higher resolution is extracted and steps (1) to (4) are repeated; (6) This process is iterated until the shell at highest resolution (Nyquist resolution) is processed; (7) The inverse Fourier transform of the isotropic amplitude normalized map is computed obtaining a new corrected map. This map restoration approach transforms map Fourier amplitudes only, without modifying map phases.

Multibody refinement and motion analysis

We used the RELION-3 implementation (Nakane et al., 2018) to perform the multibody refinement and motion analysis. The same approach was followed for all 30S subunit populations. The consensus cryo-EM map obtained for each class was divided into three bodies corresponding to the three major domains of the subunit: body, platform and head. The masks for the corresponding bodies were made using a 25 Å low-pass filtered version of the consensus map. We use available atomic models of the 30S subunit (PDB ID 4V4Q) to determine the boundaries between the bodies. By extending the binary map by 10 pixels and placing a soft edge with of 10 pixels, all three bodies overlapped with each other. The multibody refinement job was run with downsized particle images with a box size of 218 and a pixel size of 1.496 Å. The standard deviation of the Gaussian prior on the rotations was set to 10 degrees for all three bodies, and the standard deviations on the body translations were all set to 2 pixels. Based on the domain architecture of the 30S subunit, the head was set to rotate with respect to platform and body. Multibody refinement was started using an initial angular sampling rate of 1.8 degrees, an initial offset range of 3 pixels and an initial offset step of 0.25 pixels.

Map analysis and Atomic model building

Before start building the molecular models for each structure, the obtained cryo-EM maps from RELION, the connectivity of the densities of the cryo-EM maps was improved using automatic sharpening as implemented in the PHENIX suite (Adams et al., 2010). Model building of all maps started by fitting the 4V4Q.pdb (Schuwirth et al., 2005) X-ray structure of the 30S subunit into the obtained cryo-EM maps using the rigid-body docking tools in Chimera (Pettersen et al., 2004). The molecular models were then built through multiple rounds of manual model building in Coot

(Emsley and Cowtan, 2004, Emsley et al., 2010) and real space refinement using Phenix (Adams et al., 2010).

Panel for figures were prepared using Pymol program (The PyMOL Molecular Graphics System, Version 2.3.3, Schrodinger,LLC), UCSF Chimera and Chimera X. Figures were assembled using Photoshop (Adobe). Secondary structures of RNA were produced with DSSR (http://x3dna.org) (Lu and Olson, 2008) and Varna (http://varna.lri.fr) (Darty et al., 2009).

2.4 Results

High-resolution cryo-EM structure of the 30S subunit inactive conformation.

In actively growing E. coli cells, 2-5% of the existing 30S subunits remain dissociated without forming 70S ribosomes (Leong et al. 2013; Thurlow et al. 2016). Consequently, 30S subunit purification protocols typically include a step that exposes ribosomes temporarily to buffers containing 1-2 mM concentration of magnesium ions (low magnesium). This condition induces the dissociation of the ribosome into its two integrating subunits, the 30S and the 50S subunits. This is mainly triggered by a structural change in the 30S subunit termed 'inactivation', as this conformational change interferes with tRNA binding in its P site (Zamir et al. 1969; Moazed et al. 1986).

To obtain the cryo-EM structure of the inactive conformation of the 30S subunit at high resolution, we purified 30S subunits from actively growing E. coli cells using a protocol that exposed the ribosomes to 1.1 mM magnesium acetate. However, at the end of the purification, the buffer in the fractions containing purified 30S subunits was exchanged and the concentration of magnesium acetate raised to 10 mM. The obtained sample (30S-Inactivated-high-Mg2+) was then imaged by cryo-EM.

Using image classification approaches, we found that the purified 30S-Inactivated-high-Mg2+ subunits existed in two distinct conformations that we called class A and B **(Figure. 2.1A)**. The distribution of particles between these two classes was 79% and 24%, respectively. Using these particle populations, we calculated a 3.6 Å resolution cryo-EM map for class A and a 4.5 Å resolution map for class B **(Supplemental Fig. S2.1)**. The structures of the body (5' domain), platform (central domain) and head domains (3' major domain) in both structures were identical to the canonical structure of the 30S subunit obtained by X-ray crystallography (Wimberly et al.

2000) **(Figure. 2.1A)**. However, the decoding region located at the convergency point of all these three domains presented important differences in both maps **(Figure. 2.1B)**. In the canonical structure, helix 44 runs from the bottom of the body to the lower part of the head. The upper region of this helix near the platform domain is involved in the decoding process and is also a critical element in creating the subunit interface with the 50S subunit. In the map for class A, the upper domain of helix 44 adopts an alternative conformation and is not latched to the decoding center as described by the crystal structure. Instead, this entire section of the helix protrudes from the surface of the 30S subunit and distorts the interface with the 50S subunit. In class B, the entire helix 44 seemed to adopt a flexible conformation. The lower part of this helix showed a well-defined density, but the middle and upper region exhibited a highly fragmented density indicating this region is highly flexible.

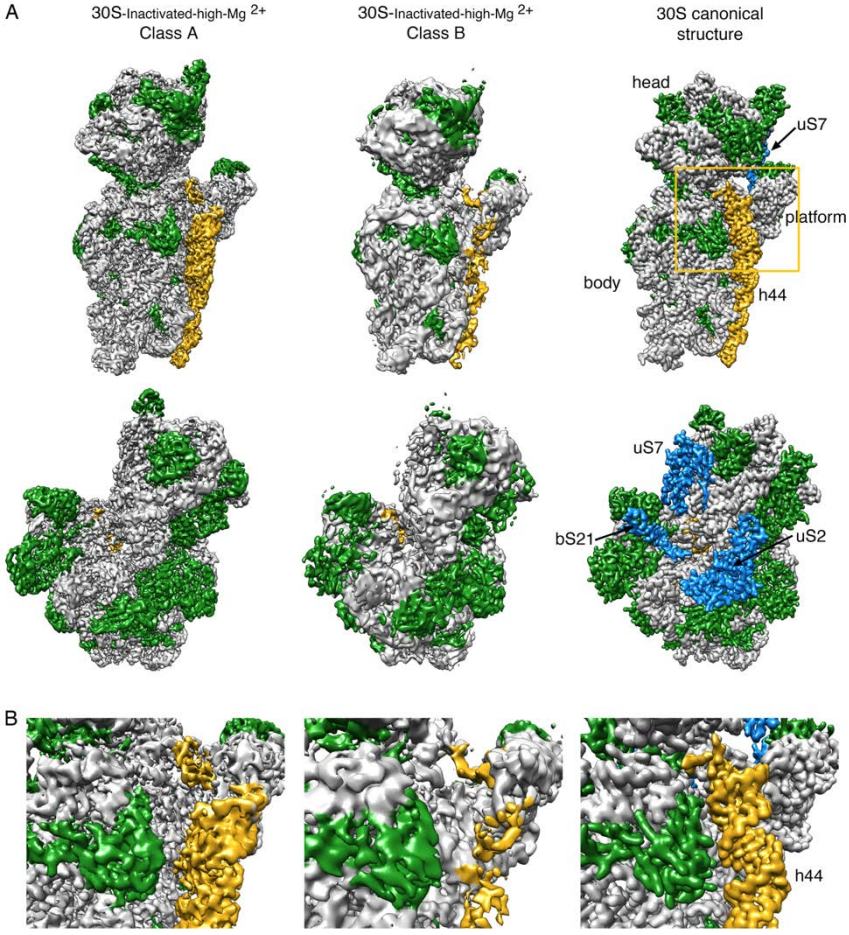


Figure 2. 1 Cryo-EM structure of the 30S-inactivated-high-Mg2+ particle.

(A) Front (top row) and back view (bottom row) of the cryo-EM maps obtained for the two subpopulations found for the 30S-inactivated-high-Mg2+ particle. The maps are shown side-by-side with the 30S subunit structure obtained by X-ray crystallography (30S canonical structure). This structure was obtained by generating a density map from PDB file 4V4Q and low-pass filtering this structure to 4 Å. The rRNA is displayed in light gray, the r-proteins in green, and helix 44 in goldenrod orange. The r-proteins uS2, uS7, and bS21 for which representative densities do not appear in the cryo-EM maps of the 30S-inactivated-high-Mg2+ particle are shown in blue in the map of the 30S

subunit obtained by X-ray crystallography. These proteins and other landmarks of the 30S subunit are labeled. (B) Zoomed-in view of the decoding region of the cryo-EM maps obtained for the two subpopulations found for the 30S-inactivated-high-Mg2+ particle and the structure of the 30S subunit obtained by X-ray crystallography. The area visualized in this panel is indicated as a frame in panel A.

To quantitatively measure the structural differences in the 30S-Inactivated-high-Mg2+ class A, we produced a molecular model from the cryo-EM map and subsequently calculated a temperature map to measure the deviation of this structure from the conformation of the 16S rRNA in the canonical structure **(Figure. 2.2A)**. The body and platform regions were highly similar and the rRNA in these structural motifs mostly overlapped with the canonical structure. In contrast, the position of the head and upper domain of helix 44 and contacting region in helix 28 diverged significantly. The head domain was tilted backwards by 16º, opening up the decoding region **(Figure. 2.2B).** The upper domain of helix 44 and bottom part of helix 28 adopted a drastically different conformation that was stabilized by a different arrangement in the base-paring of the rRNA **(Figure. 2.2C)**. The transition between both conformations involves nucleotides 1532-1534 at the 3' end of the 16S rRNA. These three nucleotides are not forming any base pairing in the conventional structure, however in class A they approach helix 28 and unfold its bottom part (region formed by nucleotides 1391-1396 and 921-925) and base pair with nucleotides 921-923 in that region. This transition also causes the partial unfolding of the top of helix 44 formed in the canonical structure by nucleotides 1397-1407 and 1494-1503, as well as the positioning of the 3' end of the 16S rRNA (distal to nucleotide 1534) in a conformation similar to that adopted by mRNA during translation. In this conformational transition helix 45 slightly shift in position but remains folded **(Supplemental Movie 1)**.

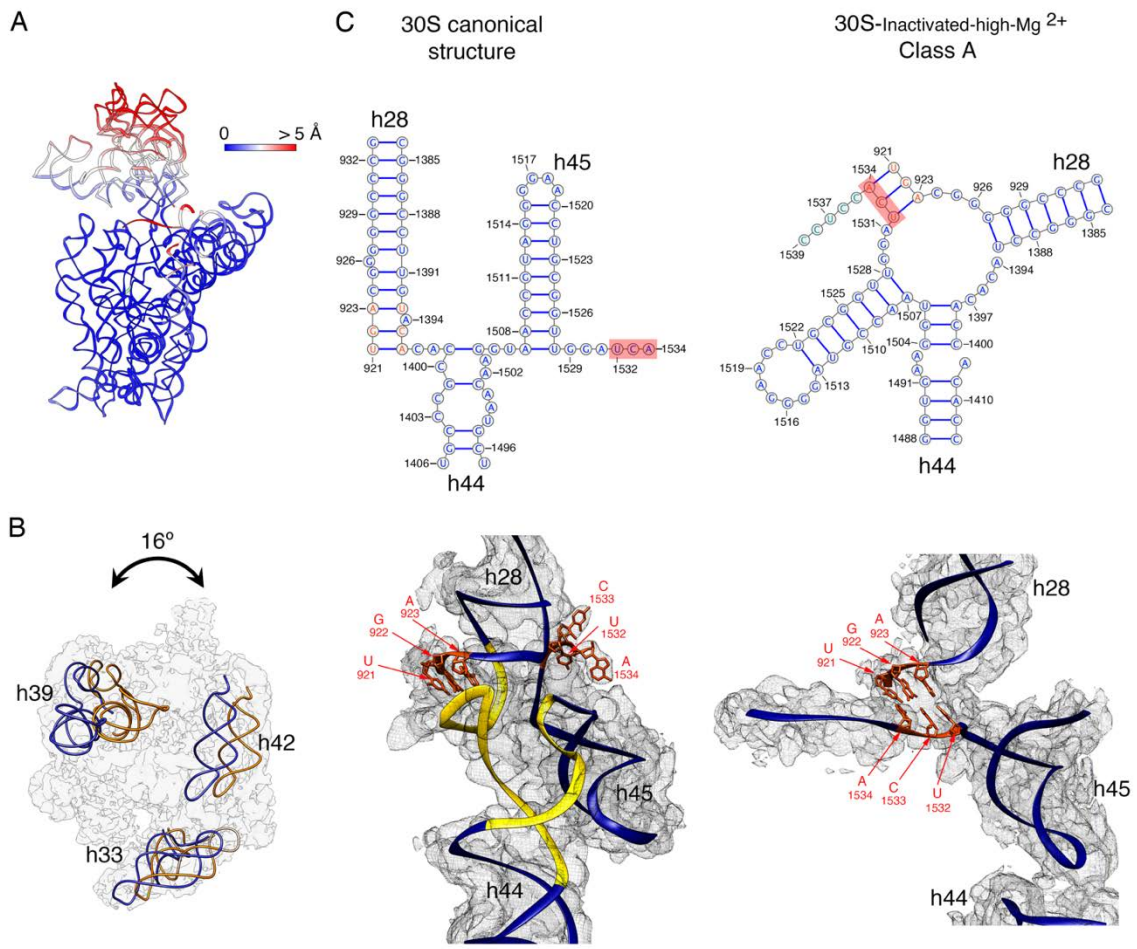


Figure 2. 2 Molecular model of the 30S-inactivated-high-Mg2+ particle.

(A) Temperature map of the 30S-inactivated-high-Mg2+ class A. The rRNA is colored according to the r.m.s.d. deviation (Å) with respect to the structure of the 30S subunit obtained by X-ray crystallography (PDB ID 4V4Q). (B) Top view of the head of the cryo-EM map of the 30S-inactivated-high-Mg2+ class A. The positions of helices 33, 39, and 42 in this structure (navy blue) and in the crystal structure of the 30S subunit (orange) are shown to illustrate the backward tilting of the head by 16° in the structure of the 30S-inactivated-high-Mg2+ class A. (C) Secondary (top panels) and tertiary (bottom panels) structures of helices 28, 44, and 45 of the 16S rRNA in the 30S subunit structure obtained by X-ray crystallography and in the molecular model derived from the cryo-EM map of the 30S-inactivated-high-Mg2+ class A. The nucleotides 1532–1534 at the 3' end of the 16S rRNA involved in the conformational transition between both structures are highlighted in red. The regions of helices 44 and 28 that become unfolded during the conformational transition are colored in yellow in the canonical 3D structure of the 30S subunit. The molecular model of the cryo-EM map of the 30S-inactivated-high-Mg2+ class A is shown overlapped with the density of the obtained cryo-EM map. In the case of the structure of the 30S subunit obtained by X-ray crystallography, a density map was generated from the atomic coordinates and overlapped with the molecular model.

Taken together, these structures indicate that 30S subunits purified through approaches that expose them to low magnesium concentration switch from the conventional structure shown by X-ray

crystallography to an inactive conformation that exhibit drastic structural differences in the decoding region. The inactive conformation is stabilized by an alternative base pairing of the nucleotides in the 3' end of the 16S rRNA molecule.

Exposure of the 30S subunits to the air-water interface in the cryo-EM grid causes the loss of r-protein uS2.

Previously reported cryo-EM structures of mature 30S subunits (Razi et al., 2019) (Datta et al., 2007) showed either fragmented or absent densities for r-proteins uS7 and bS21 suggesting that they dissociated or are intrinsically flexible when the particle is not constrained in a crystal lattice. Consistently, the densities corresponding to these two r-proteins in the cryo-EM maps obtained for the 30S-Inactivated-high-Mg^{2+} subunit class A and B were also not present **(Figure. 2.1A)**. Surprisingly, we also found the density representing uS2 to be completely absent from our cryo-EM maps. In the 30S subunit, uS2 binds in the solvent face (convex face) of the subunit stably anchoring its two domains to the 16S rRNA and typically is fully visible in previously obtained X-ray (Wimberly et al., 2000) (Schureck et al., 2016)and cryo-EM (Razi et al., 2017b) structures. In cryo-EM, to preserve the specimen in their hydrated state, one spreads the sample in a thin layer of buffer solution supported in the cryo-EM grid right before plunging the grid into liquid ethane to freeze the liquid layer into vitreous ice. In the vitrification device used in our experiments, the time that elapses between the blotting of the grid to form the thin layer and the vitrification is typically 1 second. In this time, ribosomal particles collide with the air-water interface between 100-1,000 times (Noble et al., 2018). These interactions have the potential to cause damages in the specimen.

To investigate whether the repeated interaction of the ribosomal with the air-water interface was causing the loss of uS2, we repeated the imaging of the 30S-Inactivated-high-Mg^{2+} particles by

adding an extra thin layer of continuous carbon to the grids **(Figure. 2.3A)**. We hypothesized that by adsorbing the particles to the support film, their exposure to the air-water interface would be reduced. We collected a cryo-EM dataset from these grids and particle images were subjected to a similar image classification workflow. We found that particles were present mainly as one class with helix 44 in an identical conformation to the structure of the 30S-Inactivated-high-Mg^{2+} class A particles. The cryo-EM map obtained from these particles (30S-Inactivated-Carbon-high-Mg^{2+}) refined to a resolution of 3.8 Å **(Figure. 2.3B and Supplemental Fig. S2.1)**. More importantly, the density representing uS2 was clearly visible in the cryo-EM map.

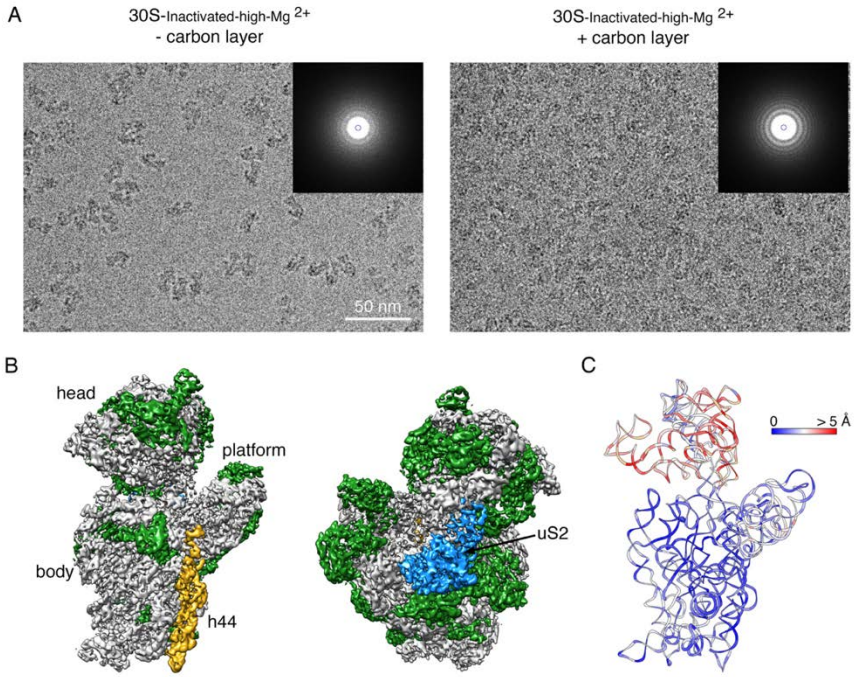


Figure 2. 3 Cryo-EM structure of the 30S-inactivated-high-Mg2+ particle in grids containing a continuous layer of carbon.

(A) Two representative electron micrographs containing 30S-inactivated-high-Mg2+ particles. The micrographs were obtained from EM grids without (left panel) and with (right panel) an extra layer of a continuous carbon. The inset shows the power spectra from each micrograph. The presence of a constant layer carbon makes the Thon rings in the power spectra more prominent and the background of the micrograph more prominent. (B) Front and back view of the cryo-EM map obtained for the 30S-inactivated-high-Mg2+ particle from grids containing a continuous carbon layer. The rRNA is shown in light gray, and the r-proteins are shown in green except uS2 that is colored in blue. (C) Temperature map of the 30S-inactivated-high-Mg2+ molecular model obtained from grids containing a continuous carbon layer. The rRNA is colored according to the r.m.s.d. deviation (Å) with respect to the structure 30S-inactivated-high-Mg2+ class A obtained from grids without a continuous carbon layer.

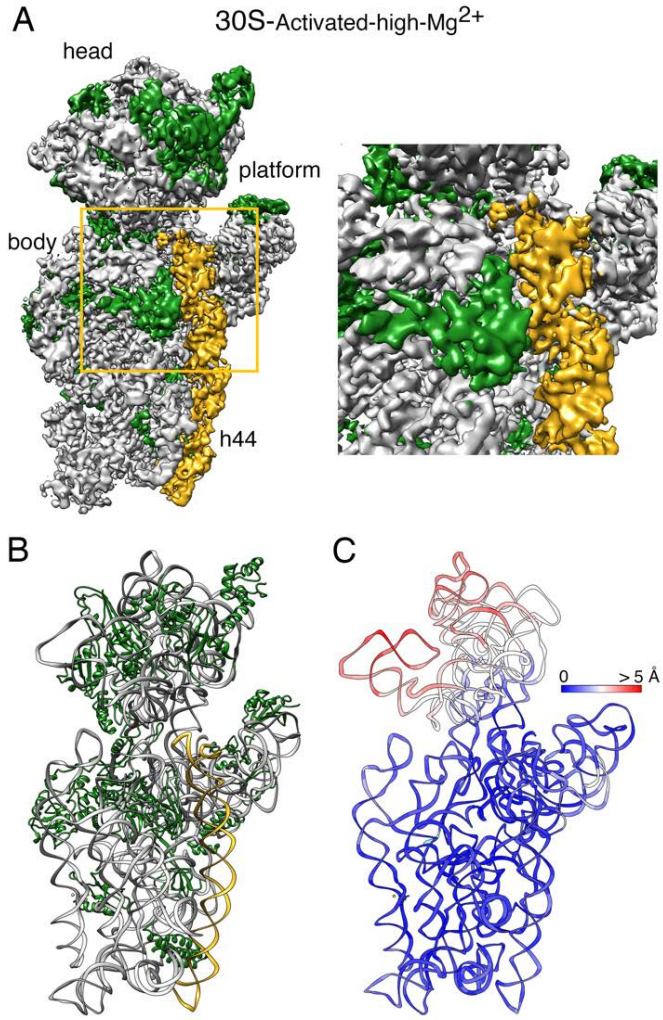
To quantitatively assess the conformational differences between the 16S rRNA in the structure obtained from grids having or lacking the extra carbon layer, we produced a molecular model from the cryo-EM map obtained from grids containing the extra carbon layer and we calculated a temperature map **(Fig. 2.3C)** with respect to the 30S-Inactivated-high-Mg^{2+} class A particles that were imaged without extra layer of carbon on the grids. We found that the conformation of the 16S rRNA in both structures was very similar, including the alternative folding adopted by the upper domain of helix 44 and 3' end of the rRNA molecule.

These results revealed that the air-water interface caused uS2 to fall-off. They also demonstrated that the non-canonical folding observed for helix 44 and 3' end of the 16S rRNA was induced by the exposure of the ribosomal particles to low magnesium concentrations and it occurs independently of the presence of uS2.

Transition from the inactive to the active state reverts the decoding center to the standard conformation.

Elson and colleagues (Zamir et al., 1969, Zamir et al., 1971) described that incubation of inactive 30S subunits at 42 ºC in the presence of 10-20 mM Mg^{2+} reverts the structural changes induced by low magnesium concentration to an 'active' conformation, in which the 30S subunits regain their ability to bind tRNA.

To structurally describe the conformational changes that this incubation triggers, purified 30S-Inactivated-high-Mg^{2+} ribosomal particles were incubated at 42 ºC and imaged by cryo-EM. These particles were called 30S-Activated-high-Mg^{2+}. Image classification revealed that the activation treatment had transformed all the 30S subunits to adopt a single conformation **(Figure. 2.4A)** that closely resembled the structure of the 30S subunit as described by X-ray crystallography **(Figure. 2.1)** (Wimberly et al., 2000). The structure refined to a resolution of 3.6 Å **(Supplemental Fig.**



S2.1) and allowed us to derive a molecular model from the cryo-EM map **(Figure. 2.4B)**. We then used this model to calculate a temperature map that compared this structure with the crystallographic structure. We observed that helix 44 and rRNA forming the decoding center closely overlapped indicating that the 42 ºC incubation treatment reverted the conformation of the functional domain induced by low magnesium concentration to that observed in the 70S ribosome and the 30S subunit structure produced by X-ray crystallography **(Figure. 2.4C)**.

Figure 2. 4 Cryo-EM structure of the 30S-activated-high-Mg2+ particle.

(A) Front view (left panel) of the cryo-EM map obtained for the 30S-activated-high-Mg2+ particle. The framed area is shown as a zoomed-in view in the right panel. The main landmarks of the 30S subunit are labeled. The rRNA is shown in light gray, and the r-proteins are shown in green. Helix 44 is shown in goldenrod orange. (B) Molecular model of the 30S-activated-high-Mg2+ particle. The rRNA and the r-proteins are colored as in panel A. (C) Temperature map of the 30S-activated-high-Mg2+ molecular model. The rRNA is colored according to the r.m.s.d. deviation (Å) with respect to the structure of the 30S subunit obtained by X-ray crystallography (PDB ID 4V4Q).

Extended exposure to low magnesium concentration induces unfolding of rRNA structural domains.

Next, we inquired about the effect of extended exposure of the 30S subunits to low concentration of magnesium ions. To this end, the purified 30S subunits were maintained in buffer containing 1.1 mM magnesium acetate before imaging them by cryo-EM. Similar classification approaches to those followed in the previous samples revealed that the 30S ribosomal subunits coexisted under low magnesium concentrations in two distinct conformations **(Figure. 2.5A)**. One of the classes represented 64% of the population and these particles generated a cryo-EM map that refined to 3.4 Å resolution (30S-Inactivated-low-Mg^{2+} class A) **(Supplemental Figure. S2.2)**. The remaining particles generated a different cryo-EM map that refined to 3.9 Å resolution (30S-Inactivated-low-Mg^{2+} class B) **(Supplemental Figure. S2.2)**. Both classes significantly diverged from the structure considered as a mature 30S subunit and contained structural features only previously observed in structures of immature 30S subunits (Jomaa et al., 2011a) (Leong et al., 2013) (Guo et al., 2013) (Razi et al., 2019). Density representing the entirety of helix 44 was missing in the two cryo-EM maps **(Figure. 2.5A)**.

Similarly, helices 23 and 24 in the platform region exhibited highly fragmented densities and the r-proteins bound to this region (bS6, uS11, bS21 and bS18) were also not observed in the cryo-EM map **(Figure. 2.5B & 2.5C)**. The EM grids used to image this sample did not contain an additional layer of continuous carbon and consequently uS2 was missing from the cryo-EM maps of both classes **(Figure. 2.5A, lower panel)**. The main difference observed between class A and B was in the head domain that was completely missing in the cryo-EM map for class B. Density for all regions of this domain was present in the map obtained for class A, except for r-protein uS7

that similar to other structures or the 30S subunit obtained in solution, may be adopting a flexible conformation **(Figure. 2.5A, top panel)**.

The molecular model derived from the 30S-Inactivated-low-Mg^{2+} class A and B cryo-EM maps were used to calculate a temperature map **(Figure. 2.5D)** to compare the conformation of the 16S rRNA in these structures with that of the crystallographic structure (Wimberly et al., 2000). The body and platform regions present in the cryo-EM map of the 30S-Inactivated-low-Mg^{2+} class A showed a close overlap. However, we found that the head domain was tilted backwards in class A, similarly to what we observed in the structures derived from the 30S-Inactivated-high-Mg^{2+} particles.

We also imaged this sample in grids containing an extra thin layer of continuous carbon. This experiment allowed us to determine whether the observed depletions of r-proteins (uS2, bS6, uS11, bS21 and bS18) and unfolding of rRNA domains in these structures was caused by the exposure of the particles to the low magnesium concentrations or because the buffer conditions made the particles easily damaged by the air-water interface. Three-dimensional classification revealed the existence of only one class of particles in the images (30S-Inactivated-Carbon-low-Mg^{2+}). From these particles, we generated a cryo-EM map that refined to 3.1 Å resolution **(Figure. 2.5E & Supplemental Fig. S2.2)**. The cryo-EM map resembled the structure of the 30S-Inactivated-low-Mg^{2+} class A, as the entirety density for helix 44 was also missing. However, the platform did not exhibit any severe distortion with helices 23 and 24 and r-proteins bS6, uS11 and bS18 were fully represented in the cryo-EM map. In addition, r-protein uS2 was also fully described. Importantly, we did not observe any particles that resemble the 30S-Inactivated-low-Mg^{2+} class B, which lacked the entire head domain.

These structures suggested that continued exposure of the mature 30S subunit to low magnesium concentrations partially destabilizes large structural motifs of the ribosomal subunit, including the central (platform), 3' major (head) and 3' minor (helix 44) domain and makes them more susceptible to unfolding when exposed to the air-water interface.

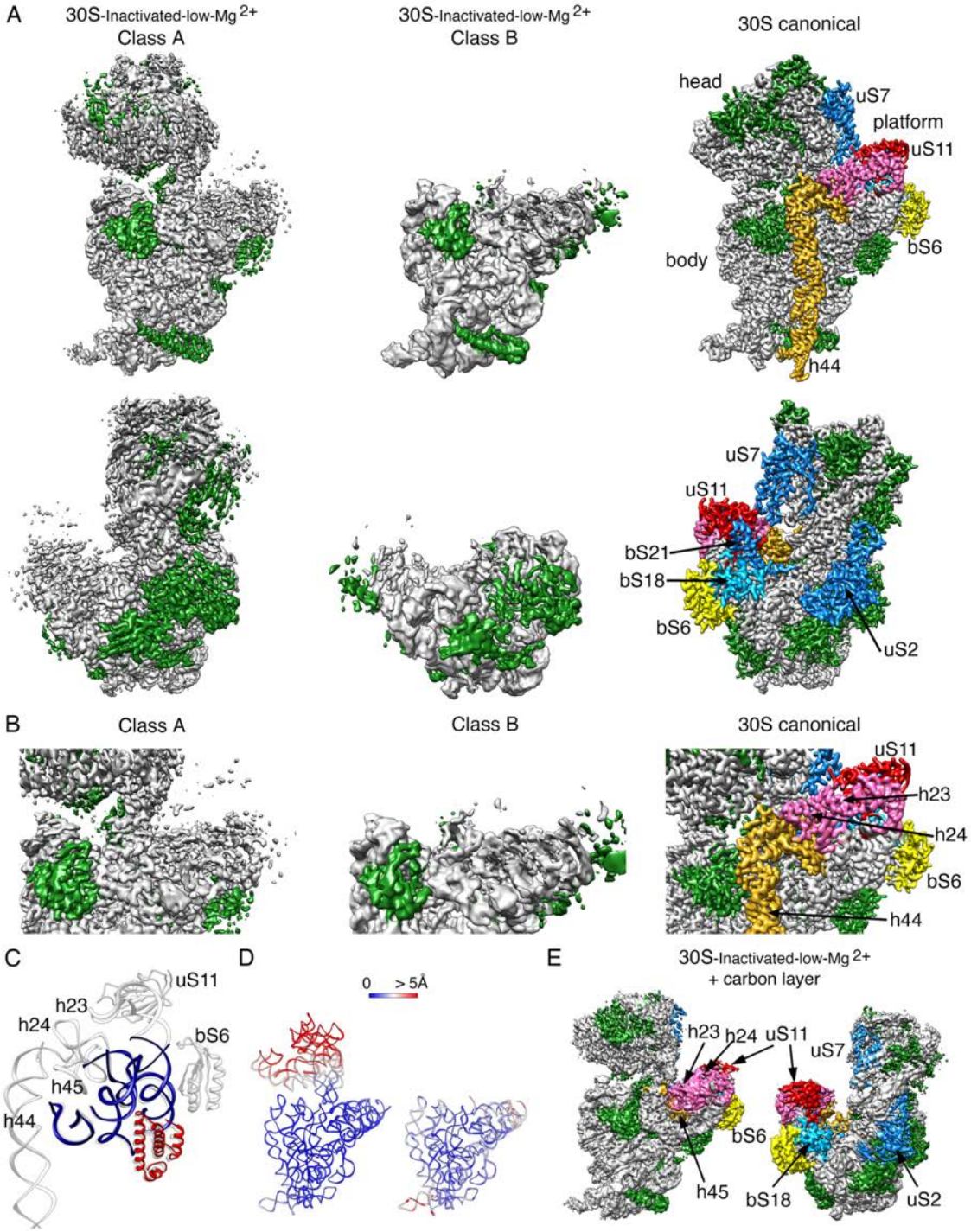


Figure 2. 5 Cryo-EM structure of 30S-inactivated-low-Mg2+ particle.

(A) Front (top panels) and back (bottom panels) views of the two conformers, class A and B of the 30S-inactivated-low-Mg2+ particle. The rRNA is shown in light gray, and the r-proteins are shown in green. Helix 44 is shown in goldenrod orange. These structures are shown side-by-side with the 30S subunit structure obtained by X-ray crystallography (PDB file 4V4Q). In this structure, the rRNA is displayed in light gray, the r-proteins in green, and helix 44 in goldenrod orange. The r-proteins uS2, bS6, uS7, uS11, uS18, and bS21 for which a representative density does not appear in the cryo-EM maps of the 30S-inactivated-low-Mg2+ particle are indicated using a color different than green. These proteins and other landmarks of the 30S subunit are labeled. (B) Zoomed-in view of the decoding and platform region of the cryo-EM maps obtained for the 30S-inactivated-low-Mg2+ class A and B using the same color coding as in (A). (C) Overlap of helix 44 and platform region of the molecular model derived from the cryo-EM structure of the 30S-inactivated-low-Mg2+ class A and the corresponding region from the structure of the 30S subunit obtained by X-ray crystallography. Parts of the structure present in the cryo-EM structure of the 30S-inactivated-low-Mg2+ class A are displayed in navy blue and red. The atomic model of the X-ray structure with all the elements of the complete 30S subunit structure is displayed in light gray. (D) Temperature maps of the 30S-inactivated-low-Mg2+ class A and B molecular models. The rRNA is colored according to the r.m.s.d. deviation (Å) with respect to the structure of the 30S subunit obtained by X-ray crystallography (PDB ID 4V4Q). (E) Front (left panel) and back (right panel) views of the 30S-inactivated-low-Mg2+ particle imaged in grids with an extra layer of continuous carbon. The r-proteins and rRNA helices missing in the same sample imaged in grids without the extra thin layer of carbon but present in this cryo-EM map are colored as in the structure of the 30S subunit obtained by X-ray crystallography shown in panel A.

Motions exhibited by the free 30S subunits.

We noticed that in all the cryo-EM maps obtained for the 30S subunit, density was clear and mostly complete for the body and platform domain but was fuzzier and fragmented in the head region **(Figure. 2.6)**. We interpreted the partial fragmentation of these densities as an indication of the motions that the head region experiences when the 30S subunits are in solution. Using multi-body refinement (Nakane et al., 2018), we investigated these motions. To this end, we split the cryo-EM maps into three bodies corresponding to three of the major domains of the 30S subunit: body, platform and head. After multi-body refinement, we compared the resulting maps obtained for the head domains with those from the consensus refinement. Visual inspection of their central section **(Figure. 2.6)** showed a small improvement, particularly in the region furthest away from the center of the map.

More importantly, this approach performs a principal component analysis of the variance in the rotations and translations of the three bodies and allowed us to generate movies describing the most important motions in the different 30S subunit populations. In all cases, this analysis revealed that between 38-43% of the variance in the rotations and translations of the three bodies is explained by the first three eigenvectors **(Figure. 2.6)**. Movies of the reconstructed body densities repositioned along these three eigenvectors revealed that they correspond to motions of the head with respect to the body and platform. In one of these motions the head rotates around an axis longitudinal to the longest dimension of the particle. The other two motions are back and forward tilting movements of the head with respect to the body. **Supplemental Movie 2.2 to 2.6** shows the main motion (represented by the first eigen vector) for each one of the main 30S structures obtained in this study. Overall, these results indicated that the most predominant motions exhibited by the head domain in the free 30S subunits are similar to those observed in the 70S ribosome during protein translation (Noeske and Cate, 2012).

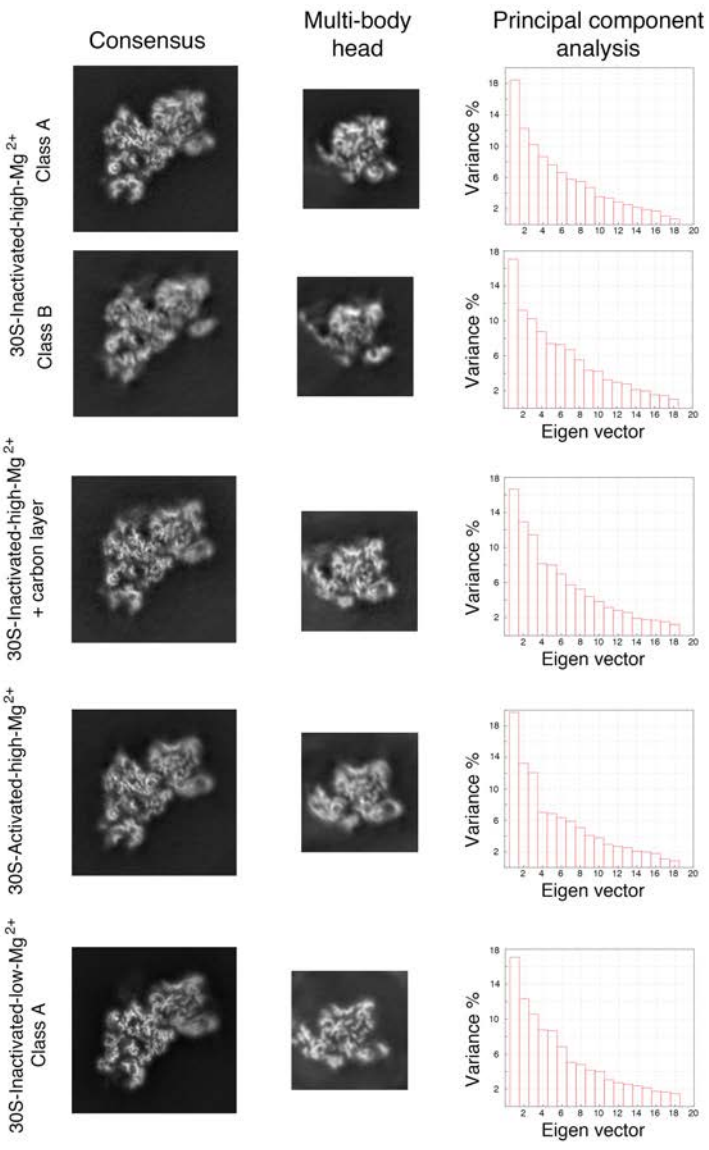


Figure 2. 6 Multibody refinement analysis of the cryo-EM structures.

The data set for each type of 30S particle was analyzed using multibody refinement to visualize their motions. The left and middle panels show, respectively, the central sections of the cryo-EM maps of the obtained 30S structures after the consensus refinement and of the head after multibody refinement. The high-resolution features of the density maps are apparent in the body region. Still, they are slightly blurred in the head domain in the central sections of the maps obtained through consensus refinement. High-resolution features of the head become more apparent in the maps of this region obtained through multibody refinement. The right panels show the principal component analysis of the different structures using the multibody refinement routine in RELION 3.0. This analysis indicated that between 38% and 43% of the variance is due to the movements of the head with respect to the body and platform.

2.5 Discussion

The transition of the decoding center in the 30S subunit from the active to the inactive state was one of the first conformational rearrangements discovered in ribosomes (Zamir et al., 1969, Zamir et al., 1971, Moazed et al., 1986). Chemical probing experiments at the time revealed that adoption of the inactive conformation involved large-scale structural changes in the neck region and the fraction of helix 44 involved in the formation of the decoding center. A precise structure could not be derived at the time from the data. However, there is a remarkable agreement between the cryo-EM structure of the inactive conformation of the 30S subunit described here and the probing data from the Noller group (Moazed et al., 1986). In particular, the decreased reactivity of nucleotides 1533-1536 suggested by the probing data correlates well with our finding of these nucleotides being involved in the base pairing with 921-923. Similarly, the enhanced reactivity found for nucleotides 924-926 and 1394-1397 suggested these nucleotides become unpaired, which is also observed in cryo-EM structure.

More recently, RNA SHAPE experiments chemically probed the RNA structure of the 30S subunits in exponentially growing *E. coli* cells (McGinnis et al., 2015, McGinnis and Weeks, 2014). An important contribution of these experiments was revealing that the inactive state of the 30S subunit is biologically relevant in regulating ribosome function. Based on the SHAPE data this group also suggested an alternative base pairing in the 16S rRNA forming the decoding center in the inactive conformation. In particular, nucleotides 1402-1408, which form an irregular helix at the top of helix 44 pairing with nucleotides 1492-1500 in the conventional structure, undergo a register shift and pair with positions 921-927 within helix 28 and nucleotides 1390-1401 form an unpaired loop. The obtained high-resolution cryo-EM map of the inactive conformation of the 30S subunit presented here diverges from the one proposed in the RNA SHAPE experiments

(McGinnis et al., 2015). In the cryo-EM map obtained for the 'inactive' conformation nucleotides 1532-1534 at the 3' end of the 16S rRNA base paired with nucleotides 921-923 in helix 28 causing the unfolding of the bottom of this helix and the top of helix 44. The newly formed base pairs force the 3' end of the 16S rRNA in a conformation that resembles the conformation adopted by mRNA during translation.

Comparison of the previous chemical probing data (Moazed et al., 1986, McGinnis et al., 2015) with our cryo-EM results shows the clear advantage of cryo-EM for the study of the alternative conformations adopted by ribosomal particles. Whereas in some cases multiple structures can be potentially derived from chemical probing data, the direct observation of the structure using cryo-EM typically allows to establish unambiguously the precise nature of the conformational rearrangement that the ribosomal particle undergoes. We found this to be a likely explanation between the disagreement of the structure proposed from the SHAPE data (McGinnis et al., 2015) and that shown by cryo-EM here.

The functional importance of the "inactive" conformation of the 30S subunit has been highlighted in previous studies (Myasnikov et al., 2009) (Karbstein, 2013). Multiple translation initiation factors and RNAses affecting 30S subunit turnover bind at sites overlapping with helices 28 and 44, suggesting that the "active-inactive" transition observed by cryo-EM could govern accessibility of these factors and regulates translation initiation and ribosome quality control. In addition, the cryo-EM structure of the 30S-Inactivated-high-Mg^{2+} class A provides a functional explanation for the inactivity of the ribosomal particle in this conformation. The 3' end of 16S rRNA of the 30S subunit containing the anti-Shine-Dalgarno sequence is typically accessible in the platform of the 30S subunit and base pairs with the Shine-Dalgarno sequence in the 5' end of the mRNA. However, in the inactive conformation the entire 3' end of 16S rRNA adopts a

conformation similar to that adopted by mRNA during translation. In this position, the anti-Shine-Dalgarno sequence is occluded and unable to recognize the Shine-Dalgarno sequence during initiation.

Our study also characterizes the motions that the different 30S subunit subpopulations exhibit when they are in solution and not constrained in a crystal lattice. X-ray crystallographic and cryo-EM structures from many laboratories have described the large-scale conformational changes the ribosome undergoes during protein synthesis (Noeske and Cate, 2012). For example, the movement of the tRNAs from the classical A/A and P/P configurations to the hybrid A/P and P/E states is driven by a rigid body rotation of the small subunit with respect to the large subunit by 10º around an axis perpendicular to the subunit interface (ratchet-like motion) (Frank and Agrawal, 2000). This movement is also accompanied with a simultaneous swiveling of the 30S subunit head domain in the direction of the E site (Spahn et al., 2004). Swiveling motions of the head also accompany the process of termination and ribosome recycling (Dunkle et al., 2011, Yokoyama et al., 2012). The head domain also exhibits other motions linked to other physiological processes such as hibernation. Binding of ribosome modulation factor or hibernation promoting factor triggers a displacement of the head away from the central protuberance (backward tilting), promoting formation of 100S ribosome dimers (Polikanov et al., 2012). All these motions have been observed in the context of the 70S ribosome. In our study, the use of cryo-EM combined with multibody refinement processing approaches (Nakane et al., 2018) allowed us to investigate the most prominent motions of the different conformational subpopulations identified for free 30S subunit. Overall, we found that the motions of swiveling and backward tilting that the head domain of the 30S subunit undergoes in the context of the 70S ribosome are also exhibited by the free 30S subunits.

An important motivation for this study was to provide a more extensive comparative reference for those studies on assembly of the 30S subunit that use single deletion or depletion strains of particular assembly factors. Drawing meaningful conclusions from the heterogeneous mixture of ribosomal particles that these strains accumulate is only possible when a comprehensive description of the conformations adopted by the mature subunit is available. Only then it is possible to separate from the observed pool of structures, those truly representing immature assembly states and derive information about the function of specific assembly factors.

Current sample preparation techniques in cryo-EM suffers from a fundamental problem. Macromolecules in solution drift during vitrification and get exposed to the air-water interface at the top and bottom of the thin films formed upon blotting. This exposure is damaging to the macromolecular assemblies and can pull their components apart and destroy them. Our results show that even for specimens considered resilient such as the ribosome or the 30S subunit, current cryo-EM sample preparation methods still cause damages. In particular, we found that r-protein uS2 is especially sensitive to the air-water interface and in our experiment caused the complete removal of the protein during the vitrification process. Most likely, this problem is exacerbated in ribosome assembly intermediates where multiple r-proteins may not be bound as tightly as in the mature subunit. Consequently, similarly to some previous studies (Li et al., 2013) (Jomaa et al., 2014) (Guo et al., 2013) (Leong et al., 2013) (Razi et al., 2017b) (Razi et al., 2019), it may be prudent that future studies analyzing ribosome assembly intermediates by cryo-EM continue to use grids containing an extra thin layer of continuous carbon to decrease exposure of the particles to the air-water interface. Importantly, the structures presented here will contribute to reinterpret previous cryo-EM work performed in grids without an additional layer of continuous carbon (Lopez-Alonso et al., 2017a) and help to differentiate *bona fide* structural defects observed on

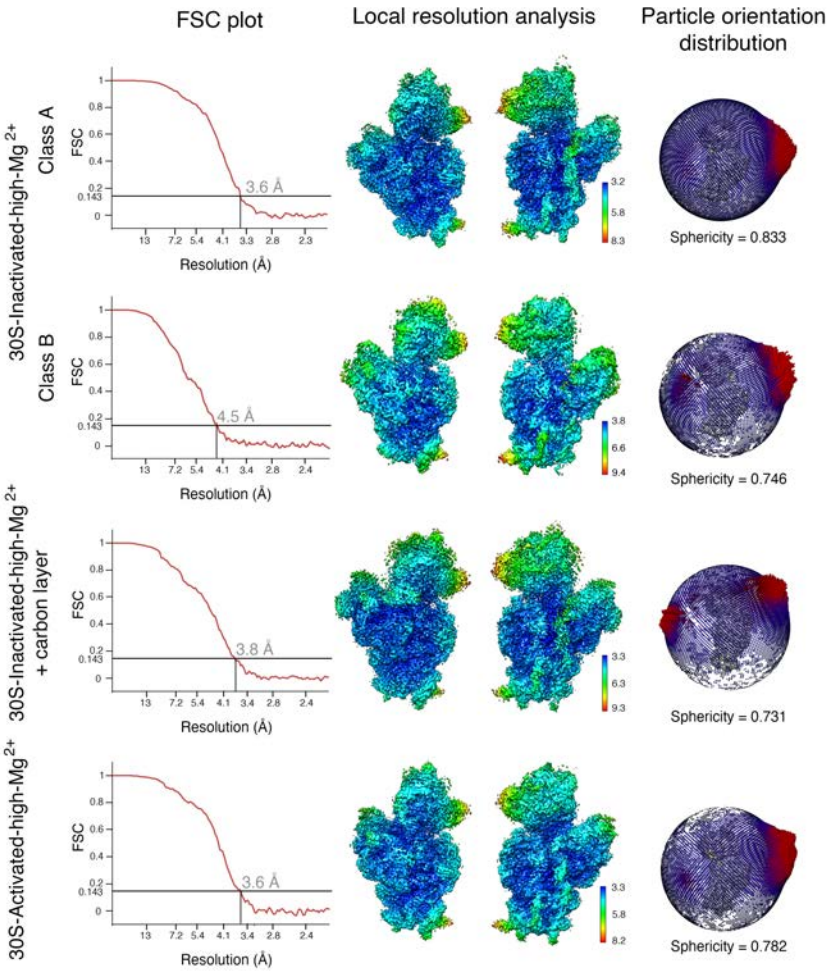
heterogeneous mixtures of 30S ribosomal particles from those caused by the sample preparation technique.

Overall, this study provides new insights into the alternative conformations and motions adopted by the 30S ribosomal subunit. The existence of one of them, named by Elson as the “inactive” conformation, was discovered over five decades ago. It has been only now with the recent improvements in electron microscopes and direct electron detector cameras that we have been able to describe these structures at high-resolution.

2.6 Supplemental Material

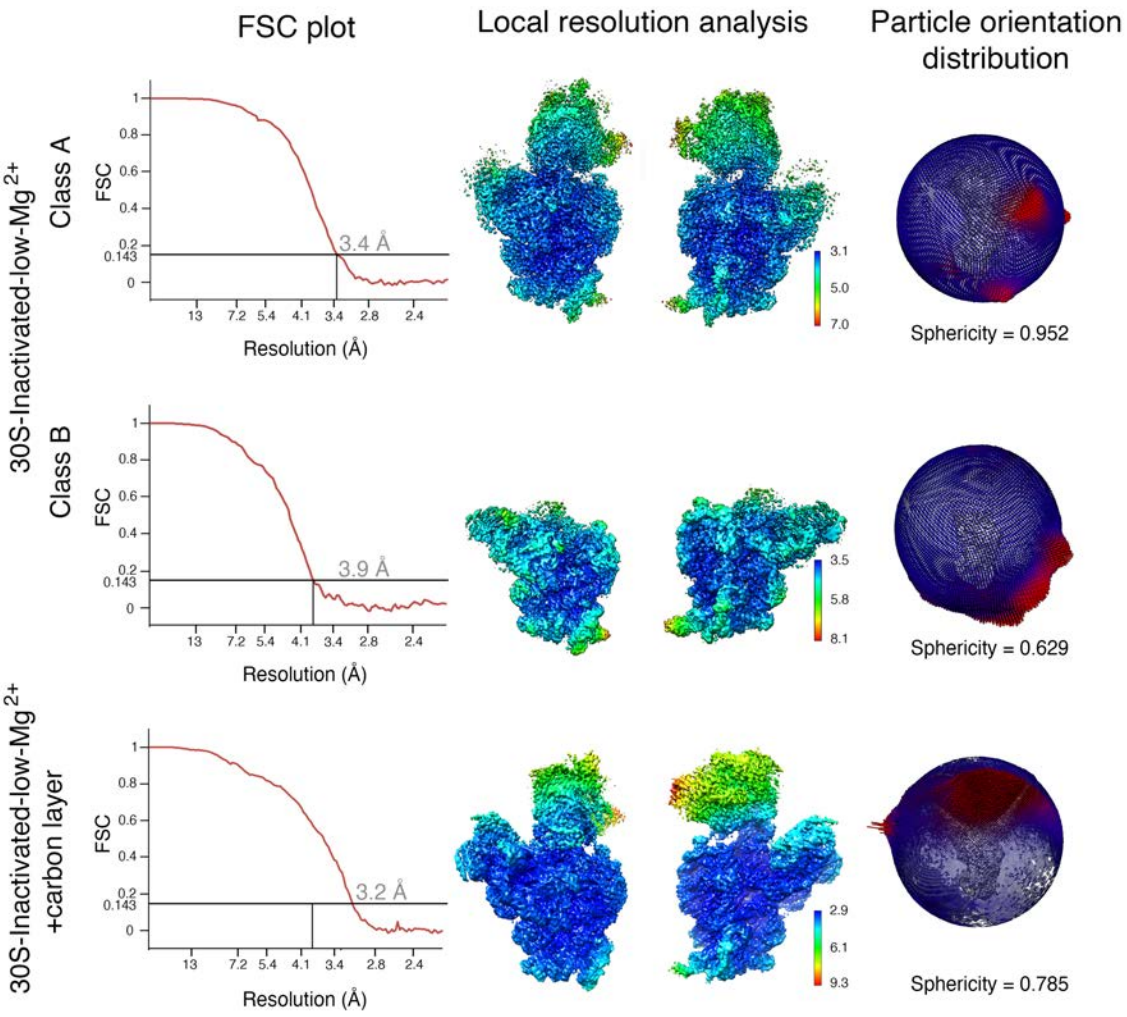
Supplemental material is available for this article.

Supplemental Figure



Supplemental Figure S2. 1 Resolution analysis of the cryo-EM maps obtained under high Mg2+ concentrations.

Left column shows the Fourier shell correlation (FSC) plots for the different cryo-EM maps obtained from samples maintained in buffer containing 10 mM magnesium acetate. Resolution is reported using FSC threshold of 0.143. The middle column displays the cryo-EM maps colored according to the local resolution analysis performed with Relion. The right column shows a sphere representing the angular distribution of the particles in each dataset. Each view angle is represented as a dot in the sphere. The height and color of the sphere relate to the number of particles representing each view. Areas in red indicate a higher number of particles representing those particular views.



Supplemental Figure S2. 2 Resolution analysis of the cryo-EM maps obtained under low Mg2+ concentrations.

Fourier shell correlation (FSC) plots, local resolution analysis and angular distribution for the cryo-EM maps obtained from samples maintained in buffer containing 1.1 mM magnesium acetate. The layout of the figure is as in Supplemental Figure S1.

Supplemental Movie Captions

Supplemental Movie 2.1. Conformational transition between the canonical structure of the 30S subunit obtained by X-ray crystallography and the 30S-Inactivated-high-Mg^{2+} particle. *The transition between both conformations involves repositioning of nucleotides 1532-1534 that are not forming any base pairing in the conventional structure to base pair with nucleotides 921-923. In this process the bottom of helix 28 (region formed by nucleotides 1391-1396 and 921-925) and top of helix 44 formed in the canonical structure by nucleotides 1397-1407 and 1494-1503 become unfolded. The 3' end of the 16S rRNA (distal to nucleotide 1534) adopts a conformation similar to that adopted by mRNA during translation. The rRNA is display in light grey, the r-proteins in red. Helices 28, 44 and 45 are colored in navy blue except the region on these helices that unfolds during the conformational change that is colored in yellow. Nucleotides 1532-1534 that upon the conformational transition form base pares with nucleotides 921-923 are shown in red.*

Supplemental Movie 2.2-2.6 Main motions of the 30S particles. *Each movie shows the main motion represented by the first eigen vector in each type of 30S particle.*
Supplemental Movies can be found using the following link:
https://rnajournal.cshlp.org/content/early/2020/09/28/rna.075846.120/suppl/DC1

Supplemental movies in the chapter are renamed for consistency. Therefore, Supplemental_Video_1.mov is renamed as Supplemental Movie 2.1.mov

Supplemental Table

Supplemental Table S2. 1 Data Deposition

MAP	EMDB code	PDB ID
30S-Inactive-high-Mg^{2+} Class A	21569	6W77
30S-Inactive-high-Mg^{2+} Class B	21570	-
30S-Inactive-high-Mg^{2+} + carbon layer	21571	6W7M
30S-Activated-high-Mg^{2+}	21558	6W6K
30S-Inactive-low-Mg^{2+} Class A	21572	6W7N
30S-Inactive-low-Mg^{2+} Class B	21573	6W7W
30S-Inactive-low-Mg^{2+} + carbon layer	XXXX	-

Supplemental Table S2. 2 Cryo-EM data acquisition, processing and map and model statistics.

		30S-Inacti high-Mg Class A	30S-Inacti high-Mg Class B	30S-Inacti high-Mg +carbo	30S-Activat high-Mg^2	30S-Inacti low-Mg^2 Class A	30S-Inacti low-Mg^2 Class B	30S-Inacti low-Mg^2 +carbo
Data collection								
Microscope		Titan Krios		Titan Krios	Titan Krios	Titan Krios		Titan Krios
Detector		Falcon II		Falcon II	Falcon II	Falcon II		K3
Nominal Magnification		75,000x		75,000x	75,000x	75,000x		81,000x
Voltage (kV)		300		300	300	300		300
Total exposure (e$^-$/Å^2)		50		52	52	50		60
Defocus range (□m)		-1.25 to -2.75		-1.25 to -2.75	-1.25 to -2.75	-1.25 to -2.75		-1 to -2.25
Calibrated physical pixel size (Å/px)		1.073		1.073	1.073	1.073		1.09
Reconstruction and refinement								
Particles		446,530	118,725	334,903	407,623	421,738	236,327	264,418
Map sharpening B factor		-116	-136	-98	-100	-121	-161	-10
Resolution (Å)		3.6	4.5	3.8	3.6	3.4	3.9	3.2
FSC Threshold		0.143	0.143	0.143	0.143	0.143	0.143	0.143
Model composition								
RNA chains		1	-	1	1	1	1	-
Protein chains		17	-	20	17	14	9	-
Model Building								
Protein Geometry	Poor rotamers	0.00%	-	0.1%	0.00%	0.00%	0.12%	-
	Favored rotamers	95.47%	-	93.90%	96.08%	95.94%	96.71%	-
	Ramachandran outliers	0.21%	-	0.31%	0.00%	0.31%	0.00%	-
	Ramachandran favored	93.73%	-	92.28%	93.07%	91.60%	95.92%	-
	Cβ deviations >0.25Å	0.00%	-	0.00%	0.00%	0.00%	0.00%	-
	Bad bonds	0.00%	-	0.00%	0.00%	0.00%	0.00%	-
	Bad angles	0.00%	-	0.00%	0.00%	0.00%	0.00%	-
Nucleic Acid Geometry	Probably wrong sugar puck	0.86%	-	1.12%	0.65%	0.58%	0.46%	-
	Bad backbone conformatio	19.42%	-	27.38%	20.08%	16.76%	20.90%	-
	Bad bonds	0.00%	-	0.00%	0.00%	0.00%	0.00%	-
	Bad angles	0.03%	-	0.01%	0.00%	0.01%	0.00%	-
Low-resolution criteria	CaBLAM outliers	4.4%	-	4.4%	5.4%	6.0%	3.2%	-
	CA Geometry outliers	0.96%	-	0.94%	1.28%	1.00%	0.99%	-
Additional validations	Chiral Volumes outliers	0/9933	-	0/10342	0/9855	0/8827	0/5606	-

Acknowledgements

We thank Kelly Sears, Mike Strauss and other staff members of the Facility for Electron Microscopy Research (FEMR) at McGill University for help with microscope operation and data collection. We acknowledge Clara Ortega for assistance with graphic design. This work was supported by grants from the Canadian Institutes of Health Research (PJT-153044) to J.O. Titan Krios cryo-EM data were collected at FEMR (McGill). FEMR is supported by the Canadian Foundation for Innovation, Quebec government and McGill University. The authors declare no competing financial interests. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

2.7 References

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., MCCOY, A. J., MORIARTY, N. W., OEFFNER, R., READ, R. J., RICHARDSON, D. C., RICHARDSON, J. S., TERWILLIGER, T. C. & ZWART, P. H. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr,* 66**,** 213-21.

BABA, T., ARA, T., HASEGAWA, M., TAKAI, Y., OKUMURA, Y., BABA, M., DATSENKO, K. A., TOMITA, M., WANNER, B. L. & MORI, H. 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol,* 2**,** 2006 0008.

CHENG, Y. 2015. Single-Particle Cryo-EM at Crystallographic Resolution. *Cell,* 161**,** 450-7.

CHENG, Y., GLAESER, R. M. & NOGALES, E. 2017. How Cryo-EM Became so Hot. *Cell,* 171**,** 1229-1231.

DAIGLE, D. M. & BROWN, E. D. 2004. Studies of the interaction of Escherichia coli YjeQ with the ribosome in vitro. *J Bacteriol,* 186**,** 1381-7.

DARTY, K., DENISE, A. & PONTY, Y. 2009. VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics,* 25**,** 1974-5.

DATTA, P. P., WILSON, D. N., KAWAZOE, M., SWAMI, N. K., KAMINISHI, T., SHARMA, M. R., BOOTH, T. M., TAKEMOTO, C., FUCINI, P., YOKOYAMA, S. & AGRAWAL, R. K. 2007. Structural aspects of RbfA action during small ribosomal subunit assembly. *Mol Cell,* 28**,** 434-45.

DE LA ROSA-TREVIN, J. M., OTON, J., MARABINI, R., ZALDIVAR, A., VARGAS, J., CARAZO, J. M. & SORZANO, C. O. 2013. Xmipp 3.0: an improved software suite for image processing in electron microscopy. *J Struct Biol,* 184**,** 321-8.

DUNKLE, J. A., WANG, L., FELDMAN, M. B., PULK, A., CHEN, V. B., KAPRAL, G. J., NOESKE, J., RICHARDSON, J. S., BLANCHARD, S. C. & CATE, J. H. 2011. Structures of the bacterial ribosome in classical and hybrid states of tRNA binding. *Science,* 332**,** 981-4.

EMSLEY, P. & COWTAN, K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr,* 60**,** 2126-32.

EMSLEY, P., LOHKAMP, B., SCOTT, W. G. & COWTAN, K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr,* 66**,** 486-501.

FRANK, J. 2017. The translation elongation cycle-capturing multiple states by cryo-electron microscopy. *Philos Trans R Soc Lond B Biol Sci,* 372.

FRANK, J. & AGRAWAL, R. K. 2000. A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature,* 406**,** 318-22.

GUO, Q., GOTO, S., CHEN, Y., FENG, B., XU, Y., MUTO, A., HIMENO, H., DENG, H., LEI, J. & GAO, N. 2013. Dissecting the in vivo assembly of the 30S ribosomal subunit reveals the role of RimM and general features of the assembly process. *Nucleic Acids Res,* 41**,** 2609-20.

HUSSAIN, T., LLACER, J. L., WIMBERLY, B. T., KIEFT, J. S. & RAMAKRISHNAN, V. 2016. Large-Scale Movements of IF3 and tRNA during Bacterial Translation Initiation. *Cell,* 167**,** 133-144 e13.

JOMAA, A., JAIN, N., DAVIS, J. H., WILLIAMSON, J. R., BRITTON, R. A. & ORTEGA, J. 2014. Functional domains of the 50S subunit mature late in the assembly process. *Nucleic Acids Res,* 42**,** 3419-35.

JOMAA, A., STEWART, G., MARTIN-BENITO, J., ZIELKE, R., CAMPBELL, T. L., MADDOCK, J. R., BROWN, E. D. & ORTEGA, J. 2011. Understanding ribosome assembly: the structure of in vivo assembled immature 30S subunits revealed by cryo-electron microscopy. *RNA,* 17**,** 697-709.

KARBSTEIN, K. 2013. Quality control mechanisms during ribosome maturation. *Trends Cell Biol,* 23**,** 242-50.

KUHLBRANDT, W. 2014. Biochemistry. The resolution revolution. *Science,* 343**,** 1443-4.

LEONG, V., KENT, M., JOMAA, A. & ORTEGA, J. 2013. Escherichia coli rimM and yjeQ null strains accumulate immature 30S subunits of similar structure and protein complement. *RNA,* 19**,** 789-802.

LI, N., CHEN, Y., GUO, Q., ZHANG, Y., YUAN, Y., MA, C., DENG, H., LEI, J. & GAO, N. 2013. Cryo-EM structures of the late-stage assembly intermediates of the bacterial 50S ribosomal subunit. *Nucleic Acids Res,* 41**,** 7073-83.

LOPEZ-ALONSO, J. P., FABBRETTI, A., KAMINISHI, T., ITURRIOZ, I., BRANDI, L., GIL-CARTON, D., GUALERZI, C. O., FUCINI, P. & CONNELL, S. R. 2017. Structure of a 30S pre-initiation complex stalled by GE81112 reveals structural parallels in bacterial and eukaryotic protein synthesis initiation pathways. *Nucleic Acids Res,* 45**,** 2179-2187.

LU, X. J. & OLSON, W. K. 2008. 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat Protoc,* 3**,** 1213-27.

MCGINNIS, J. L., LIU, Q., LAVENDER, C. A., DEVARAJ, A., MCCLORY, S. P., FREDRICK, K. & WEEKS, K. M. 2015. In-cell SHAPE reveals that free 30S ribosome subunits are in the inactive state. *Proc Natl Acad Sci U S A,* 112**,** 2425-30.

MCGINNIS, J. L. & WEEKS, K. M. 2014. Ribosome RNA assembly intermediates visualized in living cells. *Biochemistry,* 53**,** 3237-47.

MOAZED, D., VAN STOLK, B. J., DOUTHWAITE, S. & NOLLER, H. F. 1986. Interconversion of active and inactive 30 S ribosomal subunits is accompanied by a conformational change in the decoding region of 16 S rRNA. *J Mol Biol,* 191**,** 483-93.

MULDER, A. M., YOSHIOKA, C., BECK, A. H., BUNNER, A. E., MILLIGAN, R. A., POTTER, C. S., CARRAGHER, B. & WILLIAMSON, J. R. 2010. Visualizing ribosome biogenesis: parallel assembly pathways for the 30S subunit. *Science,* 330**,** 673-7.

MYASNIKOV, A. G., SIMONETTI, A., MARZI, S. & KLAHOLZ, B. P. 2009. Structure-function insights into prokaryotic and eukaryotic translation initiation. *Curr Opin Struct Biol,* 19**,** 300-9.

NAKANE, T., KIMANIUS, D., LINDAHL, E. & SCHERES, S. H. 2018. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *Elife,* 7.

NI, X., DAVIS, J. H., JAIN, N., RAZI, A., BENLEKBIR, S., MCARTHUR, A. G., RUBINSTEIN, J. L., BRITTON, R. A., WILLIAMSON, J. R. & ORTEGA, J. 2016. YphC and YsxC GTPases assist the maturation of the central protuberance, GTPase associated region and functional core of the 50S ribosomal subunit. *Nucleic Acids Res,* 44**,** 8442-55.

NIKOLAY, R., HILAL, T., QIN, B., MIELKE, T., BURGER, J., LOERKE, J., TEXTORIS-TAUBE, K., NIERHAUS, K. H. & SPAHN, C. M. T. 2018. Structural Visualization of the

Formation and Activation of the 50S Ribosomal Subunit during In Vitro Reconstitution. *Mol Cell,* 70**,** 881-893 e3.

NOBLE, A. J., WEI, H., DANDEY, V. P., ZHANG, Z., TAN, Y. Z., POTTER, C. S. & CARRAGHER, B. 2018. Reducing effects of particle adsorption to the air-water interface in cryo-EM. *Nat Methods,* 15**,** 793-795.

NOESKE, J. & CATE, J. H. 2012. Structural basis for protein synthesis: snapshots of the ribosome in motion. *Curr Opin Struct Biol,* 22**,** 743-9.

NOGALES, E. & SCHERES, S. H. 2015. Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. *Mol Cell,* 58**,** 677-89.

PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem,* 25**,** 1605-12.

POLIKANOV, Y. S., BLAHA, G. M. & STEITZ, T. A. 2012. How hibernation factors RMF, HPF, and YfiA turn off protein synthesis. *Science,* 336**,** 915-8.

RAZI, A., BRITTON, R. A. & ORTEGA, J. 2017a. The impact of recent improvements in cryo-electron microscopy technology on the understanding of bacterial ribosome assembly. *Nucleic Acids Res,* 45**,** 1027-1040.

RAZI, A., DAVIS, J. H., HAO, Y., JAHAGIRDAR, D., THURLOW, B., BASU, K., JAIN, N., GOMEZ-BLANCO, J., BRITTON, R. A., VARGAS, J., GUARNE, A., WOODSON, S. A., WILLIAMSON, J. R. & ORTEGA, J. 2019. Role of Era in assembly and homeostasis of the ribosomal small subunit. *Nucleic Acids Res,* 47**,** 8301-8317.

RAZI, A., GUARNE, A. & ORTEGA, J. 2017b. The cryo-EM structure of YjeQ bound to the 30S subunit suggests a fidelity checkpoint function for this protein in ribosome assembly. *Proc Natl Acad Sci U S A,* 114**,** E3396-E3403.

SASHITAL, D. G., GREEMAN, C. A., LYUMKIS, D., POTTER, C. S., CARRAGHER, B. & WILLIAMSON, J. R. 2014. A combined quantitative mass spectrometry and electron microscopy analysis of ribosomal 30S subunit assembly in E. coli. *Elife,* 3.

SCHORB, M., HABERBOSCH, I., HAGEN, W. J. H., SCHWAB, Y. & MASTRONARDE, D. N. 2019. Software tools for automated transmission electron microscopy. *Nat Methods,* 16**,** 471-477.

SCHURECK, M. A., MAEHIGASHI, T., MILES, S. J., MARQUEZ, J. & DUNHAM, C. M. 2016. mRNA bound to the 30S subunit is a HigB toxin substrate. *RNA,* 22**,** 1261-70.

SCHUWIRTH, B. S., BOROVINSKAYA, M. A., HAU, C. W., ZHANG, W., VILA-SANJURJO, A., HOLTON, J. M. & CATE, J. H. 2005. Structures of the bacterial ribosome at 3.5 A resolution. *Science,* 310**,** 827-34.

SEFFOUH, A., JAIN, N., JAHAGIRDAR, D., BASU, K., RAZI, A., NI, X., GUARNE, A., BRITTON, R. A. & ORTEGA, J. 2019. Structural consequences of the interaction of RbgA with a 50S ribosomal subunit assembly intermediate. *Nucleic Acids Res,* 47**,** 10414-10425.

SPAHN, C. M., GOMEZ-LORENZO, M. G., GRASSUCCI, R. A., JORGENSEN, R., ANDERSEN, G. R., BECKMANN, R., PENCZEK, P. A., BALLESTA, J. P. & FRANK, J. 2004. Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA translocation. *EMBO J,* 23**,** 1008-19.

STOKES, J. M. & BROWN, E. D. 2015. Chemical modulators of ribosome biogenesis as biological probes. *Nat Chem Biol,* 11**,** 924-32.

SWIATKOWSKA, A., WLOTZKA, W., TUCK, A., BARRASS, J. D., BEGGS, J. D. & TOLLERVEY, D. 2012. Kinetic analysis of pre-ribosome structure in vivo. *RNA,* 18**,** 2187-200.

WIMBERLY, B. T., BRODERSEN, D. E., CLEMONS, W. M., JR., MORGAN-WARREN, R. J., CARTER, A. P., VONRHEIN, C., HARTSCH, T. & RAMAKRISHNAN, V. 2000. Structure of the 30S ribosomal subunit. *Nature,* 407**,** 327-39.

YOKOYAMA, T., SHAIKH, T. R., IWAKURA, N., KAJI, H., KAJI, A. & AGRAWAL, R. K. 2012. Structural insights into initial and intermediate steps of the ribosome-recycling process. *EMBO J,* 31**,** 1836-46.

ZAMIR, A., MISKIN, R. & ELSON, D. 1969. Interconversions between inactive and active forms of ribosomal subunits. *FEBS Lett,* 3**,** 85-88.

ZAMIR, A., MISKIN, R. & ELSON, D. 1971. Inactivation and reactivation of ribosomal subunits: amino acyl-transfer RNA binding activity of the 30 s subunit of Escherichia coli. *J Mol Biol,* 60**,** 347-64.

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

ZHENG, S. Q., PALOVCAK, E., ARMACHE, J. P., VERBA, K. A., CHENG, Y. & AGARD, D. A. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods,* 14**,** 331-332.

ZIVANOV, J., NAKANE, T., FORSBERG, B. O., KIMANIUS, D., HAGEN, W. J., LINDAHL, E. & SCHERES, S. H. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife,* 7.

ZIVANOV, J., NAKANE, T. & SCHERES, S. H. W. 2019. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ,* 6**,** 5-17.

ZIVANOV, J., NAKANE, T. & SCHERES, S. H. W. 2020. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ,* 7**,** 253-267.

Connecting Text

In the previous chapter, we solved the structure of the 'inactive' 30S ribosome subunit using cryo-EM. We also visualized the transition between the 'inactive-active' states, which involve conformational changes in the 16S rRNA helices 44 and 45 that form the decoding region. In the next chapter, we investigate the potential role of YjeQ in the assembly and maturation of the 30S ribosome subunit. YjeQ binds to the 30S decoding region and induces conformational changes that convert the 30S from an 'inactive' to an 'active' state. Accordingly, YjeQ is suggested to use this functionality to test the ability of the 30S ribosome to flip critical decoding region nucleotides that ensure translational fidelity. How YjeQ achieves this and prevents the 30S from miscoding during translation elongation remains unknown. In the upcoming chapter, we use cryo-EM and fidelity assays to advance our understanding of the 30S ribosome assembly and the impacts of faulty ribosome biogenesis on ribosome function.

CHAPTER 3: Discovering the Role of YjeQ in Catalysis of Small Ribosome Subunit Assembly, Translational Fidelity, and Ribosome Rescue

Dushyant Jahagirdar[1,2], Anusha Naganathan[3], Jingyu Sun[1,2], Kaustuv Basu[1,2], Gloria Culver[3], and Joaquin Ortega[1,2] *

[1] Department of Anatomy & Cell Biology, McGill University, Montreal, Quebec H3A 0C7, Canada.

[2] Centre for Structural Biology, McGill University, 3649 Promenade Sir William Osler, Montreal, Quebec H3G 0B1, Canada.

[3] Department of Biology, University of Rochester, Rochester, New York 14627, USA

Manuscript to be submitted

Contribution of Authors

D.J created mutant YjeQ constructs, performed protein purification, wild type ribosome purifications, prepared 30S+Mutant YjeQ complex for cryo-EM, and performed all cryo-EM image analyses. J.S. purified 30S ribosome from *yjeQ* deletion strain, prepared sample for cryo-EM, and pre-processed data until particle-picking stage. K.B designed the cryo-EM data collection strategies on the Titan Krios at FEMR-McGill. A.N and G.C. designed and performed Miller's β-Galactosidase Assay. A.N., D.J., J.O analyzed Miller's β-Galactosidase assay data. D.J. and J.O. designed experiments, wrote, and edited the manuscript.

3.1 Abstract

YjeQ GTPase is a non-essential protein for the 30S ribosome subunit assembly. While *yjeQ* deletion impacts ribosome assembly and leads to the accumulation of immature 30S, the 30S assembly still proceeds to completion slowly. This study further characterizes the accumulated 30S particles using cryo-electron microscopy. We found helix 44, the last 16S rRNA component to get folded in the ribosome assembly, folds pre-maturely to promote the folding of helix 27, triggering the maturation of 30S. These observations suggest that the presence of YjeQ is critical for the timely maturation of the 30S. In the context of mature 30S, YjeQ has been proposed to play a role as a checkpoint protein. Accordingly, YjeQ induces flipping of helix 44 bases A1493, a conformational change commonly observed during mRNA decoding for protein synthesis. Combining cryo-electron microscopy with biochemical assays, we show how yjeQ deletion/mutation affects the stability of the decoding region and leads to ribosomes with more miscoding errors than ribosomes produced in the presence of YjeQ. Finally, we were able to visualize molecular interactions of the YjeQ N-Terminal Extension with the mature 30S ribosome subunit at high resolution. Our cryo-EM analysis revealed interaction of this YjeQ helix strongly resembles the C-terminal tail of YaeJ, a ribosome rescue factor. Thus, providing a preview of the potential new role of YjeQ in dissociating the 70S ribosome, possibly with ribosome rescue.

3.2 Introduction

The bacterial ribosome decodes genetic information for protein synthesis with an error rate of 1 in every ~10,000 amino acids incorporated during protein synthesis (Gallant and Lindsley, 1998) (Ogle and Ramakrishnan, 2005). Such high accuracy is solely not based on Watson-Crick base-pairing as the free energy difference between a cognate and near-cognate interaction is insignificant (≤3 kcal/mol) (Rodnina and Wintermeyer, 2001). Consequently, such a small energy difference would justify an error rate of 1 in every ~100 amino acids (Rodnina and Wintermeyer, 2001) (Moore and Steitz, 2011). Chemical foot-printing and crystallography studies later revealed that the high accuracy in recognizing the cognate mRNA: tRNA interaction is implemented by the 16S rRNA through nucleotides A1492, A1493, and G530 (Yoshizawa et al., 1999) (Wimberly et al., 2000) (Carter et al., 2000) (Ogle et al., 2001). These specific nucleotides flip out from their respective helices and monitor the geometry of interactions between the mRNA: tRNA base pairing. Interestingly, similar conformational changes have been observed in the 30S ribosome subunit upon binding of YjeQ in the GMP-PNP state (Razi et al., 2017b). This cryo-EM study suggested that YjeQ, a ribosome assembly factor, induces a translation-like conformational change in the decoding region before the 30S can associate with the 50S to engage in the translation process.

YjeQ (also known as Ribosome small-subunit-GTPase A) is a broadly conserved protein found in Gram-positive and Gram-negative bacterial species (Arigoni et al., 1998). Early sequence and structural studies showed that the YjeQ N-terminal domain comprises an uS1-like OB (oligonucleotide/oligosaccharide binding)-fold domain, followed by a circularly permuted GTPase domain, and a C-terminal zinc-finger domain (Daigle et al., 2002) (Shin et al., 2004). Extensive analysis from footprinting experiments suggests that YjeQ is nature's all-in-one quality control

package as its binding site on the 30S overlaps that of initiation factor 1 (IF1), initiation factor 3 (IF3), C-Terminal domain (Lopez-Alonso et al., 2017b). Although not essential for growth, *yjeQ* deletion impacts assembly and maturation of the small ribosome subunit (the 30S) (Jomaa et al., 2011a) (Jeganathan et al., 2015) (Leong et al., 2013) (Campbell et al., 2005) (Himeno et al., 2004). Early low-resolution cryo-EM analysis suggested that YjeQ may act as an RNA chaperone and bind to the 30S to assist in folding the decoding region (Jomaa et al., 2011a). However, binding characterization experiments revealed that YjeQ prefers binding to the mature 30S subunits over immature 30S particles ($30S_{\Delta yjeQ}$) accumulated in the *yjeQ* deletion strain (Thurlow et al., 2016). In the context of the mature 30S, medium-resolution cryo-EM studies have provided details of the interactions between YjeQ and the mature 30S ribosome (Razi et al., 2017b) (Lopez-Alonso et al., 2017b). While the cryo-EM studies from Lopez-Alonso *et al.* showed that YjeQ induced a flipping of A1493, the work from Razi et al. revealed that YjeQ induced conformational change in the A1492 the 16S rRNA helix 44. In the 'free 30S'subunits, these molecular switches (A1492 and A1493) are tucked inside the helix 44, where A1493-A1408 forms a non-canonical base pairing (Wimberly et al., 2000). These results suggested that YjeQ serves as a quality control protein by mimicking this flipping mechanism before the handover of the 30S to the translation initiation machinery (Razi et al., 2017b).

Therefore, to understand the role of YjeQ in the context of immature 30S ribosome subunits, we characterized the assembly intermediates from the 30S ribosome subunit that accumulate in the E. coli Δ*yjeQ* strain (Baba et al., 2006) using recent advances in the cryo-EM field. Subsequently, we asked which of the two adenines at 1492 and A1493 YjeQ flips out as part of the YjeQ checkpoint function. To this end, we investigated the effects of compromising this quality control mechanism by mutating the structural elements of YjeQ that seem to be involved in this checkpoint function.

The work from Razi et al. implicated that YjeQ OB-fold domain residue Phe48 in the β1/β2 hairpin loop induces the A1492 flip. Further analysis of this structure suggested that the flipped nucleotide is stabilized by another YjeQ OB-fold domain residue Arg68 in the β3/β4 hairpin loop. Therefore, we created two YjeQ variants: (1) YjeQ single mutant with Phe48Ala mutation; (2) YjeQ double mutant with Phe48Ala and Arg68Ala mutation. The YjeQ single mutant was designed to abolish the base-flip induced by Phe48. The YjeQ double mutant was designed to interfere with two aspects of YjeQ interaction with the 30S ribosome: abolish base-flipping (by Phe48) and destabilize flipped base. Using cryo-electron microscopy and in vitro fidelity assays, we show that ribosomes produced in the presence of these YjeQ variants have a decreased translational fidelity.

3.3 Materials and Methods

Cell Strains and Protein Overexpression Clones:

Parental *Escherichia coli* K-12 (BW25113) and Δ*yjeQ* (JW4122-3) strains were obtained from the Keio collection (Baba et al., 2006).

The pDEST-17-yjeQ plasmid (AmpR) (pJO241) for protein overexpression and generation of mutant pDEST-17-yjeQ constructs were obtained from the previously generated clones in the laboratory (Jomaa et al., 2011b). The plasmid pDEST-17-yjeQ-Single-Mutant expressing YjeQ Phe 48 Ala mutant was created using pDEST-17-yjeQ plasmid as a template for site-direct mutagenesis. The following primers were designed using the NEBaseChanger software. ACCORDING TO THE MANUFACTURER'S PROTOCOL, the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) was used to substitute nucleotides at positions 142-143 on the yjeQ gene.

Forward Primer: 5’ – CATCAGCCGCgcTGGTATGCAC – 3’

Reverse Primer: 5’ – ACGATACCTTCATCAGGC – 3’

Similarly, the plasmid pDEST-17-yjeQ-Double-Mutant expressing YjeQ ($Phe^{48}Ala$ $Arg^{68}Ala$) mutant was created using pDEST-17-yjeQ-Single-Mutant plasmid as a template for site-direct mutagenesis to substitute nucleotides at position 202-203 on the yjeQ gene.

Forward Primer: 5’ – CAATATTCGCgcTACCATCCGTTCGCTGGTAACCG – 3’

Reverse Primer: 5’ – CAGCGGTGAACGTCGCCA – 3’

The resulting constructs were validated by Sanger sequencing at Centre d’expertise et de services Génome Québec.

The glycerol stock for BL21DE3-TEV cells producing TEV protease to cleave the His_6 tag on the N-terminus of YjeQ wild type and mutant proteins was a gift from Dr. Alba Guarné's laboratory (McGill University).

For Miller's Assays, the kanamycin resistance cassette in the Δ*yjeQ* (*ΔrsgA787::kan*) (JW4122-3) strain was eliminated using the flanking FRT sites and the FLP helper plasmid (pCP20) (Datsenko and Wanner, 2000). The resulting kanamycin sensitive yjeQ deletion (*ΔrsgA787::scar*) strain was transformed with the pACYC series of chloramphenicol resistant plasmids used for Miller assays (O'Connor et al., 1997). To examine the effect of *yjeQ* mutants on translation fidelity, pDEST-17 plasmids carrying *yjeQ* constructs (wild-type and mutants) under the *araBAD* promoter were transformed into strains containing the fidelity plasmids. The cells were grown in the presence to maintain both plasmids.

Protein expression and Purification:

YjeQ double-mutant protein was purified exactly as YjeQ wild-type protein previously described by (Razi et al., 2017b). Before use in any experiments, purified protein was further purified by size-exclusion chromatography using a Superdex 200 Increase 10/300 GL column (GE Healthcare) equilibrated with 20mM Tris HCl pH 7.5, 10mM MgCl2, 200mM KCl, 3mM 2-mercaptoethanol for ideal reaction conditions.

Purification of $30S_{\Delta yjeQ}$ Particles:

$30S_{\Delta yjeQ}$ were purified and stored as previously described by (Jomaa et al., 2011a).

Miller's β-Galactosidase Assay:

Cells carrying the pACYC and pDEST-17 plasmids were grown in the presence of Ampicillin (100 μg/mL), Chrolamphenicol (30 μg/mL), and glucose (0.2%) or arabinose (0.2%) overnight. The next day, the cultures were diluted ($1/100^{th}$) in fresh LB with Ampicillin, Chloramphenicol, and

either Glucose or Arabinose (0.2%). The β-Galactosidase expression in each strain was calculated as Miller units (Miller, 1972). The substrate used in the Miller assay is ONPG (o-nitrophenyl-β-D-galactoside) which produces a yellow color in the presence of β-Galactosidase and is used to measure β-Galactosidase expression from different fidelity constructs (O'Connor et al., 1997). The protocol for the Miller assay was performed as described previously by (Zhang and Bremer, 1995). The miller units were normalized using the WT lacZ plasmid in the WT strain for the parental vs. knockout fidelity experiments. For the YjeQ expression experiments, the miller units were normalized using the WT lacZ plasmid in the Δ*yjeQ* strain.

30S +YjeQ Phe 48 Ala Complex Setup for Cryo-electron Microscopy:

The parental E. coli K-12 strain (BW25113) was used to purify the 70S ribosomes used in these experiments. Typically, 3 L of LB media were inoculated with 10 ml of saturated overnight culture and grown to an OD600 of 0.6. Cells were harvested by centrifugation at 3700 g for 15 min. Cell pellets were resuspended in 7 ml of buffer A (20 mM Tris-HCl at pH 7.5, 10 mM magnesium acetate, 100 mM NH4Cl, 0.5 mM EDTA, 3 mM 2-mercaptoethanol), and a protease inhibitor mixture (cOmplete Protease Inhibitor Mixture Tablets; Roche) and DnaseI (Roche). Each of the subsequent steps was performed at 4°C. The cell suspension was passed through a French pressure cell at 1400 kg/cm2 three consecutive times to lyse the cells. The lysate was spun at 59,000 g for 30 min to clear cell debris. The resulting supernatant was layered over a sucrose cushion of equal volume composed of 30% sucrose in buffer B (20 mM Tris-HCl, pH 7.5, 10 mM magnesium acetate, 500 mM NH4Cl, 0.5 mM EDTA, and 3 mM 2-mercaptoethanol), and then spun down for 4.5 h at 321 000 g. The pellet was resuspended in buffer C containing 10 mM Tris-HCl, pH 7.5, 10 mM magnesium acetate, 500 mM NH4Cl, 0.5 mM EDTA, and 3 mM 2-mercaptoethanol then spun for 16 h at 100 000 g. The washed ribosome pellet was resuspended in buffer E containing

10 mM Tris-HCl, pH 7.5, 10 mM magnesium acetate, 60 mM NH4Cl, and 3 mM 2-mercaptoethanol, which caused subunits to be associated. Approximately 120 A260 units of resuspended crude ribosomes were applied to 34 ml of 10–30% (wt/vol) sucrose gradients prepared with buffer E. The gradients were centrifuged for 16 h at 40 000 g on a Beckman Coulter SW32 Ti rotor. Gradients were fractionated using a Brandel fractionator apparatus and an AKTA Prime FPLC system (GE Healthcare). The elution profile was monitored by UV absorbance at A254, and fractions corresponding to the 70S subunit peak were pooled and spun down for another 4.5 h at 321 000 g on a Beckman SW32 Ti rotor. The pellet was resuspended in buffer E.

Prior approach to capture the 30S–YjeQ complex by incubating the two purified components together yielded 48% of 30S ribosome subunit particles bound to the YjeQ (Razi et al., 2017b). Therefore, to capture the 30S+YjeQ double-mutant complex with high occupancy, we adopted the experimental approach previously described by (Himeno et al., 2004) (Daigle and Brown, 2004) (Lopez-Alonso et al., 2017b). These studies demonstrated that in an Mg^{+2} abundant environment and access to GTP or GTP-like analogs, YjeQ could dissociate 70S ribosomes into 30S and 50S ribosomal subunits. To capture the 30S+YjeQ double-mutant complex, 1μM 70S ribosomes were incubated with 3μM of SEC-purified YjeQ double-mutant at 37°C for 15 minutes in presence of 1mM GMP-PNP, 37.5mM Tris HCl pH 7.5, 15mM MgCl2, 100mM NH4Cl, 28mM KCl and 6mM 2-mercaptoethanol. The reaction was loaded onto a 10–30% sucrose gradient and centrifuged at 15,800 rpm, 4°C for 16 hours on the SW-32 rotor. Gradients were fractionated using a Brandel fractionator apparatus and an AKTA Go FPLC system (Cytiva). Fractions containing the 30S subunits were selected based on the UV absorbance at A254 and were further analyzed in a 4-12% Bis-Tris Gel (Biorad) to confirm the presence of YjeQ double-mutant. Subsequently, fractions were spun down for another 20 hours at 321,000g on a Beckman 70Ti rotor. The supernatant

containing sucrose was discarded, and the pellets were washed with 700µL of the reaction buffer twice before resuspending the pellets in the same buffer. The concentration of the resuspended complex was measured at A260 to be 696.01 nM. Aliquots were flash-frozen in liquid nitrogen and stored at –80°C until use.

Cryo-electron Microscopy

30S+ YjeQ double-mutant complex: The sample vitrification was performed using Vitrobot Mark IV (Thermo Fisher Scientific Inc.) with its climate control chamber set at 25 ºC and 100% relative humidity. Cryo-EM grids (CF-1.2/1.3-3Cu-T) were prepared by evaporating a continuous layer of carbon (5–10 nm) to reduce exposure of the ribosomal particles to the air-water interface. Grids were then soaked with chloroform for 2 hours and treated with glow discharged in the air at 5 mA for 15 seconds right before sample application. A 300nM of 3.6 µL sample was applied to each grid and blotted once in the Vitrobot for 3 seconds and with a blot force +1 before plunge freezing in liquid ethane.

Automated data acquisition was performed using SerialEM software (Schorb et al., 2019) at FEMR-McGill using a Titan Krios microscope at 300 kV equipped with a Gatan BioQuantum K3 detector. The energy-filter slit was centered on the zero-loss peak, with the slit width set to 20 eV. Data acquisition was set up to perform focusing once per stage position on the carbon area at the center of 4 holes and recording 2 images per hole. Movies were collected with a total dose of 50 $e^{-}/Å^{2}$ divided over 30 frames acquired in 3.156 seconds of exposure at a magnification of 105,000x, producing images with a calibrated pixel size of 0.855 Å. The nominal defocus range used during data collection was between -1.00 to -2.50 µm at an interval of -0.25 µm.

$30S_{\Delta yjeQ}$: The sample vitrification was performed using Vitrobot Mark IV (Thermo Fisher Scientific Inc.) with its climate control chamber set at 25 ºC and 100% relative humidity. Cryo-

EM grids (CF-2/2-2Cu-T) were soaked in chloroform overnight and treated with glow discharged in the air at 5 mA for 15 seconds right before sample application. A 300nM of 3.6 μL sample was applied to each grid and blotted once in the Vitrobot for 3 seconds and with a blot force +1 before plunge freezing in liquid ethane.

Data acquisition was performed using EPU software at FEMR-McGill using a Titan Krios microscope at 300 kV equipped with a Falcon II direct electron detector (Thermo Fisher Scientific). Movies were collected with a total dose of 49 $e^{-}/Å^2$ divided over 7 frames acquired in 1.0 second exposure at a magnification of 75,000x, producing images with a calibrated pixel size of 1.073 Å. The nominal defocus range used to collect all these data sets was −1.25 to −2.75 μm.

Image Processing:

30S+YjeQ F48A R68A complex: Image processing was performed with cryoSPARC v3.3.1 **(Supplementary Figure S3.1)** (Punjani et al., 2017). Imported corrected for stage drift, beam-induced motion, and frame alignment using Patch Motion Correction. Patch CTF estimation estimated the contrast Transfer Function (CTF) parameters for these micrographs. Using the Manually Curate Exposures job, micrographs were manually curated via motion trajectories, CTF fit values, scores, etc., using the Manually Curate Exposures job. Best micrographs from the previous step were selected and were grouped according to their beam-tilt positions for future CTF parameter refinements. 1,750,698 particles were picked from 7,039 micrographs subjected to two rounds of 2D classification prior, resulting in the final set of 1,182,653 particles after selecting good classes. These particles were subjected to beam-induced motion correction at the individual micrograph level and re-extracted at 384px. These particles were downsampled to 128px and subjected to a homogenous refinement, generating a 5.34Å consensus cryo-EM map for 3D classification using 3D Variability. Accordingly, particles from the downsampled consensus

refinement were subjected to 3 principal modes of eigenvectors. The results of this 3D classification were analyzed using 3D Variability Analysis by specifying the program to plot the particle groups in 15 cluster modes. Results from this first round of 3D Variability-based classification were analyzed using UCSF Chimera v1.15 (Pettersen et al., 2004). 940,536 particles showing YjeQ bound to the 30S ribosome subunit decoding region were subjected to the second round of 3D Variability-based classification. While the remaining 241,841 particles with no YjeQ bound were also subjected to another round of 3D Variability-based classification. These classifications yielded 791,453 particles with 'unfragmented' cryo-EM density representing YjeQ at the 30S ribosome subunit decoding region. Next, we isolated 149,083 particles with 'fragmented' cryo-EM density representing YjeQ at the 30S ribosome subunit decoding region. In the case of the 30S ribosome subunit particles that didn't have the cryo-EM density for YjeQ, 86,059 particles represented the 30S ribosome subunit in the previously defined 'inactive' state. In comparison, the rest 155,782 particles represented the 'active' state. The particles from each identified class were subjected to Non-Uniform Refinement with jobs settings activated for Local and Global CTF Refinements to estimate per-particle defocus and correct the higher-order aberrations introduced by the microscope. These refinements yielded 2.68 Å resolution for the 30S+YjeQ$_{\text{Phe48Ala Arg68Ala}}$ complex, 3.11 Å resolution for the Inactive 30S, 2.99 Å for the Active 30S, and 3.25 Å resolution for the 30S with fragment YjeQ. Resolutions for the cryo-EM volumes from these high-resolution refinements were obtained with independently refined half-maps using Fourier shell correlation (FSC) curves at the 0.143 criterion (Scheres and Chen, 2012).

30S$_{\Delta yjeQ}$: Motion correction and contrast transfer function (CTF) estimation for each collected movie were done with MotionCor2 and Gctf programs. 1679813 particle stacks were selected and extracted from the micrographs using auto-picking. These processing steps were done with

RELION 2.1. To produce a 'clean' dataset, extracted particles were subjected to two rounds of reference-free 2D classification in cryoSPARC v3.3.1. These 1,016,541 particles were subjected to a homogenous refinement which generated a 3.11 Å consensus cryo-EM map for 3D classification using cryoDRGN v0.3.2. The poses and CTF parameters of the particle stacks from the consensus refinement were extracted using the *cryodrgn parse_pose_csparc* and *cryodrgn parse_ctf_csparc* commands, respectively. Particle images for training neural networks were Fourier cropped (downsampled) using the *cryodrgn preprocess* command. The low-resolution training particles were downsampled to 72 pixels (4.530 Å/pix), while the particles for high-resolution training were downsampled to 152 pixels (2.146 Å/pix). The initial low-resolution training was performed using the *cryodrgn train_vae* command using an 8-dimensional latent variable and 256x3 network architecture encoder and decoder for 50 epochs. These trained particles were filtered to exclude junk particles using the z-score filtering in *cryoDRGN_filtering.ipynb* Jupyter Notebook. 976,253 particle stacks resulting from this clean-up were then subjected to low-resolution training using a 10-dimensional latent variable and 1024x3 neural network architecture encoder and decoder for 75 epochs.

Map analysis and Model Building:

Before building the molecular models for each structure, the connectivity of the Coulomb potential densities of the cryo-EM refinement maps was improved using DeepEMhancer (Sanchez-Garcia et al., 2018) available on the COSMIC² cloud platform. The resulting .mrc map was then converted to .ccp4 using the Autosharpen map tool from Phenix suite v1.20-4459 (Liebschner et al., 2019); no sharpening was applied during this conversion. Model building of all maps started by fitting the 5UZ4.pdb (Razi et al., 2017b), 7NAR.pdb and 7BOI.pdb (Schedlbauer et al., 2021). The molecular models were then built through multiple rounds of manual model building in Coot

v0.9.6 (Emsley and Cowtan, 2004) (Emsley et al., 2010) from the ccp4-7.1 suite (Winn et al., 2011), and real-space refinement (Afonine et al., 2018b).

3.4 Results

Pre-mature folding of helix 44 rescues 30S assembly in the absence of YjeQ.

To understand why YjeQ cannot recognize immature 30S ribosome subunits accumulated in the *E. coli* Δ*yjeQ* strain (Baba et al., 2006), we analyzed over 976,253 particle images from our $30S_{\Delta yjeQ}$ dataset using cryoDRGN (Zhong et al., 2021). Our cryoDRGN analysis of the 20 classes revealed four major groups **(Figure 3.1)**. Group A was composed of classes 0, 17, 18, and 19, which contributed to 20.58% and represented the most immature 30S particles that accumulate. While the cryo-EM density for the head and the platform domains was completely absent, the body domain represented severe unfolding of the 16S rRNA helices that make the core of the 30S. Group B represented classes 2, 3, 5, 6, 7, 8, and 16, contributing to 31.74% of the particles. This group was characterized by the lack of cryo-EM densities for the 30S head, variable occupancy of the platform region, absence of helix 44, and mature body domain marked by the completely folded helix 27.

The most striking observation came from Group C, as classes 4, 9, 10, 11, 12, 13, 14, and 15 made up 40.9% of the particles. This group was characterized by pre-mature folding of 16S rRNA helix 44 and the absence of the head and the platform domains. This group had two sub-populations based on the folding or unfolding of helix 27. Accordingly, Group C1, composed of classes 11, 12, 13, 14, and 15, had demonstrated completely unfolded helix 27 and helix 44 in a displaced state, falling away from the body domain. Group C2, composed of classes 4, 9, and 10, demonstrated helix 44 docked into the body domain and complete folding of helix 27 and a mature platform region compared to Group C1.

Only 6.78% of the particles contributed to Group D, composed of class 1 representing mature-like 30S but lacked cryo-EM density for r-proteins uS2, uS7, uS13, and bS21. These observations

suggest that the maturation of helix 27 is a prerequisite for folding and stability of the 30S head domain since helix 27 continues to form helix 28, which is the only connection between the head and the rest of the 30S (Mohan et al., 2014). As a result, the appropriate maturation of helix 27 is promoted by the pre-mature folding of helix 44, which may further trigger the folding of the 30S platform and the 30S head domains.

Previous quantitative mass spectrometry analysis (qMS) and iTRAQ data show that the purified $30S_{\Delta yjeQ}$ particles only lack r-proteins uS2 and bS21 (Jomaa et al., 2011a) (Thurlow et al., 2016). This suggests that most of the r-proteins are bound to their respective 16S rRNA helices but have not been able to stabilize the 30S domains in their respective positions. Combined with our cryo-EM analysis, these observations demonstrate that: (1) although YjeQ is not-essential, it catalyzes the folding of the critical 16S rRNA helix in the 30S body domain in a timely manner; (2) ensures the folding of 16S rRNA domains follow the canonical order of assembly: body, platform, head, and helix44.

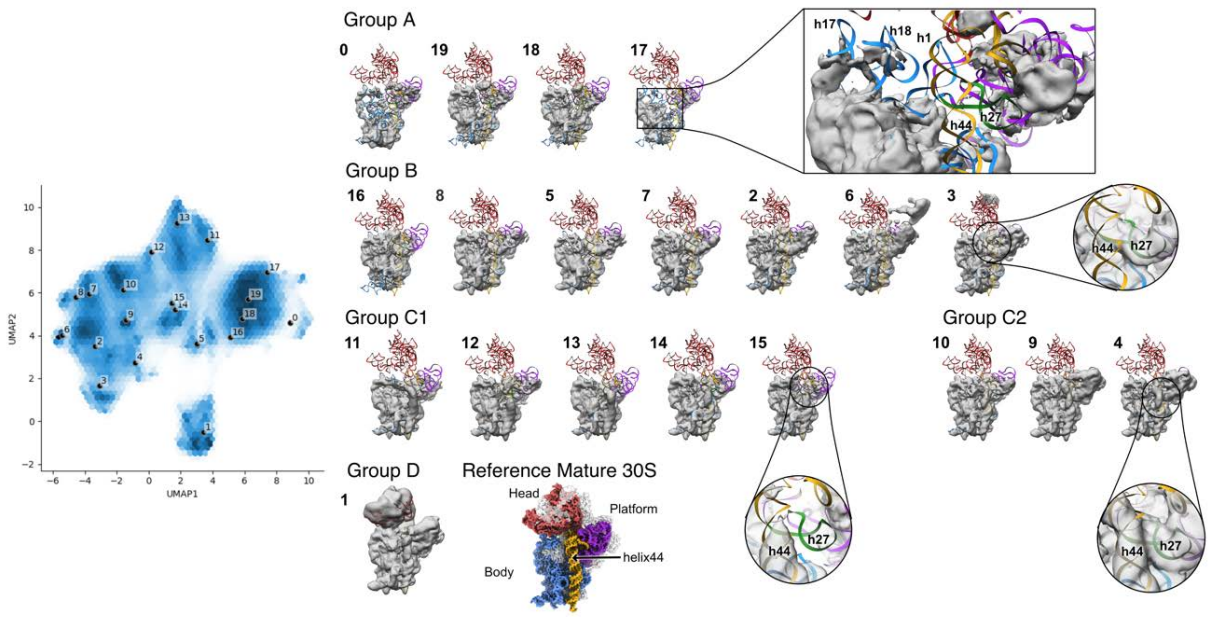


Figure 3. 1 Various 16S rRNA helices remain unfolded in the 30SΔyjeQ particles.

cryoDRGN analysis revealed four major groups that can be identified by docking the 16S rRNA helices from the crystal structure of the 30S ribosome subunit (PDB ID: 2AVY). The 3D classes obtained by kmeans clustering from cryoDRGN analysis are oriented to show the interface view of the 30S ribosome subunit. Each point in the UMAP embedding shows where the decoder created each 3D volume. The domains of the 30S ribosome subunit are colored as follows: body domain (rRNA residues 1- 561 colored in blue); platform domain (rRNA residues in 562- 916 colored in purple); head domain (rRNA residues in 562- 916 colored in purple 917 – 1396 colored in red), helix 44 and 45 (1397-1534) colored in golden and helix 27 (887 -917) colored in green. Zoom-in view of the core of the 30S ribosome subunit (body domain) for a representative class within each group focuses on the occupancy of helix 44 and helix 27. For comparison with the 30SΔyjeQ classes, the 30S atomic model was converted to volume using the e2pdb2mrc.py script from EMAN2; the 16S rRNA helices were colored as mentioned above.

The binding of the YjeQ double mutant has a disruptive effect on the 30S decoding region.

To understand the role of YjeQ as a checkpoint protein, we utilized the ability of YjeQ to dissociate wild-type 70S ribosomes into 50S subunits and bind to the 30S subunit in the presence GMP-PNP and a magnesium-abundant environment. Our sucrose density gradient experiments showed that alanine substitution mutations at YjeQ residues 48 and 68 did not alter this YjeQ ability **(Supplementary Figure S3.2)**. This allowed us to capture the YjeQ double mutant bound 30S subunit complex for structural analysis using cryo-EM. The cryo-EM map for this complex was refined to a resolution of 2.68 Å. This allowed us to visualize interactions of YjeQ with the 30S

ribosome subunit with a local resolution range of 2.3 Å - 3.1 Å **(Supplementary Figure S3.3)** and build a molecular model with high confidence.

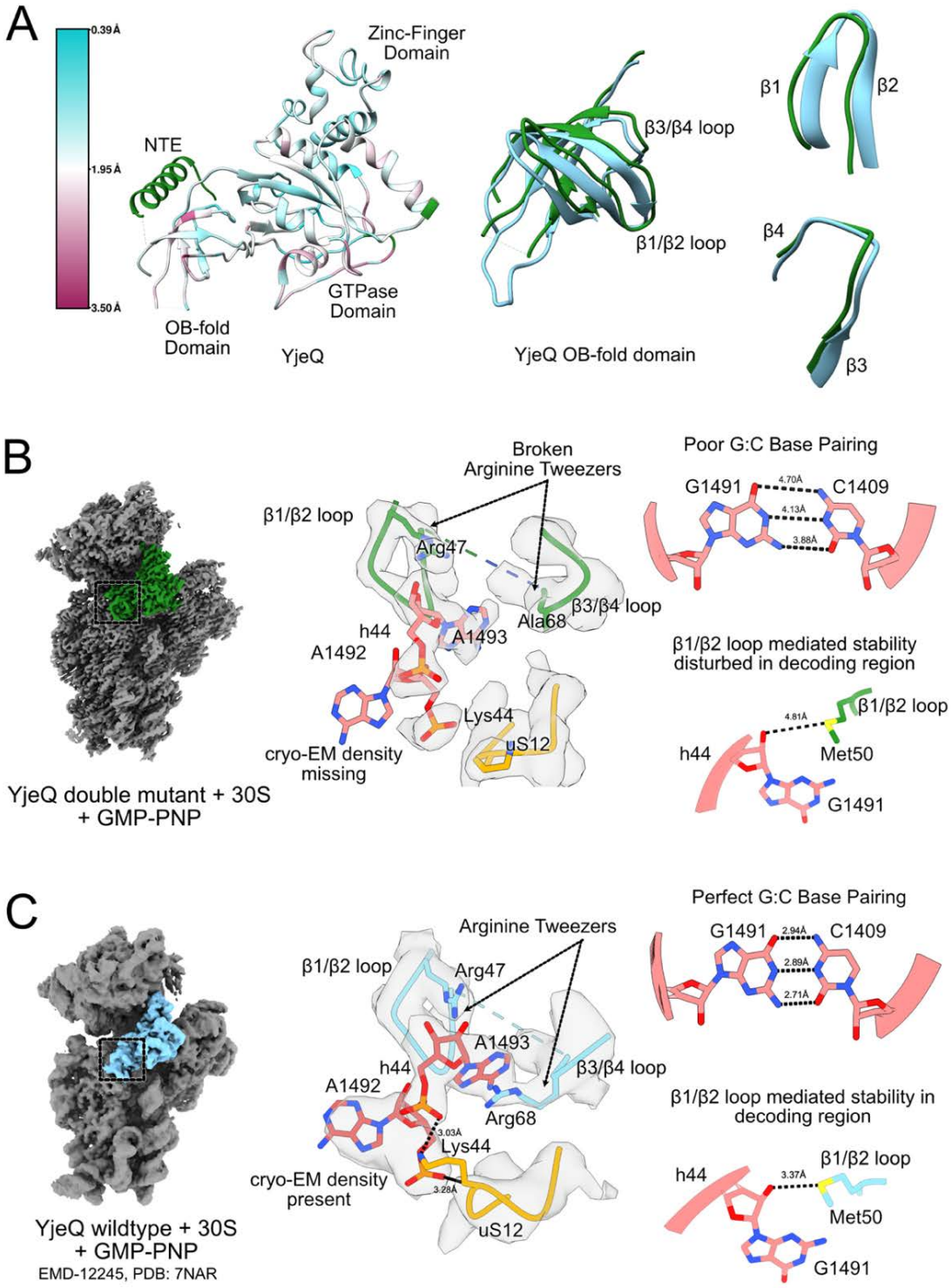


Figure 3. 2 Cryo-EM structure and analysis of YjeQ double mutant bound to wildtype mature

30S ribosome subunit in the presence of GMP-PNP. (A) The temperature map of the YjeQ double mutant was colored according to the r.m.s.d. deviation (Å) with respect to the structure of the wildtype YjeQ obtained by cryo-EM (PDB ID: 7NAR). r.m.s.d. deviations for YjeQ were calculated by aligning 16Ss rRNA helices. The relative position of YjeQ OB-fold domains and specific beta-hairpin loops containing residues of interest for this study. YjeQ double mutant structure is colored green, and the YjeQ wildtype structure is colored blue. (B) Cryo-EM structure of 30S+YjeQ DM+GMP-PNP with YjeQ segmented in green. Zoom view shows the effect of YjeQ double mutant with the decoding region of the 30S ribosome subunit. Helix 44 nucleotides are colored in salmon pink. YjeQ OB-fold domain loops are colored in green, and the uS12 loop is colored in golden. (C) For comparative purposes, the cryo-EM structure of 30S+YjeQ YjeQ +GMP-PNP with YjeQ is segmented in blue. Zoom view shows the interaction of YjeQ wildtype with the decoding region of the 30S ribosome subunit. Other components in the YjeQ OB-fold domain and 30S decoding region zoom-in view are colored similarly as in panel (B).

Comparing our YjeQ double mutant structure with the wildtype YjeQ cryo-EM structure revealed that overall YjeQ conformation does not dramatically change upon binding to the 30S ribosome subunit. This comparison produced an r.m.s.d. value of ∼1.65 in the GMP-PNP state **(Figure 3.2A)**. This difference clearly reflects in the binding of the OB-fold domain, which hold the key to its interaction with the 30S ribosome subunit and probing of the nucleotides in the decoding region. . We also note that the moderate resolution cryo-EM study from Razi et al. 2017, implicated that A1492 flips outside the helix 44 when YjeQ in the GMP-PNP state binds to 30S. However, our higher resolution cryo-EM map shows that A1493 is flipped during this interaction and is consistent with the observations from Schedlbauer et al. Further structural analysis revealed that cryo-EM density corresponds to helix44 nucleotide A1492 was absent when YjeQ double mutant was bound to the 30S **(Figure 3.2B).** However, in the wildtype YjeQ bound 30S complex from Schedlbauer et al., A1492 had a well-defined cryo-EM density **(Figure 3.2C)**. In the wildtype-YjeQ, the OB-fold domain β1/β2-hairpin loop that causes the A1493 to flip out is composed of residues Phe48, Gly49, Met50, and His51. Further, A1493 is held in place by the OB-fold domain residues Arg47 and Arg68 in the flipped-out position due to the cation-π interaction of these 'Arginine-tweezers.' Accordingly, the Arg68Ala mutation in this study results in broken

'Arginine-tweezers' and contributes to the instability of A1492 **(Figure 3.2B)**. These YjeQ mutations also affect other nucleotides in the decoding region. Helix 44 nucleotides G1491 and C1409 commonly form a perfect GC base pair when wildtype YjeQ interacts with the 30S **(Figure 3.2C)**. However, in our structure, the identical nucleotides have weaker base pairing **(Figure 3.2B)** and can be attributed to the far-positioning of Met50 in the YjeQ OB-fold domain β1/β2-hairpin loop. Typically, Met50 forms a hydrogen bond with the ribose sugar group of G1491 **(Figure 3.2C)**. Similar observations have been reported in the case of mt-SAF39 (mt-IF-3 like assembly factor) interaction with the mitochondrial small ribosome subunit (Lenarcic et al., 2022). Additional observations in the 30S subunit decoding region show that binding of wild type YjeQ does not affect the uS12 residue Lys44 interaction with the backbone of A1492 **(Figure 3.2C)**. However, the binding of the YjeQ double mutant to the 30S disrupts the same hydrogen bond **(Figure 3.2B)**. These observations show that mutations at positions 48 and 68 in the YjeQ OB-fold hinder the ability of YjeQ to probe critical nucleotides in the 30S ribosome subunit decoding region.

YjeQ double mutant treated free 30S ribosomes demonstrate distortion in the 30S decoding region.

Extensive 3D classification of the 30S+YjeQ mutant dataset allowed us to separate YjeQ free 30S ribosome subunit particles. Previous biochemical and structural studies have shown that free 30S ribosome subunits exist in 'active' and 'inactive' states that depend upon the conformation of helix 44 and helix 45 (Moazed et al., 1986) (McGinnis et al., 2015) (Jahagirdar et al., 2020). In this study, we separated 155,782 particles representing the 'active' state, accounting for about 13% of the particle population. In our control experiment, the sucrose density gradient profile of wildtype 70S ribosome purification in an abundant magnesium (15mM Mg^{2+}) environment resulted only in

2% of the free 30S population **(Supplementary Figure S3.2C)**. These observations suggest that free 30S ribosome particles identified in our 30S+YjeQ double mutant complex dataset are YjeQ treated.

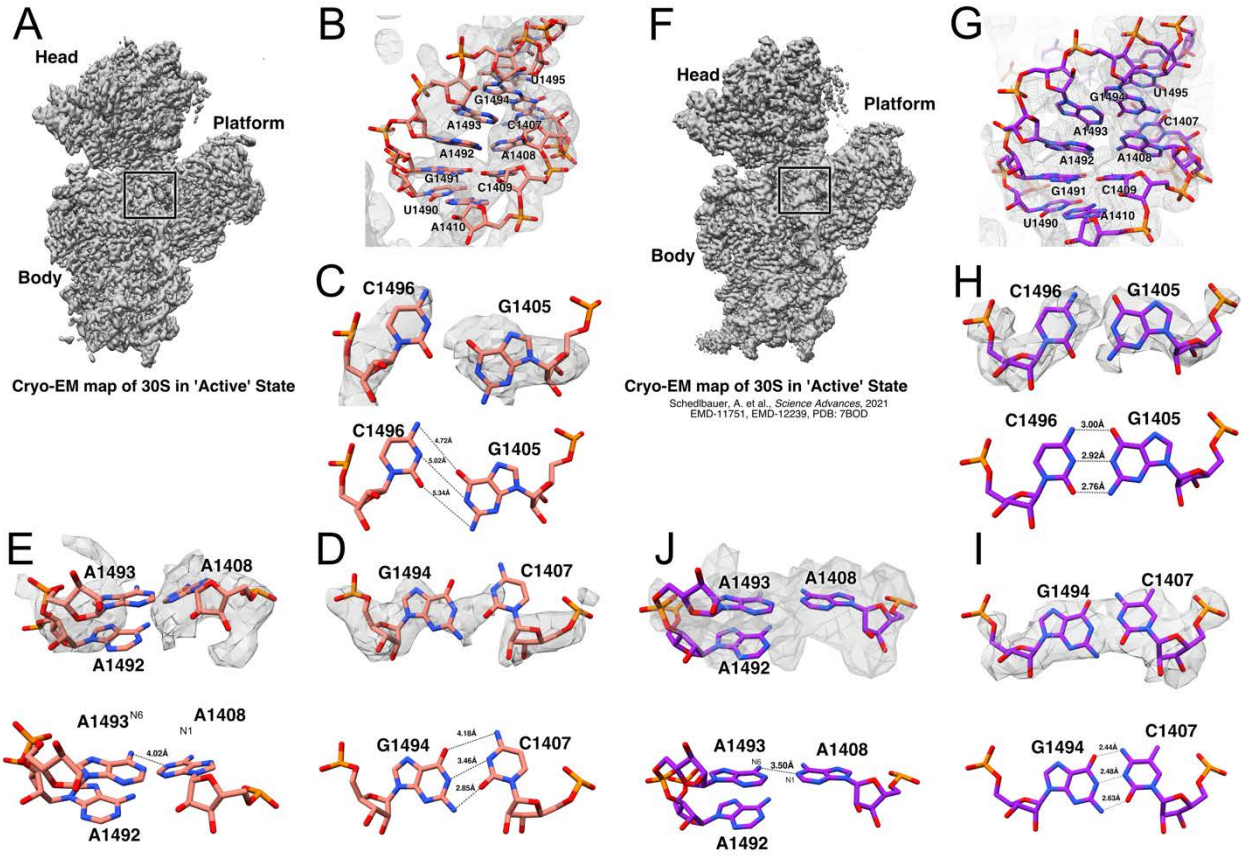


Figure 3. 3 Cryo-EM structure of YjeQ double mutant treated the 30S in the 'active' state.

(A) Cryo-EM map of 'active 30S' oriented in the interface view compared to a cryo-EM map of the 30S obtained in a different study (EMDB: 11751). (B) The atomic model of the YjeQ treated the 30S decoding region with a cryo-EM map represented as the mesh showing minor distortion in the helix 44 backbone; panel (G) shows the helix 44 decoding region without any fragmentation. The atomic model from this study is colored as coral color, while the atomic model with PDB ID: 7BOD is colored purple. Panels (E) and (J) show contrast distortion and normal positioning at helix 44 residues A1492, A1493, and A1408. (C) The 30S decoding region residues C1496 and G1405 fail to form Watson-Crick base pairing compared to the identical residues as shown in (H). Panels (D) and (I) show distortions in base pairing residues C1496 and G1405 and perfect base pairing of the identical residues in the reference 30S cryo-EM structure.

Molecular model-building from this cryo-EM map **(Figure 3.3A)** revealed a disorder in the cryo-EM density for the backbone of nucleotides A1491-C1494 **(Figure 3.3B)**. These nucleotides are

part of the helix 44 minor groove that continues to form helix 45. Crystallography and cryo-EM structures of the 30S ribosome subunit have shown that A1492 and A1493 are usually tucked inside helix 44, and A1493 forms non-canonical base-pairing A1408 through N6 of A1493 and N1 of A1492 **(Figure 3.3J)**. However, in our 'active' 30S structure, adenine nucleotides at positions 1408, 1492, and 1493 were disordered and did not form the hydrogen bond between N6 of A1493 and N1 of A1492 **(Figure 3.3E)**. Further analysis of the decoding region revealed that bases 1496 and 1405 that form CG base pairs in a typical active 30S **(Figure 3.3H)** did not form Watson-Crick base pairs in our structure **(Figure 3.3C)**. Similarly, bases 1494 and 1407 also form a perfect GC base-pairing in a typical 'active 30S' ribosome subunit **(Figure 3.3I)**. However, the cryo-EM analysis of this study revealed broken cryo-EM density for C1407 causing the GC base-pair to have weaker interactions **(Figure 3.3D)**. This structural analysis indicates that the interaction of the YjeQ double mutant with the free 30S ribosome subunits induces subtle distortions in the helix 44 that are localized to the decoding region.

Translation defects are comparable in strains expressing *yjeQ* mutants and *yjeQ* knockout strains.

Observations from our cryo-EM data prompted us to investigate whether these mutations in YjeQ influence translational fidelity *in vivo*. Previous pulse-chase experiments and in vitro maturation have shown that immature 30S particles that accumulate in the Δ*yjeQ* strain ($30S_{\Delta yjeQ}$) are not dead-end products. Instead, these $30S_{\Delta yjeQ}$ are capable of maturation and assembling into 70S ribosomes (Thurlow et al., 2016). Consistent with these experiments, previous studies have observed a large proportion of the 70S ribosomes in the sucrose density gradient profiles (Leong et al., 2013) (Thurlow et al., 2016). However, whether these assembled $70S_{\Delta yjeQ}$ ribosomes are

fully functional or whether they have comparable fidelity to ribosomes assembled in the presence of a complete set of assembly factors is unknown.

To identify translational fidelity defects specific to initiation, elongation, and termination, the β-galactosidase reporter gene has been designed with an unusual start, stop, or frameshift codons at the open reading frame (O'Connor et al., 1997). Similar experiments aiming to understand the effect of *yjeQ* deletion on translation initiation have reported minor defects in the translation initiation mechanism (Sharma and Anand, 2019). Therefore, in this study, we assessed the ability of the Δ*yjeQ* strain to detect the premature presence of +1 frameshift and -1 frameshift (elongation) and stop codons (termination). Accordingly, our β-Galactosidase assays revealed that the frameshifting levels in the Δ*yjeQ* strain were elevated compared to the wild-type strain **(Figure 3.4A)**. However, these assays revealed that the Δ*yjeQ* strain retained the ability to detect stop-codon. Since our experiments clearly show that *yjeQ* deletion mainly compromises translation elongation, we investigated the effects of expressing YjeQ single mutant and YjeQ double mutants in the Δ*yjeQ* strain. These follow-up β-Galactosidase assays revealed elevated frameshifting errors consistent with the Δ*yjeQ* strain **(Figure 3.4B)**. Accordingly, the Δ*yjeQ* strain expressing YjeQ single mutant fails to detect +1 frameshift, while the YjeQ double mutant still retains the ability to see this defect. Analysis of the Miller's assay to understand the effects of yjeQ mutations on detecting -1 frameshift revealed compromised fidelity in strains expression yjeQ variants. This proves that YjeQ residues Phe48 and Arg68 ensure that the decoding region is accurately folded. Previous β-galactosidase experiments to understand the effect of mutations in the 30S ribosome subunit revealed that elongation-related defects result from mutations in the 530 loop of helix 18 and helix 34 in the head (O'Connor et al., 1997).

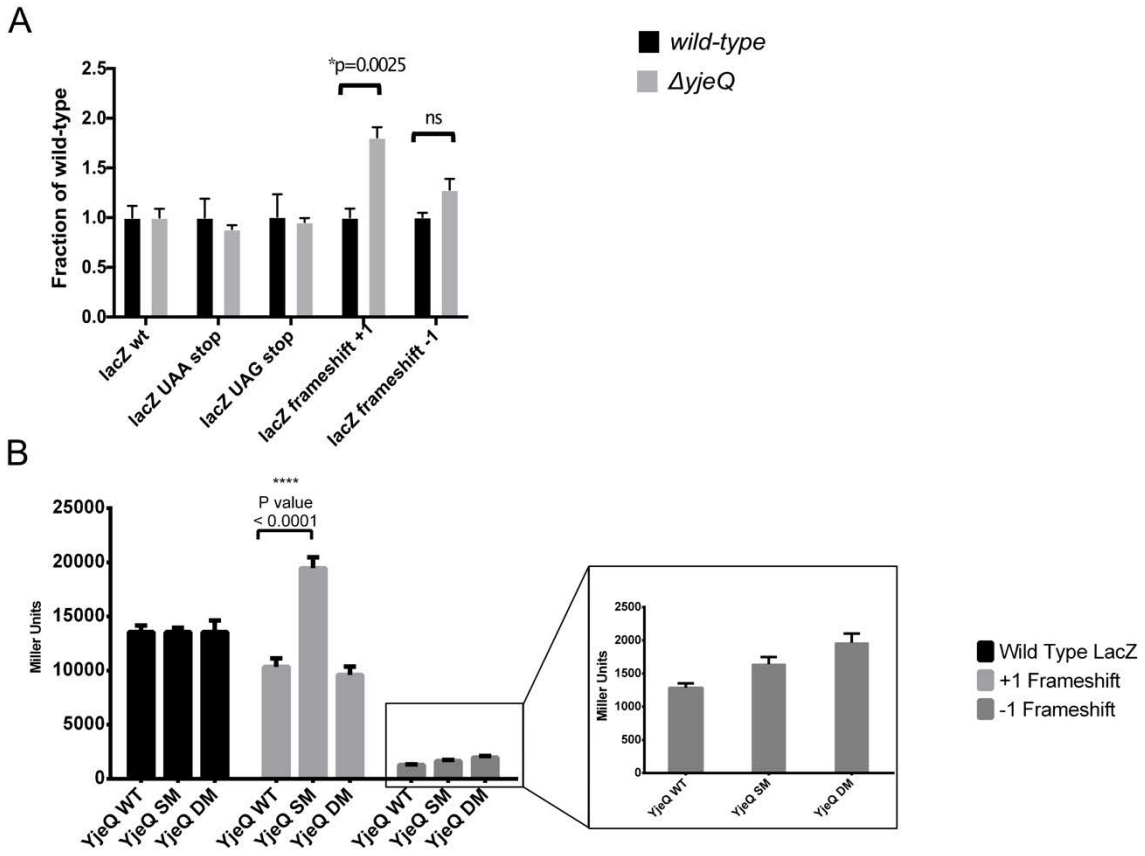


Figure 3. 4 (A) Effect of deletion of yjeQ on frameshifting errors and stop-codon readthrough in vivo.

Fidelity of translation at termination (premature UAG and UAA stop codon) and elongation (−1 frameshift and +1 frameshift) stages were assayed in wildtype E. coli and ΔyjeQ strains. (B) Effect of overexpression of YjeQ wildtype and mutants in the ΔyjeQ strain on frameshifting errors in vivo. The β-galactosidase units (obtained with the mutant lacZ reporters) are normalized to the β-galactosidase units (obtained with WT lacZ) in the ΔyjeQ strain. Error bars represent SD.

Characterization of YjeQ N-Terminal Extension reveals a possible role of YjeQ as a rescue factor.

Bioinformatic studies have shown that YjeQ has an additional weakly conserved N-terminal extension (NTE) that precedes the OB-fold domain (Nichols et al., 2007). Previous biochemical experiments have shown that deletion of the first 20 amino acids from the YjeQ NTE does not affect the YjeQ binding ability to the 30S. However, this deletion slightly reduces GTPase activity compared to the wildtype YjeQ (Daigle and Brown, 2004). Although previous cryo-EM studies

have described that this extension of YjeQ interacts with helix 18 and helix 34, we have only visualized the interaction of these components in moderate resolution (Razi et al., 2017b) (Lopez-Alonso et al., 2017b). While the helix 44 decoding region interacts with the first two codon-anticodon positions, helix 18 π- stacking nucleotides C518 and G530 and helix 34 base C1054 are involved in stabilizing the third codon, which forms the 'wobble' base pair with its respective anticodon. This suggests that YjeQ may also be probing the nucleotides involved in recognizing the wobble base pair through its NTE.

The high resolution for the YjeQ NTE in our 3D reconstruction allowed us to build a molecular model and analyze its interaction with the 30S head and the body in detail **(Figure 3.5A)**. In our structure, the NTE forms an alpha helix that can be visualized from residues 4-25 and is mainly composed of positively charged (Arg, Lys) and neutral polar (Asn, Ser) amino acids making it highly basic with a pI of 12.48. The first visible residue in our structure's NTE Asn4 forms van der Waals interactions with the nucleotide C1397, suggesting that the first few NTE amino acids probe the first nucleotides on the helix 44 **(Figure 3.5C)**. As the NTE continues, Leu6 and Gly10 form strong hydrogen bonds with O6 and N2 atoms of G530, respectively **(Figure 3.5D)**. Next, Gln10 and Arg13 interact with the phosphate backbone of U531 on helix 18 **(Figure 3.5E)**. In the case of NTE interactions with the helix 34 in the 30S head region, the side chain atoms on Lys8 form stable interactions with G1053 on helix 34. The neighboring C1054 is contacted by Lys8 through its main chain atoms to interact with the base O2 atom, while the side-chain atoms on Arg13 form hydrogen bonds with the sugar group **(Figure 3.5B)**.

Interestingly, YaeJ, a ribosome rescue factor, has comparable interactions with the small ribosome subunit. YaeJ works by hydrolyzing peptidyl-tRNA on ribosomes stalled by mRNAs without stop codon, mRNAs containing rare codons clusters, or ribosomes stalled abruptly in the middle of

mRNA (Chadani et al., 2011) (Shimizu, 2012). There are structural, compositional, and binding sites with striking similarities between YjeQ NTE and YaeJ C-terminal tail **(Figure 3.5F)**. Previously reported crystal structure of YaeJ bound to *T. thermophilus* 70S ribosome shows that the YaeJ C-terminal tail occupies the mRNA entry channel in the 30S A-site (Gagnon et al., 2012). The C-terminal positively charged alpha helix interacts with the 16S rRNA helices 18, 34, and 44, which allows it to recognize stalled ribosomes. This YaeJ functionality is comparable to YjeQ since, in GMPPNP or GTP bound state, YjeQ can dissociate 70S ribosomes with empty A-site. Consequently, YjeQ may potentially act as a rescue factor in mature ribosomes. Alternatively, during the late-stage ribosome assembly, YjeQ, through its NTE, probes if the binding site for rescue factors is appropriately folded.

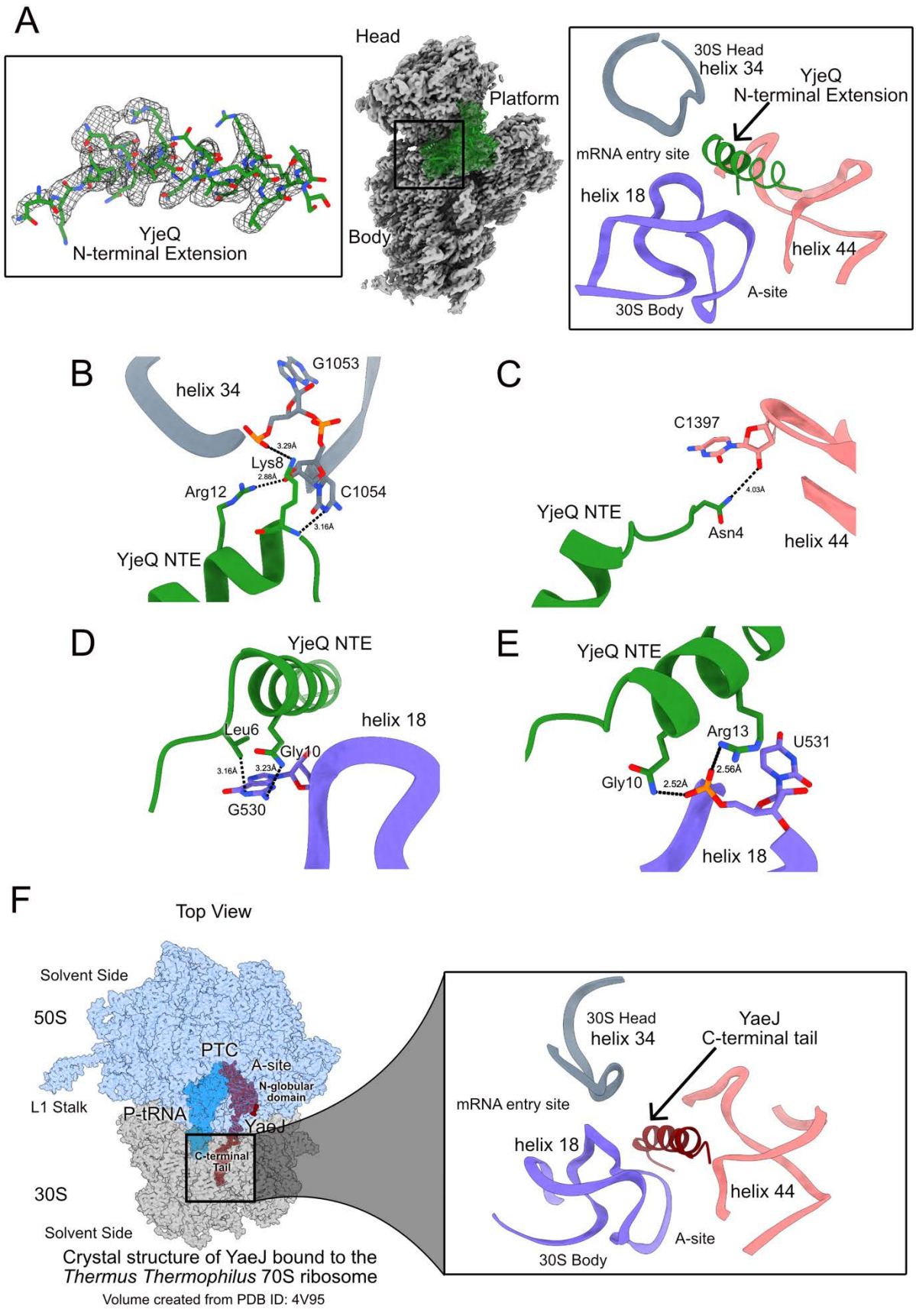

A
YjeQ
N-terminal Extension
Head
Platform
Body
30S Head
helix 34
YjeQ
N-terminal Extension
mRNA entry site
helix 18
helix 44
A-site
30S Body
B
G1053
helix 34
3.29Å
Lys8
2.88Å
Arg12
C1054
3.16Å
YjeQ NTE
C
C1397
4.03Å
YjeQ NTE
Asn4
helix 44
D
YjeQ NTE
helix 18
Leu6
Gly10
3.16Å
3.23Å
G530
E
YjeQ NTE
Arg13
U531
2.56Å
Gly10
2.52Å
helix 18
F
Top View
Solvent Side
50S
PTC
A-site
N-globular domain
L1 Stalk
P-tRNA
YaeJ
C-terminal Tail
30S
Solvent Side
30S Head
helix 34
YaeJ
C-terminal tail
mRNA entry site
helix 18
helix 44
A-site
30S Body
Crystal structure of YaeJ bound to the
Thermus Thermophilus 70S ribosome
Volume created from PDB ID: 4V95

Figure 3. 5 Cryo-EM characterization of YjeQ N-terminal extension.

(A) Atomic model of the YjeQ N-terminal residues 4-25 depicted in green inside the cryo-EM mesh. This short helix is positioned at the 30S A-site approaching the mRNA entry channel between 16S rRNA helices 18, 34, and 44. (B) Structural analysis shows the interaction of residues Lys8 and Arg12 with helix 34 nucleotides G1053 and C1054. (C) First visible residue Asn4 approaching the beginning of helix 44 at nucleotide C1397. (D and E) YjeQ N-terminal extension probing helix 18 residues G530 (responsible for monitoring wobble base-pairing during mRNA decoding) and the backbone of U531. (F) The binding site of the C-terminal tail of the YaeJ, a ribosome rescue factor, is like the YjeQ NTE. T. thermophilus 70S ribosome subunit shown via the top view and sliced until 30S decoding and the 50S peptidyl transferase center are visible. YaeJ is colored dark red. Volume was created using PDB ID: 4V95. Zoom-in view of the YaeJ C-terminal tail shows identical positioning of the C-terminal tail approaching the 16S rRNA helices 18, 34, and 44.

3.5 Discussion

Structural implications of *yjeQ* deletion in the maturation of the 30S ribosome subunit

In the context of ribosome assembly, we explore how the 30S ribosome subunit assembly proceeds in the absence of YjeQ. The 3D classification maps obtained from the cryo-DRGN analysis of the $30S_{\Delta yjeQ}$ revealed that in the majority of the 30S classes, the head and the platform region were unfolded **(Figure 3.1)**. In addition, we observed that helix 44 was partially or entirely folded even in the absence of head/platform domains. A previous study from our laboratory focused on understanding the role of Era GTPase has shown that Era is essential for the folding of the 30S platform region (Razi et al., 2019). However, in the depletion state, the 30S ribosome assembly intermediates continue to fold and mature the body and the head domains. This is consistent with the previous observation that 30S ribosome assembly can proceed through multiple parallel pathways (Mulder et al., 2010) (Woodson, 2011). However, no studies have reported even the slightest folding of helix 44 before the maturation of the head or the platform region. Genetic interaction screening experiments show that overexpression of Initiation-Factor 2 (IF-2) and Era GTPase (Era) in the Δ*yjeQ* strain may rescue growth defects in bacteria due to the *yjeQ* deletion (Campbell and Brown, 2008).

Consequently, IF-2 and Era may be involved in the premature folding of helix 44. These observations led us to propose a model for the maturation of the 30S in the absence of YjeQ **(Figure 3.6)**. Accordingly, Group 1 particles that resemble highly immature 30S may proceed to the assembly via at least two parallel pathways. In the first pathway, helix 44 folds pre-maturely, which triggers the folding of helix 27 in the body of the 30S and, as a result, promotes the folding of helices in the platform and the head domain as represented by Group3A and Group 3B classes. In the second pathway, Group 1 particles seem to proceed with maturation via the ‘canonical’

pathway by folding the body first, the platform second, and the head in the end. Observations from our cryoDRGN dataset may explain previous pulse-chase labeling experiments demonstrating that the assembly of the 30S in the absence of YjeQ proceeds to maturation, although slowly (Thurlow et al., 2016).

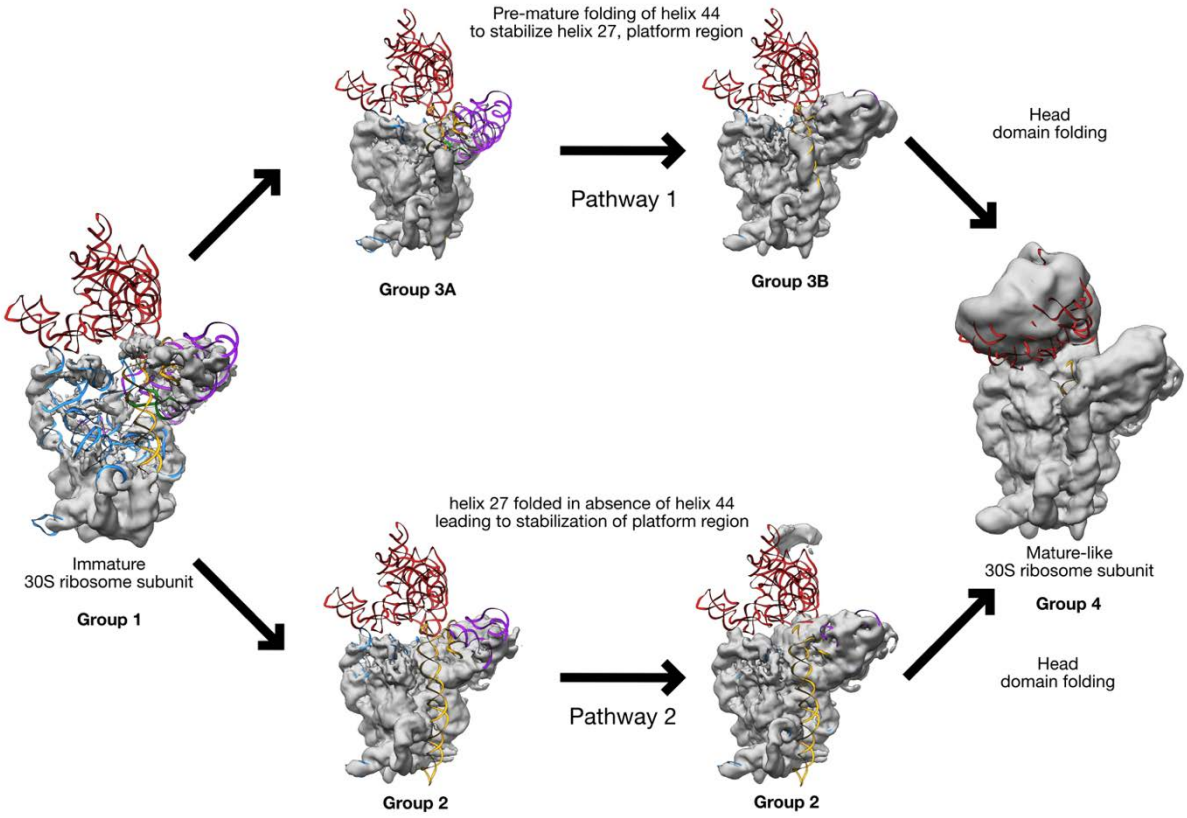


Figure 3. 6 Proposed model for the role of YjeQ in the maturation of the 30S subunit.

cryoDRGN analysis revealed four major groups of the 30SΔyjeQ classes. Immature 30S (Group 1) may either proceed through pathway 1 via pre-mature folding of helix 44 (Group 3A) to stabilize helix 27 and trigger the folding of the platform (Group 3B) and eventually complete the maturation (Group 4). Alternatively, Immature 30S (Group 1) may proceed through pathway 2 via pre-folded helix 27, which leads to the folding of the head and the platform region (Group 2) eventually to complete the maturation (Group 4). Maturation of the 30S in the absence of YjeQ follows the classical pathway to fold body, platform, and the head domain in this order.

The study from (Thurlow et al., 2016) suggested that immature ribosome subunits that accumulate due to deletion or depletion of a specific assembly factor may be trapped in the local energy

minimum. As a result, assembly intermediates might adopt conformational changes that are unrecognizable from that assembly factor.

Observations from our $30S_{\Delta yjeQ}$ cryo-EM dataset provide a possible explanation for the poor affinity of YjeQ towards the $30S_{\Delta yjeQ}$ particles *in vitro*. Accordingly, particles that contribute to Group 4 (mature like the 30S) only make up ~9% of the population. Therefore, *in vitro,* only 9% of the $30S_{\Delta yjeQ}$ particles may be recognizable by YjeQ.

Role of YjeQ in quality control of translational fidelity

Quality control mechanisms that prevent the premature entry of ribosome assembly intermediates into the translation cycle in eukaryotic cells have been studied extensively (Karbstein, 2013). However, in the case of the prokaryotes, we have only recently started to uncover similar quality control mechanisms. This study establishes how YjeQ acts as a quality control protein, ensuring that assembled 30S subunits maintain high translational fidelity. Our cryo-EM analysis of the 30S+YjeQ double mutant complex and the 30S in the 'active' state reveal how superficial probing of helix 44 by YjeQ double mjutant contributes to the disorder in the decoding region, mainly in the A-site of the 30S. Accordingly, YjeQ with mutations in its OB-fold domain may loosely bind to the 30S ribosome subunit yet maintain essential interactions with the 30S ribosome subunit. For example, the YjeQ double mutant, through its crucial GTPase domain residues, interacts with the 16S rRNA nucleotides in the same way as the YjeQ wildtype does **(Figure S3.5)**. This residue is essential for triggering YjeQ GTP hydrolysis. As a result, YjeQ with OB-fold domain mutations binds and releases from the 30S, mimicking a flawed translation-like cycle that wild-type YjeQ has been suggested to do. Therefore *in vivo*, YjeQ mutants fail to probe the conformation of the helix44 decoding region and might still be able to hydrolyze GTP, which explains why *E. coli yjeQ* deletion strain also demonstrates frameshift errors in our β-galactosidase assays.

Further, the fact that Initiation Factor 3 gets recruited to form the pre-initiation complex with mRNA and f-Met-tRNA suggests that these 30S ribosome subunits may be carried to the initiation stage with a local disorder decoding region. This is consistent with the previously observed phenomenon that the ribosomes from the Δ*yjeQ* strain have not compromised their translation initiation mechanism (Sharma and Anand, 2019). Overall, these observations prove that at the level of ribosome assembly, YjeQ ensures the appropriate folding/maturation of the decoding region helices so that ribosomes can accurately perform the 'initial decoding' step as part of the two-step decoding and proofreading mechanism that allows the ribosomes to maintain high translational fidelity.

Recruitment of YjeQ towards the end of ribosome assembly

One of several roles that YjeQ plays during ribosome assembly is to release ribosome assembly factor RbfA from the 30S (Jeganathan et al., 2015) (Goto et al., 2011). Follow-up cryo-EM studies from our laboratory revealed that YjeQ acts as an RNA chaperone and alters the conformation of the decoding region helices 44 and 45, which may help remove RbfA and promote the 30S transition towards translation initiation machinery (Razi et al., 2017b). These exact conformation changes in helices 44 and 45 are now known as the transition from inactive to active state (Jahagirdar et al., 2020). However, recent biochemical studies have clarified this functional interplay between RbfA and YjeQ (Sharma and Woodson, 2020). According to this study, when YjeQ releases RbfA, helices 44 and 45 can fall back into the inactive state, which provides an opportunity for RbfA to bind again to these inactive 30S ribosome subunits, leading to the failure of 30S ribosome subunits to transition from the assembly stage to translation initiation stage. This correlates with our finding that ~7% of the 30S population we identified in our 30S + YjeQ double-mutant dataset was inactive **(Figure S2 bottom row)**. Consistent with this biochemical data, recent

cryo-EM studies have proved that RbfA alone can promote the transition of the 30S ribosome subunit from an inactive to an active state (Schedlbauer et al., 2020). These findings imply that the recruitment of YjeQ towards the end of ribosome assembly to alter the conformation of 16S rRNA helices 44 and 45 is primarily for quality control purposes.

3.6 Supplemental Material

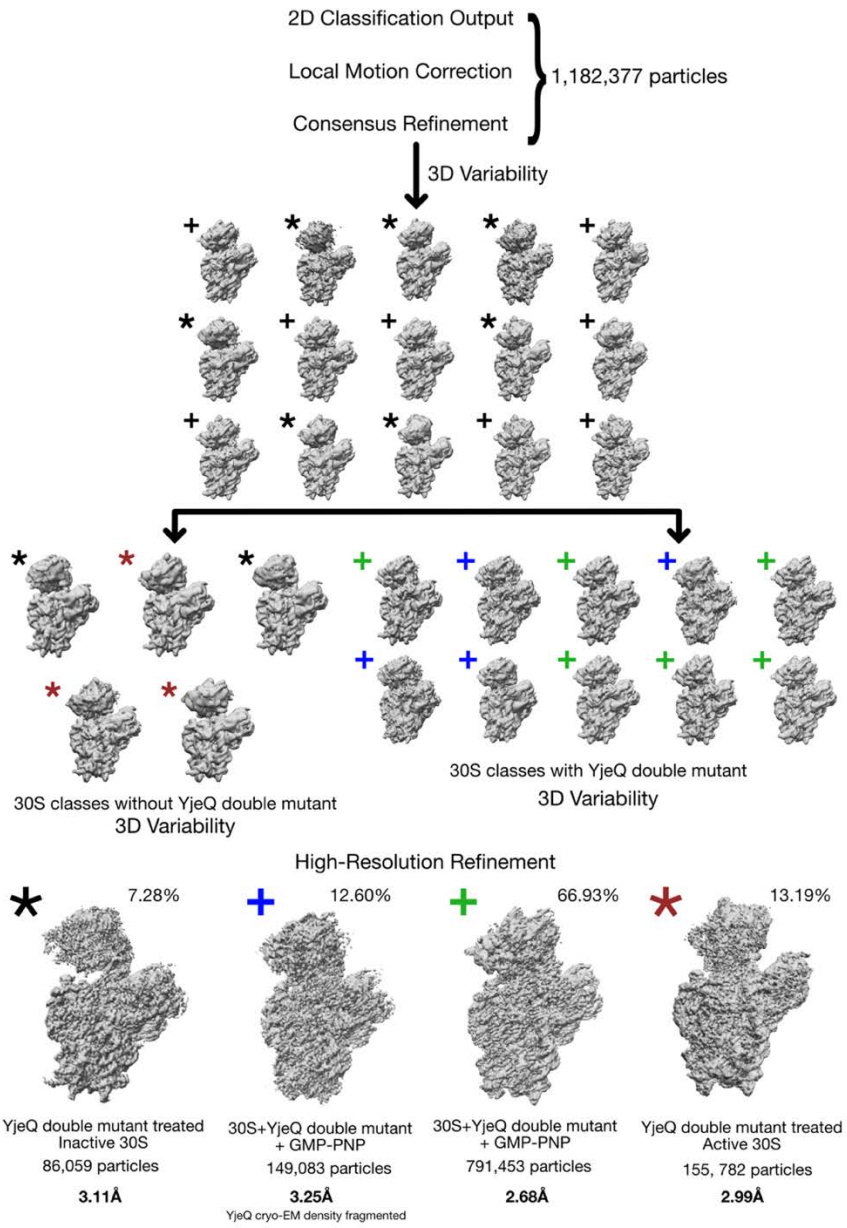


Figure S3. 1 Image processing workflow for the 30S+YjeQ double mutant dataset.

Purified 30S+YjeQ double mutant particles were imaged by cryo-EM and images were subjected to multiple layers of 3D classification using cryoSPARC 3D Variability. Percentages for each class are calculated within each subgroup. The resulting subclasses were grouped into four main classes indicated by a red asterisk, black asterisk, blue cross, and green cross. Particles in these two groups were used to produce the high-resolution cryo-EM maps shown at the bottom. The maps shown are unsharpened. The number of particles in each class or used for the final refinement of each class is indicated. The resolution value obtained for the cryo-EM maps of each one of the classes is also indicated.

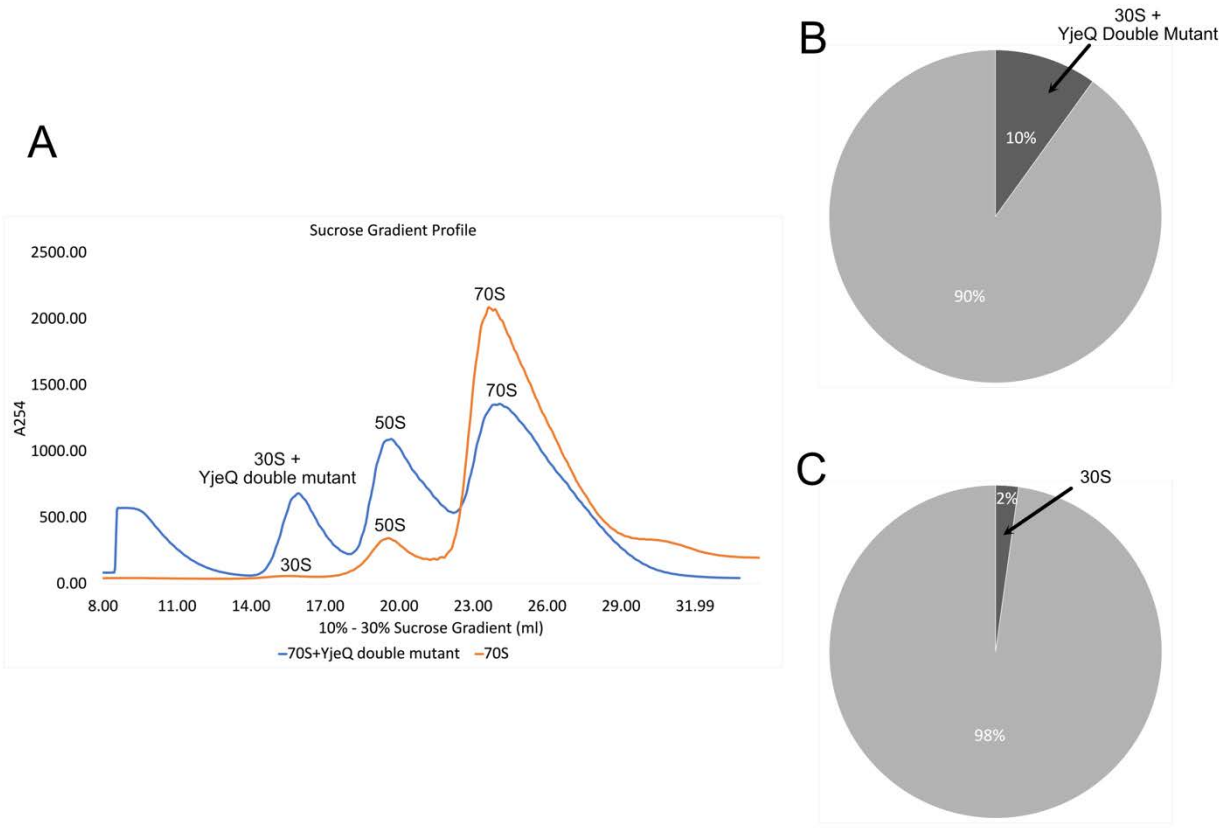


Figure S3. 2 YjeQ double mutant in presence of GMP-PNP can dissociate 70S ribosomes under associating conditions

Panel A shows ribosome purification profiles for 70S ribosome under associating conditions (15 mM Mg^{2+} in typically Buffer E). Curves are in orange color. Ribosome profile for 70S+YjeQ double mutant reaction under associating conditions (15 mM Mg^{2+} in typically Buffer E) and in presence of GMP-PNP is blue color. Panel B shows pie chart highlighting the area covered by the 30S+YjeQ double mutant in dark grey. Panel C shows pie chart highlighting the area covered by the 30S in dark grey.

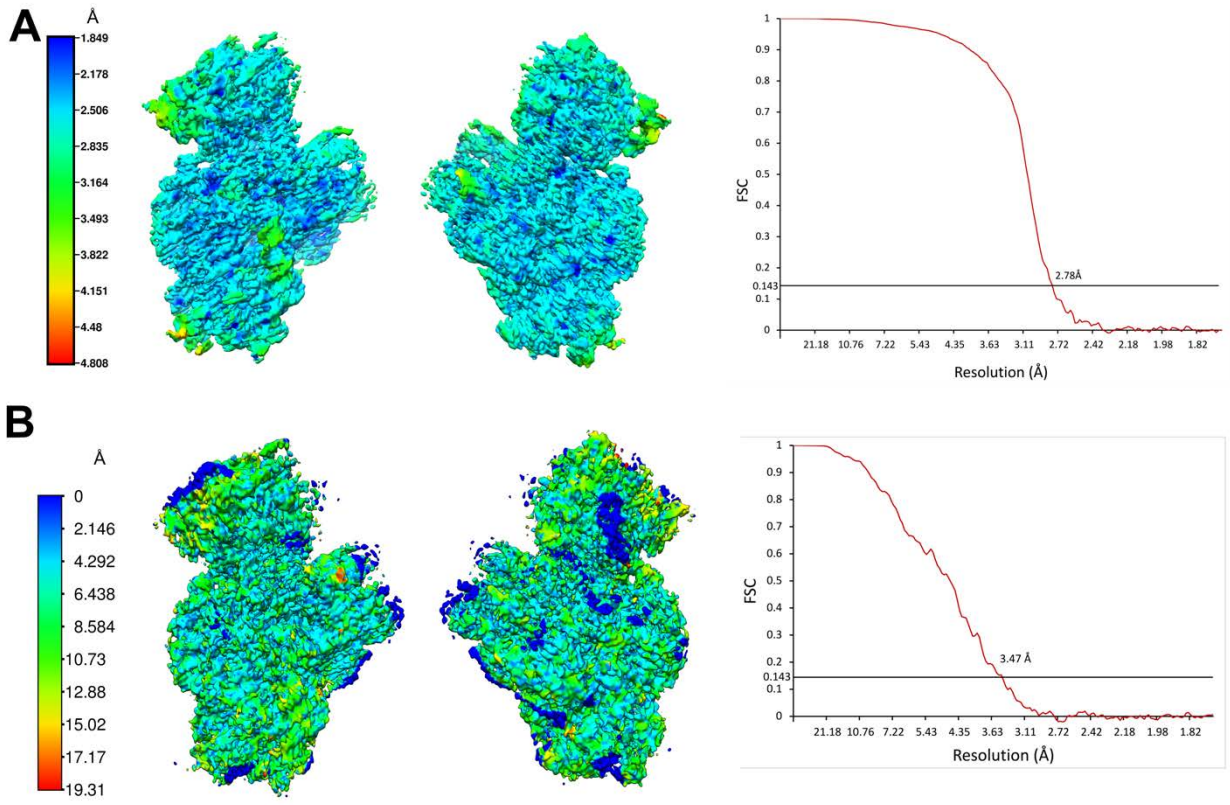


Figure S3. 3 Resolution analysis and Fourier Shell Correlation (FSC) of the 30S+YjeQ double mutant complex.

Panel A and B show the local resolution analysis and FSC curves for the 30S+YjeQ double mutant and 30S+YjeQ double mutant (fragmented YjeQ) cryo-EM for the interface and the solvent views respectively. Maps are colored according to their local resolution using the color coding indicated in the scale bars.

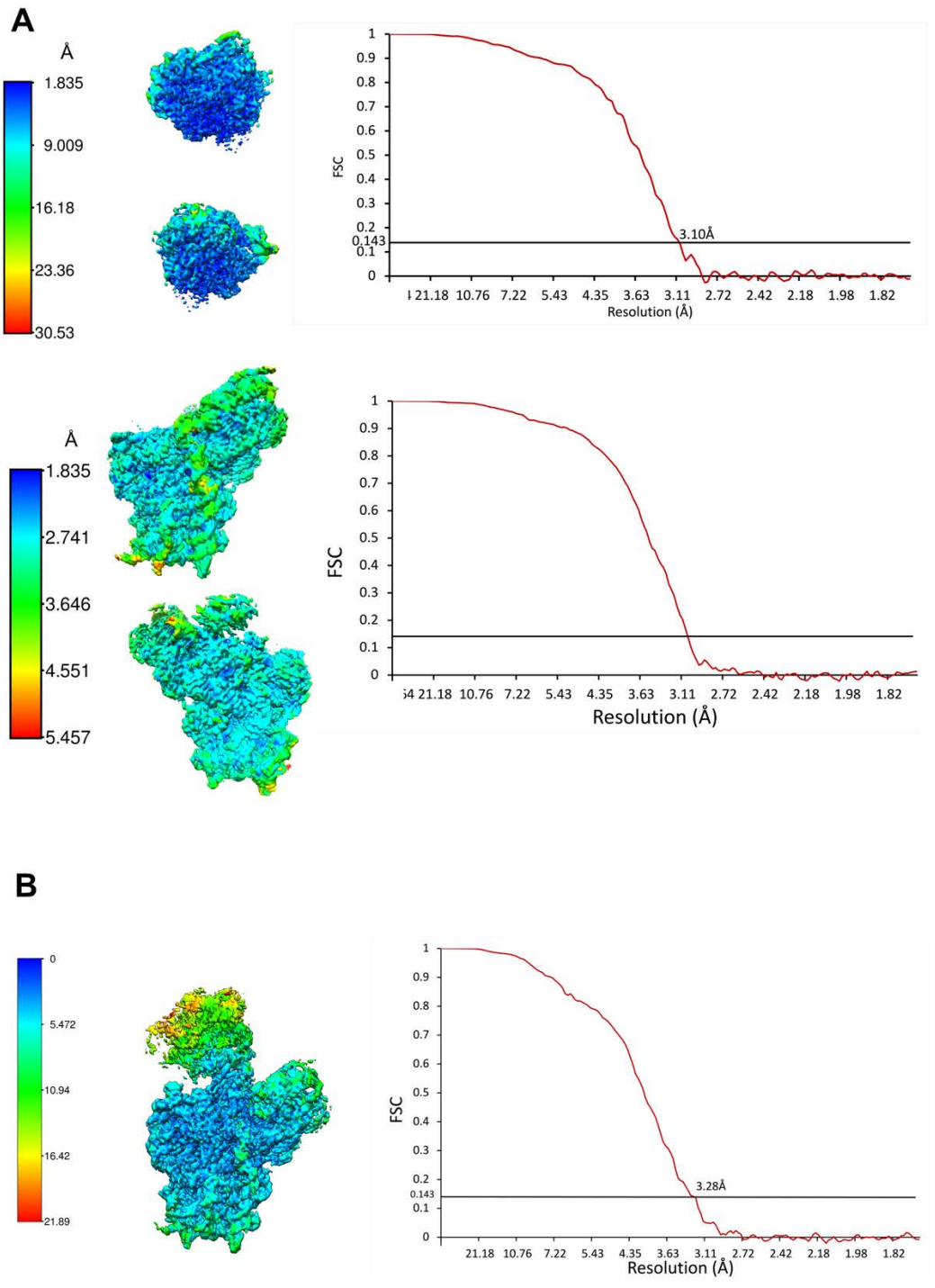


Figure S3. 4 Resolution analysis and Fourier Shell Correlation (FSC) of the 'active' and 'inactive' 30S.

Panel A and B show the local resolution analysis and FSC curves for the 'active' and 'inactive' 30S cryo-EM maps. Maps are colored according to their local resolution using the color coding indicated in the scale bars.

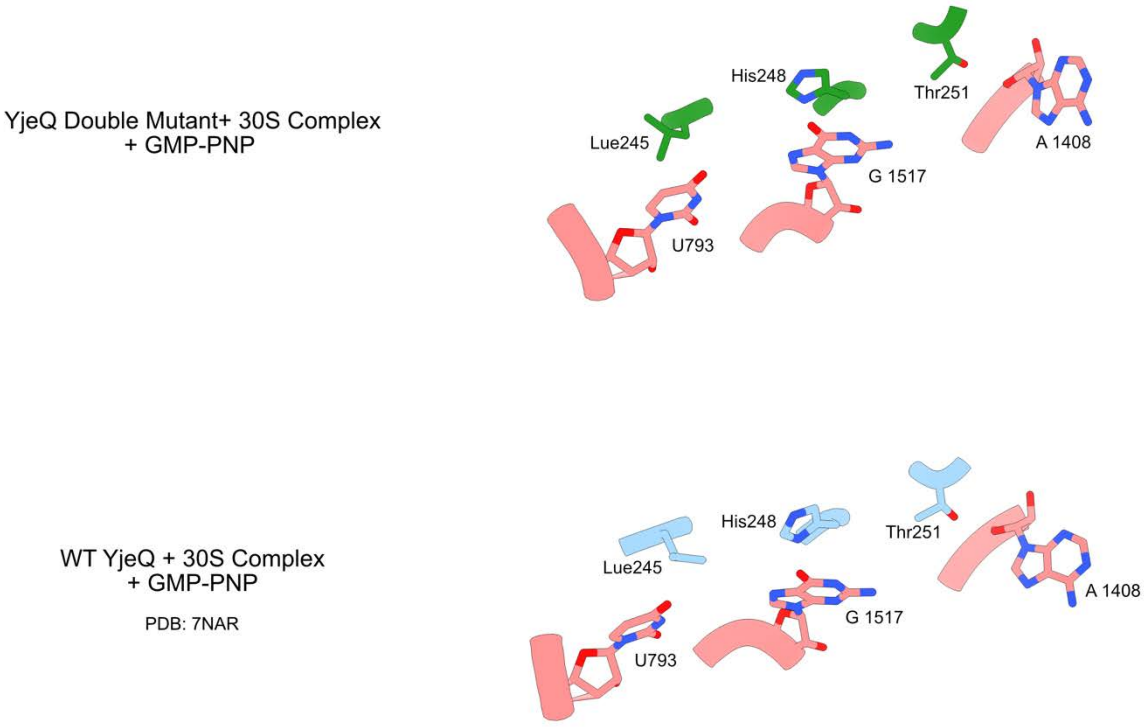


Figure S3. 5 Comparison between interaction of YjeQ double mutant and wildtype GTPase domain residues with 16S rRNA.

Top panel shows interaction of residues from YjeQ double mutant GTPase domain (green) with 16S rRNA. Bottom panel shows interaction of residues from YjeQ wildtype GTPase domain (blue) with 16S rRNA. Interaction of these YjeQ double mutant residues remain the same despite mutation in the OB-fold domain.

3.7 References

AFONINE, P. V., POON, B. K., READ, R. J., SOBOLEV, O. V., TERWILLIGER, T. C., URZHUMTSEV, A. & ADAMS, P. D. 2018. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol,* 74**,** 531-544.

ARIGONI, F., TALABOT, F., PEITSCH, M., EDGERTON, M. D., MELDRUM, E., ALLET, E., FISH, R., JAMOTTE, T., CURCHOD, M. L. & LOFERER, H. 1998. A genome-based approach for the identification of essential bacterial genes. *Nat Biotechnol,* 16**,** 851-6.

BABA, T., ARA, T., HASEGAWA, M., TAKAI, Y., OKUMURA, Y., BABA, M., DATSENKO, K. A., TOMITA, M., WANNER, B. L. & MORI, H. 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol,* 2**,** 2006 0008.

CAMPBELL, T. L. & BROWN, E. D. 2008. Genetic interaction screens with ordered overexpression and deletion clone sets implicate the Escherichia coli GTPase YjeQ in late ribosome biogenesis. *J Bacteriol,* 190**,** 2537-45.

CAMPBELL, T. L., DAIGLE, D. M. & BROWN, E. D. 2005. Characterization of the Bacillus subtilis GTPase YloQ and its role in ribosome function. *Biochem J,* 389**,** 843-52.

CARTER, A. P., CLEMONS, W. M., BRODERSEN, D. E., MORGAN-WARREN, R. J., WIMBERLY, B. T. & RAMAKRISHNAN, V. 2000. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature,* 407**,** 340-8.

CARTER, A. P., CLEMONS, W. M., JR., BRODERSEN, D. E., MORGAN-WARREN, R. J., HARTSCH, T., WIMBERLY, B. T. & RAMAKRISHNAN, V. 2001. Crystal structure of an initiation factor bound to the 30S ribosomal subunit. *Science,* 291**,** 498-501.

CHADANI, Y., ONO, K., KUTSUKAKE, K. & ABO, T. 2011. Escherichia coli YaeJ protein mediates a novel ribosome-rescue pathway distinct from SsrA- and ArfA-mediated pathways. *Mol Microbiol,* 80**,** 772-85.

CIANFROCCO, M. A., WONG-BARNUM, M., YOUN, C., WAGNER, R. & LESCHZINER, A. 2017. COSMIC2: A Science Gateway for Cryo-Electron Microscopy Structure Determination. *Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact.* New Orleans, LA, USA: Association for Computing Machinery.

CUNNINGHAM, P. R., NURSE, K., WEITZMANN, C. J. & OFENGAND, J. 1993. Functional effects of base changes which further define the decoding center of Escherichia coli 16S ribosomal RNA: mutation of C1404, G1405, C1496, G1497, and U1498. *Biochemistry,* 32**,** 7172-80.

DAIGLE, D. M. & BROWN, E. D. 2004. Studies of the interaction of Escherichia coli YjeQ with the ribosome in vitro. *J Bacteriol,* 186**,** 1381-7.

DAIGLE, D. M., ROSSI, L., BERGHUIS, A. M., ARAVIND, L., KOONIN, E. V. & BROWN, E. D. 2002. YjeQ, an essential, conserved, uncharacterized protein from Escherichia coli, is an unusual GTPase with circularly permuted G-motifs and marked burst kinetics. *Biochemistry,* 41**,** 11109-17.

DATSENKO, K. A. & WANNER, B. L. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A,* 97**,** 6640-5.

EMSLEY, P. & COWTAN, K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr,* 60**,** 2126-32.

EMSLEY, P., LOHKAMP, B., SCOTT, W. G. & COWTAN, K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr,* 66**,** 486-501.

GAGNON, M. G., SEETHARAMAN, S. V., BULKLEY, D. & STEITZ, T. A. 2012. Structural basis for the rescue of stalled ribosomes: structure of YaeJ bound to the ribosome. *Science,* 335**,** 1370-2.

GALLANT, J. A. & LINDSLEY, D. 1998. Ribosomes can slide over and beyond "hungry" codons, resuming protein chain elongation many nucleotides downstream. *Proc Natl Acad Sci U S A,* 95**,** 13771-6.

HANDA, Y., INAHO, N. & NAMEKI, N. 2011. YaeJ is a novel ribosome-associated protein in Escherichia coli that can hydrolyze peptidyl-tRNA on stalled ribosomes. *Nucleic Acids Res,* 39**,** 1739-48.

HIMENO, H., HANAWA-SUETSUGU, K., KIMURA, T., TAKAGI, K., SUGIYAMA, W., SHIRATA, S., MIKAMI, T., ODAGIRI, F., OSANAI, Y., WATANABE, D., GOTO, S., KALACHNYUK, L., USHIDA, C. & MUTO, A. 2004. A novel GTPase activated by the small subunit of ribosome. *Nucleic Acids Res,* 32**,** 5303-9.

JAHAGIRDAR, D., JHA, V., BASU, K., GOMEZ-BLANCO, J., VARGAS, J. & ORTEGA, J. 2020. Alternative conformations and motions adopted by 30S ribosomal subunits visualized by cryo-electron microscopy. *RNA,* 26**,** 2017-2030.

JEGANATHAN, A., RAZI, A., THURLOW, B. & ORTEGA, J. 2015. The C-terminal helix in the YjeQ zinc-finger domain catalyzes the release of RbfA during 30S ribosome subunit assembly. *RNA,* 21**,** 1203-16.

JOMAA, A., STEWART, G., MARTIN-BENITO, J., ZIELKE, R., CAMPBELL, T. L., MADDOCK, J. R., BROWN, E. D. & ORTEGA, J. 2011a. Understanding ribosome assembly: the structure of in vivo assembled immature 30S subunits revealed by cryo-electron microscopy. *RNA,* 17**,** 697-709.

JOMAA, A., STEWART, G., MEARS, J. A., KIREEVA, I., BROWN, E. D. & ORTEGA, J. 2011b. Cryo-electron microscopy structure of the 30S subunit in complex with the YjeQ biogenesis factor. *RNA,* 17**,** 2026-38.

LENARCIC, T., NIEMANN, M., RAMRATH, D. J. F., CALDERARO, S., FLUGEL, T., SAURER, M., LEIBUNDGUT, M., BOEHRINGER, D., PRANGE, C., HORN, E. K., SCHNEIDER, A. & BAN, N. 2022. Mitoribosomal small subunit maturation involves formation of initiation-like complexes. *Proc Natl Acad Sci U S A,* 119.

LEONG, V., KENT, M., JOMAA, A. & ORTEGA, J. 2013. Escherichia coli rimM and yjeQ null strains accumulate immature 30S subunits of similar structure and protein complement. *RNA,* 19**,** 789-802.

LIEBSCHNER, D., AFONINE, P. V., BAKER, M. L., BUNKOCZI, G., CHEN, V. B., CROLL, T. I., HINTZE, B., HUNG, L. W., JAIN, S., MCCOY, A. J., MORIARTY, N. W., OEFFNER, R. D., POON, B. K., PRISANT, M. G., READ, R. J., RICHARDSON, J. S., RICHARDSON, D. C., SAMMITO, M. D., SOBOLEV, O. V., STOCKWELL, D. H., TERWILLIGER, T. C., URZHUMTSEV, A. G., VIDEAU, L. L., WILLIAMS, C. J. & ADAMS, P. D. 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol,* 75**,** 861-877.

LOPEZ-ALONSO, J. P., KAMINISHI, T., KIKUCHI, T., HIRATA, Y., ITURRIOZ, I., DHIMOLE, N., SCHEDLBAUER, A., HASE, Y., GOTO, S., KURITA, D., MUTO, A., ZHOU, S., NAOE, C., MILLS, D. J., GIL-CARTON, D., TAKEMOTO, C., HIMENO, H., FUCINI, P. & CONNELL, S. R. 2017. RsgA couples the maturation state of the 30S

ribosomal decoding center to activation of its GTPase pocket. *Nucleic Acids Res,* 45**,** 6945-6959.
MCGINNIS, J. L., LIU, Q., LAVENDER, C. A., DEVARAJ, A., MCCLORY, S. P., FREDRICK, K. & WEEKS, K. M. 2015. In-cell SHAPE reveals that free 30S ribosome subunits are in the inactive state. *Proc Natl Acad Sci U S A,* 112**,** 2425-30.
MILLER, J. H. 1972. *Experiments in molecular genetics,* Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory.
MOAZED, D., VAN STOLK, B. J., DOUTHWAITE, S. & NOLLER, H. F. 1986. Interconversion of active and inactive 30 S ribosomal subunits is accompanied by a conformational change in the decoding region of 16 S rRNA. *J Mol Biol,* 191**,** 483-93.
MOHAN, S., DONOHUE, J. P. & NOLLER, H. F. 2014. Molecular mechanics of 30S subunit head rotation. *Proc Natl Acad Sci U S A,* 111**,** 13325-30.
MOORE, P. B. & STEITZ, T. A. 2011. The roles of RNA in the synthesis of protein. *Cold Spring Harb Perspect Biol,* 3**,** a003780.
MULDER, A. M., YOSHIOKA, C., BECK, A. H., BUNNER, A. E., MILLIGAN, R. A., POTTER, C. S., CARRAGHER, B. & WILLIAMSON, J. R. 2010. Visualizing ribosome biogenesis: parallel assembly pathways for the 30S subunit. *Science,* 330**,** 673-7.
NICHOLS, C. E., JOHNSON, C., LAMB, H. K., LOCKYER, M., CHARLES, I. G., HAWKINS, A. R. & STAMMERS, D. K. 2007. Structure of the ribosomal interacting GTPase YjeQ from the enterobacterial species Salmonella typhimurium. *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 63**,** 922-8.
O'CONNOR, M., THOMAS, C. L., ZIMMERMANN, R. A. & DAHLBERG, A. E. 1997. Decoding fidelity at the ribosomal A and P sites: influence of mutations in three different regions of the decoding domain in 16S rRNA. *Nucleic Acids Res,* 25**,** 1185-93.
OGLE, J. M., BRODERSEN, D. E., CLEMONS, W. M., JR., TARRY, M. J., CARTER, A. P. & RAMAKRISHNAN, V. 2001. Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science,* 292**,** 897-902.
OGLE, J. M. & RAMAKRISHNAN, V. 2005. Structural insights into translational fidelity. *Annu Rev Biochem,* 74**,** 129-77.
PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem,* 25**,** 1605-12.
PUNJANI, A., RUBINSTEIN, J. L., FLEET, D. J. & BRUBAKER, M. A. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods,* 14**,** 290-296.
RAZI, A., DAVIS, J. H., HAO, Y., JAHAGIRDAR, D., THURLOW, B., BASU, K., JAIN, N., GOMEZ-BLANCO, J., BRITTON, R. A., VARGAS, J., GUARNE, A., WOODSON, S. A., WILLIAMSON, J. R. & ORTEGA, J. 2019. Role of Era in assembly and homeostasis of the ribosomal small subunit. *Nucleic Acids Res,* 47**,** 8301-8317.
RAZI, A., GUARNE, A. & ORTEGA, J. 2017. The cryo-EM structure of YjeQ bound to the 30S subunit suggests a fidelity checkpoint function for this protein in ribosome assembly. *Proc Natl Acad Sci U S A,* 114**,** E3396-E3403.
SANCHEZ-GARCIA, R., GOMEZ-BLANCO, J., CUERVO, A., CARAZO, J. M., SORZANO, C. O. S. & VARGAS, J. 2021. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun Biol,* 4**,** 874.

SCHEDLBAUER, A., ITURRIOZ, I., OCHOA-LIZARRALDE, B., CAPUNI, R., HAN, X., DE ASTIGARRAGA, E., DIERCKS, T., FUCINI, P. & CONNELL, S. R. 2020. Backbone and sidechain NMR assignments for the ribosome maturation factor RbfA from Escherichia coli. *Biomol NMR Assign,* 14**,** 317-321.

SCHEDLBAUER, A., ITURRIOZ, I., OCHOA-LIZARRALDE, B., DIERCKS, T., LOPEZ-ALONSO, J. P., LAVIN, J. L., KAMINISHI, T., CAPUNI, R., DHIMOLE, N., DE ASTIGARRAGA, E., GIL-CARTON, D., FUCINI, P. & CONNELL, S. R. 2021. A conserved rRNA switch is central to decoding site maturation on the small ribosomal subunit. *Sci Adv,* 7.

SCHERES, S. H. & CHEN, S. 2012. Prevention of overfitting in cryo-EM structure determination. *Nat Methods,* 9**,** 853-4.

SCHORB, M., HABERBOSCH, I., HAGEN, W. J. H., SCHWAB, Y. & MASTRONARDE, D. N. 2019. Software tools for automated transmission electron microscopy. *Nat Methods,* 16**,** 471-477.

SHARMA, H. & ANAND, B. 2019. Ribosome assembly defects subvert initiation Factor3 mediated scrutiny of bona fide start signal. *Nucleic Acids Res,* 47**,** 11368-11386.

SHARMA, I. M. & WOODSON, S. A. 2020. RbfA and IF3 couple ribosome biogenesis and translation initiation to increase stress tolerance. *Nucleic Acids Res,* 48**,** 359-372.

SHIMIZU, Y. 2012. ArfA recruits RF2 into stalled ribosomes. *J Mol Biol,* 423**,** 624-31.

SHIN, D. H., LOU, Y., JANCARIK, J., YOKOTA, H., KIM, R. & KIM, S. H. 2004. Crystal structure of YjeQ from Thermotoga maritima contains a circularly permuted GTPase domain. *Proc Natl Acad Sci U S A,* 101**,** 13198-203.

THURLOW, B., DAVIS, J. H., LEONG, V., MORAES, T. F., WILLIAMSON, J. R. & ORTEGA, J. 2016. Binding properties of YjeQ (RsgA), RbfA, RimM and Era to assembly intermediates of the 30S subunit. *Nucleic Acids Res,* 44**,** 9918-9932.

WIMBERLY, B. T., BRODERSEN, D. E., CLEMONS, W. M., JR., MORGAN-WARREN, R. J., CARTER, A. P., VONRHEIN, C., HARTSCH, T. & RAMAKRISHNAN, V. 2000. Structure of the 30S ribosomal subunit. *Nature,* 407**,** 327-39.

WINN, M. D., BALLARD, C. C., COWTAN, K. D., DODSON, E. J., EMSLEY, P., EVANS, P. R., KEEGAN, R. M., KRISSINEL, E. B., LESLIE, A. G., MCCOY, A., MCNICHOLAS, S. J., MURSHUDOV, G. N., PANNU, N. S., POTTERTON, E. A., POWELL, H. R., READ, R. J., VAGIN, A. & WILSON, K. S. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr,* 67**,** 235-42.

WOODSON, S. A. 2011. RNA folding pathways and the self-assembly of ribosomes. *Acc Chem Res,* 44**,** 1312-9.

YOSHIZAWA, S., FOURMY, D. & PUGLISI, J. D. 1999. Recognition of the codon-anticodon helix by ribosomal RNA. *Science,* 285**,** 1722-5.

ZHANG, X. & BREMER, H. 1995. Control of the Escherichia coli rrnB P1 promoter strength by ppGpp. *J Biol Chem,* 270**,** 11181-9.

ZHONG, E. D., BEPLER, T., BERGER, B. & DAVIS, J. H. 2021. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nat Methods,* 18**,** 176-185.

Connecting Text

Our focus in the previous two chapters was to explore aspects of bacterial ribosomes that have yet to be exploited for the development of novel antibiotics. Accordingly, we first solved the structure of the 'inactive' 30S, which represents the native conformation and could not be solved using X-ray crystallography. We also explored the direct or indirect involvement of YjeQ in ribosome assembly, translational fidelity, and possibly ribosome rescue. Thus, highlighting components of ribosome biogenesis as an attractive target for developing new antibiotics. However, in both the studies we used *E. coli* as a model organism mainly because, a significant amount of prior knowledge came from this model organism. Similarly, much of our understanding of ribosome biology comes either from *E. coli* or *B. subtilis* since there are good genetic tools to exploit these bacteria.

Therefore, in the upcoming chapter, we chose to explore aspects of ribosome biology from bacteria that are not commonly used as model organisms for understanding ribosome biology. Accordingly, we chose to study ribosomes from bacteria belonging to the Bacteroidetes phylum for two reasons:

(1) They are the dominant human gut microbiota and are implicated in obesity and may act as opportunistic pathogens causing infection at a broad range of sites like the abdomen, central nervous systems, oral cavity, etc.

(2) These bacteria can successfully and effectively initiate translation mechanisms without the need for SD-ASD recognition. Therefore, making the translation initiation mechanism in Bacteroidetes unique, unexplored, and an attractive target for the development of new antibiotics.

CHAPTER 4: Structural basis of sequestration of the Anti-Shine-Dalgarno sequence in the Bacteroidetes ribosome

Vikash Jha[1,2 ±], Bappaditya Roy[3,4 ±], Dushyant Jahagirdar[1,2 ±], Zakkary A. McNutt[4,5], Elan A. Shatoff[4,6], Bethany L. Boleratz[5], Dean E. Watkins[3], Ralf Bundschuh[4,6,7], Kaustuv Basu[1,2], Joaquin Ortega[1,2 *], Kurt Fredrick[3,4,5 *]

[1]Department of Anatomy and Cell Biology, McGill University, Montreal, Quebec H3A 0C7, Canada, [2]Centre for Structural Biology, McGill University, Montreal, Quebec H3G 0B1, Canada, [3]Department of Microbiology, The Ohio State University, Columbus, Ohio 43210, USA, [4]Center for RNA Biology, The Ohio State University, Columbus, Ohio 43210, USA, [5]Ohio State Biochemistry Program, The Ohio State University, Columbus, Ohio 43210, USA, [6]Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA, [7]Department of Chemistry & Biochemistry, Division of Hematology, The Ohio State University, Columbus, Ohio 43210

Published

Nucleic Acids Research, Volume 49, Issue 1, 11 January 2021, Pages 547–567

Contribution of Authors

D.J., V.J., and B.R. have equal contributions to this work as First Authors. D.J. performed cryo-EM sample preparation, EM data collection and image analysis; V.J. performed model building of the atomic models; B.P. performed most biochemical experiments; Z.A.M. performed multiple sequence alignments and contributed to study design; E.A.S. performed bioinformatics analyses; B.L.B. collected the filter-binding data; D.E.W. determined the N-terminus of F. johnsoniae IF3; R.B. designed and supervised the bioinformatics analyses; K.B. designed and implemented strategies for cryo-EM data collection in the Titan Krios TEM (FEMR-McGill); J.O. and K.F. designed and supervised the study. V.J., J.O., and K.F. wrote the first draft of the manuscript and all authors contributed to editing.

4.1 Abstract

Genomic studies have indicated that certain bacterial lineages such as the Bacteroidetes lack Shine-Dalgarno (SD) sequences, and yet with few exceptions ribosomes of these organisms carry the canonical anti-SD (ASD) sequence. Here, we show that ribosomes purified from Flavobacterium johnsoniae, a representative of the Bacteroidetes, fail to recognize the SD sequence of mRNA in vitro. A cryo-electron microscopy structure of the complete 70S ribosome from F. johnsoniae at 2.8 Å resolution reveals that the ASD is sequestered by ribosomal proteins bS21, bS18 and bS6, explaining the basis of ASD inhibition. The structure also uncovers a novel ribosomal protein—bL38. Remarkably, in F. johnsoniae and many other Flavobacteriia, the gene encoding bS21 contains a strong SD, unlike virtually all other genes. A subset of Flavobacteriia have an alternative ASD, and in these organisms the fully complementary sequence lies upstream of the bS21 gene, indicative of natural covariation. In other Bacteroidetes classes, strong SDs are frequently found upstream of the genes for bS21 and/or bS18. We propose that these SDs are used as regulatory elements, enabling bS21 and bS18 to translationally control their own production.

4.2 Introduction

The translation machinery must identify the correct start codon among all other AUG (and similar) trinucleotides to establish the reading frame for protein synthesis. In all cells, start codon selection is guided by intrinsic features of the mRNA. For prokaryotes, one such feature is the Shine Dalgarno (SD) sequence, a purine-rich element (e.g., GGAGG) that lies 7-9 nucleotides (nt) upstream from the start codon (Shine and Dalgarno, 1974) (Shultzaberger et al., 2001) (Vellanoweth and Rabinowitz, 1992). The SD base-pairs with the 30S subunit's anti-Shine-Dalgarno (ASD) sequence, a pyrimidine-rich element (CCUCC) near the 3' end of 16S rRNA. SD-ASD interaction positions the start codon in the P site and stabilizes mRNA in the 30S initiation complex (IC) (Steitz and Jakes, 1975) (Jenner et al., 2010) (Korostelev et al., 2007) (Hussain et al., 2016) (Kaminishi et al., 2007). The SD motif is commonly used to help delineate genes (Zhu et al., 2004) (Delcher et al., 2007) (Besemer et al., 2001) (Suzek et al., 2001) (Ou et al., 2004), and often the SD and start codon are collectively referred to as the ribosome binding site (RBS) (Osada et al., 1999) (Ringquist et al., 1992) (Schurr et al., 1993). Numerous studies have shown that the SD, when present, plays an important role in initiation (de Smit and van Duin, 1994) (Hui and de Boer, 1987, Jacob et al., 1987). However, faithful and efficient initiation can also occur on bacterial mRNAs that naturally have no SD (Li et al., 2014) (Schrader et al., 2014) (Skorski et al., 2006).

In 2010, Nakagawa and coworkers analyzed the genomes of >200 representative bacteria and found that SD usage varies substantially (Nakagawa et al., 2010). Remarkably, certain lineages including the Bacteroidetes and a subset of cyanobacteria appear to lack SD sequences completely, a result confirmed in subsequent studies (Nakagawa et al., 2017) (Wegmann et al., 2013) (Accetto and Avgustin, 2011). With few exceptions (Lim et al., 2012) (Amin et al., 2018), these bacteria

retain the canonical anti-Shine-Dalgarno (ASD) sequence at the 3' end of 16S rRNA. Yet, based on reporter gene studies in representative species, the corresponding ribosomes fail to recognize SD sequences *in vivo* (Wegmann et al., 2013) (Accetto and Avgustin, 2011). These observations imply that other mRNA determinants contribute to start codon selection in the Bacteroidetes and that some mechanism prevents SD-ASD pairing (or makes it inconsequential) during initiation.

The Bacteroidetes represent a large and understudied group of bacteria (Hahnke et al., 2016). They inhabit diverse environments, including the soil, ocean, and animal gut. Well known for their ability to import and degrade complex polysaccharides, members of the phylum account for a substantial portion of the intestinal microbiome of mammals (Ley et al., 2008) (Martens et al., 2009b) (Martens et al., 2009a) (Martens et al., 2008). There, they play a crucial role in catabolic processes, impacting nutrient availability and uptake by the host intestinal epithelium. Human health problems such as irritable bowel syndrome, inflammatory bowel disease, obesity, type 2 diabetes, and autism-spectrum disorders have been associated with under- or over-abundances of Bacteroidetes species in the gut microbiome (Johnson et al., 2017) (Rinninella et al., 2019). Gene expression in the Bacteroidetes differs substantially from that of well-studied *E. coli*, with distinct mechanisms involved in not only translation initiation as discussed above but also transcription initiation (Bayley et al., 2000) (Chen et al., 2007b) (Chen et al., 2007a) (Vingadassalom et al., 2005). Consequently, genes moved from these organisms into *E. coli* yield no protein products and *vice versa*.

Baez et al. 2019 used ribosome profiling to identify mRNA determinants of translation initiation in *Flavobacterium johnsoniae*, a representative of the Bacteroidetes (Baez et al., 2019). They found that initiation is enhanced by reduced secondary structure in the translation initiation region (TIR) and by adenines at positions -3, -6, -12, and -13. Comparative analysis of *Escherichia coli*

showed that A-3 and A-6 also act as positive determinants of translation. Notably, A-3 is the key feature of the Kozak sequence (Kozak, 1986) (Nakagawa et al., 2008) (Yamauchi, 1991), which in eukaryotes promotes start codon selection. Baez at al. also found that AUG trinucleotides other than the start codon are significantly underrepresented in the TIR, and the degree of underrepresentation follows the trend Bacteroidetes > Proteobacteria > Firmicutes. The opposite trend is seen for SD prevalence (Nakagawa et al., 2017) (Nakagawa et al., 2010), suggesting that elimination of AUG trinucleotides from the TIR is one means by which Bacteroidetes compensate for the absence of SD-ASD pairing (Baez et al., 2019).

Genomic studies indicate that the ASD sequence is conserved in most Bacteroidetes and in tested cases the mature 16S rRNA includes the ASD. Yet, the Bacteroidetes ribosome seems blind to SD sequences in the cell, an apparent paradox that has defied explanation. Here, we show that *F. johnsoniae* ribosomes are intrinsically recalcitrant to SD recognition. A cryo-electron microscopy (cryo-EM) structure of the complete 70S ribosome from *F. johnsoniae* at 2.8 Å resolution shows that the ASD sequence is sequestered by ribosomal proteins (r proteins) bS21, bS18 and bS6. We also find that even though most genes lack SD sequences in the Bacteroidetes, *rpsU* and/or *rpsR* often contain strong SDs. These genes encode bS21 and bS18, respectively, the same proteins responsible for ASD occlusion, suggesting straightforward mechanisms of translational autoregulation.

4.3 Material and Methods

Purification of ribosomes and subunits

Ribosomes were purified from *F. johnsoniae* strain UW101 and *E. coli* strain MRE600. Cells were grown in rich media CYE (McBride and Kempf, 1996), *F. johnsoniae*; LB (Sezonov et al., 2007), *E. coli*] at optimum temperature (30°C, *F. johnsoniae*; 37°C, *E. coli*) to mid-logarithmic phase, cooled on ice, and pelleted. Ribosomes or individual subunits were then purified via conventional sedimentation methods as detailed previously (Qin et al., 2007) (Lancaster et al., 2002).

Determination of the complete coding region of the IF3 gene in *F. johnsoniae*

In *E. coli*, translation of the IF3 gene (*infC*) begins at the noncanonical start codon AUU (Butler et al., 1986) (Gold et al., 1984). This enables an autoregulatory circuit because IF3 reduces initiation from near-cognate start codons (Sacerdot et al., 1996) (Sussman et al., 1996). Based on its sequence, the IF3 gene of *F. johnsoniae* appears to be similarly autoregulated. The *infC* open reading frame extends 57 codons upstream of the (mis)annotated start codon (genome position 32,505) (McBride et al., 2009) and is predicted to encode conserved amino acids of IF3 (Supplementary Figure S1A). To elucidate the correct *N*-terminus of *F. johnsoniae* IF3, we first cloned the *infC* gene with ample upstream DNA into the *F. johnsoniae* expression vector pSCH710 (Baez et al., 2019), such that the resulting protein would be *C*-terminally tagged with hexahistidine (His6). *F. johnsoniae* cells carrying this plasmid were induced to overexpress IF3-His6, and the protein was partially purified via Ni-NTA (Qiagen) affinity chromatography. IF3-His6 was resolved using SDS-PAGE, isolated, and subjected to LC/MS-MS analysis (Mass Spectrometry and Proteomics Facility, Campus Chemical Instrument Center, Ohio State University). One of the peptides identified was SNRGFQPRVEKK, narrowing down the start codon to two possible near-cognate codons: ATA (genome position 32,340) or ATA (genome position 32,346). To distinguish

which of these codons corresponds to the natural start codon, we generated an *infC-gfp* fusion construct and targeted each candidate start codon by mutagenesis **(Supplementary Figure S4.1B)**. *F. johnsoniae* cells carrying the WT *infC-gfp* construct produced strong fluorescence (350,000 RFU), measured as described previously (Baez et al., 2019). Changing the upstream ATA codon (32,340) to GCT nearly eliminated fluorescence **(Mut_1; Supplementary Figure S4.1B)**, whereas changing the downstream ATA codon (32,346) to GCT failed to reduce fluorescence **(Mut_2; Supplementary Figure S4.1B)**. Thus, we infer that ATA (genome position 32,340) corresponds to the natural *infC* start codon. This assignment is consistent with the annotated ATG start codon of *infC* of *C. hutchinsonii* (Xie et al., 2007), another member of the Bacteroidetes **(Supplementary Figure S4.1A)**.

Overexpression and purification of initiation factors Supplementary Figure

Each initiation factor gene was amplified from the *F. johnsoniae* chromosome and cloned into pET28b (Novagen). Sequences of the primers used are listed in Supplementary Table S1. The resulting overexpression constructs—pBR6, pBR7, and pBR8—encode IF1 with no tag, IF2 with a *C*-terminal His6 tag, and IF3 with removable *N*-terminal His6 tag, respectively. The latter construct contains the complete coding region of *infC* (32,340-32,891), determined as described above, with codon 1 changed from ATA to ATG.

Plasmids pBR6, pBR7, and pBR8 were each transformed into *E. coli* strain BL21/DE3. Transformed cells were grown in LB at 37°C to mid-logarithmic phase, induced with IPTG (1 mM), and grown for an additional 5 hours to allow ample protein overproduction. In each case, the overproduced protein was purified from the soluble lysate. IF1 was purified using a cation exchange column (Macro-Prep High S, Bio-RAD) and dialyzed against storage buffer SBA (20 mM Tris-HCl pH 7.5, 100 mM NH_4Cl, 1 mM EDTA, 2 mM β-ME, 15% glycerol). IF2-His6 was

purified in three sequential chromatography steps. The protein was partially purified on a Ni-NTA (Qiagen) affinity column, further purified on a heparin column (HiTrap, GE Healthcare), passed through a size exclusion column (Superdex 75, GE Healthcare), and finally dialyzed against storage buffer SBB (20 mM Tris-HCl pH 8.0, 100 mM NH_4Cl, 1 mM EDTA, 2 mM β-ME, 5% glycerol). IF3 containing a *N*-terminal His6-tag was partially purified using Ni-NTA (Qiagen) and then subjected to thrombin cleavage. The cleavage reaction was performed in 20 mM Tris-HCl (pH 7.0), 300 mM NH_4Cl, 1 mM EDTA, and 2 mM β-ME at room temperature for 4 h, using 1U of thrombin (GE Healthcare) per 400 μg of protein. Following digestion, the reaction mixture was incubated with an excess of Ni-NTA resin at 4°C for 1 h, and the trimmed IF3 was recovered as unbound protein. IF3 was further purified using a cation exchange column (Macro-Prep High S, Bio-RAD) and dialyzed against storage buffer SBC (20 mM Tris-HCl pH 7.0, 300 mM NH_4Cl, 1 mM EDTA, 2 mM β-ME, 15% glycerol). Each protein was >95% pure as judged by SDS-PAGE. Small aliquots were flash-frozen and stored at -80°C.

E. coli initiation factors were overexpressed and purified as described (Qin and Fredrick, 2009) (Dallas and Noller, 2001).

Preparations of fMet-tRNA

A DNA template containing a T7 promoter upstream of the *F. johnsoniae* $tRNA^{fMet}$ gene was generated using overlapping primers 201467 and 201468 (listed in Supplementary Table S1) in a PCR reaction. This DNA template (25 μg/μL) was incubated with T7 RNA polymerase in a 1 mL reaction containing 5 mM ATP, 5 mM CTP, 5 mM GTP, 5 mM UTP, 60 mM Tris-HCl (pH 8.0), 25 mM $MgCl_2$, 1 mM spermidine, 30 mM DTT, and 0.01% Triton X-100 for 4 h at 37 °C. The $tRNA^{fMet}$ transcript was purified, charged, and formylated as described (Walker and Fredrick, 2008). Because T7 RNA polymerase strongly favors initiation with GTP, the 5' nucleotide of the

tRNA transcript was G2 (rather than C1). However, the absence of C1 caused no apparent problems for aminoacylation, formylation, or initiation complex formation. Pure native *E. coli* $tRNA^{fMet1}$ was purchased (Chemical Block Ltd) and charged and formylated as described previously (Walker and Fredrick, 2008).

Preparations of mRNA

Various model mRNAs from *F. johnsoniae* and *E. coli* were made by *in vitro* transcription, templated with BamHI-linearized plasmids, and gel purified (Supplementary Table S2). Model mRNAs from *F. johnsoniae* were named based on the gene (Fjoh) number, with an "m" prefix. To make each plasmid template, a region of DNA from the predicted transcriptional start site (Baez et al., 2019) to a position downstream of codon 30 was amplified and cloned into the EcoRI and BamHI sites of pUC19. A T7 promoter sequence was included in the 5' portion of the forward primer and, in some cases, a primer-binding site for primer 132 was included in the 5' portion of the reverse primer (see sequences in Supplementary Table S1). Model *E. coli* mRNAs derive from *gene 32* of bacteriophage T4 and have been described previously (Fredrick and Noller, 2002) (Shoji et al., 2006).

Toeprinting experiments

Initiation complexes were detected using the toeprinting technique (Hartz et al., 1989). Typically, 5' ^{32}P-labeled primer 132 or 201445 (see sequences in Supplementary Table S1) was annealed to mRNA (0.1 μM) in 50 mM Tris-HCl (pH 7.5), 110 mM NH_4Cl, and 30 mM KCl by incubating at 37°C for 5 min and placing on ice. $MgCl_2$ (7 mM), DTT (1 mM), GTP (0.1 mM), fMet-tRNA (2 μM), and ribosomes (1 μM) or heat-activated 30S subunits (1 μM) were then added, with or without IFs (2 μM each, as indicated) and further incubated for 1 h (without IFs) or 3 min (with IFs). To evaluate complexes formed, dNTPs (0.2 mM each) and AMV reverse transcriptase (2 U,

Life Sciences Advanced Technologies, Inc.) were added, and after 3 min, cDNA products were resolved by 6% denaturing PAGE. Gel imaging and quantification were performed with a Typhoon FLA 9000 phosphorimager (GE Healthcare) and associated software (ImageQuant 5.2).

The reaction conditions described above were established empirically for 70S IC formation, by systematically varying NH_4Cl concentration (50, 70, 110, 150, and 170 mM), $MgCl_2$ concentration (7 and 11 mM), and temperature (30°C, optimum growth temperature of *F. johnsoniae*; 37°C, optimum growth temperature of *E. coli*). Polymix buffer (Shoji et al., 2006) was also tested but did not enhance 70S IC formation. 30S IC formation was also compared at 30° C and 37° C, and little-to-no difference was seen.

The overall equilibrium association constant for 30S IC formation was measured using toeprinting as described previously (Qin et al., 2012) (Roy et al., 2018) **(Supplementary Figure S4.2)**. In these experiments, the concentration of mRNA (0.02 μM), subunits (0.1 μM, Eco; 1 μM Fjo), and factors (0.5 μM, Eco 30S case; 1.5 μM, Fjo 30S case) were held constant, while the concentration of fMet-tRNA was varied (0.05 to 1 μM). Binding reactions at 37°C were allowed to equilibrate for 1 h (absence of IFs) or 3 min (presence of IFs). The fraction of bound complex (F) was calculated as F = *toeprint*/(*toeprint*+*runoff*), where *toeprint* and *runoff* represent signal intensities of the corresponding bands. Data were plotted as a function of fMet-tRNA concentration and fit to the equation $F = F_{max}\ \{bc/(bc+1/K_A)\}$, where b is the input tRNA concentration, c is the input subunit concentration, and K_A is the equilibrium association constant. F_{max} corresponds to the maximal level of detected complex and presumably reflects the probability that the complex resists disruption by reverse transcriptase. Experiments to measure the extent of 70S IC formation as a function of fMet-tRNA concentration were performed in an analogous way, except that 70S ribosomes were used instead of 30S subunits.

Puromycin reactions

70S ICs were formed with either *F. johnsoniae* or *E. coli* components by incubating mRNA (0.5 μM), formyl-^{35}S-Met-tRNA^{fMet} (0.1 μM), heat-activated 30S subunits (1 μM), native IFs (1.5 μM each), and 50S subunits (3 μM) in the presence of 50 mM Tris-HCl (pH 7.5), 110 mM NH_4Cl, 30 mM KCl, 7 mM $MgCl_2$, 1 mM GTP, and 1 mM DTT for 10 min at 37°C. In each case, the complex formed was rapidly mixed with puromycin (1 mM) in a quench-flow machine (KinTek Laboratories, Inc.), and the reaction was quenched with 0.5 M KOH at various time points. The product, formyl-^{35}S-Met-puromycin (fMet-Pmn), was extracted with ethyl acetate and quantified in a liquid scintillation counter (Beringer et al., 2005).

Filter-binding experiments

Binding of RNA oligonucleotides to 30S subunits was measured using a filter-binding assay described previously (Fahlman and Uhlenbeck, 2004) (Shoji et al., 2009). Briefly, radiolabeled RNA (0.5 nM; 5'^{32}P-AGAAAGGAGGU-3', SD underscored; or 5'^{32}P-ACCUCCUUUCU-3', negative control) and activated 30S subunits (from *F. johnsoniae* or *E. coli*; various concentrations) were separately preequilibrated in 20 mM K-HEPES (pH 7.6), 100 mM NH_4Cl, 6 mM $MgCl_2$, 1 mM GTP, 4 mM β -ME at 25°C. At time t = 0, 100 μL of labeled RNA was mixed with 100 μL of 30S subunits. Aliquots (20 μL) were removed at various time points, filtered through a bi-layer of nitrocellulose and positively charged nylon membranes (NitroBind and Hybond-N+, GE Healthcare), and immediately washed with 150 μL of the same buffer. Membranes were dried, and radioactive spots were quantified with a phosphorimager to determine the fraction of RNA bound as a function of time. Apparent rates were plotted versus 30S concentration, and the data were fit to a linear equation to deduce k_{on} (slope) and k_{off} (Y intercept).

Determination of the mature ends of 16S, 23S, and 5S rRNA in F. johnsoniae

The mature ends of each rRNA molecule were identified using *F. johnsoniae* RNA-seq data collected previously (Baez et al., 2019). These RNA-seq libraries were made from 30-40 nucleotide RNA fragments, obtained after limited base hydrolysis of total *F. johnsoniae* RNA, without any rRNA-removal steps. Reads were mapped back to the genome, and 5'-read-end coverage and 3'-read-end-coverage were each plotted with respect to the corresponding genome-annotated gene ends. For each 5' and 3' coverage plot, we observed a large peak corresponding to the mature 5' and 3' ends, respectively. This enrichment is due to a higher probability that a small RNA fragment generated by limited hydrolysis comes from the terminus of the source molecule (Lalanne et al., 2018).

Mass Spectrometry

Purified *F. johnsoniae* ribosomes and 50S subunits were subjected to LC-MS/MS analysis as described (Zeng-Elmore et al., 2014).

Cryo-electron microscopy

Cryo-EM grids (c-flat CF-2/2-2C-T) were prepared by washing them in chloroform for two hours and treating them with glow discharged in air at 5 mA for 15 seconds, right before the sample was applied. A total volume of 3.6 µL of purified *F. johnsoniae* 70S ribosomes at a concentration of 170 nM in dilution buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 100 mM NH_4Cl, 6 mM b-mercaptoethanol) was deposited on the grid. The b-mercaptoethanol was added to the buffer right before the dilution was performed. Sample vitrification was performed in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific Inc.) using one blotting time for 3 seconds and with a blot force +1. The Vitrobot chamber was set to 25 ºC and 100% relative humidity. Data acquisition was performed using EPU software at FEMR-McGill using a Titan Krios microscope

at 300 kV equipped with a Falcon II direct electron detector (Thermo Fisher Scientific Inc.). Movies were collected with seven frames acquired in 1 second exposure at a magnification of 75,000x, producing images with a calibrated pixel size of 1.073 Å. The nominal defocus range used during data collection was between -1.25 to -2.75 μm, and the total dose used per movie was 50 $e^-/Å^2$.

Image processing

The movies comprising the *F. johnsoniae* 70S dataset were corrected for beam-induced motion using RELION's implementation of the MotionCor 2 algorithm (Zheng et al., 2017) (Zivanov et al., 2018). Estimation of the Contrast Transfer Function (CTF) was done using the Gctf program (Zhang, 2016). From here, all processing was done with RELION 3.0. A description of the image processing workflow is shown in Supplementary Figure S3. Particle images were selected and extracted from the micrographs using auto-picking and subsequently subjected to one cycle of reference-free 2D classification to remove false positive and damaged particles. The clean dataset output by the 2D classification contained 1,138,048 particle images. These images were subjected to 3D classification and subsequent refinement. The initial 3D reference used for the classification and refinements was a 60 Å low pass filtered map of the mature 70S subunit created from 4V4Q.cif (Schuwirth et al., 2005) using the Xmipp program (de la Rosa-Trevin et al., 2013) or the intermediate cryo-EM maps obtained during classification and refinement. To speed up the computer calculations, the 2D and 3D classifications were performed using particle images binned by 4 and a pixel size of 4.292 Å per pixel. However, we use the full-size images with a pixel size of 1.073 Å per pixel in the refinement steps. A soft-mask created with 'relion_mask_create' command was applied to all refinements. In this process, we extended the binary mask by two pixels and created a soft-edge with a width of ten pixels. Refinement was performed in three stages.

In the first stage we used the full-size particle images as they come out from 3D classification. In stage two of refinement, the particle images were first subjected to CTF refinement. In this process, we selected 'Fit per-particle defocus' as 'Yes' but selected 'No' for all the astigmatisms fits and phase-shift. We also did not correct for beam tilt estimation, as our data was not collected using this approach. In the last step particle images were also corrected using Bayesian polishing before subjecting them to final refinement. Bayesian polishing was performed using sigma values of 0.4515, 13545 and 0.21 for velocity, divergence and acceleration, respectively. Sharpening of the final cryo-EM maps and the local analysis was done with RELION (Zivanov et al., 2018). The average resolution for the structure of the *F. johnsoniae* 70S ribosome was estimated by gold-standard Fourier shell correlation. Resolution estimation is reported using a FSC threshold value of 0.143.

Map analysis and Atomic model building

The starting point of the structural modeling for the *E. coli* 70S ribosome was the atomic model of the *E. coli* 70S ribosome (PDB ID 4V4Q) (Schuwirth et al., 2005). The models for the 50S and 30S subunits were individually docked using rigid-body approaches into the obtained high-resolution cryo-EM map of the *F. johnsoniae* 70S ribosome using 'dock-in-map tool' in Phenix (Adams et al., 2010). The atomic model building and coordinate refinement for the *F. johnsoniae* 70S ribosome were performed with successive rounds of real space refinement in Phenix (Adams et al., 2010) and manual model building in Coot (Emsley and Cowtan, 2004, Emsley et al., 2010).

F. johnsoniae r proteins and rRNA sequences were obtained from UniProt (UniProt, 2019) and gene-NCBI databases, respectively. *F. johnsoniae* sequences were aligned with *E. coli* sequences using Multiple Sequence Alignment program ClustalW (Larkin et al., 2007). The conserved regions in the rRNA helices of the docked subunits were first located in the density map, and the

rest of the segments were traced from these anchor points. The numbering in the model obtained for *F. johnsoniae* 16S, 5S and 23S rRNA follow the *E. coli* numbering. The rRNA expansion segments and segments that differed from the *E. coli* 70S ribosome structure were built manually into the density using Coot (Emsley and Cowtan, 2004, Emsley et al., 2010).

The r proteins were also built using a similar approach. Conserved sequences in the homologous proteins were located in the density map, and the positions of bulky sidechains were used both to determine and validate the correct assignment. The structure of the 70S ribosome from *B. subtilis* (PDB ID 3J9W) (Sohmen et al., 2015) was used as a template to model the bL31 protein, as this protein is not present in the PDB 4V4Q, representing the *E. coli* 70S ribosome. Homologous r proteins were mutated to the correct sequences for *F. johnsoniae,* and specific regions were manually built into the density. Two unassigned regions of density corresponding to proteins were observed in the EM map. One of the areas was a well-resolved protein density, and it was possible to unambiguously build the *F. johnsoniae* protein with unknown function annotated with GenBank ID: ABQ04565.1. The resolution of the second area of density was lower. However, after identification of the Fjoh_4981 gene product by mass spectrometry as a new component of the *F. johnsoniae* 70S ribosome, it was possible to fit the sequence of the protein encoded by this gene into this second density in the map.

Finally, the atomic coordinates of the 50S and 30S models were combined to obtain the entire 70S. After fixing the regions of intersubunit contacts, the model was refined using Phenix real-space and B-factor refinement (Afonine et al., 2018b). The final atomic coordinates of the *F. johnsoniae* 70S ribosome were validated using Phenix cryo-EM Comprehensive validation tool (Afonine et al., 2018a, Williams et al., 2018) and the Molprobity server (Williams et al., 2018) **(Supplementary Table S4.3)**.

Graphical representations of molecular models and cryo-EM density maps were generated using PyMOL (DeLano, 2002), UCSF Chimera (Pettersen et al., 2004), and UCSF ChimeraX (Goddard et al., 2018).

Secondary structure diagram of the rRNA

The secondary structure diagrams of the 16S, 5S, and 23S rRNA were prepared by extracting base pairs information from the refined model using DSSR (Lu et al., 2015) and were drawn in VARNA (Darty et al., 2009).

Multiple sequence alignments

Peptide sequences for bS21, bS18, and bS6 were retrieved from the NCBI database **(Supplementary Table S4.4)**. For each phylum, the organisms selected were diverse and well distributed. Sequences were analyzed in UGENE (Okonechnikov et al., 2012) using the MAFFT (Katoh et al., 2002) multiple sequence alignment with default settings. Gaps were removed when present in 33% or more of the aligned sequences. Sequence logos were generated from edited MAFFT alignments using the WebLogo 3 application (Crooks et al., 2004). Logos were aligned manually, based on alignments of phylum consensus sequences, generated in UGENE.

Comparisons of bS21 C-terminal domains

The isoelectric point (pI) of the C-terminal domain (CTD; defined as residues C-terminal of the conserved KPS/T motif) was determined using the Isoelectric Point Calculator (Kozlowski, 2016). Because these are short peptide sequences (< 40 residues), the IPC2 peptide scale was used.

Detection of SD sequences in the Bacteroidetes

A diverse and representative collection of almost 300 Bacteroidetes was chosen for analysis **(Supplementary Table S4.5)**. Genomes and annotations were obtained from NCBI using the

Assembly search feature (Kitts et al., 2016) from BioProject numbers used in Hahnke *et al*. (Hahnke et al., 2016). To determine which genes contained SD sequences, sections of the TIR were analyzed by the free2bind suite of programs (Starmer et al., 2006). Relevant sections of the TIR for SD detection were defined such that the spacing between the 5' end of the ASD and the first nucleotide of the start codon was between 0 and 15 nucleotides, as described (Salis et al., 2009). Sequences were fed to the free_scan.pl program in the free2bind suite, which calculates the free energy of pairing between two RNA strands as a function of register / relative position. The ASD sequence 3'-UUCCUCCA-5' was tested incrementally across each TIR, and if the ASD paired with a free energy of (for example) -7 kcal/mol or lower at any position, a SD sequence was assigned. As a control, the same analysis was performed on a window 25-40 nt upstream of the start codon, too distal for an authentic SD. The resulting "hits" were termed mock SD (MSD) sequences, which served as a proxy for false positive SD assignments in the various organisms analyzed **(Supplementary Table S4.5)**. An analogous approach was taken to screen for SD2 and MSD2 sequences, using the ASD2 sequence 3'-UUACUCUA-5'.

Data Availability

The electron microscopy map and the model for the F. johnsoniae 70S ribosome have been deposited in the Electron Microscopy Data Bank (EMDB) and in the Protein Data Bank (PDB) with accession codes EMD-22345 and PDB 7JIL, respectively.

4.4 Results

Reconstitution of *Flavobacterium johnsoniae* initiation *in vitro*

We purified *F. johnsoniae* and *E. coli* components of translation initiation (ribosomes, subunits, IF1, IF2, IF3, fMet-tRNA^{fMet}, various mRNAs) to compare complex formation in the two systems. Using *F. johnsoniae* components and the toeprinting technique, we observed efficient 70S initiation complex (IC) formation at the cognate start codon of multiple tested mRNAs **(Figure 4.1A)**. Virtually identical results were seen with either fully modified *E. coli* fMet-tRNA^{fMet1} or unmodified *F. johnsoniae* fMet-tRNA^{fMet} **(Supplementary Figure S4)**. Once formed, 70S ICs of *F. johnsoniae* and *E. coli* exhibit comparable puromycin reactivity **(Supplementary Figure S5)**, indicating that fMet-tRNA is bound in the P/P site. In all tested cases, IC formation was strictly factor dependent. Surprisingly, 30S ICs were not readily detected with *F. johnsoniae* components in this set of experiments. However, replacement of *F. johnsoniae* factors with *E. coli* factors enabled 30S IC formation, an effect attributable to *E. coli* IF3 specifically **(Figure 4.1B)**. Systematic mix-and-match experiments showed that swapping *F. johnsoniae* IF3 with *E. coli* IF3 is sufficient to stabilize the 30S IC **(Supplementary Figure S6)**. The reason *F. johnsoniae* IF3 failed to stabilize the *F. johnsoniae* 30S IC remains unclear. This preparation of *F. johnsoniae* IF3 was active, based on its ability to promote *F. johnsoniae* 70S IC formation **(Supplementary Figure S7)**.

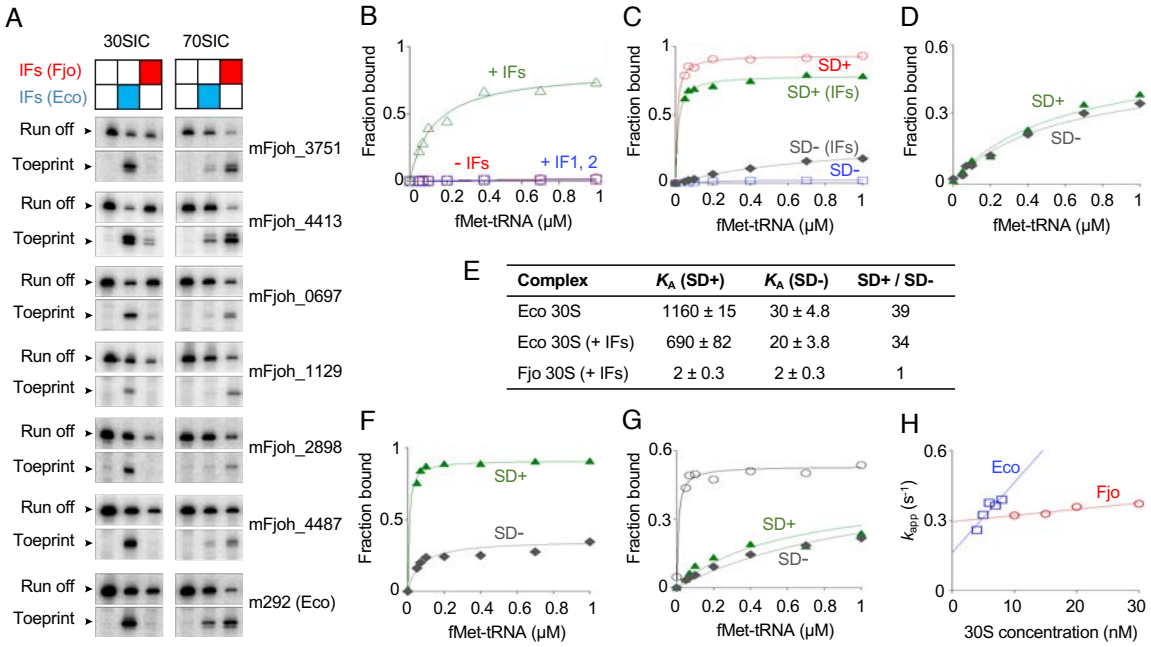


Figure 4. 1 Biochemical evidence that E. coli and F. johnsoniae ribosomes differ in ASD function.

(A) F. johnsoniae 30S and 70S complexes formed on various mRNAs in the presence of initiation factors (IFs) from F. johnsoniae (Fjo) or E. coli (Eco) as indicated. The toeprint band indicates the cognate complex (with start codon in P site), and the run-off band reflects unbound mRNA. Messages mFjoh_xxxx, are named based on corresponding F. johnsoniae genes; m292, a model E. coli mRNA(Shoji et al., 2006). (B-D) Experiments to measure the overall equilibrium association constant (K_A) for 30S IC formation. For each experiment, the amount of complex formed was quantified and plotted as a function of fMet-tRNA concentration. (B) F. johnsoniae 30S subunits (1 μM) were incubated with mFjoh_4413 and fMet-tRNA (various concentrations, as indicated), in the presence of all E. coli factors (+IFs, 1.5 μM each; open green triangles), IF1 and IF2 only (+IF1,2; open blue squares), or no factors (-IFs; open red circles). (C) E. coli 30S subunits (0.1 μM) were incubated with mRNA (m291, SD+ (Shoji et al., 2006); or its derivative m295, SD-; as indicated) and fMet-tRNA (various concentrations, as indicated), in the absence or presence of E. coli factors (0.5 μM each), as indicated. Messages m291 and m295 are identical except that the latter has cytosines in place of guanines at positions -9 and -10. (D) F. johnsoniae 30S subunits (1 μM) were incubated with mRNA (m291, SD+; or m295, SD-; as indicated) and fMet-tRNA (various concentrations, as indicated), in presence of E. coli factors (1.5 μM each). (E) Summary of K_A values (in units of μM^{-2}) obtained from experiments like those of panels C-D. (F) 70S ribosomes (0.1 μM) from E. coli were incubated with mRNA (m291, SD+; or m295, SD-; as indicated) and fMet-tNA (various concentrations, as indicated) in presence of E. coli initiation factors (0.5 μM each). (G) 70S ribosomes (1 μM) from F. johnsoniae were incubated with mRNA (m291, SD+, filled green triangles; m295, SD-, filled gray diamonds; or mFjoh_4413, open black circles) and fMet-tRNA (various concentrations, as indicated) in the presence of F. johnsoniae initiation factors (1.5 μM each). (H) The rate of binding of a SD-containing RNA oligonucleotide to the 30S subunit was measured using a double-membrane filtration method (Shoji et al., 2006). The apparent rate (k_{app}) was plotted versus 30S concentration (E. coli, open blue squares; F. johnsoniae, open red circles), to estimate k_{on} (E. coli, 30 $\mu M^{-1}s^{-1}$; F. johnsoniae, 2.7 $\mu M^{-1}s^{-1}$) and k_{off} (E. coli, 0.16 s^{-1}; F. johnsoniae, 0.29 s^{-1}).

Flavobacterium johnsoniae Ribosomes Fail to Recognize the Shine Dalgarno Sequence of mRNA

The overall equilibrium association constant for 30S·mRNA·fMet-tRNA formation (K_A) was measured as described previously (Qin et al., 2012) (Roy et al., 2018), using message m291 (SD+) and its variant m295 (SD-) (Fredrick and Noller, 2002). For *E. coli* 30S, the SD increased K_A by 30- to 40-fold in either the absence or presence of factors **(Figure 4.1C & 4.1E)**. In contrast, for *F. johnsoniae* 30S, the SD had virtually no effect on K_A **(Figure 4.1D and 4.1E)**. Because the latter experiment involved (by necessity) *E. coli* factors, we also measured the extent of 70S IC formation as a function of fMet-tRNA concentration, employing native factors in each case. Again, we observed that the SD promotes complex formation in the *E. coli* case but not the *F. johnsoniae* case **(Figure 4.1F & 4.1G)**. *F. johnsoniae* ribosomes generally exhibited lower efficiency in complex formation in these experiments than *E. coli* ribosomes. This is due at least in part the model mRNAs (derived from *E. coli*) because IC formation is considerably more efficient on mFjo_4413 (Figure 4.1G, open circles; K_A = 130 μM^{-2}, F_{max} = 0.52). Collectively, these data suggest that *F. johnsoniae* ribosomes are inherently refractory to SD recognition.

To further compare ASD function in *E. coli* versus *F. johnsoniae*, we used a double-membrane filtration method (Fahlman and Uhlenbeck, 2004) (Shoji et al., 2009) to measure binding between 30S subunits and an RNA oligo containing a "perfect" SD (5'-AGAAAGGAGGU-3'; SD underscored). Subunits from either organism were able to bind this oligo, but the kinetics of binding differed **(Figure 4.1H & Supplementary Figure S8)**. The rate of binding (k_{on}) was 11-fold larger for *E. coli* 30S, and the dissociation rate (k_{off}) was ~2-fold smaller. These data suggest that the ASD in *F. johnsoniae* 30S is constrained in some way, making mRNA pairing slower and less favorable.

The Structure of the *Flavobacterium johnsoniae* 70S ribosome

Ribosomes were purified from *F. johnsoniae* strain UW101 and their structure solved by cryo-EM and single particle techniques. Image classification revealed that all the ribosomes in the sample were empty and did not contain any tRNA molecule bound to their A, P or E site. The 30S subunit in all the ribosomal particles was in the same ratcheting position and exhibited the same rotation angle with respect to the 50S subunit, indicating that the sample represented a conformationally homogeneous population of 70S ribosomes **(Supplementary Figure S3)**. The cryo-EM map **(Figure 4.2A & 4.2B; Supplementary Video 1)** refined to 2.8 Å resolution **(Supplementary Figure S9)**. Local resolution analysis revealed that the core of the ribosomal particle was the most defined region reaching a resolution of 2.5 Å. Peripherical regions that are known to be more dynamic, including the beak and shoulder in the 30S subunit and the two stalks in the 50S subunit showed a resolution ranging between 3-3.8 Å.

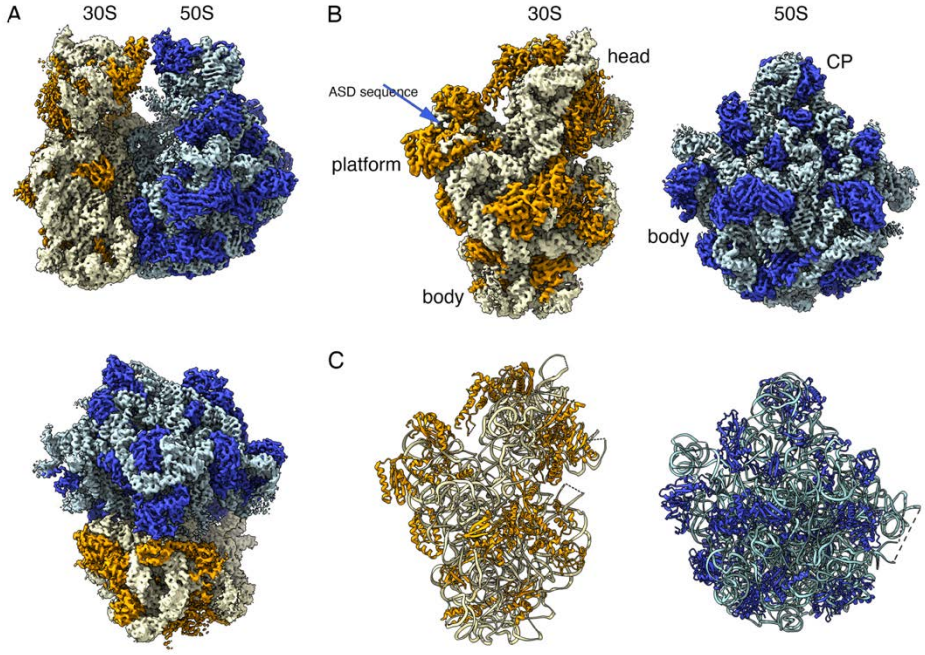


Figure 4. 2 Structure of the Flavobacterium johnsoniae ribosome.

(A) Side and top views of the cryo-EM map of the F. johnsoniae ribosome at 2.8 A resolution. The rRNA and r proteins in the 30S subunit are shown in yellow and orange, respectively. In the 50S subunit, the rRNA is colored in pale blue (rRNA) and the r proteins in dark blue. (B) Solvent view of the 30S and 50S subunit forming the F. johnsoniae ribosome. The main landmarks of the 30S and 50S subunits and the ASD sequence in the small subunit are indicated. CP indicates the central protuberance. (C) Solvent view of the molecular models derived for the 30S and 50S subunits derived from the cryo-EM map of the F. johnsoniae ribosome. The rRNA and r proteins and panel B and C are colored following the same scheme as in panel A.

The quality of the electron density map enabled us to build a molecular model of the 70S ribosome with confidence **(Figure 4.2C and Supplementary Table S4.3)**. We were able to unambiguously distinguish most of the nucleobases in the rRNA helices and side chains in the r proteins, respectively **(Supplementary Figure S4.10)**. A few regions of the rRNA and r proteins in both 50S and 30S subunits, in particular, the regions which are known to be flexible, had either weaker or no density and were therefore not built in the model. In line with common convention, helices of the small and large subunits were denoted with a lowercase "h" and uppercase "H", respectively.

To model the rRNA, we complemented the information from the density map with RNA-seq data that unambiguously identified the mature 5' and 3' ends of each rRNA molecule **(Supplementary Figure S4.11)**. With one exception (5S, 3'), the termini differed from the annotated-genome predictions and showed the lengths of mature 16S, 23S, and 5S rRNA to be 1519, 2862, and 111 nucleotides, respectively. Notably, the mature 23S rRNA of *F. johnsoniae* lacks nucleotides 1-8 and nucleotides 2895-2902, which in other bacteria pair to form helix H1. Based on the genomic sequence, RNA strands corresponding to nucleotides 2-7 and 2895-2900 are complementary. Hence, a short (6 base-pair) helix 1 may form in the precursor rRNA during the 50S assembly and then be cleaved off during final maturation.

We were able to model two additional proteins. One of them is located in the 30S subunit and the other in the 50S subunit (see description below). The final three-dimensional structure of the *F. johnsoniae* 70S ribosome contains all three rRNAs (16S in the 30S subunit and 23S and 5S in 50S subunit), 20 r-proteins in the 30S subunit and 30 r-proteins in the 50S subunit (Supplementary Table S6 & S7). No density was observed for bS1 and uS2 in the small subunit. Association of bS1 in most species is weak and reversible (Subramanian and van Duin, 1977) and uS2 frequently dissociates upon exposure to the air-water interface during sample vitrification in cryo-EM

(Jahagirdar et al., 2020). In the large subunit, density representing uL1, uL7/L12, uL10 and uL11 was also not present or highly fragmented due to the intrinsic flexibility of these ribosome components. Therefore, these proteins were excluded from the model.

To gain some insight on the basis of the missing small subunit proteins, we compared the protein composition of isolated subunits and ribosomes using tricine SDS-PAGE **(Supplementary Figure S4.12)**. Protein bS1 is clearly underrepresented in the 30S subunits and is further underrepresented in the 70S ribosomes. Protein uS2 is present at stoichiometric levels in the 30S subunits but is somewhat underrepresented in the 70S ribosomes. One difference between the purification methods was a second high-salt wash in the 70S case, which likely explains the larger degree of protein depletion observed. We infer that loss of bS1 occurred mainly during the 70S purification, while loss of uS2 occurred during purification and sample vitrification.

Structural differences in the 30S subunit

Overall, the structure of *F. johnsoniae* 70S ribosome is very similar to other bacterial ribosomes **(Figure 4.2 and Supplementary Video 4.1)**. However, several structural motifs in both the 30S and 50S subunits displayed differences compared to other known 70S ribosome structures.

Comparison of the 30S subunit from our 70S structure with the 30S subunit from 70S structures from *E. coli* (Kaledhonkar et al., 2019), *Thermus thermophilus* (Polikanov et al., 2015), *Mycobacterium smegmatis* (Hentschel et al., 2017), *Pseudomonas aeruginosa* (Halfon et al., 2019), *Bacillus subtilis* (Sohmen et al., 2015), *Staphylococcus aureus* (Khusainov et al., 2017) **(Supplementary Table S4.8)** revealed structural differences in the 16S rRNA helices highlighted in Figure 4.3A. Significant differences were observed between helices h7-h10. Whereas helices h7-h8 were similar in length in all the structures, helices h9 and h10 were different, both in length and fold, compared to *E. coli* (Figure 4.3B) and other ribosomes **(Supplementary Figure S4.13A)**.

Helix h9 is much longer in *F. johnsoniae*, *T. thermophilus*, *M. smegmatis*, *B. subtilis*, and *S. aureus* than in *E. coli* and *P. aeruginosa*; whereas helix h10 is shortest in *F. johnsoniae* and *M. smegmatis* compared to all others. Overall, helices h9 and h10 in *F. johnsoniae* resemble more closely to *M. smegmatis*. We also observed subtle differences in helix h17, which is comparatively shorter in *F. johnsoniae* than in *E coli* **(Figure 4.3C)** and others, except in *M. smegmatis*, which has the shortest helix h17 **(Supplementary Figure S4.13B)**. Apparent differences were also present in helices h26 and h44, where these helices are the shortest and less extended in *F. johnsoniae* compared to those in *E. coli* **(Figure 4.3D and 4.3E)** and other species **(Supplementary Figure S4.13C & 4.13D)**. Whether any of these structural variations have functional consequences remains unclear.

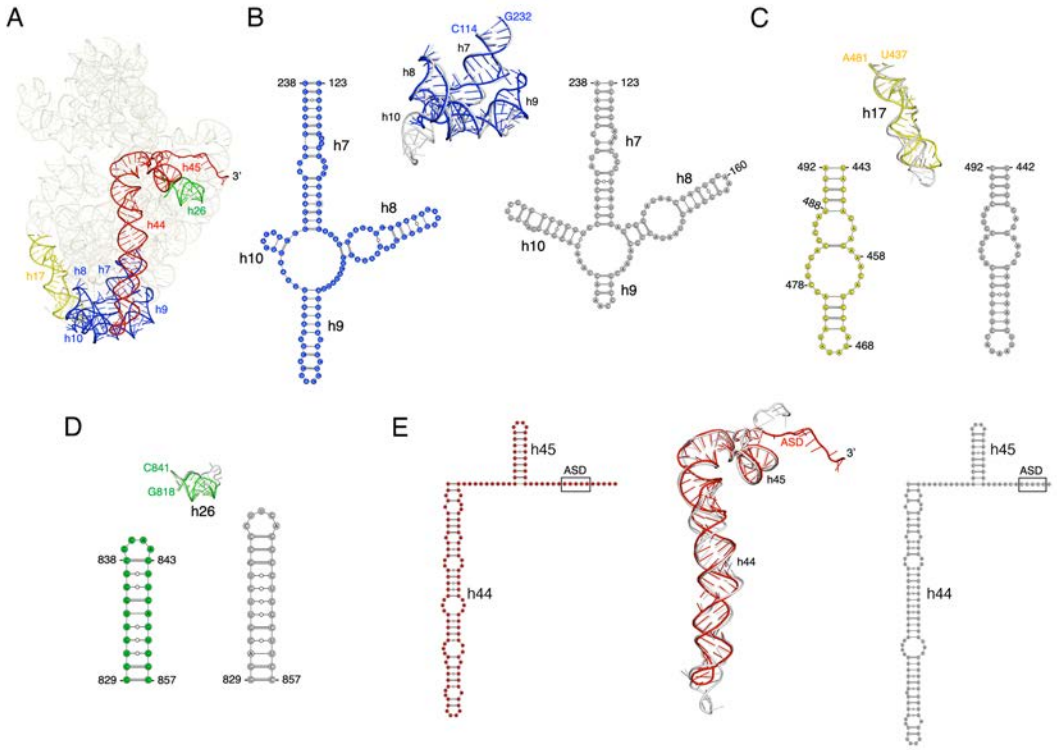


***Figure 4. 3 Comparison of structural differences in the 16S rRNA between* F. johnsoniae *and* E. coli.**

*(**A**) Tertiary structure of 16S rRNA from our refined model. The regions of structural variability (in terms of length, fold, and/or loop size) between* F. johnsoniae *and* E. coli *(PDB ID: 6O9K), are labelled and highlighted in different colors. (**B–E**) Superpositions of h7–10, h17, h26, and h44–45, respectively from* F. johnsoniae *and* E. coli *along with their secondary structure diagram. Color codes for* F. johnsoniae *16S rRNA elements are same as in (A), and* E. coli *motifs are shown in grey. Comparisons with* T. thermophilus, M. smegmatis, P. aeruginosa, B. subtilis *and* S. aureus *are shown in Supplementary Figure S4.13.*

Identification of a structural homolog of bS22

In the process of modelling the *F. johnsoniae* ribosome, we identified an unassigned area of density in the cryo-EM map in the 30S subunit. The clarity of density allowed us to build a *de novo* molecular model for this region **(Figure 4.4A)**. The primary sequence best matched gene Fjoh_1533, annotated to encode a hypothetical protein. The protein was comprised of 30 amino acids and was highly basic with a pI of 12.7, owing to its high arginine and lysine content. The amino acid sequence of this protein is highly conserved among members of Bacteroidetes **(Figure 4.4D)**.

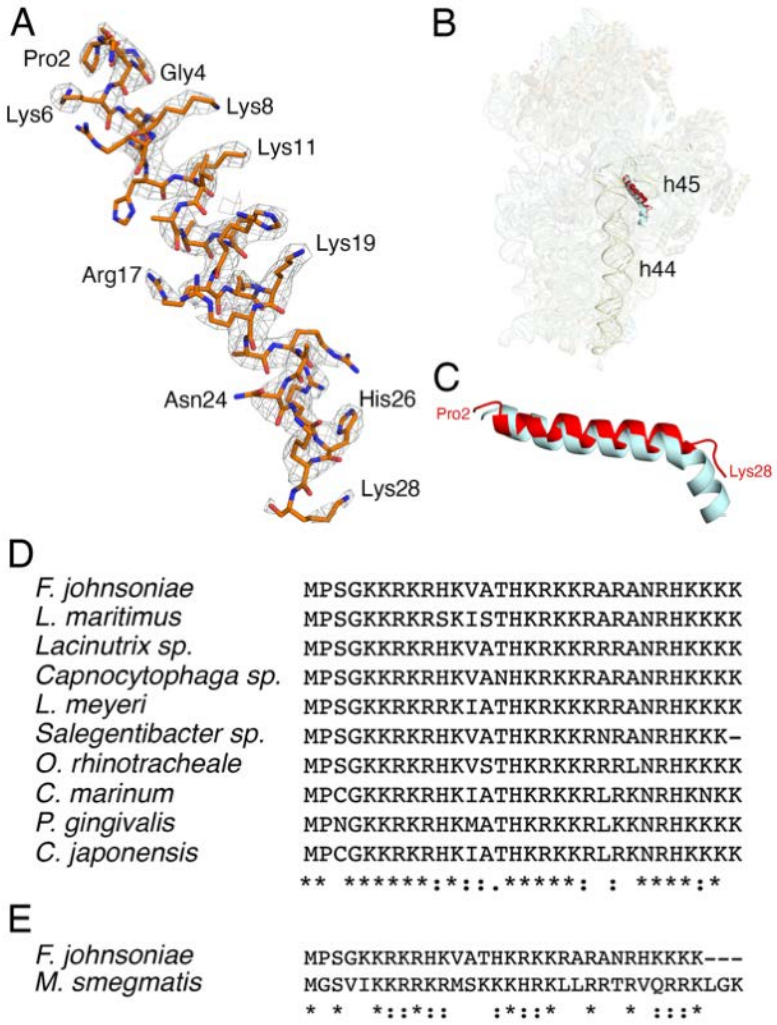


Figure 4. 4 Homolog of bS22 in the 30S subunit of F. johnsoniae ribosome.

*(**A**) Electron density map and the derived molecular model of the Fjoh_1533 protein. (**B**) Structural comparison of the Fjoh_1533 protein in F. johnsoniae (highlighted in red) and bS22 in M. smegmatis (cyan) in the 30S subunit. Helices h44 and h45 in F. johnsoniae are highlighted in light yellow. (**C**) close-up view of C, showing the structural homology between the two proteins. (**D**) Amino acid sequence alignment of the uncharacterized protein in various Bacteroidetes, as indicated. (**E**) Amino acid sequence alignment of the uncharacterized protein of F. johnsoniae and bS22 of M. smegmatis.*

Upon comparison with other ribosome structures, we found that protein Fjoh_1533 occupies the same position as bS22 in the *M. smegmatis* ribosome (Hentschel et al., 2017) **(Figure 4.4B)**. The

protein is similarly inserted between helices h44 and h45 of the 16S rRNA and in contact with helix H70 of the 50S subunit. Both proteins superimposed well and form an alpha helix **(Figure 4.4C)**. Considering the structural homology between the two proteins **(Figure 4.4E)**, we infer that this *F. johnsoniae* protein is the functional counterpart to bS22 in *M. smegmatis* (Hentschel et al., 2017) and hence name it bS22. Notably, SD sequences are quite prevalent in Mycobacteria (Nakagawa et al., 2010), arguing against any role for bS22 in ASD occlusion.

Sequestration of the ASD Sequence within the 30S subunit by ribosomal proteins bS21, bS18, and bS6

In all the structures previously obtained for the 30S subunit and 70S ribosome without bound mRNA, the 3' end of the 16S rRNA, including the ASD region, remains flexible and is mostly unmodelled. In our structure, the ASD sequence presented an apparent density, and we were able to trace the entire 3' sequence of the 16S rRNA with confidence **(Figure 4.5A and 4.5B)**. Only the bases of C1535 and C1536 exhibited weak electron density, so the assigned orientation of these bases in the current model might be variable.

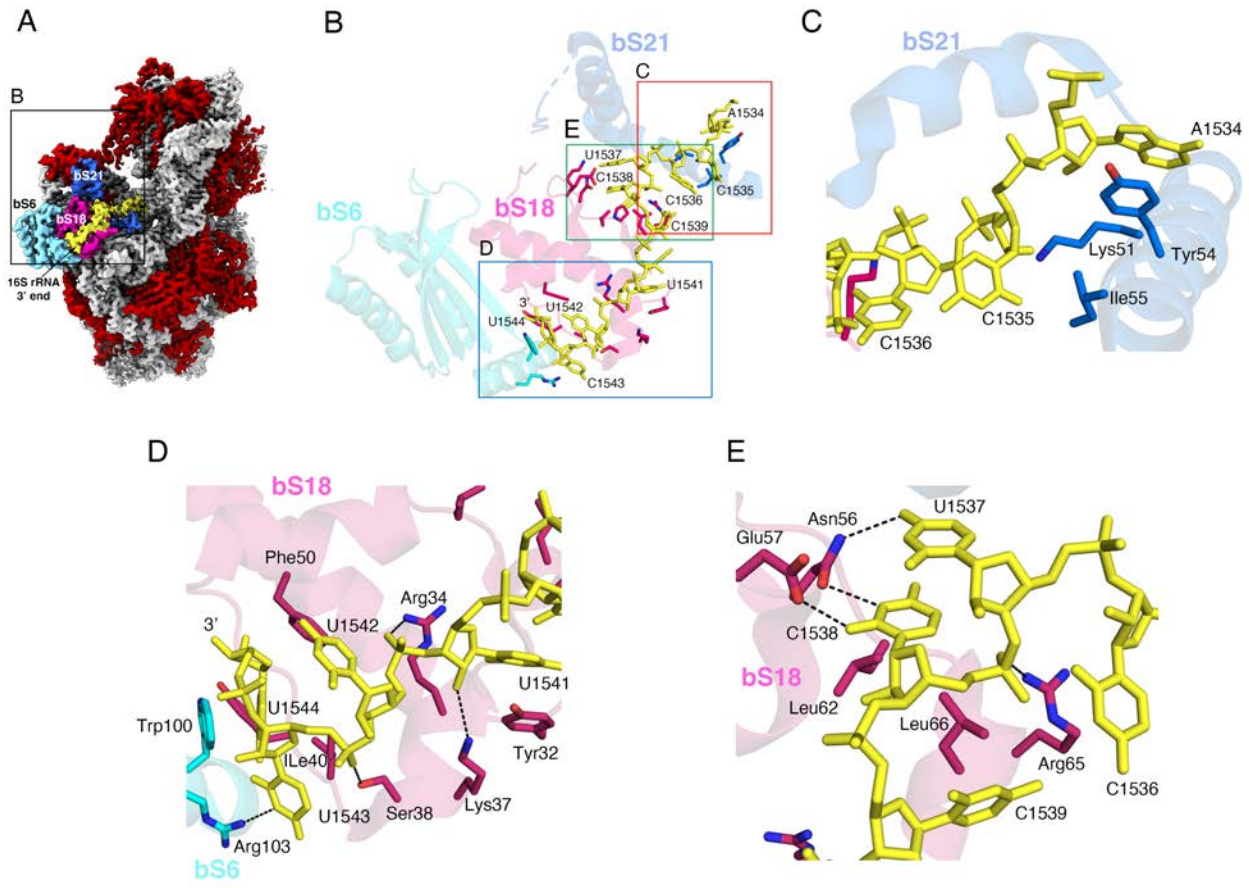


Figure 4.5 Structural Basis for the Sequestration of the ASD sequence within the F. johnsoniae 30S subunit.

*(**A**) Solvent view of the cryo-EM structure of the F. johnsoniae 30S subunit showing the 3' end segment of 16S rRNA (highlighted in yellow). The remaining 16S rRNA is shown in light grey. The r proteins bS6, bS18 and bS21 involved in the sequestration of the SD sequence are colored in cyan, pink and blue, respectively. The remaining r proteins are shown in red. Zoomed area shown in (B) is framed in the structure. (**B**) Interaction details between the 3' end nucleotides (including the ASD sequence) of 16S rRNA (yellow sticks) and r proteins bS6 (cyan), bS18 (pink), and bS21 (dark blue). Zoomed areas shown in (**C**), (**D**) and (**E**) are indicated in panel (B).*

We found that the 3’ end of the 16S rRNA, including the ASD sequences folds back away from the mRNA exit channel and binds along a groove formed by r proteins bS6 and bS18 **(Figure 4.3E & Figure 4.5A)**. The terminal 3’ nucleotides are stabilized through an extensive network of H-bonds and van der Waals contacts by the residues from bS6, bS18, and bS21 **(Figure 4.5B-4.5E)**. For example, A1534 and U1541 are stabilized through stacking interaction with the side chains of

Tyr54 and Tyr32 from bS21 and bS18, respectively **(Figure 4.5C and 4.5D)**. Similarly, the base and sugar of U1544 stack with the side chains of Trp100 from bS6 and Phe50 from bS18, respectively **(Figure 4.5D)**. In addition, the side chains of Ser38, Arg34, Lys37, Asn56, Glu57, and Arg65 from bS18 and Arg103 from bS6 form H-bonds with U1537, C1538, U1541, U1542, and C1543. The van der Waals contacts from the side chains of Ile40, Leu62, and Leu66 from bS18 and Ile55 from bS21 contribute to further stabilize the terminal 3' nucleotides **(Figure 4.5C-4.5E)**.

Multiple residues contributing to ASD occlusion are uniquely conserved in the Bacteroidetes

To further evaluate the ASD-binding site of the Bacteroidetes ribosome, we compared the conservation of amino acids in bS21, bS18, and bS6. Sequences from many representative species of the Bacteroidetes were aligned, as were those from the Firmicutes, g-Proteobacteria, and Spirochaetae **(Figure 4.6, Supplementary Figure S4.14, Supplementary Table S4.4)**. Organisms of these latter groups exhibit high SD prevalence, unlike the Bacteroidetes (Nakagawa et al., 2010). Comparisons of the resulting logos/consensus sequences revealed features of all three proteins unique to the Bacteroidetes.

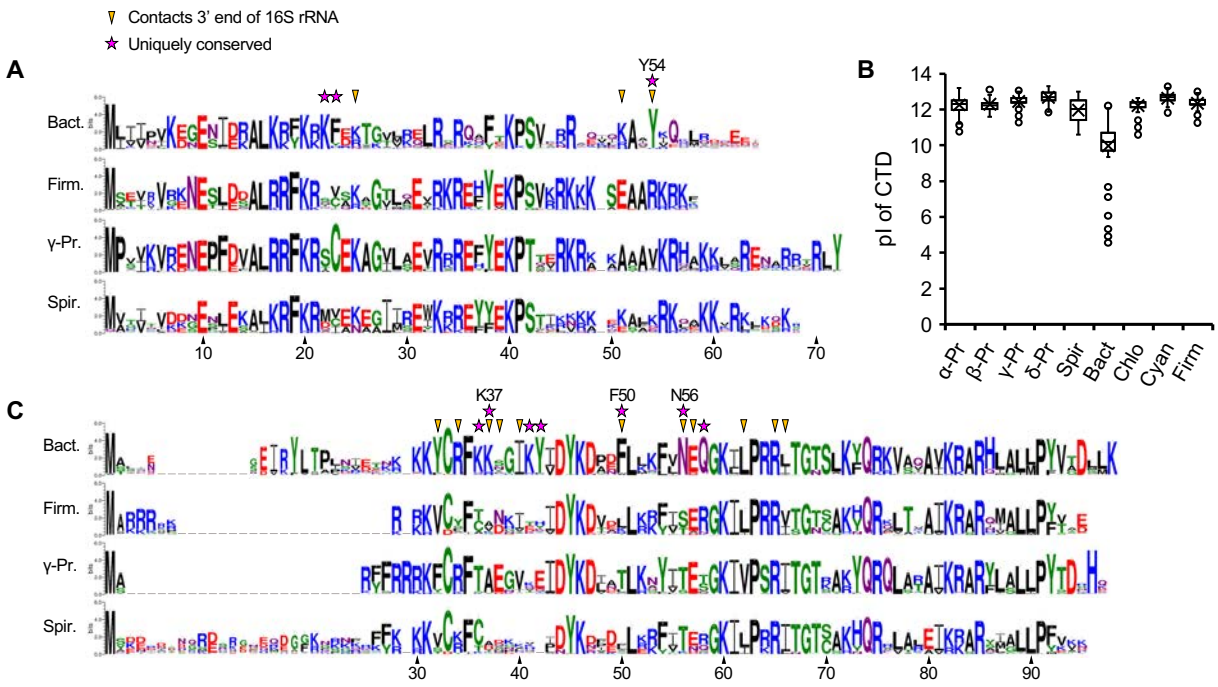


Figure 4. 6 Unique features of bS21 and bS18 of the Bacteroidetes.

*(**A**) Sequence logos for bS21 of the Bacteroidetes (Bact, n = 357), Firmicutes (Firm, n = 300), γ-Proteobacteria (γ-Pr, n = 247), and Spirochaetes (Spir, n = 112). Numbering based on F. johnsoniae. Arrowheads indicate residues that contact the 3' end of 16S rRNA in the F. johnsoniae 70S structure. Stars indicate residues uniquely conserved in the Bacteroidetes, defined as any identical or similar (K/R; F/Y; S/T; D/E; N/Q) residue that occurs with a frequency of >0.9 in the Bacteroidetes and < 0.5 in each of the other groups analyzed. (**B**) Box plot comparing the isoelectric point (pI) of the C-terminal domain (CTD) of bS21 (residues C-terminal of S42). Based on two-tailed t tests, the Bacteroidetes data clearly differ from those of each other group ($P < 10^{-70}$). α-Pr, α-Proteobacteria (n = 516); β-Pr, β-Proteobacteria (n = 172); γ-Pr, γ-Proteobacteria (n = 247); δ-Pr, δ-Proteobacteria (n = 164); Spir, Spirochaetes (n = 112); Bact, Bacteroidetes (n = 357); Chlo, Chorobi (n = 62); Firm, Firmicutes (n = 300); Cyan, Cyanobacteria (n = 312). (**C**) Sequence logos for bS18, annotated as described above. Bact, n = 317; Firm, n = 299; γ-Pr, n = 239; Spir, n = 111.*

For bS21, the C-terminal portion of the protein is considerably less basic than in other organisms and contains a signature tyrosine (Tyr54, *F. johnsoniae* numbering), highly conserved in the phylum **(Figure 4.6A & 4.6B)**. In the *F. johnsoniae* 70S structure, Tyr54 stacks onto A1534 of 16S rRNA **(Figure 4.5C)**. This interaction occurs right where the 3' tail begins to diverge from that seen in the *E. coli* and *T. thermophilus* ribosome and may be crucial for reorienting the ASD towards the platform in the Bacteroidetes ribosome. For bS18, residues uniquely conserved in the

Bacteroidetes include Phe50, Asn56, Gln58, and Lys/Arg at position 37 **(Figure 4.6C)**. In the *F. johnsoniae* ribosome structure, Asn56 contacts U1537 and C1538 **(Figure 4.5E)**, Lys37 contacts U1541, and Phe50 contacts U1544 **(Figure 4.5D)**. bS6 of the Bacteroidetes contains a phylum-specific C-terminal extension, which folds into an alpha helix in the *F. johnsoniae* ribosome **(Supplementary Figure S4.14)**. Three residues within this helix (Arg103, Trp100, and Lys107) interact with C1543 and U1544, the last nucleotides of *F. johnsoniae* 16S rRNA **(Figure 4.5D)**.

To evaluate whether the described protein-RNA interactions stabilizing the ASD sequence in *F. johnsoniae* could potentially occur in species of other phyla, we structurally aligned bS6, bS18 and bS21 of *E. coli* and *T. thermophilus* onto the structure of the *F. johnsoniae* ribosome. We found that critical interactions contributing to ASD occlusion in *F. johnsoniae* (including those established through Tyr54 of bS21; Phe50, Asn56 and Lys37 of bS18; and Arg103 and Trp100 of bS6) cannot occur with the residues occupying the structurally equivalent positions in the *E. coli* or *T. thermophilus* ribosome **(Supplementary Figure S4.15)**.

Structural differences in the 50S subunit

The most notable differences in 23S rRNA were observed in helices H10, H16-H17, H28, H63 and H98 **(Figure 4.7A & Supplementary Figure S4.16)**, which deviate in length, fold, and/or orientation compared to six other species. Helices H10 and H63 are the shortest in *F. johnsoniae* **(Figure 4.7B & 4.7D, and Supplementary Figure S4.16A & S4.16B)**. Helix H28 is longer in *F. johnsoniae* and *T. thermophilus* than in *E. coli* and in all the other structures we compared **(Figure 4.7C & Supplementary Figure S4.16C)**. Helices H16-H17 adopt a different conformation from the equivalent helices in *E. coli* **(Figure 4.7E)**. The overall orientation of helices H16-H17 in *F. johnsoniae* was somewhat similar to that in the other four structures. However, obvious differences do exist. For example, *P. aeruginosa* contains the shortest and *M. smegmatis* the longest of these

helices **(Supplementary Figure S4.16D)**. Finally, the length of H98 in *F. johnsoniae* is much reduced, essentially eliminating the helix **(Figure 4.7F & Supplementary Figure S16E)**.

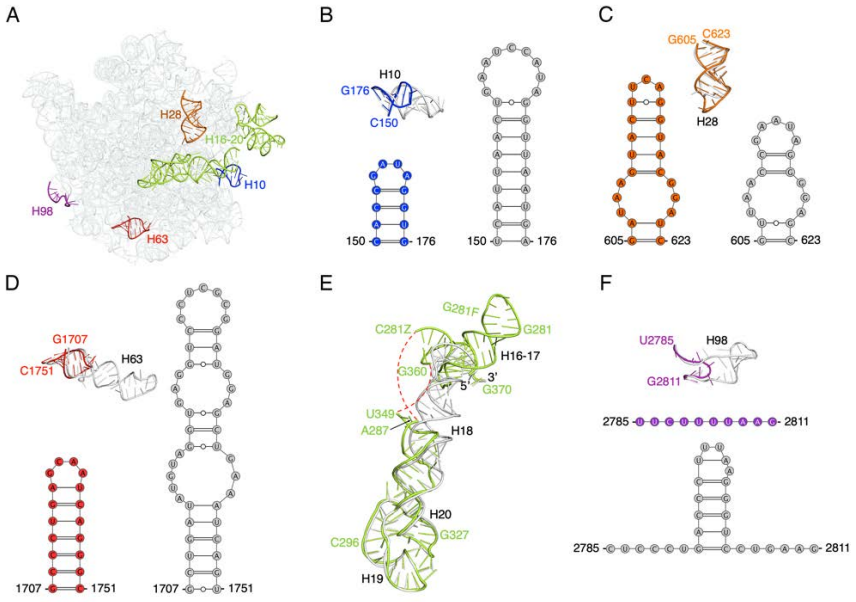


Figure 4. 7 Comparison of structural differences in the 23S rRNA between the 50S subunit in F. johnsoniae and E. coli.

(A) Solvent view of the 23S rRNA from F. johnsoniae with the regions distinct from E. coli (PDB ID: 6O9K) labelled and highlighted in different colors. (B–F) Comparison of helices H10, H28, H63, H16-H17, and H98, respectively from F. johnsoniae and E. coli along with their tertiary structure derived secondary structure diagram (except for helices H16-H17). Color codes for F. johnsoniae 23S rRNA elements are same as in (A), and E. coli's elements are shown in grey. The red dashed lines in (E) indicate the segment which could not be modeled due to weaker densities. Comparisons with T. thermophilus, M. smegmatis, P. aeruginosa, B. subtilis, and S. aureus are shown in Supplementary Figure S4.16.

Discovery of a novel ribosomal protein: bL38

Besides the differences in the rRNA, we identified a region of protein density in the 50S subunit, which to our knowledge has not been assigned to any known protein in any ribosome. To identify the protein, purified ribosomes and 50S subunits were subjected to LC/MS-MS analysis. One uncharacterized protein (UniProt A5F9Y9; pfam 14128, DUF4295) was found in both samples (Supplementary Table S9). This small (5.6 kDa), basic (pI = 10.5) polypeptide is encoded by Fjoh_4981, which lies immediately downstream of *rpmB* (Fjoh_4983) and *rpmG* (Fjoh_4982), the

genes for bL28 and bL33, respectively (Figure 4.8A & 4.8B). RNA-seq and ribo-seq read coverage is similar across these three genes (**Figure 4.8B**,(Baez et al., 2019)), suggesting that they belong to the same operon and are translated at comparable levels. While absent from other phyla, genes homologous to Fjoh_4981 are widely distributed across the Bacteroidetes **(Supplementary Figure S4.17A)**. Among representative species of the phylum, the gene neighborhood of Fjoh_4981 looks well conserved **(Supplementary Figure S17B)**. These observations suggest that Fjoh_4981 encodes a bona fide ribosomal protein, unique to the Bacteroidetes. Accordingly, we name the protein bL38 and the corresponding gene *rpmL*, in line with modern nomenclature (Ban et al., 2014, Li et al., 2018).

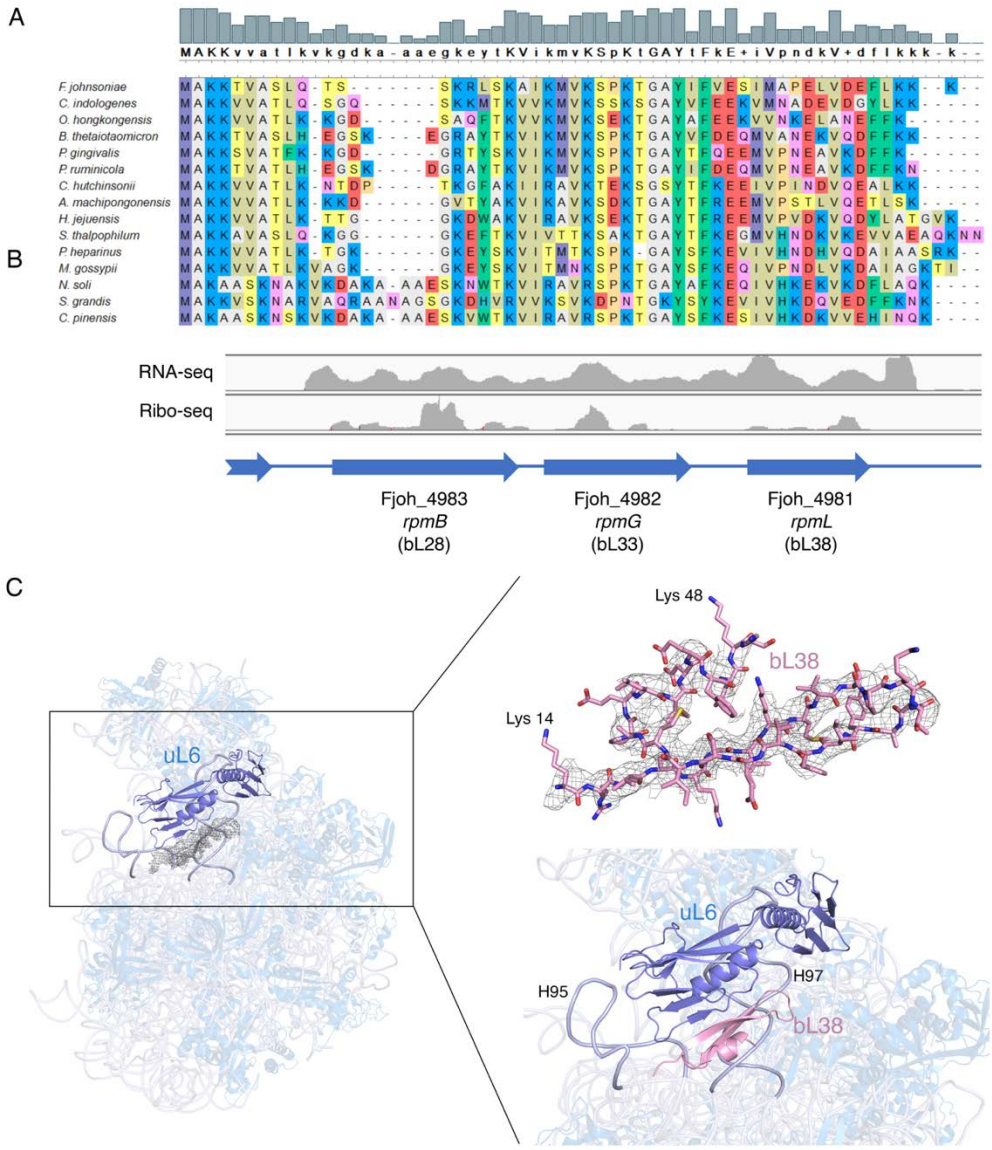


***Figure 4. 8** A novel ribosomal protein: bL38.*

*(**A**) Sequence alignment of bL38 proteins from fifteen Bacteroidetes, three representatives of each class Flavobacteriia, Bacteroidia, Cytophagia, Sphingobacteriia, and Chitinophagia (top to bottom, species listed). (**B**) RNA-seq and ribo-seq read coverage across rpmB (bL28), rpmG (bL33), and rpmL (bL38) in F. johnsoniae. Ratios of average per-nucleotide coverage are 1.0:1.1:1.3 (RNA-seq) and 1.0:0.6:0.5 (ribo-seq), respectively. (**C**) Ribbon representation of the overall 50S subunit (side view) from F. johnsoniae, highlighting the location of the unassigned protein density with respect to uL6 (marine blue; left panel). The modeled amino acid sequence for bL38 (from Lys 14 to Lys 48) is shown fitted into the density of the cryo-EM map (right panel; top) along with a close-up view of the left panel (right panel; bottom). The close-up view shows a cartoon representation of the newly described r protein, bL38 folding into a β-hairpin that forms a mixed β-sheet with r protein uL6. It also shows its close association with helices H95 and H97 (labelled and highlighted in light blue).*

The sequence of the protein encoded by Fjoh_4981 was used to build the molecular model using the unassigned density of the cryo-EM map **(Figure 4.8C)**. The quality of density was sufficient to assign side chains with confidence. We found that the observed density represented the C-terminal region of the protein from Lys 14 to Lys 48. There was no observed density for the first thirteen residues of the N-terminus or the last two residues of the C-terminus **(Figure 4.8C, right top panel)**. The newly described r protein bL38 folds into a b-hairpin that forms a mixed b-sheet with r protein uL6, located just above. bL38 also interacts with rRNA helices H95 and H97 **(Figure 4.8C, right bottom panel).**

Strong SD sequences lie upstream of Bacteroidetes genes encoding bS21 and/or bS18

We noticed that the bS21 gene (*rpsU*) of *F. johnsoniae* contains a "perfect" SD sequence (Figure 4.9A), whereas all the other *rps* genes lack an obvious SD. To further investigate this, we analyzed the TIRs of all *rps* genes in 293 representative species of the Bacteroidetes. Using the program free_scan (Starmer et al., 2006), the free energy of pairing between the canonical ASD (3'-UUCCUCCA-5') and each 8-nt stretch of mRNA within the TIR was calculated and plotted **(Supplementary Figures S4.18-S22)**. For many members of class Flavobacteriia, including *F. johnsoniae*, a prominent trough was observed near position -8 of *rpsU* **(Supplementary Figure S4.18)**, indicative of an authentic SD. Based on a stringent threshold of -7 kcal/mol, 60% of Flavobacteriia harbor a strong SD upstream of *rpsU* **(Table 4.1)**. This value increases to 77% when only those organisms with a canonical ASD are considered. By contrast, other *rps* genes of Flavobacteriia lack SD sequences, with only a few potential exceptions. For class Bacteroidia, SDs were rarely seen for *rpsU* but commonly seen for *rpsR* (encoding S18) (Supplementary Figure S19). In fact, 66% of the Bacteroidia analyzed exhibit a strong SD (≤ -7 kcal/mol) upstream of *rpsR* **(Table 4.1)**. Interestingly, the gene for uS11, *rpsK*, also contains a strong SD in a subset of

the Bacteroidia (14%). For class Chitinophagia, SDs were found to be most prevalent upstream of *rpsU* and less common upstream of *rpsR* and/or *rpsA* **(Table 1, Supplementary Figure S4.20)**. In the Cytophagia and Sphingobacteriia, strong SDs were commonly seen for *rpsU* and/or *rpsR* **(Table 1, Supplementary Figure S4.21-S4.22)**. In Sphingobacteriia, strong SDs were also evident for *rpsB* (encoding uS2), *rpsG* (encoding uS7), and *rpsS* (encoding uS19). Thus, even though the vast majority of Bacteroidetes genes lack SD sequences (Nakagawa et al., 2017) (Nakagawa et al., 2010) (Wegmann et al., 2013) (Accetto and Avgustin, 2011), certain ribosomal genes clearly have them. Remarkably, overall in the phylum, *rpsU* and/or *rpsR* commonly contain a strong SD **(Table 1)**, and the corresponding proteins bS21 and bS18 interact with the 16S rRNA in a way that occludes the ASD **(Figure 4.5)**.

Figure 4. 9 Sequences near the start codon of rpsU in various Flavobacteriia

(A) Potential base pairing between rRNA and rpsU mRNA in various Flavobacteriia that have the canonical ASD (as indicated). (B) Potential base pairing between rRNA and rpsU mRNA in various Flavobacteriia that have an alternative ASD (ASD2, as indicated). Bold, start codon. Red, mRNA nucleotides predicted to pair to 16S rRNA.

TABLES

Table 4. 1 Occurrences of strong SD sequences in the Bacteroidetes.

Gene	Flavobacteriia	Bacteroidia	Chitinophagia	Cytophagia	Sphingobacteriia
Small subunit genes					
rpsA (uS1)	0	0	0.182	0.022	0
rpsB (uS2)	0	0	0	0	0.769
rpsC (uS3)	0	0	0	0	0
rpsD (uS4)	0	0	0	0	0
rpsE (uS5)	0	0	0	0	0
rpsF (bS6)	0	0	0	0	0
rpsG (uS7)	0	0	0	0	0.462
rpsH (uS8)	0	0	0	0	0
rpsI (uS9)	0	0	0	0	0
rpsJ (uS10)	0	0	0	0	0
rpsK (uS11)	0	0.144	0	0	0
rpsL (uS12)	0	0	0	0	0
rpsM (uS13)	0	0	0	0	0
rpsN (uS14)	0	0	0	0	0
rpsO (uS15)	0	0	0	0	0
rpsP (bS16)	0	0	0	0	0
rpsQ (uS17)	0	0	0	0	0
rpsR (bS18)	0.008 (0.011) [1]	0.660	0.091	0.696	0.615
rpsS (uS19)	0	0	0	0	0.154
rpsT (bS20)	0	0	0	0	0
rpsU (bS21)	0.597 (0.772) [1]	0	0.455	0.478	0.846
Large subunit genes					
rplA (uL1)	0	0	0	0.261	0
rplB (uL2)	0	0	0	0	0
rplC (uL3)	0	0	0	0	0
rplD (uL4)	0	0.010	0	0	0
rplE (uL5)	0.008 (0.011) [1]	0.481	0.364	0.413	0.769
rplF (uL6)	0	0.539	0	0.065	0.462
rplI (bL9)	0	0.825	0.182	0.152	0.692

rplJ (uL10)	0	0.126	0	0	0.077
rplK (uL11)	0	0	0	0	0
rplL (uL7/12)	0	0	0	0	0
rplM (uL13)	0	0	0	0	0
rplN (uL14)	0	0	0	0	0
rplO (uL15)	0	0	0	0.022	0
rplP (uL16)	0	0	0	0	0
rplQ (bL17)	0	0	0	0	0
rplR (uL18)	0	0	0	0	0
rplS (bL19)	0	0	0	0	0
rplT (bL20)	0	0	0	0	0
rplU (bL21)	0	0	0	0	0
rplV (uL22)	0.050 (0) [1]	0	0	0	0
rplW (uL23)	0	0.385	0	0	0.692
rplX (uL24)	0	0	0.091	0.022	1
rplY (bL25)	0	0	0	0	N
rpmA (bL27)	0	0.327	0	0	0
rpmB (bL28)	0	0	0	0	0
rpmC (uL29)	0	0	0	0	0
rpmD (uL30)	0	0.385	0	0	0
rpmE (bL31)	0	N	N	N	N
rpmF (bL32)	0	0	0	0	0
rpmG (bL33)	0	0.087	0	0	0
rpmH (bL34)	0	0	0	0	0
rpmI (bL35)	0	0	0	0	0
rpmJ (bL36)	0	0	0	0	0
All genes[2]	0.002 (0.002) [1]	0.017	0.006	0.007	0.011

Data represent the fraction of genes that contain a strong SD, defined by a free energy of pairing threshold of -7 kcal/mol. N, no such gene identified. Numbers of genomes (*n*) analyzed: Flavobacteriia, *n* = 119; Bacteroidia, *n* = 104; Chitinophagia, *n* = 11; Cytophagia, *n* = 46; Sphingobacteriia, *n* = 13.

[1]Values in parentheses derive from Flavobacteriia with the canonical ASD (*n* = 92).

[2]All annotated protein coding genes, including ribosomal genes.

The fact that several genes of Sphingobacteriia besides *rpsU* and *rpsR* contain SD sequences prompted us to screen the large-subunit genes as well (Table 1). In Flavobacteriia, SDs are rarely

seen in any *rpl/rpm* genes. By contrast, SDs are commonly seen for a subset of *rpl/rpm* genes in the other classes. In Bacteroidia, strong SDs occur with a frequency of > 0.1 upstream of genes encoding uL5, uL6, bL9, uL10, uL23, bL27, and uL30. A similar but smaller set of large-subunit genes (uL5, uL6, bL9, uL10, and uL23) exhibit SDs in the Sphingobacteriia. In Chitinophagia and Cytophagia, genes for uL5 and bL9 also exhibit SDs, albeit at lower frequency. SDs also lie upstream of *rplA* (uL1) in about one fourth of the Cytophagia species analyzed **(Table 4.1)**.

We next applied the same screen to all TIRs in all genomes. This allowed us to compute the SD frequency for all genes per organism and per class (Supplementary Table S5). To assess the expected rate of false positives, we also screened an upstream control window (-25 to -40; too far from the start codon) and termed the resulting "hits" mock SD (MSD) sequences. Three different free energy thresholds (-7, -6, and -5 kcal/mol) were used for these screens, and consistent trends emerged regardless of the threshold. The analysis revealed that Flavobacteriia exhibit the lowest average SD frequency **(Table 4.1, bottom row; Supplementary Table S4.5)**. In most Flavobacteriia species, fewer SDs were identified than MSDs, suggesting that selective pressure has acted to largely eliminate SDs in these organisms **(Supplementary Table S4.5)**. In Chitinophagia and Cytophagia, average SD frequencies are slightly higher, on par with MSD frequencies **(Table 1, bottom row; Supplementary Table S4.5)**. Bacteroidia and Sphingobacteriia exhibit somewhat larger average SD frequencies. For most organisms in these classes, the number of SDs modestly exceeds the number of MSDs. These genome-wide metrics are in line with the ribosomal gene results **(Table 4.1)**, and collectively the data reveal differences in SD usage among the Bacteroidetes classes.

Covariation between mRNA and rRNA in a subset of Flavobacteriia

Futcher and coworkers showed that 16S rRNA genes of Chryseobacteria, Riemerella, and related genera encode an alternative ASD sequence, 5'-AUCUCAUU-3' (henceforth termed ASD2), with two substitutions (C1535U and C1539A, underscored above) (Amin et al., 2018). We looked upstream of *rpsU* in these organisms and in all cases found a sequence fully complementary to ASD2 and able form an extended mRNA-rRNA helix **(Figure 4.9B)**. This prompted us to look more systematically in these organisms for "SD2" sequences, predicted to pair with ASD2, using free_scan. No other ribosomal gene exhibits a convincing SD2 **(Supplementary Figure S4.23)**. We then screened all genes and compared the frequencies of SD2 and mock SD2 (MSD2) sequences, using various free energy thresholds **(Supplementary Table S4.5)**. We found that the SD2 sequences are generally rare in these organisms. At the -7 kcal/mol threshold, very few (if any) genes besides *rpsU* were identified. At the -6 and -5 kcal/mol thresholds, frequencies of SD2 sequences remained quite low, similar to MSD2 frequencies. Relative to MSD2, SD2 sequences seemed to be modestly underrepresented in the former case and modestly overrepresented in the latter case, although the biological relevance of this observation remains unclear. Regardless, it is evident that translation initiation rarely entails mRNA-rRNA pairing in the Flavobacteriia, whether their ribosomes harbor ASD or ASD2. Importantly, these data provide compelling evidence for natural covariation in the SD-ASD helix and further implicate SD-ASD pairing in translation of *rpsU* in the Flavobacteriia.

4.5 Discussion

It has been known for a decade that Bacteroidetes lack SD sequences, even though ribosomes of most of these organisms retain the conserved ASD. Introduced SD elements cause no stimulation of translation *in vivo*, suggesting that pairing between 16S rRNA and mRNA is somehow prevented or made inconsequential. In this work, we uncover the basis of ASD inhibition in the Bacteroidetes. The 30S subunits purified from *F. johnsoniae* fail to recognize the SD *in vitro*, indicating that intra-subunit interactions are responsible for ASD occlusion. A high-resolution cryo-EM structure of the *F. johnsoniae* 70S ribosome reveals that the 3' end of 16S rRNA interacts with bS21, bS18, and bS6 on the 30S platform, contacts which physically sequester the ASD bases. Many of these contacts involve amino acids uniquely conserved in the Bacteroidetes, suggesting that the same mechanism of ASD occlusion is operational across the phylum. The entire 3' end of 16S rRNA is well resolved in the cryo-EM map. This differs from all other vacant bacterial ribosome structures (Schuwirth et al., 2005, Cocozaki et al., 2016, Borovinskaya et al., 2007), which exhibit weak density beyond nucleotide 1534. We infer that the 3' tail of 16S rRNA binds stably to the 30S platform in the Bacteroidetes ribosome but is free and mobile in other bacterial ribosomes. Consequently, the Bacteroidetes ribosome is unable to recognize SD sequences as other ribosomes do.

Absent from the imaged 70S particles were ribosomal proteins bS1 and uS2. Based on SDS-PAGE analysis and previous cryo-EM work (Sengupta et al., 2001), we infer that bS1 was mainly lost during ribosome purification while uS2 was lost during purification and sample vitrification. Importantly, there is no indication that either protein affects ASD function. Levels of both proteins are higher in our 30S preparation (which contains stoichiometrically-bound uS2). Yet, the 30S and 70S particles similarly fail to recognize the SD in assays of IC formation **(Figure 4.1)**. Moreover,

occlusion of the ASD is directly observed in 70S particles lacking bS1 and uS2, indicating that neither protein is required for the mechanism. That being said, we do suspect that bS1 depletion contributes to the generally low efficiency of IC formation seen for *F. johnsoniae* ribosomes. Protein bS1 is known to interact with mRNA, and bS1-mRNA contacts may be particularly important for translation initiation in organisms like the Bacteroidetes, which generally lack SD sequences (Salah et al., 2009) (Baez et al., 2019) (Nakagawa et al., 2010). Future work will be needed to clarify the role of bS1 in *F. johnsoniae* IC formation.

We also report that in the Bacteroidetes certain ribosomal genes often contain SD sequences. In many Flavobacteriia, *rpsU* is the only ribosomal gene with an obvious SD. Flavobacteriia with an alternative ASD (ASD2) also have the complementary sequence (SD2) upstream of *rpsU*, corroborating the functional importance of rRNA-mRNA pairing in translation of *rpsU*. The corresponding protein, bS21, makes specific contacts to the 3' tail of 16S rRNA and orients it toward bS18 and bS6. These observations suggest a simple mechanism of translational autoregulation. We propose that replete and bS21-deficient ribosomes are both active in translation; however, the latter ribosomes initiate translation of *rpsU* mRNA at a higher rate because the ASD is functionally liberated **(Figure 4.10)**. As levels of bS21 increase in the cell, so does the proportion of replete ribosomes, damping down further synthesis of bS21. Protein bS21 is one of the last proteins to be incorporated during 30S assembly (Chen and Williamson, 2013) (Sashital et al., 2014). Thus, such feedback regulation may help ensure the completion of ribosome biogenesis, for example when cells rapidly enter an environment without nutrients. While future work will be needed to test this hypothesis, the basic concept is consistent with the following observations. First, bS21 is naturally absent from some bacteria (Yutin et al., 2012) and nonessential in others (Klein et al., 2012, Takada et al., 2014, Akanuma et al., 2012, Metselaar et

al., 2015), so there are ample examples of active bS21-deficient ribosomes. Second, homologs of bS21 are encoded by many marine bacteriophages (Mizuno et al., 2019). In tested cases, these viral proteins can assemble into host ribosomes, possibly to shunt those ribosomes to phage mRNA.

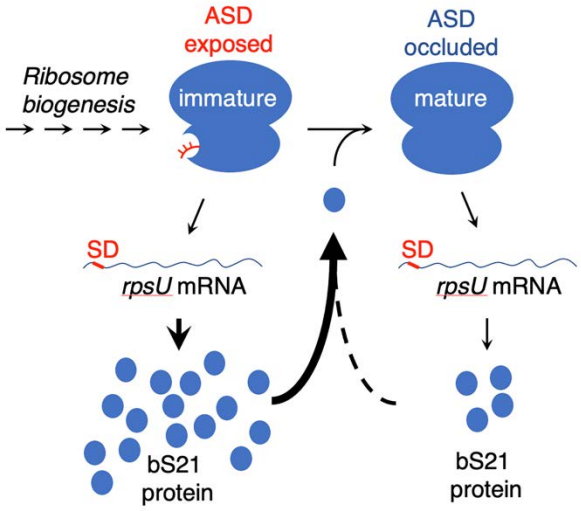


Figure 4. 10 A potential mechanism of bS21 autoregulation.

Incorporation of bS21 is one of the last steps in 30S biogenesis. Ribosomes lacking bS21 (immature) contain a liberated ASD, allowing for high-level translation of the rpsU mRNA. Product bS21 binds to generate the replete (mature) ribosome, in which the ASD is occluded. Replete ribosomes will also translate rpsU mRNA but at a reduced rate.

In Bacteroidia, SDs are absent from *rpsU* but frequently found upstream of *rpsR*, the gene encoding bS18. Like bS21, bS18 contributes to ASD occlusion and hence could regulate its own production through a mechanism analogous to that described above. Namely, we hypothesize that bS18-depleted subunits, which can readily engage SD sequences, increase the overall rate of translation of *rpsR* mRNA. The bS18 gene is nonessential in *Porphyromonas gingivalis* (Klein et al., 2012), supporting the plausibility of this hypothesis. Interestingly, SDs are prevalent in *rpsU*

and/or *rpsR* in many Chitinophagia, Cytophagia, and Sphingobacteriia, raising the possibility that bS21- and/or bS18-depleted subunits contribute to translational control in these organisms.

In Bacteroidia, Chitinophagia, Cytophagia, and Sphingobacteriia, other ribosomal genes frequently contain SDs. While the purpose of these SDs remains unclear, they are likely involved in some form of translational control. Intriguingly, proteins of the 30S encoded by these genes (bS1, uS2, uS7, uS11) bind near bS21 and bS18 in the mature subunit, and there exist multiple functional associations between these proteins during 30S assembly (Mizushima and Nomura, 1970, Mulder et al., 2010, Sashital et al., 2014). If our hypothesis above is correct, the functional difference between replete and bS21-depleted (or bS18-depleted) ribosomes could be exploited to regulate ribosome assembly. These depleted subunits, acting as proxies of assembly intermediates, might adjust translation of multiple mRNAs to ensure efficient ribosome biogenesis under various growth conditions. We envisage that this sort of feedback control has been elaborated to varying degrees, depending on the class of Bacteroidetes. A clear advantage gained by occlusion of the ASD in the vast majority of cellular ribosomes is an opportunity to use SDs as regulatory elements.

One of the most conserved features of the bacterial ribosome is the ASD. Recent genomic studies have identified the rare cases of ASD sequence variation, which occur almost exclusively in a subset of Flavobacteriia (Amin et al., 2018, Lim et al., 2012). These natural substitutions occur mainly at positions 1535 and/or 1539, as exemplified by ASD2 of Chyseobacteria and related organisms **(Figure 4.9B)**. Notably, in the *F. johnsoniae* ribosome structure, the base of C1535 is oriented toward the solvent and neither C1535 nor C1539 makes specific contacts with the ribosome platform. Hence, substitutions at these positions are predicted to have little or no effect on the mechanism of ASD occlusion, in contrast to substitutions at neighboring positions. This

hints that, in the Bacteroidetes, platform interactions impose stronger evolutionary constraints on the 3' end of 16S rRNA than mRNA interactions do.

Finally, we show that *F. johnsoniae* initiation complexes can be reconstituted *in vitro* using purified components. Complex formation is strictly factor-dependent and occurs at the canonical start codon for all tested model mRNAs (from *F. johnsoniae* or *E. coli*). While 70S ICs were readily detected by toeprinting, 30S ICs were not (or gave weak toeprints) in parallel experiments. Either these 30S ICs were too labile to halt reverse transcriptase or they simply did not form under the conditions employed. It is possible that these 30S ICs fail to form due to loss of factor or subunit activity. However, efficient 70S IC formation requires all three *F. johnsoniae* initiation factors **(Supplementary Figure S4.4)**, suggesting that each purified factor exhibits activity. Moreover, replacement of *F. johnsoniae* IF3 with *E. coli* IF3 results in a heterologous 30S IC, readily detected by toeprinting, indicating that 30S subunits prepared from *F. johnsoniae* are capable of complex formation. Additionally, 70S ICs can be formed from isolated *F. johnsoniae* subunits (rather than tight-couple ribosomes), providing further evidence that these 30S subunits exhibit activity. Still, further work will be needed to determine whether the anomalous behavior of the *F. johnsoniae* 30S IC holds biological relevance or has a more trivial basis.

The ability of *E. coli* IF3 to stabilize *F. johnsoniae* 30S complexes came as a surprise, because studies of *E. coli* initiation have shown that IF3 has a net destabilizing effect on the 30S IC (Qin et al., 2012) (Antoun et al., 2006). IF3 increases both the binding rate (k_{on}) and dissociation rate (k_{off}) of fMet-tRNA, but the effect on k_{off} is larger (Antoun et al., 2006). Notably, the *E. coli* studies have employed SD-containing mRNA, and to our knowledge analogous studies using SD-lacking mRNA have yet to be performed. IF3 is a two-domain protein that exhibits multiple modes of binding during initiation (Hussain et al., 2016, Elvekrog and Gonzalez, 2013). The *C*-terminal

domain interacts with distinct sites on the 30S platform, and the *N*-terminal domain contacts the elbow region of fMet-tRNA. IF3 lies near the SD-ASD helix, and IF3 and IF1 affect the positioning of the SD-ASD helix (Hussain et al., 2016). These observations raise the following questions, which remain open: (1) Does the role of IF3 differ in the two bacterial systems? (2) Does SD-ASD pairing influence IF3 dynamics and its interplay with fMet-tRNA? Additional experiments will be needed to address these questions and further advance our understanding of translation initiation in bacteria.

4.6 Supplementary Data

This section contains enormous amounts of supplementary material and is openly accessible at NAR journal online link: https://academic.oup.com/nar/article/49/1/547/6039927#supplementary-data

Acknowledgements

We thank S. McNutt and F. Chu at the University of New Hampshire for performing the LC-MS/MS analysis; K. Sears, M. Strauss and other staff members of the Facility for Electron Microscopy Research (FEMR) at McGill University for help with microscope operation and data collection; and R. Green, A. Buskirk, and T. Dever for feedback on the manuscript.

Funding

This work was supported by grants from the National Science Foundation (MCB-2029502 to K.F.) and the NSERC Discovery Program (RGPIN-2019-05799 to J.O.). Titan Krios cryo-EM data were collected at FEMR (McGill). FEMR is supported by the Canadian Foundation for Innovation, Quebec Government and McGill University.

Conflict of interest statement.

None declared.

4.7 References

ACCETTO, T. & AVGUSTIN, G. 2011. Inability of Prevotella bryantii to form a functional Shine-Dalgarno interaction reflects unique evolution of ribosome binding sites in Bacteroidetes. *PLoS One,* 6**,** e22914.

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., MCCOY, A. J., MORIARTY, N. W., OEFFNER, R., READ, R. J., RICHARDSON, D. C., RICHARDSON, J. S., TERWILLIGER, T. C. & ZWART, P. H. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr,* 66**,** 213-21.

AFONINE, P. V., KLAHOLZ, B. P., MORIARTY, N. W., POON, B. K., SOBOLEV, O. V., TERWILLIGER, T. C., ADAMS, P. D. & URZHUMTSEV, A. 2018a. New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr D Struct Biol,* 74**,** 814-840.

AFONINE, P. V., POON, B. K., READ, R. J., SOBOLEV, O. V., TERWILLIGER, T. C., URZHUMTSEV, A. & ADAMS, P. D. 2018b. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol,* 74**,** 531-544.

AKANUMA, G., NANAMIYA, H., NATORI, Y., YANO, K., SUZUKI, S., OMATA, S., ISHIZUKA, M., SEKINE, Y. & KAWAMURA, F. 2012. Inactivation of ribosomal protein genes in Bacillus subtilis reveals importance of each ribosomal protein for cell proliferation and cell differentiation. *J Bacteriol,* 194**,** 6282-91.

AMIN, M. R., YUROVSKY, A., CHEN, Y., SKIENA, S. & FUTCHER, B. 2018. Re-annotation of 12,495 prokaryotic 16S rRNA 3' ends and analysis of Shine-Dalgarno and anti-Shine-Dalgarno sequences. *PLoS One,* 13**,** e0202767.

ANTOUN, A., PAVLOV, M. Y., LOVMAR, M. & EHRENBERG, M. 2006. How initiation factors tune the rate of initiation of protein synthesis in bacteria. *EMBO J,* 25**,** 2539-50.

BAEZ, W. D., ROY, B., MCNUTT, Z. A., SHATOFF, E. A., CHEN, S., BUNDSCHUH, R. & FREDRICK, K. 2019. Global analysis of protein synthesis in Flavobacterium johnsoniae reveals the use of Kozak-like sequences in diverse bacteria. *Nucleic Acids Res,* 47**,** 10477-10488.

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., MCALEAR, M. A., MOORE, P. B., NOLLER, H. F., ORTEGA, J., PANSE, V. G., RAMAKRISHNAN, V., SPAHN, C. M., STEITZ, T. A., TCHORZEWSKI, M., TOLLERVEY, D., WARREN, A. J., WILLIAMSON, J. R., WILSON, D., YONATH, A. & YUSUPOV, M. 2014. A new system for naming ribosomal proteins. *Curr Opin Struct Biol,* 24**,** 165-9.

BAYLEY, D. P., ROCHA, E. R. & SMITH, C. J. 2000. Analysis of cepA and other Bacteroides fragilis genes reveals a unique promoter structure. *FEMS Microbiol Lett,* 193**,** 149-54.

BERINGER, M., BRUELL, C., XIONG, L., PFISTER, P., BIELING, P., KATUNIN, V. I., MANKIN, A. S., BOTTGER, E. C. & RODNINA, M. V. 2005. Essential mechanisms in the catalysis of peptide bond formation on the ribosome. *J Biol Chem,* 280**,** 36065-72.

BESEMER, J., LOMSADZE, A. & BORODOVSKY, M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res,* 29**,** 2607-18.

BOROVINSKAYA, M. A., PAI, R. D., ZHANG, W., SCHUWIRTH, B. S., HOLTON, J. M., HIROKAWA, G., KAJI, H., KAJI, A. & CATE, J. H. 2007. Structural basis for aminoglycoside inhibition of bacterial ribosome recycling. *Nat Struct Mol Biol,* 14**,** 727-32.

BUTLER, J. S., SPRINGER, M., DONDON, J., GRAFFE, M. & GRUNBERG-MANAGO, M. 1986. Escherichia coli protein synthesis initiation factor IF3 controls its own gene expression at the translational level in vivo. *J Mol Biol,* 192**,** 767-80.

CHEN, S., BAGDASARIAN, M., KAUFMAN, M. G., BATES, A. K. & WALKER, E. D. 2007a. Mutational analysis of the ompA promoter from Flavobacterium johnsoniae. *J Bacteriol,* 189**,** 5108-18.

CHEN, S., BAGDASARIAN, M., KAUFMAN, M. G. & WALKER, E. D. 2007b. Characterization of strong promoters from an environmental Flavobacterium hibernum strain by using a green fluorescent protein-based reporter system. *Appl Environ Microbiol,* 73**,** 1089-100.

CHEN, S. S. & WILLIAMSON, J. R. 2013. Characterization of the Ribosome Biogenesis Landscape in E. coli Using Quantitative Mass Spectrometry. *J Mol Biol,* 425**,** 767-79.

COCOZAKI, A. I., ALTMAN, R. B., HUANG, J., BUURMAN, E. T., KAZMIRSKI, S. L., DOIG, P., PRINCE, D. B., BLANCHARD, S. C., CATE, J. H. & FERGUSON, A. D. 2016. Resistance mutations generate divergent antibiotic susceptibility profiles against translation inhibitors. *Proc Natl Acad Sci U S A,* 113**,** 8188-93.

CROOKS, G. E., HON, G., CHANDONIA, J. M. & BRENNER, S. E. 2004. WebLogo: a sequence logo generator. *Genome Res,* 14**,** 1188-90.

DALLAS, A. & NOLLER, H. F. 2001. Interaction of translation initiation factor 3 with the 30S ribosomal subunit. *Mol Cell,* 8**,** 855-64.

DARTY, K., DENISE, A. & PONTY, Y. 2009. VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics,* 25**,** 1974-5.

DE LA ROSA-TREVIN, J. M., OTON, J., MARABINI, R., ZALDIVAR, A., VARGAS, J., CARAZO, J. M. & SORZANO, C. O. 2013. Xmipp 3.0: an improved software suite for image processing in electron microscopy. *J Struct Biol,* 184**,** 321-8.

DE SMIT, M. H. & VAN DUIN, J. 1994. Translational initiation on structured messengers. Another role for the Shine-Dalgarno interaction. *J Mol Biol,* 235**,** 173-84.

DELANO, W. L. 2002. *The PyMOL Molecular Graphic Systems,* Palo Alto, CA, USA, DeLano Scientific.

DELCHER, A. L., BRATKE, K. A., POWERS, E. C. & SALZBERG, S. L. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics,* 23**,** 673-9.

ELVEKROG, M. M. & GONZALEZ, R. L., JR. 2013. Conformational selection of translation initiation factor 3 signals proper substrate selection. *Nat Struct Mol Biol,* 20**,** 628-33.

EMSLEY, P. & COWTAN, K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr,* 60**,** 2126-32.

EMSLEY, P., LOHKAMP, B., SCOTT, W. G. & COWTAN, K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr,* 66**,** 486-501.

FAHLMAN, R. P. & UHLENBECK, O. C. 2004. Contribution of the esterified amino acid to the binding of aminoacylated tRNAs to the ribosomal P- and A-sites. *Biochemistry,* 43**,** 7575-83.

FREDRICK, K. & NOLLER, H. F. 2002. Accurate translocation of mRNA by the ribosome requires a peptidyl group or its analog on the tRNA moving into the 30S P site. *Mol Cell,* 9**,** 1125-31.

GODDARD, T. D., HUANG, C. C., MENG, E. C., PETTERSEN, E. F., COUCH, G. S., MORRIS, J. H. & FERRIN, T. E. 2018. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci,* 27**,** 14-25.

GOLD, L., STORMO, G. & SAUNDERS, R. 1984. Escherichia coli translational initiation factor IF3: a unique case of translational regulation. *Proc Natl Acad Sci U S A,* 81**,** 7061-5.

HAHNKE, R. L., MEIER-KOLTHOFF, J. P., GARCIA-LOPEZ, M., MUKHERJEE, S., HUNTEMANN, M., IVANOVA, N. N., WOYKE, T., KYRPIDES, N. C., KLENK, H. P. & GOKER, M. 2016. Genome-Based Taxonomic Classification of Bacteroidetes. *Front Microbiol,* 7**,** 2003.

HALFON, Y., JIMENEZ-FERNANDEZ, A., LA ROSA, R., ESPINOSA PORTERO, R., KROGH JOHANSEN, H., MATZOV, D., EYAL, Z., BASHAN, A., ZIMMERMAN, E., BELOUSOFF, M., MOLIN, S. & YONATH, A. 2019. Structure of Pseudomonas aeruginosa ribosomes from an aminoglycoside-resistant clinical isolate. *Proc Natl Acad Sci U S A,* 116**,** 22275-22281.

HARTZ, D., MCPHEETERS, D. S. & GOLD, L. 1989. Selection of the initiator tRNA by Escherichia coli initiation factors. *Genes Dev,* 3**,** 1899-912.

HENTSCHEL, J., BURNSIDE, C., MIGNOT, I., LEIBUNDGUT, M., BOEHRINGER, D. & BAN, N. 2017. The Complete Structure of the Mycobacterium smegmatis 70S Ribosome. *Cell Rep,* 20**,** 149-160.

HUI, A. & DE BOER, H. A. 1987. Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in Escherichia coli. *Proc Natl Acad Sci U S A,* 84**,** 4762-6.

HUSSAIN, T., LLACER, J. L., WIMBERLY, B. T., KIEFT, J. S. & RAMAKRISHNAN, V. 2016. Large-Scale Movements of IF3 and tRNA during Bacterial Translation Initiation. *Cell,* 167**,** 133-144 e13.

JACOB, W. F., SANTER, M. & DAHLBERG, A. E. 1987. A single base change in the Shine-Dalgarno region of 16S rRNA of Escherichia coli affects translation of many proteins. *Proc Natl Acad Sci U S A,* 84**,** 4757-61.

JAHAGIRDAR, D., JHA, V., BASU, K., GOMEZ-BLANCO, J., VARGAS, J. & ORTEGA, J. 2020. Alternative conformations and motions adopted by 30S ribosomal subunits visualized by cryo-electron microscopy. *RNA,* 26**,** 2017-2030.

JENNER, L. B., DEMESHKINA, N., YUSUPOVA, G. & YUSUPOV, M. 2010. Structural aspects of messenger RNA reading frame maintenance by the ribosome. *Nat Struct Mol Biol,* 17**,** 555-60.

JOHNSON, E. L., HEAVER, S. L., WALTERS, W. A. & LEY, R. E. 2017. Microbiome and metabolic disease: revisiting the bacterial phylum Bacteroidetes. *J Mol Med (Berl),* 95**,** 1-8.

KALEDHONKAR, S., FU, Z., CABAN, K., LI, W., CHEN, B., SUN, M., GONZALEZ, R. L., JR. & FRANK, J. 2019. Late steps in bacterial translation initiation visualized using time-resolved cryo-EM. *Nature,* 570**,** 400-404.

KAMINISHI, T., WILSON, D. N., TAKEMOTO, C., HARMS, J. M., KAWAZOE, M., SCHLUENZEN, F., HANAWA-SUETSUGU, K., SHIROUZU, M., FUCINI, P. &

YOKOYAMA, S. 2007. A snapshot of the 30S ribosomal subunit capturing mRNA via the Shine-Dalgarno interaction. *Structure,* 15**,** 289-97.
KATOH, K., MISAWA, K., KUMA, K. & MIYATA, T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res,* 30**,** 3059-66.
KHUSAINOV, I., VICENS, Q., BOCHLER, A., GROSSE, F., MYASNIKOV, A., MENETRET, J. F., CHICHER, J., MARZI, S., ROMBY, P., YUSUPOVA, G., YUSUPOV, M. & HASHEM, Y. 2017. Structure of the 70S ribosome from human pathogen Staphylococcus aureus. *Nucleic Acids Res,* 45**,** 1026.
KITTS, P. A., CHURCH, D. M., THIBAUD-NISSEN, F., CHOI, J., HEM, V., SAPOJNIKOV, V., SMITH, R. G., TATUSOVA, T., XIANG, C., ZHERIKOV, A., DICUCCIO, M., MURPHY, T. D., PRUITT, K. D. & KIMCHI, A. 2016. Assembly: a resource for assembled genomes at NCBI. *Nucleic Acids Res,* 44**,** D73-80.
KLEIN, B. A., TENORIO, E. L., LAZINSKI, D. W., CAMILLI, A., DUNCAN, M. J. & HU, L. T. 2012. Identification of essential genes of the periodontal pathogen Porphyromonas gingivalis. *BMC Genomics,* 13**,** 578.
KOROSTELEV, A., TRAKHANOV, S., ASAHARA, H., LAURBERG, M., LANCASTER, L. & NOLLER, H. F. 2007. Interactions and dynamics of the Shine Dalgarno helix in the 70S ribosome. *Proc Natl Acad Sci U S A,* 104**,** 16840-3.
KOZAK, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell,* 44**,** 283-92.
KOZLOWSKI, L. P. 2016. IPC - Isoelectric Point Calculator. *Biol Direct,* 11**,** 55.
LALANNE, J. B., TAGGART, J. C., GUO, M. S., HERZEL, L., SCHIELER, A. & LI, G. W. 2018. Evolutionary Convergence of Pathway-Specific Enzyme Expression Stoichiometry. *Cell,* 173**,** 749-761 e38.
LANCASTER, L., KIEL, M. C., KAJI, A. & NOLLER, H. F. 2002. Orientation of ribosome recycling factor in the ribosome from directed hydroxyl radical probing. *Cell,* 111**,** 129-40.
LARKIN, M. A., BLACKSHIELDS, G., BROWN, N. P., CHENNA, R., MCGETTIGAN, P. A., MCWILLIAM, H., VALENTIN, F., WALLACE, I. M., WILM, A., LOPEZ, R., THOMPSON, J. D., GIBSON, T. J. & HIGGINS, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics,* 23**,** 2947-8.
LEY, R. E., HAMADY, M., LOZUPONE, C., TURNBAUGH, P. J., RAMEY, R. R., BIRCHER, J. S., SCHLEGEL, M. L., TUCKER, T. A., SCHRENZEL, M. D., KNIGHT, R. & GORDON, J. I. 2008. Evolution of mammals and their gut microbes. *Science,* 320**,** 1647-51.
LI, G. W., BURKHARDT, D., GROSS, C. & WEISSMAN, J. S. 2014. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell,* 157**,** 624-35.
LI, Z., GE, X., ZHANG, Y., ZHENG, L., SANYAL, S. & GAO, N. 2018. Cryo-EM structure of Mycobacterium smegmatis ribosome reveals two unidentified ribosomal proteins close to the functional centers. *Protein Cell,* 9**,** 384-388.
LIM, K., FURUTA, Y. & KOBAYASHI, I. 2012. Large variations in bacterial ribosomal RNA genes. *Mol Biol Evol,* 29**,** 2937-48.
LU, X. J., BUSSEMAKER, H. J. & OLSON, W. K. 2015. DSSR: an integrated software tool for dissecting the spatial structure of RNA. *Nucleic Acids Res,* 43**,** e142.

MARTENS, E. C., CHIANG, H. C. & GORDON, J. I. 2008. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host Microbe,* 4**,** 447-57.

MARTENS, E. C., KOROPATKIN, N. M., SMITH, T. J. & GORDON, J. I. 2009a. Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. *J Biol Chem,* 284**,** 24673-7.

MARTENS, E. C., ROTH, R., HEUSER, J. E. & GORDON, J. I. 2009b. Coordinate regulation of glycan degradation and polysaccharide capsule biosynthesis by a prominent human gut symbiont. *J Biol Chem,* 284**,** 18445-57.

MCBRIDE, M. J. & KEMPF, M. J. 1996. Development of techniques for the genetic manipulation of the gliding bacterium Cytophaga johnsonae. *J Bacteriol,* 178**,** 583-90.

MCBRIDE, M. J., XIE, G., MARTENS, E. C., LAPIDUS, A., HENRISSAT, B., RHODES, R. G., GOLTSMAN, E., WANG, W., XU, J., HUNNICUTT, D. W., STAROSCIK, A. M., HOOVER, T. R., CHENG, Y. Q. & STEIN, J. L. 2009. Novel features of the polysaccharide-digesting gliding bacterium Flavobacterium johnsoniae as revealed by genome sequence analysis. *Appl Environ Microbiol,* 75**,** 6864-75.

METSELAAR, K. I., DEN BESTEN, H. M., BOEKHORST, J., VAN HIJUM, S. A., ZWIETERING, M. H. & ABEE, T. 2015. Diversity of acid stress resistant variants of Listeria monocytogenes and the potential role of ribosomal protein S21 encoded by rpsU. *Front Microbiol,* 6**,** 422.

MIZUNO, C. M., GUYOMAR, C., ROUX, S., LAVIGNE, R., RODRIGUEZ-VALERA, F., SULLIVAN, M. B., GILLET, R., FORTERRE, P. & KRUPOVIC, M. 2019. Numerous cultivated and uncultivated viruses encode ribosomal proteins. *Nat Commun,* 10**,** 752.

MIZUSHIMA, S. & NOMURA, M. 1970. Assembly mapping of 30S ribosomal proteins from E. coli. *Nature,* 226**,** 1214.

MULDER, A. M., YOSHIOKA, C., BECK, A. H., BUNNER, A. E., MILLIGAN, R. A., POTTER, C. S., CARRAGHER, B. & WILLIAMSON, J. R. 2010. Visualizing ribosome biogenesis: parallel assembly pathways for the 30S subunit. *Science,* 330**,** 673-7.

NAKAGAWA, S., NIIMURA, Y. & GOJOBORI, T. 2017. Comparative genomic analysis of translation initiation mechanisms for genes lacking the Shine-Dalgarno sequence in prokaryotes. *Nucleic Acids Res,* 45**,** 3922-3931.

NAKAGAWA, S., NIIMURA, Y., GOJOBORI, T., TANAKA, H. & MIURA, K. 2008. Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. *Nucleic Acids Res,* 36**,** 861-71.

NAKAGAWA, S., NIIMURA, Y., MIURA, K. & GOJOBORI, T. 2010. Dynamic evolution of translation initiation mechanisms in prokaryotes. *Proc Natl Acad Sci U S A,* 107**,** 6382-7.

OKONECHNIKOV, K., GOLOSOVA, O., FURSOV, M. & TEAM, U. 2012. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics,* 28**,** 1166-7.

OSADA, Y., SAITO, R. & TOMITA, M. 1999. Analysis of base-pairing potentials between 16S rRNA and 5' UTR for translation initiation in various prokaryotes. *Bioinformatics,* 15**,** 578-81.

OU, H. Y., GUO, F. B. & ZHANG, C. T. 2004. GS-Finder: a program to find bacterial gene start sites with a self-training method. *Int J Biochem Cell Biol,* 36**,** 535-44.

PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem,* 25**,** 1605-12.

POLIKANOV, Y. S., STAROSTA, A. L., JUETTE, M. F., ALTMAN, R. B., TERRY, D. S., LU, W., BURNETT, B. J., DINOS, G., REYNOLDS, K. A., BLANCHARD, S. C., STEITZ, T. A. & WILSON, D. N. 2015. Distinct tRNA Accommodation Intermediates Observed on the Ribosome with the Antibiotics Hygromycin A and A201A. *Mol Cell,* 58**,** 832-44.

QIN, D., ABDI, N. M. & FREDRICK, K. 2007. Characterization of 16S rRNA mutations that decrease the fidelity of translation initiation. *RNA,* 13**,** 2348-55.

QIN, D. & FREDRICK, K. 2009. Control of translation initiation involves a factor-induced rearrangement of helix 44 of 16S ribosomal RNA. *Mol Microbiol,* 71**,** 1239-49.

QIN, D., LIU, Q., DEVARAJ, A. & FREDRICK, K. 2012. Role of helix 44 of 16S rRNA in the fidelity of translation initiation. *RNA,* 18**,** 485-95.

RINGQUIST, S., SHINEDLING, S., BARRICK, D., GREEN, L., BINKLEY, J., STORMO, G. D. & GOLD, L. 1992. Translation initiation in Escherichia coli: sequences within the ribosome-binding site. *Mol Microbiol,* 6**,** 1219-29.

RINNINELLA, E., RAOUL, P., CINTONI, M., FRANCESCHI, F., MIGGIANO, G. A. D., GASBARRINI, A. & MELE, M. C. 2019. What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. *Microorganisms,* 7.

ROY, B., LIU, Q., SHOJI, S. & FREDRICK, K. 2018. IF2 and unique features of initiator tRNA(fMet) help establish the translational reading frame. *RNA Biol,* 15**,** 604-613.

SACERDOT, C., CHIARUTTINI, C., ENGST, K., GRAFFE, M., MILET, M., MATHY, N., DONDON, J. & SPRINGER, M. 1996. The role of the AUU initiation codon in the negative feedback regulation of the gene for translation initiation factor IF3 in Escherichia coli. *Mol Microbiol,* 21**,** 331-46.

SALAH, P., BISAGLIA, M., ALIPRANDI, P., UZAN, M., SIZUN, C. & BONTEMS, F. 2009. Probing the relationship between Gram-negative and Gram-positive S1 proteins by sequence analysis. *Nucleic Acids Res,* 37**,** 5578-88.

SALIS, H. M., MIRSKY, E. A. & VOIGT, C. A. 2009. Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol,* 27**,** 946-50.

SASHITAL, D. G., GREEMAN, C. A., LYUMKIS, D., POTTER, C. S., CARRAGHER, B. & WILLIAMSON, J. R. 2014. A combined quantitative mass spectrometry and electron microscopy analysis of ribosomal 30S subunit assembly in E. coli. *Elife,* 3.

SCHRADER, J. M., ZHOU, B., LI, G. W., LASKER, K., CHILDERS, W. S., WILLIAMS, B., LONG, T., CROSSON, S., MCADAMS, H. H., WEISSMAN, J. S. & SHAPIRO, L. 2014. The coding and noncoding architecture of the Caulobacter crescentus genome. *PLoS Genet,* 10**,** e1004463.

SCHURR, T., NADIR, E. & MARGALIT, H. 1993. Identification and characterization of E.coli ribosomal binding sites by free energy computation. *Nucleic Acids Res,* 21**,** 4019-23.

SCHUWIRTH, B. S., BOROVINSKAYA, M. A., HAU, C. W., ZHANG, W., VILA-SANJURJO, A., HOLTON, J. M. & CATE, J. H. 2005. Structures of the bacterial ribosome at 3.5 A resolution. *Science,* 310**,** 827-34.

SENGUPTA, J., AGRAWAL, R. K. & 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**,** 11991-6.

SEZONOV, G., JOSELEAU-PETIT, D. & D'ARI, R. 2007. Escherichia coli physiology in Luria-Bertani broth. *J Bacteriol,* 189**,** 8746-9.

SHINE, J. & 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**,** 1342-6.
SHOJI, S., ABDI, N. M., BUNDSCHUH, R. & FREDRICK, K. 2009. Contribution of ribosomal residues to P-site tRNA binding. *Nucleic Acids Res,* 37**,** 4033-42.
SHOJI, S., WALKER, S. E. & FREDRICK, K. 2006. Reverse translocation of tRNA in the ribosome. *Mol Cell,* 24**,** 931-42.
SHULTZABERGER, R. K., BUCHEIMER, R. E., RUDD, K. E. & SCHNEIDER, T. D. 2001. Anatomy of Escherichia coli ribosome binding sites. *J Mol Biol,* 313**,** 215-28.
SKORSKI, P., LEROY, P., FAYET, O., DREYFUS, M. & HERMANN-LE DENMAT, S. 2006. The highly efficient translation initiation region from the Escherichia coli rpsA gene lacks a shine-dalgarno element. *J Bacteriol,* 188**,** 6277-85.
SOHMEN, D., CHIBA, S., SHIMOKAWA-CHIBA, N., INNIS, C. A., BERNINGHAUSEN, O., BECKMANN, R., ITO, K. & WILSON, D. N. 2015. Structure of the Bacillus subtilis 70S ribosome reveals the basis for species-specific stalling. *Nat Commun,* 6**,** 6941.
STARMER, J., STOMP, A., VOUK, M. & BITZER, D. 2006. Predicting Shine-Dalgarno sequence locations exposes genome annotation errors. *PLoS Comput Biol,* 2**,** e57.
STEITZ, J. A. & JAKES, K. 1975. How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in Escherichia coli. *Proc Natl Acad Sci U S A,* 72**,** 4734-8.
SUBRAMANIAN, A. R. & VAN DUIN, J. 1977. Exchange of individual ribosomal proteins between ribosomes as studied by heavy isotope-transfer experiments. *Mol Gen Genet,* 158**,** 1-9.
SUSSMAN, J. K., SIMONS, E. L. & SIMONS, R. W. 1996. Escherichia coli translation initiation factor 3 discriminates the initiation codon in vivo. *Mol Microbiol,* 21**,** 347-60.
SUZEK, B. E., ERMOLAEVA, M. D., SCHREIBER, M. & SALZBERG, S. L. 2001. A probabilistic method for identifying start codons in bacterial genomes. *Bioinformatics,* 17**,** 1123-30.
TAKADA, H., MORITA, M., SHIWA, Y., SUGIMOTO, R., SUZUKI, S., KAWAMURA, F. & YOSHIKAWA, H. 2014. Cell motility and biofilm formation in Bacillus subtilis are affected by the ribosomal proteins, S11 and S21. *Biosci Biotechnol Biochem,* 78**,** 898-907.
UNIPROT, C. 2019. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res,* 47**,** D506-D515.
VELLANOWETH, R. L. & RABINOWITZ, J. C. 1992. The influence of ribosome-binding-site elements on translational efficiency in Bacillus subtilis and Escherichia coli in vivo. *Mol Microbiol,* 6**,** 1105-14.
VINGADASSALOM, D., KOLB, A., MAYER, C., RYBKINE, T., COLLATZ, E. & PODGLAJEN, I. 2005. An unusual primary sigma factor in the Bacteroidetes phylum. *Mol Microbiol,* 56**,** 888-902.
WALKER, S. E. & FREDRICK, K. 2008. Preparation and evaluation of acylated tRNAs. *Methods,* 44**,** 81-6.
WEGMANN, U., HORN, N. & CARDING, S. R. 2013. Defining the bacteroides ribosomal binding site. *Appl Environ Microbiol,* 79**,** 1980-9.
WILLIAMS, C. J., HEADD, J. J., MORIARTY, N. W., PRISANT, M. G., VIDEAU, L. L., DEIS, L. N., VERMA, V., KEEDY, D. A., HINTZE, B. J., CHEN, V. B., JAIN, S., LEWIS, S. M., ARENDALL, W. B., 3RD, SNOEYINK, J., ADAMS, P. D., LOVELL, S. C.,

RICHARDSON, J. S. & RICHARDSON, D. C. 2018. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci,* 27**,** 293-315.

XIE, G., BRUCE, D. C., CHALLACOMBE, J. F., CHERTKOV, O., DETTER, J. C., GILNA, P., HAN, C. S., LUCAS, S., MISRA, M., MYERS, G. L., RICHARDSON, P., TAPIA, R., THAYER, N., THOMPSON, L. S., BRETTIN, T. S., HENRISSAT, B., WILSON, D. B. & MCBRIDE, M. J. 2007. Genome sequence of the cellulolytic gliding bacterium Cytophaga hutchinsonii. *Appl Environ Microbiol,* 73**,** 3536-46.

YAMAUCHI, K. 1991. The sequence flanking translational initiation site in protozoa. *Nucleic Acids Res,* 19**,** 2715-20.

YUTIN, N., PUIGBO, P., KOONIN, E. V. & WOLF, Y. I. 2012. Phylogenomics of prokaryotic ribosomal proteins. *PLoS One,* 7**,** e36972.

ZENG-ELMORE, X., GAO, X. Z., PELLARIN, R., SCHNEIDMAN-DUHOVNY, D., ZHANG, X. J., KOZACKA, K. A., TANG, Y., SALI, A., CHALKLEY, R. J., COTE, R. H. & CHU, F. 2014. Molecular architecture of photoreceptor phosphodiesterase elucidated by chemical cross-linking and integrative modeling. *J Mol Biol,* 426**,** 3713-3728.

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

ZHENG, S. Q., PALOVCAK, E., ARMACHE, J. P., VERBA, K. A., CHENG, Y. & AGARD, D. A. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods,* 14**,** 331-332.

ZHU, H. Q., HU, G. Q., OUYANG, Z. Q., WANG, J. & SHE, Z. S. 2004. Accuracy improvement for identifying translation initiation sites in microbial genomes. *Bioinformatics,* 20**,** 3308-17.

ZIVANOV, J., NAKANE, T., FORSBERG, B. O., KIMANIUS, D., HAGEN, W. J., LINDAHL, E. & SCHERES, S. H. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife,* 7.

CHAPTER 5: Discussion and Conclusion

5.1 Role of 'inactive' 30S in ribosome assembly

Our cryo-EM structure of the 'inactive' 30S ribosome subunit provides an opportunity to advance our understanding of the assembly and maturation of the 30S ribosome subunit. This is because the 'inactive' conformation of the 30S ribosome subunit acts as a substrate for several ribosome assembly factors and is perfectly demonstrated in the recent cryo-EM studies by (Stephan et al., 2021) and (Schedlbauer et al., 2021). Accordingly, both studies used the 30S ribosome subunits purified by dissociating 70S into individual subunits by depleting Mg^{2+} during the ribosome purification and treated the 'inactive' 30S ribosome subunits with one or several late-stage ribosome assembly factors like RbfA, YjeQ (RsgA), RimP, and RsmA (KsgA). Although we had prior knowledge of where RbfA, YjeQ, and KsgA bind to the 30S, this study using the 'inactive' 30S subunit as substrate was able to demonstrate the binding site of RimP on the 30S and how it plays a role in the maturation of the decoding region by rearranging the helix 44 and making the binding site accessible to KsgA. In the case of RbfA (14kDa protein composed of a single KH-Domain), an early low-resolution cryo-EM study showed that this ribosome assembly factor binds to the 30S between the A and P (Datta et al., 2007). However, the study from *Schedlbauer et al.* revealed RbfA binds in the platform region near the 3' end of the 16S rRNA and not the A- and P-sites as was thought earlier. We now understand the role of RbfA in the maturation of the 30S ribosome subunit because this study demonstrated that RbfA promotes the conversion of the 30S from an 'inactive' to an 'active' state. This role of RbfA is comparable to treating the 30S with heat (42 °C) in an Mg^{2+} abundant environment. Therefore, it can be stated that the role of RbfA *in vivo* is to provide the final 'activation energy' so that the 30S with a mature decoding center becomes competent for translation initiation. The study from *Schedlbauer et al.* revealed that the

‘inactive’ 30S represents an ultimate kinetically trapped intermediate that can be directed on the correct path to maturation via concerted efforts of one or several assembly factors. The cryo-EM study by *Stephan et al*. also demonstrated that due to the 16S rRNA helices 44 and 45 conformations, the ‘inactive’ 30S resembles a near-native substrate for KsgA (a methyltransferase broadly conserved across all kingdoms of life). Consequently, they captured the 30S+KsgA complex and subsequently obtained a high-resolution 3D reconstruction providing insights into how KsgA recognizes, binds, and methylates *E. coli* 16S rRNA nucleotides A1518 and A1519. Phylogenetic analysis by *Schedlbauer et al*. also revealed the existence of the ‘inactive’ conformation of the 30S conserved across all kingdoms of life. Accordingly, a recent cryo-EM analysis of the human pre-40S particles in the late stages of ribosome assembly also contains helix 44 in a displaced position, which partially resembles the ‘inactive 30S’ identified in Chapter 2 of this thesis (Plassart et al., 2021). Another cryo-EM study also recently reported the cryo-EM structure of the ‘inactive’ 30S purified from gram-positive bacteria *Staphylococcus aureus* and its conversion to the ‘active’ state using spermidine in a nominal Mg^{2+} environment (Belinite et al., 2021). Therefore, co-incubation of ‘inactive’ 30S with spermidine has been used as an alternative mechanism instead of heat treatment to activate 30S and study different stages of the translation cycle.

5.2 30S head dynamics and ribosome mRNA helicase activity

Our cryo-EM analysis of the free 30S ribosome subunits reported that the head domain demonstrates large-scale movements such as tilting and swiveling via its 16S rRNA helices in the head region. While these motions are essential for the translocation of the mRNA-tRNA, antibiotics like spectinomycin and sparsomycin have been discovered to trap the 30S head in the swiveled state or promote translocation independent of EF-G, respectively (Mohan et al., 2014)

(Takyar et al., 2005). However, resistance to these antibiotics has also been observed. An additional role of the 30S head movement during translation is to unwind duplex mRNAs at the mRNA entrance channel. Extensive biochemical studies have shown that the ribosome acts as a helicase through its r-proteins uS4, and uS3 in the 30S body and head respectively (Takyar et al., 2005). Accordingly, the 30S head rotation places the r-protein uS3 near the mRNA entrance channel already lined up by r-proteins uS4 and uS5. This placement allows the r-proteins uS3, uS4, and uS5 to form a helicase processivity clamp and help unwind an mRNA duplex at the entrance channel. No antibiotics have been discovered that can inhibit ribosome helicase activity making the uS3, uS4, and uS5 helicase clamp an attractive target for developing novel antibiotics. This may also be achieved by targeting ribosome assembly factors such as YjeQ, RimM, GTPase Era, and RimP since quantitative mass spectrometry analysis has shown depleted levels of uS3 in the absence of these assembly factors (Leong et al., 2013) (Thurlow et al., 2016).

5.3 YjeQ – a true quality control manager

Cryo-EM analysis of the $30S_{\Delta yjeQ}$ dataset also provides a unique opportunity to understand how other ribosome assembly factors work. Since YjeQ is a non-essential ribosome assembly factor, its absence does not stop ribosome assembly but slows it, which allows us to capture these assembly intermediates. While the particle accumulated in the $30S_{\Delta yjeQ}$ dataset may not be the true substrate for YjeQ as previously shown by with the Microscale Thermophoresis experiments, these particles may be bona fide substrates for other assembly factors (Thurlow et al., 2016). Accordingly, ribosome assembly factor KsgA may also play a direct role in stabilizing 16S rRNA helices near its binding site around the decoding center and the platform and promoting maturation of the 30S ribosome subunit. Previous cryo-EM studies have shown that the N-terminal domain of KsgA interacts with 16S rRNA helix 45, where it methylates A1518 and A1519, and its C-terminal

domain interacts with 16S rRNA helices 24 and 27 (Boehringer et al., 2012) (Stephan et al., 2021) (Schedlbauer et al., 2021). These cryo-EM studies show that a 30S particle with displaced helix 44 is a true substrate for KsgA. Therefore, the particles contributing to Group 2 in our $30S_{\Delta yjeQ}$ dataset represent a perfect substrate for KsgA. Consequently, it can be suggested that apart from the methylation function, KsgA may assist the folding of helix 27. Therefore, the maturation of helix 27 may help stabilize the platform and head domains due to the proximity and continuity of helix 27 to the platform and the head regions, respectively. These observations make 16S rRNA helix 27 an attractive target for developing novel antibiotics since it's accurate folding and placement governs the timely maturation of the 30S ribosome subunit.

Biochemical studies from the Woodson lab have demonstrated that the ability of YjeQ to dock the helix 44 from an 'inactive' state to an 'active' conformation is not meant for the handover of the mature 30S from the ribosome assembly to the translation initiation mechanism (Sharma and Woodson, 2020).To this end, we demonstrated that the helix 44 docking function of YjeQ serves to probe 16S rRNA nucleotide A1493 which during translation flips along with A1492, to monitor the geometry of the codon-anticodon helix. This helix 44 docking function may also be necessary to facilitate the binding of NpmA, a protein belonging to the kanamycin–apramycin methyltransferase (Kam) family. This methyltransferase is known to methylate the N1 position of A1408 in 16S rRNA and provide resistance to aminoglycoside. However, the binding of NpmA to the 30S strongly depends on recognizing the sugar-phosphates in the backbone of helices 24, 27, 44, and 45 (Dunkle et al., 2014). Thus, suggesting that the rRNA chaperoning in the decoding region by YjeQ may be necessary for the methylation function of NpmA.

Understanding the defects that arise due to deletion or depletion of assembly factors can provide a window to predict defects that may occur during the several steps involved in the translation cycle.

For example, the β-galactosidase assays have shown that the Δ*ksgA* strain has compromised the ability to detect start codons accurately (Connolly and Culver, 2013). This is because the binding sites of KsgA and IF3 overlap; therefore, KsgA deletion impairs the binding of initiation factor 3 (IF3) to the $30S_{\Delta ksgA}$ particles. Likewise, RbfA deletion also leads to a compromised start codon selection (Sharma and Anand, 2019). Cryo-EM studies have shown that the $30S_{\Delta rbfA}$ particles 16S rRNA helices 1, 2, and 3 forming the pseudoknot region are unstable (Maksimova et al., 2021). As a result, this instability makes the 30S head unstable and impairs the 70S initiation complex formation (Poot et al., 1996). Similarly, the frameshift defects observed in the Δ*yjeQ* strain correlate with the disorder marked in the A-site of the 30S. Due to this defect in the decoding region, elongating ribosomes tend to accept a non-cognate tRNA which triggers conformational changes in the 30S head and promotes frameshifting of the tRNA and mRNA via EF-G mediated translocation (Demo et al., 2021).

Structural analysis of YjeQ N-terminal extension (NTE) provided insights into an uncharacterized role of YjeQ as a rescue factor due to its resemblance to YaeJ (also known as ArfA) (Gagnon et al., 2012). SmpB, BrfA, ArfB, and ICT1 are several other rescue factors with a structural and compositional similarity of their C-termini with the YjeQ NTE and recognize an empty A-site but specifically when ribosomes are stalled at the 3’ end of truncated mRNA (Kurita and Himeno, 2022). This suggests that the YjeQ quality control function is not limited to translational fidelity but may also ensure that future ribosome rescue mechanisms will work accurately. In conclusion, targeting YjeQ for its role in ribosome assembly, translation fidelity, and potential role in ribosome rescue offers several incentives for developing new antibiotics.

5.4 Do 'inactive' 30S ribosome subunits exist in *Flavobacterium johnsoniae*?

Our cryo-EM analysis of the *F. johnsoniae* 70S (Fjo 70S) ribosome revealed that the 3' end of the 16S rRNA containing the anti-Shine Dalgarno sequence is sequestered to a network of interactions that it forms with neighboring ribosome proteins bS21, bS18, and bS6. However, it is essential to note that 16S 3' end positioning observed in our cryo-EM structure is in the context of the 70S ribosome subunit. Since our work does not include the cryo-EM structure of the free 30S ribosome, the positioning of the 3'end of the 16S rRNA in the context of the free 30S ribosome subunits is unknown. However, information obtained from our cryo-EM analysis, multiple sequencing alignments, and the prior knowledge from the Nomura assembly maps suggest that *F. johnsoniae* 30S (Fjo 30S) may never adopt an 'inactive' conformation that is typically observed in the *E. coli* 30S ribosome subunits and could be due to following reasons:

(1) If the Fjo 30S assembly follows the r-protein binding hierarchy based on the classical Nomura Assembly map, then the binding of secondary ribosome proteins bS6 and bS18 to the platform region creates a binding pocket for tertiary protein bS21. As a result, the binding site for the anti-Shine-Dalgarno sequence is ready before the 3' end of the 16S rRNA is transcribed by RNA Polymerase.

(2) Since bS6, bS18, and bS21 also have Fjo-specific residues that form an extensive network of H-bonds and van der Waals contacts with the anti-Shine-Dalgarno sequence, the nascently transcribed 3' end of the 16S rRNA may not get an opportunity to fall back and occupy the mRNA channel as it is typically observed in the *E. coli* 'inactive' 30S. Further, cryo-EM analysis of the purified Fjo 30S particles will be required to confirm this hypothesis.

5.5 Structural differences between *Flavobacterium johnsoniae* and other bacteria provide evolutionary insights

Our study reported overall structural differences between the 30S from the Fjo 70S and the 30S from 70S ribosomes of several other bacterial species. A relevant difference in the context of the 30S platform region was observed in 16S rRNA helix 26 as this helix was shorter and less extended in the *F. johnsoniae* than in other bacteria. Typically, bacteria that use the SD-antiSD recognition mechanism to position the start codon form the SD-antiSD helix in the 30S platform region. The SD-antiSD helix is accommodated by interactions of the anti-SD phosphate backbone with the helix 26, which restricts its movement in the platform region and may contribute to the accurate positioning of the start codon in the 30S P-site. Since *F. johnsoniae* translation initiation does not require the recognition of anti-SD by the SD for the correct position of the start codon in the P-site, these bacteria may not require helix 26 for accommodation and stability in the platform region (Kaminishi et al., 2007) (Korostelev et al., 2007). Accordingly, *F. johnsoniae* may have evolved to shorten the length of the rRNA helices that may not be essential for ribosome function or structural integrity. We report several such shortened rRNA for both subunits in our study. Such an approach is beneficial for avoiding unnecessary energy expenditure since ribosome biogenesis is energetically the most expensive process in all cells.

5.6 Need to adapt for next-generation cryo-EM sample preparation

The biggest drawback of studying ribosomes using cryo-EM is that, like most particles, ribosomes too are susceptible to migrating to the air-water interface (AWI), which leads to partial denaturation of one or several ribosomal components. Migration of particles to the air-water interface is promoted due to the blotting time required by the Vitrobot (between 3 to 6 seconds) to remove the excess sample applied to the grid before plunge freezing. Even ribosomes that are

generally considered robust specimens for cryo-EM are susceptible to the AWI problem. In our case, r-protein uS2 was mainly sensitive to the dissociation or denaturation and results in the loss from the *E. coli* 30S and the *F. Johnsoniae* 70S. Other groups have obtained similar results studying the 30S ribosome subunit on a cryo-EM grid with no support layer (typically 3-4 nm of continuous carbon) (Klebl et al., 2020). While using the continuous carbon layer is a standard solution to negate the effects of the AWI, this method introduces unnecessary electron scattering in the background, which lowers the signal-to-noise ratio in the micrographs, steals contrast from particles. Future ribosome studies using cryo-EM will benefit from recently developed vitrification equipment, for example, Chameleon, Shake-it-off, or time-resolved cryo-EM devices based on the microfluidics (Kaledhonkar et al., 2018) (Rubinstein et al., 2019) (Dandey et al., 2020). While Chameleon is commercially available, the Shake-it-off or time-resolved cryo-EM devices can be manufactured in-house. These devices can deposit optimized amounts of samples on the cryo-EM grids, plunge freezes them as fast as 6 milliseconds, and prevents the loss of r-proteins (at least from the 30S) at the AWI (Klebl et al., 2020). Since this approach does not involve using an additional support layer, micrographs from these cryo-EM grids provide particles with a high signal-to-noise ratio, thus helping boost resolution for 3D reconstruction. In addition, the microfluidics-based time-resolved sample preparation devices from the Frank lab also capture short-lived ribosome-factor complexes. Significant advancements in the hardware aspect of cryo-EM have made it possible to achieve an atomic resolution (Yip et al., 2020) (Nakane et al., 2020). However, sample preparation remains a major barrier for the observation of complexes or specific intermediates that are more labile and fall apart more easily upon exposure to AWI.

Conclusion and Summary

The focus of this thesis was to explore aspects of ribosome biology that are poorly understood and constitute potential targets for the development of new antibiotics. Accordingly:

(1) In chapter 2, we solved the structure of the 'inactive' 30S that was unknown for more than five decades. We visualized the conformational transition between the 'active' and 'inactive' state of the 30S and different motions exhibited by a free 30S in solution. Consequently, this study will facilitate our understanding of how the 'active–inactive' transition promotes the binding and displacement of ribosome assembly and translation initiation factors.

(2) In chapter 3, we demonstrate how the absence of YjeQ causes chaos and delays in the assembly of the 30S ribosome subunit. We also show that YjeQ acts as a quality control protein by ensuring that the assembled 30S can accurately perform the 'initial decoding' and thus guarantees high translational fidelity. We also identified a potential role for YjeQ in ribosome rescue.

(3) In chapter 4, we described why Bacteroidetes ribosomes fail to recognize SD sequence on the mRNA. Our high-resolution cryo-EM structure revealed that r-proteins bS21, bS18, and bS6 have residues unique to Bacteroidetes that form a binding pocket and sequester the ASD on ribosomes. We also proposed that Bacteroidetes use this ASD occlusion mechanism to autoregulate ribosome biogenesis. Finally, this study shows that mRNA selection mechanisms are not conserved throughout bacterial species.

In conclusion, my thesis provides insights into several untapped aspects of the ribosome that can serve as targets for developing new antibiotics, which we desperately need to combat rapidly rising antibiotic resistance.

Bibliography

ACCETTO, T. & AVGUSTIN, G. 2011. Inability of Prevotella bryantii to form a functional Shine-Dalgarno interaction reflects unique evolution of ribosome binding sites in Bacteroidetes. *PLoS One,* 6**,** e22914.

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., MCCOY, A. J., MORIARTY, N. W., OEFFNER, R., READ, R. J., RICHARDSON, D. C., RICHARDSON, J. S., TERWILLIGER, T. C. & ZWART, P. H. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr,* 66**,** 213-21.

ADILAKSHMI, T., BELLUR, D. L. & WOODSON, S. A. 2008. Concurrent nucleation of 16S folding and induced fit in 30S ribosome assembly. *Nature,* 455**,** 1268-72.

AFONINE, P. V., KLAHOLZ, B. P., MORIARTY, N. W., POON, B. K., SOBOLEV, O. V., TERWILLIGER, T. C., ADAMS, P. D. & URZHUMTSEV, A. 2018a. New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr D Struct Biol,* 74**,** 814-840.

AFONINE, P. V., POON, B. K., READ, R. J., SOBOLEV, O. V., TERWILLIGER, T. C., URZHUMTSEV, A. & ADAMS, P. D. 2018b. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol,* 74**,** 531-544.

AKANUMA, G., KOBAYASHI, A., SUZUKI, S., KAWAMURA, F., SHIWA, Y., WATANABE, S., YOSHIKAWA, H., HANAI, R. & ISHIZUKA, M. 2014. Defect in the formation of 70S ribosomes caused by lack of ribosomal protein L34 can be suppressed by magnesium. *J Bacteriol,* 196**,** 3820-30.

AKANUMA, G., NANAMIYA, H., NATORI, Y., YANO, K., SUZUKI, S., OMATA, S., ISHIZUKA, M., SEKINE, Y. & KAWAMURA, F. 2012. Inactivation of ribosomal protein genes in Bacillus subtilis reveals importance of each ribosomal protein for cell proliferation and cell differentiation. *J Bacteriol,* 194**,** 6282-91.

AKANUMA, G., YAMAZAKI, K., YAGISHI, Y., IIZUKA, Y., ISHIZUKA, M., KAWAMURA, F. & KATO-YAMADA, Y. 2018. Magnesium Suppresses Defects in the Formation of 70S Ribosomes as Well as in Sporulation Caused by Lack of Several Individual Ribosomal Proteins. *J Bacteriol,* 200.

AMIN, M. R., YUROVSKY, A., CHEN, Y., SKIENA, S. & FUTCHER, B. 2018. Re-annotation of 12,495 prokaryotic 16S rRNA 3' ends and analysis of Shine-Dalgarno and anti-Shine-Dalgarno sequences. *PLoS One,* 13**,** e0202767.

ANDREWS, E. S. V. & PATRICK, W. M. 2022. The hypothesized role of YbeZ in 16S rRNA maturation. *Arch Microbiol,* 204**,** 114.

ANTOUN, A., PAVLOV, M. Y., LOVMAR, M. & EHRENBERG, M. 2006. How initiation factors tune the rate of initiation of protein synthesis in bacteria. *EMBO J,* 25**,** 2539-50.

ARIGONI, F., TALABOT, F., PEITSCH, M., EDGERTON, M. D., MELDRUM, E., ALLET, E., FISH, R., JAMOTTE, T., CURCHOD, M. L. & LOFERER, H. 1998. A genome-based approach for the identification of essential bacterial genes. *Nat Biotechnol,* 16**,** 851-6.

BABA, T., ARA, T., HASEGAWA, M., TAKAI, Y., OKUMURA, Y., BABA, M., DATSENKO, K. A., TOMITA, M., WANNER, B. L. & MORI, H. 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol,* 2**,** 2006 0008.

BAEZ, W. D., ROY, B., MCNUTT, Z. A., SHATOFF, E. A., CHEN, S., BUNDSCHUH, R. & FREDRICK, K. 2019. Global analysis of protein synthesis in Flavobacterium johnsoniae reveals the use of Kozak-like sequences in diverse bacteria. *Nucleic Acids Res,* 47**,** 10477-10488.

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., MCALEAR, M. A., MOORE, P. B., NOLLER, H. F., ORTEGA, J., PANSE, V. G., RAMAKRISHNAN, V., SPAHN, C. M., STEITZ, T. A., TCHORZEWSKI, M., TOLLERVEY, D., WARREN, A. J., WILLIAMSON, J. R., WILSON, D., YONATH, A. & YUSUPOV, M. 2014. A new system for naming ribosomal proteins. *Curr Opin Struct Biol,* 24**,** 165-9.

BAN, N., NISSEN, P., HANSEN, J., MOORE, P. B. & STEITZ, T. A. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science,* 289**,** 905-20.

BAYLEY, D. P., ROCHA, E. R. & SMITH, C. J. 2000. Analysis of cepA and other Bacteroides fragilis genes reveals a unique promoter structure. *FEMS Microbiol Lett,* 193**,** 149-54.

BELINITE, M., KHUSAINOV, I., SOUFARI, H., MARZI, S., ROMBY, P., YUSUPOV, M. & HASHEM, Y. 2021. Stabilization of Ribosomal RNA of the Small Subunit by Spermidine in Staphylococcus aureus. *Front Mol Biosci,* 8**,** 738752.

BERINGER, M., BRUELL, C., XIONG, L., PFISTER, P., BIELING, P., KATUNIN, V. I., MANKIN, A. S., BOTTGER, E. C. & RODNINA, M. V. 2005. Essential mechanisms in the catalysis of peptide bond formation on the ribosome. *J Biol Chem,* 280**,** 36065-72.

BESEMER, J., LOMSADZE, A. & BORODOVSKY, M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res,* 29**,** 2607-18.

BOEHRINGER, D., O'FARRELL, H. C., RIFE, J. P. & BAN, N. 2012. Structural insights into methyltransferase KsgA function in 30S ribosomal subunit biogenesis. *J Biol Chem,* 287**,** 10453-10459.

BOROVINSKAYA, M. A., PAI, R. D., ZHANG, W., SCHUWIRTH, B. S., HOLTON, J. M., HIROKAWA, G., KAJI, H., KAJI, A. & CATE, J. H. 2007. Structural basis for aminoglycoside inhibition of bacterial ribosome recycling. *Nat Struct Mol Biol,* 14**,** 727-32.

BRITTON, R. A. 2009. Role of GTPases in bacterial ribosome assembly. *Annu Rev Microbiol,* 63**,** 155-76.

BUNNER, A. E., BECK, A. H. & WILLIAMSON, J. R. 2010. Kinetic cooperativity in Escherichia coli 30S ribosomal subunit reconstitution reveals additional complexity in the assembly landscape. *Proc Natl Acad Sci U S A,* 107**,** 5417-22.

BUTLER, J. S., SPRINGER, M., DONDON, J., GRAFFE, M. & GRUNBERG-MANAGO, M. 1986. Escherichia coli protein synthesis initiation factor IF3 controls its own gene expression at the translational level in vivo. *J Mol Biol,* 192**,** 767-80.

CAMPBELL, T. L. & BROWN, E. D. 2008. Genetic interaction screens with ordered overexpression and deletion clone sets implicate the Escherichia coli GTPase YjeQ in late ribosome biogenesis. *J Bacteriol,* 190**,** 2537-45.

CAMPBELL, T. L., DAIGLE, D. M. & BROWN, E. D. 2005. Characterization of the Bacillus subtilis GTPase YloQ and its role in ribosome function. *Biochem J,* 389**,** 843-52.

CARBONE, C. E., LOVELAND, A. B., GAMPER, H. B., JR., HOU, Y. M., DEMO, G. & KOROSTELEV, A. A. 2021. Time-resolved cryo-EM visualizes ribosomal translocation with EF-G and GTP. *Nat Commun,* 12**,** 7236.

CARTER, A. P., CLEMONS, W. M., BRODERSEN, D. E., MORGAN-WARREN, R. J., WIMBERLY, B. T. & RAMAKRISHNAN, V. 2000. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature,* 407**,** 340-8.

CARTER, A. P., CLEMONS, W. M., JR., BRODERSEN, D. E., MORGAN-WARREN, R. J., HARTSCH, T., WIMBERLY, B. T. & RAMAKRISHNAN, V. 2001. Crystal structure of an initiation factor bound to the 30S ribosomal subunit. *Science,* 291**,** 498-501.

CHADANI, Y., MATSUMOTO, E., ASO, H., WADA, T., KUTSUKAKE, K., SUTOU, S. & ABO, T. 2011. trans-translation-mediated tight regulation of the expression of the alternative ribosome-rescue factor ArfA in Escherichia coli. *Genes Genet Syst,* 86**,** 151-63.

CHEN, S., BAGDASARIAN, M., KAUFMAN, M. G., BATES, A. K. & WALKER, E. D. 2007a. Mutational analysis of the ompA promoter from Flavobacterium johnsoniae. *J Bacteriol,* 189**,** 5108-18.

CHEN, S., BAGDASARIAN, M., KAUFMAN, M. G. & WALKER, E. D. 2007b. Characterization of strong promoters from an environmental Flavobacterium hibernum strain by using a green fluorescent protein-based reporter system. *Appl Environ Microbiol,* 73**,** 1089-100.

CHEN, S. S., SPERLING, E., SILVERMAN, J. M., DAVIS, J. H. & WILLIAMSON, J. R. 2012. Measuring the dynamics of E. coli ribosome biogenesis using pulse-labeling and quantitative mass spectrometry. *Mol Biosyst,* 8**,** 3325-34.

CHEN, S. S. & WILLIAMSON, J. R. 2013. Characterization of the Ribosome Biogenesis Landscape in E. coli Using Quantitative Mass Spectrometry. *J Mol Biol,* 425**,** 767-79.

CHENG, Y. 2015. Single-Particle Cryo-EM at Crystallographic Resolution. *Cell,* 161**,** 450-7.

CHENG, Y., GLAESER, R. M. & NOGALES, E. 2017. How Cryo-EM Became so Hot. *Cell,* 171**,** 1229-1231.

CHRISTIANSEN, J. 1988. The 9S RNA precursor of Escherichia coli 5S RNA has three structural domains: implications for processing. *Nucleic Acids Res,* 16**,** 7457-76.

COCOZAKI, A. I., ALTMAN, R. B., HUANG, J., BUURMAN, E. T., KAZMIRSKI, S. L., DOIG, P., PRINCE, D. B., BLANCHARD, S. C., CATE, J. H. & FERGUSON, A. D. 2016. Resistance mutations generate divergent antibiotic susceptibility profiles against translation inhibitors. *Proc Natl Acad Sci U S A,* 113**,** 8188-93.

CONNOLLY, K. & CULVER, G. 2009. Deconstructing ribosome construction. *Trends Biochem Sci,* 34**,** 256-63.

CONNOLLY, K. & CULVER, G. 2013. Overexpression of RbfA in the absence of the KsgA checkpoint results in impaired translation initiation. *Mol Microbiol,* 87**,** 968-81.

CROOKS, G. E., HON, G., CHANDONIA, J. M. & BRENNER, S. E. 2004. WebLogo: a sequence logo generator. *Genome Res,* 14**,** 1188-90.

D'IMPRIMA, E., FLORIS, D., JOPPE, M., SANCHEZ, R., GRININGER, M. & KUHLBRANDT, W. 2019. Protein denaturation at the air-water interface and how to prevent it. *Elife,* 8.

DAIGLE, D. M. & BROWN, E. D. 2004. Studies of the interaction of Escherichia coli YjeQ with the ribosome in vitro. *J Bacteriol,* 186**,** 1381-7.

DAIGLE, D. M., ROSSI, L., BERGHUIS, A. M., ARAVIND, L., KOONIN, E. V. & BROWN, E. D. 2002. YjeQ, an essential, conserved, uncharacterized protein from Escherichia coli, is an unusual GTPase with circularly permuted G-motifs and marked burst kinetics. *Biochemistry,* 41**,** 11109-17.

DALLAS, A. & NOLLER, H. F. 2001. Interaction of translation initiation factor 3 with the 30S ribosomal subunit. *Mol Cell,* 8**,** 855-64.

DANDEY, V. P., BUDELL, W. C., WEI, H., BOBE, D., MARUTHI, K., KOPYLOV, M., ENG, E. T., KAHN, P. A., HINSHAW, J. E., KUNDU, N., NIMIGEAN, C. M., FAN, C., SUKOMON, N., DARST, S. A., SAECKER, R. M., CHEN, J., MALONE, B., POTTER, C. S. & CARRAGHER, B. 2020. Time-resolved cryo-EM using Spotiton. *Nat Methods,* 17**,** 897-900.

DARTY, K., DENISE, A. & PONTY, Y. 2009. VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics,* 25**,** 1974-5.

DATSENKO, K. A. & WANNER, B. L. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A,* 97**,** 6640-5.

DATTA, P. P., WILSON, D. N., KAWAZOE, M., SWAMI, N. K., KAMINISHI, T., SHARMA, M. R., BOOTH, T. M., TAKEMOTO, C., FUCINI, P., YOKOYAMA, S. & AGRAWAL, R. K. 2007. Structural aspects of RbfA action during small ribosomal subunit assembly. *Mol Cell,* 28**,** 434-45.

DAVIES, B. W., KOHRER, C., JACOB, A. I., SIMMONS, L. A., ZHU, J., ALEMAN, L. M., RAJBHANDARY, U. L. & WALKER, G. C. 2010. Role of Escherichia coli YbeY, a highly conserved protein, in rRNA processing. *Mol Microbiol,* 78**,** 506-18.

DAVIS, J. H., TAN, Y. Z., CARRAGHER, B., POTTER, C. S., LYUMKIS, D. & WILLIAMSON, J. R. 2016. Modular Assembly of the Bacterial Large Ribosomal Subunit. *Cell,* 167**,** 1610-1622 e15.

DAVIS, J. H. & WILLIAMSON, J. R. 2017. Structure and dynamics of bacterial ribosome biogenesis. *Philos Trans R Soc Lond B Biol Sci,* 372.

DE LA ROSA-TREVIN, J. M., OTON, J., MARABINI, R., ZALDIVAR, A., VARGAS, J., CARAZO, J. M. & SORZANO, C. O. 2013. Xmipp 3.0: an improved software suite for image processing in electron microscopy. *J Struct Biol,* 184**,** 321-8.

DE SMIT, M. H. & VAN DUIN, J. 1994. Translational initiation on structured messengers. Another role for the Shine-Dalgarno interaction. *J Mol Biol,* 235**,** 173-84.

DELANO, W. L. 2002. *The PyMOL Molecular Graphic Systems,* Palo Alto, CA, USA, DeLano Scientific.

DELCHER, A. L., BRATKE, K. A., POWERS, E. C. & SALZBERG, S. L. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics,* 23**,** 673-9.

DEMO, G., GAMPER, H. B., LOVELAND, A. B., MASUDA, I., CARBONE, C. E., SVIDRITSKIY, E., HOU, Y. M. & KOROSTELEV, A. A. 2021. Structural basis for +1 ribosomal frameshifting during EF-G-catalyzed translocation. *Nat Commun,* 12**,** 4644.

DIACONU, M., KOTHE, U., SCHLUNZEN, F., FISCHER, N., HARMS, J. M., TONEVITSKY, A. G., STARK, H., RODNINA, M. V. & WAHL, M. C. 2005. Structural basis for the function of the ribosomal L7/12 stalk in factor binding and GTPase activation. *Cell,* 121**,** 991-1004.

DUBOCHET, J., ADRIAN, M., CHANG, J. J., HOMO, J. C., LEPAULT, J., MCDOWALL, A. W. & SCHULTZ, P. 1988. Cryo-electron microscopy of vitrified specimens. *Q Rev Biophys,* 21**,** 129-228.

DUNKLE, J. A., VINAL, K., DESAI, P. M., ZELINSKAYA, N., SAVIC, M., WEST, D. M., CONN, G. L. & DUNHAM, C. M. 2014. Molecular recognition and modification of the 30S ribosome by the aminoglycoside-resistance methyltransferase NpmA. *Proc Natl Acad Sci U S A,* 111**,** 6275-80.

DUNKLE, J. A., WANG, L., FELDMAN, M. B., PULK, A., CHEN, V. B., KAPRAL, G. J., NOESKE, J., RICHARDSON, J. S., BLANCHARD, S. C. & CATE, J. H. 2011. Structures of the bacterial ribosome in classical and hybrid states of tRNA binding. *Science,* 332**,** 981-4.

DUSS, O., STEPANYUK, G. A., PUGLISI, J. D. & WILLIAMSON, J. R. 2019. Transient Protein-RNA Interactions Guide Nascent Ribosomal RNA Folding. *Cell,* 179**,** 1357-1369 e16.

ELVEKROG, M. M. & GONZALEZ, R. L., JR. 2013. Conformational selection of translation initiation factor 3 signals proper substrate selection. *Nat Struct Mol Biol,* 20**,** 628-33.

EMSLEY, P. & COWTAN, K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr,* 60**,** 2126-32.

EMSLEY, P., LOHKAMP, B., SCOTT, W. G. & COWTAN, K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr,* 66**,** 486-501.

FAHLMAN, R. P. & UHLENBECK, O. C. 2004. Contribution of the esterified amino acid to the binding of aminoacylated tRNAs to the ribosomal P- and A-sites. *Biochemistry,* 43**,** 7575-83.

FRANK, J. 2017. The translation elongation cycle-capturing multiple states by cryo-electron microscopy. *Philos Trans R Soc Lond B Biol Sci,* 372.

FRANK, J. & AGRAWAL, R. K. 2000. A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature,* 406**,** 318-22.

FREDRICK, K. & NOLLER, H. F. 2002. Accurate translocation of mRNA by the ribosome requires a peptidyl group or its analog on the tRNA moving into the 30S P site. *Mol Cell,* 9**,** 1125-31.

FU, Z., INDRISIUNAITE, G., KALEDHONKAR, S., SHAH, B., SUN, M., CHEN, B., GRASSUCCI, R. A., EHRENBERG, M. & FRANK, J. 2019. The structural basis for release-factor activation during translation termination revealed by time-resolved cryogenic electron microscopy. *Nat Commun,* 10**,** 2579.

GAGNON, M. G., SEETHARAMAN, S. V., BULKLEY, D. & STEITZ, T. A. 2012. Structural basis for the rescue of stalled ribosomes: structure of YaeJ bound to the ribosome. *Science,* 335**,** 1370-2.

GALLANT, J. A. & LINDSLEY, D. 1998. Ribosomes can slide over and beyond "hungry" codons, resuming protein chain elongation many nucleotides downstream. *Proc Natl Acad Sci U S A,* 95**,** 13771-6.

GESTELAND, R. F. 1966. Unfolding of Escherichia coli ribosomes by removal of magnesium. *J Mol Biol,* 18**,** 356-71.

GINSBURG, D. & STEITZ, J. A. 1975. The 30 S ribosomal precursor RNA from Escherichia coli. A primary transcript containing 23 S, 16 S, and 5 S sequences. *J Biol Chem,* 250**,** 5647-54.

GODDARD, T. D., HUANG, C. C., MENG, E. C., PETTERSEN, E. F., COUCH, G. S., MORRIS, J. H. & FERRIN, T. E. 2018. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci,* 27**,** 14-25.

GOLD, L., STORMO, G. & SAUNDERS, R. 1984. Escherichia coli translational initiation factor IF3: a unique case of translational regulation. *Proc Natl Acad Sci U S A,* 81**,** 7061-5.

GOTO, S., KATO, S., KIMURA, T., MUTO, A. & HIMENO, H. 2011. RsgA releases RbfA from 30S ribosome during a late stage of ribosome biosynthesis. *EMBO J,* 30**,** 104-14.

GREGORY, S. T., CARR, J. F., RODRIGUEZ-CORREA, D. & DAHLBERG, A. E. 2005. Mutational analysis of 16S and 23S rRNA genes of Thermus thermophilus. *J Bacteriol,* 187**,** 4804-12.

GUO, Q., GOTO, S., CHEN, Y., FENG, B., XU, Y., MUTO, A., HIMENO, H., DENG, H., LEI, J. & GAO, N. 2013. Dissecting the in vivo assembly of the 30S ribosomal subunit reveals the role of RimM and general features of the assembly process. *Nucleic Acids Res,* 41**,** 2609-20.

HAHNKE, R. L., MEIER-KOLTHOFF, J. P., GARCIA-LOPEZ, M., MUKHERJEE, S., HUNTEMANN, M., IVANOVA, N. N., WOYKE, T., KYRPIDES, N. C., KLENK, H. P. & GOKER, M. 2016. Genome-Based Taxonomic Classification of Bacteroidetes. *Front Microbiol,* 7**,** 2003.

HALFON, Y., JIMENEZ-FERNANDEZ, A., LA ROSA, R., ESPINOSA PORTERO, R., KROGH JOHANSEN, H., MATZOV, D., EYAL, Z., BASHAN, A., ZIMMERMAN, E., BELOUSOFF, M., MOLIN, S. & YONATH, A. 2019. Structure of Pseudomonas aeruginosa ribosomes from an aminoglycoside-resistant clinical isolate. *Proc Natl Acad Sci U S A,* 116**,** 22275-22281.

HARTZ, D., MCPHEETERS, D. S. & GOLD, L. 1989. Selection of the initiator tRNA by Escherichia coli initiation factors. *Genes Dev,* 3**,** 1899-912.

HENTSCHEL, J., BURNSIDE, C., MIGNOT, I., LEIBUNDGUT, M., BOEHRINGER, D. & BAN, N. 2017. The Complete Structure of the Mycobacterium smegmatis 70S Ribosome. *Cell Rep,* 20**,** 149-160.

HILLEBRAND, A., WURM, R., MENZEL, A. & WAGNER, R. 2005. The seven E. coli ribosomal RNA operon upstream regulatory regions differ in structure and transcription factor binding efficiencies. *Biol Chem,* 386**,** 523-34.

HIMENO, H., HANAWA-SUETSUGU, K., KIMURA, T., TAKAGI, K., SUGIYAMA, W., SHIRATA, S., MIKAMI, T., ODAGIRI, F., OSANAI, Y., WATANABE, D., GOTO, S., KALACHNYUK, L., USHIDA, C. & MUTO, A. 2004. A novel GTPase activated by the small subunit of ribosome. *Nucleic Acids Res,* 32**,** 5303-9.

HUI, A. & DE BOER, H. A. 1987. Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in Escherichia coli. *Proc Natl Acad Sci U S A,* 84**,** 4762-6.

HUSSAIN, T., LLACER, J. L., WIMBERLY, B. T., KIEFT, J. S. & RAMAKRISHNAN, V. 2016. Large-Scale Movements of IF3 and tRNA during Bacterial Translation Initiation. *Cell,* 167**,** 133-144 e13.

JACOB, A. I., KOHRER, C., DAVIES, B. W., RAJBHANDARY, U. L. & WALKER, G. C. 2013. Conserved bacterial RNase YbeY plays key roles in 70S ribosome quality control and 16S rRNA maturation. *Mol Cell,* 49**,** 427-38.

JACOB, W. F., SANTER, M. & DAHLBERG, A. E. 1987. A single base change in the Shine-Dalgarno region of 16S rRNA of Escherichia coli affects translation of many proteins. *Proc Natl Acad Sci U S A,* 84**,** 4757-61.

JAHAGIRDAR, D., JHA, V., BASU, K., GOMEZ-BLANCO, J., VARGAS, J. & ORTEGA, J. 2020. Alternative conformations and motions adopted by 30S ribosomal subunits visualized by cryo-electron microscopy. *RNA,* 26**,** 2017-2030.

JANOWSKI, M., ZOSCHKE, R., SCHARFF, L. B., MARTINEZ JAIME, S., FERRARI, C., PROOST, S., NG WEI XIONG, J., OMRANIAN, N., MUSIALAK-LANGE, M., NIKOLOSKI, Z., GRAF, A., SCHOTTLER, M. A., SAMPATHKUMAR, A., VAID, N.

& MUTWIL, M. 2018. AtRsgA from Arabidopsis thaliana is important for maturation of the small subunit of the chloroplast ribosome. *Plant J,* 96**,** 404-420.
JEGANATHAN, A., RAZI, A., THURLOW, B. & ORTEGA, J. 2015. The C-terminal helix in the YjeQ zinc-finger domain catalyzes the release of RbfA during 30S ribosome subunit assembly. *RNA,* 21**,** 1203-16.
JENNER, L. B., DEMESHKINA, N., YUSUPOVA, G. & YUSUPOV, M. 2010. Structural aspects of messenger RNA reading frame maintenance by the ribosome. *Nat Struct Mol Biol,* 17**,** 555-60.
JOHNSON, E. L., HEAVER, S. L., WALTERS, W. A. & LEY, R. E. 2017. Microbiome and metabolic disease: revisiting the bacterial phylum Bacteroidetes. *J Mol Med (Berl),* 95**,** 1-8.
JOMAA, A., JAIN, N., DAVIS, J. H., WILLIAMSON, J. R., BRITTON, R. A. & ORTEGA, J. 2014. Functional domains of the 50S subunit mature late in the assembly process. *Nucleic Acids Res,* 42**,** 3419-35.
JOMAA, A., STEWART, G., MARTIN-BENITO, J., ZIELKE, R., CAMPBELL, T. L., MADDOCK, J. R., BROWN, E. D. & ORTEGA, J. 2011a. Understanding ribosome assembly: the structure of in vivo assembled immature 30S subunits revealed by cryo-electron microscopy. *RNA,* 17**,** 697-709.
JOMAA, A., STEWART, G., MEARS, J. A., KIREEVA, I., BROWN, E. D. & ORTEGA, J. 2011b. Cryo-electron microscopy structure of the 30S subunit in complex with the YjeQ biogenesis factor. *RNA,* 17**,** 2026-38.
JULIAN, P., MILON, P., AGIRREZABALA, X., LASSO, G., GIL, D., RODNINA, M. V. & VALLE, M. 2011. The Cryo-EM structure of a complete 30S translation initiation complex from Escherichia coli. *PLoS Biol,* 9**,** e1001095.
KALEDHONKAR, S., FU, Z., CABAN, K., LI, W., CHEN, B., SUN, M., GONZALEZ, R. L., JR. & FRANK, J. 2019. Late steps in bacterial translation initiation visualized using time-resolved cryo-EM. *Nature,* 570**,** 400-404.
KALEDHONKAR, S., FU, Z., WHITE, H. & FRANK, J. 2018. Time-Resolved Cryo-electron Microscopy Using a Microfluidic Chip. *Methods Mol Biol,* 1764**,** 59-71.
KAMINISHI, T., WILSON, D. N., TAKEMOTO, C., HARMS, J. M., KAWAZOE, M., SCHLUENZEN, F., HANAWA-SUETSUGU, K., SHIROUZU, M., FUCINI, P. & YOKOYAMA, S. 2007. A snapshot of the 30S ribosomal subunit capturing mRNA via the Shine-Dalgarno interaction. *Structure,* 15**,** 289-97.
KARBSTEIN, K. 2007. Role of GTPases in ribosome assembly. *Biopolymers,* 87**,** 1-11.
KARBSTEIN, K. 2013. Quality control mechanisms during ribosome maturation. *Trends Cell Biol,* 23**,** 242-50.
KATOH, K., MISAWA, K., KUMA, K. & MIYATA, T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res,* 30**,** 3059-66.
KHUSAINOV, I., VICENS, Q., BOCHLER, A., GROSSE, F., MYASNIKOV, A., MENETRET, J. F., CHICHER, J., MARZI, S., ROMBY, P., YUSUPOVA, G., YUSUPOV, M. & HASHEM, Y. 2017. Structure of the 70S ribosome from human pathogen Staphylococcus aureus. *Nucleic Acids Res,* 45**,** 1026.
KIPPER, K., HETENYI, C., SILD, S., REMME, J. & LIIV, A. 2009. Ribosomal intersubunit bridge B2a is involved in factor-dependent translation initiation and translational processivity. *J Mol Biol,* 385**,** 405-22.

KITTS, P. A., CHURCH, D. M., THIBAUD-NISSEN, F., CHOI, J., HEM, V., SAPOJNIKOV, V., SMITH, R. G., TATUSOVA, T., XIANG, C., ZHERIKOV, A., DICUCCIO, M., MURPHY, T. D., PRUITT, K. D. & KIMCHI, A. 2016. Assembly: a resource for assembled genomes at NCBI. *Nucleic Acids Res,* 44**,** D73-80.

KLEBL, D. P., GRAVETT, M. S. C., KONTZIAMPASIS, D., WRIGHT, D. J., BON, R. S., MONTEIRO, D. C. F., TREBBIN, M., SOBOTT, F., WHITE, H. D., DARROW, M. C., THOMPSON, R. F. & MUENCH, S. P. 2020. Need for Speed: Examining Protein Behavior during CryoEM Grid Preparation at Different Timescales. *Structure,* 28**,** 1238-1248 e4.

KLEIN, B. A., TENORIO, E. L., LAZINSKI, D. W., CAMILLI, A., DUNCAN, M. J. & HU, L. T. 2012. Identification of essential genes of the periodontal pathogen Porphyromonas gingivalis. *BMC Genomics,* 13**,** 578.

KLEIN, D. J., MOORE, P. B. & STEITZ, T. A. 2004. The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA,* 10**,** 1366-79.

KLIMOVA, M., SENYUSHKINA, T., SAMATOVA, E., PENG, B. Z., PEARSON, M., PESKE, F. & RODNINA, M. V. 2019. EF-G-induced ribosome sliding along the noncoding mRNA. *Sci Adv,* 5**,** eaaw9049.

KONEVEGA, A. L., SOBOLEVA, N. G., MAKHNO, V. I., SEMENKOV, Y. P., WINTERMEYER, W., RODNINA, M. V. & KATUNIN, V. I. 2004. Purine bases at position 37 of tRNA stabilize codon-anticodon interaction in the ribosomal A site by stacking and Mg2+-dependent interactions. *RNA,* 10**,** 90-101.

KOROSTELEV, A., TRAKHANOV, S., ASAHARA, H., LAURBERG, M., LANCASTER, L. & NOLLER, H. F. 2007. Interactions and dynamics of the Shine Dalgarno helix in the 70S ribosome. *Proc Natl Acad Sci U S A,* 104**,** 16840-3.

KOROSTELEV, A., ZHU, J., ASAHARA, H. & NOLLER, H. F. 2010. Recognition of the amber UAG stop codon by release factor RF1. *EMBO J,* 29**,** 2577-85.

KOROSTELEV, A. A. 2022. The Structural Dynamics of Translation. *Annu Rev Biochem*.

KOTHE, U., WIEDEN, H. J., MOHR, D. & RODNINA, M. V. 2004. Interaction of helix D of elongation factor Tu with helices 4 and 5 of protein L7/12 on the ribosome. *J Mol Biol,* 336**,** 1011-21.

KOZAK, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell,* 44**,** 283-92.

KOZLOWSKI, L. P. 2016. IPC - Isoelectric Point Calculator. *Biol Direct,* 11**,** 55.

KUDVA, R., TIAN, P., PARDO-AVILA, F., CARRONI, M., BEST, R. B., BERNSTEIN, H. D. & VON HEIJNE, G. 2018. The shape of the bacterial ribosome exit tunnel affects cotranslational protein folding. *Elife,* 7.

KUHLBRANDT, W. 2014. Biochemistry. The resolution revolution. *Science,* 343**,** 1443-4.

KURITA, D. & HIMENO, H. 2022. Bacterial Ribosome Rescue Systems. *Microorganisms,* 10.

LALANNE, J. B., TAGGART, J. C., GUO, M. S., HERZEL, L., SCHIELER, A. & LI, G. W. 2018. Evolutionary Convergence of Pathway-Specific Enzyme Expression Stoichiometry. *Cell,* 173**,** 749-761 e38.

LANCASTER, L., KIEL, M. C., KAJI, A. & NOLLER, H. F. 2002. Orientation of ribosome recycling factor in the ribosome from directed hydroxyl radical probing. *Cell,* 111**,** 129-40.

LARKIN, M. A., BLACKSHIELDS, G., BROWN, N. P., CHENNA, R., MCGETTIGAN, P. A., MCWILLIAM, H., VALENTIN, F., WALLACE, I. M., WILM, A., LOPEZ, R.,

THOMPSON, J. D., GIBSON, T. J. & HIGGINS, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics,* 23**,** 2947-8.

LEONG, V., KENT, M., JOMAA, A. & ORTEGA, J. 2013. Escherichia coli rimM and yjeQ null strains accumulate immature 30S subunits of similar structure and protein complement. *RNA,* 19**,** 789-802.

LEY, R. E., HAMADY, M., LOZUPONE, C., TURNBAUGH, P. J., RAMEY, R. R., BIRCHER, J. S., SCHLEGEL, M. L., TUCKER, T. A., SCHRENZEL, M. D., KNIGHT, R. & GORDON, J. I. 2008. Evolution of mammals and their gut microbes. *Science,* 320**,** 1647-51.

LI, G. W., BURKHARDT, D., GROSS, C. & WEISSMAN, J. S. 2014. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell,* 157**,** 624-35.

LI, N., CHEN, Y., GUO, Q., ZHANG, Y., YUAN, Y., MA, C., DENG, H., LEI, J. & GAO, N. 2013. Cryo-EM structures of the late-stage assembly intermediates of the bacterial 50S ribosomal subunit. *Nucleic Acids Res,* 41**,** 7073-83.

LI, W., LIU, Z., KORIPELLA, R. K., LANGLOIS, R., SANYAL, S. & FRANK, J. 2015. Activation of GTP hydrolysis in mRNA-tRNA translocation by elongation factor G. *Sci Adv,* 1.

LI, Z., GE, X., ZHANG, Y., ZHENG, L., SANYAL, S. & GAO, N. 2018. Cryo-EM structure of Mycobacterium smegmatis ribosome reveals two unidentified ribosomal proteins close to the functional centers. *Protein Cell,* 9**,** 384-388.

LI, Z., PANDIT, S. & DEUTSCHER, M. P. 1999. Maturation of 23S ribosomal RNA requires the exoribonuclease RNase T. *RNA,* 5**,** 139-46.

LIEBSCHNER, D., AFONINE, P. V., BAKER, M. L., BUNKOCZI, G., CHEN, V. B., CROLL, T. I., HINTZE, B., HUNG, L. W., JAIN, S., MCCOY, A. J., MORIARTY, N. W., OEFFNER, R. D., POON, B. K., PRISANT, M. G., READ, R. J., RICHARDSON, J. S., RICHARDSON, D. C., SAMMITO, M. D., SOBOLEV, O. V., STOCKWELL, D. H., TERWILLIGER, T. C., URZHUMTSEV, A. G., VIDEAU, L. L., WILLIAMS, C. J. & ADAMS, P. D. 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol,* 75**,** 861-877.

LIM, K., FURUTA, Y. & KOBAYASHI, I. 2012. Large variations in bacterial ribosomal RNA genes. *Mol Biol Evol,* 29**,** 2937-48.

LIN, J., ZHOU, D., STEITZ, T. A., POLIKANOV, Y. S. & GAGNON, M. G. 2018. Ribosome-Targeting Antibiotics: Modes of Action, Mechanisms of Resistance, and Implications for Drug Design. *Annu Rev Biochem,* 87**,** 451-478.

LOPEZ-ALONSO, J. P., FABBRETTI, A., KAMINISHI, T., ITURRIOZ, I., BRANDI, L., GIL-CARTON, D., GUALERZI, C. O., FUCINI, P. & CONNELL, S. R. 2017a. Structure of a 30S pre-initiation complex stalled by GE81112 reveals structural parallels in bacterial and eukaryotic protein synthesis initiation pathways. *Nucleic Acids Res,* 45**,** 2179-2187.

LOPEZ-ALONSO, J. P., KAMINISHI, T., KIKUCHI, T., HIRATA, Y., ITURRIOZ, I., DHIMOLE, N., SCHEDLBAUER, A., HASE, Y., GOTO, S., KURITA, D., MUTO, A., ZHOU, S., NAOE, C., MILLS, D. J., GIL-CARTON, D., TAKEMOTO, C., HIMENO, H., FUCINI, P. & CONNELL, S. R. 2017b. RsgA couples the maturation state of the 30S ribosomal decoding center to activation of its GTPase pocket. *Nucleic Acids Res,* 45**,** 6945-6959.

LOVELAND, A. B., DEMO, G., GRIGORIEFF, N. & KOROSTELEV, A. A. 2017. Ensemble cryo-EM elucidates the mechanism of translation fidelity. *Nature,* 546**,** 113-117.

LOVELAND, A. B., DEMO, G. & KOROSTELEV, A. A. 2020. Cryo-EM of elongating ribosome with EF-Tu*GTP elucidates tRNA proofreading. *Nature,* 584**,** 640-645.

LU, X. J., BUSSEMAKER, H. J. & OLSON, W. K. 2015. DSSR: an integrated software tool for dissecting the spatial structure of RNA. *Nucleic Acids Res,* 43**,** e142.

LU, X. J. & OLSON, W. K. 2008. 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat Protoc,* 3**,** 1213-27.

MAGUIRE, B. A. 2009. Inhibition of bacterial ribosome assembly: a suitable drug target? *Microbiol Mol Biol Rev,* 73**,** 22-35.

MAITI, P., LAVDOVSKAIA, E., BARRIENTOS, A. & RICHTER-DENNERLEIN, R. 2021. Role of GTPases in Driving Mitoribosome Assembly. *Trends Cell Biol,* 31**,** 284-297.

MAKSIMOVA, E. M., KOREPANOV, A. P., KRAVCHENKO, O. V., BAYMUKHAMETOV, T. N., MYASNIKOV, A. G., VASSILENKO, K. S., AFONINA, Z. A. & STOLBOUSHKINA, E. A. 2021. RbfA Is Involved in Two Important Stages of 30S Subunit Assembly: Formation of the Central Pseudoknot and Docking of Helix 44 to the Decoding Center. *Int J Mol Sci,* 22.

MARTENS, E. C., CHIANG, H. C. & GORDON, J. I. 2008. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host Microbe,* 4**,** 447-57.

MARTENS, E. C., KOROPATKIN, N. M., SMITH, T. J. & GORDON, J. I. 2009a. Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. *J Biol Chem,* 284**,** 24673-7.

MARTENS, E. C., ROTH, R., HEUSER, J. E. & GORDON, J. I. 2009b. Coordinate regulation of glycan degradation and polysaccharide capsule biosynthesis by a prominent human gut symbiont. *J Biol Chem,* 284**,** 18445-57.

MCBRIDE, M. J. & KEMPF, M. J. 1996. Development of techniques for the genetic manipulation of the gliding bacterium Cytophaga johnsonae. *J Bacteriol,* 178**,** 583-90.

MCBRIDE, M. J., XIE, G., MARTENS, E. C., LAPIDUS, A., HENRISSAT, B., RHODES, R. G., GOLTSMAN, E., WANG, W., XU, J., HUNNICUTT, D. W., STAROSCIK, A. M., HOOVER, T. R., CHENG, Y. Q. & STEIN, J. L. 2009. Novel features of the polysaccharide-digesting gliding bacterium Flavobacterium johnsoniae as revealed by genome sequence analysis. *Appl Environ Microbiol,* 75**,** 6864-75.

MCCALL, K. A., HUANG, C. & FIERKE, C. A. 2000. Function and mechanism of zinc metalloenzymes. *J Nutr,* 130**,** 1437S-46S.

MCCARTHY, B. J. 1962. The effects of magnesium starvation on the ribosome content of Escherichia coli. *Biochimica et Biophysica Acta,* 55**,** 880-889.

MCGINNIS, J. L., LIU, Q., LAVENDER, C. A., DEVARAJ, A., MCCLORY, S. P., FREDRICK, K. & WEEKS, K. M. 2015. In-cell SHAPE reveals that free 30S ribosome subunits are in the inactive state. *Proc Natl Acad Sci U S A,* 112**,** 2425-30.

MCGINNIS, J. L. & WEEKS, K. M. 2014. Ribosome RNA assembly intermediates visualized in living cells. *Biochemistry,* 53**,** 3237-47.

METSELAAR, K. I., DEN BESTEN, H. M., BOEKHORST, J., VAN HIJUM, S. A., ZWIETERING, M. H. & ABEE, T. 2015. Diversity of acid stress resistant variants of

Listeria monocytogenes and the potential role of ribosomal protein S21 encoded by rpsU. *Front Microbiol,* 6**,** 422.
MILLER, J. H. 1972. *Experiments in molecular genetics,* Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory.
MILON, P. & RODNINA, M. V. 2012. Kinetic control of translation initiation in bacteria. *Crit Rev Biochem Mol Biol,* 47**,** 334-48.
MISKIN, R., ZAMIR, A. & ELSON, D. 1970. Inactivation and reactivation of ribosomal subunits: the peptidyl transferase activity of the 50 s subunit of Escherihia coli. *J Mol Biol,* 54**,** 355-78.
MISRA, T. K. & APIRION, D. 1979. RNase E, an RNA processing enzyme from Escherichia coli. *J Biol Chem,* 254**,** 11154-9.
MIZUNO, C. M., GUYOMAR, C., ROUX, S., LAVIGNE, R., RODRIGUEZ-VALERA, F., SULLIVAN, M. B., GILLET, R., FORTERRE, P. & KRUPOVIC, M. 2019. Numerous cultivated and uncultivated viruses encode ribosomal proteins. *Nat Commun,* 10**,** 752.
MIZUSHIMA, S. & NOMURA, M. 1970. Assembly mapping of 30S ribosomal proteins from E. coli. *Nature,* 226**,** 1214.
MOAZED, D., VAN STOLK, B. J., DOUTHWAITE, S. & NOLLER, H. F. 1986. Interconversion of active and inactive 30 S ribosomal subunits is accompanied by a conformational change in the decoding region of 16 S rRNA. *J Mol Biol,* 191**,** 483-93.
MOHAN, S., DONOHUE, J. P. & NOLLER, H. F. 2014. Molecular mechanics of 30S subunit head rotation. *Proc Natl Acad Sci U S A,* 111**,** 13325-30.
MOHLER, K. & IBBA, M. 2017. Translational fidelity and mistranslation in the cellular response to stress. *Nat Microbiol,* 2**,** 17117.
MOORE, P. B. & STEITZ, T. A. 2011. The roles of RNA in the synthesis of protein. *Cold Spring Harb Perspect Biol,* 3**,** a003780.
MORSE, J. C., GIRODAT, D., BURNETT, B. J., HOLM, M., ALTMAN, R. B., SANBONMATSU, K. Y., WIEDEN, H. J. & BLANCHARD, S. C. 2020. Elongation factor-Tu can repetitively engage aminoacyl-tRNA within the ribosome during the proofreading stage of tRNA selection. *Proc Natl Acad Sci U S A,* 117**,** 3610-3620.
MULDER, A. M., YOSHIOKA, C., BECK, A. H., BUNNER, A. E., MILLIGAN, R. A., POTTER, C. S., CARRAGHER, B. & WILLIAMSON, J. R. 2010. Visualizing ribosome biogenesis: parallel assembly pathways for the 30S subunit. *Science,* 330**,** 673-7.
MURPHY, F. V. T. & RAMAKRISHNAN, V. 2004. Structure of a purine-purine wobble base pair in the decoding center of the ribosome. *Nat Struct Mol Biol,* 11**,** 1251-2.
MYASNIKOV, A. G., SIMONETTI, A., MARZI, S. & KLAHOLZ, B. P. 2009. Structure-function insights into prokaryotic and eukaryotic translation initiation. *Curr Opin Struct Biol,* 19**,** 300-9.
NAKAGAWA, S., NIIMURA, Y. & GOJOBORI, T. 2017. Comparative genomic analysis of translation initiation mechanisms for genes lacking the Shine-Dalgarno sequence in prokaryotes. *Nucleic Acids Res,* 45**,** 3922-3931.
NAKAGAWA, S., NIIMURA, Y., GOJOBORI, T., TANAKA, H. & MIURA, K. 2008. Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. *Nucleic Acids Res,* 36**,** 861-71.
NAKAGAWA, S., NIIMURA, Y., MIURA, K. & GOJOBORI, T. 2010. Dynamic evolution of translation initiation mechanisms in prokaryotes. *Proc Natl Acad Sci U S A,* 107**,** 6382-7.

NAKANE, T., KIMANIUS, D., LINDAHL, E. & SCHERES, S. H. 2018. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *Elife,* 7.

NAKANE, T., KOTECHA, A., SENTE, A., MCMULLAN, G., MASIULIS, S., BROWN, P., GRIGORAS, I. T., MALINAUSKAITE, L., MALINAUSKAS, T., MIEHLING, J., UCHANSKI, T., YU, L., KARIA, D., PECHNIKOVA, E. V., DE JONG, E., KEIZER, J., BISCHOFF, M., MCCORMACK, J., TIEMEIJER, P., HARDWICK, S. W., CHIRGADZE, D. Y., MURSHUDOV, G., ARICESCU, A. R. & SCHERES, S. H. W. 2020. Single-particle cryo-EM at atomic resolution. *Nature,* 587**,** 152-156.

NI, X., DAVIS, J. H., JAIN, N., RAZI, A., BENLEKBIR, S., MCARTHUR, A. G., RUBINSTEIN, J. L., BRITTON, R. A., WILLIAMSON, J. R. & ORTEGA, J. 2016. YphC and YsxC GTPases assist the maturation of the central protuberance, GTPase associated region and functional core of the 50S ribosomal subunit. *Nucleic Acids Res,* 44**,** 8442-55.

NICHOLS, C. E., JOHNSON, C., LAMB, H. K., LOCKYER, M., CHARLES, I. G., HAWKINS, A. R. & STAMMERS, D. K. 2007. Structure of the ribosomal interacting GTPase YjeQ from the enterobacterial species Salmonella typhimurium. *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 63**,** 922-8.

NIERHAUS, K. H. 2014. Mg2+, K+, and the ribosome. *J Bacteriol,* 196**,** 3817-9.

NIKOLAY, R., HILAL, T., QIN, B., MIELKE, T., BURGER, J., LOERKE, J., TEXTORIS-TAUBE, K., NIERHAUS, K. H. & SPAHN, C. M. T. 2018. Structural Visualization of the Formation and Activation of the 50S Ribosomal Subunit during In Vitro Reconstitution. *Mol Cell,* 70**,** 881-893 e3.

NISSEN, P., HANSEN, J., BAN, N., MOORE, P. B. & STEITZ, T. A. 2000. The structural basis of ribosome activity in peptide bond synthesis. *Science,* 289**,** 920-30.

NOBLE, A. J., WEI, H., DANDEY, V. P., ZHANG, Z., TAN, Y. Z., POTTER, C. S. & CARRAGHER, B. 2018. Reducing effects of particle adsorption to the air-water interface in cryo-EM. *Nat Methods,* 15**,** 793-795.

NOESKE, J. & CATE, J. H. 2012. Structural basis for protein synthesis: snapshots of the ribosome in motion. *Curr Opin Struct Biol,* 22**,** 743-9.

NOGALES, E. & SCHERES, S. H. 2015. Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. *Mol Cell,* 58**,** 677-89.

NOLLER, H. F. & WOESE, C. R. 1981. Secondary structure of 16S ribosomal RNA. *Science,* 212**,** 403-11.

NOMURA, M., GOURSE, R. & BAUGHMAN, G. 1984. Regulation of the synthesis of ribosomes and ribosomal components. *Annu Rev Biochem,* 53**,** 75-117.

NOMURA, M. & TRAUB, P. 1968. Structure and function of Escherichia coli ribosomes. 3. Stoichiometry and rate of the reconstitution of ribosomes from subribosomal particles and split proteins. *J Mol Biol,* 34**,** 609-19.

NOMURA, M., TRAUB, P., GUTHRIE, C. & NASHIMOTO, H. 1969. The assembly of ribosomes. *J Cell Physiol,* 74**,** Suppl 1:241+.

O'CONNOR, M., THOMAS, C. L., ZIMMERMANN, R. A. & DAHLBERG, A. E. 1997. Decoding fidelity at the ribosomal A and P sites: influence of mutations in three different regions of the decoding domain in 16S rRNA. *Nucleic Acids Res,* 25**,** 1185-93.

OGLE, J. M., BRODERSEN, D. E., CLEMONS, W. M., JR., TARRY, M. J., CARTER, A. P. & RAMAKRISHNAN, V. 2001. Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science,* 292**,** 897-902.

OGLE, J. M. & RAMAKRISHNAN, V. 2005. Structural insights into translational fidelity. *Annu Rev Biochem,* 74**,** 129-77.
OKONECHNIKOV, K., GOLOSOVA, O., FURSOV, M. & TEAM, U. 2012. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics,* 28**,** 1166-7.
OPARINA, N. J., KALININA, O. V., GELFAND, M. S. & KISSELEV, L. L. 2005. Common and specific amino acid residues in the prokaryotic polypeptide release factors RF1 and RF2: possible functional implications. *Nucleic Acids Res,* 33**,** 5226-34.
OSADA, Y., SAITO, R. & TOMITA, M. 1999. Analysis of base-pairing potentials between 16S rRNA and 5' UTR for translation initiation in various prokaryotes. *Bioinformatics,* 15**,** 578-81.
OU, H. Y., GUO, F. B. & ZHANG, C. T. 2004. GS-Finder: a program to find bacterial gene start sites with a self-training method. *Int J Biochem Cell Biol,* 36**,** 535-44.
PETROV, A. S., BERNIER, C. R., HSIAO, C., OKAFOR, C. D., TANNENBAUM, E., STERN, J., GAUCHER, E., SCHNEIDER, D., HUD, N. V., HARVEY, S. C. & WILLIAMS, L. D. 2012. RNA-magnesium-protein interactions in large ribosomal subunit. *J Phys Chem B,* 116**,** 8113-20.
PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem,* 25**,** 1605-12.
PLASSART, L., SHAYAN, R., MONTELLESE, C., RINALDI, D., LARBURU, N., PICHEREAUX, C., FROMENT, C., LEBARON, S., O'DONOHUE, M. F., KUTAY, U., MARCOUX, J., GLEIZES, P. E. & PLISSON-CHASTANG, C. 2021. The final step of 40S ribosomal subunit maturation is controlled by a dual key lock. *Elife,* 10.
POLIKANOV, Y. S., BLAHA, G. M. & STEITZ, T. A. 2012. How hibernation factors RMF, HPF, and YfiA turn off protein synthesis. *Science,* 336**,** 915-8.
POLIKANOV, Y. S., STAROSTA, A. L., JUETTE, M. F., ALTMAN, R. B., TERRY, D. S., LU, W., BURNETT, B. J., DINOS, G., REYNOLDS, K. A., BLANCHARD, S. C., STEITZ, T. A. & WILSON, D. N. 2015. Distinct tRNA Accommodation Intermediates Observed on the Ribosome with the Antibiotics Hygromycin A and A201A. *Mol Cell,* 58**,** 832-44.
POLIKANOV, Y. S., STEITZ, T. A. & INNIS, C. A. 2014. A proton wire to couple aminoacyl-tRNA accommodation and peptide-bond formation on the ribosome. *Nat Struct Mol Biol,* 21**,** 787-93.
PONTES, M. H., SEVOSTYANOVA, A. & GROISMAN, E. A. 2015. When Too Much ATP Is Bad for Protein Synthesis. *J Mol Biol,* 427**,** 2586-2594.
POOT, R. A., PLEIJ, C. W. & VAN DUIN, J. 1996. The central pseudoknot in 16S ribosomal RNA is needed for ribosome stability but is not essential for 30S initiation complex formation. *Nucleic Acids Res,* 24**,** 3670-6.
PUNJANI, A., RUBINSTEIN, J. L., FLEET, D. J. & BRUBAKER, M. A. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods,* 14**,** 290-296.
QIN, D., ABDI, N. M. & FREDRICK, K. 2007. Characterization of 16S rRNA mutations that decrease the fidelity of translation initiation. *RNA,* 13**,** 2348-55.
QIN, D. & FREDRICK, K. 2009. Control of translation initiation involves a factor-induced rearrangement of helix 44 of 16S ribosomal RNA. *Mol Microbiol,* 71**,** 1239-49.
QIN, D., LIU, Q., DEVARAJ, A. & FREDRICK, K. 2012. Role of helix 44 of 16S rRNA in the fidelity of translation initiation. *RNA,* 18**,** 485-95.

RAZI, A., BRITTON, R. A. & ORTEGA, J. 2017a. The impact of recent improvements in cryo-electron microscopy technology on the understanding of bacterial ribosome assembly. *Nucleic Acids Res,* 45**,** 1027-1040.

RAZI, A., DAVIS, J. H., HAO, Y., JAHAGIRDAR, D., THURLOW, B., BASU, K., JAIN, N., GOMEZ-BLANCO, J., BRITTON, R. A., VARGAS, J., GUARNE, A., WOODSON, S. A., WILLIAMSON, J. R. & ORTEGA, J. 2019. Role of Era in assembly and homeostasis of the ribosomal small subunit. *Nucleic Acids Res,* 47**,** 8301-8317.

RAZI, A., GUARNE, A. & ORTEGA, J. 2017b. The cryo-EM structure of YjeQ bound to the 30S subunit suggests a fidelity checkpoint function for this protein in ribosome assembly. *Proc Natl Acad Sci U S A,* 114**,** E3396-E3403.

REDKO, Y., BECHHOFER, D. H. & CONDON, C. 2008. Mini-III, an unusual member of the RNase III family of enzymes, catalyses 23S ribosomal RNA maturation in B. subtilis. *Mol Microbiol,* 68**,** 1096-106.

RINGQUIST, S., SHINEDLING, S., BARRICK, D., GREEN, L., BINKLEY, J., STORMO, G. D. & GOLD, L. 1992. Translation initiation in Escherichia coli: sequences within the ribosome-binding site. *Mol Microbiol,* 6**,** 1219-29.

RINNINELLA, E., RAOUL, P., CINTONI, M., FRANCESCHI, F., MIGGIANO, G. A. D., GASBARRINI, A. & MELE, M. C. 2019. What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. *Microorganisms,* 7.

RODGERS, M. L., HAO, Y. & WOODSON, S. A. 2019. A newborn RNA switches its fate. *Nat Chem Biol,* 15**,** 1031-1032.

RODNINA, M. V. & WINTERMEYER, W. 2001. Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms. *Annu Rev Biochem,* 70**,** 415-35.

ROY, B., LIU, Q., SHOJI, S. & FREDRICK, K. 2018. IF2 and unique features of initiator tRNA(fMet) help establish the translational reading frame. *RNA Biol,* 15**,** 604-613.

ROZOV, A., KHUSAINOV, I., EL OMARI, K., DUMAN, R., MYKHAYLYK, V., YUSUPOV, M., WESTHOF, E., WAGNER, A. & YUSUPOVA, G. 2019. Importance of potassium ions for ribosome structure and function revealed by long-wavelength X-ray diffraction. *Nat Commun,* 10**,** 2519.

RUBINSTEIN, J. L., GUO, H., RIPSTEIN, Z. A., HAYDAROGLU, A., AU, A., YIP, C. M., DI TRANI, J. M., BENLEKBIR, S. & KWOK, T. 2019. Shake-it-off: a simple ultrasonic cryo-EM specimen-preparation device. *Acta Crystallogr D Struct Biol,* 75**,** 1063-1070.

RUSSO, C. J. & PASSMORE, L. A. 2014. Controlling protein adsorption on graphene for cryo-EM using low-energy hydrogen plasmas. *Nat Methods,* 11**,** 649-52.

SACERDOT, C., CHIARUTTINI, C., ENGST, K., GRAFFE, M., MILET, M., MATHY, N., DONDON, J. & SPRINGER, M. 1996. The role of the AUU initiation codon in the negative feedback regulation of the gene for translation initiation factor IF3 in Escherichia coli. *Mol Microbiol,* 21**,** 331-46.

SADER, K., STOPPS, M., CALDER, L. J. & ROSENTHAL, P. B. 2013. Cryomicroscopy of radiation sensitive specimens on unmodified graphene sheets: reduction of electron-optical effects of charging. *J Struct Biol,* 183**,** 531-536.

SALAH, P., BISAGLIA, M., ALIPRANDI, P., UZAN, M., SIZUN, C. & BONTEMS, F. 2009. Probing the relationship between Gram-negative and Gram-positive S1 proteins by sequence analysis. *Nucleic Acids Res,* 37**,** 5578-88.

SALIS, H. M., MIRSKY, E. A. & VOIGT, C. A. 2009. Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol,* 27**,** 946-50.

SANCHEZ-GARCIA, R., SEGURA, J., MALUENDA, D., CARAZO, J. M. & SORZANO, C. O. S. 2018. Deep Consensus, a deep learning-based approach for particle pruning in cryo-electron microscopy. *IUCrJ,* 5**,** 854-865.

SASHITAL, D. G., GREEMAN, C. A., LYUMKIS, D., POTTER, C. S., CARRAGHER, B. & WILLIAMSON, J. R. 2014. A combined quantitative mass spectrometry and electron microscopy analysis of ribosomal 30S subunit assembly in E. coli. *Elife,* 3.

SCHEDLBAUER, A., ITURRIOZ, I., OCHOA-LIZARRALDE, B., CAPUNI, R., HAN, X., DE ASTIGARRAGA, E., DIERCKS, T., FUCINI, P. & CONNELL, S. R. 2020. Backbone and sidechain NMR assignments for the ribosome maturation factor RbfA from Escherichia coli. *Biomol NMR Assign,* 14**,** 317-321.

SCHEDLBAUER, A., ITURRIOZ, I., OCHOA-LIZARRALDE, B., DIERCKS, T., LOPEZ-ALONSO, J. P., LAVIN, J. L., KAMINISHI, T., CAPUNI, R., DHIMOLE, N., DE ASTIGARRAGA, E., GIL-CARTON, D., FUCINI, P. & CONNELL, S. R. 2021. A conserved rRNA switch is central to decoding site maturation on the small ribosomal subunit. *Sci Adv,* 7.

SCHERES, S. H. & CHEN, S. 2012. Prevention of overfitting in cryo-EM structure determination. *Nat Methods,* 9**,** 853-4.

SCHMEING, T. M., HUANG, K. S., KITCHEN, D. E., STROBEL, S. A. & STEITZ, T. A. 2005. Structural insights into the roles of water and the 2' hydroxyl of the P site tRNA in the peptidyl transferase reaction. *Mol Cell,* 20**,** 437-48.

SCHMEING, T. M., VOORHEES, R. M., KELLEY, A. C., GAO, Y. G., MURPHY, F. V. T., WEIR, J. R. & RAMAKRISHNAN, V. 2009. The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science,* 326**,** 688-94.

SCHORB, M., HABERBOSCH, I., HAGEN, W. J. H., SCHWAB, Y. & MASTRONARDE, D. N. 2019. Software tools for automated transmission electron microscopy. *Nat Methods,* 16**,** 471-477.

SCHRADER, J. M., ZHOU, B., LI, G. W., LASKER, K., CHILDERS, W. S., WILLIAMS, B., LONG, T., CROSSON, S., MCADAMS, H. H., WEISSMAN, J. S. & SHAPIRO, L. 2014. The coding and noncoding architecture of the Caulobacter crescentus genome. *PLoS Genet,* 10**,** e1004463.

SCHUETTE, J. C., MURPHY, F. V. T., KELLEY, A. C., WEIR, J. R., GIESEBRECHT, J., CONNELL, S. R., LOERKE, J., MIELKE, T., ZHANG, W., PENCZEK, P. A., RAMAKRISHNAN, V. & SPAHN, C. M. 2009. GTPase activation of elongation factor EF-Tu by the ribosome during decoding. *EMBO J,* 28**,** 755-65.

SCHURECK, M. A., MAEHIGASHI, T., MILES, S. J., MARQUEZ, J. & DUNHAM, C. M. 2016. mRNA bound to the 30S subunit is a HigB toxin substrate. *RNA,* 22**,** 1261-70.

SCHURR, T., NADIR, E. & MARGALIT, H. 1993. Identification and characterization of E.coli ribosomal binding sites by free energy computation. *Nucleic Acids Res,* 21**,** 4019-23.

SCHUWIRTH, B. S., BOROVINSKAYA, M. A., HAU, C. W., ZHANG, W., VILA-SANJURJO, A., HOLTON, J. M. & CATE, J. H. 2005. Structures of the bacterial ribosome at 3.5 A resolution. *Science,* 310**,** 827-34.

SEFFOUH, A., JAIN, N., JAHAGIRDAR, D., BASU, K., RAZI, A., NI, X., GUARNE, A., BRITTON, R. A. & ORTEGA, J. 2019. Structural consequences of the interaction of RbgA with a 50S ribosomal subunit assembly intermediate. *Nucleic Acids Res,* 47**,** 10414-10425.

SELMER, M., DUNHAM, C. M., MURPHY, F. V. T., WEIXLBAUMER, A., PETRY, S., KELLEY, A. C., WEIR, J. R. & RAMAKRISHNAN, V. 2006. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science,* 313**,** 1935-42.

SENGUPTA, J., AGRAWAL, R. K. & 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**,** 11991-6.

SEZONOV, G., JOSELEAU-PETIT, D. & D'ARI, R. 2007. Escherichia coli physiology in Luria-Bertani broth. *J Bacteriol,* 189**,** 8746-9.

SHAJANI, Z., SYKES, M. T. & WILLIAMSON, J. R. 2011. Assembly of bacterial ribosomes. *Annu Rev Biochem,* 80**,** 501-26.

SHARMA, H. & ANAND, B. 2019. Ribosome assembly defects subvert initiation Factor3 mediated scrutiny of bona fide start signal. *Nucleic Acids Res,* 47**,** 11368-11386.

SHARMA, I. M. & WOODSON, S. A. 2020. RbfA and IF3 couple ribosome biogenesis and translation initiation to increase stress tolerance. *Nucleic Acids Res,* 48**,** 359-372.

SHIMIZU, Y. 2012. ArfA recruits RF2 into stalled ribosomes. *J Mol Biol,* 423**,** 624-31.

SHIN, D. H., LOU, Y., JANCARIK, J., YOKOTA, H., KIM, R. & KIM, S. H. 2004. Crystal structure of YjeQ from Thermotoga maritima contains a circularly permuted GTPase domain. *Proc Natl Acad Sci U S A,* 101**,** 13198-203.

SHINE, J. & 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**,** 1342-6.

SHOJI, S., ABDI, N. M., BUNDSCHUH, R. & FREDRICK, K. 2009. Contribution of ribosomal residues to P-site tRNA binding. *Nucleic Acids Res,* 37**,** 4033-42.

SHOJI, S., WALKER, S. E. & FREDRICK, K. 2006. Reverse translocation of tRNA in the ribosome. *Mol Cell,* 24**,** 931-42.

SHULTZABERGER, R. K., BUCHEIMER, R. E., RUDD, K. E. & SCHNEIDER, T. D. 2001. Anatomy of Escherichia coli ribosome binding sites. *J Mol Biol,* 313**,** 215-28.

SISSI, C. & PALUMBO, M. 2009. Effects of magnesium and related divalent metal ions in topoisomerase structure and function. *Nucleic Acids Res,* 37**,** 702-11.

SKORSKI, P., LEROY, P., FAYET, O., DREYFUS, M. & HERMANN-LE DENMAT, S. 2006. The highly efficient translation initiation region from the Escherichia coli rpsA gene lacks a shine-dalgarno element. *J Bacteriol,* 188**,** 6277-85.

SOHMEN, D., CHIBA, S., SHIMOKAWA-CHIBA, N., INNIS, C. A., BERNINGHAUSEN, O., BECKMANN, R., ITO, K. & WILSON, D. N. 2015. Structure of the Bacillus subtilis 70S ribosome reveals the basis for species-specific stalling. *Nat Commun,* 6**,** 6941.

SOHMEN, D., HARMS, J. M., SCHLUNZEN, F. & WILSON, D. N. 2009. Enhanced SnapShot: Antibiotic inhibition of protein synthesis II. *Cell,* 139**,** 212-212 e1.

SPAHN, C. M., GOMEZ-LORENZO, M. G., GRASSUCCI, R. A., JORGENSEN, R., ANDERSEN, G. R., BECKMANN, R., PENCZEK, P. A., BALLESTA, J. P. & FRANK, J. 2004. Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA translocation. *EMBO J,* 23**,** 1008-19.

STARMER, J., STOMP, A., VOUK, M. & BITZER, D. 2006. Predicting Shine-Dalgarno sequence locations exposes genome annotation errors. *PLoS Comput Biol,* 2**,** e57.

STEITZ, J. A. & JAKES, K. 1975. How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in Escherichia coli. *Proc Natl Acad Sci U S A,* 72**,** 4734-8.

STEPHAN, N. C., RIES, A. B., BOEHRINGER, D. & BAN, N. 2021. Structural basis of successive adenosine modifications by the conserved ribosomal methyltransferase KsgA. *Nucleic Acids Res,* 49**,** 6389-6398.
STOKES, J. M. & BROWN, E. D. 2015. Chemical modulators of ribosome biogenesis as biological probes. *Nat Chem Biol,* 11**,** 924-32.
SUBRAMANIAN, A. R. & VAN DUIN, J. 1977. Exchange of individual ribosomal proteins between ribosomes as studied by heavy isotope-transfer experiments. *Mol Gen Genet,* 158**,** 1-9.
SULTHANA, S. & DEUTSCHER, M. P. 2013. Multiple exoribonucleases catalyze maturation of the 3' terminus of 16S ribosomal RNA (rRNA). *J Biol Chem,* 288**,** 12574-9.
SUSSMAN, J. K., SIMONS, E. L. & SIMONS, R. W. 1996. Escherichia coli translation initiation factor 3 discriminates the initiation codon in vivo. *Mol Microbiol,* 21**,** 347-60.
SUZEK, B. E., ERMOLAEVA, M. D., SCHREIBER, M. & SALZBERG, S. L. 2001. A probabilistic method for identifying start codons in bacterial genomes. *Bioinformatics,* 17**,** 1123-30.
SWIATKOWSKA, A., WLOTZKA, W., TUCK, A., BARRASS, J. D., BEGGS, J. D. & TOLLERVEY, D. 2012. Kinetic analysis of pre-ribosome structure in vivo. *RNA,* 18**,** 2187-200.
SYKES, M. T. & WILLIAMSON, J. R. 2009. A complex assembly landscape for the 30S ribosomal subunit. *Annu Rev Biophys,* 38**,** 197-215.
TAKADA, H., MORITA, M., SHIWA, Y., SUGIMOTO, R., SUZUKI, S., KAWAMURA, F. & YOSHIKAWA, H. 2014. Cell motility and biofilm formation in Bacillus subtilis are affected by the ribosomal proteins, S11 and S21. *Biosci Biotechnol Biochem,* 78**,** 898-907.
TAKYAR, S., HICKERSON, R. P. & NOLLER, H. F. 2005. mRNA helicase activity of the ribosome. *Cell,* 120**,** 49-58.
TALKINGTON, M. W., SIUZDAK, G. & WILLIAMSON, J. R. 2005. An assembly landscape for the 30S ribosomal subunit. *Nature,* 438**,** 628-32.
THOMPSON, R. C. & STONE, P. J. 1977. Proofreading of the codon-anticodon interaction on ribosomes. *Proc Natl Acad Sci U S A,* 74**,** 198-202.
THURLOW, B., DAVIS, J. H., LEONG, V., MORAES, T. F., WILLIAMSON, J. R. & ORTEGA, J. 2016. Binding properties of YjeQ (RsgA), RbfA, RimM and Era to assembly intermediates of the 30S subunit. *Nucleic Acids Res,* 44**,** 9918-9932.
TRAUB, P. & NOMURA, M. 1969. Studies on the assembly of ribosomes in vitro. *Cold Spring Harb Symp Quant Biol,* 34**,** 63-7.
UNIPROT, C. 2019. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res,* 47**,** D506-D515.
VELLANOWETH, R. L. & RABINOWITZ, J. C. 1992. The influence of ribosome-binding-site elements on translational efficiency in Bacillus subtilis and Escherichia coli in vivo. *Mol Microbiol,* 6**,** 1105-14.
VERCRUYSSE, M., KOHRER, C., SHEN, Y., PROULX, S., GHOSAL, A., DAVIES, B. W., RAJBHANDARY, U. L. & WALKER, G. C. 2016. Identification of YbeY-Protein Interactions Involved in 16S rRNA Maturation and Stress Regulation in Escherichia coli. *mBio,* 7.
VINGADASSALOM, D., KOLB, A., MAYER, C., RYBKINE, T., COLLATZ, E. & PODGLAJEN, I. 2005. An unusual primary sigma factor in the Bacteroidetes phylum. *Mol Microbiol,* 56**,** 888-902.

WACKER, W. E. 1969. The biochemistry of magnesium. *Ann N Y Acad Sci,* 162**,** 717-26.
WALKER, S. E. & FREDRICK, K. 2008. Preparation and evaluation of acylated tRNAs. *Methods,* 44**,** 81-6.
WEGMANN, U., HORN, N. & CARDING, S. R. 2013. Defining the bacteroides ribosomal binding site. *Appl Environ Microbiol,* 79**,** 1980-9.
WIEDEN, H. J., WINTERMEYER, W. & RODNINA, M. V. 2001. A common structural motif in elongation factor Ts and ribosomal protein L7/12 may be involved in the interaction with elongation factor Tu. *J Mol Evol,* 52**,** 129-36.
WILLIAMS, C. J., HEADD, J. J., MORIARTY, N. W., PRISANT, M. G., VIDEAU, L. L., DEIS, L. N., VERMA, V., KEEDY, D. A., HINTZE, B. J., CHEN, V. B., JAIN, S., LEWIS, S. M., ARENDALL, W. B., 3RD, SNOEYINK, J., ADAMS, P. D., LOVELL, S. C., RICHARDSON, J. S. & RICHARDSON, D. C. 2018. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci,* 27**,** 293-315.
WILSON, D. N. 2014. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nat Rev Microbiol,* 12**,** 35-48.
WILSON, D. N. & NIERHAUS, K. H. 2005. Ribosomal proteins in the spotlight. *Crit Rev Biochem Mol Biol,* 40**,** 243-67.
WIMBERLY, B. T., BRODERSEN, D. E., CLEMONS, W. M., JR., MORGAN-WARREN, R. J., CARTER, A. P., VONRHEIN, C., HARTSCH, T. & RAMAKRISHNAN, V. 2000. Structure of the 30S ribosomal subunit. *Nature,* 407**,** 327-39.
WINN, M. D., BALLARD, C. C., COWTAN, K. D., DODSON, E. J., EMSLEY, P., EVANS, P. R., KEEGAN, R. M., KRISSINEL, E. B., LESLIE, A. G., MCCOY, A., MCNICHOLAS, S. J., MURSHUDOV, G. N., PANNU, N. S., POTTERTON, E. A., POWELL, H. R., READ, R. J., VAGIN, A. & WILSON, K. S. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr,* 67**,** 235-42.
WOODSON, S. A. 2011. RNA folding pathways and the self-assembly of ribosomes. *Acc Chem Res,* 44**,** 1312-9.
XIE, G., BRUCE, D. C., CHALLACOMBE, J. F., CHERTKOV, O., DETTER, J. C., GILNA, P., HAN, C. S., LUCAS, S., MISRA, M., MYERS, G. L., RICHARDSON, P., TAPIA, R., THAYER, N., THOMPSON, L. S., BRETTIN, T. S., HENRISSAT, B., WILSON, D. B. & MCBRIDE, M. J. 2007. Genome sequence of the cellulolytic gliding bacterium Cytophaga hutchinsonii. *Appl Environ Microbiol,* 73**,** 3536-46.
YAMAUCHI, K. 1991. The sequence flanking translational initiation site in protozoa. *Nucleic Acids Res,* 19**,** 2715-20.
YIP, K. M., FISCHER, N., PAKNIA, E., CHARI, A. & STARK, H. 2020. Atomic-resolution protein structure determination by cryo-EM. *Nature,* 587**,** 157-161.
YOKOYAMA, T., SHAIKH, T. R., IWAKURA, N., KAJI, H., KAJI, A. & AGRAWAL, R. K. 2012. Structural insights into initial and intermediate steps of the ribosome-recycling process. *EMBO J,* 31**,** 1836-46.
YOSHIZAWA, S., FOURMY, D. & PUGLISI, J. D. 1999. Recognition of the codon-anticodon helix by ribosomal RNA. *Science,* 285**,** 1722-5.
YUTIN, N., PUIGBO, P., KOONIN, E. V. & WOLF, Y. I. 2012. Phylogenomics of prokaryotic ribosomal proteins. *PLoS One,* 7**,** e36972.
ZAHER, H. S. & GREEN, R. 2011. A primary role for release factor 3 in quality control during translation elongation in Escherichia coli. *Cell,* 147**,** 396-408.

ZAMIR, A., MISKIN, R. & ELSON, D. 1969. Interconversions between inactive and active forms of ribosomal subunits. *FEBS Lett,* 3**,** 85-88.

ZAMIR, A., MISKIN, R. & ELSON, D. 1971. Inactivation and reactivation of ribosomal subunits: amino acyl-transfer RNA binding activity of the 30 s subunit of Escherichia coli. *J Mol Biol,* 60**,** 347-64.

ZAVIALOV, A. V., MORA, L., BUCKINGHAM, R. H. & EHRENBERG, M. 2002. Release of peptide promoted by the GGQ motif of class 1 release factors regulates the GTPase activity of RF3. *Mol Cell,* 10**,** 789-98.

ZENG-ELMORE, X., GAO, X. Z., PELLARIN, R., SCHNEIDMAN-DUHOVNY, D., ZHANG, X. J., KOZACKA, K. A., TANG, Y., SALI, A., CHALKLEY, R. J., COTE, R. H. & CHU, F. 2014. Molecular architecture of photoreceptor phosphodiesterase elucidated by chemical cross-linking and integrative modeling. *J Mol Biol,* 426**,** 3713-3728.

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

ZHANG, X. & BREMER, H. 1995. Control of the Escherichia coli rrnB P1 promoter strength by ppGpp. *J Biol Chem,* 270**,** 11181-9.

ZHENG, S. Q., PALOVCAK, E., ARMACHE, J. P., VERBA, K. A., CHENG, Y. & AGARD, D. A. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods,* 14**,** 331-332.

ZHONG, E. D., BEPLER, T., BERGER, B. & DAVIS, J. H. 2021. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nat Methods,* 18**,** 176-185.

ZHU, H. Q., HU, G. Q., OUYANG, Z. Q., WANG, J. & SHE, Z. S. 2004. Accuracy improvement for identifying translation initiation sites in microbial genomes. *Bioinformatics,* 20**,** 3308-17.

ZIVANOV, J., NAKANE, T., FORSBERG, B. O., KIMANIUS, D., HAGEN, W. J., LINDAHL, E. & SCHERES, S. H. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife,* 7.

ZIVANOV, J., NAKANE, T. & SCHERES, S. H. W. 2019. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ,* 6**,** 5-17.

ZIVANOV, J., NAKANE, T. & SCHERES, S. H. W. 2020. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ,* 7**,** 253-267.