STRATEGIES AND APPROACHES FOR MECHANISTIC UNDERSTANDING OF BACTERIAL RIBOSOME ASSEMBLY

A THESIS

submitted in partial fulfilment of the requirements
for the award of the degree
of

DOCTOR OF PHILOSOPHY

by

HIMANSHU SHARMA

DEPARTMENT OF BIOSCIENCES AND BIOENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

NOVEMBER 2018

Dedicated to maa, pa and shilpu

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

DEPARTMENT OF BIOSCIENCES AND BIOENGINEERING

STATEMENT

I do hereby declare that the matter embodied in this thesis entitled "**Strategies and approaches for the mechanistic understanding of bacterial ribosome assembly**" is the result of work carried out in the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, India, under the supervision of **Dr. B. Anand**.

In keeping with the general practice of reporting scientific observations, due acknowledgement has been made wherever the work described is based on the findings of other investigators.

28/06/19

Himanshu Sharma

Roll no. 11610608

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

DEPARTMENT OF BIOSCIENCES AND BIOENGINEERING

CERTIFICATE

It is certified that the work described in this thesis entitled **"Strategies and approaches for mechanistic understanding of bacterial ribosome assembly"** by Mr. Himanshu Sharma for the award of degree of Doctor of Philosophy is an authentic record of the results obtained from the research work carried out under my supervision in the Department of Biosciences and Bioengineering, IITG. The work embodied in this thesis has not been submitted elsewhere for a degree.

28/06/2019

Dr. B. Anand

Thesis Supervisor

Department of Biosciences and Bioengineering

Indian Institute of Technology Guwahati

Assam 781039, India

Acknowledgement

I feel the opportunity to pen down an acknowledgement, is one of the finest privileges and an equally great responsibility that one can have. A privilege to approach the finish line of this exciting part of the scientific journey and a responsibility to remember all with whose help, this journey of learning, improvement and introspection has been possible.

My heartfelt gratitude goes out to the head of the MAB lab, and my thesis supervisor Dr. B. Anand, for his guidance and support during every step of my work. His great passion to pursue the most challenging problems in science and to promote scientific temper can only be matched by the patience, care and independence he has bestowed upon me during my stay at IITG. I am grateful for his constant advice, follow-ups and knowledge because of which this thesis has become a reality.

I would also extend my gratitude to the members of my Doctoral committee, Dr. Biplab Bose, Dr. Vishal Trivedi, Dr. Nitin Chaudhary of the Department of Biosciences and Bioengineering and Dr. S. R. Mahadeva Prasanna of the Department of Electronics and Electrical Engineering for their valuable suggestions. I would also like to thank Dr. Senthil Kumar Sivaprakasam for his valuable suggestions during annual evaluations.

I would like to thank the Heads of the Department, Dr. Arun Goyal, Dr. V. Venkata Dasu and Dr. Kannan Pakshirajan as well as the other faculty members their support and advice.

I would also like to thank all members of the MAB lab who have given me a genial and edifying environment in the lab. I would like to acknowledge the help I got from Rohan, Sumit and Kiran for the work performed in the chapter IV and V. I would also like to thank Sunanda for the help in preparing the figures and Manasa for her help in proofreading the thesis. I also appreciate the important suggestions from Yoganand and Siddharth during the thesis preparation. I am grateful for their constant support, encouragement and suggestions

for this work. We have gathered some great memories while working together as colleagues and even more recollections as friends. I would always cherish the memories I made with my friends at IITG who have played varied roles of colleagues, life coaches, teachers, family members and mischief-makers, thus making the journey an eventful experience.

I would also like to thank the Department of Biosciences and Bioengineering as well as its staff for providing me all the logistical support essential to perform my research. I would also like to thank the Central Instrumentation Facility and Department of Chemistry, IITG for allowing the use of scientific instruments. I would also like to appreciate the geniality of all the investigators who shared their research material and reagents.

My next thanks go to my alma mater, IIT Guwahati which gave me an opportunity of embarking on this fine endeavour of learning with ample facilities and an equally enchanting natural bounty of the North-east. My interactions with the fraternity at IITG has given me a first-hand experience of the vast cultural diversity of our country and this has been a great personal learning.

Next, I would thank the funding agencies that provided the financial support to conduct my research. I would like to acknowledge the support from Department of Biotechnology, Govt. of India, Department of Science and Technology, GOI for the research grants provided to our lab. I would also like to acknowledge MHRD, GOI for the fellowship during my Ph.D. tenure.

Finally, I would like to remember my parents and family for their never-ending love, support and patience towards me. It is their numerous personal sacrifices that have enabled me to reach this juncture in life.

Thank you and Godspeed!!

Himanshu Sharma

Table of Content

List of Figures

List of Tables

Abbreviations

AP-MS	Affinity Purification coupled Mass Spectrometry
A-site	Acceptor Site
ATP	Adenosine Tri-Phosphate
BAP	Biotin Acceptor Peptide
BccP	Biotin Carboxyl Carrier Protein
BiFC	Bimolecular Fluorescence Complementation
BioID	Proximity-dependent Biotin Identification
bp	Base pair
BSA	Bovine Serum Albumin
DNA	De-oxy Ribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EFs	Elongation factors
E-site	Exit site
GFP	Green fluorescent Protein
GTP	Guanosine Tri-Phosphate
hr/hrs	hour/ hours
IFs	Initiation Factors
IMAC	Immobilized Metal Affinity Chromatography
iTRAQ	Isobaric tags for relative and absolute quantitation
kb	Kilo bases
kDa	Kilo Dalton
MDa	Mega Dalton
Mg^{2+}	Magnesium ion
mins	Minutes
mM	millimolar
MNase	Micrococcal Nuclease
mRNA	Messenger RNA
ONPG	Ortho-Nitrophenyl-β-Galactoside
PAGE	Poly acrylamide Gel Electrophoresis

PCR	Polymerase Chain Reaction
PMSF	Phenylmethane Sulfonyl Fluoride
PPI	Protein-Protein Interactions
P-site	Peptidyl Transferase Site
FP	Ribosome Footprint
RFs	Release Factors
Ribo-seq	Ribosome Footprint Sequencing
RNA	Ribonucleic Acid
RNAP	RNA polymerase
RNase	Ribonuclease
r-Protein	Ribosomal Protein
RRF	Ribosome Recycling Factor
rRNA	Ribosomal RNA
RT	Reverse Transcriptase/ Room Temperature
SD sequence	Shine-Dalgarno Sequence
SDS	Sodium Dodecyl Sulphate
secs	Seconds
SILAC	Stable Isotope Labeling by Amino acids in Cell culture
SPA	Sequential Peptide Affinity
TAP	Tandem Affinity Peptide
tRNA	Transfer RNA
tRNAe	Elongator tRNA
$tRNA^{i}_{fmet}$	Initiator Transfer RNA
Y2H	Yeast two hybrid
β-ME	2-Beta-Mercaptoethanol

Chapter I

Introduction

CHAPTER I

Chapter 1

1.1 Introduction

A central aspect to understand the functioning of living systems is to unravel the multilayer complexity involved in implementing and regulating basic chemical reactions in the cell. Synthesis of proteins is one such omnipresent reaction. The significance of this process to life can be understood by the fact that the protein synthesis machinery has been the most conserved set of cellular arsenals over the course of evolution (Fox and Naik, 2004; Hsiao et al., 2009; Mushegian and Koonin, 1996; Sousa et al., 2016). At the heart of this process are ribosomes, which are large ribonucleoprotein complexes that polymerize amino acids as directed by a messenger RNA (mRNA) template.

Ribosomes were discovered in the mid-1950s as RNA-rich, spherical embellishments on the endoplasmic reticulum of rat tissues (Palade, 1955). They were later shown to be composed of two asymmetric subunits (Tissieres and Watson, 1958) and associated with the protein synthesis mechanism (Loftfield, 1957; Roberts, 1958; Ts'o, 1962). For long, the property to catalyse peptide bond synthesis was believed to rest in its protein component (Nierhaus and Nierhaus, 1973; Nierhaus and Montejo, 1973). However, it was only during the early-1990s that RNA was proposed to be the catalytic core of this macromolecular complex (Noller et al., 1992), only to be later confirmed by the crystal structure of the large ribosomal subunit (Nissen et al., 2000). The discovery of the RNA rather than the protein as the active component of this ribonucleoprotein assemblage propelled RNA further to the centre stage of studies regarding RNA mediated catalysis and the origins of life (Cech and Bass, 1986; Robertson and Joyce, 2012).

In the year 2000, ribosome research took another giant leap when the first high-resolution crystal structures of the bacterial ribosomes were published (Ban et al., 2000; Schluenzen et al., 2000; Wimberly et al., 2000). The crystal structures epitomized four decades of arduous efforts to gather critical structural insights by means of negative staining electron microscopy (Lake, 1976), immuno-electron microscopy (Stoffler and Stoffler-

Meilicke, 1984), Cryo-Electron Microscopy (Mitra and Frank, 2006; Penczek et al., 1994), neutron scattering analysis (Nowotny et al., 1986) and X-ray crystallography (Pennisi, 1999).

These breakthroughs opened up new frontiers of ribosome structural studies and revolutionized our understanding of protein synthesis (Carter et al., 2000; Trobro and Åqvist, 2005).

1.2 Structure of the bacterial ribosome and their functional implications

The previous decade has seen a surge in the wealth of structural information that enriched our knowledge and understanding about structure and function of the ribosome across all domains of life (Ban et al., 2000; Ben-Shem et al., 2011; Greber et al., 2012; Khatter et al., 2015; Schluenzen et al., 2000; Wimberly et al., 2000). However, given the vast expanse of the subject matter and a discrete focus of the work described in this thesis, this chapter will focus on the ribosome structure, function and assembly principles pertaining to the model organism *Escherichia coli* (*E. coli*).

1.2.1 Ribosomal subunits share some structural attributes

The bacterial ribosome settles as a 70S particle during sedimentation analysis and is composed of two asymmetrical subunits that sediment as 30S and 50S particles. The 30S subunit consists of 16S ribosomal RNA (rRNA) and 21 ribosomal proteins (r-Proteins). The larger, 50S particle is made up of 23S rRNA, 5S rRNA, and 33 r-Proteins. The eukaryotic counterpart of 70S sediments as an 80S particle that harbours extended rRNA composition (Hassouna et al., 1984; Ware et al., 1983) and 20-30 additional r-Proteins (Spahn et al., 2001). In spite of these gross compositional differences, the two particles share a common functional core whose structure is dominated by the rRNA (Spahn et al., 2001; Wilson and Nierhaus, 2003; Wimberly et al., 2000). The rRNA folds into extensive helical structures whose minor grooves closely pack against each other, similar to those observed in group I intron structures (Cate et al., 1996; Szewczak et al., 1998). The r-Proteins bind to the rRNA core with long helical extensions to stabilize long-range interactions of the rRNA (Klein et al., 2004;

Wimberly et al., 2000). On the contrary, their globular domains are exposed on the subunit surface. Strikingly, the subunit interface of both the subunits is essentially free of r-Proteins and harbours 12 inter-subunit bridges serving as anchor points for the association of subunits. The interface also harbours regions of the subunit functional core with sites that accommodate transfer-RNA (tRNA) and mRNA. Three tRNA accommodation sites namely, the **A**cceptor (A-site), **P**eptidyl transferase (P-site) and **E**xit (E-site) are located on both subunits. During translation tRNAs enter from the A-site and adopt hybrid conformations while moving from A to P to E site and are eventually ejected from the ribosomes (Dorner et al., 2006).

1.2.2 The 30S subunit harbours the decoding centre

The 30S subunit has a molecular weight of about 0.8 MDa, three fourth of which rests in the 16 rRNA (Ehresmann et al., 1975; Ortega and Hill, 1973). The 16S rRNA folds into distinct helical structure numbered from helix 1 to helix 45 (h1-h45). These secondary structures are packed using minor groove to minor groove and phosphate to minor groove interactions forming four distinct domains: 5', central, 3' major and 3' minor (Fig. 1.1A) (Wimberly et al., 2000). The four domains confer distinct morphological features identified as the body, platform, and the head region. The 5' domain forms the wedge-shaped part of the body and spur regions (Fig. 1.1B) and the platform region is formed by the central domain, where the h21-h23 form the outer boundary of the platform region. Unlike the 5' and the central domains, components of the 3' domain are rather loosely packed. The 3' domain is connected to the rest of the structure by two hinges in the neck region formed by h28, h34, and h35, which serve as the points of rotation during elongation (Hagerman, 1997; Schuwirth et al., 2005). The 3' domain can be divided into two sub-domains, the 3' major domain that forms a distinctly recognizable head region with a protruding "beak". The head is separated from the platform by a deep cleft region that binds to the Shine-Dalgarno sequence (SD) of the incoming mRNA (Julián et al., 2011). The other part, 3' minor domain is made up of only two helices (h44 and h45). The helix 44 spans from the head to the entire length of the body making it the longest helix in of the 16S rRNA. The h44 also contains part of the A-site, the two highly conserved tRNA binding residues A1492 and A1493 (Fig 1.2A). Additionally, the

base G530 and helical extensions from at least five r-Proteins stabilize the A-site-tRNA interactions. Adjoining to the A-site lies the P-site formed by the conserved bases G926, G966, G1338, A1339, and C1400. The 30S subunit closely surveys the bases pairing at the A and P site while the E-site, formed by regions from h29, loops 690 and 790 is scrutinized to a lesser extent.

1.2.3 The catalytic core rests in the 50S subunit

The 50S subunit is a 1.5 MDa, hemispherical particle of roughly 250Å in diameter. The larger subunit is nearly a monolithic structure in contrast to well-spaced, distinct morphological domains of the smaller subunit. The 23S rRNA is divided into 6 (I-VI) domains with Helix 1-Helix 101 (H1-H101) and the 5S rRNA forms the seventh rRNA domain of the 50S subunit (Fig. 1.1A) (Ban et al., 2000). When observed from the subunit interface, the 50S displays a crown-like structure called the central protuberance (CP) that is formed by the 5S rRNA. At roughly -60° and +60° of CP, lies the left and the right protuberance or stalk formed by L1 and L7/L12 proteins, respectively (Fig.1.1C). At the centre of the 50S subunits, below the CP lies the P-site that contains the peptidyl-transferase centre (PTC) or the catalytic base A2451. Other regions of the P-site comprise of the residues G2251-G2252 and C2063 (Fig. 1.2B). The residue A2451 provides a hydroxyl group to the P-site tRNA thus effectively initiating a proton shuttle mechanism leading to the peptide bond synthesis reaction (Katunin et al., 2002; Lang et al., 2008; Schuwirth et al., 2005; Simonović and Steitz, 2008). The PTC core is virtually free of r-Proteins and L1, L33, L27, L16, L5 are known to perform the auxiliary function of stabilizing the P-site tRNA (Fig. 1.2B). Although passive to the catalytic process, mutation and truncation of proteins that are adjacent to PTC have shown a significant drop in rates of peptide bond synthesis (Kirillov et al., 2002; Maguire et al., 2005; Selmer et al., 2006). In close vicinity, lies the A-site, formed by the residues G2553, C2573, G2583, U2584 and U2506 (Fig. 1.2C). Beneath the PTC, lies a 100Å narrow peptide exit tunnel lined by the rRNA, r-Proteins L4, and L22.

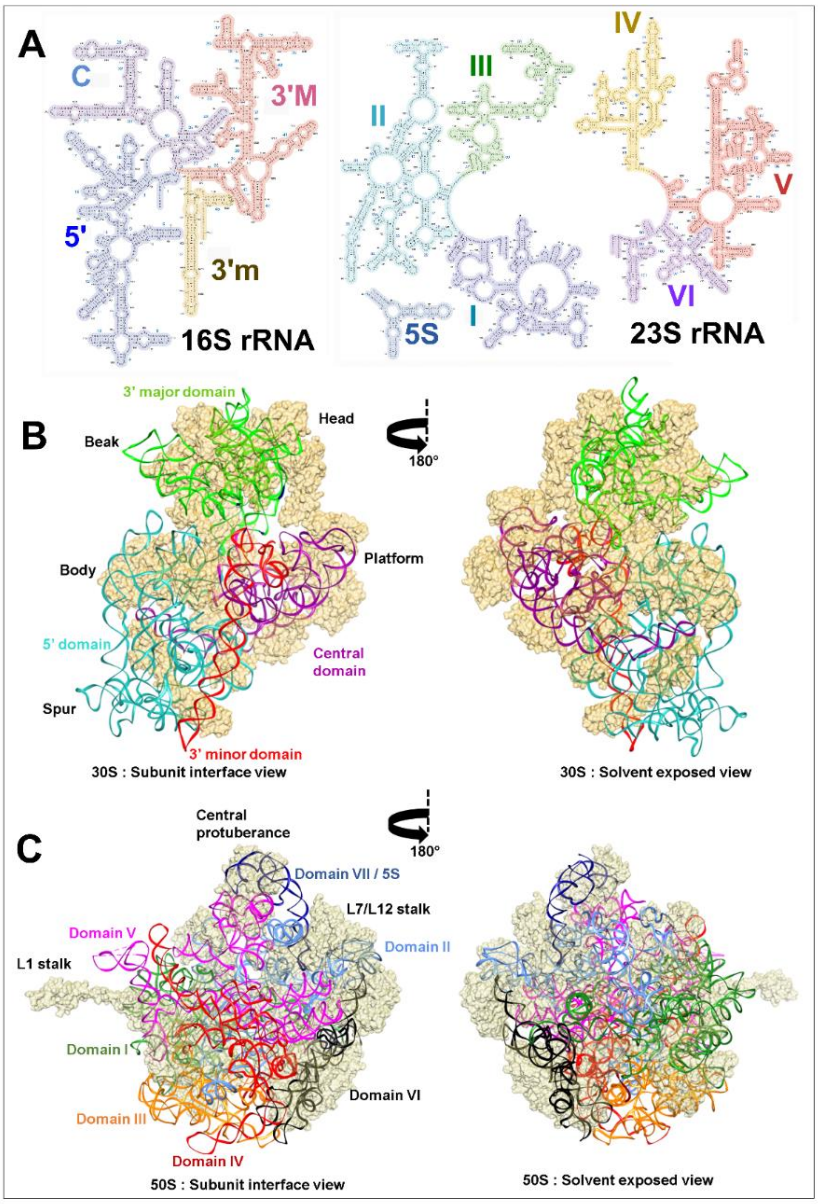

Figure 1.1: Ribosomal RNA defines the structure of the ribosomal subunits

(A) The secondary structure of 16S rRNA (left) and 23S rRNA (Right) from *E. coli.* The domains of 16S rRNA (5': 5' major, C: Central, 3'M: 3' major, 3'm: 3' minor) and 23S rRNA (Domain I-VI and VII formed by the 5S rRNA) are highlighted.

(B) A structural representation of the 30S subunit from *E. coli* (PDB ID: 2QAL) viewed from the subunit interface and the solvent-exposed side (as indicated). RNA is shown as coloured ribbons with indications for the corresponding domains. The r-Proteins are shown in a surface representation in brown. The morphological units of the subunit are indicated.

(C) A structural representation of the 50S subunit from *E. coli* (PDB ID: 2QAL) observed from the subunit interface as well as the solvent-exposed side. RNA is shown as coloured ribbons with indications for the corresponding domains. The r-Proteins are shown as surface representation in wheat. The morphological units of the subunit are indicated.

*****Secondary structures were generated using XRNA and Inkscape. Molecular models were generated using Chimera molecular viewer software.

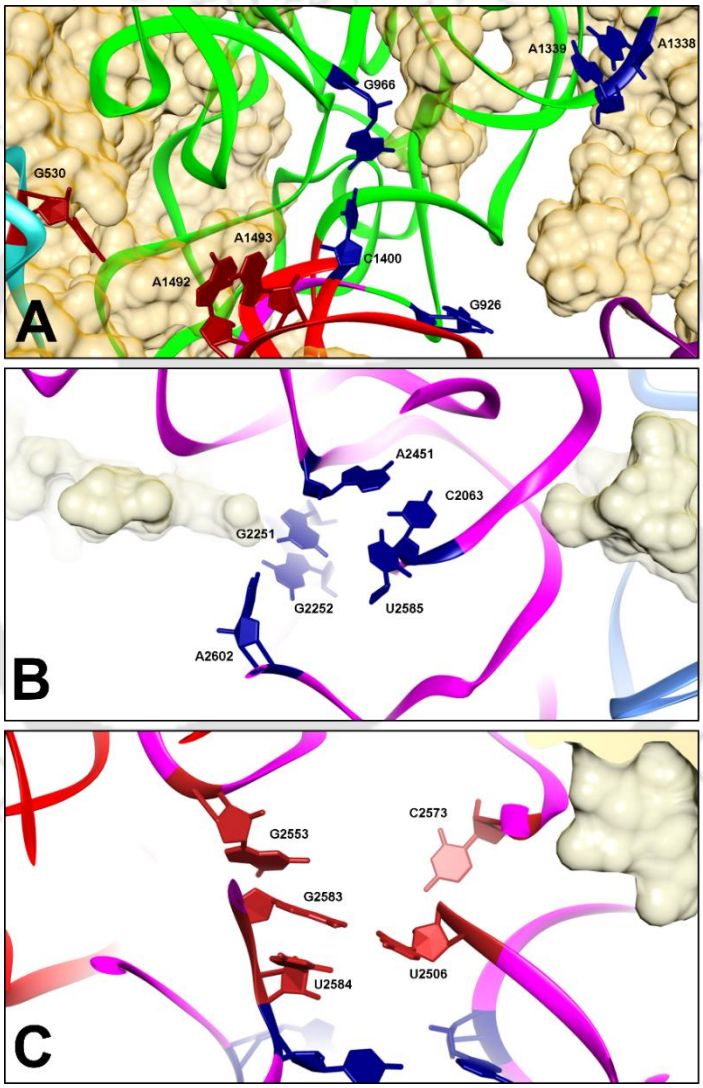

Figure 1.2: Ribosomal RNA forms the functional core of the ribosome

(A) View of the P-site and the A-site of the 30S subunit. RNA backbone is represented as ribbons and r-Proteins are in surface representation. Residues of the 30S P-site are shown as purple-filled rings and A-site are shown as dark red-filled rings.

(B) View of the 50S subunit P-site. RNA backbone is represented as ribbons and r-Proteins are in the surface representation. Active site residues of the P-site are shown as purple –filled rings. The residue A2451 is the catalytic residue in *E. coli*.

(C) View of the 50S subunit A-site. RNA backbone is represented as ribbons and r-Proteins are represented in the surface representation. Residues of the A-site are shown as dark red –filled rings.

*****Models were generated using Chimera molecular viewer software.

1.3 **The translation cycle**

The reaction catalysed by the ribosome is the aminolysis of an ester bond. The α-amino group of the A-site aminoacyl tRNA (aa-tRNA) attacks the carbonyl carbon of the elongating peptide attached to the P-site tRNA (Satterthwait and Jencks, 1974). Unlike protein-mediated enzymatic catalysis that involves catalytic residues with near neutral pk_a values (Alexov and Gunner, 1997; Bashford and Karplus, 1990; Gaylord and Gibbs, 1962; Herries, 1985), the ribosome catalyzes the peptide bond synthesis by orienting the reactants in an orientation that initiates a charge relay (Hansen et al., 2002; Moore and Steitz, 2003; Sievers et al., 2004; Trobro and Åqvist, 2005).

The translation cycle can be divided into four major events: (i) Initiation, (ii) Elongation, (iii) Termination and (iv) Recycling (Fig. 1.3). In prokaryotes, translation is closely coupled with transcription events as 30S subunits directly bind to the SD sequence of the nascent RNA emerging from the RNAP complex to initiate translation (Demo et al., 2017; Steitz and Jakes, 1975). Our current knowledge of initiation process comes mainly from the translation by monosomes as nature of initiation events involving polysomes is only starting to emerge (Andreeva et al., 2018). Three factors *viz*., Initiation factor 1 (IF-1), IF-2 and IF-3 mediate the initiation process, where IF-1 blocks the A-site and seems to enhance the activity of IF-2 and IF-3 (Carter et al., 2001; Gualerzi and Pon, 1990). IF-2 binds to the P-site and also recruits the initiator tRNA ($tRNA^{i}_{fmet}$) (Gualerzi and Pon, 1990) whereas IF-3 blocks the E-site and later verifies the codon-anticodon (CO-AC) pairing between the mRNA and the $tRNA^{i}_{fmet}$ (Hussain et al., 2016; Laursen et al., 2005). The order of arrival of IFs is not expected to be linear but precede tRNA and mRNA binding states (Hussain et al., 2016; Tsai et al., 2012). The 30S-IFs-tRNA and mRNA form a ternary complex called the 30S-initiation

complex (30S-IC) (Fig. 1.3A) (Gualerzi et al., 1977; Milon et al., 2012) which subsequently binds to the 50S subunit to form a 70S initiation complex (70S-IC) (Allen and Frank, 2006; Milon et al., 2008) (Fig. 1.3A). The 70S-IC eventually recruits the charged elongator tRNA (tRNAe) carried by the Elongation factor Tu (EF-Tu) at the A-site (Kavaliauskas et al., 2018; Kavaliauskas et al., 2012) to initiate the process of decoding (Fig. 1.3B). The CO-AC pairing is also checked here, followed by the reorientation of the A & P-site tRNAs for the peptidyl synthesis reaction. Next, the Elongation Factor G (EF-G) catalyses translocation during which the P-site resting tRNA shifts to the E-site and the peptidyl chain carrying tRNA moves to the P-site vacating the A-site for the next incoming tRNAe (Fig. 1.3B) (Pulk and Cate, 2013; Tourigny et al., 2013). Once completed, the ribosomes enter the next cycle of elongation that involves multiple cycles of decoding, translocation, and movement of the ribosome on the mRNA, one codon at a time. Additionally, Elongation Factor P (EF-P) assists the ribosomes during the elongation steps at proline-rich peptides (Doerfel et al., 2013). Finally, the stop codon signals the end of the translation process as Release Factor 1 or 2 (RF1/2) bind the A-site to catalyse the release of the peptide chain (Korostelev, 2011; Youngman et al., 2008). RF-3, EF-G, and ribosome recycling factor (RRF) help the release of the RFs and dissociate the 70S particles for the next cycle of initiation (Fig. 1.3C) (Chen et al., 2017; Freistroffer et al., 1997; Hirashima and Kaji, 1972; Karimi et al., 1999).

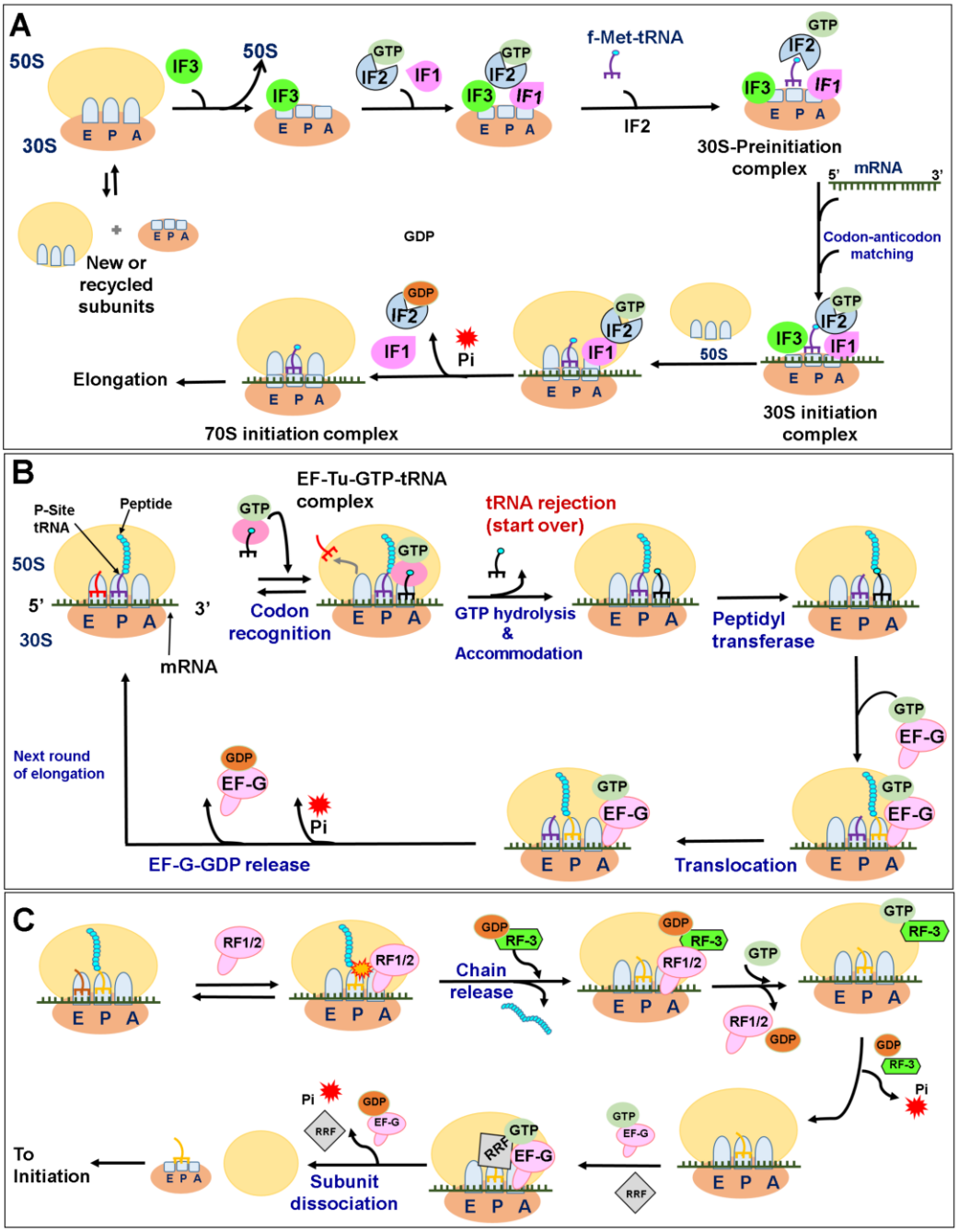

Figure 1.3: A simplified view of the translation cycle

(A) A schematic flow showing the events of translation initiation in bacteria.
(B) A representation of the events of translation elongation in bacteria.
(C) Events leading up to translation termination and ribosome recycling.
*The figure has been recreated using the following references (Laursen et al., 2005; Steitz, 2008; Watson, 2004)

1.4 Biogenesis of bacterial ribosomes

The synthesis of bacterial ribosomes is a complicated and energy-intensive process. Creation of this functional superstructure involves multiple steps, involving the synthesis of the rRNA precursors and r-Proteins from independent operons spread across the genome. This follows the processing and modification of rRNA and its subsequent assembly with r-Proteins to form functional subunits. Given the numbers of players involved and the functional significance, the process needs to be efficient by demand and therefore is tightly regulated (Maitra and Dill, 2015). Studies into ribosome biogenesis have been closely intertwined with the empirical and quantitative model defining bacterial physiology. The bacterial "growth law" proposes a linear relation between the ribosome content and the growth rate of cells at fast to medium growth rates (Neidhardt, 1999; Schaechter et al., 1958; Scott and Hwa, 2011). This simple model, however, deceives the intricate regulation and coordination the cells employ to synthesize the protein factories in a regulated yet efficient manner to cope up with the nutritional shifts (Dai et al., 2016) or stress (Starosta et al., 2014).

1.4.1 Ribosomal RNA synthesis and regulation

Ribosomal RNA is encoded from seven operons (*rrnA, rrnB, rrnC, rrnD, rrnE, rrnG, rrnH*) in *E. coli* (Fig. 1.4). The operons share a common structural pattern with each operon containing genes encoding 16S, 23S and 5S rRNA along with at least one tRNA. An interesting feature of the *rrn* operons is the intricate architecture that confers the strong yet highly regulated expression. A typical *rrn* operon encodes a single transcript driven under two promoters *rrn*P1 and *rrn*P2 (Bartlett and Gourse, 1994; Gafny et al., 1994; Murray et al., 2003). Upstream to the core promoter region one or two UP elements, recognized by the α-subunit of the RNA Polymerase (RNAP) stimulate the promoter activity by up to 50-fold (Estrem et al., 1998; Hirvonen et al., 2001). Upstream to the UP site, the operons house three to five transcription factor Fis (**F**actor for **i**nversion **s**timulation) binding sites (Bokal et al., 1995; Ross et al., 1990) which are also used to stimulate the transcription. The operon also employs, "anti-terminator" systems like Nus factors and r-Proteins binding sites that allow read through at rho-dependent terminators for synthesis of large RNA molecules (Das et al.,

1985; DeVito and Das, 1994; Sharrock et al., 1985; Torres et al., 2004; Turtola and Belogurov, 2016). Additionally, mechanisms involving H-NS mediated downregulation of *rrn*P1 (Afflerbach et al., 1998), Guanosine tetraphosphate (ppGpp) and pentaphosphate (pppGpp) mediated inhibition of RNAP (Ryals et al., 1982; Srivatsan and Wang, 2008) and limiting transcription initiation at low NTP concentration (Gaal et al., 1997) are used to downregulate rRNA synthesis.

1.4.2 Processing and modification of rRNA

The single rRNA transcripts need further processing and modifications to be functional. The exact role of synthesizing precursors still remains unclear, however, mutational studies suggest that precursor sequences may play a role in coupling rRNA processing with various stages of the assembly (Gutgsell and Jain, 2012). Initial insights into rRNA processing came from studies on null mutants of RNase III that accumulated a single rRNA molecule that migrates as a 30S particle (Dunn and Studier, 1973). Hence, RNase III was identified as the first enzyme to initiate the multistep rRNA maturation process. RNase III, a double-stranded RNA targeting nuclease, executes the first cleavage on the primary rRNA transcript to generate precursor pre-16S (17S), pre-23S (p23S) and a pre-5S (9S) fragments (Bram et al., 1980; Ghora and Apirion, 1978; Young and Steitz, 1978). The precursors are then directed for maturation by specific RNases for each transcript (Table 1.1) (figure1.4). The progress in identification of the RNases has been rather slow due to the redundant and obscure nature of the maturation process (Li et al., 1999b; Sulthana and Deutscher, 2013). However, the general theme that emerges from the current knowledge on rRNA processing indicated that the precursor secondary structures dominate the order of the processing (Deutscher, 2009). It is also noteworthy that many of the nucleases deployed in rRNA maturation also play a critical role in general RNA turnover (Deutscher, 2006).

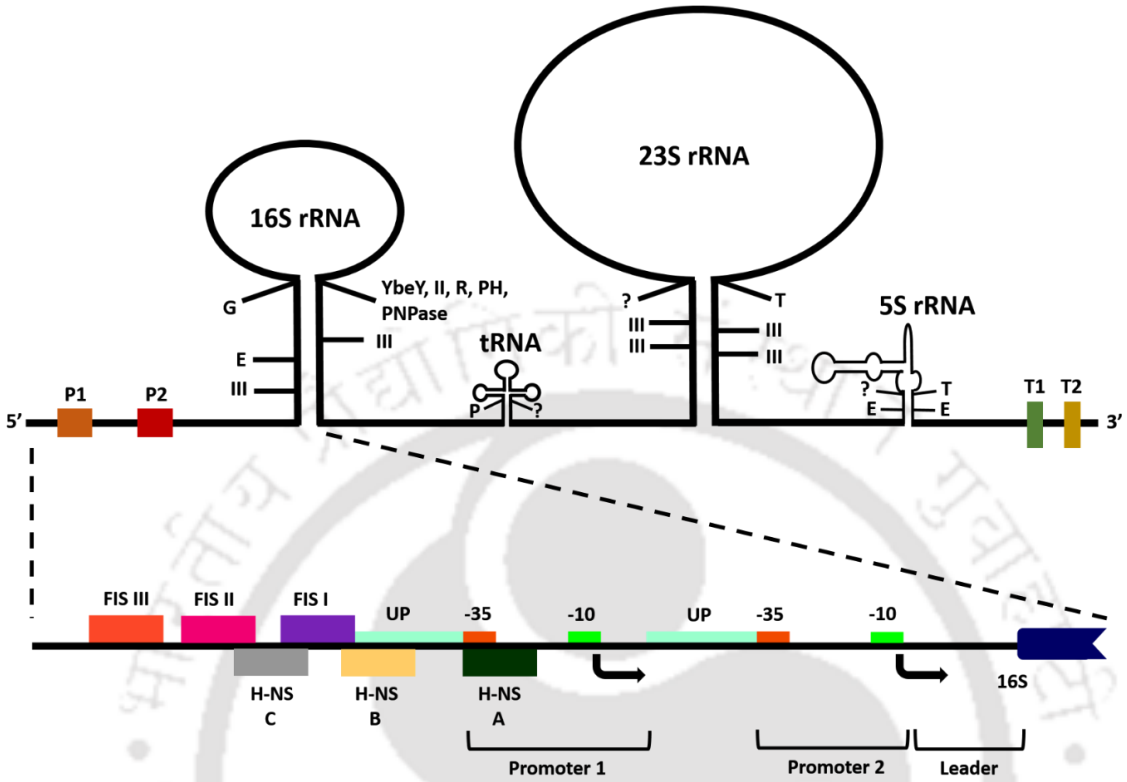

Figure 1.4: The ribosomal RNA operons share common structural architecture and processing scheme

Representation of the structure of an *rrn* operon encoding for 16S, 23S, and 5S rRNA. The promoter *rrn*P1 and *rrn*P2 are shown with the transcription stimulator regions UP and Fis binding sites. H-NS binding regions are also represented. Various RNases acting on the primary transcripts and premature rRNA species are also indicated.

Table 1.1: Enzymes catalysing the maturation of rRNA in *E. coli*

Substrate	Enzyme	Product	References
Primary 30S rRNA transcript	RNase III	17S, p23S, 9S	(Bram et al., 1980; Ghora and Apirion, 1978; Young and Steitz, 1978)
17S: 5' ends	RNase E	16.3S (pre-16S with 66 ucleotides (nt) on the 5' end)	(Dahlberg et al., 1978)
16.3S: 66 nt on the 5' ends	RNase G	16S with mature 5' end	(Li et al., 1999b; Wachi et al., 1999)
pre-16S: 33 nt on 3' end	RNases II, R, and PH, PNPase, YbeY	Mature 16S rRNA	(Jacob et al., 2013; Sulthana and Deutscher, 2013)
p23S: 7 or 9 nt on 3' end	RNase T	p23S with mature 3' end	(Li et al., 1999a)
p23S: 7 or 3 nt on the 5' end	Unknown	---	(Shajani et al., 2011)
9S: pre-5S with 5' and 3' extensions	RNase E	pre-5S with 3 nt on 5' and 3' ends	(Ghora and Apirion, 1978; Misra and Apirion, 1979)
pre-5S : 3 nt on the 3' end	RNase T	pre-5S with 3 nt on 5'end	(Li and Deutscher, 1995)
pre-5S : 3 nt on the 5' end	Unknown	---	---

Ribosomal RNA is modified in primarily three ways to expand its functional repertoire: (i) Conversion of uridines to pseudouridine (Penzo and Montanaro, 2018), (ii) Methylation of 2' OH group of the ribose sugar and (iii) Methylation of the nucleotide bases (Sergiev et al., 2018). The 16S rRNA is known to undergo 11 modifications (1 pseudouridylation, 10 methylations) and 23S rRNA is known to be modified at 25 positions (14 pseudouridylations, 9 methylations, 1 methylated pseudouridylation, and 1 unknown modification) (Shajani et al., 2011). In *E. coli*, most of the modifications (~95%) cluster around the functional centres (Decatur and Fournier, 2002) and seem to enhance the ability of rRNA to form extensive secondary structures, bind regulatory proteins with higher

affinities and even stabilize subunit interactions (Polikanov et al., 2015). Modifications also seem to play a critical role in ribosome functions, as unmodified 16S forms 30S subunits with suboptimal tRNA binding and translating capacities (Denman et al., 1989). However, these requirements seem more stringent for 23S rRNA modifications which are essential for subunit assembly (Green and Noller, 1996).

1.4.3 Ribosomal protein synthesis and regulation

Similar to rRNA, genes encoding r-Proteins are also organized into operons that are conspicuously regulated by intricate feedback mechanisms (Draper, 1987). However, operons encoding r-Proteins do not share a common structural architecture and also contain genes encoding components of the translation machinery, chaperones, nucleases and modification enzymes (Bedwell et al., 1985; Cerretti et al., 1983; Coenye and Vandamme, 2005). The operons display primary and secondary promoter regions along with transcription read-through mechanisms between operons (Zengel and Lindahl, 1994). Ribosomal proteins nucleating on rRNA during early stages of assembly bind to mRNA encoded by the respective operon, thus asserting a feedback loop (Nomura et al., 1984; Nomura et al., 1980; Zengel and Lindahl, 1994) (Fig. 1.5). Additionally, recent studies have also unravelled the indirect regulation of r-Protein synthesis by external factors like pppGpp and DskA under nutrient limiting conditions (Aseev et al., 2016; Lemke et al., 2011)

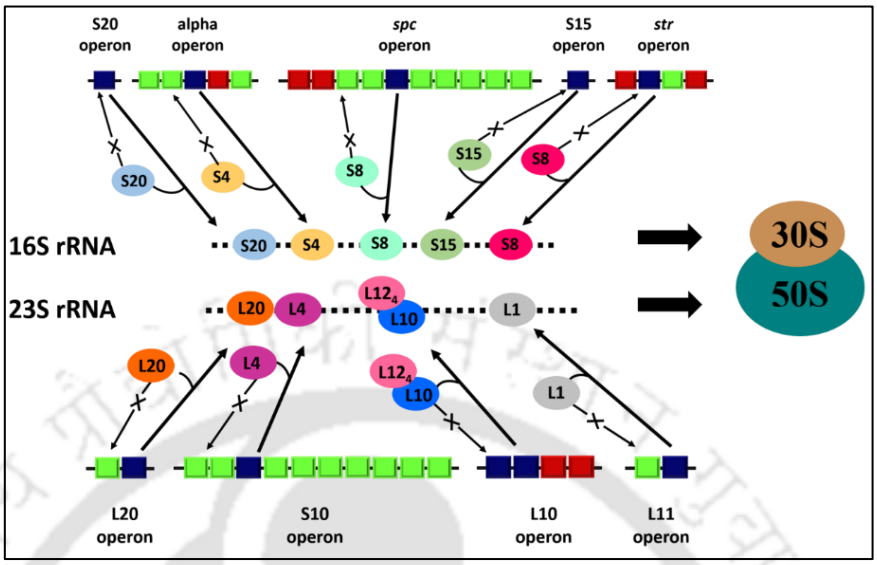

Figure 1.5: Ribosomal proteins regulate their expression autogenously

Representation of the feedback regulation exerted by r-Proteins on the mRNA transcript. Operons encoding r-Proteins are shown as boxes with the regulatory protein shown in purple. The regulatory protein binds to the transcribed mRNA and the expression of genes they regulate is shown in green. Unregulated genes are shown in red boxes. Ribosomal RNA is represented as dashed lines.

1.4.4 Early investigations into ribosome assembly revealed its hierarchical nature

The timeline of studies into ribosome assembly runs in parallel with investigations into their structure and composition (Mangiarotti et al., 1968; McCarthy et al., 1962; Nomura, 1973). Early investigations relied on pulse-chase experiments to measure the integration of rRNA and r-Proteins into ribosome precursors (then named as “eosome” and “neosome”) that could mature into functional subunits. The resolution of these studies although limited, established that ribosome maturation progressed at a rapid rate in actively growing bacterial cells (Leive, 1965). Further *in vitro* studies employed partial dissociation of ribosomes to generate “core particles” and “split protein” fractions followed by re-association into a functional subunit (Staehelin and Meselson, 1966). A major advancement on this was achieved by Nomura and co-workers (Traub and Nomura, 1968a) by deciphering the order of r-Protein binding to the rRNA. The study involved the stepwise addition of r-Proteins to the rRNA and thus by withholding certain r-Proteins, intermediates generated in their absence

could be used to derive their preferential order of nucleation on the naked rRNA chain (Fig. 1.6A) (Mizushima and Nomura, 1970). As it turned out, certain r-Proteins could bind the nascent rRNA chain directly and be accordingly classified as primary binding proteins. Similarly, other proteins required the binding of primary binders to facilitate their nucleation and such proteins formed the secondary binders followed by tertiary binding proteins. This suggested thermodynamically independent binding of certain r-Proteins but hierarchical binding of others to the rRNA. Similar experiments by Nierhaus and co-workers (Fig. 1.6A) (Nierhaus and Dohme, 1974) established similar observations for the larger subunit. However, the map is obviously more complicated on account of the large number of proteins and larger rRNA being assembled. These findings reinforced then evolving views of the spontaneous and self-driven ribosome assembly process (Nomura and Traub, 1966). The *in vitro* reconstitution experiments, however, differed from the findings of *in vivo* studies, as it was extremely slow and therefore the test tube reconstitution of functional ribosomes required heat to raise the temperature to non-physiological growth conditions and high salt concentrations (Nomura, 1973; Sieber and Nierhaus, 1978; Spillmann et al., 1977).

Owing to the rapid assembly events in *vivo*, accretion of ribosome intermediates was scant in rapidly growing cells. Therefore, in order to capture ribosome intermediates alternative strategies were employed such as generation of cold-sensitive mutants via chemical mutagenesis (Guthrie et al., 1969), isolation of fragile mutants (Mangiarotti et al., 1968) and genetic manipulations (Nikolaev et al., 1973; Spillmann et al., 1977). Such mutants ensued precursor particles for the 30S and 50S subunits which varied from mature subunits in terms of composition (Hayes and Hayes, 1971; Lindahl, 1973b) or structure (Lindahl, 1975; Sieber and Nierhaus, 1978) (Fig. 1.6B) suggesting that *in vivo* assembly also involved multiple intermediates (Lindahl, 1973a, 1975). Additionally, assembly studies in higher organisms displayed binding of nucleolar proteins to late-stage assembly intermediates, suggesting a role of non-ribosomal proteins in ribosome assembly (Hadjiolov and Nikolaev, 1978). These observations gave way to then fledgling views supporting a multi-step, cooperative assembly process to generate ribosome precursors that would be acted upon by auxiliary proteins to overcome activation energy barriers during rRNA folding events (Sieber and Nierhaus, 1978).

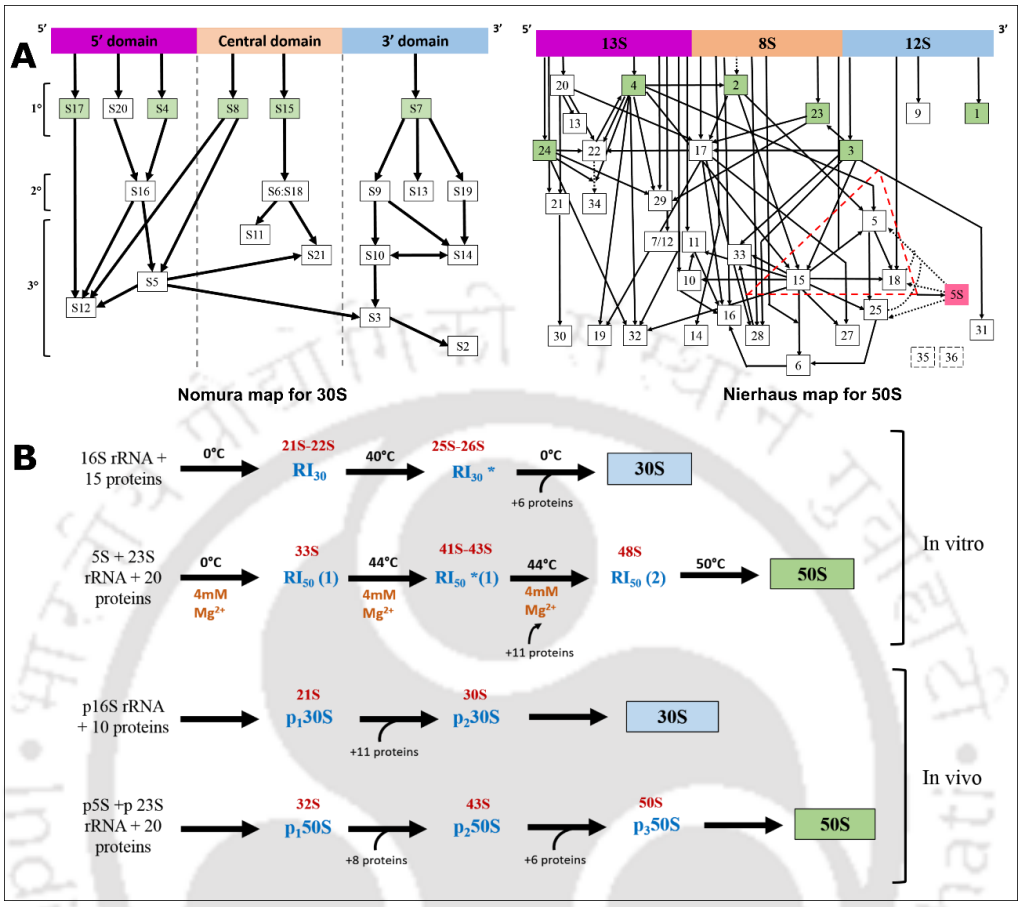

Figure 1.6: Ribosome assembly is a multistep-hierarchical process

(A) *In vitro* assembly map for ribosomal subunits derived by Nomura for the 30S subunit (left) and Nierhaus for the 50S subunit (right). The rRNA is shown as a bar on the top with respective domains and dependencies of r-Protein binding are indicated by directional arrows. Ribosomal proteins shown in filled boxes represent most ancient proteins and possibly co-evolved with the rRNA core (Hury et al., 2006; Korobeinikova et al., 2012).

(B) A comparison of the subunit assembly intermediates observed from *in vitro* and *in vivo* investigations. The scheme of *in vitro* assembly also shows the non-physiological conditions required to drive the assembly (Shajani et al., 2011).

1.4.5 Ribosome assembly is an "RNA folding" problem

Early structural studies of the ribosome suggested a major contribution of internal interactions in the folding of the 16S rRNA, and the requirement of r-Proteins to provide additional compactness and stability (Ramakrishnan, 1986; Serdyuk et al., 1983). These predictions were eventually confirmed by the observations of intricate packing of the RNA components and stabilization provided by r-Proteins in the crystal structure of both subunits (Ban et al., 2000; Wimberly et al., 2000). Additionally, kinetic studies of rRNA folding and co-transcriptional assembly supported an RNA guided assembly process (Adilakshmi et al., 2005; de Narvaez and Schaup, 1979; Stern et al., 1989). Further, ribosome reconstitution experiments established that all three domains of the 16S rRNA could independently drive their assembly *in vitro*, only to be stabilized by r-Proteins (Agalarov et al., 2000; Agalarov and Williamson, 2000; Agalarov et al., 1998; Ramaswamy and Woodson, 2009; Samaha et al., 1994). Similar observations were also made for the assembly of the 50S using *in vitro* transcribed 23S rRNA, that becomes effective in presence of certain osmolytes (Semrad and Green, 2002), suggesting a significant bearing of proteins and ions in solving a classical "RNA folding" problem for this large ribonucleoprotein complex (Herschlag, 1995; Shcherbakova et al., 2008; Tan and Chen, 2011). The notion is also corroborated evolutionarily by idiosyncratic extensions of r-Proteins that "lock-in" the rRNA in productive folds (Klein et al., 2004).

1.4.6 Transiently binding proteins expedite ribosome assembly

Apart from the structural knits provided by the r-Proteins, rRNA folding events are also catalysed by numerous accessory, non-ribosomal proteins called ribosome assembly factors (RAFs). The factors are broadly classified as (i) GTPases, (ii) Chaperones and maturation enzymes and (iii) Helicases (Shajani et al., 2011). These factors catalyse ribosome assembly by avoiding the kinetic traps during rRNA folding (Treiber and Williamson, 1999; Woodson, 2000) that may be inevitable during folding of large RNA molecules like the rRNA. In higher organisms, RAFs have also been known to serve as ribosome quality control checkpoints (Strunk et al., 2012) as they avert premature subunit joining and entry of

premature subunits into translation (Karbstein, 2013). Currently, we know of roughly 200 RAFs for eukaryotic organisms whereas only about 50 for bacteria (Shajani et al., 2011). However, the continuous emergence of new RAFs (Andrade et al., 2018; Gibbs et al., 2017) negates the possibility that the current list is exhaustive.

1.4.7 DExD/H box helicases

The DExD/H box proteins are members of the large family of DEAD-box helicases known to unwind RNA in an ATP-dependent manner (Bleichert and Baserga, 2007; Cordin et al., 2006). These RAFs are expected to direct ribosome assembly by assisting in RNA folding events, association and dissociation of RNA binding proteins and relieve kinetic traps during RNA folding (Martin et al., 2013). Helicase assisted folding seems more pertinent in the case of the larger subunit as five helicases such as DbpA, RhlB, SrmB, DeaD (CsdA), and RhlE are known to drive this process (Iost and Dreyfus, 2006). None of these proteins seem to be essential for growth but their deletion confers significant growth defects accompanied by accumulation of premature 50S particles (Jagessar and Jain, 2010b). Characterization of pre-50S particles accumulated in null mutants of the helicases suggests that they act during early stages of assembly but their exact roles are still to be identified (Charollais et al., 2003a; Peil et al., 2008). Additionally, there seems to be some level of functional interaction between these cofactors as overexpression of certain helicase can relieve or exacerbate the growth defects in null mutants of other helicases (Cartier et al., 2010).

1.4.8 Chaperones and maturation factors

This broad class of RAFs contains several protein chaperones like DnaK, DnaJ, GrpE in late stages of assembly (Alix and Guerin, 1993; El Hage et al., 2001). Null mutants of all three chaperones accumulate precursors of 50S subunit but their actual role in the assembly process is still unknown and debated (Alix and Nierhaus, 2003). A possible role of these chaperones may be to assist protein folding thus playing an indirect role in assembly (Alix and Nierhaus, 2003).

Other RAFs in this class include RbfA, RimM, RimJ, RimP, KsgA and Hfq. RbfA (Ribosome-binding factor A), a KH domain containing, cold shock protein is known to assist in assembly of the 30S subunit and its absence confers significant growth defect especially at colder temperatures (Jones and Inouye, 1996). RbfA growth defects can partially be rescued by overexpression of Era GTPase (Inoue et al., 2003).

RimM, RimP, and RimJ (Ribosome maturation factors) bind to various regions on the 16S rRNA and are expected to expedite the binding of r-Proteins (Lovgren et al., 2004; Nord et al., 2009; Roy-Chaudhuri et al., 2008). Deletion of maturation factors often results in accumulation of pre-30S particles with unprocessed ends (Bylund et al., 1998a).

KsgA a methyltransferase, modifies two highly conserved adjacent adenosines (A1518 and A1519) on the 16S rRNA. Its deletion is known to compromise the fidelity of translation and confer a growth defect that can be aggravated by overexpression of RbfA (Connolly and Culver, 2013). The small RNA binding protein, Hfq, a member of the Lsm family of proteins (Wilusz and Wilusz, 2013) was recently shown to affect ribosome biogenesis and fidelity of protein synthesis (Andrade et al., 2018). It proposed roles include chaperoning rRNA folding or facilitation r-Protein-rRNA binding.

1.4.9 GTPases

Era, RsgA, Der/EngA, ObgE and LepA form a class of conserved switch proteins that may regulate critical assembly events during late-stage assembly of both subunits (Gibbs et al., 2017; Shajani et al., 2011). Era and Der/EngA are essential for survival and mutations in these proteins generate defects in growth and chromosome segregation, earning them a collateral role in cell cycle regulation and chromosome partitioning (Britton et al., 1997; Hwang and Inouye, 2006b). RsgA, a dispensable GTPase, was recently discovered to release RbfA from the 30S subunit (Goto et al., 2011b). ObgE is another universally conserved protein that is essential for efficient assembly of 50S subunit, proper chromosome partitioning and its overexpression imparts growth advantage under stress (Feng et al., 2014; Verstraeten et al., 2015).

1.5 The evolving understanding of assembly: 30S subunit as a case study

Historically, RAFs have been spotted as multi-copy suppressors of ribosome related genetic defects (Lu and Inouye, 1998; Nishi et al., 1988; Toone et al., 1991) but later genetic perturbations and sub-optimal growth conditions were used to populate assembly intermediates. This pattern of fortuitous, trial-and-error based discovery coupled with an apparent functional redundancy between RAFs has made the progress in this field relatively slow. When the author embarked on this project, understanding in the field came from growth-related studies, analysis of premature subunits by rate-zonal centrifugation (Mehta et al., 2012), chemical probing experiments (Powers and Noller, 1995), low resolution Cryo-EM derived RAF bound or fitted ribosome structures (Sharma et al., 2005) and intensive quantitative mass spectrometry measurements that were just taking off (Talkington et al., 2005). Further, kinetic studies using time-resolved X-ray foot-printing and mass spectrometry of *in vitro* assembled particles established the existence of multiple rate-limiting pathways for assembly of 30S subunits. The studies also underscored the rates of independent rRNA folding events and r-Protein binding that suggested eventual conversion of multiple assembly intermediates into a mature 30S subunits (Fig. 1.7) (Sykes and Williamson, 2009).

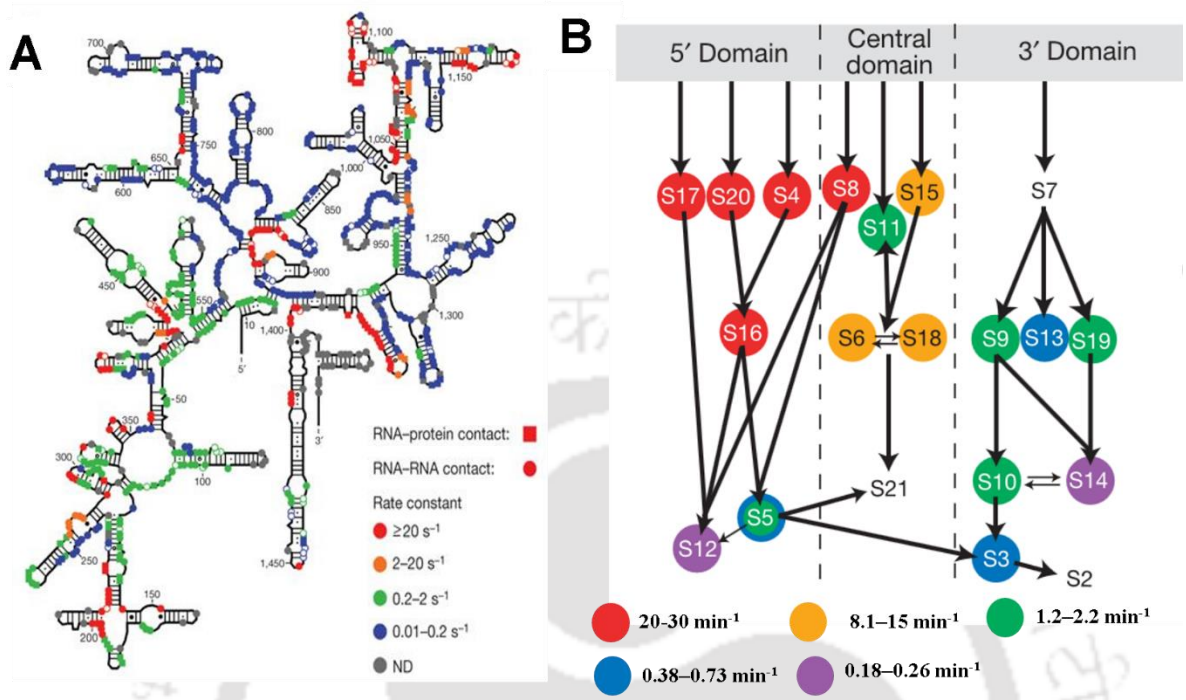

Figure 1.7: Kinetics of rRNA folding and r-Protein binding

(A) A representation of the secondary structure of the 16S rRNA with color markings indicating the folding rate constants of different regions of the RNA as calculated by Adilakshmi and co-workers* (Adilakshmi et al., 2008).

(B) A redefined Nomura map with kinetic measurements for nucleation of r-Proteins on the rRNA as measured by Talkington and co-workers ** (Talkington et al., 2005). Ribosomal proteins are shown in filled circles in varied colours indicating their rate of binding to rRNA.

*Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature, *Concurrent nucleation of 16S folding and induced fit in 30S ribosome Assembly*, Tadepalli Adilakshmi, Deepti L. Bellur, Sarah A. Woodson, ©, 2008

** Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature, *An assembly landscape for the 30S ribosomal subunit*, Megan W. T. Talkington, Gary Siuzdak, James R. Williamson, ©, 2005

In order to understand the contribution of RAFs, attempts were made to characterize *in vivo* assembly intermediates accumulating in absence of RAFs. Interestingly enough, assembly intermediates characterized by Cryo-EM and Mass spectrometry pointed towards the prevalence of heterogeneous pre-30S particles with similar composition and architecture across many RAFs. Additionally, it was also observed that even in absence of the respective

RAFs, mature and near mature ribosomal subunits were formed (Leong et al., 2013; Sashital et al., 2014).

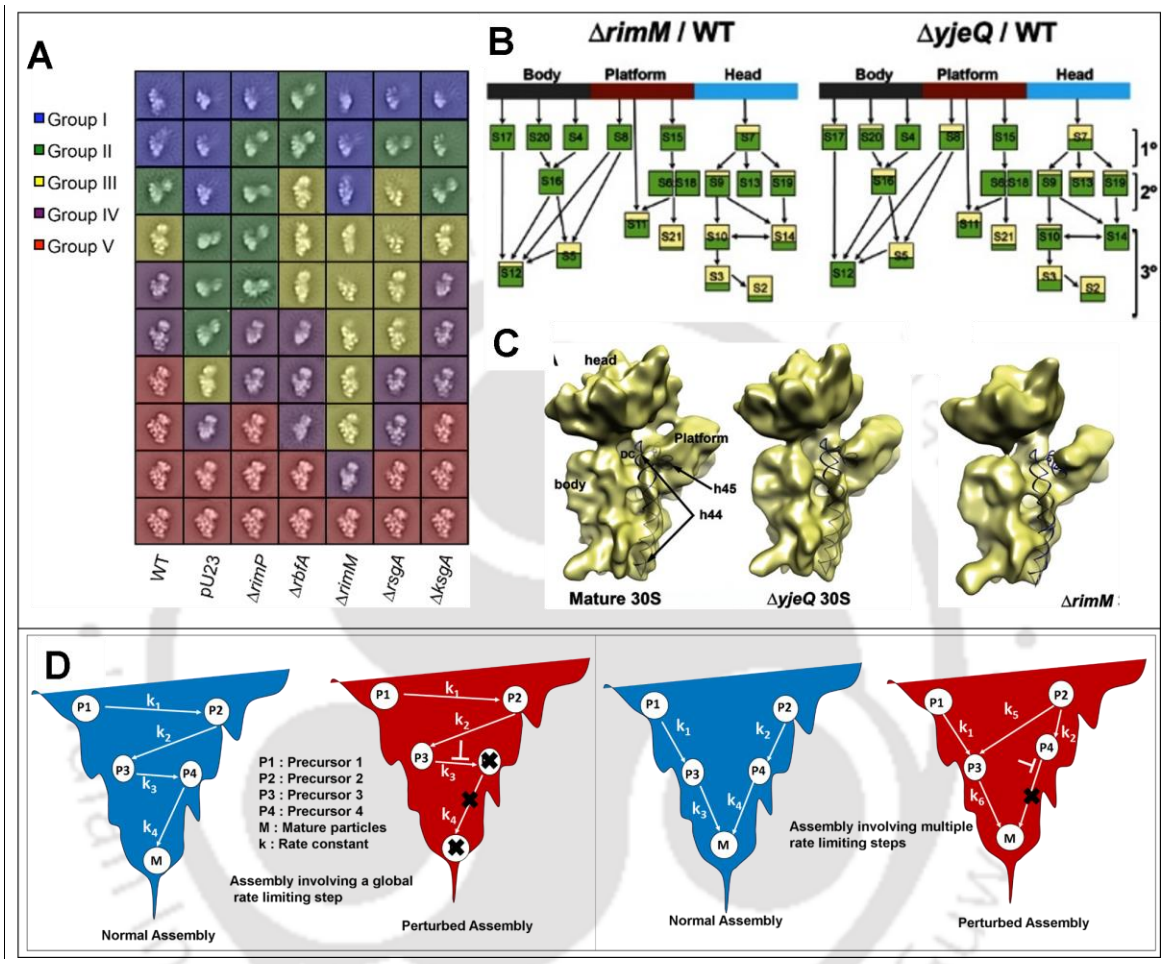

Figure 1. 8: Perturbed assembly generates similar assembly intermediates

(A) Cryo-EM class average structures of 30S assembly intermediates from *E.coli* null mutants for the indicated assembly factors and cells carrying a plasmid encoding a mutated *rrn* operon (pU23) as observed by Sashital et al. (Sashital et al., 2014)

(B) Results of mass spectrometric analysis of 30S intermediates isolated from *ΔrimM* and *ΔyjeQ* shown as a Nomura map as observed by Leong and co-workers* (Leong et al., 2013). The proportion of the yellow colour indicated the underrepresentation of the respective r-Protein.

(C) Cryo-EM structures of 30S particles isolated by Leong and co-workers* (Leong et al., 2013) from *ΔrimM* and *ΔyjeQ*. Structure for the Wt 30S subunit is also shown for reference.

*Images A, B & C were adapted from the cited literature and used according to the terms of a Creative Commons License.

(D) A schematic representation of two possible energy landscapes of ribosome assembly. A normal assembly landscape is shown in blue while perturbed assembly due to RAF deletion is shown in red. Ribosomal precursors are shown as P1-P4 while mature ribosomes are shown as "M". Perturbations in the assembly are marked above the arrows that indicate the direction of the reaction and black crosses indicate failure to form the respective intermediates due to perturbations. In the first scenario (shown on left), Ribosome precursors P1-P4, converge to form mature ribosomes indicated as "M" at the end of the energy landscape. Energy landscapes involving a single maturation pathway with consecutive reactions having rates k_1-k_4 is susceptible to disruption and failure to produce mature ribosomes. On the contrary, landscapes with parallel independent assembly pathways (right, rates shown as k_1-k_6) allow for rerouting assembly intermediates to alternative pathways thus imparting robustness to the assembly process.

These findings also suggested that RAF assisted folding also followed multiple pathways towards maturation and raise important concerns about deriving the rate of action of RAFs solely by analysing the composition and shape of the assembly intermediates. The alignment between RAF dependent and independent modes of assembly reiterates the seminal findings of Nomura and Nierhaus that the r-Proteins and rRNA intrinsically possess all that are needed to make functional ribosomes and also extends to predict how this may happen. Driving assembly through multiple independent pathways without a single global rate-limiting step (Fig.1.8D) allows cells to manoeuvre assembly in perturbed conditions by diverting the flux of the perturbed pathway to alternative, functional pathways. This, however, would lead to the accumulation of intermediates for the rerouted pathways leading to an apparent decrease in assembly rate, but still enables cells to make functional ribosomes thus providing a significant metabolic advantage. The appearance of similar ribosomal precursors, as well as mature ribosomes in RAF deficient strains, supports this theory as RAFs may assist in driving these seemingly redundant pathways that eventually lead to maturation. Support for this theory also comes from the old and recent literatures that support a critical role of the translation machinery in assisting assembly. Previous reports by Srivastava and co-workers suggested maturation of 23S rRNA in polysomes (Srivastava and Schlessinger, 1988). Similar findings by Shetty and co-workers (Shetty and Varshney, 2016b) suggested that maturation of 16S rRNA was closely coupled with $tRNA^{i}_{fmet}$ binding i.e. in the initial stages of translation

initiation. These findings create a possible unearthed link between the assembly and subsequent fate of ribosomes in cells. That said, our understanding of how this is achieved is still far from complete and requires revisiting the field with new strategies and approaches. Attempts have been made to measure the kinetics of RAF mediated remodelling of 30S precursors but, this understanding is still incomplete (Thurlow et al., 2016). Recent strides in the area of quantitive measurement of gene expression using ribo-seq and Mass spectrometry have underlined a significant distortion between stochiometry of RAFs and total ribosome content (Occupancy of RAF was as low as 5% on purified ribosomes as well as cell lysates) (Li et al., 2014; Thurlow et al., 2016). These findings although enlightening, raise serious concerns about the thoroughness of our knowledge about RAFs and their mechanisms of catalysis. Additionally, the recent realization that certain proteins such as Hfq and EF-4/LepA, which are earlier known to play a role in sRNA maturation and translation, respectively, were found to contribute to the assembly by unknown mechanisms only consolidates this belief (Andrade et al., 2018; Gibbs et al., 2017) and calls for advanced methodologies to characterize existing as well as novel RAFs. Adding to this, the emerging understanding of close intertwining between translation machinery and ribosome assembly raises exciting questions about regulation of protein synthesis rates and quality. The knowledge of these mechanisms in bacteria would have far-reaching consequences on dissecting the evolution of even more complicated assembly and translation regulation mechanisms seen in higher organisms.

1.6 Definition of the problem

A wide variety of assembly factors chaperone rRNA folding, chemically modify rRNA or signal the completion of a molecular event during assembly. Therefore, based on the functional roles they play during assembly, they are broadly classified into:

1. DEAD-box helicases
2. Chaperones and Maturation enzymes
3. GTPases

Although, studies on ribosome biogenesis date as old as the discovery of ribosomes itself, it is comparatively recent that the role of assembly factors has come to the fore. The discovery of assembly factors has been sporadic and fuelled by serendipity. In bacteria, the systematic identification of assembly factors is stymied by pleiotropy associated with loss of functional phenotype. Additionally, the methods employed for monitoring ribosome assembly have not been amenable to large-scale investigations mainly owing to operose cycles of analysis involving rounds of centrifugation and biochemical characterization of assembly intermediates. In order to address the aforementioned shortcomings, the current work aims to systematically identify and characterize hitherto poorly defined putative assembly factors by devising novel strategies and approaches.

1.7 Objectives of the study

- Development and validation of novel methodologies to identify RAF candidates.
- Development and application of a scalable assay system to validate RAF candidates.
- Understanding RAF mediated ribosome quality control mechanisms in bacteria and its implications on protein synthesis.

Chapter II

Identification of candidate ribosome assembly factors using BioID

Chapter 2

2.1 Introduction

The timeline of milestones achieved in understanding ribosome assembly events spans nearly six decades (Andrade et al., 2018; Dohme and Nierhaus, 1976; Gibbs et al., 2017; Lewandowski and Brownstein, 1966; Nierhaus, 1991; Shajani et al., 2011; Traub and Nomura, 1968b, 1969). The breakthrough experiments by Nomura and Nierhaus (Dohme and Nierhaus, 1976; Traub and Nomura, 1968b) established that the ribosomes contain within themselves all that is needed for their assembly *in vitro*. However, the second striking conclusion arising from this study was a set of non-physiological conditions such as high temperature and high salt concentration needed to assemble ribosomes *in vitro*. This highlighted the stark contrast between the intracellular and extracellular requirements of ribosome assembly. Further investigations strengthened the central role of RAFs in driving ribosome assembly (Formenoy et al., 1994; Fuller-Pace et al., 1993; Gustafsson and Persson, 1998; Nikolaev et al., 1973; Sirdeshmukh and Schlessinger, 1985; Toone et al., 1991). However, efforts towards mapping the assembly landscape of bacterial ribosomes have been moving at a snail's pace as identification of RAFs has been fuelled – to a large extent – purely by serendipity. Hitherto, the conventional methodologies involved the pipeline that begins with the deletion of prospective RAF candidates and the subsequent following of the phenotypes, this produces in the aftermath. Typically, this includes screening for growth defects, presence of premature rRNA and an altered ribosome profile following sedimentation analysis (Hayes and Hayes, 1971; Lindahl, 1973b). Incidentally, deletion of vital genes in bacteria often confers pleiotropy (Lerner and Inouye, 1991) that further clouds the landscape of ribosome assembly events. These limitations coupled with the lack of systematic screening strategies have slackened the search and characterization of RAFs.

In order to advance the current understanding of bacterial ribosome biogenesis, we concerted our efforts towards developing a robust, high-throughput approach for identification of RAF candidates. We reasoned that as RAFs interact transiently with ribosome intermediates during assembly, capturing the **P**rotein-**P**rotein **I**nteraction (PPI) events could provide a snapshot of assembly related events. Elucidating these interaction networks can help

pin down the molecular partners and events driving the assembly process. Conventional PPI discovery tools include the yeast two-hybrid (Y2H) system (Brückner et al., 2009) and **A**ffinity **P**urification coupled **M**ass **S**pectrometry strategies (AP-MS) like Tandem affinity peptide (TAP) (Rigaut et al., 1999) and Sequential peptide affinity (SPA) (Zeghouf et al., 2004). The Y2H system lacks a fundamental ability to fish out interactions in a truly physiological environment and has a limited scalability. On the other hand, the AP-MS methodologies require a cross-linking step to monitor any short-lived interactions (Gingras et al., 2007).

An alternate methodology that was devised in the recent past to probe interactions of macromolecular complexes and study subcellular protein localization is referred as Proximity-dependent **Bio**tin **Id**entification (BioID) (Liu et al., 2018; Roux et al., 2012). The technique exploits the promiscuous, proximity dependent biotinylation activity of an R118G mutant of Biotin ligase (BirA*) from *E. coli* (Choi-Rhee et al., 2004). The Wild-type protein, BirA forms an ATP-Biotin adduct (bioAMP) which is transferred specifically to a lysine residue on the Biotin Acceptor Peptide (BAP) of the Biotin Carboxyl Carrier Protein (BccP). However, the R118G mutation decreases the affinity of BirA* towards the bioAMP adduct (Fig. 2.1A), resulting in its premature release from the active site. The diffused reactive species readily biotinylates proximal proteins within a radius of ~10 nm (Kim et al., 2014). Once biotinylated, the modified protein can be captured later under denaturing conditions by exploiting the extraordinarily specific and strong streptavidin-biotin interaction (Weber et al., 1989). This also makes BioID the method of choice for studying interactions between membrane proteins that often display limited solubility (Holthusen et al., 2015; Mehus et al., 2016; Schnider et al., 2018). Relatively simple, BioID involves the creation of a BioID bait by fusing the protein of interest to BirA*. The bait once expressed in its cellular environment, modifies vicinal proteins, which can be affinity purified and later identified by mass spectrometry (Fig. 2.1B). The technique has now evolved as a discovery tool to identify novel PPIs in a reliable high-throughput manner (Kim et al., 2014; Roux et al., 2018; Roux et al., 2012) in eukaryotic systems but its application in prokaryotic systems has not been as widespread.

In this chapter, we have attempted to develop and apply BioID as a discovery tool to identify prospective RAFs with a long-term objective of dissecting the events leading to the

assembly of ribosomes. Towards this, we have focused our efforts on the 30S subunit. Our initial studies focused on optimizing and validating the application of BioID in the context of 30S assembly by overexpressing the BioID bait. Upon validation of the design principles, we utilized the technology to identify the interactome of the widely conserved assembly factor Era and also counter-probed its contacts with the vital head-platform cleft of the 30S subunit. Overall, this study is a proof-of-concept of the first step in our efforts aimed at the systematic identification of RAFs in bacteria.

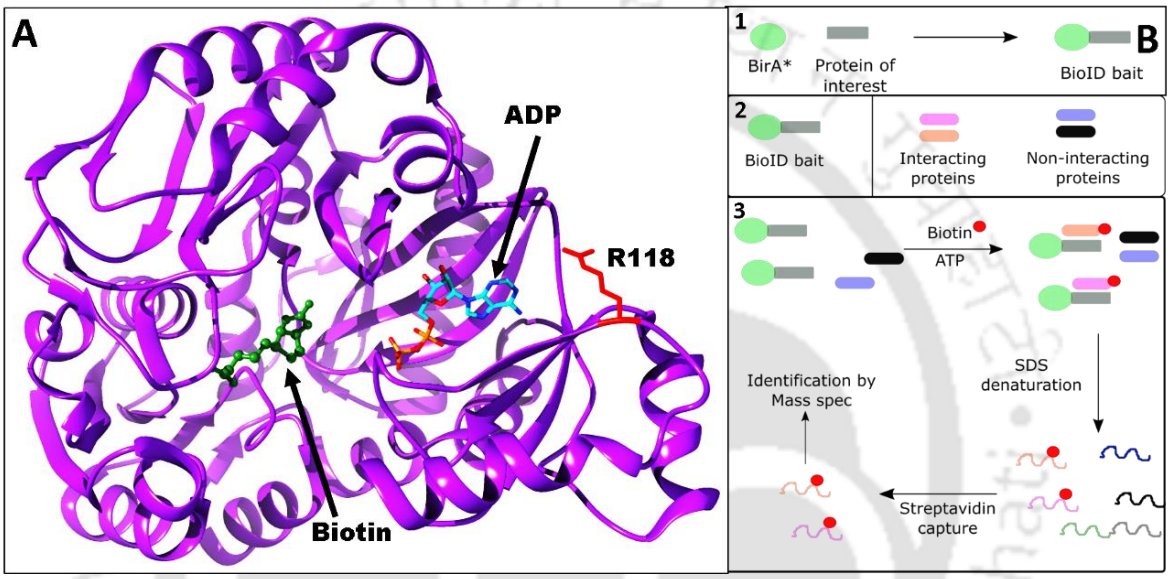

Figure 2.1: BioID exploits the proximity-dependent biotinylation by BirA*

(A) The molecular structure of BirA from *E. coli* in complex with biotin (green) and ADP (blue-orange sticks). The Arginine residue at position 118 is also highlighted in red. Mutation of R118 to G converts the protein to a promiscuously biotinylating enzyme BirA*.

(B) An outline of the BioID methodology to identify PPIs. A BioID bait is generated by fusing the protein of interest to BirA* followed by introduction of the bait in its physiological condition. This is followed by proximity dependent biotinylation, purification and identification of the target proteins by mass spectrometry.

2.2 Materials & Methods

2.2.1 Cloning, expression and purification

The gene encoding Biotin ligase (*birA*) was amplified from E. *coli* K-12 MG1655 (Wt) genomic DNA using gene specific primers with Pfu DNA polymerase (Fermentas). The gene was cloned into vector 1B using Ligation independent cloning (LIC). The vector was used as a template for creating the R118G point mutant of *birA* (*birA**) using mega primers based PCR method. The *birA** gene was fused to 5' end of the gene encoding the r-Protein uS13 (*rpsM)* by overlap extension PCR to create the gene *b*s13.* The construct was subsequently cloned into 1B to create the vector pB*S13. This placed the *b*s13* under a T7 based, IPTG inducible promoter to encode the B*S13 protein with an N-terminal His-Tag. The cloned constructs were verified by sequencing.

E. coli BL21 (DE3) cells were transformed with pB*S13 and were grown until OD_{600} reached 0.6 in LB medium supplemented with 2 µM D-biotin at 37 ℃, 180 rpm. Expression was induced with 0.3 mM IPTG and growth was continued for another 3 hours (hrs) at 37 °C with shaking at 180 rpm. Cells were harvested by centrifugation at 10,000xg for 10 minutes (mins) and pellets were washed followed by lysis using sonication in Lysis buffer (20 mM Tris-Cl (pH-8.0 at 25 °C), 300 mM NaCl and 6 mM β-mercaptoethanol). The fusion protein was found to express in the soluble fraction of the cell lysate.

For purification, expression was scaled up to 400 ml culture volume followed by **I**mmobilized **M**etal **A**ffinity **C**hromatography (IMAC). Briefly, harvested cells were washed and resuspended in 20 mL Lysis buffer and subsequently, cells were lysed using sonication. Cell lysate was clarified by centrifugation at 25,000xg for 30 mins and loaded onto a 5 ml HiTrap IMAC HP column (GE healthcare). Post binding, the column was washed with 10 column volumes (CV) of Wash buffer (Lysis buffer supplemented with 50 mM Imidazole) following which the protein was eluted using a linear gradient of imidazole concentration using Lysis buffer and Elution buffer (Lysis buffer supplemented with 500 mM imidazole). Affinity purified samples were checked for purity using SDS-PAGE followed by dialysis to remove imidazole and reduce the NaCl concentration to 150 mM. Subsequently, small aliquots of the proteins were flash frozen and stored at -80 °C until required.

2.2.2 Assay for *in vitro* biotinylation

In vitro biotinylation assay was performed using BSA as a substrate protein. The reaction contained 2 mM substrate suspended in the reaction buffer (40 mM Tris-Cl (pH-8.0 at 25 ℃), 5.5 mM $MgCl_2$, 100 mM KCl, 1.4 mM β-ME, 3 mM ATP & 5 µM D-Biotin). The reaction was supplemented with 10 µM B*S13 protein and incubated at 37 ˚C for 6 hrs. After 6 hrs, the reaction was dialyzed against reaction buffer deviod of ATP and D-biotin. The dialyzed reaction was then incubated with Streptavidin sepharose beads (GE Healthcare) overnight on mild shaking to capture the biotinylated proteins. Subsequently, the beads were recovered by centrifugation and washed by the following strategy:

Washing strategy:

(A) Two washes with Solution 1 (2 % SDS in deionized H_2O)

(B) One wash with Buffer 1 (1 % Triton X-100, 300 mM NaCl, 1 mM EDTA and 50 mM HEPES (pH 7.5 at 25 ℃))

(C) One wash with Solution 2 (250 mM LiCl and 0.5 mM NP-40)

(D) One wash with Buffer 2 (50 mM Tris-Cl (pH 7.4 at 25 ℃), 300 mM NaCl, 1 mM DTT and 0.05 % Tween-20)

Steps A-D were performed at room temperature and the beads were recovered by centrifugation between each wash step. Washes A-C were done with 5 bead volumes of solutions for 8 mins & wash D was done with 250 bead volumes for 2 hrs. Post washing the beads were recovered and resuspended in Buffer 3 (50 mM Tris-Cl (pH 7.4 at 25 ℃), 300 mM NaCl and 1 mM DTT) and analyzed using SDS-PAGE.

2.2.3 Recombineering for BirA* knock-in into *E. coli* genome

The Wt strain was used as a parent strain for all genetic manipulation experiments using λ- red recombination method (Datsenko and Wanner, 2000c). The *birA** gene was fused to a *cat* gene encoding chloramphenicol acetyl transferase (CAT) by overlap extension PCR to create the *birA*-cat* gene. Flanking regions (50 bases each) specific to 3' end of *era* or *rpsR*

were added to the *birA**-*cat* using PCR. The construct was generated so as to encode Era and bS18 in frame with BirA*. The amplified product was electroporated into *E. coli* cells harbouring vector pkD46. The colonies were screened for the in-frame integration of the *birA**-*cat* into the locus of *era* and *rpsR* using colony PCR, thus creating the strains Era.BirA* and sR.BirA*.

2.2.4 Ribosome pulldown using IMAC

E. coli BL21(DE3) cells transformed with pB*S13 were grown in LB medium supplemented with 2 µM D-Biotin till OD_{600} reached 0.6. The culture was induced with 0.3 mM IPTG and allowed to grow for another 3 hrs. Cells were harvested by centrifugation at 8,000xg for 10 mins. Cells were washed and resuspended in Buffer R (20 mM Tris-Cl (pH 7.4 at 25 °C), 150 mM NH_4Cl, 10.5 mM $MgCl_2$, 0.5 mM EDTA and 6 mM β-ME) supplemented with Lysozyme (0.5 mg/ml), DNase (10 µg/ml), Protease inhibitor 1 mM and lysed using French-press at 20 kpsi. Cell lysate was clarified by centrifugation at 25,000xg for 30 mins and loaded onto a 5 ml HiTrap IMAC HP column. Post binding, column was washed with 10 CV of Buffer W (Buffer R supplemented with 5 mM Imidazole). The protein was eluted using a linear gradient of imidazole concentration using Buffer R and Buffer E (Buffer R supplemented with 500 mM Imidazole). Affinity purified samples were instantly dialyzed against Buffer R with a buffer change every 15 mins for four rounds. The dialyzed samples were divided into aliquots, flash frozen and stored at -80 °C until further use. The dialyzed samples were analysed using SDS-PAGE for their protein content. In parallel, RNA was isolated using phenol-chloroform method followed by ethanol precipitation. The purified RNA was analysed using agarose gel electrophoresis.

2.2.5 *In vivo* biotinylation and purification of biotinylated proteins

Wt**,** Era.BirA* or sR.BirA* cells were grown in 400 ml LB medium supplemented with 1 mM D-Biotin until OD_{600} reached 0.6. Cells were harvested by centrifugation at 8,000xg for 10 mins. Cells were washed and resuspended in Buffer V (50 mM Tris-Cl (pH:7.4 at 25 °C), 300 mM NaCl and 1 mM DTT) supplemented with 2.5 mM phenylmethanesulfonyl fluoride (PMSF), DNase (10 µg/ml) and lysed with Frech press with 20 kpsi. The clarified

cell lysate was incubated with streptavidin sepharose beads overnight on gentle shaking.The beads were recovered and washed as described in the section 2.2.2. Purified fractions were later analyzed by SDS-PAGE.

2.2.6 Mass spectrometric analysis

Purified protein-bound beads that were resuspended in SDS-gel loading buffer were heated to 95 ℃ for 15 mins and loaded onto 15% SDS-PAGE. The sample was allowed to run just enough to enter the resolving gel after which the run was stopped and the gel was stained with Coomassie Brilliant Blue. The gel was placed on a clean platform in a laminar hood and the stained region of the gel was sliced out using a clean scalpel. The gel pieces were added in fresh mass spec compatible tubes and supplemented with 200 µL of deionized water. The Mass spectrometric analysis was outsourced to V proteomics, New Delhi for the identification of purified proteins. The excised band was used for in-gel trypsin digestion, followed by C18 cleaning and analysed by nano-ESI LC-MS/MS using the QExactive platform. The database search was performed with SEQUEST software with 1% FDR for identified proteins.

2.3 Results

2.3.1 Engineering the BioID bait to probe ribosome assembly

In order to establish the BioID methodology in bacteria, we first set out to test if BioID worked in the bacterial physiological setting. We reasoned that in order to trace PPI events on the ribosome, the BioID bait should integrate into the ribosome while retaining its biotinylation activity. Further, the integration should not affect subunit joining as well as ribosome function in any respect. Hence, we analysed the structure of the 30S and its assembly landscape to screen for an appropriate BioID bait candidate. The following stringent set of criteria was framed to screen for a viable bait protein: (i) The protein should not be an essential r-Protein; (ii) It should integrate on the ribosome in a stable manner; (iii) The candidate should not affect the nucleation of other r-Proteins; (iv) The candidate should have a free N- or C-terminus that is amenable to tagging.

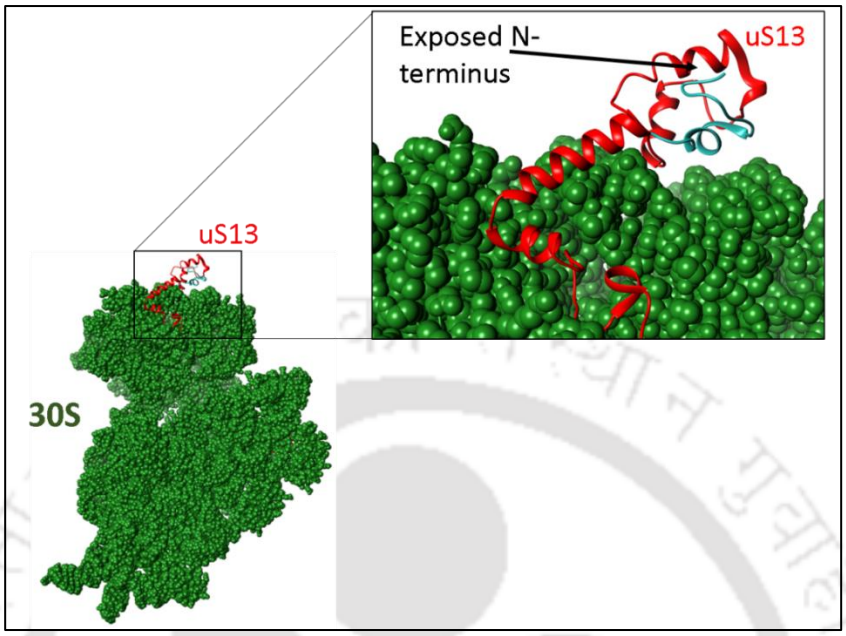

Figure 2.2: Selection of the BioID candidate

The solvent occluded side of the 30S subunit is shown in green spheres. Protein uS13 is shown in red ribbons in the 30S-head region. A zoomed in view of the head region showing the exposed N-terminus of the uS13 highlighted in blue.

Upon thorough structural analysis, we found the r-Proteins uS11, uS13 and bS18 to fulfil most selection parameters. However, the N-terminus of the r-Protein uS11 was found to approach the subunit interface raising the possibility of interfering with subunit association if tagged with BirA*. Thus, we utilized uS13 to create the BioID bait (Fig. 2.2).

In order to create the bait, we first engineered the *E. coli* Biotin ligase (BirA) to create a R118G mutant (BirA*) using site directed mutagenesis. A fusion construct containing the *rpsM* gene, a region encoding the linker Gly-Gly-Gly-Ser (GGGS) and *birA** was created (Fig. 2.3A). The linker region was inserted to avoid stearic hindrance once the bait integrated onto ribosomes. Additionally, introduction of the linker would help in independent folding of the two proteins. Subsequently, the fusion construct was cloned into an IPTG inducible expression system in vector 1B (Fig. 2.3B). The orientation of the insert was confirmed by restriction digestion and sequencing. The cloned construct expressed the BioID bait in the soluble fraction, which was later purified by IMAC (Fig. 2.3C).

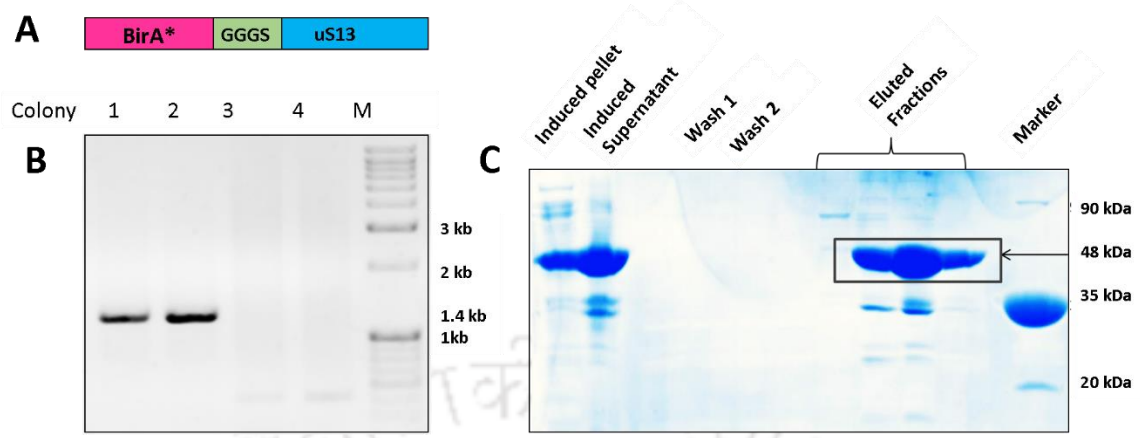

Figure 2.3: Engineering and purification of the BioID bait

(A) A schematic representation of the BioID bait B*S13.
(B) Colony PCR to confirm the insertion of *b*s13* into vector 1B (to create pB*S13), using vector specific primers. The screened colonies are indicated on the top and DNA markers are loaded in the lane 'M'. Size of the marker is indicated on the right. Expected size of the *b*s13* fusion construct (1.4 kb) is also indicated
(C) An SDS-PAGE analysis to test the purification of B*S13 by IMAC. The induced supernatant and pellet fractions are loaded along with the wash and eluted fractions. Protein size standard is loaded in the lane 'Marker' and the respective molecular weights of standard proteins are shown. The eluted factions of B*S13 are seen to migrate at the expected size of 48 kDa (indicated by a box).

2.3.2 *In vitro* testing of the BioID bait

The purified bait was checked for its ability to biotinylate vicinal protein in solution. Towards this, we utilized BSA as a substrate protein. We reasoned that active B*S13 should biotinylate vicinal protein *in vitro* by virtue of molecular crowding. Hence, B*S13 was supplemented with ATP, Biotin and BSA in a reaction mixture. The biotinylated proteins were further purified using Streptavidin beads and analysed in SDS-PAGE. The reaction mixture containing B*S13 and BSA showed significant BSA biotinylation indicating proficient activity of the bait (Fig. 2.4). As anticipated, the bait also displayed significant self-biotinylation.

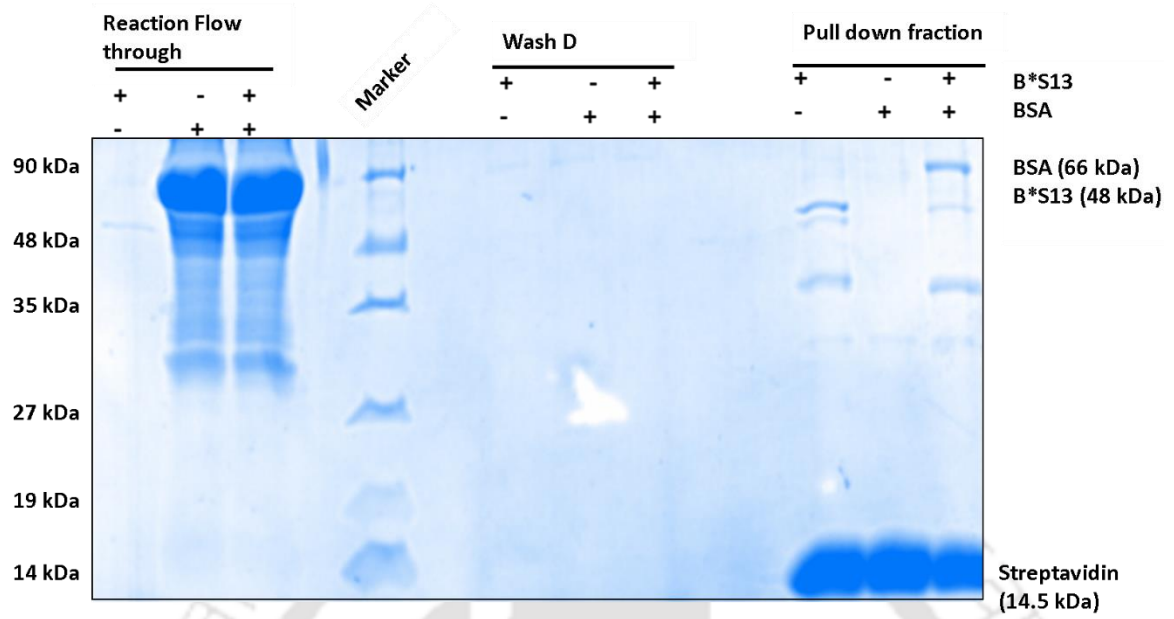

Figure 2.4: B*S13 actively biotinylates proteins in a proximity dependent manner

An SDS-PAGE analysis to test activity of the B*S13 bait. The reaction flow through, Wash flow through from washing step D (as described in section 2.2.2) and Streptavidin bound-pull down fraction are loaded in the respective lanes. A protein standard is loaded in the lane 'Marker' and the respective molecular weights of standard proteins are shown. Biotinylation of the target protein is seen wherever the bait B*S13 is present (in the last three lanes towards the right of the gel).

2.3.3 Integration of the BioID bait into assembling ribosomes

In order to perform BioID in a truly physiological setting it was essential to ensure localization of the bait on ribosomes. To test this, we reasoned that co-purification of the bait with ribosomes would indicate its stable integration into ribosomes. Towards this, we utilized the N-terminal His_6-Tag of B*S13 (Fig. 2.5A). Upon elution, two distinct ribonucleoprotein fractions eluting at 100 mM (Frac 1) and 250 mM (Frac 2) imidazole concentration were obtained (Fig. 2.5B). The fractions contained multiple proteins as would be expected for ribosome containing fractions. To further confirm the presence of ribosomes, we purified RNA from the two fractions (Fig. 2.5C). Both IMAC purified fractions contained significant amounts of 16S rRNA, strongly indicating the presence of 30S subunit. Interestingly, Frac 1 was also found to have minor quantities of 23S rRNA indicating the presence of intact 70S ribosomes.

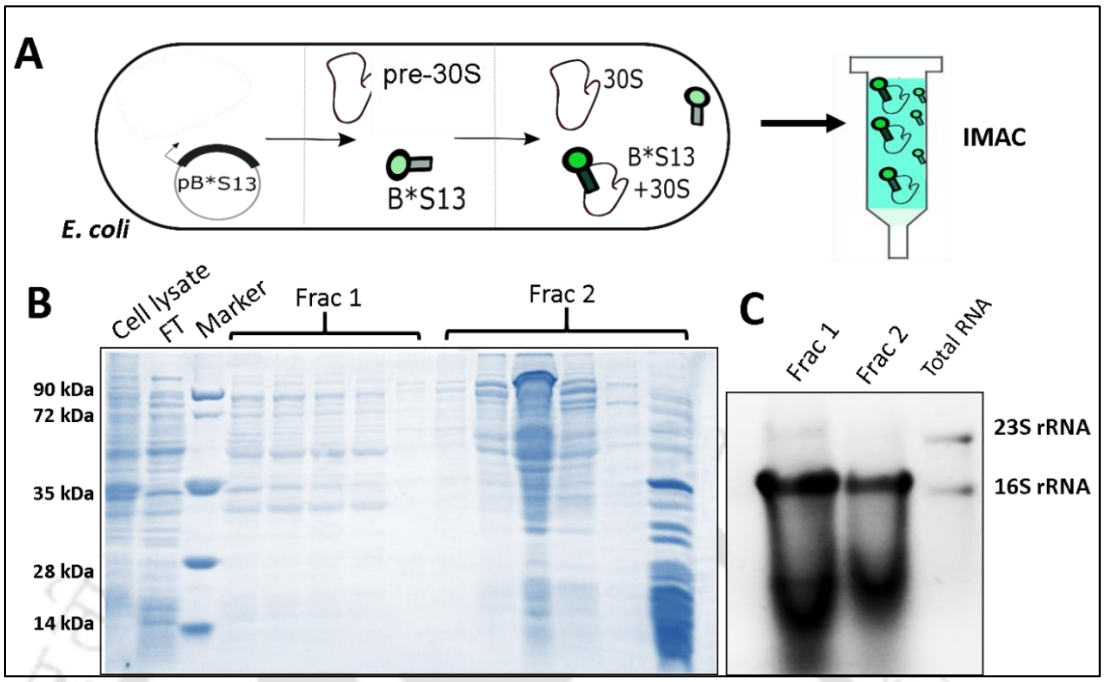

Figure 2.5: BioID bait integrates into ribosomes

(A) A schema to evaluate the integration of BioID bait into ribosomes. B*S13 integrates into assembling 30S subunits which can be purified using the N-terminal His_6-Tag of B*S13 by IMAC.

(B) An SDS-PAGE analysis for the affinity purified fractions using B*S13. Two fractions corresponding to 100 mM imidazole (Frac 1) and 250 mM imidazole (Frac 2) are shown. Clarified cell lysate, column flow through and Marker are loaded in the indicated lanes. Molecular weights of the respective standards are also indicated on the left.

(C) Agarose gel electrophoresis for RNA isolated from affinity purified fractions of B*S13. Total cellular RNA is loaded for size reference along with indications for bands corresponding to 23S rRNA and 16S rRNA.

2.3.4 Engineering of the strain to install the BioID bait on the ribosome

Till this end, our experiments provided a proof-of-concept that the BioID bait could firmly localize on the ribosomes, thus fulfilling an essential condition to capture authentic PPIs. Motivated by these observations, we performed another round of structural analysis of the ribosome to survey the regions that are acted upon by RAFs. Given the much better understanding of the 30S assembly landscape, we concentrated our efforts to dissect 30S late stage assembly events. These events include the formation of the 30S head region and are

accelerated by the concerted action of numerous RAFs such as RbfA, KsgA, Era, RimM and RimP (Shajani et al., 2011). A cleft between the head and the 30S platform region stabilizes the incoming mRNA and also binds to the N-terminal end of Initiation factor-3 (Hussain et al., 2016; Yusupova et al., 2001).

We decided to probe the events leading to the formation of this cleft using BioID. Based on our earlier analysis, bS18 a candidate that matched all of our bait selection criteria (Fig. 2.6A) was chosen for BirA* fusion. We reasoned that since bS18 binds near the cleft region any assembly factors that transiently interact with the head-platform cleft should be biotinylated by the S18-BioID bait. Further, in order to bolster this hypothesis, we also decided to create a RAF-BioID bait that would in turn biotinylate any r-Proteins located in the vicinity of the cleft region. Interestingly, Era an essential GTPase is proposed to bind on the solvent exposed side of the 30S near the 3' minor domain of the 16S rRNA on this cleft region (Sharma et al., 2005) (Fig. 2.6B). Era is also proposed to be involved in regulation of cell cycle events in bacteria (Britton et al., 2002), but its actual function is still elusive. Keeping these observations in mind, bS18 and Era seemed to be ideal candidates to probe important assembly related events as well as to decipher the function of Era. In order to minimize the artefacts that may arise due to overexpression of the BioID baits, we decided to introduce the *birA** gene in-frame at native locus of *era* and *rpsR* by λ-red recombination (Fig. 2.6C & 2.6D). The two strains sR.BirA* & Era.BirA* would produce the BioID bait under the native promoters of *era* and *rpsR*. The sR.BirA* cells were also checked for integration of the bait onto ribosomes. Towards this, ribosomes were purified from sR.BirA* strain and analysed by SDS-PAGE. We found a distinct band indicating the presence of S18-BirA* bait in crude ribosomes purified from sR.BirA* cells (Fig. 2.6E)

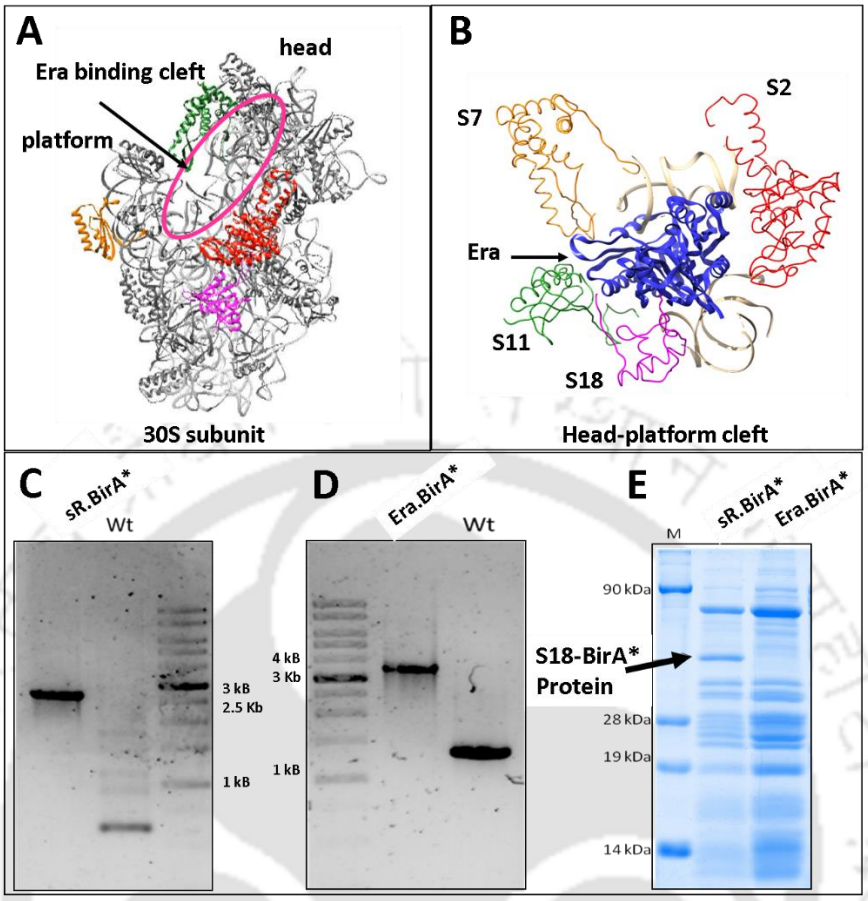

Figure 2.6: Engineering of Era and S18 BioID baits

(A) A representation of the solvent exposed side of the 30S subunit. The 30S subunit is shown in grey ribbons with proteins uS11, uS7, uS2 & bS18 are shown in green, orange, red and pink, respectively. The region between the proteins forms the head-platform cleft and is also the proposed binding site of Era (PDB ID: 4WZO).

(B) A close up view of Era fitted into the head-platform cleft of 30S. Era is represented in blue ribbons (PDB ID: 1X18, 1X1L). The other r-Proteins surrounding the cleft are also marked.

(C) PCR verification for integration of *birA*-cat* at the 3' end of *rpsR*. Locus flanking primers specific to *rpsR* were used to highlight the difference in the size of Wt locus (0.8 kB) and the engineered locus (2.6 kB).

(D) PCR verificaion for integration of *birA*-cat* at the 3' end of *era* . Locus flanking primers specific to *era* were used to highlight the difference in the size of Wt locus (1.4 kB) and the engineered locus (3.5 kB).

(E) An SDS-PAGE analysis to check for the presence of S18-BirA* in ribosomes isolated from sR.BirA* cells. Crude ribosomes purified from sR.BirA* and Era.BirA* were loaded in the respective lanes. The protein marker is loaded in the lane 'M' and the sizes are indicated. A distinct band of roughly 42 kDa representing the S18-BirA* was observed only in sR.BirA* derived ribosomes (indicated by an arrow).

2.3.5 BioID probes the interaction between Era and bS18

Excited by these observations, we also wanted to further confirm if the baits retained their promiscuous biotinylation activity. For this, we purified cell lysates from Wt, sR.BirA* & Era.BirA* cells and probed for presence of biotin modification on cellular proteins using anti-biotin antibody. The cell lysates from Wt cells failed to generate any signal upon immunoblotting (Fig. 2.7A) indicating the absence of any biotinylated proteins. However, both sR.BirA* & Era.BirA* showed significantly higher signals indicating the presence of many biotinylated proteins thus confirming the promiscuous biotinylaion activity of the respective baits. We were also curious to see if the localization of bS18 and Era on ribosomes would lead to biotinylation of r-Proteins as per our hypothesis. In order to address this, we probed for the presence of biotin modification in r-Proteins derived from crude ribosomes of sR.BirA* & Era.BirA*. The immunoblotting experiments clearly indicated presence of biotin modification on numerous r-Proteins in ribosomes from sR.BirA*. However, the number of modified r-Proteins was much lesser in case of Era.BirA* (Fig 2.7B).

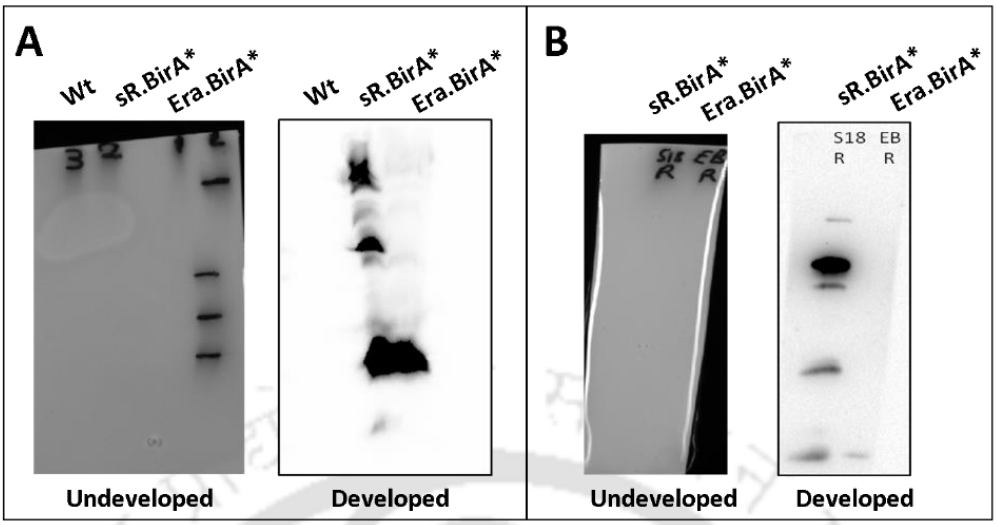

Figure 2.7: Chromosomally integrated baits retain their biotinylation activity

(A) An immunoblot with cell lysates from Wt, sR.BirA* and Era.BirA* cells. The undeveloped blot is shown on the left with the respective lanes marked on the top. The blots were probed with an anti-Biotin antibody, which was visualized by HRP-generated chemiluminescence (right).

(B) An immunoblot for crude ribosomes from sR.BirA* and Era.BirA* cells. The undeveloped blot is shown on the left with the respective lanes marked on the top. The blots were probed with an anti-Biotin antibody, which was visualized by HRP-generated chemiluminescence (right).

This differential pattern of biotinylation is in line with our proposed hypothesis as bS18 bait is anticipated to localize on the ribosomes for much longer period of time in contrast to transient interactions between Era and ribosomes. These studies also bolster the credence of the approach, as it is evident from the differential biotinylation that the baits do not simply generate a population of biotinylated proteins but tag only the relevant proximal proteins.

Motivated by these observations, we decided to characterize these PPIs by mass spectrometry. We purified cells lysates from sR.BirA* & Era.BirA* cells grown in presence of 1 mM D-biotin. Biotinylated proteins were purified using streptavidin beads and identified by mass spectrometry. Mass spectrometry derived interactome of Era and bS18 contained 449 and 495 proteins, respectively (Fig. 2.8). Keeping in mind that enrichment by biotinylation may not follow a linear relationship with number of interaction events, any proteins having a Mass spectrometric score > 0 were considered as potential interaction partners. The interaction partners were classified into 7 broad functional classes, namely: (i) Cellular metabolism and

transport, (ii) Cell division and shape determination, (iii) Nucleic acid synthesis and regulation, (iv) Translation regulation, (v) Ribosomal proteins, (vi) Ribosome assembly factors, and (vii) Uncharacterized.

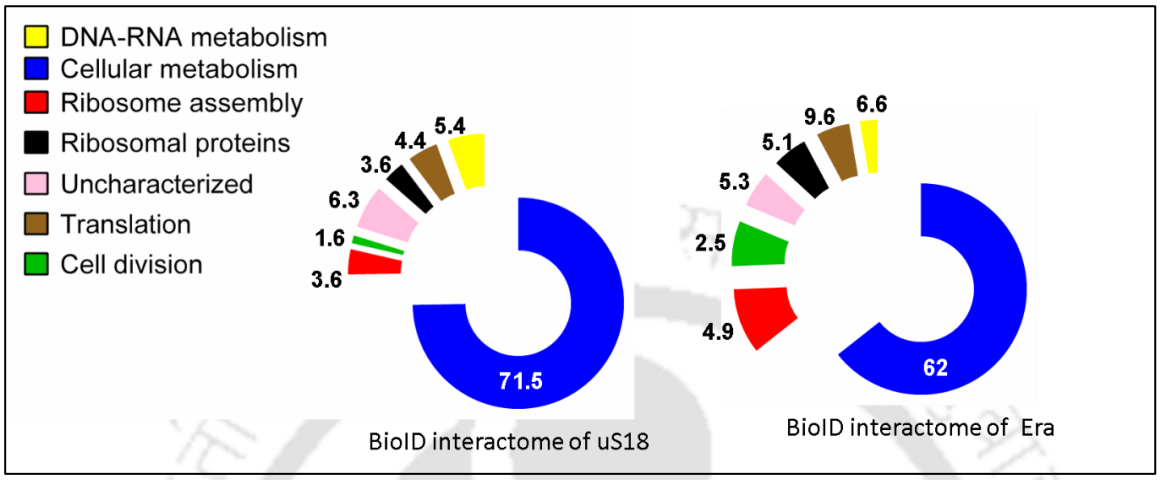

Figure 2.8: The BioID interactome of bS18 and Era

A distribution representing the interaction partners of bS18 and Era as captured by BioID. The classifications were made on basis of known or predicted function of interacting proteins. The percentage of each representative class is indicated in each section of the pie-chart.

One of the most important highlights of the interactomes was that the mutual biotinylation event between bS18 and Era were captured with high confidence scores, clearly indicating the vicinity dependent biotinylation on the ribosome. The interactome of Era also comprised of r-Proteins uS1, uS2, uS3, uS5 and uS7 with high confidence scores indicating its localization within the cleft region. Additionally, Era was also seen to interact with other RAFs like Der, YchF, BipA, essential Initiation factor-2 and transcription termination factor Rho. These results indicate that the BioID bait does indeed capture molecular contacts with precision. The interactome also contained some uncharacterized proteins and the underlying molecular basis of these novel interactions requires further validation and can serve as starting point for identifying new RAFs.

2.4 Discussion

Here, we provide a proof-of–concept establishing the application of a high-throughput assay system to identify transient PPIs of the ribosome that can be extended to identify RAFs. Relatively simple, the methodology makes use of the proximity dependent promiscuous biotinylaiton activity of BirA*. The only requirements that seem essential for application of BioID in context of ribosome assembly would be a careful choice of the bait candidate coupled with expression of the BioID bait preferably at true physiological concentrations of the candidate protein. Although widely applied in multicellular and unicellular eukaryotic organisms, this study, to our knowledge is the first application of BioID to a bacterial system. The limited application of BioID to the prokaryotic systems may be attributed to the lack of cellular compartmentalization in bacteria, which further randomizes the localization of a single protein in the bacterial cell. This important lacuna must be considered while designing BioID experiments. In the context of ribosomes, we have tackled this problem by thoroughly verifying the localization of the respective baits on the ribosome.

The BioID bait once installed on the ribosomes would record the PPIs of the respective r-Protein or RAF, thus giving a glimpse of its lifetime inside the cells. Given the ability of the BioID bait to capture real-time and transient interactions on a high-throughput basis, it provides a significant advancement over a reductionist approach towards identifying indivisual interaction events during ribosome assembly. However, it must be noted that all interactions recorded by the BioID bait may not be meaningful as they may originate due to cellular crowding and the omnipresence of ribosomes in the bacterial cytoplasm. It is also possible that some interactions recorded by bait may represent its lifecycle before getting installed on the ribosome and thus may lead to false positives. The problem of false positives due to molecular crowding can be overcome by increasing the number of biological replicates used to generate the interactome. It is also possible to cross verify some of the captured PPIs using the existing large scale PPI screens before ursuing any further validation (Ooi, 2010; Rajagopala et al., 2014; von Mering et al., 2005; Wuchty and Uetz, 2014). Further, coupling of BioID with a quantitative mass spectrometric methodology like SILAC (Chen et al., 2015) or iTRAQ (Rauniyar and Yates, 2014) can provide better quantitative metrics of PPIs. Recent advancements in the field have led to the development of new smaller, engineered versions of

BirA* (Kim et al., 2016) that can decrease the molecular radius of biotinylation, further ascertaining better localization of the BioID bait. Finally, it should also be deliberated that BioID is essentially a discovery tool and does not always describe the nature of interactions in a biological setting. Careful consideration and cross-validation by other techniques are essential to correctly interpret a novel PPI event.

2.5 Summary

This chapter describes the development of a technology directed towards identifying RAF candidates by capturing the protein-protein interactions during ribosome assembly. The PPIs were captured using proximity dependent biotinylation (BioID) that employs a variant of the *E. coli* BirA protein. Using a combination of *in vitro* and *in vivo* experiments, we demonstrate the first ever application of BioID to the bacterial ribosomes and captured the molecular events leading up to the formation of the cleft between the head and the 30S platform region. This entailed determination of the BioID interactome of the cleft localizing r-Protein bS18 and the essential RAF Era. Interestingly, BioID established the prevalence of stable interactions between Era, bS18 and other components of the cleft region. Additionally, novel contacts of both Era and bS18 were also identified. This proof-of-concept established the first stage of our systematic efforts directed towards identifying and characterizing RAFs in bacteria.

Chapter III

Bimolecular fluorescence complementation assisted characterization of putative ribosome assembly factors

The work embodied in this chapter is published.

Himanshu Sharma, B. Anand (2016). Fluorescence Bimolecular Complementation Enables Facile Detection of Ribosome Assembly Defects in *Escherichia coli***.** *RNA Biology*, 13, 872-882.

Chapter 3

3.1 Introduction

Assembly factors are known to transiently associate with premature ribosomal subunits to facilitate and organize the seemingly interrelated events ranging from RNA processing, folding and coordinated binding of r-Proteins. It is, therefore, the loss of assembly factors which ensues premature subunits that are inept to associate to form functional ribosomes competent for protein synthesis and this state is captured as an altered ribosome profile in the density gradient fractionation. Cold-sensitive phenotype coupled with an altered ribosome profile in density gradient fractionation has been considered as a symptom to identify the involvement of candidate assembly factors in ribosome maturation (Stokes et al., 2014a). The profiles are obtained by employing laborious and time-consuming sucrose density gradient ultracentrifugation followed by fractionation of ribosomal subunits (Mehta et al., 2012). Therefore, this approach though robust doesn't lend itself for scalability to identify potential assembly factors. This has been a long-standing deterrent in characterizing factors affecting assembly and thus poses a potential setback of overlooking bona fide candidate RAFs. Our pursuit of systematic characterization of RAFs calls for a technique that can complement BioID to validate the candidates in a scalable manner. Here, we have adopted the concept of **Bi**molecular **F**luorescence **C**omplementation (BiFC) and designed and deployed it as a facile tool to capture assembly defect in *E. coli*. The potential utility of this tool is validated by capturing the assembly defects induced by the loss of known assembly factors, *viz.*, RsgA and SrmB in *E. coli*.

3.2 Materials and Methods

3.2.1 Construction of bacterial strains and plasmids

E. coli K-12 MG1655 (Wt) was used as a parent strain for all genome manipulation experiments. Genes encoding uS13 and uL5 were amplified from Wt cells; araC-ara regulatory regions were amplified from vector 8R; V_N (amino acids 1-176) and V_C (amino acids 155-238) coding regions were amplified from vector pET21b-RL027A with primers containing 25-30 bases overlap between the respective constructs. Overlap

extension PCR was performed to create different combinations of V_N-uS13 and V_C-uL5 fusion cassettes. V_N-uS13 fusion cassette was cloned into pOSIP-CT using Kpn1/Xho1 sites and V_C-uL5 fusion cassette was cloned into pOSIP-KO using Kpn1/Pst1 sites. The cassettes were integrated into the genome of Wt strain by a one-step process for cloning and integrating genes into bacterial attB sites referred to as "clonetegration" as described elsewhere (St-Pierre et al., 2013). After selection against the respective antibiotics, the cells were PCR screened for the integration of the given construct using gene-specific primers.

SLV_{NC} strain was used as a parent strain to create null mutants of *srmB*, *rsgA* and *rpsM* using λ-red recombination method (Datsenko and Wanner, 2000a). A Tetracycline cassette amplified from *E. coli* TKC cells with 50 bases flanks for respective genes was used to create the null mutants of SLV_{NC} *ΔsrmB*: *tetA* and SLV_{NC} *ΔrsgA*: *tetA*. Similarly, in order to create SLV_{NC} *ΔrpsM*:*cat*, the Chloramphenicol cassette was amplified from DH10B *ΔrpsM* pCDSSara-S13 strain. Genes encoding SrmB and RsgA were amplified from Wt cells and cloned into vector 8R using ligation-independent cloning to create vectors p*srmB* and p*rsgA*. Plasmid *psrmB* was used to transform SLV_{NC} and SLV_{NC} *ΔsrmB* strains and plasmid p*rsgA* was used to transform SLV_{NC} and SLV_{NC} *ΔrsgA* strains by chemical transformation method. All constructs were verified by sequencing.

3.2.2 Culture growth conditions and growth curve measurement

SLV_{NC} strain or variants were maintained and propagated in LB medium, supplemented with respective antibiotics: Chloramphenicol (15 µg/ml), Kanamycin (15 µg/ml), Tetracycline (10 µg/ml) and Ampicillin (100 µg/ml). All growth related studies were done in 3 independent experimental replicates. Overnight grown cultures were diluted 100 times with fresh LB medium with or without respective antibiotics followed by vigorous shaking at permissible (37 °C) or non-permissible (20 °C) temperatures. Aliquots of 200 µL each were drawn at regular intervals and OD_{600} measurements were taken in Tecan M200 microplate reader (Tecan instruments, Mannedorf, Switzerland) in 96 well transparent flat bottom plates.

3.2.3 RT-PCR for confirming gene expression

Strain SLV_{NC} was grown in presence of 2 mM arabinose till $0D_{600}$ reached 0.8 and mixed with equal volume of RNAlater (Sigma-Aldrich) for 5 mins followed by centrifugation to recover the cells. RNA from recovered cells was isolated by successive extractions using acid phenol (pH 4.5)-chloroform followed by ethanol precipitation. The extracted RNA was treated with RNase free DNase for 2 hrs at 37 °C followed by ethanol precipitation. Subsequently, it was resuspended in RNase free water. Reverse transcription (RT) reaction was performed using SuperScript III Reverse Transcriptase (Invitrogen) according to the manufacturer specifications. PCR was performed using the RT reaction product as a template with forward and reverse primers specific to the corresponding Venus fused uS13 or uL5.

3.2.4 Antibiotic and Lamotrigine treatment

Overnight grown SLV_{NC} strain was diluted into fresh LB medium supplemented with 2 mM arabinose and was allowed to grow at 37˚C and 20˚C till OD_{600} reached 0.2. Subsequently, sublethal concentrations of respective antibiotics were added and the cells were grown at 37˚C or 20˚C till OD_{600} reached 0.6-0.8. Chloramphenicol (7µg/ml), Erythromycin (100µg/ml) and Tetracycline (7µg/ml) were obtained from Sigma-Aldrich, St. Louis, USA and Ampicillin (15µg/ml), Kanamycin (7µg/ml) and Neomycin (7µg/ml) were obtained from Himedia Laboratories (Mumbai, India). After the growth, cells were harvested for fluorescence complementation studies. In order to administer Lamotrigine (Sigma-Aldrich) treatment, cells were grown in fresh LB medium supplemented with 2 mM arabinose and 10 µM Lamotrigine (dissolved in DMSO) at respective temperatures till OD_{600} reached 0.6-0.8.

3.2.5 Immunoblotting

Equal amounts of purified ribosomes from Wt, SLV_{NC} or SLV_{NC} *ΔrpsM* cells were fractionated by SDS-PAGE and transferred to a Polyvinylidene difluoride (PVDF) membrane according to the manufacturer's protocols (Bio-Rad). After incubation with 5% non-fat milk in TBST (10 mM Tris-HCl, (pH 8.0 at 25 °C), 130 mM, NaCl, 0.1% Tween 20) for 60 min at 4 °C, the membrane was washed thrice (10 mins each) with TBST and incubated with antibodies against GFP (Rabbit Anti-GFP antibody, Biobharti

Lifesciences, Kolkata) according to the manufacturers recommendations. Membranes were washed three times for 10 min and incubated with a 1:10,000 dilution of horseradish peroxidase-conjugated Goat anti-rabbit antibody (Merck life sciences) for 1 hour at room temperature. The membrane was then washed thrice with an excess of TBST for 10 mins each and developed with the Clarity™ Western ECL Blotting Substrate (Bio-Rad) according to the manufacturer's recommendations.

Densitometry to calculate the ratio of Venus tagged ribosomal proteins to that of the untagged ribosomal proteins was performed using ImageJ software (Schneider et al., 2012). Calculation of pixel intensities for tagged ribosomal proteins in Wt and SLV_{NC} was done using immunoblots while the intensities for untagged ribosomal proteins were estimated using Coomassie Brilliant Blue (CBB) stained gel. These were compared against standard plot generated for immunoblots and CBB stained gels to estimate the total amount of Venus tagged or untagged ribosomal proteins. The fraction of the tagged and untagged ribosomal proteins was calculated based on the total protein content loaded in each lane.

3.2.6 Isolation of crude ribosomes

Overnight grown cultures were diluted 100 times with fresh LB medium with or without the addition of the respective antibiotic in presence of 2 mM arabinose. Wt, SLV_{NC} or variant strains were grown at 37 °C for permissible and 20 °C for non-permissible temperatures with vigorous shaking. Lamotrigine treatment was administered by growing cells at 37 °C or 20 °C with or without 10 µM Lamotrigine or only DMSO. Cells were grown till OD_{600} reached 0.8 and incubated on ice for 15 mins to allow the formation of "run-off" ribosomes. The following procedure was adopted to ensure preparation of compact and highly active tight-coupled ribosomes (Mehta et al., 2012). The cells were harvested by centrifugation, washed and resuspended in 1 ml chilled buffer A (20 mM Tris–HCl (pH 7.4 at 25 °C), 100 mM NH_4Cl, 10.5 mM Mg $(CH_3COO)_2$, 0.5 mM EDTA, and 6 mM β-ME). Lysis was initiated by incubating cells with 0.2 mg/ml Lysozyme, 40 U RNase free DNase, 1 mM PMSF and RNaseOUT (Invitrogen) for 1 hour on ice followed by 5-8 cycles of freeze-thaw lysis. The lysate was cleared by centrifugation at 30,000xg for 30 mins. The cleared lysate was loaded onto equal volume of buffer B containing 20 mM Tris–HCl (pH 7.4 at 25 °C), 1 M NH_4Cl, 1.1 M sucrose, 10.5 mM Mg $(CH_3COO)_2$, 0.5 mM EDTA and 6 mM β-ME and centrifuged at 220,000xg for 80 min using TLN-100 or Type 70.1 Ti rotors on the Optima TLX or Optima LX-100 XP ultracentrifuge

(Beckman Coulter). The precipitated ribosomes were washed twice with buffer A and resuspended in 200 µL pellet resuspension buffer C (20 mM Tris–HCl, (pH 7.4 at 25 °C), 60 mM NH_4Cl, 10.5 mM Mg $(CH_3COO)_2$, 0.5 mM EDTA, and 6 mM β-ME) and used immediately for further analysis.

3.2.7 Ribosome profile analysis

Ribosome profile analysis was performed as described previously (Mašek et al., 2011) with minor modifications. Briefly, overnight grown cultures were diluted 100 times with fresh LB medium with or without the addition of the respective antibiotic in presence of 2 mM arabinose. Cells were grown till OD_{600} reaches 0.6-0.8. Following this, Chloramphenicol was added to a final concentration of 100 µg/ml and cells were allowed to grow for 15 mins. Subsequently, the cells were chilled rapidly and harvested by centrifugation. Harvested cells were used for isolation of crude ribosomes as described above. Total ribosomes concentration was quantified by measuring the absorbance at 254 nm (A_{254}). To analyse the polysome profiles, crude ribosomes from different strains having A_{254} = 100 were loaded onto a 10-50 % (w/v) linear sucrose gradient in buffer E (20 mM Tris–HCl, (pH 7.4 at 25 °C), 100 mM NH_4Cl, 10.5 mM Mg $(CH_3COO)_2$, 0.5mM EDTA and 6 mM β-ME) followed by centrifugation at 210,000x g for 2 hrs in TLS-55 rotor on Optima TLX ultracentrifuge. Gradients were analysed by manually separating fractions of 100 µL each followed by measurement of absorbance at 254 nm. The subunit peaks were ascertained by isolating RNAs from the fractions. The separated fractions were also analysed for fluorescence complementation wherever required.

3.2.8 Analysis of subunit dissociation by BiFC

Crude ribosomes were incubated with 50 mM EDTA or 10 mg/ml RNase A (Sigma-Aldrich) on ice for 2 hrs. After the completion of the treatment, fluorescence complementation was analysed from RNase A or EDTA treated and untreated ribosomes. In order to decouple the ribosomal subunits gently by employing a low concentration of Mg^{2+}, purified crude ribosomes were resuspended or diluted in buffer D (20 mM Tris–HCl, (pH 7.4 at 25 °C), 60 mM NH_4Cl, 1 mM $Mg(CH_3COO)_2$, and 6 mM β-ME) along with RNaseOUT (Invitrogen) and dialyzed for 4 hrs and later analysed for fluorescence complementation.

3.2.9 Measurement of fluorescence intensities

Quantification of purified ribosomes was carried out by measuring absorbance at 260 nm as a direct indicator of ribosome concentration using GeneQuant 1300 spectrophotometer (GE Healthcare). In order to assess fluorescence complementation without a concentration artefact or any inner filter effect, the ribosome concentration was normalized to 55 nM for all fluorescence measurements. All fluorescence spectrums were generated by exciting the sample at 514 nm and scanning for emission from 524-700 nm, with averaging over three scans after baseline correction in a FluoroMax-4 spectrofluorometer (Horiba Scientific, Edison, NJ). The slit width used for excitation and emission was 2 nm and 5 nm, respectively. All fluorescence experiments were performed with at least three independent trials and the plots are shown for one such representative experiment. The measured fluorescence intensity is further corrected for background fluorescence that arises due to other cellular components. Unless specified otherwise, the relative fluorescence indicates that the fluorescence intensity of the subject is normalized with respect to that of SLV_{NC} making the relative fluorescence for SLV_{NC} to be 1.

3.3 Results

3.3.1 Design of a BiFC system to monitor ribosome assembly in *E. coli*

The involvement of RAFs in assembly is typically probed by whether its deficiency ensues an altered ribosome profile in sedimentation analysis along with a cold-sensitive growth phenotype (Stokes et al., 2014a). An alternative facile assay system that offers scope in scalability for systematic screening of candidate assembly factors is desirable to probe the comprehensive roles of assembly factors during the maturation. In this direction, we exploited the split-Venus fluorescence complementation (Kerppola, 2008; Robida and Kerppola, 2009) to design the assay system. Here, the fluorescent Venus protein is split into two complementary parts – the N (amino acids 1-175) and C-terminal (amino acids 155-238) fragments. When these fragments are brought together by intermolecular interactions, they associate to form the full complement that gives rise to fluorescence.

In order to adopt this concept for our assay system, from each subunit of the ribosome, one ribosomal protein (r-Protein) was chosen based on its location at the interface of the 30S and 50S subunits as well as its accessibility to tagging with Venus (Fig. 3.1A). uS13 and uL5 conformed to such requirements and these r-Proteins were fused to N (V_N) and C (V_C) terminal fragments of the Venus, respectively. We reasoned that when the r-Proteins assemble onto the ribosome, the two complementary Venus fragments might come into close spatial proximity such that the fluorescence is restored. These Venus tagged r-Proteins were integrated into two different loci in *E. coli* genome by a one-step process for cloning and integrating genes into bacterial attB sites referred as clonetegration (St-Pierre et al., 2013) and were placed under an arabinose-inducible promoter for tuneable as well as tight expression (*vide.* Fig. 3.1C and methods). Wt cells were used as a parent for all genome manipulation experiments as well as for all comparative analyses with respect to growth phenotype and BiFC. We refer to the modified strain as SLV_{NC} when both the Venus complements are tagged to uS13 and uL5, respectively and SV_N or LV_C when one of the complements is tagged to either uS13 or uL5 (*vide.* Fig. 3.1B for the nomenclature).

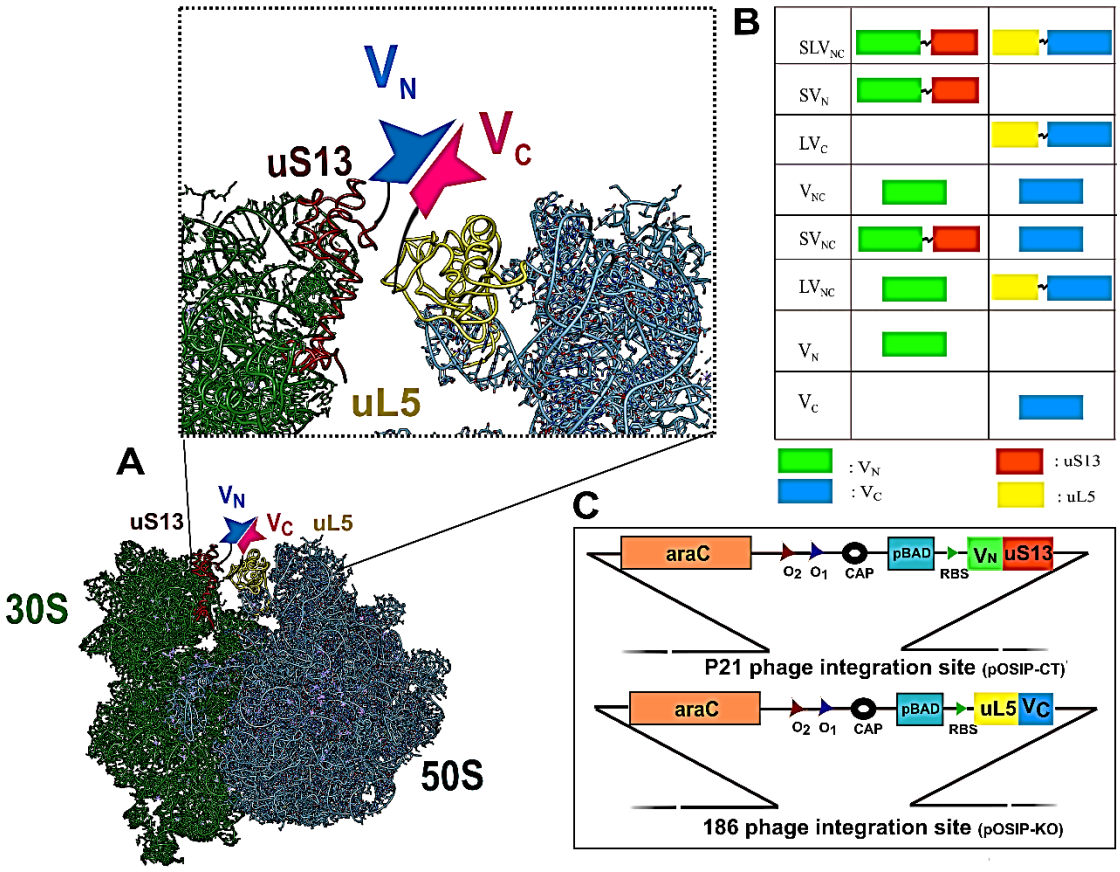

Figure. 3.1: BiFC scheme to monitor ribosome assembly in *E. coli*

(A) The 30S and 50S subunits are shown as ribbons in green and cyan, respectively. uS13 and uL5 are depicted on the ribosome as red and yellow ribbons, respectively. The Split-Venus fragments V_N and V_C are shown as half complementary stars in blue and magenta. uS13 from the 30S and uL5 from 50S subunits were chosen for their location at the interface of the subunits and accessibility of their N and C termini for tagging with Venus fragments.

(B) Nomenclature for various strains created using different combinations of Split-Venus, uS13, and uL5 to establish the BiFC system is shown.

(C) Split-Venus fused uS13 and uL5 were cloned and integrated into the P21 and 186 phage integrations sites, respectively, in *E. coli*. This places the expression of these constructs under the arabinose-inducible promoter (pBAD).

3.3.2 BiFC signals the maturation of ribosome

Having uS13 and uL5 tagged with Venus, we intended to assess whether the presence of tag would alter the growth of SLV_{NC} with respect to Wt. We observed no apparent cell growth defect for SLV_{NC} under the given experimental conditions suggesting that the tag doesn't impair the growth (Fig. 3.2A). Following this, SLV_{NC} cells were screened for the expression of Venus tagged uS13 and uL5 by RT-PCR (Fig. 3.2B). After confirming the expression by RT-PCR, we also tested for the integration of the Venus

tagged uS13 and uL5 onto ribosomes from SLV_{NC} cells by immunoblotting (Fig. 3.2C). It turns out that about 25% of the ribosomes harbour Venus tagged uS13 and uL5. Upon confirming the expression and integration, we proceeded to test whether the fluorescence complementation would occur in strains other than SLV_{NC}. We quantified the relative fluorescence (rf) that is the ratio of fluorescence from cell lysate of Venus tagged variant to fluorescence from cell lysate of SLV_{NC} (Fig. 3.2E). It turned out that though lysates from SV_N and LV_C emitted low background compared to SLV_{NC} (rf < 1), other variants such as V_{NC}, $SV_{NC,}$ and LV_{NC} were fluorescing as good as SLV_{NC} (rf ≥ 1). This suggests that even in the absence of uS13 or uL5, the association between V_N and V_C fragments occurs in a location other than the ribosome leading to fluorescence complementation. This, in turn, would lead to high background signal if one were to assess the complementation directly from the cell lysate.

Since we were interested in monitoring the association of the Venus tagged ribosomal subunits alone, we reasoned that the background noise could be mitigated if the free Venus tagged uS13 and uL5 were segregated from the ribosome-bound counterparts. It is worth mentioning that in order to achieve this, it was just sufficient to purify crude ribosomes by clarifying the cell lysate on an isocratic sucrose cushion. Subsequent steps of density gradient centrifugation and fractionation of ribosomal subunits, as performed for ribosomal profile generation, were obviated (*vide.* Methods). After purifying the crude ribosomes from $SLV_{NC,}$ we tested for the presence of fluorescence signal. Purified crude ribosomes from Wt and V_{NC} were analysed along with SLV_{NC} for the fluorescence signal (Fig. 3.2D). Wt lacks the Venus tag; no complementation is expected to take place on the ribosome for V_{NC}. Therefore, BiFC is not expected from these variants. Indeed, though fluorescence from SLV_{NC} lysates was comparable to that measured from V_{NC} lysates (*vide.* Fig. 3.2D inset), in the modified approach, no significant fluorescence was observed for crude ribosomes purified from Wt and V_{NC} (Fig. 3.2D). However, crude ribosomes isolated from SLV_{NC} alone exhibited significant fluorescent signal (Fig. 3.2D). This suggests that the background noise can be largely mitigated for BiFC analysis even using crude ribosomes that can be rapidly purified from cell lysates.

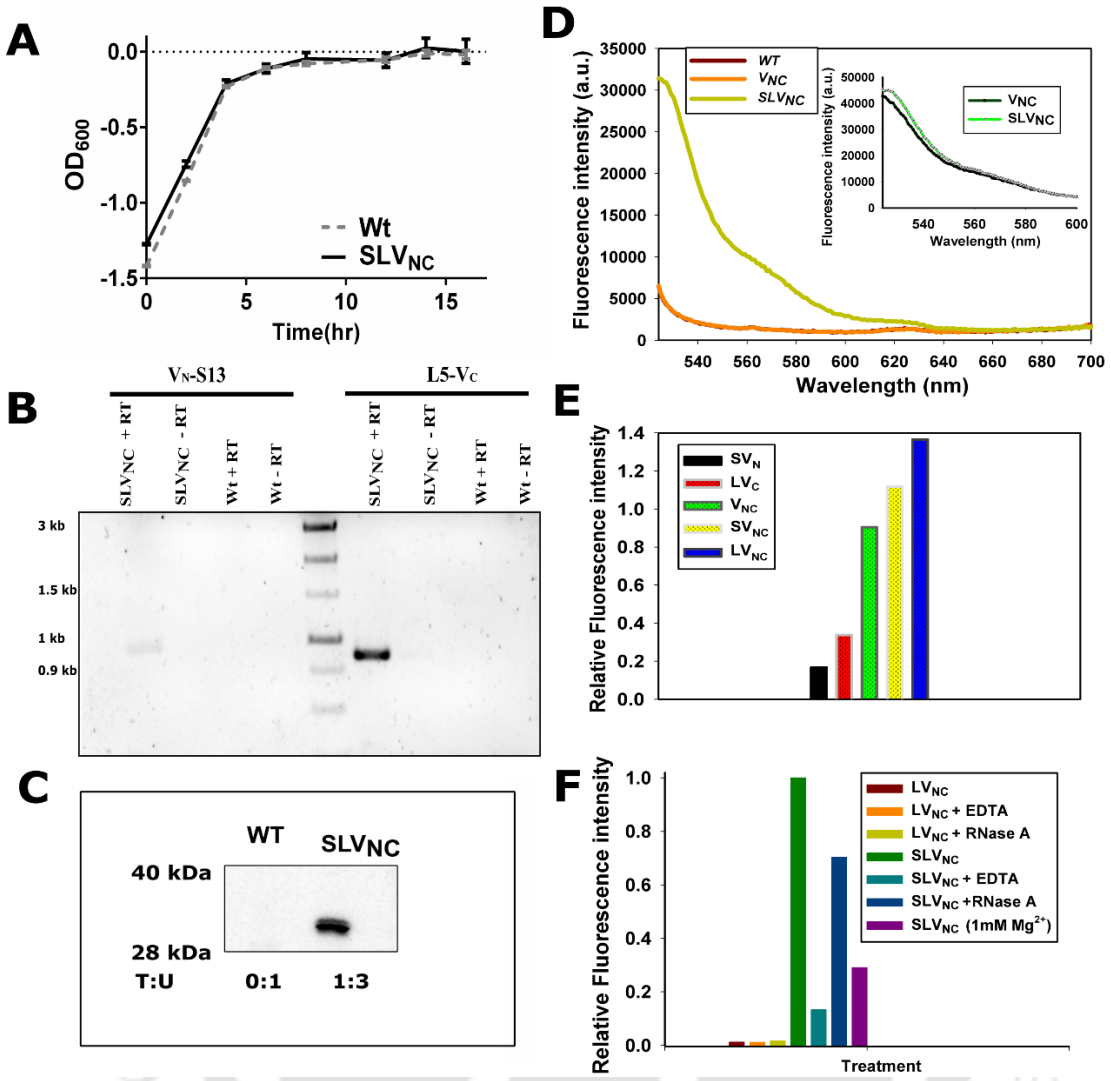

Figure 3.2 Ribosomes isolated from BiFC competent strain are fluorescent

(A) A growth comparison between the Wt and SLV_{NC} strains is shown.

(B) Confirmation of expression of the cloned constructs was done using RT-PCR. The marker is shown in the middle of the gel and the corresponding sizes of the bands are indicated.

(C) Immunoblotting to confirm the expression and integration of Venus tagged uS13 and uL5 onto ribosomes is shown. These were probed with an anti-GFP antibody. The upper and lower bound range of the protein marker with respect to Venus tagged uS13 (31 kDa) and uL5 (34 kDa) is shown. Lanes corresponding to ribosomes from WT and SLV_{NC} are indicated. The ratio of the fraction of Venus tagged ribosomal proteins uS13 and uL5 (denoted as T) to Venus untagged uS13 and uL5 (denoted as U) is indicated at the bottom of each lane.

(D) Fluorescence complementation is shown for crude ribosomes purified from Wt, SLV and V_{NC} strains. Lysate from V_{NC} and SLV_{NC} showed comparable fluorescence (inset).

(E) Fluorescence emissions from cell lysates of SV_N, LV_C, V_{NC}, SV_{NC} and LV_{NC} strains relative to lysate from SLV_{NC} strain were measured (*vide.* text).

(F) Relative fluorescence intensities from LV_{NC} and SLV_{NC} ribosomes are shown after the crude ribosomes were treated with EDTA, RNase A and buffer

containing 1mM Mg^{2+} to dissociate the subunits. The fluorescence intensity normalized with respect to untreated ribosomes from SLV_{NC} is shown.

In order to further reinforce that the BiFC signal from the purified crude ribosomes is due to the association of the 30S and 50S, we destabilized the subunit association by the addition of EDTA and RNase A. This led to the reduction in the fluorescence intensity significantly (Fig.3.2F). EDTA chelates the Mg^{2+} ions essential for subunit association and RNase A compromises the integrity of rRNA – both leading to the dissociation of the 30S and 50S to a varying extent – that was proficiently captured by the fluorescence complementation (Fig. 3.2F). A similar test was performed under low Mg^{2+} concentration that favors the dissociation of subunits. Here, the crude ribosomes were first purified and dialyzed against buffer containing 1 mM Mg^{2+}. This showed a reduction in the BiFC signal (Fig. 3.2F). This suggests that indeed BiFC is highly sensitive and correlates well with the dissociation of the subunits.

Having reduced the background noise and confirmed the subunit association, we set out to ask whether the presence of tagged uS13 and uL5 alleles would impair the assembly. Therefore, we performed sedimentation analysis for the ribosomes prepared from the cell lysates of SLV_{NC} and Wt under conditions that favor association of ribosomal subunits (10 mM Mg^{2+}). For Wt, we observed majorly 70S and polysomes with minor amounts of 30S and 50S indicating that the association between 30S and 50S is robust. The ribosome profile for SLV_{NC} was similar to that of Wt suggesting that the assembly is not impaired by the presence of Venus tag on uS13 and uL5 (Fig. 3.3A & 3.3B).

Further, the BiFC measured from the fractionated crude ribosomes correlated well with the sedimentation profile displaying significantly high fluorescence signal for fractions corresponding to 70S and polysomes than those related to 30S and 50S. This suggests that BiFC indeed captures the association of the subunits (Fig. 3.3C). It is worthy to note that when the sedimentation analysis of the ribosomes was performed by loading the cell lysates directly onto the sucrose gradient, noticeable fluorescence corresponding to 30S and 50S fractions was observed (data not shown). This might have stemmed from the co-sedimentation of trace amounts of free Venus-tagged uS13 and uL5 along with the 30S and 50S subunits. In line with our earlier observation, however, when we separately purify crude ribosomes and then subjected them to sedimentation analysis, this background noise is largely alleviated (Fig. 3.3C).

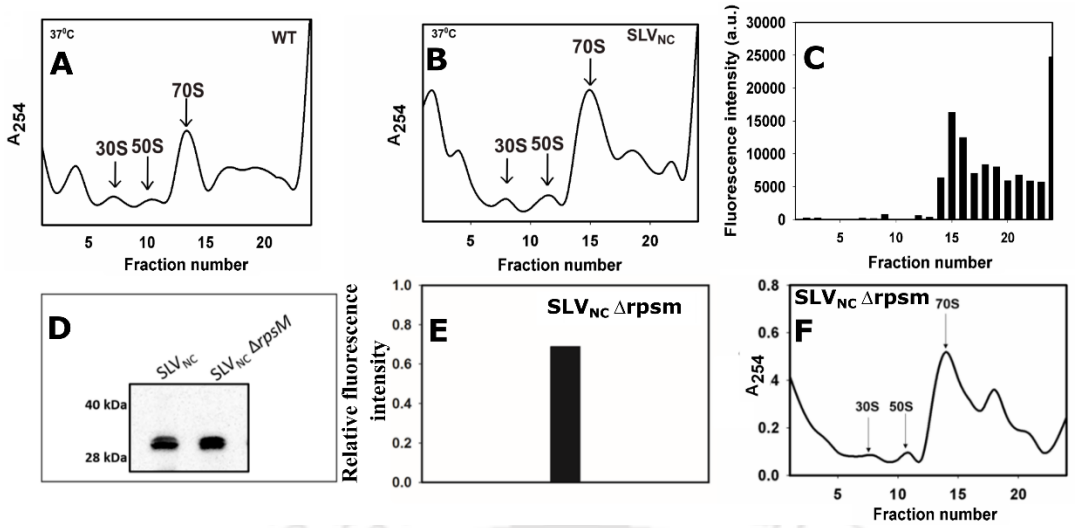

Figure 3.3: BiFC profiles correlate with conventional ribosome profiles

(A) Ribosome profile for Wt is presented. The peaks pertaining to 30S, 50S, and 70S are indicated.

(B) Ribosome profile for SLV_{NC} is presented.

(C) The BiFC measured from the corresponding fractions (in A) is shown for SLV_{NC}. The BiFC profile is in agreement with the sedimentation profile of the ribosomes.

(D) An immunoblot is shown for equal amounts of ribosomes isolated from SLV_{NC} and SLV_{NC} *ΔrpsM* cells. These were probed with the anti-GFP antibody. The upper and lower bound range of the protein marker with respect to Venus tagged uS13 (34.4 kDa) and uL5 (31.4 kDa) is shown (Left). This relative fluorescence intensity of crude ribosomes prepared from SLV_{NC} Δ*rpsM* that was normalized with respect to the crude ribosomes from SLV_{NC} (Right).

(E) Relative fluorescence intensities for crude ribosomes from SLV_{NC} *ΔrpsM* that was normalized with respect to the crude ribosomes from SLV_{NC}.

(F) The ribosome profile obtained from SLV_{NC} *ΔrpsM* cell lysates. The peaks corresponding to 30S, 50S, and 70S are indicated.

In addition, we also created a null mutant of *rpsM* that codes for the non-essential S13 (Bubunenko et al., 2007) against the SLV_{NC} background (SLV_{NC} *ΔrpsM*). In this modified strain, the Venus tagged uS13 allele remains as the only source for uS13 and that allowed us to assess whether the Venus tag on uS13 would impair the assembly. The integration of Venus tagged uS13 onto the ribosomes was confirmed and that it displayed significant fluorescence complementation (Fig. 3.3D). Further, based on the correspondence between the ribosome profile and the BiFC signal, we confirmed that the Venus tag doesn't alter the assembly (Fig. 3.3E & 3.3F). Similar such efforts to create a null mutant of *rplE* that codes for the essential L5 (Korepanov et al., 2007; Shoji et al., 2011) didn't fructify despite several meticulous attempts. Therefore, it wasn't possible to

assess the effect of tag against the SLV_{NC} *ΔrplM* background. However, based on SLV_{NC} and SLV_{NC} *Δrps*M, we believe that it is unlikely that the tag on uL5 will impact the assembly significantly.

3.3.3 BiFC detects maturation defects that are prompted by the loss of assembly factors

Encouraged by the ability of the BiFC system to monitor the state of assembly (Fig. 3.2F), we proceeded to assess this in case of assembly defects due to the loss of known assembly factors. RsgA is a circularly permuted GTPase that is known to be involved in the assembly of small subunit (30S) in *E. coli*. The loss of RsgA produces slow growth and severe assembly defect that is characterized by the accumulation of premature 30S and a concomitant reduction in the 70S particles (Goto et al., 2011a; Guo et al., 2011; Himeno et al., 2004; Jomaa et al., 2011; López-Alonso et al., 2017). Similarly, the loss of SrmB – a DEAD box helicase – is shown to produce a 50S assembly defect in *E. coli* (Cartier et al., 2010; Charollais et al., 2003b; Proux et al., 2011; Trubetskoy et al., 2009). We chose these factors as they represent a different class of proteins – one is a GTPase and another is a helicase – that are reasonably well characterized and shown to participate in the maturation of different subunits (30S or 50S). We generated null mutants of *rsgA* and *srmB* against the genetic background of SLV_{NC}. The fluorescence complementation from these mutants was assessed relative to that of SLV_{NC}. This showed that compared to the crude ribosomes purified from SLV_{NC}, those from SLV_{NC} *ΔrsgA* and SLV_{NC} *ΔsrmB* strains showed a drastic reduction in fluorescence complementation (less than 50% that of SLV_{NC}) that was more pronounced at 20°C (Fig. 3.4A). In line with this, both SLV_{NC} *ΔrsgA* and *ΔsrmB* strains showed a slow growth phenotype at 37°C that was further exacerbated at 20°C reinforcing the essential requirement of RsgA especially at low temperature (Fig. 3.4B & 3.4C).

The ribosome profile generated from SLV_{NC} *ΔrsgA* recapitulated the earlier observation, which showed the accumulation of 30S at the expense of 70S (Fig. 3.4D). In tune with this, the BiFC profile that was generated from the corresponding fractions from the sedimentation analysis mirrored this altered ribosome profile with drastically reduced fluorescence signal for 70S relative to SLV_{NC} (compare with Fig. 3.4E inset) highlighting the assembly defect. In contrast to SLV_{NC} *ΔrsgA*, despite the reduction in BiFC, the SLV_{NC} *ΔsrmB* exhibited minor growth defect both at 37°C and 20°C, which is consistent with

previous reports (Charollais et al., 2003b; Proux et al., 2011; Trubetskoy et al., 2009) (Fig. 3.4B & 3.4C). The ribosome profile generated for this strain grown at 37°C showed noticeable amounts of 40S whose appearance represents the hallmark of *ΔsrmB* indicating the 50S assembly defect (Fig. 3.4F) (Charollais et al., 2003b; Jagessar and Jain, 2010a) . We measured BiFC for the fractionated ribosomes from this strain that showed drastically reduced fluorescence consistent with the altered ribosome profile (Fig. 3.4G). Strikingly, the fractions corresponding to 70S from SLV_{NC} *ΔsrmB* displayed significantly reduced fluorescence than those from SLV_{NC} *ΔrsgA* suggesting that the cause for the defect in subunit association differs between the two (*vide*. Discussion).

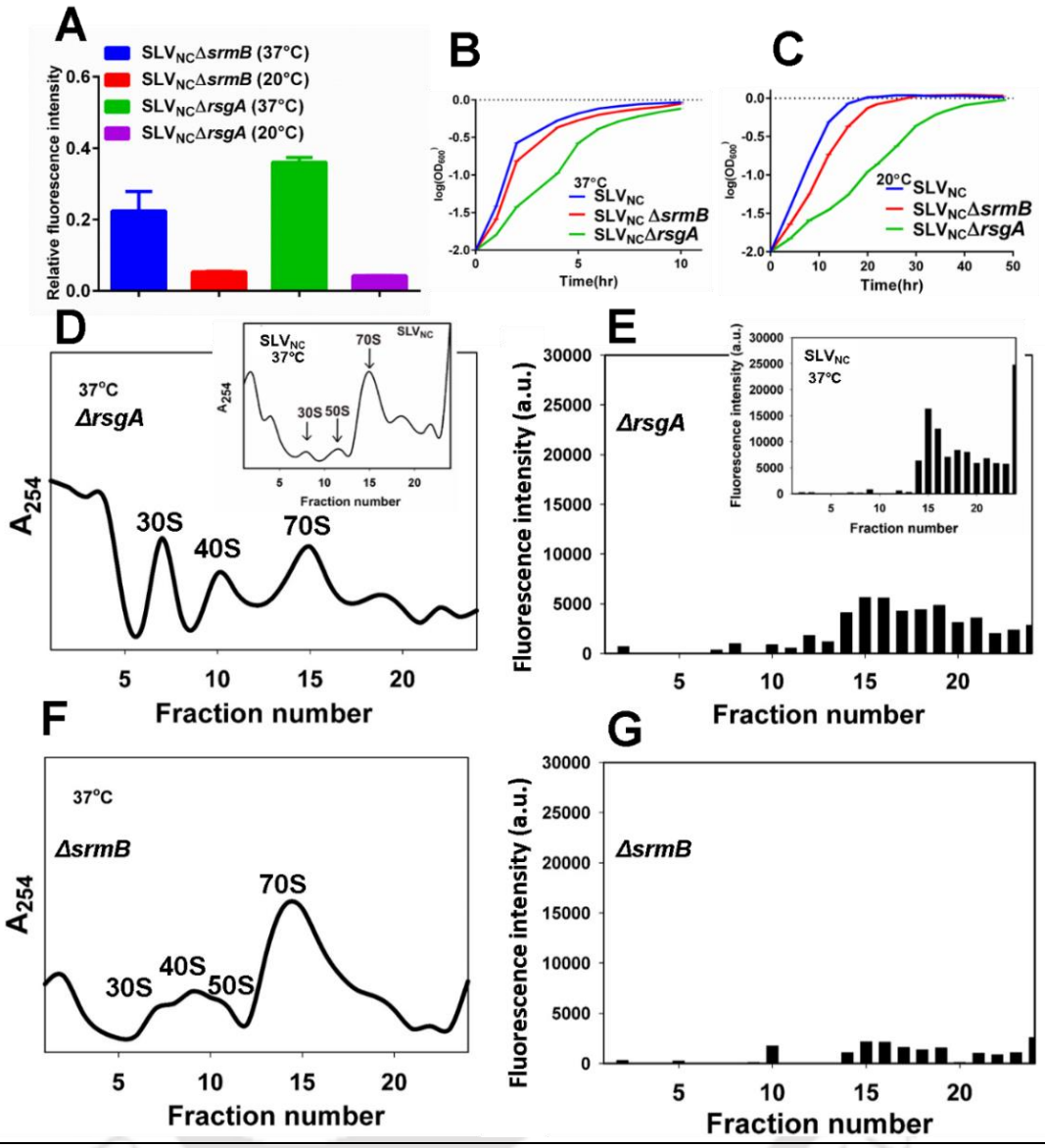

Figure 3.4: BiFC detects assembly defects that are induced by the loss of assembly factors

(A) Relative fluorescence complementation is shown for purified ribosomes from SLV$_{NC}$ *ΔsrmB* and SLV$_{NC}$ *ΔrsgA* strains that were grown at 37˚C and 20˚C. Drastic differences in the fluorescence complementation profiles were observed at 20˚C for SLV$_{NC}$ *ΔsrmB* and SLV$_{NC}$ *ΔrsgA* strains.

(B) Growth profiles of SLV$_{NC}$ *ΔsrmB* and SLV$_{NC}$ *ΔrsgA* strains are shown at permissible temperatures (37˚C).

(C) Growth profiles of SLV$_{NC}$ *ΔsrmB* and SLV$_{NC}$ *ΔrsgA* strains are shown at non-permissible temperatures (20˚C). Both SLV$_{NC}$ *ΔsrmB* and SLV$_{NC}$ *ΔrsgA* strains exhibited growth defect that was more prominent at 20˚C.

(D) Ribosome profiles for SLV$_{NC}$ *ΔrsgA* cells are shown. The peaks corresponding to 30S, 40S and 70S are indicated. Ribosome profile for SLV$_{NC}$ that corresponds to wild type is shown in the inset.

(E) BiFC measured for SLV_{NC} *ΔrsgA* from the corresponding fractions of the sedimentation analysis is presented. BiFC profile for SLV_{NC} is shown in the inset.

(F) Ribosome profiles for SLV_{NC} *ΔsrmB* cells are shown. The peaks corresponding to 30S, 40S, 50S and 70S are indicated.

(G) BiFC measured for SLV_{NC} *ΔsrmB* from the corresponding fractions of the sedimentation analysis is presented.

To further reaffirm the notion that the reduced fluorescence complementation in SLV_{NC} *ΔrsgA* and SLV_{NC} *ΔsrmB* strains was due to the assembly defect that was induced by the loss of RsgA and SrmB, respectively, we exogenously complemented their expression at 20°C. The fluorescence complementation from these ribosomes was assessed relative to that of SLV_{NC}. This showed that the restoration of fluorescence complementation corresponds to 30% for SrmB and 60% for RsgA (Fig.3.5A). In tune with this, the slow growth defect was relieved by exogenous expression of *rsgA* and *srmB* (Fig. 3.5B & 3.5C). Intriguingly, we observed in the gene complementation that though the restoration of growth defect is comparable to the wild type, the restoration of BiFC, especially for SLV_{NC} Δ*srmB*, is not appreciable (Fig. 3.5A & 3.5C). This may be attributed to the fact that the exogenous dosage in the gene complementation experiments lacks the regulatory mechanism the corresponding chromosomal counterpart is subjected to under varying environmental cues. This is likely to impact the temporal expression of *srmB* in SLV_{NC} *ΔsrmB* compared to the wild type. Secondly, the over-expression of *rsgA* and *srmB* from multi-copy plasmid as against the single copy in native chromosomal locus might have some undesirable effects on the assembly that could lead to low fluorescence. It is possible that these factors could exert negligible effect on RsgA compared to SrmB giving rise to better restoration of BiFC for RsgA. Given this, it is indeed encouraging to note that BiFC can highlight such subtle differences in the roles of the assembly factors.

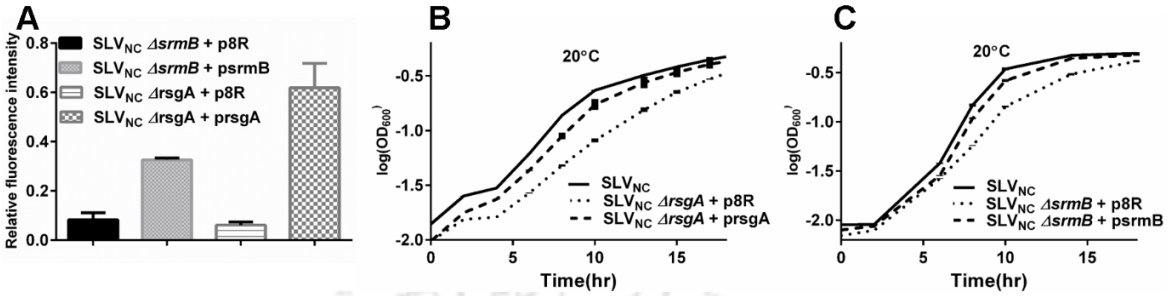

Figure 3.5: BiFC profiles mimic conventional ribosome profiles

(A) Null mutants complemented with respective assembly factors by exogenous expression depicted significant restoration of BiFC for the purified ribosomes at 20˚C. BiFC for gene complementation is shown relative to that of SLV_{NC} and the error bars depict the standard deviation. The empty vector is depicted as p8R and the vector harbouring the respective assembly factor is indicated.

(B) & (C) The stress on growth due to deletion of assembly factors at 20˚C was relieved by expression of the assembly factors from an inducible vector. The empty vector is depicted as p8R and the vector harbouring the respective assembly factor is indicated. The complementation experiment was conducted against wild type and deletion background of the respective assembly factor in SLV_{NC}.

3.3.4 BiFC highlights assembly defects that are provoked by chemical perturbation too

Apart from the loss of assembly factors, a few small molecule inhibitors are shown to induce assembly defects. Therefore, we intended to assess the feasibility of deploying the BiFC system for screening small molecule inhibitors of ribosome maturation. Towards this, we utilized Lamotrigine, an anticonvulsant drug that is recently shown to reversibly induce assembly defects especially at low temperatures in bacteria (Stokes et al., 2014a). Crude ribosomes purified from Lamotrigine treated SLV_{NC} exhibited low fluorescence complementation with respect to that of untreated SLV_{NC} at 37°C (Fig. 3.6A). However, those grown at 20°C showed drastic reduction in BiFC suggesting that the Lamotrigine induced assembly defect is indeed proficiently captured by BiFC (Fig. 3.6A). In line with the maturation defect at 20°C, Lamotrigine treated cells exhibited a slow growth phenotype at 20°C than at 37°C (Fig. 3.6B). These observations are in consonance with an earlier report (Stokes et al., 2014a). Encouraged by this, we proceeded to probe whether BiFC would detect

antibiotics induced assembly defects too. Antibiotics *viz.*, Chloramphenicol, Kanamycin, Neomycin, Erythromycin and Tetracycline target the protein synthesizing machinery and render it inactive. This leads to the cascading effect of diminishing the production of r-Proteins whose dosage insufficiency becomes rate limiting thus, provoking the ribosome assembly defects indirectly (Nikolay et al., 2014; Siibak et al., 2009). On the contrary, Ampicillin, which is an inhibitor of transpeptidases prevents cell wall synthesis and therefore is anticipated to not induce any ribosome assembly defects.

We assessed BiFC from crude ribosomes purified from antibiotics treated SLV_{NC} relative to that of untreated SLV_{NC}. Indeed, the ribosomal antibiotics drastically reduced BiFC at 20°C than at 37°C; however, Ampicillin remained neutral to BiFC both at 20°C and 37°C (Fig. 3.6C). This underscores that BiFC is specific for monitoring the assembly defect induced by antibiotics that target ribosomes. This further suggests that though BiFC highlights the deficiency in the ribosome maturation, it doesn't deconvolute whether the defect is due to an error in the assembly process or indirectly caused by an error in protein synthesis. This is a conceptual limitation wherein the role of BiFC is restricted to just as a "read-out" of ribosome assembly defect – like the density gradient sedimentation analysis – and it lacks the scope to uncover the underlying causal relationships. Resolving this necessitates further analyses as shown in earlier studies (Stokes et al., 2014a). Overall, the applicability of BiFC seems to be scalable for medium throughput platforms for screening chemical inhibitors that target the ribosome assembly either directly or indirectly and has the potential to uncover new classes of assembly inhibitors and antibiotics.

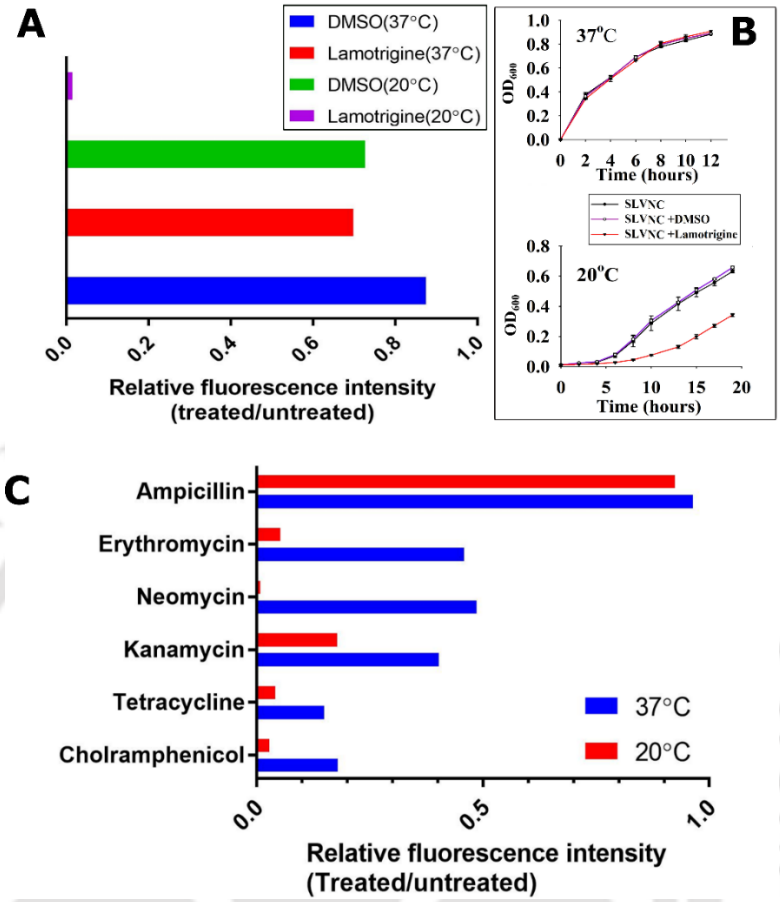

Figure 3.6: BiFC displays assembly defects that are provoked by chemical inhibitors and antibiotics.

(A) Fluorescence complementation profile of SLV_{NC} cells treated with Lamotrigine at 37˚C and 20˚C is presented. Major deviation in the levels of fluorescence complementation was observed at 20˚C than at 37˚C. BiFC for ribosomes from DMSO treated cells is shown as a mock experiment. BiFC is shown relative to that of untreated SLV_{NC}.

(B) The effect of Lamotrigine on the growth of SLV_{NC} was also monitored at 37˚C and 20˚C. In line with the BiFC profile, the growth defect was more pronounced at 20˚C.

(C) The effect of Chloramphenicol, Kanamycin, Neomycin, Erythromycin, Tetracycline and Ampicillin on the fluorescence complementation profile at sublethal concentrations is shown at 37˚C and 20˚C. BiFC is shown relative to that of untreated SLV_{NC}. BiFC highlights that the defects are more pronounced at 20˚C for ribosome targeting antibiotics.

3.4 **Discussion**

Exploiting the modular nature of fluorophore maturation in Venus, we have presented an ingenious approach to track the maturation of ribosomes in *E. coli*. Our earlier attempts to detect BiFC directly from cell lysates didn't fructify due to the presence of background noise. This arises owing to the intrinsic limitation of segregating the free Venus tagged uS13 and uL5 from the ribosome bound counterparts. Despite this, in our modified approach, in order to detect BiFC, it is sufficient to purify the crude ribosomes alone that can be accomplished rather quickly. We have further demonstrated that BiFC emanating from crude ribosomes has negligible contributions from the premature subunits (Fig. 3.3C, 3.4E & 3.4G). This suggests that subsequent long stages of fractionation of the ribosomal subunits under the density gradient centrifugation to generate the profile for assessing the maturation can be circumvented. This, in turn, minimizes the considerable time and effort to detect assembly defects. At this point, though the requirement to purify the crude ribosomes limits the scalability to be medium throughput, we anticipate that further investigation to alleviate this background can potentially scale-up the application.

Though Venus is a stable fluorescence protein with high quantum yield, it has been reported that the presence of high concentration of chloride and nitrate ions at mildly acidic (pH 6.0 at 25 °C) condition reduce the fluorescence drastically (Hsu et al., 2010). Further, it was observed for YFP that the maturation of complements for BiFC is impacted around pH 7.2 (Ottmann et al., 2008). In our study, we observed that pH 7.4 and 60 mM NH_4Cl seem to be suitable for the fluorophore maturation and therefore we suggest that the pH range of 7.4-8.0 may be optimal for detection of ribosome assembly defect using BiFC.

Our choice of tagging uS13 and uL5 was based on their location at the interface of the ribosomal subunits. In our diverse validation experiments such as the loss of assembly factors RsgA and SrmB as well as the presence of chemical inhibitors such as Lamotrigine and antibiotics – that target the ribosome at different sites – induce assembly defects that are proficiently captured by BiFC. Nevertheless, it raises a remote possibility wherein the defects induced by hitherto unidentified assembly factors that have no effect on subunit association, may go unnoticed. However, it appears that the quality control mechanism

ensures that only the mature subunits are allowed to interact and enter the translational cycle while the premature subunits are rapidly cleared from the cell (Karbstein, 2013; Strunk et al., 2012). In such scenario, only the mature 30S and 50S subunits can be stable long enough to associate and turn on BiFC. Further, BiFC is highly sensitive to the spatial proximity of the chromophore complements and therefore any mild assembly defect that impacts the orientation of uS13 and uL5, despite their presence on the ribosome, can render the BiFC inactive. For example, in the case of SLV_{NC} *ΔsrmB*, we noticed that though the population of the 70S is not altered as deduced from the ribosome profile, the BiFC corresponding to these fractions is significantly low (Fig. 3.4G). We speculate that such possibility could arise if the positioning of uS13 and uL5 is distorted in those 70S particles owing to some minor assembly defect of the subunits that seems to be tolerated with no major perturbation to growth rate. Therefore, it is highly likely that only the perfect positioning of uS13 and uL5 on the mature subunits are competent to facilitate the spatial proximity of the Split-Venus to turn on BiFC with high efficiency. On the other hand, the reduction in BiFC for the 70S in SLV_{NC} *ΔrsgA* is concomitant with the reduction seen in the ribosomal profile (Fig. 3.4E). This suggests that the 70S population accrued in SLV_{NC} *ΔrsgA* harbors ribosomal subunits wherein the spatial proximity between uS13 and uL5 remains unperturbed. Since RsgA and SrmB participate in the maturation of different subunits (i.e., 30S and 50S, respectively), their mode of action is expected to differ as well. In such scenarios, it is highly encouraging to note that BiFC could capture those subtle defects that may go unnoticed in sedimentation analysis. This further suggests that in conjunction with the conventional sucrose gradient sedimentation analysis, BiFC could bring forth new insights into the roles of the assembly factors. Finally, the choice of uS13 and uL5 is not sacrosanct; it is possible that in lieu of uS13 and uL5, another r-Protein pair that is located at a different region can be chosen based on their proximity on the ribosome and assayed for their ability to activate BiFC. We also posit that an extension of the SLV_{NC} strategy could be designed for monitoring maturation of individual subunits. The Venus halves, when fused with a pair of r-Proteins from a ribosomal subunit would give rise to BiFC upon assembly of the respective subunit.

Earlier attempts to capture assembly defects in *E. coli* based on fluorescence relied on dual colour fluorescence. This necessitated that the ratio of the intensity of both colours are

characterized which is further subjected to variability under different experimental conditions (Nikolay et al., 2015; Nikolay et al., 2014). In our design, this is circumvented by adopting BiFC as a surrogate measure to monitor the state of the ribosome assembly. Since the fluorophore maturation in Venus – owing to the requirement of spatial proximity of the complementary parts – takes place on the fully assembled ribosomes, BiFC analysis based on purified crude ribosomes alleviates false positives to a large extent without compromising the sensitivity. This work shows that in most instances, the BiFC "OFF" state signals the abrogation of assembly while the BiFC "ON" state indicates the proficient assembly. This simplicity enables one to assess the status of the ribosome maturation with ease and precision based on the binary state of BiFC – "ON" or "OFF". Thus, this facile assay system can be readily adopted for systematic screening to probe the involvement of RAF candidates identified from the BioID approach and small molecule effectors of ribosome maturation both in a quantitative and qualitative framework.

3.5 **Summary**

Here, we employ bimolecular fluorescence complementation (BiFC) to detect ribosome assembly defects. The fusion of N and C-terminal fragments of Venus fluorescent protein to the ribosomal proteins uS13 and uL5, respectively, in *E. coli* facilitated the incorporation of the tagged uS13 and uL5 onto the respective ribosomal subunits. When the ribosomal subunits associated to form the 70S particle, the complementary fragments of Venus are brought into proximity and render the Venus protein fluorescent. Genetic or chemical perturbations that prompt assembly defects inhibited subunit association and showed abrogation of fluorescence complementation. This suggests that BiFC can be employed as a surrogate measure to detect ribosome assembly defects thus complementing the BioID methodology by offering a facile platform for a systematically screen towards validating potential RAFs.

Chapter IV

Characterization of essential ribosome assembly factors

Chapter 4

4.1 Introduction

In pursuit of developing an inclusive model of ribosome biogenesis, we have displayed the applicability of a versatile BiFC based validation approach. The methodology calls for the creation of null mutants of the respective RAF candidate. In order to achieve this, gene knock-out or expression interference of target genes can be achieved using conventional genetic approaches like transposon mutagenesis (Lin et al., 2014), recombination-mediated gene replacement (Datsenko and Wanner, 2000b) and RNA-interference (Shan, 2010). However, some proteins play an indispensable role in the maintenance of fundamental cellular functions and form the *essentialome* of an organism (Gerdes et al., 2003; Giaever et al., 2002; Hutchison et al., 1999; Kamath et al., 2003). The application of conventional genetic manipulation tools on them is generally met with lethality. Over the years, observation of decisive phenotypes of survival or lethality has given rise to the binary theory of gene essentiality where a gene can only be essential or non-essential (Yu et al., 2008). Similarly, many r-Proteins and RAF encoding genes are also known to be essential in nature and repeated futile attempts to delete these genes have undermined the efforts to fully characterize their cellular function. Excitingly, constantly evolving models of gene function have also proposed an alternative – a contextual model of essentiality (D'Elia et al., 2009; Rancati et al., 2017) where intracellular or extracellular factors may dictate essentiality of a gene (Formstone and Errington, 2005; Reuber et al., 1991; Reuber and Walker, 1993). The theory also seems plausible as RAFs like RbfA are known to be essential to relieve growth defects arising at lower temperatures in a dosage-dependent manner (Connolly et al., 2008; Dammel and Noller, 1995b).

Inspired by these concepts, we reasoned that depletion of the dosage of a gene rather than its absolute deletion could be an alternative method to better characterize the function of essential RAFs. Conditional depletion of such factors at the protein level would provide a giant leap towards functional characterization of RAFs like Era (Shimamoto and Inouye, 1996), Der (Hwang and Inouye, 2006a) and ObgE (Arigoni et al., 1998) by BiFC based validation. In contrast to perturbation at the genetic level – which is subjected to extensive layers of regulatory control and feedback circuits – intervention at the protein level is subtle,

yet very selective. In order to achieve this, we surveyed the current methodologies used for conditional depletion of proteins. We reasoned that an ideal strategy should mediate conditional protein depletion without compromising the integrity of the respective genetic loci and also not interrupt the expression of neighbouring genes at transcriptional level. Ensuring this would circumvent any pleiotropic effects at the global level. A targeted protein depletion strategy could fulfill both criteria and also provide extraordinary specificity for depletion. The recent application of targeted protein depletion strategies in eukaryotic (Jungbluth et al., 2010) and prokaryotic systems (Carr et al., 2012) have revolutionized the spectrum of conditional protein depletion. Among these, the application of components of the ribosome rescue system from *E. coli* (Carr et al., 2012; Davis et al., 2009) and *Mesoplasma florum (M. florum* or *Mf)* (Gur and Sauer, 2008a) has shown great promise of achieving specific and regulated protein depletion in bacteria. Moreover, importing *Mf*- Lon proteases (Lon) does not seem to interfere with the protein degradation machinery in host organism such as *E. coli* (Cameron and Collins, 2014). Importing this system into *E. coli* requires fusion of the *Mf*-ssrA tag to the target protein and expression of *Mf*-Lon. Fusion of ssrA tag can be achieved using a one-step homologous recombination method and further the system requires the introduction of a single *Mf*-Lon encoding plasmid, making it relatively easy to implement this system. Thus in this chapter, we have attempted to deploy the *M. florum* protein depletion machinery to conditionally deplete the essential RAF Era in *E. coli*. The choice of Era partly stems from our experience in manipulation of the essential *rnc* operon and the central role of Era in the ribosome assembly landscape.

4.2 Materials and Methods

4.2.1 Construction of bacterial strains and plasmids

E. coli K-12 MG1655 (Wt) was used as a parent strain for all genome manipulation experiments. A DNA cassette encoding HA tag followed by an optimized *Mf*-ssrA tag (Cameron and Collins, 2014) was amplified by overlap extension PCR. The cassette was further fused with a chloramphenicol acetyltransferase (CAT) gene flanked with *loxP* sites on both ends using another cycle of overlap extension PCR. Finally, flanking regions (50 bases each) specific to 3'end of *era* were added to this cassette for knock-in using λ-red mediated

recombination (Datsenko and Wanner, 2000b) to create the strain Era^{m}. The construct was generated so as to fuse the HA-ssrA tag in-frame with Era. The transformed colonies were screened using colony PCR and confirmed by sequencing. Upon confirmation of integration, CAT gene was cured by overexpressing Cre recombinase from a heat curable plasmid. Era^{m} was transformed with a plasmid overexpressing *Mf*-Lon protease (p*Mf*-Lon) to generate the strain Era^{Lon}.

4.2.2 Induction of protein depletion and cell growth arrest

Freshly prepared Era^{Lon} cells were diluted 100 times with fresh LB medium supplemented with 25 µg/mL Chloramphenicol (Cmp), 13 mM Arabinose (Ara) and incubated at 37 °C with shaking at 180 rpm. The OD_{600} was measured at regular interval and cells were harvested aseptically when OD_{600} reached 0.2. The harvested cells were diluted 10 fold into fresh medium supplemented with same Cmp and Ara concentrations. The procedure was repeated for one more cycle with close monitoring of OD_{600}. Cells were harvested after achieving growth arrest.

4.2.3 Growth profiling

The primary inoculum was grown overnight from freshly transformed colonies in LB medium with the respective antibiotics at 37 °C, with shaking at 180 rpm. The next day, OD_{600} of primary cultures were normalized. These cultures were then diluted into fresh 1 ml LB medium with appropriate antibiotics in 24 well plates following incubation at 37 °C with regular shaking. OD_{600} measurements were taken in a Tecan infinite M200 multimode plate reader at every 30-minute (mins) interval until cultures reached saturation.

4.2.4 Ribosome profiling

Wt, Era^{m} or Era^{Lon} were grown as described above till growth arrest was observed. The cells were harvested by centrifugation at 4 °C, 8000xg for 10 mins. Cells were washed once with Buffer CL (20 mM Tris-HCl (pH 7.6 at 25 °C), 150 mM NH_4Cl, 10.5 mM $Mg(OAc)_2$, 0.5 mM EDTA and 6 mM β-ME). The washed cells were resuspended in 1 mL

buffer CL supplemented with 1 mg/mL Lysozyme, 1 mM PMSF, 0.5 x CellLyticB reagent and incubated on ice for 1 hour (hr). The cells were then lysed by 5 cycles of freeze-thawing in liquid nitrogen. The lysate was clarified by centrifugation at 30,000xg followed by A_{260} (Absorbance at 260 nm) quantification in Implen Nanospectrophotometer. 250 µL of 100 A_{260} was loaded on a 10-50% sucrose gradient prepared in buffer CL. The samples were centrifuged in rotor TH-641 in Thermo-Sorvall WX^{+} ultracentrifuge at 102,000xg for 16 hrs and subsequently fractionated using a syringe pump connected to Akta Pure system for measurement of A_{260} profiles.

4.2.5 Translation kinetics

Translation kinetics was studied using a β-galactosidase **(**bgal) production assay (Dalbow and Bremer, 1975; Dalbow and Young, 1975; Miller, 1992). Briefly, cells were freshly transformed with a plasmid containing an IPTG inducible, bgal encoding gene. Overnight grown cultures were diluted into fresh 5 ml LB medium with 100 µg/ml ampicillin and allowed to grow at 37 °C and 180 rpm till OD_{600} reached 0.6 or until growth arrest was observed. These cultures were then shifted to a water bath at 37 °C. Bgal production was induced with 1 mM IPTG followed by instant mixing for 5 seconds (secs). Following this, aliquots of 200 µL were taken every 30 secs and added to tubes containing 300 µL of Z-buffer (60 mM $Na_2HPO_4{\cdot}2\ H_2O$, 40 mM $NaH_2PO_4{\cdot}\ H_2O$ (pH 7.0 at 25 °C), 10 mM KCl, 1mM $MgCl_2$, 5 mM β-ME), 20 µL of chloroform, and 10 µL of 0.1% SDS. The tubes were instantly vortexed for 20 secs and transferred to ice. The samples were drawn till 10 mins from the time of induction followed by incubation on ice for 5 mins and then at room temperature (RT) for 5 mins. Finally, 500 µL of 1 mg/mL ONPG (o-nitro-phenyl galactopyranoside) dissolved in Z buffer was added to the reaction. ONPG hydrolysis was allowed to proceed for 15 mins at RT (25 °C) after which the reaction was stopped by adding 500 µL of 1 M Na_2CO_3. The colour was developed for 15 mins and then 300 µL of the mix was transferred to 96 well plates for measurement. The absorbance of each reaction was then measured at 420 nm and 550 nm using a Tecan infinite M200 multimode plate reader.

4.3 **Results**

4.3.1 Strain engineering for regulated depletion of Era

In order to create the synthetic genetic circuit for Era depletion, we imported components of the *Mf*-Lon protease system into *E. coli*. The *Mf*-ssrA tag was fused onto the C-terminus of an HA (hemagglutinin) epitope tag to create a HA-ssrA tag (Fig. 4.1A). The cassette was integrated in-frame to the 3' end of *era*, which is part of the native *rnc* operon using λ-red recombination (Fig. 4.1B). The integration was confirmed by flanking primers and further validated by sequencing.

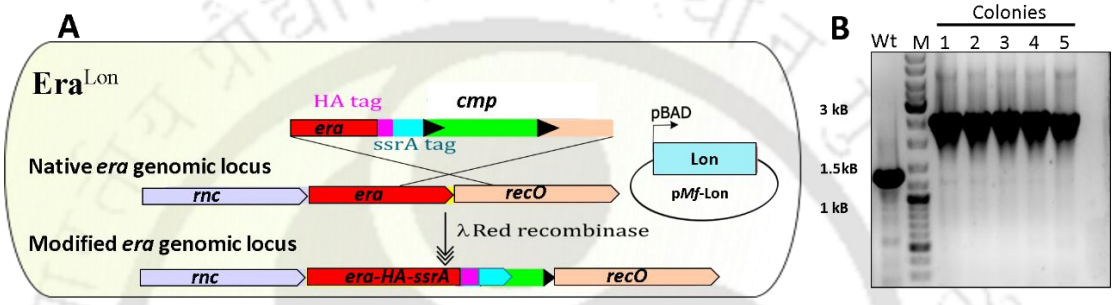

Figure 4.1: Creation of a synthetic genetic circuit for regulated depletion of Era

(A) A schematic representation of the methodology used to import the *Mf*-ssrA-Lon protease system in *E. coli* for depletion of Era. The HA-ssrA tandem tags integrated onto the C-terminus of Era are depicted. The vector p*Mf*-Lon was introduced into these cells to create Era^{Lon}

(B) A colony PCR confirmation for integration of HA-ssrA-CAT cassette into the *rnc* operon using primers annealing to 200 bases upstream and downstream of *era*. The parent cells, Wt is shown with a native locus around 1.4 kB while the Era^{m} cells containing the modified *rnc* operon around 2.5 kB are numbered 1-5 kB The DNA marker is loaded in the lane marked M.

4.3.2 Selective Era depletion arrests cell growth

In order to test the engineered protein depletion system, we decided to investigate if Era depletion had any effect on cell growth. The depletion of RAFs is known to confer cold sensitivity and slow growth phenotypes (Stokes et al., 2014a). Similarly, mutants of Era have been known to display a slow growth phenotype and disturbed cellular morphology due to its probable involvement in cell cycle regulation (Lerner and Inouye, 1991). Thus to confirm Era depletion, Era^{m} cells were transformed with p*Mf*-Lon to create the strain Era^{Lon} (Fig. 4.1A) and Lon production was induced using varying concentrations of arabinose, followed by measurement of growth rate. The uninduced cells displayed a normal growth profile with a

doubling time of about 36 mins, but to our surprise, Era^{Lon} when induced with arabinose concentrations as low as 1 pM failed to show any growth even after 5 hrs of incubation at 37 °C (Fig. 4.2A). These observations clearly indicate that Lon protease in Era^{Lon} cells leads to Era depletion resulting in a slow growth phenotype. The absence of a growth defect in uninduced cells also reinforces the specificity of protein degradation.

Excited by these observations, we wanted to test whether depletion of Era would perturb cell division in a manner as described previously (Lerner and Inouye, 1991). Wt^{Lon} or Era^{Lon} cells were grown in presence of arabinose and Gram stained for microscopic examination. In line with the previous investigations performed with Era mutants, Era depleted cells were severely elongated as compared to Wt^{Lon} (Fig.4.2B). The Wt cells displayed a homogenous population of small rod-shaped cells while the Era depleted cells displayed a heterogeneous population of elongated rods that seemed to be arrested in various stages of growth or cell division (Fig 4.2B).

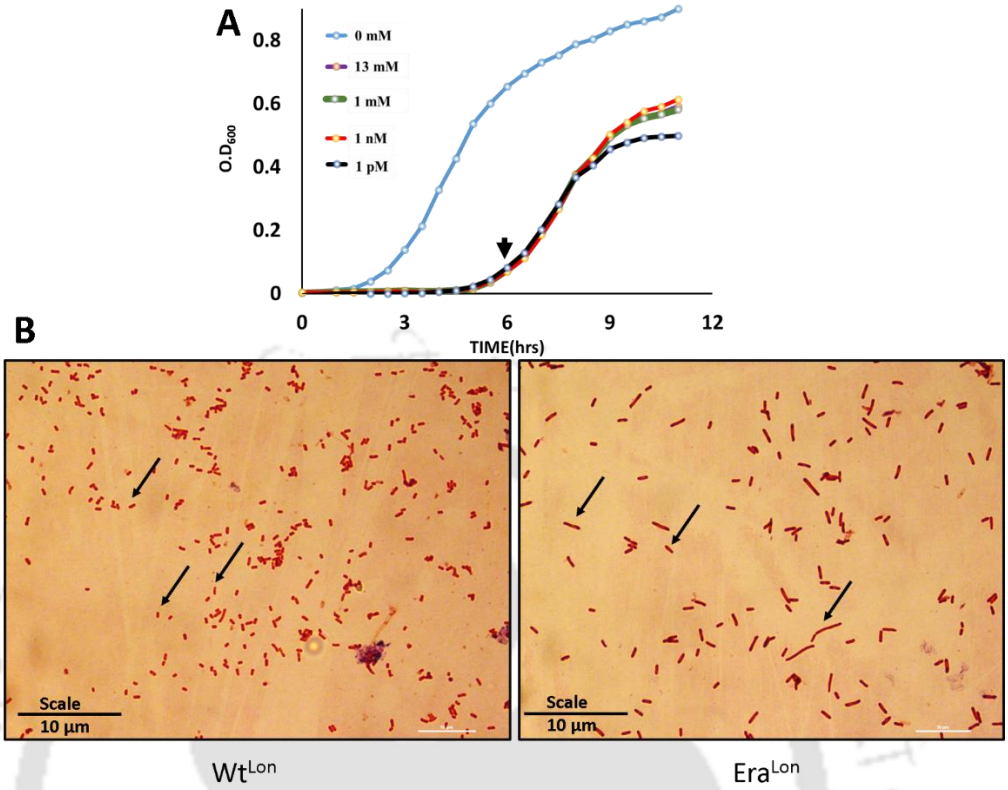

Figure 4.2: Targeted Era depletion arrests cell growth

(A) The growth curve for Era^{Lon} cells producing Lon upon induction with varying concentrations of arabinose. The growth of uninduced cells is shown as a control.

(B) Microscopic examination of Gram-stained Wt^{Lon} or Era^{Lon} cells. In Wt cells, a uniform distribution of rod-shaped cells is highlighted with arrows. Varied stages of growth arrest is also highlighted for Era^{Lon}.

4.3.3 Suppressor mutants overcome Era depletion

Era depleted cells displayed a drastic growth defect. However, after prolonged incubation (about 6 hrs in test tube conditions) a sudden growth burst was observed (see arrow in Figure 4.1A) and subsequently, this population achieved growth saturation similar to cells lacking Lon protease. Interestingly, such spontaneous derivatives are reminiscent of suppressor mutations that are known to compensate for drastic cellular defects (Agarwal et al., 2015; Gardner et al., 2017; Ramesh et al., 1997). The appearance of such mutants could nullify the defects generated due to Era depletion leading to ambiguous phenotypes. Thus, in order to confirm if the growth burst was indeed due to the rise of suppressor mutants, we

performed another round of growth measurement with stationary phase cultures of Era^{Lon} (referred as Era^{SUP}). We reasoned that a suppressor mutant should grow similar to Era^{m}. In line with the hypothesis, Era^{m} and Era^{SUP} showed similar growth pattern with nearly equal doubling times (Fig. 4.3A). Era^{SUP} induced for Lon production did not display the initial lag in growth as previously observed for Era^{Lon} cells (Fig. 4.3A) indicating that Era^{SUP} were neutral to Lon mediated Era degradation and may indeed be suppressors of Era depletion.

4.3.4 Suppressor mutants were traced back to the *lon* containing plasmid

Next, in order to understand the possible origins of the suppressor mutants, we postulated three possible scenarios leading to compensatory mutations: (i) Disruptive mutations in p*Mf*-Lon leading to repression of *lon* or inactivation of Lon protease. (ii) Mutations in the Era-ssrA tag that avert its recognition and subsequent degradation by Lon protease (iii) Compensatory mutations in the genome that could nullify the effect of Era depletion.

In order to resolve the above scenarios, we first tried to test if the suppressors arose due to disruptive mutations in the p*Mf*-Lon. Towards this, we isolated p*Mf*-Lon from Era^{SUP} cells and introduced into Era^{m} to create the strain Era^{Lon*}. Era^{Lon*} cells induced for Lon production grew similar to the uninduced Era^{Lon*} as well as the parent Era^{m} cells (Fig. 4.3B), indicating that the presence of p*Mf*-Lon had little or no effect on cell growth (Table 4.1). This could be attributed to mutations that disrupt *lon* expression (Belin, 2003).

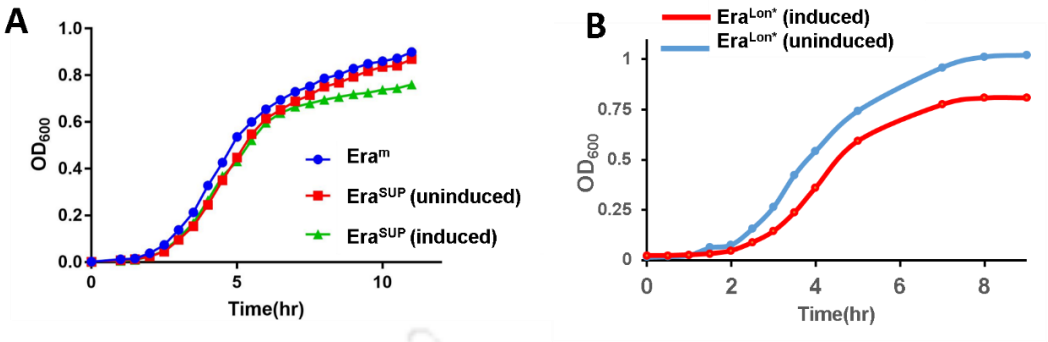

Figure 4.3: Suppressor mutants compensate Era depletion

(A) Growth comparison for Era^{SUP} (induced or uninduced) with parent Era^{m} cells. Induced Era^{SUP} cells and parent cells are seen to grow in a similar manner.

(B) A growth comparison between Era^{Lon*} grown in presence or absence of the inducer. Cells seem to be neutral to Lon production indicating a suppressor.

Table 4.1: Doubling times for suppressor mutants in presence or absence of Lon
(related to figure 4.3B)

Strain	Doubling time (mins)
Era^{Lon*}(uninduced)	36.4
Era^{Lon*}(induced)	38.4

4.3.5 Multiple passaging with shortened log phase averts the rise of suppressor mutants

The characterization of the growth pattern of Era depleted cells resulted in two main observations: (i) The culture produced negligible biomass in initial phases of growth (ii) This slump in growth was followed by a sudden burst due to the appearance of suppressor mutants. The two growth patterns did not yield cells that would be amenable to characterization. In order to avoid this, it was essential to device an alternate strategy to deplete Era without allowing the rise of compensatory mutations. To achieve this, we took cues from our observations that the rise of suppressor mutants was prominent just after Era depleted cells entered log phase. We reasoned that the appearance of suppressor mutants could be avoided by providing multiple passages to the culture, thus shortening or completely avoiding the entry of cells into the log phase (Fig. 4.4A) and in turn avoiding the outgrowth of suppressor mutants.

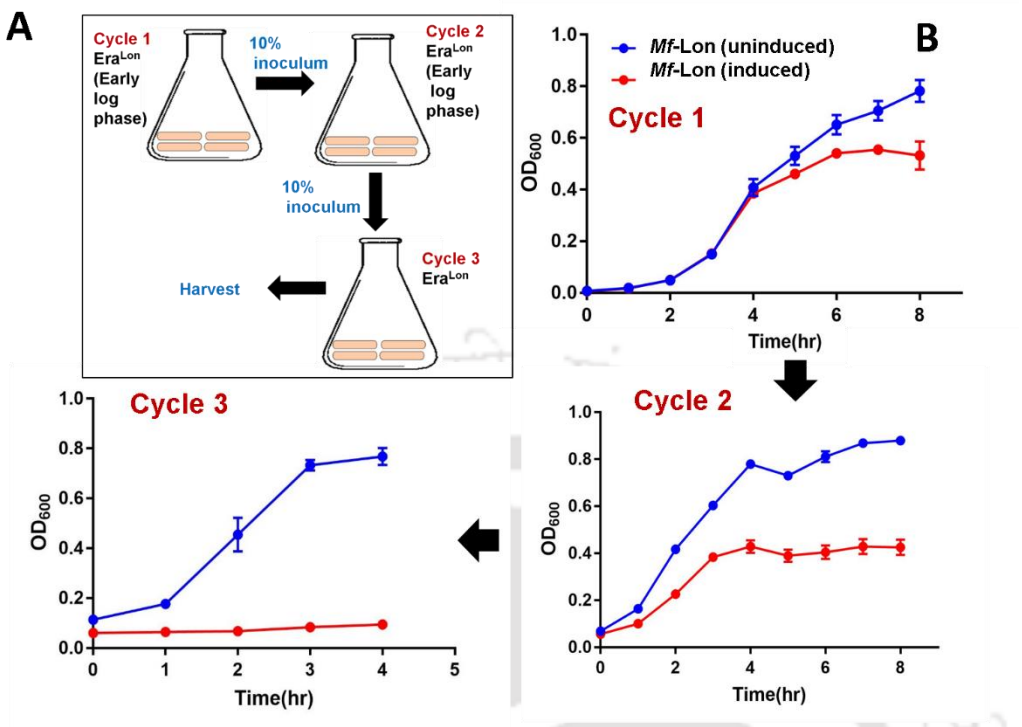

Figure 4.4: A shortened log phase averts the rise of suppressor mutants

(A) A schematic representation of the growth scheme used to shorten the log phase in Era depleted cells

(B) Growth curve for Era^{Lon} grown in presence or absence of the inducer for 3 cycles for a shortened log phase. Cycle 1 and 2 were harvested at Time = 2 hrs followed by the harvest of Cycle 3 cells at Time = 4 hrs.

Table 4.2: Doubling times observed for Era^{Lon} for 3 consecutive cycles of dilution (Related to figure 4.4)

Dilution cycle	Doubling time for Era^{Lon} (uninduced)	Doubling time for Era^{Lon} (induced)
1	39.9	41.5
2	45.5	62.3
3	60	~600

(Time in mins)

In order to test this method, we cultured Era^{Lon} cells till the early log phase followed by harvest in an aseptic environment. These early log phase cells were then used to inoculate a second cycle of growth followed by harvesting when the culture entered log phase. The repeated dilution of cells was continued for another cycle (Fig. 4.4A). Interestingly, this dilution displayed a consistent fall in growth rates over every dilution cycle for induced Era^{Lon}

(Fig. 4.4B). In the third cycle, a stable population of cells, with a doubling time of approximately 600 mins was obtained (Fig. 4.4B, Table 4.2). It is also noteworthy that, if the third cycle of growth was prolonged, suppressor mutants were eventually obtained. The biomass obtained from the initial growth phase of the third dilution cycle were ample for characterization and thus no further dilution cycles were performed. It is also worth mentioning that the uninduced Era^{Lon} cells also displayed a steady increase in the doubling time over each dilution cycle which could be attributed to low, leaky expression of Lon protease.

4.3.6 Era depleted cells display compromised translation capacity

Once a stationary population of Era depleted cells was obtained, we wanted to evaluate the effect of depletion on ribosome maturation. Towards this, we first performed ribosome profiling to study the ribosome population in Era depleted cells. The Wt^{Lon} cells displayed a normal ribosome distribution with significant amounts of 70S particles and minor amounts of 30S and 50S particles (Fig.4.5A). To our surprise, like Wt^{Lon}, Era^{Lon} cells also displayed a significant fraction of 70S particles and minor quantities of 30S and 50S population (Fig. 4.5B). These observations were in stark contrast with the observed growth defect. In order to understand these contrasting observations, we tested the translation capacities of the Era depleted cells by studying the protein production kinetics of β-galactosidase (bgal) (Dalbow and Young, 1975; Schleif et al., 1973). Towards this, Wt^{Lon} or Era^{Lon} cells were transformed with a vector carrying an IPTG inducible copy of bgal encoding gene. Wt^{Lon} cells displayed significant protein build up within minutes of protein induction (Fig.4. 5C). Surprisingly, Era^{Lon} failed to show any build-up of bgal activity (Fig. 4.5C), indicating lack of protein production in these cells. These observations indicate that Era depletion somehow compromises the translation capacity of cells. These observations also explain the arrest of cell growth even with a significant 70S population.

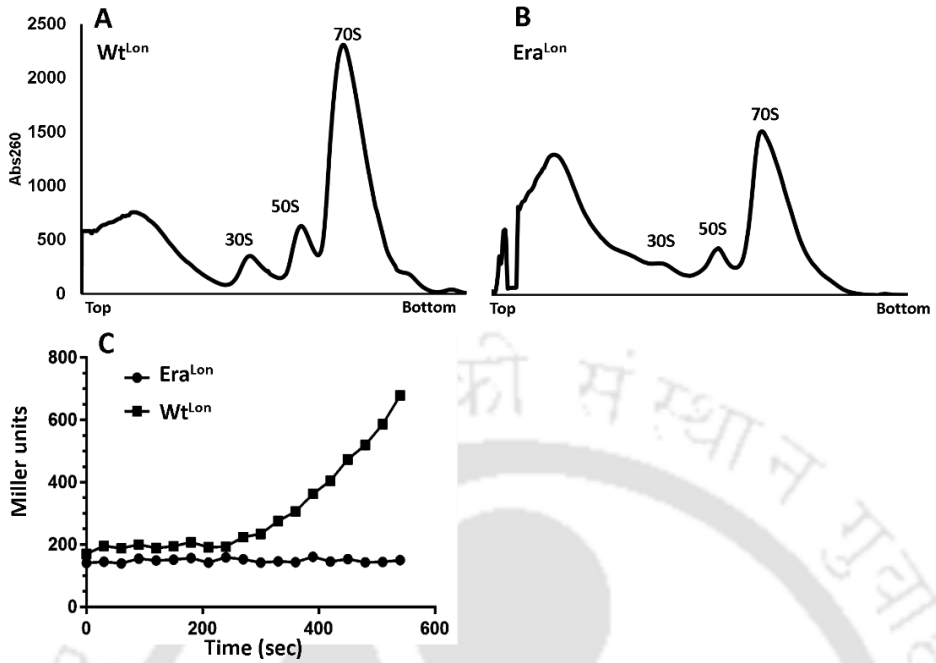

Figure 4.5: Era depletion compromises cellular translation capacity

(A) Ribosome profiles for Wt^{Lon} cells. Peaks observed for the respective particles have been indicated.
(B) Ribosome profiles for Era^{Lon} cells. Peaks observed for the respective particles have been indicated.
(C) Translation kinetics for the production of bgal in Wt^{Lon} or Era^{Lon} cells. The activity measured in Miller units (Y axis) is traced over time (X-axis).

4.4 Discussion

In this part of the thesis, we presented a methodology to conditionally deplete essential proteins that can enable functional characterization in a true physiological setting. Utilizing components of the *M. florum* ribosome rescue system (Gur and Sauer, 2008b; Janssen and Hayes, 2012), we offer a proof-of-concept by depleting the essential RAF Era in a tunable manner. The strategy relies on the use of the ssrA tag which, when recognized by the *Mf*-Lon protease triggers degradation of the tagged protein. Conventional methods used to deplete essential genes involve depletion by modifying the regulatory regions of the genes of interest (Kaufmann and Knop, 2011). These engineered regulatory regions provide altered expression levels or the possibility of conditional expression using appropriate inducers or triggers (Cruz-Bustos et al., 2017; Minden et al., 2000; Petrillo et al., 2014; Seeliger et al.,

2012). In addition, emerging methods to block gene expression in bacteria employ Cas9 mediated transcription repression (Peters et al., 2016). The Cas9-targeting RNA complex binds to the promoter regions of the target gene occluding the entry of RNA polymerase thus stalling transcription. The aforementioned methods although established, produce artefacts like off-target hits or creates a polarity effect, affecting the expression of downstream genes. Additionally, the methods require extensive optimization for designing target binding RNA. Overcoming these difficulties, the methodology devised here, mediates targeted protein exhaustion after its translation, leaving the other genetic circuits and the mRNA intact, avoiding leaky expression or any polarity effects. The relatively simple method required the creation of a simple genetic circuit that is turned ON by adding an inducer to trigger the production of Lon protease. Based on previous reports, the system can achieve nearly complete degradation of a protein within 2.5-4 hrs (Cameron and Collins, 2014) of Lon protease induction. However, given the culture conditions, the authors did not find viable cell counts after depletion of essential genes. A major advancement made in our study lies in achieving stable populations of cells lacking the essential proteins. Our protein based depletion strategy circumvents undesirable side effects of Era depletion previously achieved by deleting the *rnc* locus (Britton et al., 1997; Britton et al., 2002) and produces a stable cell line for focused study. Secondly, Lon mediated depletion also avoids the previously reported pleiotropic effects rising due to a collateral depletion of RNase III, an enzyme essential for processing of rRNA, tRNA and some mRNAs in *E. coli* (Nikolaev et al., 1974; Régnier and Grunberg-Manago, 1989, 1990).

Although currently it is difficult to comment on the exact stage of cellular arrest achieved here because of the unresolved nature of cell cycle regulation in bacteria (Zheng et al., 2016). The microscopic examination hints towards a mixed population which consists of severely elongated as well as cells in two cell stage (Fig. 4.2) indicating disruption of the cell division mechanisms. Further, the phenotypic characterization (Fig. 4.2) and translation kinetics (Fig. 4.5) of Era deficient cells reinforced its role in translation regulation, ribosome biogenesis as well as cell cycle progression, which are in line with the previous reports(Britton et al., 1997; Britton et al., 2002; Lerner and Inouye, 1991). However, the translation arrest observed despite a significant 70S population is indeed intriguing and calls for structural characterisation of the 70S particles accumulating upon Era depletion. It is possible that these

particles contain structural defects and thus are unable to engage in protein synthesis. Alternatively, it is also plausible that, due to the extensive regulatory mechanisms that govern translation under biotic or abiotic stress, a pool of 70S particles are sequestered from protein synthesis. Such ribosomes would also be recalcitrant to ribosome degradation pathways and thus can be employed for restoring protein synthesis when stress is relieved. A tantalizing fate of such ribosomes could be to drive protein synthesis during the rise of suppressor mutants that were ultimately observed in Era depleted cells.

Here, our initial efforts have utilized an arabinose operated genetic circuit but greater control of *lon* expression can be achieved using other inducers like rhamnose, which confer greater linearity between inducer concentration and expressed protein (Giacalone et al., 2006; Kelly et al., 2018). In conclusion, this chapter marks the completion of the pipeline for RAF discovery and characterization. The methodology developed in this chapter complements the BioID and BiFC pipeline by creating conditional null mutants of essential RAFs.

4.5 **Summary**

A thorough characterization of RAFs involves studying ribosome assembly intermediates that accumulate in cells devoid of the respective RAFs. However, this analysis is not feasible for RAFs that are essential for survival of an organism. Hence, in this chapter we have attempted to address this problem by developing a methodology to specifically degrade essential RAFs using a Lon protease. We demonstrate regulatable degradation of Era to achieve a stable population of Era deficient cells. These cells displayed a severe slowdown in growth and protein synthesis further highlighting its role in cell cycle control and regulation of protein synthesis. This Lon mediated depletion of essential RAFs can complement the BioID and BiFC mediated characterization of RAFs in bacteria.

Chapter V

Characterization of ribosome quality control mechanisms in bacteria

The work embodied in this chapter is under preparation for publication.

Chapter 5

5.1 Introduction

Our efforts to revisit the discovery and characterization of RAFs have been set in motion in the form of assay systems presented in the preceding chapters. However, another aspect of ribosome biogenesis that is widely unaddressed is the physiological implications of a defective assembly process. In contrast to the precise allocation of cellular resources in healthy cells (Li et al., 2014), inefficient conversion of premature ribosomes to functional particles may lead to grave flux imbalances in a cell. To counter this, cellular resources need to be rerouted towards slow maturation of ribosomes by alternate pathways and a concomitant decrease in the components of the translation machinery.

In bacterial cells, assembly defects arising due to deletion of RAFs are usually seen to be manifested as compromised growth phenotype, sensitivity to cold temperatures as well as accumulation of ribosome assembly intermediates (Alix and Guerin, 1993; Boehringer et al., 2012; Britton, 2009b; Bylund et al., 1998b; Charollais et al., 2004; Charollais et al., 2003c; Davis and Williamson, 2017; Guthrie et al., 1969; Shajani et al., 2011). However, little is known about the metabolic cost of harbouring premature subunits and their ultimate fate in bacterial cells. Although studies have shed light on how cells regulate the number of active ribosomes and maintain fitness in nutrient-limiting conditions (Dai et al., 2018; Dai et al., 2017; Koch and Deppe, 1971; Neidhardt and Magasanik, 1960; Towbin et al., 2017), such knowledge is virtually unknown for cells with compromised ribosome structure and composition.

It is believed that subunit association may be impaired in premature ribosomes which lack fully formed bridging sites and this, in turn, may block their entry into translation (Boehringer et al., 2012; Datta et al., 2007; Razi et al., 2017). In addition, it has been observed that mutations in rRNA, lack of ribosomal proteins or the absence of assembly factors like KsgA and Hfq induces defects during translation (Andrade et al., 2018; Connolly and Culver, 2013; O'Connor et al., 2004; Qin et al., 2012). Studies have also hypothesized that perturbed ribosome maturation may create a bottleneck that decreases new translation initiation events

(Gibbs et al., 2017). This would result in the suppression of translation or alternate scenario wherein premature subunits may compete with mature subunits for the mRNA thus averting translation once again. The spectrum of studies suggests that participation of premature ribosomes in translation may have implications that affect the cellular fitness. However, mechanistic details of such relationships are still poorly defined. Thus, we set out to address some of the primary outstanding questions: Do premature ribosomes enter translation? If so, how such ribosomes bypass the quality control mechanisms in bacteria? To address these questions, we initiated this study with a hypothesis that premature ribosomes entering translation may display defects in decoding (30S subunit) or peptide bond synthesis (50S subunits). Herein, we probed how the assembly defects conferred upon by the deletion of 30S subunit specific late-stage assembly factors influence translation initiation and elongation. We also investigated the mechanism by which premature ribosomes elude translation quality control processes in bacteria. Thus by probing the participation of premature ribosomes in the translation process, it would be possible to appraise the relevance of ribosome quality control checkpoints in bacteria.

5.2 Materials & Methods

5.2.1 Creation of strains and plasmids

The Keio collection parent strain BW25113 referred as Wt was used as the parental strain for all genetic manipulations and reference measurements. Null mutant for LepA was procured from the Coli genetic stock centre (CGSC). Additionally, null mutants for RbfA, RsgA and KsgA were created using λ Red recombineering method (Datsenko and Wanner, 2000c).

A gene encoding GFP, *gfp* was amplified from the vector 1GFP (Addgene #29663) with four different start codons i.e. AUG, CUG, AUA and AGG. Additionally, frameshift mutations (+1 base and -1 base) were introduced at codon 7 of the GFP construct using oligonucleotides. Similarly, codon 7 and 8 were replaced with UAG and UAA codons. The modified GFP constructs were individually cloned into vector 8R (Addgene # 37506) using Xba1/Nhe1 and BamH1 specific restriction sites. Similarly, the constructs used for studying

translation kinetics were created by amplifying the gene encoding β-galactosidase (bgal) from *E. coli* B cells. Modifications were also introduced into this gene by incorporating different start codons (O'Connor, 1997; O'Connor et al., 2004). The amplified cassettes were cloned into vector pQE2 using XhoI/HindIII sites.

In order to complement assembly factors (RAFs) or initiation factors (IFs), respective genes were amplified from Wt cells and cloned into a p15A vector backbone under an Anhydro-tetracycline (Atc) inducible promoter.

5.2.2 Growth analysis

Primary inoculums were always prepared by growing freshly transformed colonies in LB medium with the respective antibiotics at 37 °C, with shaking at 180 rpm. The next day, OD_{600} of primary cultures was normalized. These cultures were then diluted into 1 ml fresh LB medium with appropriate antibiotics in a 24 well plate followed by incubation at 37 °C with regular shaking. OD_{600} measurements were taken in Tecan infinite M200 multimode plate reader at every 30-minute interval until cultures reached saturation. Additionally, in order to study growth in presence of elevated levels of IFs or RAFs, LB medium was supplemented with 50 ng/ml Atc at time of inoculation to trigger protein production. The experiment was repeated at least three times to derive the average growth curves.

5.2.3 GFP fluorescence measurements and calculation of indices

Wt or null mutant were freshly transformed using the desirable plasmid carrying the gene encoding GFP before initiating the experiments. Colonies were picked in fresh LB medium supplemented with 100 µg/ml ampicillin (Amp) and allowed to grow at 37 °C and 180 rpm till OD_{600} reached 0.6. At this point, *gfp* expression was induced by adding 2 mM arabinose and cells were allowed to grow for another 3 hrs. Upon completion, cells were harvested and lysed in Buffer G (20mM Tris-HCl (pH 7.5 at 25 °C), 500 mM NaCl, 1 mM EDTA, 1 mM PMSF, 6 mM β-ME) supplemented with 1X CelLytic B solution (Sigma Aldrich). The lysates were normalized for total protein content and were taken for fluorescence measurements. All fluorescence spectrums were generated by exciting at 488 nm

and scanning for emission from 500-600 nm, with averaging over three scans after baseline correction in a FluoroMax-4 spectrofluorometer (Horiba Scientific, Edison, NJ). The slit width used for excitation and emission was 2 nm and 7 nm, respectively. All fluorescence experiments were performed with five independent trials. For measurement of Initiation Error Index, GFP emission from null mutants was compared with that of Wt cells. Similarly, Rescue Index was calculated as the ratio of GFP fluorescence after complementing the null mutants or Wt cells. Briefly, cells were co-transformed using plasmids carrying genes coding for GFP and assembly or initiation factors, respectively. The cells were grown on a dual selection marker at 37 °C and induced with Atc to trigger production of the RAFs or IFs when OD_{600} reached 0.3. Subsequently, GFP production was induced with 2 mM arabinose when OD_{600} of the Atc induced cells reached 0.6. Further, the cells were processed for fluorescence measurements as described above. The indices were defined as following:

$$Initiation\ Error\ Index = \frac{\text{GFP expression in null mutants}}{\text{GFP expression in Wt}}$$

$$Initiation\ Rescue\ Index = \frac{\text{GFP expression in complemented null mutants}}{\text{GFP expression in complemented Wt}}$$

5.2.4 Determination of initiation and elongation rates from β-galactosidase (bgal) production kinetics

In order to calculate the translation initiation and elongation rates, the bgal production assay (Dalbow and Bremer, 1975; Dalbow and Young, 1975; Miller, 1992) was performed with minor modifications. Briefly, Wt or null mutants were transformed using respective constructs that encoded bgal (variants containing different start codons: pAUG-bgal or pAGG-bgal). Overnight grown cultures were diluted into fresh 5 ml LB medium supplemented with 100 µg/ml ampicillin and allowed to grow at 37 °C and 180 rpm till OD_{600} reached 0.6. These cultures were then shifted to a water bath at 37 °C or 25 °C (RT). Bgal production was induced with 1 mM IPTG followed by instant mixing for 5 secs. Following this, aliquots of 200 µL each were taken every 30 secs for the assays done at 37°C and every 120 secs for assays done at RT. The aliquots were added to tubes containing 300 µL of Z-

buffer (60 mM $Na_2HPO_4 \cdot 2\ H_2O$, 40 mM $NaH_2PO_4 \cdot\ H_2O$ (pH 7.0 at 25 °C), 10 mM KCl, 1 mM $MgCl_2$, 5 mM β-ME), 20 µL of chloroform, and 10 µL of 0.1% SDS. The tubes were then instantly vortexed for 20 secs and transferred to ice. At the end, all the tubes were incubated for 5 mins on ice and then for 5 mins at RT after which, 500 µL of 1 mg/ml ONPG (o-nitro-phenyl galactopyranoside) was added to the reaction. ONPG hydrolysis for cells with pAUG-bgal plasmid was allowed to proceed for 15 mins at RT. Similarly, for cells with pAGG-bgal the reaction was allowed to proceed for 1 hour at RT. After completion, the reaction was stopped by adding 500 µL of 1 M Na_2CO_3. Colour was allowed to develop for 15 mins and then 300 µL of the reaction mix was transferred to 96 well plates. Absorbance was recorded at 420 nm and 550 nm using a Tecan infinite M200 multimode plate reader. The readings were used to derive the plot representing residual enzymatic activity for calculating elongation rates and rates of initiation (Schleif et al., 1973). Length of the protein (1,024 amino acids) divided by the lag time for the appearance of enzymatic activity after induction was used to calculate the elongation rates. The square root of the residual enzyme activity ($\sqrt{(E_t-E_0)}$, where E_t signifies the miller units at time t and E_0 signifies the miller units recorded at t=0) was plotted against time (t) to derive a linear plot for calculating the rate of translation initiation.

5.2.5 Polysome profiling

Wt, *ΔrbfA* or *ΔksgA* cells transformed with a plasmid carrying the gene *rbfA*, *ksgA*, *infA*, *infB* or *infC* were grown overnight in LB medium supplemented with 25 µg/ml Chloramphenicol (Cmp). The next day, cultures were diluted in 100 ml fresh LB medium supplemented with Cmp and grown at 37 °C with vigorous shaking. After OD_{600} reached 0.3, the protein production was induced with Atc (50 ng/ml) and allowed to grow till the OD_{600} of the cultures was 0.6. At this point, the culture was chilled rapidly on ice with addition of ice cubes and 200 µg/ml Cmp following incubation for 15 mins. The cells were harvested by centrifugation at 4 °C, 8000xg for 10 mins. Cells were then washed once with Buffer CL (20 mM Tris-HCl (pH 7.6 at 25 °C), 150 mM NH_4Cl, 10.5 mM $Mg(OAc)_2$, 0.5 mM EDTA and 6 mM β-ME) and later re-suspended in 1 ml buffer CL supplemented with 1 mg/ml Lysozyme, 1 mM PMSF, 0.5x CellLyticB reagent followed by incubation on ice for 1 hour (hr). The cells were then ruptured by 5 cycles of freeze-thaw using liquid nitrogen. The lysates were clarified

by centrifugation at 30,000xg followed by A_{260} (Absorbance at 260 nm) quantification in Implen Nanospectrophotometer. For ultracentrifugation, 300µL of 10 A_{260} was loaded on a 10-50% sucrose gradient prepared in buffer CL and spun at 102,000xg in rotor TH-641 (Thermo-Sorvall WX^+ ultracentrifuge) for 16 hrs. Subsequently, fractionation was performed using a syringe pump connected to Akta Pure system (GE healthcare).

5.2.6 Growth, Crosslinking and cell lysis conditions for TCP-seq

Preparation of ribosome and initiation complex footprints were done essentially according to the original protocol (Shirokikh et al., 2017) with minor modification as required for bacterial cells. Overnight cultures of Wt, *ΔrbfA* or *ΔksgA* cells carrying the plasmids pAGG-GFP-8R were diluted into 2 litres of fresh LB medium supplemented with 100 µg/ml Amp and incubated at 25 °C with shaking at 180 rpm. Minimal *gfp* expression was induced using 50 µM arabinose and growth was allowed to proceed till OD_{600} 0.4. At this point, cells were rapidly chilled by adding ice cubes accompanied by addition of 150 ml freshly prepared 30% formaldehyde. Cultures were mixed vigorously and placed on ice for 15 mins followed by addition of 200 ml 2.5 M glycine. Following this, cells were harvested by centrifugation at 8000xg for 10 mins. Pellets were washed with 40 ml of Buffer TW (20 mM HEPES-KOH (pH 7.4 at 25 °C), 100 mM KCl and 2 mM $MgCl_2$) followed by resuspension in 0.4 ml of buffer TL (20 mM HEPES-KOH (pH 7.4 at 25 °C), 100 mM KCl, 2.5 mM $MgCl_2$ and 0.5 mM EDTA) supplemented with 2 mg/ml lysozyme, 6 mM β-ME, 1 Unit/µl RNaseOUT inhibitor (ThermoFisher Scientific), 40 Units DNase and 1 mM PMSF. Cells were placed on ice for 1 hr followed by rupturing using 5 cycles of freeze-thaw in liquid nitrogen. Cell lysates was clarified by centrifugation at 30,000xg for 30 mins at 4 °C.

5.2.7 Isolation of mRNA bound fractions and MNase treatment

Clarified cell lysate were layered onto a 10-20% sucrose gradient prepared in Buffer TP (50 mM Tris-HCl (pH 7.0 at 25 °C), 50 mM NH_4Cl, 4.5 mM $MgCl_2$, 0.5 mM EDTA, 6 mM β-ME) and centrifuged at 52,000 rpm for 80 mins at 4 °C using a TLS-55 rotor in an Optima TLX 120 ultracentrifuge (Beckman-Coulter). Ribosome pellets derived after centrifugation were re-suspended in 1 ml of buffer TL supplemented with 1 Unit/µl RNaseOUT and 1 mM PMSF. These ribosomes were then treated with 30 Units of MNase per 1 A_{260} of ribosomes for 1 hr at RT. The reaction was stopped by addition of 2 mM EGTA and the nuclease treated complexes was immediately loaded onto a 10%-40% sucrose gradient prepared with buffer TP. Samples were spun at 40,000 rpm for 4 hrs at 4 °C using a TLS-55 rotor. Upon completion, fractions of 100 µl each were collected manually and later processed for footprint isolation.

5.2.8 De-crosslinking and footprint isolation

Fractionated 30S or 70S fractions were supplemented with 1% w/v sodium dodecyl sulphate (SDS), 10 mM EDTA, 10 mM Tris-HCl (pH 7.4 at 25 °C), 10 mM glycine and 1 mg/ml Proteinase K (Sigma-Aldrich). The samples were incubated at 50 °C for 30 mins with regular mixing. Following this, equal volume of 5: 1: acid phenol: chloroform was added to this mix and incubated at 65 °C for another 45 mins with regular vortexing. The sample was then centrifuged at 16,000xg for 30 mins at 4 °C followed by separation of the aqueous phase. Equal volume of chloroform was then added to the aqueous phase and centrifuged at 16,000xg for 30 mins at 4 °C. The aqueous phase was again separated in a fresh tube and supplemented with 0.1 volume of 3M sodium acetate (pH 5.2 at 25 °C), 1 mg/ml glycol blue (Ambion) and 2.5 volumes of chilled ethanol. Precipitation was done overnight at -20 °C following which the sample was centrifuged at 16,000xg for 45 mins at 4 °C. RNA pellets were washed twice with chilled 70% ethanol followed by air drying. Finally, the pellets were re-suspended in nuclease free water and checked for integrity. Footprints corresponding to 30S or 70S fractions were size selected from 10-80 bases by gel elution using denaturing Polyacrylamide

gel electrophoresis (PAGE) in Protean ii XL setup (Bio-Rad). The extracted RNA was end repaired using T4 Poly nucleotide kinase (PNK) and purified by ethanol precipitation.

5.2.9 Library preparation and sequencing

Library preparation and sequencing was outsourced to Genotypic Technology Pvt. Ltd., Bengaluru, India. Libraries were prepared strictly according to the manufacturers recommendations using the TruSeq Small RNA Sample Preparation kit (Illumina, U.S.A). Briefly, 50 ng of RNA was used as starting material for ligation of 3' and 5' adapters. Specific index sequence was added to each sample for identification during sequencing. The Illumina Universal Adapter used in the study was:

AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGA and the Index Adapters were

CAAGCAGAAGACGGCATACGAGAT[INDEX]GTGACTGGAGTTCCTTGGCACCCG AGAATTCC.

Adapter ligated fragments were reverse transcribed with Superscript III Reverse transcriptase (Invitrogen). The cDNA was enriched and barcoded by 15 cycles of PCR amplification and the amplified library was size selected using denaturing PAGE. The library was size selected in the range of 140 bp– 210 bp followed by overnight gel elution and finally re-suspended in nuclease free water. Illumina compatible sequencing library was initially quantified by Qubit fluorimeter (Thermo Fisher Scientific) and its fragment size distribution was analysed on Agilent TapeStation. Sequencing was performed on the Illumina NextSeq 500 platform.

5.2.10 Analysis of sequencing data

The reads were subjected to several pre-processing steps described as follows. Firstly, reads with Phred score less than 20 were removed by utilizing fastq_quality_trimmer from the FASTX-toolkit-version-0.0.13. The remaining reads were trimmed for 3'-end adapter sequences [5'-TGGAATTCTCGGGTGCCAAGGAACTC-3'] using Cutadapt-1.5 (Martin, 2011). Following these, in order to filter and remove the reads derived from rRNA, tRNA,

sRNA or any other non-coding RNA, reads were aligned using STAR-2.5.3 (Dobin et al., 2013) against non-coding RNAs from *E. coli* K-12 MG1655 genome (ftp://ftp.ncbi.nlm.nih.gov/genomes/refseq/bacteria/Escherichia_coli/reference/GCF_000005845.2_ASM584v2/GCF_000005845.2_ASM584v2_rna_from_genomic.fna). The unmatched reads were aligned against the *E. coli* K-12 MG1655 reference genome (NCBI: NC_000913.3;ftp://ftp.ncbi.nlm.nih.gov/genomes/refseq/bacteria/Escherichia_coli/reference/GCF_000005845.2_ASM584v2/GCF_000005845.2_ASM584v2_genomic.fna) using STAR-2.5.3 with the following parameters: spliced alignment turned off (--alignIntronMax 1), forced end to end alignment (--alignEndsType EndToEnd) and allowing 8 mismatches per 100 nt (--outFilterMismatchNoverLmax 0.08). Sam to Bam conversion was accomplished using Samtools-1.4.1 (Li et al., 2009). Only uniquely matching reads were considered for further analyses. Metagene analysis for mapping the reads to START and STOP codons was performed by employing custom written shell scripts utilizing Samtools-1.4.1. Plots were generated using R.

5.3 Results

5.3.1 Defects in late 30S assembly disrupt translation fidelity

We wondered whether the premature 30S subunits bypass the quality control checkpoints to enter the translation cycle. To address this, we created several variants of *E. coli* that harboured null mutants for genes encoding late-stage 30S specific RAFs such as RsgA, RbfA, KsgA and LepA (Fig. 5.1A). Premature ribosomes from these strains were tested for their ability to recognize AUG (cognate) and non-AUG (near cognate and non-cognate) initiation codons for translating the gene encoding GFP *in vivo* (Fig. 5.1B). In order to quantify the extent of error in decoding the start codon due to an assembly defect in the 30S subunit, we formulated two indices: (i) Initiation Error Index (IEI) and (ii) Initiation Rescue Index (IRI). IEI was defined as the ratio of GFP fluorescence from null mutants of the respective RAF to Wild type (Wt). We have chosen four different initiation signals of variable strength: AUG>CUG>AUA>AGG to assess the potency of premature 30S subunits in discriminating cognate from non-cognate codons (Hecht et al., 2017; Sussman et al., 1996).

Deletion of RsgA did not display any increase in the IEI, indicating no defects in the codon recognition (Fig. 5.1C). Deletion of the recently implicated RAF, EF-4/LepA led to misrecognition of CUG and AUA as start codons (Fig. 5.1D). However, expression from AUG and AGG start codons remained unchanged in *ΔlepA* strain (Fig. 5.1D). In order to further investigate if the decoding errors were specifically due to deletion of the respective assembly factor, we complemented the loss of RsgA and LepA from a plasmid-borne copy. Initiation Rescue Index (IRI) was calculated as the ratio of GFP fluorescence in complemented null mutants to that of complemented Wt cells. If the translation errors were indeed caused due to the absence of the respective RAF, a shift from IEI >1 to IRI ~1 would hint towards a rescue from mis-initiation at non-canonical start codons. Upon complementation, IRI~1 indicated that translation defects were ameliorated in *ΔlepA* cells (Fig.5.1D), on the other hand, the IRI remained unaffected for *ΔrsgA* cells (Fig. 5.1C). In contrast, loss of KsgA and RbfA displayed a higher IEI for near-cognate as well as non-cognate start codons indicating significantly compromised initiation codon recognition (Fig. 5.1E & 5.1F). In order to probe this phenomenon further, we also complemented the loss of KsgA and RbfA with a plasmid-borne copy to measure the IRI values. The rescue was indicated by IRI ~ 1 in RbfA and KsgA complemented strains (Fig. 5.1E & 5.1F). These findings highlight the fact that the premature ribosomes accumulating due to loss of LepA, KsgA and RbfA have the potency to enter translation even with a significantly compromised native structure, if the initiation signal is non-AUG.

Given the aberrant effect on the initiation stage exerted by the premature ribosomes, we tried to investigate further if the assembly defects in *ΔrbfA* and *ΔksgA* strains also manifest during translation elongation stage. Towards this, we employed variants of *gfp* that harboured frameshift mutations or premature stop codons and measured the Elongation Error Index (EEI) (Fig. 5.1G). EEI >1 would indicate decoding errors leading to compromised frame maintenance or bypass of stop codons with respect to Wt. Both *ΔrbfA* and *ΔksgA* showed slightly elevated EEI values, indicating that premature ribosomes in both strains have compromised fidelity during elongation and termination. However, the observations also suggest that these bypasses are less frequent in comparison to the erroneous decoding of the initiation signals.

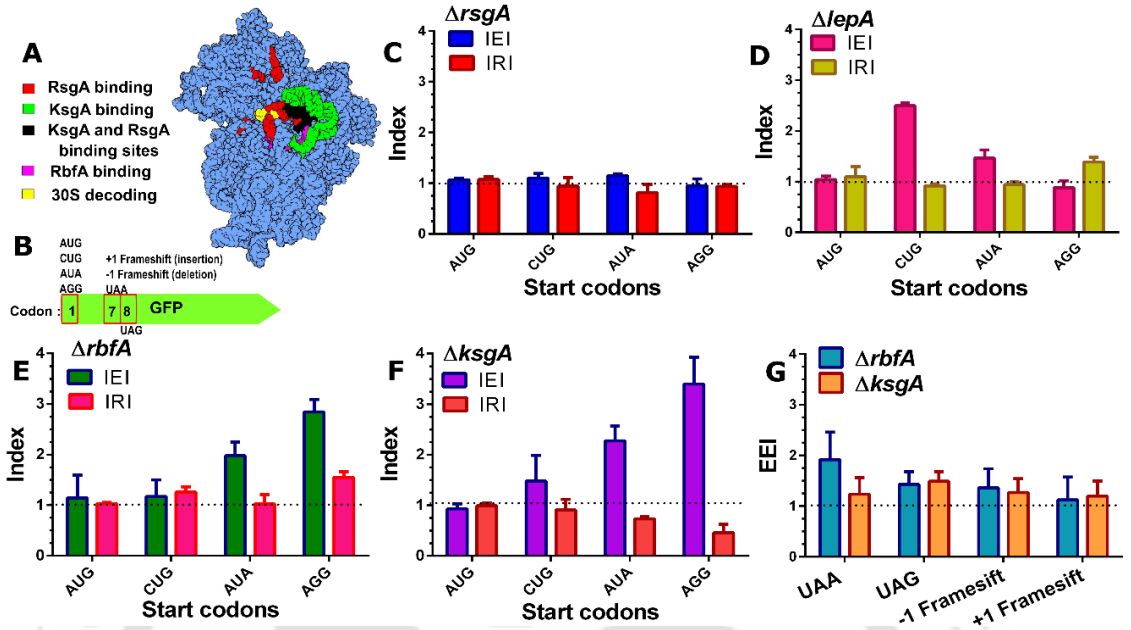

Figure 5.1: Assembly defects in 30S subunit compromise decoding fidelity

(A) Structure of the *E. coli* 30S subunit represented in blue spheres [PDB ID: 4V52] showing the binding sites of late stage acting RAFs as observed by Cryo-Em studies (Boehringer et al., 2012; Datta et al., 2007; Razi et al., 2017) . Colour coding for each RAF binding site is indicated along with the 30S decoding centre.

(B) A representation of the GFP construct used for measuring translation fidelity. Codon variations at codon 1, 7 and 8 are indicated.

(C)– (F) Errors and rescue from misrecognition of start codon, measured as IEI and IRI are shown for *ΔrsgA* (C), *ΔlepA* (D), *ΔrbfA* (E) and *ΔksgA* (F), respectively. Dashed lines marking IEI =1, signifies similar expression in null mutants and Wt, indicating optimal recognition of the start codon. Dashed lines marking IRI =1, signifies rescue from erroneous translation initiation at near or non-canonical start codons.

(G) Evaluation of frame maintenance and fidelity of stop codon recognition measured as EEI, is shown for *ΔksgA* and *ΔrbfA*. The dashed line indicating EEI =1 indicates similar level of GFP expression between null mutants and Wt cells.

5.3.2 Translation defects correlate with assembly defects

In order to further dissect the link between assembly and translation, we studied translation by tracing the kinetics of β-galactosidase (bgal) synthesis *in vivo*. For this, we

employed an IPTG inducible construct, pAUG-bgal (AUG as the start codon). The trace signifying accumulation of bgal was used to calculate the peptide chain elongation rates (ER) as described previously (Dalbow and Young, 1975; Schleif et al., 1973). We measured ERs for Wt, *ΔrbfA* and *ΔksgA* cells at 37 °C (Fig. 5.2A; Table 5.1). Wt cells were found to have ERs that are in close agreement with the previous investigation (Dai et al., 2017). However, the ERs were slightly higher in case of *ΔksgA* cells, whereas *ΔrbfA* displayed an approximately three-fold decrease relative to Wt (Fig. 5.2A; Table 5.1) suggesting that perturbed assembly may indeed upset translation. This was even more striking at 25°C and was in line with the assembly defects getting exacerbated at lower temperature (Fig. 5.2B; Table 5.1).

Using the same kinetics data, we also tried to understand how the assembly defects affect the translation initiation process. We measured the translation initiation rates (IR), as indicated by the slope of the bgal accumulation trace (Fig. 5.2C & 5.2D) (Dalbow and Young, 1975; Schleif et al., 1973). The rate of translation initiation from the AUG start codon, followed a pattern similar to the ERs. Rates for initiation were similar for Wt and *ΔksgA* (Fig. 5.2C; Table 5.2) and much lower in case of *ΔrbfA* at 37 °C (Fig. 5.2C; Table 5.2). In line with ER, IR for Wt, *ΔrbfA* and *ΔksgA* was significantly reduced at 25°C (Fig. 5.2D; Table 5.2). Notably, IR for *ΔrbfA* was much lower, reinforcing that premature ribosomes entering translation in *ΔrbfA* have impaired translation capacity.

It was also important to address a possibility whether the lower ER and IR were specifically due to errors in decoding or just a general slowdown of the translation machinery that may arise due to a deficit in number of mature ribosomes entering translation. In order to address this, we hypothesized that, if the lower ERs and IRs were caused by a general slowdown of the translation machinery, the respective rates should remain unaffected if measured from non-canonical start codons. However, if the decrease was indeed caused by erroneous decoding due to assembly defects, the non-canonical start codons must be misread as canonical ones preferentially in null mutants in comparison to Wt. To test this notion, we performed bgal translation kinetics using constructs with an AGG start codon (pAGG-bgal) at 37 °C. Remarkably, the IR from AGG start codon, displayed an inverse pattern in comparison to AUG start codon for the three strains. Null mutants of RbfA displayed elevated translation initiation events over time in comparison to Wt (Fig. 5.2E & 5.2F; Table 5.2).

Similar reversal, but with a lesser intensity, was also observed for *ΔksgA* (Fig. 5.2F; Table 5.2). These observations strongly indicate that the errors in assembly lead to erroneous translation initiation events. Collectively, these findings hint at a scenario where evasion of quality control checkpoints permits entry of premature ribosomes into translation. Such ribosomes not only attenuate the intensity of the translation process but also elude the quality control checkpoints during translation.

Table 5.1: Peptide chain elongation rates (ER) measured from Wt, *ΔksgA*, and *ΔrbfA* cells using bgal translation kinetics. All measurements are in amino acids/second (aa/sec).

Start codon	Growth temperature	Wt	*ΔksgA*	*ΔrbfA*
AUG	**37 °C**	16.1 ± 2.5	17.9 ± 2.9	4.9 ± 0.5
AUG	**25 °C**	7.4 ± 1.9	3.5 ± 0.5	1.7 ± 0.1
AGG	**37 °C**	16.1 ± 3	9.94 ± 0.7	4.3 ± 0.2

Table 5.2: Translation initiation rates (IR) measured from Wt, *ΔksgA*, and *ΔrbfA* cells using bgal translation kinetics. The IR values are derived as the slope of the curve marking accumulation of bgal over time (units: Miller units$^{1/2}$/sec).

Start codon	Growth temperature	Wt	*ΔksgA*	*ΔrbfA*
AUG	**37 °C**	0.2104 ± 0.008	0.2029 ± 0.007	0.1660 ± 0.008
AUG	**25 °C**	0.05746 ± 0.002	0.06675 ± 0.003	0.04020 ± 0.001
AGG	**37 °C**	0.05594 ± 0.001	0.08348 ± 0.002	0.1270 ± 0.004

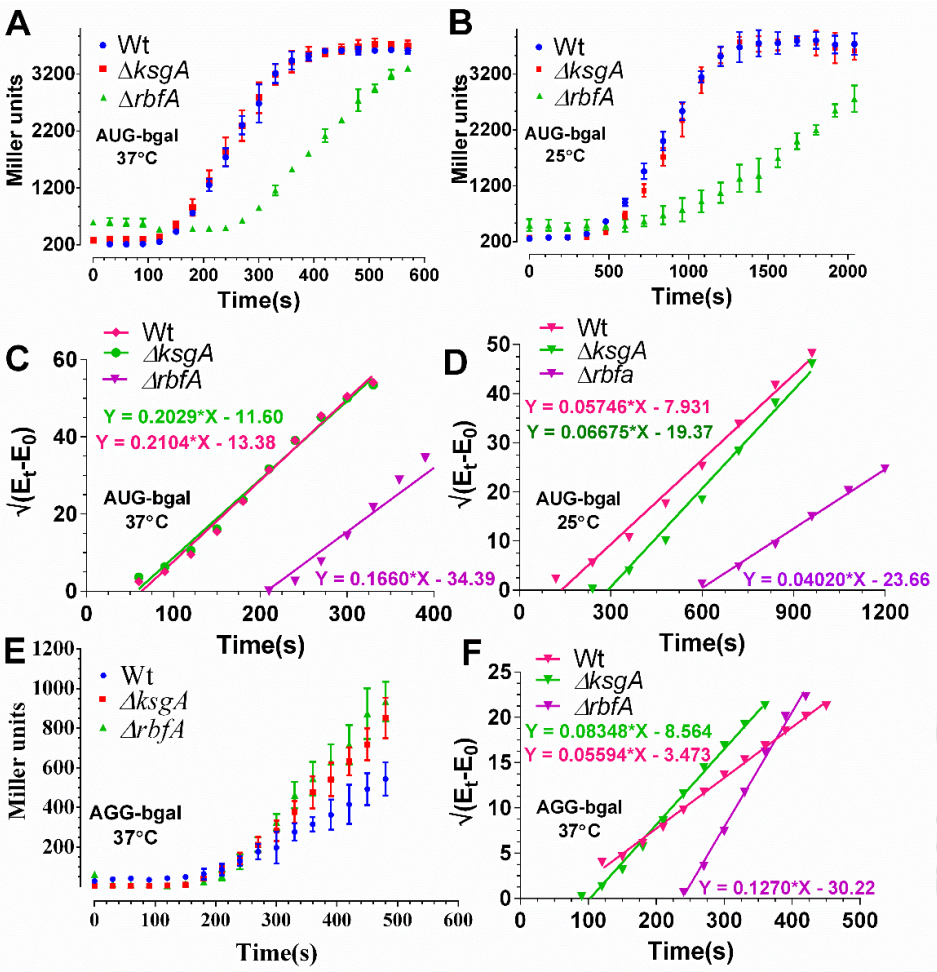

Figure 5.2: Translation kinetics of premature ribosomes

(A)–(B) Translation kinetics of bgal production with AUG start codon in Wt, Δ*rbfA* and Δ*ksgA*. Protein production kinetics was performed at 37 °C (A) and 25 °C (B), respectively. The plots were drawn as averaged measurement from three independent time course experiments. The error bars represent standard deviation of the three independent trials.

(C)-(D) Schleif plot for bgal translation kinetics derived by plotting the square root of residual bgal enzyme activity against time from data in (A) & (B) (*vide.* Methods*)*. The plotted data was fitted to a linear regression model. Equations representing the trend line for each plot have been shown. The slope of the line indicates the rate of translation initiation and the X-intercept marks the first appearance of the enzyme activity that can be further used to calculate the rate of peptide chain elongation.

(E) Translation kinetics of bgal production with AGG start codon in Wt, *ΔrbfA* and *ΔksgA*. The kinetics was performed at 37 °C.

(F) Schleif plot for bgal translation kinetics with AGG start codon derived from the data in (E).

5.3.3 Defects in ribosome assembly upset translation at a global scale

In order to probe how the assembly defects impact the formation of initiation complexes at global level, we turned to TCP-seq (Archer et al., 2016; Shirokikh et al., 2017), a modified Ribo-seq strategy that has been devised to probe translation initiation, elongation and termination complexes in *Saccharomyces cerevisiae* (Fig. 5.3A). The technique uses formaldehyde crosslinking to trap ribosomes during various stages of translation and thus, we have adopted this technique for probing the translational complexes in *E. coli*.

Widely used bacterial translation elongation inhibitors like Chloramphenicol do not trap ribosomes in the initiation stage and also introduces a codon specific bias when used to arrest ribosomes during elongation (Marks et al., 2016). Thus, in order to arrest the ribosome from Wt, *ΔksgA* and *ΔrbfA* (Fig. 5.3B) during various phases of translation in an unbiased manner, we snap chilled the cells and cross-linked mRNA bound translating complexes using formaldehyde (Archer et al., 2016; Valášek et al., 2007). After cross-linking, free 30S particles were depleted using sedimentation analysis and mRNA cross-linked ribosomes were released by treating with Micrococcal nuclease (MNase). The released particles were further partitioned using density gradient sedimentation and then used to extract the mRNA footprint (FP) bound to 30S and 70S particles. For library preparation, a wide range of FPs of roughly 10-80 nucleotides (nt) were size selected to accommodate FPs arising from translating complexes as well as the tRNA molecules.

The length distribution of FPs that are mapped to the coding region was congruent with the size selection criteria used during library preparation, harbouring reads ranging from as low as 12 nt through 75 nt in length (Fig 5.3C & 5.3D). The 30S subunit derived FPs (30S_FP) contained a distinct population of ~ 14 nt in length (Fig. 5.3C). Whereas, we spotted three distinct population of FPs from 70S particles (70S_FP) centred on 15, 30 and 40 nt, respectively (Fig. 5.3D). Deletion of both KsgA and RbfA led to a significant drop in the number of FPs suggesting dampened initiation and elongation events (Fig. 5.3C & 5.3D). The size of the FPs mirrors the conformational rearrangements the ribosomes undergo as well as their interactions with translation factors during different stages of translation (Archer et al., 2016; Lareau et al.).To track these, we mapped the 5’ and 3’end of the FPs of given sizes to the known START and STOP codons of the coding sequences (CDS) in *E. coli* (Fig. 5.4 and

5.5). In line with the characteristic size distribution of FPs for 30S and 70S, the majority of the mapped FPs for 30S fell in the smaller size regime (12 – 20 nt) whereas they correspond to larger size regime (30 nt and above) for 70S (Fig. 5.4 & 5.5). Each FP of a given length shows a range of characteristic offset distances from the start/stop codons. 30S_FPs are densely populated near the start/stop codons as against 70S_FPs that recapitulates the fact that 30S is subjected to extensive regulatory events during initiation and termination stages. At the 5' UTR, the trailing edge of 30S_FP (5' end of FPs in Fig. 5.4A) shows graded increase in the offset distance with increase in the length of FPs than the leading edge (3' end of FPs in Fig. 5.4 C). In fact, the leading edge is arrested around the start codon suggesting that there is a queuing of 30S ribosomes at the 5' UTR due to the fact that the residence time of 30S at the start codon is protracted owing to the extensive regulatory controls. Interestingly, as expected, these trends are reversed for stop codons (Fig. 5.5). However, these signatures are significantly lost in case of *ΔksgA* and *ΔrbfA* suggesting that the premature ribosomes are actively blocked by the translation machinery to surpass the initiation stage (Fig. 5.4 & 5.5). Further, these trends are in line with the translation kinetics studied using bgal assay (Fig. 5.2).

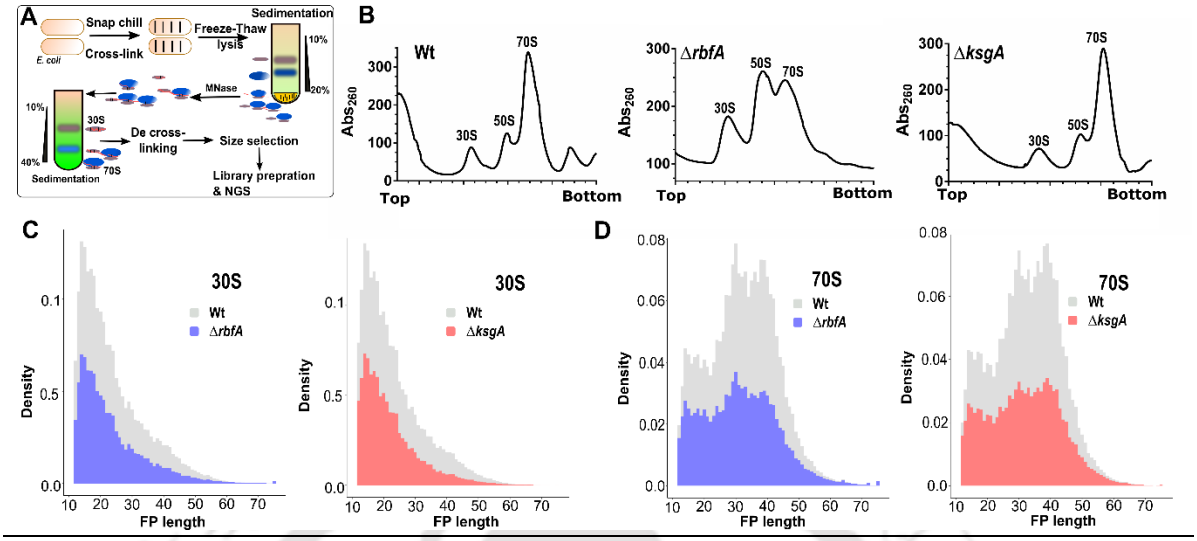

Figure 5.3: Translation complex profiling for studying initiation and elongation

(A) An outline of the strategy used to capture the ribosomes at various stages of translation using TCP-seq.

(B) Representative polysome profiles for Wt, *ΔrbfA*, and *ΔksgA* derived by performing sucrose density gradient sedimentation. Samples for TCP-seq were derived from 30S and 70S fractions, respectively.

(C) The distribution of 30S protected fragments isolated from Wt, *ΔrbfA* and *ΔksgA* is shown by plotting against the FP size (X-axis) versus the footprint density (Y-axis), respectively.

(D) The distribution of 70S protected fragments isolated from Wt, *ΔrbfA* and *ΔksgA* is shown.

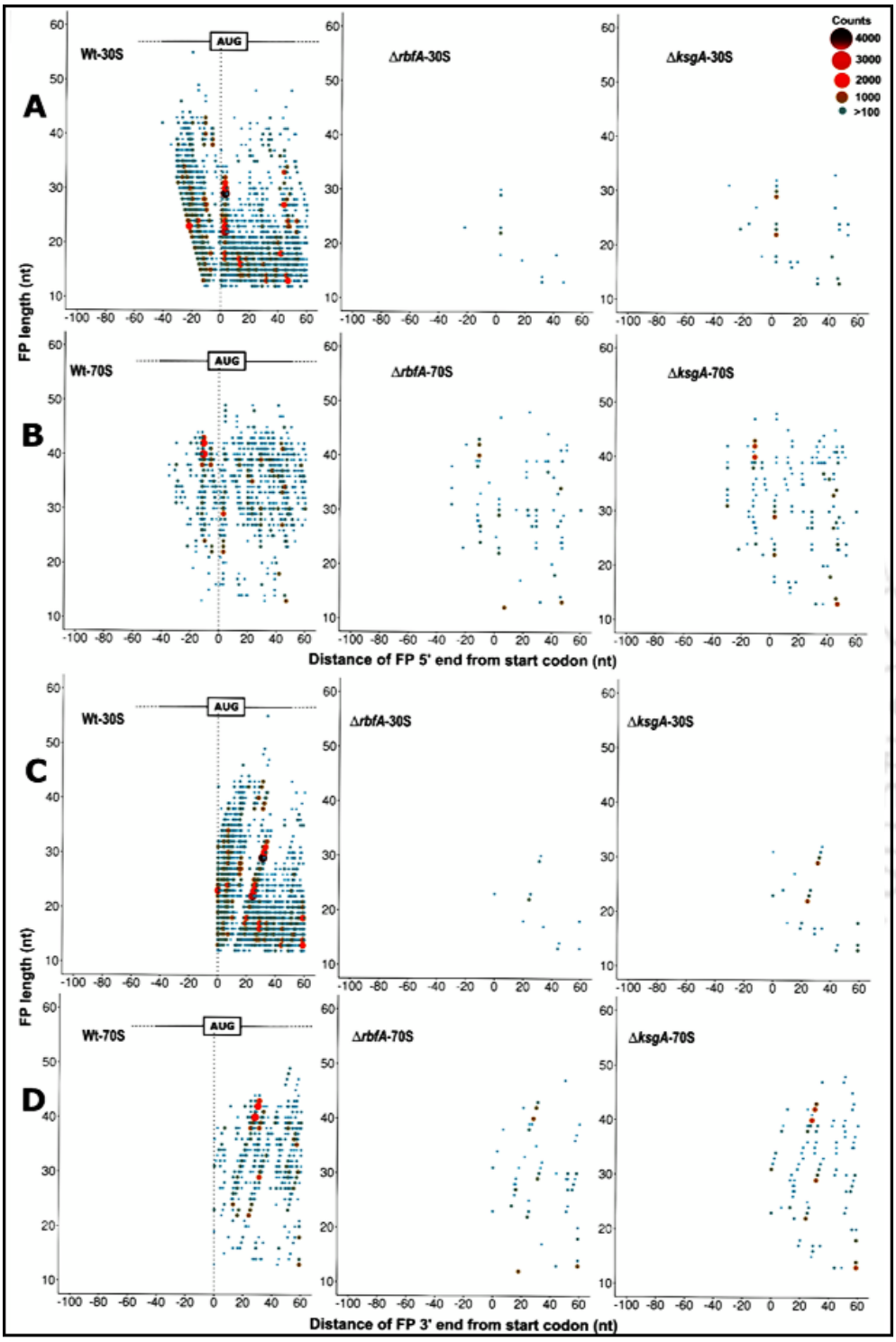

Figure 5.4: Metagene analysis for translation initiation complex

5' ends of the FPs derived from 30S **(A)** and 70S **(B)** bound mRNA as well as the 3' ends of the FPs derived from 30S **(C)** and 70S **(D)** from Wt, *ΔrbfA* and, *ΔksgA*, respectively, are

mapped against the known start codon. The first base of the start codon is aligned to the ribosome P site. The size and colour of the points correspond to the frequency of occurrence of FPs. Only those FPs with at least 100 reads are represented here.

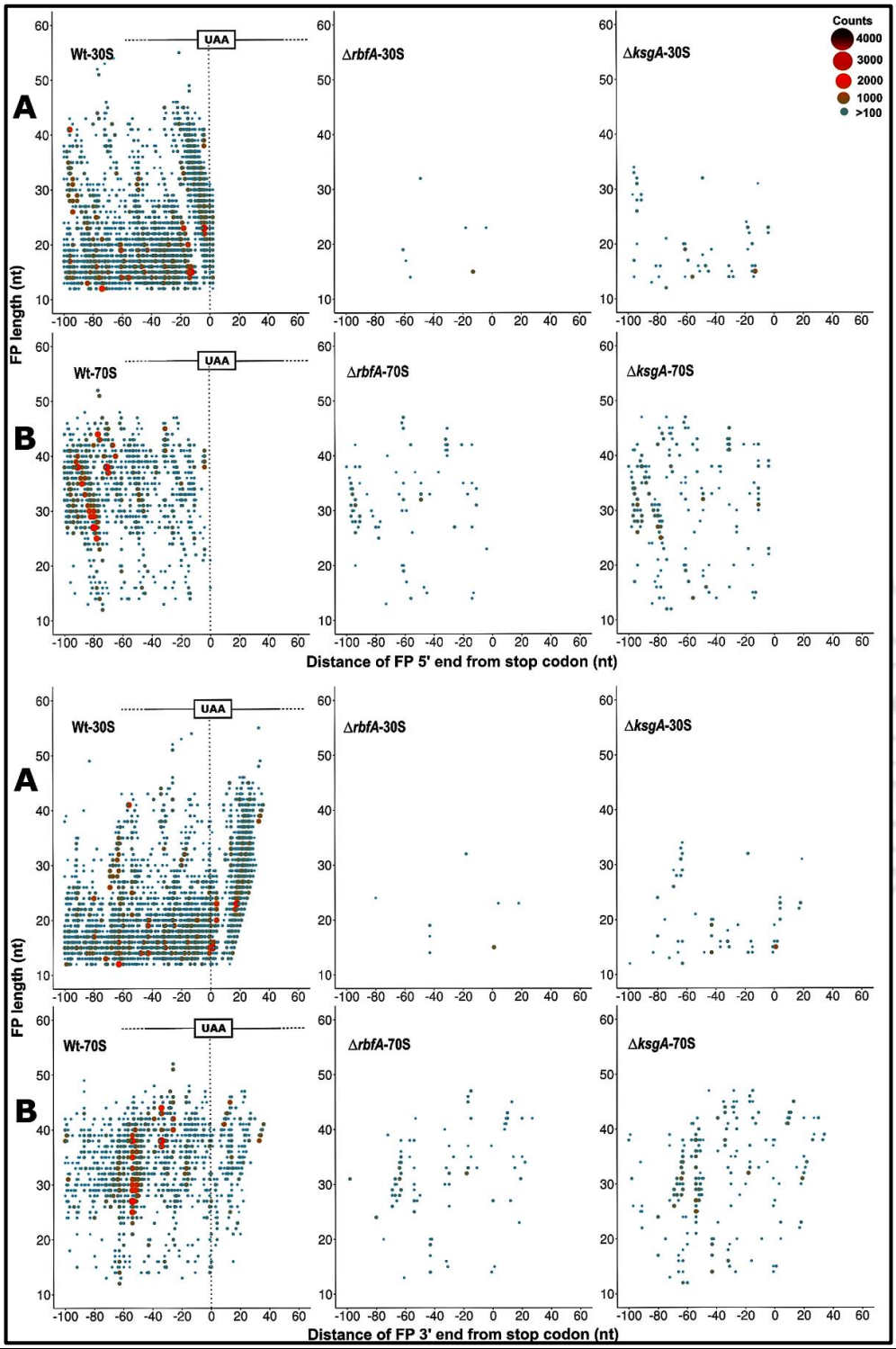

Figure 5.5: Metagene analysis of translation termination complex

5' ends of the FPs derived from 30S **(A)** and 70S **(B)** bound mRNA as well as the 3' ends of the FPs derived from 30S **(C)** and 70S **(D)** bound mRNA from Wt, *ΔrbfA* and, *ΔksgA*,

respectively, are mapped against the known stop codon. The first base of the stop codon is aligned to the ribosome A site. The size and colour of the points correspond to the frequency of occurrence of FPs. Only those FPs with at least 100 reads are represented here.

5.3.4 Premature 30S form short-lived preinitiation complexes

In order to further understand the nature of initiation and elongation complexes formed by the premature subunits, we studied the distribution of N-Formylmethionine initiator –tRNA ($tRNA^{i}_{fmet}$) and the elongator tRNA (tRNAe) that cross-link and co-purify with the 30S and 70S fractions (Fig. 5.3A). We reasoned that the tRNA distribution ($tRNA^{i}_{fmet}$ and tRNAe) would act as an indicator to pinpoint the ribosomes that are at initiation and elongation phases of the translation. The initiation stage being the rate-limiting step of translation is expected to be overpopulated by $tRNA^{i}_{fmet}$ during the transition of 30S-PIC to 30-IC (Archer et al., 2016; Milon et al., 2012) as compared to other stages associated with the elongating 70S particles. Similarly, an inverse distribution of tRNAe is expected to be observed with elongating 70S than the 30S-PIC/IC. In line with our hypothesis, Wt 30S ribosomes were found to harbour more number of $tRNA^{i}_{fmet}$ than the 70S (Fig.5.6A). Surprisingly, the distribution of $tRNA^{i}_{fmet}$ was reversed in both *ΔksgA* and *ΔrbfA* such that its level in 70S was significantly higher in comparison to that in 30S. We posit that the reduced distribution of $tRNA^{i}_{fmet}$ in the 30S fraction of *ΔksgA* and *ΔrbfA* may arise from decreased binding affinities of 30S towards the $tRNA^{i}_{fmet}$ or due to a short lived 30S-PIC/IC formed in *ΔksgA* and *ΔrbfA*, which rapidly associate with the 50S subunit to form 70S-IC. Since the transition from 30S-IC to 70S IC occurs through a well-orchestrated sequence of events, the association between $tRNA^{i}_{fmet}$ and 70S should have occurred during the formation of 70S-IC. Therefore, this alludes to the possibility that there is a rapid transition of 30S-IC to 70S-IC in *ΔksgA* and *ΔrbfA* and the prospect of low affinity between 30S and $tRNA^{i}_{fmet}$ can thus be ruled out. Further, tRNAe was distributed equally between Wt 30S and 70S particles, suggesting the dynamic interaction between tRNAe and ribosome subunits. However, its distribution was skewed for *ΔksgA* and *ΔrbfA* and showed more preference for 70S than the 30S (5.6A).

Intrigued by the preferential blocking of the translation initiation complex formed by premature ribosomes bound to mRNA containing cognate start (AUG) codon (Fig. 5.4), we asked whether the identity of the cognate start codon has any role in limiting the progression

towards elongation stage. To address this, we studied the distribution of 30S and 70S specific FPs from AGG-*gfp (*Fig. 5.6B). We reasoned that the distribution of FPs should mirror the observations from GFP and bgal expression studies (Fig. 5.1 & 5.2). In line with our conjecture, FPs derived from Wt 30S were densely populated around the start codon region (Fig. 5.6C), whereas 30S from *ΔrbfA* and *ΔksgA* showed reduced distribution of FPs around the start codon. This suggests that unlike the 30S-IC from *ΔrbfA* and *ΔksgA*, 30S-IC from Wt does not efficiently progress towards the formation of 70S-IC. Next, in order to understand how this 30S-PIC/IC transformed into 70S-IC or an elongation complex, we analysed the distribution of FPs from the 70S fraction. Interestingly, we noted that despite the significant distribution of FPs around the initiation site from the Wt 30S and 70S fraction, their distribution throughout the length of GFP is abysmal suggesting that the AGG-*gfp* is poorly translated. Strikingly, 70S FPs from *ΔrbfA* were evenly distributed in higher proportion than those from *ΔksgA* and Wt suggesting that premature ribosomes from *ΔrbfA* undergo rapid conversion of 30S-IC to 70S-IC leading to efficient translation of AGG-*gfp*. This agrees with the observations from GFP and bgal expression studies (Fig. 5.1 & 5.2).

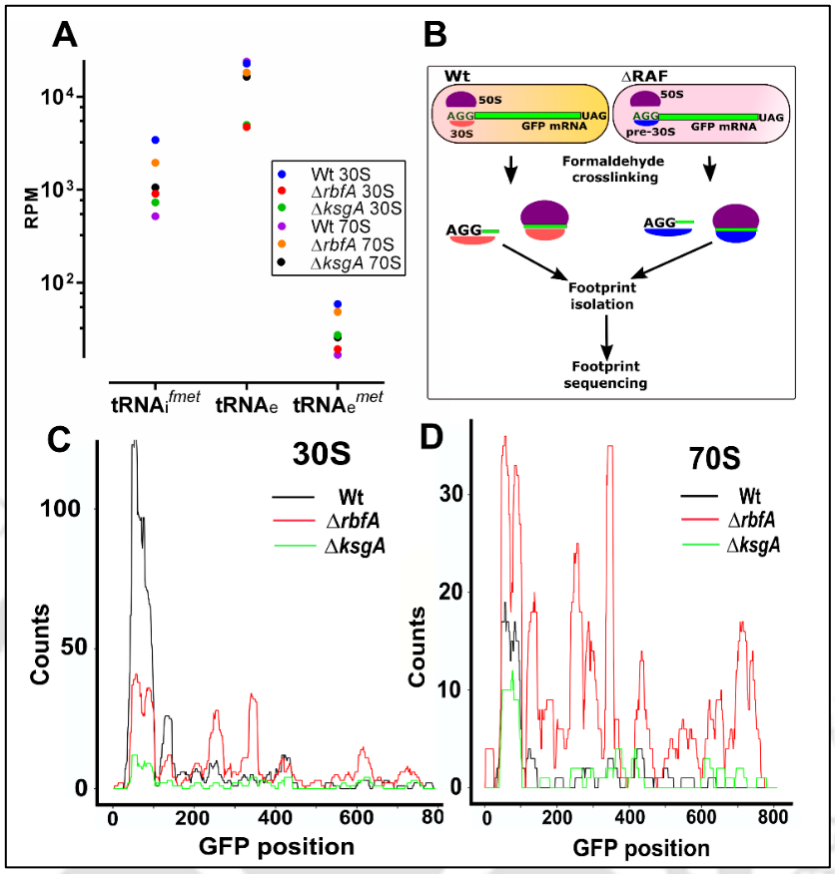

Figure 5.6: Premature ribosomes form distinct initiation complexes

(A) Distribution of initiator (tRNA$^{i}_{fmet}$) and elongator (tRNAe) tRNAs in 30S and 70S derived fractions as observed in the TCP-seq data. Read counts were converted to Reads per million (RPM). Distribution of Methionine specific elongator tRNA (tRNAe^{met}) is also shown for the respective strains.

(B) An outline of the methodology used to capture translation initiation and elongation complexes on an mRNA harbouring AGG start codon.

(C) & (D) Distribution of the AGG-GFP mRNA derived FPs from the 30S (C) & 70S fractions (D) of Wt, *ΔrbfA* and *ΔksgA*. The FP counts are shown along the entire length of the gene encoding GFP.

5.3.5 Hotspots for maintaining translation integrity are compromised in premature ribosomes

Our experiments clearly outline that although premature, these ribosomes possess essential hotspots to engage with the translation machinery. How these non-cognate PICs bypassed the scrutiny during the translation initiation is still a mystery. To solve this, we performed a survey of the structures of ribosome in complex with Initiation factors or RAFs

(Boehringer et al., 2012; Hussain et al., 2016; Razi et al., 2017). Taking cues from our analysis and previous observations, we found a significant overlap in the binding site of Initiation factors (IFs) and late stage assembly factors on the 30S subunit. We hypothesized that due to structural distortions, premature ribosomes may possess reduced affinities for the IFs. Such ribosomes, when entering translation with non-cognate initiation signals would bypass the fidelity checkpoints posed by IFs, thus leading to an inaccurate translation event. To test this *in vivo*, we simply elevated the cellular concentration of all three IFs individually by ectopic expression, expecting that an increased cellular concentration of IFs may restore the skewed mass balance ratio between IFs and premature 30S subunits. IRI values were calculated as a ratio between GFP fluorescence from cells with elevated IF concentration to that of cells with basal concentration of IF. An IRI value ~1 would mean that the expression levels of GFP were unaffected upon overexpressing IFs, whereas a value <1 would indicate repression of translation upon expression of the respective IF. The IRI values were measured for constructs with either AUG or AGG start codon individually. In order to achieve this, IF production was triggered when cells just entered the log phase (OD_{600} reached 0.2-0.3) followed by induction of GFP production at mid-log phase (OD_{600} reached 0.6) (*vide.* Methods section 5.2.3), thus allowing for cellular levels of the respective IFs to be elevated before *gfp* expression could initiate.

Elevated levels of IF-1 had no significant effects on expression of GFP from AGG or AUG start codons for Wt or *ΔksgA* cells (Fig. 5.7A). In *ΔrbfA*, IF-1 specifically suppressed expression from AGG and not AUG start codon (Fig. 5.7A). Unusually, an indiscriminate repression of GFP expression from AGG and AUG start codons containing constructs was observed when cellular levels of IF-2 were increased (Fig. 5.7B). It would seem plausible that overexpression of IF-2 somehow inhibited translation. Similar to IF-1, raising IF-3 specifically inhibited expression from AGG start codon rather than AUG (Fig. 5.7C). However unlike IF-1, IF-3 mediated inhibition of translation from AGG start codons is equally strong even in case of Wt or assembly deficient strains highlighting the role of IF-3 in discriminating the CO-AC interactions on the 30S-PIC/IC. Collectively, these observations reaffirm our hypothesis that the restoration of the pre-30S-IF mass balance may avert the participation of premature ribosomal particles in translation. To further test the probable IF

mediated moderation of the ribosome populations we studied the polysome profiles from Wt, *ΔksgA* and *ΔrbfA* cells containing elevated concentrations of the respective IFs and by complementing RAFs using sucrose density gradient centrifugation (Fig. 5.7D-F). In line with our expectations, complementation of RbfA and KsgA restored the distorted polysomes profiles in the respective null mutants (inset in Fig. 5.7E & 5.7F), suggesting efficient rescue from the assembly defects. Further, elevated levels of IF-1 did not seem to alter the distribution of ribosomes in Wt, *ΔksgA* and *ΔrbfA*. However, IF-2 overexpression led to a stabilization effect on the 70S particles that seemed drastic for *ΔrbfA* (Fig. 5.7E, compare *ΔrbfA* with *ΔrbfA* + IF-2) and subtler for Wt and *ΔksgA* (Fig. 5.7D & 5.7F). On the contrary, elevated levels of IF-3 displayed a pronounced subunit dissociation activity, where a decrease in the 70S population was accompanied by a concomitant increase in the 50S and 30S populations (Fig. 5.7D-F, compare *ΔrbfA* with *ΔrbfA* + IF-3). Similar to IF-2, the effect of IF-3 overexpression was more prominent in *ΔrbfA* as compared to Wt and *ΔksgA*.

In parallel, we also checked if ectopic expression of IFs improved cellular fitness, which could in turn lead to decrease in error-prone translation. To test this, Wt, *ΔrbfA* and *ΔksgA* cells transformed with plasmids carrying genes encoding IFs were tested for their growth characteristics after inducing the expression of the respective IF encoding genes. For the purpose of uniformity OD_{600} of all cultures was normalized at time of starting the experiment and IF expression was also triggered at the same time (time = 0) (*vide* methods section 5.2.2). The absence of a growth advantage (Fig. 5.8) indicated that the rescue was only a result of the translation quality control and not any improvement in the overall pool of premature ribosomes. As observed, overexpression of IF-1 or IF-3 did not improve cellular fitness in any manner but a radical effect of IF-2 overexpression was observed for all the three strains as they failed to grow even after prolonged incubation (Fig. 5.8). These observations suggest a central role of IF-2 in regulation of translation (Milon et al., 2006), that may be very sensitive to cellular concentrations of IF-2.

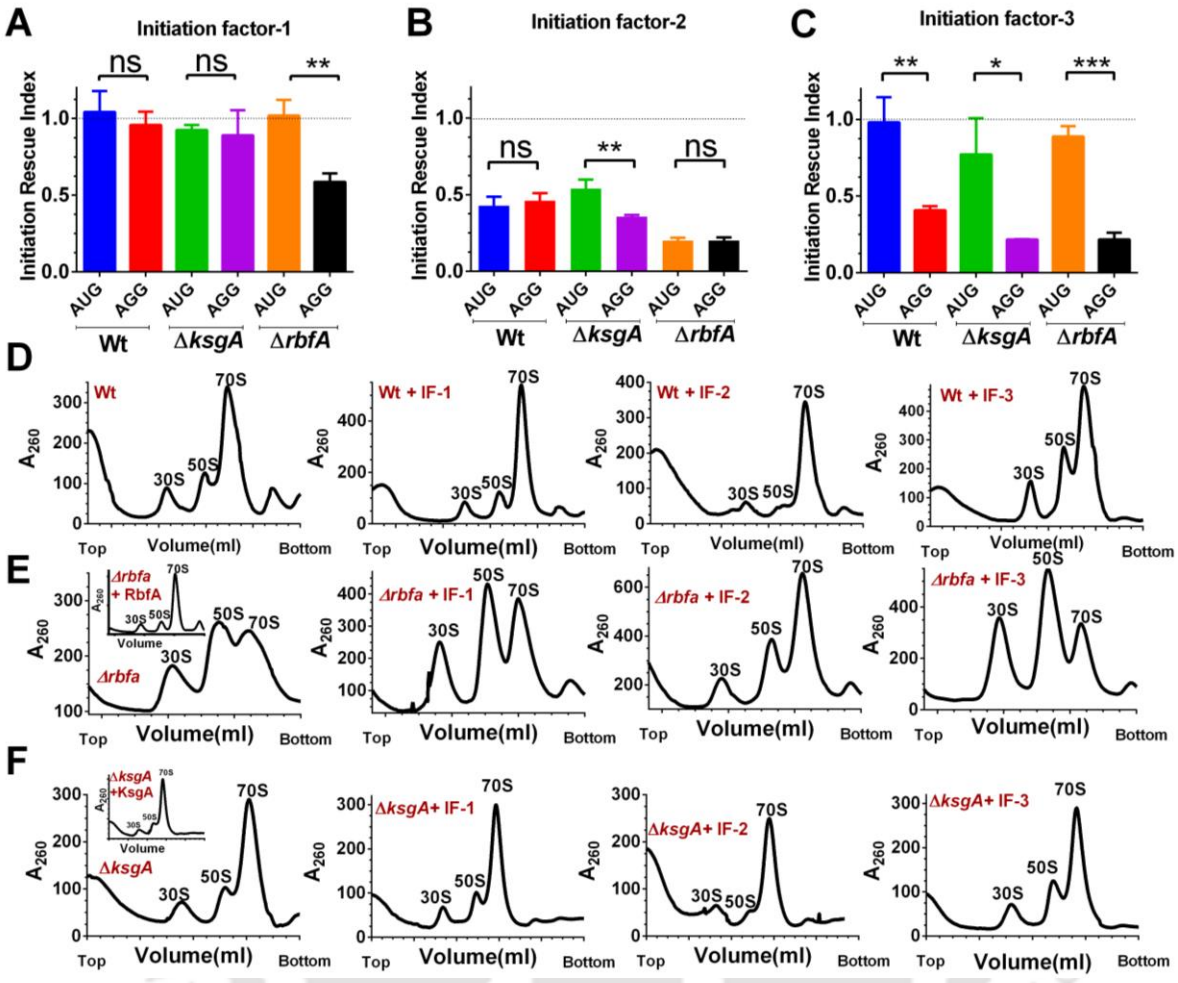

Figure 5.7: Elevated levels of initiation factors avert premature ribosomes from engaging in translation

(A) IRI measurements for Wt, *ΔrbfA* and *ΔksgA* from a *gfp* encoded mRNA with AUG or AGG start codon in presence of elevated cellular concentrations of IF-1. The rescue index is calculated as a ratio of GFP fluorescence in the respective strains with elevated IF-1 levels to that in cells with basal levels. The statistical significance was tested using a paired t-test analysis with a 95% confidence interval.

(B) IRI measurements for Wt, *ΔrbfA* and *ΔksgA* from a *gfp* encoded mRNA with AUG or AGG start codon in presence of elevated cellular concentrations of IF-2.

(C) IRI measurements for Wt, *ΔrbfA* and *ΔksgA* from a *gfp* encoded mRNA with AUG or AGG start codon in presence of elevated cellular concentrations of IF-3.

(D) Representative polysome profiles for Wt cells carrying an empty p15A vector (only Wt) or p15A vector carrying genes encoding IF-1, IF-2 or IF-3. The

profiles were drawn from mid-log phase cells overexpressing the respective initiation factors.

(E)–(F) Representative polysome profiles for *ΔrbfA* or *ΔksgA* cells carrying an empty p15A vector (only *ΔrbfA* or *ΔksgA)*. The inset shows the profile for the respective null mutants carrying p15A vector carrying a gene encoding RbfA or KsgA. Similar profiles are also shown for the respective null mutants carrying the p15A vector encoding IF-1, IF-2 or IF-3.

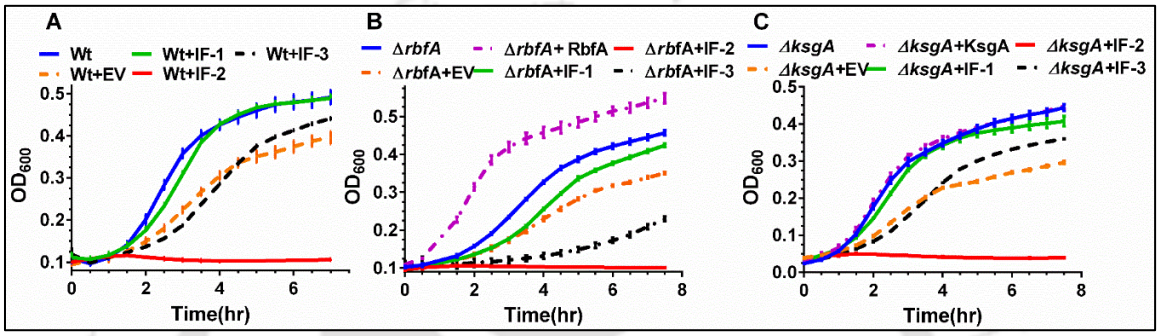

Fig. 5.8 Overexpression of IFs does not provide a growth advantage

Growth comparison for Wt **(A)**, *ΔrbfA* **(B)** and *ΔksgA* **(C)** transformed with an empty vector (EV) or a vector carrying genes encoding IF-1, IF-2, IF-3, RbfA and KsgA. Protein expression from the respective plasmids was induced at Time = 0.

5.3.6 Repression of translation from mRNA with non-cognate start codon is at the initiation stage

It is evident from the above experiments that IFs pre-empt premature ribosomes from participating in translation. However, in order to understand the mechanism of this inhibition we studied translation kinetics in strains with elevated levels of IFs. To perform this, IF production was triggered before measuring bgal production kinetics. Elevated IF-1 levels displayed appreciable disruption of translation initiation events only from AGG start codon in *ΔrbfA* (Fig. 5.9A & 5.9D) (Table 5.3). This was in line with the GFP repression from AGG start codon for *ΔrbfA* during the steady state measurements for GFP expression (Fig. 5.7A). However, in case of IF-2 overexpression, the initiation events dampened for AUG start codon and fell abysmally for AGG start codon (Fig. 5.9B & 5.9E) (Table 5.3). A completely different scenario unfolded when the cellular concentration of IF-3 was increased. Here, the initiation events were not repressed for AUG start codon but a graduated decrease was seen in

translation initiation events for AGG start codon (Fig. 5.9C & 5.9F) (Table 5.3). This IF-3 mediated, progressive inhibition of translation from non-canonical initiation signals again reinforces the positive effects of restoration of the mass balance between IFs and premature subunits to varying degrees in null mutants and Wt, as null mutants would harbour a higher proportion of premature subunits than Wt.

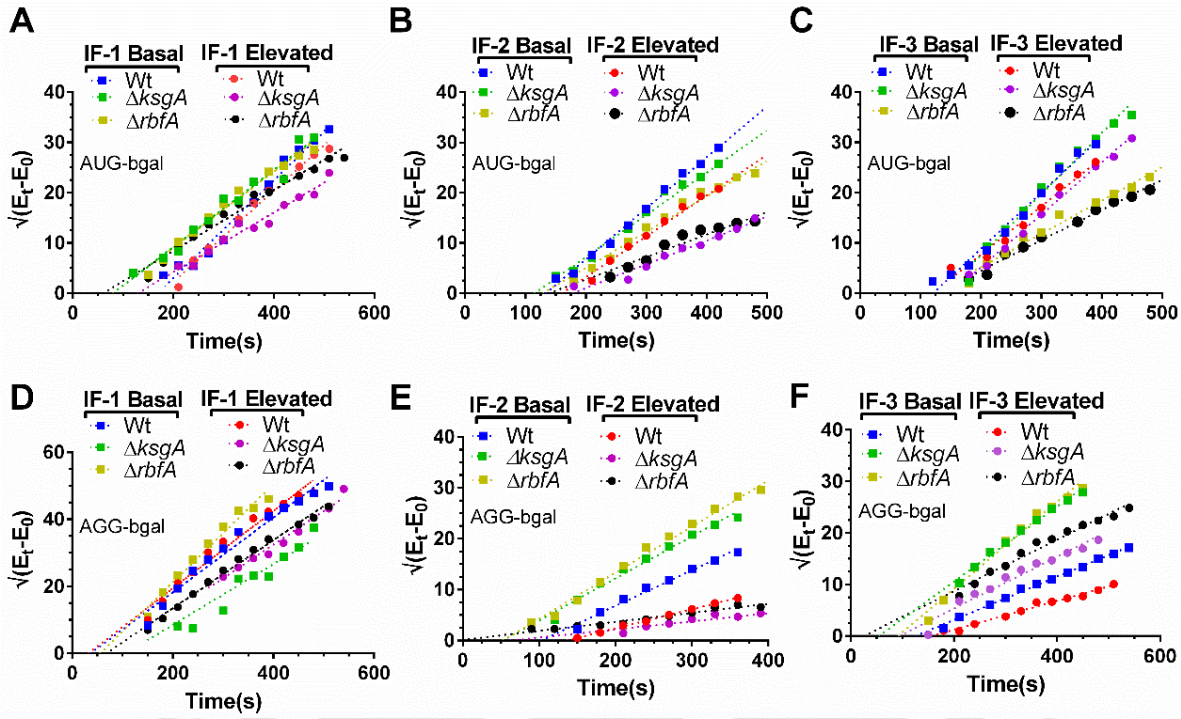

Figure 5.9: Elevated concentrations of Initiation factors prohibit translation initiation by premature ribosomes

(A)–(C) Translation kinetics of bgal production for Wt, *ΔrbfA* and *ΔksgA* from an AUG-bgal mRNA in elevated or basal concentrations of IF-1 (A), IF-2 (B) and IF-3(C)
(D)–(F) Translation kinetics of bgal production for Wt, *ΔrbfA* and *ΔksgA* from an AGG-bgal mRNA in elevated or basal concentrations of IF-1 (D), IF-2 (E) and IF-3 (F).

Table 5.3: Translation initiation rates (IR) measured from Wt, *ΔksgA*, and *ΔrbfA* cells after elevating cellular concentrations of IFs. Translation initiation from the respective start codons was studied from cells induced (elevated) or uninduced (basal) for IF production. The IR values are derived as the slope of the curve marking accumulation of bgal over time (units: Miller units$^{1/2}$/sec).

Strains	Start codon	IF-1 basal	IF-1 elevated	IF-2 basal	IF-2 elevated	IF-3 basal	IF-3 elevated
Wt	AUG	0.09065 ± 0.003	0.09714 ± 0.004	0.1014 ± 0.002	0.08108 ± 0.003	0.1113 ± 0.005	0.09598 ± 0.005
Wt	AGG	0.1115 ± 0.006	0.1156 ± 0.006	0.07005 ± 0.003	0.03771 ± 0.000	0.04192 ± 0.000	0.02794 ± 0.001
ΔksgA	AUG	0.07674 ± 0.004	0.05927 ± 0.003	0.08411 ± 0.002	0.04730 ± 0.002	0.1213 ± 0.006	0.1044 ± 0.003
ΔksgA	AGG	0.1001 ± 0.006	0.09060 ± 0.011	0.08254 ± 0.004	0.01587 ± 0.001	0.07259 ± 0.002	0.05026 ± 0.003
ΔrbfA	AUG	0.07567 ± 0.003	0.06036 ± 0.002	0.07105 ± 0.003	0.04475 ± 0.003	0.06621 ± 0.002	0.06027 ± 0.001
ΔrbfA	AGG	0.1465 ± 0.009	0.1017 ± 0.003	0.09060 ± 0.002	0.01766 ± 0.001	0.08442± 0.003	0.04989 ± 0.002

5.4 Discussion

Genetic disruptions of ribosome assembly (Dragon et al., 2002; Ganapathi et al., 2007; Sun and Woolford, 1994) and production of misfolded proteins by a compromised translation machinery have been associated with cell death and neurodegeneration (Lee et al., 2006). In order to counter this, eukaryotic systems deploy comprehensive mechanisms to avert the entry of premature ribosomes into the translation cycle (Strunk et al., 2012). However, parallel prokaryotic mechanisms that may ensure entry of only mature ribosomes into translation still elude discovery. Interestingly for *E. coli*, late stages of 30S subunit maturations seem to be closely coupled with translation initiation (Shetty and Varshney, 2016a). Our data suggests that, quality control checkpoints are also present in bacterial cells and the premature ribosomes generated in absence of RAFs successfully evade these checkpoints. Here, we show that premature ribosomes possessing suboptimal fidelity of codon recognition, participate in the translation cycle (Fig. 5.1). Consistent with previous observations of mistranslation events in *ΔksgA* (Connolly and Culver, 2013), our observations for *ΔrbfA*, *ΔksgA* and *ΔlepA* also reiterate a direct correlation between assembly and impaired translation machinery. These observations are also reminiscent of previous reports of start codon misrecognition upon mutation in the *rrn* (Qin et al., 2012) and *infC* genes (Haggerty and Lovett, 1997). It is important to reflect upon, whether the mis-translation events observed here arise from CO-AC interactions between non-cognate mRNA and $tRNA^{i}_{fmet}$ or incorporation of tRNAe into the 30S-ICs. Our observations of significant $tRNA^{i}_{fmet}$ localization in 70S particles from *ΔrbfA and ΔksgA* (Fig. 5.6A) and previous reports of N-terminal sequencing of mis-translated products (O'Connor et al., 1997) suggest that such mis-initiation events arise due to interactions between the $tRNA^{i}_{fmet}$ and non-AUG start codons. This further attributes mistranslation to a failure in discriminating non-canonical CO-AC interactions during formation of the 30S-IC.

During assembly, RbfA, a cold shock protein, is proposed to assist in formation of h1 (Dammel and Noller, 1995a) and later to reorient h44 and h45 that form the 3'minor domain of 16S rRNA. Also during these late stages of assembly, KsgA modifies two consecutive residues on the h45 region (Helser et al., 1972; Xu et al., 2008). Incidentally, mutations in h44 also promote initiation from non-AUG start codon by altering the movement of the 3' major

domain that forms the 30S-head region (Qin and Fredrick, 2009). All these findings, hint towards an essential need of efficient assembly of the 16S 3' domain in the decoding step. In this study, we find that recognition of canonical start signals seems to remain unaffected but there is a graduated increase in the propensity to recognize near-cognate (CUG and AUA) or a non-cognate (AGG) start codon (Fig. 5.1) for both *ΔksgA* and *ΔrbfA*. Elevated expression from AGG start codon is particularly noteworthy given the marked conservation of U among other near-cognate start codons (i.e., GUG, UUG, AUU, AUC, and AUA) and indicates severely compromised decoding capabilities at the 30S P-site. However, milder defects in recognizing frameshifts and stop codons (Fig. 5.1G) indicate a limited effect of assembly defects on decoding at the A-site too. Further, our experiments to gauge the mechanistic implications of impaired P-site decoding also reveal wide spread effects on protein production due to a decrease in rate of translation initiation and lowered rates of peptide bond synthesis (Fig. 5.2C, 5.2D & 5.2F). However, it is not possible to infer from the current data, if this decrease in the peptide chain elongation rate is solely due to defects in the P-site or other structural deformities that decrease the speed of the translocation step (Shi et al., 2009), thus slowing down the overall process of translation. It also seems plausible to extend that, as defects in assembly exacerbate, defects in codon recognition and translation impairment also aggravate (Fig. 5.3C, 5.3D & 5.4). However, it is also clearly outlined that these defects are not an outcome of translation cessation in RAF null mutants but rather a bona fide compromise in the fidelity (Fig. 5.2F).

Furthermore, our efforts to understand the nature and mechanism of faulty initiation is clearly highlighted by the distinct distribution of $tRNA^{i}_{fmet}$ with 30S-PIC/IC fractions (Fig. 5.6A) and the expression of *gfp* harbouring AGG start codon that suggest rapid transition of initiating complexes to elongating complexes (Fig. 5.6C & 5.6D), leading to evasion of the translation quality control checkpoints. These notions have been reaffirmed by the initiation rescue observed when cellular concentrations of IFs were increased to compensate for a possible decrease in pre-30S-IF binding due to structural defects in the 30S (Fig.5.7A-C). It is however noteworthy, that IF-1 and IF-3 seem to specifically inhibit translation initiation (Fig. 5.9) while the reason for non-specific inhibition of translation initiation after IF-2 overexpression is still unclear (Fig. 5.9). One plausible scenario for this could be that the

elevated IF-2 levels are not appropriately accounted for by the concomitant rise in the levels of $tRNA^{i}_{fmet,}$ leading to suboptimal charging of IF-2 with $tRNA^{i}_{fmet}$. This may give rise to vacant IF-2 binding to 30S and stalling of initiation. IF-3 is already known to discriminate the CO-AC interactions while formation of the 30S-PIC and its subsequent conversion to 30S-IC by a series of precisely controlled structural rearrangements (Hussain et al., 2016; Milon et al., 2008; Milon et al., 2012). Its subunit anti-association activity (Fig. 5.7D-F) also seems to play critical roles in ensuring fidelity of initiation by rapid ejection of non-canonical mRNA-tRNA in the P-site which could be abolished in case of premature 30S (Petrelli et al., 2001). However, the mechanism of IF-1 mediated rescue is subtler (Fig. 5.7D-F) as it is thought to enhance the function of IF-3 and IF-2 (Hussain et al., 2016). Thus it seems possible to derive from these findings, that a speedy conversion of initiating pre-30S particles to elongating 70S particles takes place by skipping the IF mediated quality control of protein synthesis. Further structural and biochemical evidence would be needed to understand the nature of these structural defects to fully understand the factors that contribute to a compromised decoding fidelity in premature subunits.

4.5 **Summary**

In this chapter, we show that ribosomes with assembly defects evade the proofreading steps during translation initiation and participate in the translation cycle. Such ribosomes show severely compromised decoding capabilities that give rise to errors in initiation and elongation. Tracing the genesis, we discovered that the assembly defects compromise the binding of initiation factors especially IF-3, thus licensing the rapid transitioning of 30S (pre) initiation complex to 70S initiation complex by tempering the codon recognition mechanism. Overall, our work highlights that a mass balance deficit between premature ribosomes and initiation factors steers the entry of premature ribosomes into the translation cycle.

Chapter VI

Conclusion, Significance and Future Directions

Chapter 6

6.1 Towards a brighter, broader and mechanistic understanding of ribosome assembly

After a golden age marked by structural characterization of ribosome function, our understanding of how these macromolecular machines are assembled from their constituents is entering its long-awaited renaissance. The observations derived from high-resolution structures and intensive biochemical approaches mostly yielded a "frozen" state information, that is now advancing into a more dynamic and high-throughput fashion to highlight new details about ribosome structure (Frank, 2017), composition (Mehmood et al., 2015) and modification (Mehmood et al., 2015) using Cryo-EM, Ribo-seq and mass spectrometry. In parallel, the ribosome assembly has also emerged as a rational drug target (Nikolay et al., 2016; Stokes et al., 2014b) to address the menace of multi-drug resistant pathogens and development of anti-cancer therapies (Ghalei et al., 2015). In an attempt to complement these advances, we set out to identify and address some of the long-standing problems in ribosome assembly.

One of the first problems, that we perceived as a major deterrent in this area was the slow discovery of RAFs and an equally perplexing pattern of seemingly similar ribosomal particles that accumulate in their absence (Shajani et al., 2011). We reasoned that, in order to resolve this, it would be necessary to understand the pathways of ribosome maturation and the functional overlap between RAFs. To address this, we set out to develop systematic methodologies that would facilitate the discovery of a minimal set of RAFs that are essential to drive assembly of ribosome. Thus, keeping in mind the incomplete knowledge of RAFs and their probable function, our first step was focused on developing a strategy to discover novel RAFs and study the interactions of known ones using BioID (Roux et al., 2018). In a first, we have demonstrated a proof-of-concept in bacterial system by probing the molecular interactions of essential GTPase Era. Through these validations, we have successfully applied BioID to the context of bacterial ribosome assembly and also unravelled new interactions of Era. Its novel interactions with components of the 50S subunit and members of translation

regulation machinery hint at some previously overlooked roles of Era and also calls for revisiting the roles of other essential GTPases involved in ribosome assembly (Britton, 2009a). The notion has been further bolstered by the disruptions in translation and cell growth observed upon Era depletion using Lon protease. Additionally, the uncharacterized proteins observed in the interactome of Era and S18 could serve as primers to start the next phase of RAF discovery using the BiFC based assay system. As demonstrated, the attempts would simply require the creation of null mutants or conditional depletion of the target proteins in the SLV_{NC} strain. A simplistic assessment of the participation of a candidate protein in ribosome assembly can be garnered using this methodology. Similarly, as demonstrated by the ability of BiFC to specifically highlight ribosome targeting by antibiotics, the technique shows great promise for the screening of ribosome targeting drugs (Wilson, 2014). This would be instrumental in augmenting the emerging area of exploiting ribosome assembly as a drug target. Finally, this report along with the works by Al-Jubran and coworkers, (Al-Jubran et al., 2013) have expanded the boundaries of utilizing BiFC, by employing it to study the assembly of large macromolecular complexes.

The implications of our next work entailing *in vivo* conditional depletion of essential proteins promise a major advancement in the area of ribosome biology in particular and in the area of genetic engineering in general. As demonstrated, it is now possible to avail a lineage of cells devoid of essential proteins, and this could be instrumental in dissecting their cellular roles. The methodology is particularly encouraging in attempts to study the function of major housekeeping genes for which deletion is not a viable option. Additionally, the phenotypes observed in Era depleted cells, establish its suspected role in cell cycle regulation and hint at unknown functional roles in regulating the translation. Further, structural characterization of ribosomal intermediates accumulating in absence of Era will shed more light on its function.

At last, taking hints from the disruption in translation due to Era depletion, we wondered about the true physiological implication of impaired assembly. To address this, we studied protein synthesis in absence of RAFs acting in the late stages of 30S maturation. Our results indicate that in RAF null mutants, there exist, serious flux imbalances in the cellular translational apparatus. As ribosome maturation slows down, premature ribosomes make their way into the translation cycle with a compromised structural complement. Consequently,

translation quality control checkpoints are impaired leading to a compromise in the fidelity of protein synthesis. This opens up new avenues for characterization of the regulatory and structural aspects leading to these observations. It would also be interesting to understand the physiological consequences of these observations given the grave penalties of erroneous protein synthesis in higher organisms (Belin et al., 2009). It also remains to be seen if this is a calculated bargain that the cellular machinery takes to balance the flux of ribosomes needed to match cellular protein demand thus highlighting the resilience of living systems to perturbations (Stelling et al., 2004; Wollman, 2018) or simply a failure of the ribosome and translation quality control mechanisms. These questions also hold broader implications for the evolution of the protein synthesis machinery and principles underlying the distribution of cellular resources.

References

Adilakshmi, T., Bellur, D.L., and Woodson, S.A. (2008). Concurrent nucleation of 16S folding and induced fit in 30S ribosome assembly. Nature *455*, 1268-1272.

Adilakshmi, T., Ramaswamy, P., and Woodson, S.A. (2005). Protein-independent folding pathway of the 16S rRNA 5' domain. J Mol Biol *351*, 508-519.

Afflerbach, H., Schroder, O., and Wagner, R. (1998). Effects of the Escherichia coli DNA-binding protein H-NS on rRNA synthesis in vivo. Molecular microbiology *28*, 641-653.

Agalarov, S.C., Sridhar Prasad, G., Funke, P.M., Stout, C.D., and Williamson, J.R. (2000). Structure of the S15,S6,S18-rRNA complex: assembly of the 30S ribosome central domain. Science *288*, 107-113.

Agalarov, S.C., and Williamson, J.R. (2000). A hierarchy of RNA subdomains in assembly of the central domain of the 30 S ribosomal subunit. Rna *6*, 402-408.

Agalarov, S.C., Zheleznyakova, E.N., Selivanova, O.M., Zheleznaya, L.A., Matvienko, N.I., Vasiliev, V.D., and Spirin, A.S. (1998). In vitro assembly of a ribonucleoprotein particle corresponding to the platform domain of the 30S ribosomal subunit. Proceedings of the National Academy of Sciences of the United States of America *95*, 999-1003.

Agarwal, D., Kamath, D., Gregory, S.T., and Connor, M. (2015). Modulation of Decoding Fidelity by Ribosomal Proteins S4 and S5. Journal of Bacteriology *197*, 1017.

Al-Jubran, K., Wen, J., Abdullahi, A., Roy Chaudhury, S., Li, M., Ramanathan, P., Matina, A., De, S., Piechocki, K., Rugjee, K.N., *et al.* (2013). Visualization of the joining of ribosomal subunits reveals the presence of 80S ribosomes in the nucleus. RNA *19*, 1669-1683.

Alexov, E.G., and Gunner, M.R. (1997). Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. Biophysical Journal *72*, 2075-2093.

Alix, J., and Nierhaus, K.H. (2003). DnaK-facilitated ribosome assembly in Escherichia coli revisited. Rna *9*, 787-793.

Alix, J.H., and Guerin, M.F. (1993). Mutant DnaK chaperones cause ribosome assembly defects in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America *90*, 9725-9729.

Allen, G.S., and Frank, J. (2006). Structural insights on the translation initiation complex: ghosts of a universal initiation complex. Molecular microbiology *63*, 941-950.

Andrade, J.M., Dos Santos, R.F., Chelysheva, I., Ignatova, Z., and Arraiano, C.M. (2018). The RNA-binding protein Hfq is important for ribosome biogenesis and affects translation fidelity. The EMBO journal *37*, e97631.

Andreeva, I., Belardinelli, R., and Rodnina, M.V. (2018). Translation initiation in bacterial polysomes through ribosome loading on a standby site on a highly translated mRNA. Proceedings of the National Academy of Sciences of the United States of America *115*, 4411-4416.

Archer, S.K., Shirokikh, N.E., Beilharz, T.H., and Preiss, T. (2016). Dynamics of ribosome scanning and recycling revealed by translation complex profiling. Nature *535*, 570.

Arigoni, F., Talabot, F., Peitsch, M., Edgerton, M.D., Meldrum, E., Allet, E., Fish, R., Jamotte, T., Curchod, M.L., and Loferer, H. (1998). A genome-based approach for the identification of essential bacterial genes. Nature biotechnology *16*, 851-856.

Aseev, L.V., Koledinskaya, L.S., and Boni, I.V. (2016). Regulation of Ribosomal Protein Operons rplM-rpsI, rpmB-rpmG, and rplU-rpmA at the Transcriptional and Translational Levels. Journal of Bacteriology *198*, 2494.

Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 Å Resolution. Science *289*, 905.

Bartlett, M.S., and Gourse, R.L. (1994). Growth rate-dependent control of the rrnB P1 core promoter in Escherichia coli. Journal of Bacteriology *176*, 5560-5564.

Bashford, D., and Karplus, M. (1990). pKa's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. Biochemistry *29*, 10219-10225.

Bedwell, D., Davis, G., Gosink, M., Post, L., Nomura, M., Kestler, H., Zengel, J.M., and Lindahl, L. (1985). Nucleotide sequence of the alpha ribosomal protein operon of Escherichia coli. Nucleic Acids Res *13*, 3891-3903.

Belin, D. (2003). Why Are Suppressors of Amber Mutations So Frequent Among Escherichia coli K12 Strains?: A Plausible Explanation for a Long-Lasting Puzzle. Genetics *165*, 455.

Belin, S., Beghin, A., Solano-Gonzàlez, E., Bezin, L., Brunet-Manquat, S., Textoris, J., Prats, A.-C., Mertani, H.C., Dumontet, C., and Diaz, J.-J. (2009). Dysregulation of Ribosome Biogenesis and Translational Capacity Is Associated with Tumor Progression of Human Breast Cancer Cells. PLOS ONE *4*, e7147.

Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G., and Yusupov, M. (2011). The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution. Science *334*, 1524.

Bleichert, F., and Baserga, S.J. (2007). The long unwinding road of RNA helicases. Mol Cell *27*, 339-352.

Boehringer, D., O'Farrell, H.C., Rife, J.P., and Ban, N. (2012). Structural insights into methyltransferase KsgA function in 30S ribosomal subunit biogenesis. J Biol Chem *287*, 10453-10459.

Bokal, A.J.t., Ross, W., and Gourse, R.L. (1995). The transcriptional activator protein FIS: DNA interactions and cooperative interactions with RNA polymerase at the Escherichia coli rrnB P1 promoter. J Mol Biol *245*, 197-207.

Bram, R.J., Young, R.A., and Steitz, J.A. (1980). The ribonuclease III site flanking 23S sequences in the 30S ribosomal precursor RNA of E. coli. Cell *19*, 393-401.

Britton, R.A. (2009a). Role of GTPases in bacterial ribosome assembly. Annual review of microbiology *63*, 155-176.

Britton, R.A. (2009b). Role of GTPases in bacterial ribosome assembly. Annu Rev Microbiol *63*, 155-176.

Britton, R.A., Powell, B.S., Court, D.L., and Lupski, J.R. (1997). Characterization of mutations affecting the Escherichia coli essential GTPase era that suppress two temperature-sensitive dnaG alleles. Journal of Bacteriology *179*, 4575.

Britton, R.A., Powell, B.S., Dasgupta, S., Sun, Q., Margolin, W., Lupski, J.R., and Court, D.L. (2002). Cell cycle arrest in Era GTPase mutants: a potential growth rate-regulated checkpoint in Escherichia coli. Molecular Microbiology *27*, 739-750.

Brückner, A., Polge, C., Lentze, N., Auerbach, D., and Schlattner, U. (2009). Yeast Two-Hybrid, a Powerful Tool for Systems Biology. International Journal of Molecular Sciences *10*, 2763-2788.

Bubunenko, M., Baker, T., and Court, D.L. (2007). Essentiality of ribosomal and transcription antitermination proteins analyzed by systematic gene replacement in Escherichia coli. J Bacteriol *189*, 2844-2853.

Bylund, G.O., Wipemo, L.C., Lundberg, L.A., and Wikstrom, P.M. (1998a). RimM and RbfA are essential for efficient processing of 16S rRNA in Escherichia coli. Journal of bacteriology *180*, 73-82.

Bylund, G.O., Wipemo, L.C., Lundberg, L.A., and Wikström, P.M. (1998b). RimM and RbfA are essential for efficient processing of 16S rRNA in Escherichia coli. Journal of Bacteriology *180*, 73-82.

Cameron, D.E., and Collins, J.J. (2014). Tunable protein degradation in bacteria. Nature Biotechnology *32*, 1276.

Carr, A.C., Taylor, K.L., Osborne, M.S., Belous, B.T., Myerson, J.P., and Moore, S.D. (2012). Rapid Depletion of Target Proteins Allows Identification of Coincident Physiological Responses. Journal of Bacteriology *194*, 5932.

Carter, A.P., Clemons, W.M., Brodersen, D.E., Morgan-Warren, R.J., Wimberly, B.T., and Ramakrishnan, V. (2000). Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. Nature *407*, 340.

Carter, A.P., Clemons, W.M., Jr., Brodersen, D.E., Morgan-Warren, R.J., Hartsch, T., Wimberly, B.T., and Ramakrishnan, V. (2001). Crystal structure of an initiation factor bound to the 30S ribosomal subunit. Science *291*, 498-501.

Cartier, G., Lorieux, F., Allemand, F., Dreyfus, M., and Bizebard, T. (2010). Cold adaptation in DEAD-box proteins. Biochemistry *49*, 2636-2646.

Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R., and Doudna, J.A. (1996). Crystal structure of a group I ribozyme domain: principles of RNA packing. Science *273*, 1678-1685.

Cech, T.R., and Bass, B.L. (1986). Biological Catalysis by RNA. Annual Review of Biochemistry *55*, 599-629.

Cerretti, D.P., Dean, D., Davis, G.R., Bedwell, D.M., and Nomura, M. (1983). The spc ribosomal protein operon of Escherichia coli: sequence and cotranscription of the ribosomal protein genes and a protein export gene. Nucleic Acids Res *11*, 2599-2616.

Charollais, J., Dreyfus, M., and Iost, I. (2004). CsdA, a cold-shock RNA helicase from Escherichia coli, is involved in the biogenesis of 50S ribosomal subunit. Nucleic Acids Research *32*, 2751-2759.

Charollais, J., Pflieger, D., Vinh, J., Dreyfus, M., and Iost, I. (2003a). The DEAD-box RNA helicase SrmB is involved in the assembly of 50S ribosomal subunits in Escherichia coli. Molecular microbiology *48*, 1253-1265.

Charollais, J., Pflieger, D., Vinh, J., Dreyfus, M., and Iost, I. (2003b). The DEAD-box RNA helicase SrmB is involved in the assembly of 50S ribosomal subunits in Escherichia coli. Molecular Microbiology *48*, 1253-1265.

Charollais, J., Pflieger, D., Vinh, J., Dreyfus, M., and Iost, I. (2003c). The DEAD-box RNA helicase SrmB is involved in the assembly of 50S ribosomal subunits in Escherichia coli: E. coli ribosome assembly involves a RNA helicase. Molecular Microbiology *48*, 1253-1265.

Chen, X., Wei, S., Ji, Y., Guo, X., and Yang, F. (2015). Quantitative proteomics using SILAC: Principles, applications, and developments. PROTEOMICS *15*, 3175-3192.

Chen, Y., Kaji, A., Kaji, H., and Cooperman, B.S. (2017). The kinetic mechanism of bacterial ribosome recycling. Nucleic Acids Research *45*, 10168-10177.

Choi-Rhee, E., Schulman, H., and Cronan, J.E. (2004). Promiscuous protein biotinylation by Escherichia coli biotin protein ligase. Protein Science : A Publication of the Protein Society *13*, 3043-3050.

Coenye, T., and Vandamme, P. (2005). Organisation of the S10, spc and alpha ribosomal protein gene clusters in prokaryotic genomes. FEMS microbiology letters *242*, 117-126.

Connolly, K., and Culver, G. (2013). Overexpression of RbfA in the absence of the KsgA checkpoint results in impaired translation initiation. Molecular microbiology *87*, 968-981.

Connolly, K., Rife, J.P., and Culver, G. (2008). Mechanistic insight into the ribosome biogenesis functions of the ancient protein KsgA. Molecular Microbiology *70*, 1062-1075.

Cordin, O., Banroques, J., Tanner, N.K., and Linder, P. (2006). The DEAD-box protein family of RNA helicases. Gene *367*, 17-37.

Cruz-Bustos, T., Ramakrishnan, S., Cordeiro, C.D., Ahmed, M.A., and Docampo, R. (2017). A Riboswitch-based Inducible Gene Expression System for Trypanosoma brucei. Journal of Eukaryotic Microbiology *65*, 412-421.

D'Elia, M.A., Pereira, M.P., and Brown, E.D. (2009). Are essential genes really essential? Trends in Microbiology *17*, 433-438.

Dahlberg, A.E., Dahlberg, J.E., Lund, E., Tokimatsu, H., Rabson, A.B., Calvert, P.C., Reynolds, F., and Zahalak, M. (1978). Processing of the 5' end of Escherichia coli 16S ribosomal RNA. Proceedings of the National Academy of Sciences of the United States of America *75*, 3598-3602.

Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Okano, H., Williamson, J.R., Fredrick, K., Hwa, T., and Lee, S.Y. (2018). Slowdown of Translational Elongation in Escherichia coli under Hyperosmotic Stress. mBio *9*, e02375-02317.

Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J.R., Fredrick, K., Wang, Y.-P., and Hwa, T. (2016). Reduction of translating ribosomes enables Escherichia coli to maintain elongation rates during slow growth. Nature Microbiology *2*, 16231.

Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J.R., Fredrick, K., Wang, Y.-P., and Hwa, T. (2017). Reduction of translating ribosomes enables Escherichia coli to maintain elongation rates during slow growth. Nature Microbiology *2*, 16231.

Dalbow, D.G., and Bremer, H. (1975). Metabolic regulation of beta-galactosidase synthesis in Escherichia coli. A test for constitutive ribosome synthesis. Biochem. J. *150*, 1-8.

Dalbow, D.G., and Young, R. (1975). Synthesis time of beta-galactosidase in Escherichia coli B/r as a function of growth rate. Biochem. J. *150*, 13-20.

Dammel, C.S., and Noller, H.F. (1995a). Suppression of a cold-sensitive mutation in 16S rRNA by overexpression of a novel ribosome-binding factor, RbfA. Genes Dev *9*, 626-637.

Dammel, C.S., and Noller, H.F. (1995b). Suppression of a cold-sensitive mutation in 16s rRNA by overexpression of a novel ribosome-binding factor, RbfA. Genes and Development *9*, 626-637.

Das, A., Ghosh, B., Barik, S., and Wolska, K. (1985). Evidence that ribosomal protein S10 itself is a cellular component necessary for transcription antitermination by phage lambda N protein. Proceedings of the National Academy of Sciences *82*, 4070.

Datsenko, K.A., and Wanner, B.L. (2000a). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proceedings of the National Academy of Sciences of the United States of America *97*, 6640-6645.

Datsenko, K.A., and Wanner, B.L. (2000b). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proceedings of the National Academy of Sciences of the United States of America *97*, 6640-6645.

Datsenko, K.A., and Wanner, B.L. (2000c). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proceedings of the National Academy of Sciences *97*, 6640-6645.

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., *et al.* (2007). Structural aspects of RbfA action during small ribosomal subunit assembly. Mol Cell *28*, 434-445.

Davis, J.H., Baker, T.A., and Sauer, R.T. (2009). Engineering Synthetic Adaptors and Substrates for Controlled ClpXP Degradation. Journal of Biological Chemistry *284*, 21848-21855.

Davis, J.H., and Williamson, J.R. (2017). Structure and dynamics of bacterial ribosome biogenesis. Philos Trans R Soc Lond B Biol Sci *372*.

de Narvaez, C.C., and Schaup, H.W. (1979). In vivo transcriptionally coupled assembly of Escherichia coli ribosomal subunits. J Mol Biol *134*, 1-22.

Decatur, W.A., and Fournier, M.J. (2002). rRNA modifications and ribosome function. Trends in biochemical sciences *27*, 344-351.

Demo, G., Rasouly, A., Vasilyev, N., Svetlov, V., Loveland, A.B., Diaz-Avalos, R., Grigorieff, N., Nudler, E., and Korostelev, A.A. (2017). Structure of RNA polymerase bound to ribosomal 30S subunit. eLife *6*, e28560.

Denman, R., Weitzmann, C., Cunningham, P.R., Negre, D., Nurse, K., Colgan, J., Pan, Y.C., Miedel, M., and Ofengand, J. (1989). In vitro assembly of 30S and 70S bacterial ribosomes from 16S RNA containing single base substitutions, insertions, and deletions around the decoding site (C1400). Biochemistry *28*, 1002-1011.

Deutscher, M.P. (2006). Degradation of RNA in bacteria: comparison of mRNA and stable RNA. Nucleic Acids Res *34*, 659-666.

Deutscher, M.P. (2009). Maturation and degradation of ribosomal RNA in bacteria. Progress in molecular biology and translational science *85*, 369-391.

DeVito, J., and Das, A. (1994). Control of transcription processivity in phage lambda: Nus factors strengthen the termination-resistant state of RNA polymerase induced by N antiterminator. Proceedings of the National Academy of Sciences *91*, 8660.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., and Rodnina, M.V. (2013). EF-P Is Essential for Rapid Synthesis of Proteins Containing Consecutive Proline Residues. Science *339*, 85.

Dohme, F., and Nierhaus, K.H. (1976). Total reconstitution and assembly of 50 S subunits from Escherichia coli Ribosomes in vitro. Journal of molecular biology *107*, 585-599.

Dorner, S., Brunelle, J.L., Sharma, D., and Green, R. (2006). The hybrid state of tRNA binding is an authentic translation elongation intermediate. Nature structural & molecular biology *13*, 234-241.

Dragon, F., Gallagher, J.E., Compagnone-Post, P.A., Mitchell, B.M., Porwancher, K.A., Wehner, K.A., Wormsley, S., Settlage, R.E., Shabanowitz, J., Osheim, Y., *et al.* (2002). A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. Nature *417*, 967-970.

Draper, D.E. (1987). Translational Regulation of Ribosomal Proteins in Escherichia coli. In Translational Regulation of Gene Expression, J. Ilan, ed. (Boston, MA: Springer US), pp. 1-26.

Dunn, J.J., and Studier, F.W. (1973). T7 early RNAs and Escherichia coli ribosomal RNAs are cut from large precursor RNAs in vivo by ribonuclease 3. Proceedings of the National Academy of Sciences of the United States of America *70*, 3296-3300.

Ehresmann, C., Stiegler, P., Mackie, G.A., Zimmermann, R.A., Ebel, J.P., and Fellner, P. (1975). Primary sequence of the 16S ribosomal RNA of Escherichia coli. Nucleic Acids Research *2*, 265-278.

El Hage, A., Sbai, M., and Alix, J.H. (2001). The chaperonin GroEL and other heat-shock proteins, besides DnaK, participate in ribosome biogenesis in Escherichia coli. Molecular & general genetics : MGG *264*, 796-808.

Estrem, S.T., Gaal, T., Ross, W., and Gourse, R.L. (1998). Identification of an UP element consensus sequence for bacterial promoters. Proceedings of the National Academy of Sciences of the United States of America *95*, 9761-9766.

Feng, B., Mandava, C.S., Guo, Q., Wang, J., Cao, W., Li, N., Zhang, Y., Zhang, Y., Wang, Z., Wu, J., *et al.* (2014). Structural and functional insights into the mode of action of a universally conserved Obg GTPase. PLoS Biol *12*, e1001866.

Formenoy, L.J., Cunningham, P.R., Nurse, K., Pleij, C.W., and Ofengand, J. (1994). Methylation of the conserved A1518-A1519 in Escherichia coli 16S ribosomal RNA by the ksgA methyltransferase is influenced by methylations around the similarly conserved U1512.G1523 base pair in the 3' terminal hairpin. Biochimie *76*, 1123-1128.

Formstone, A., and Errington, J. (2005). A magnesium-dependent mreB null mutant: implications for the role of mreB in Bacillus subtilis. Molecular Microbiology *55*, 1646-1657.

Fox, G.E., and Naik, A.K. (2004). The Evolutionary History of the Translation Machinery. In The Genetic Code and the Origin of Life, L. Ribas de Pouplana, ed. (Boston, MA: Springer US), pp. 92-105.

Frank, J. (2017). Advances in the field of single-particle cryo-electron microscopy over the last decade. Nature protocols *12*, 209-212.

Freistroffer, D.V., Pavlov, M.Y., MacDougall, J., Buckingham, R.H., and Ehrenberg, M. (1997). Release factor RF3 in E.coli accelerates the dissociation of release factors RF1 and RF2 from the ribosome in a GTP-dependent manner. The EMBO journal *16*, 4126-4133.

Fuller-Pace, F.V., Nicol, S.M., Reid, A.D., and Lane, D.P. (1993). DbpA: a DEAD box protein specifically activated by 23s rRNA. Embo j *12*, 3619-3626.

Gaal, T., Bartlett, M.S., Ross, W., Turnbough, C.L., Jr., and Gourse, R.L. (1997). Transcription regulation by initiating NTP concentration: rRNA synthesis in bacteria. Science *278*, 2092-2097.

Gafny, R., Cohen, S., Nachaliel, N., and Glaser, G. (1994). Isolated P2 rRNA promoters of Escherichia coli are strong promoters that are subject to stringent control. J Mol Biol *243*, 152-156.

Ganapathi, K.A., Austin, K.M., Lee, C.S., Dias, A., Malsch, M.M., Reed, R., and Shimamura, A. (2007). The human Shwachman-Diamond syndrome protein, SBDS, associates with ribosomal RNA. Blood *110*, 1458-1465.

Gardner, K.A.J.A., Osawa, M., and Erickson, H.P. (2017). Whole genome re-sequencing to identify suppressor mutations of mutant and foreign Escherichia coli FtsZ. PLOS ONE *12*, e0176643.

Gaylord, N.G., and Gibbs, J.H. (1962). Physical chemistry of macromolecules. . Journal of Polymer Science *62*, S22-S23.

Gerdes, S.Y., Scholle, M.D., Campbell, J.W., Balázsi, G., Ravasz, E., Daugherty, M.D., Somera, A.L., Kyrpides, N.C., Anderson, I., Gelfand, M.S., *et al.* (2003). Experimental determination and system level analysis of essential genes in Escherichia coli MG1655. Journal of Bacteriology *185*, 5673.

Ghalei, H., Schaub, F.X., Doherty, J.R., Noguchi, Y., Roush, W.R., Cleveland, J.L., Stroupe, M.E., and Karbstein, K. (2015). Hrr25/CK1δ-directed release of Ltv1 from pre-40S ribosomes is necessary for ribosome assembly and cell growth. The Journal of Cell Biology *208*, 745.

Ghora, B.K., and Apirion, D. (1978). Structural analysis and in vitro processing to p5 rRNA of a 9S RNA molecule isolated from an rne mutant of E. coli. Cell *15*, 1055-1066.

Giacalone, M.J., Gentile, A.M., Lovitt, B.T., Berkley, N.L., Gunderson, C.W., and Surber, M.W. (2006). Toxic protein expression in Escherichia coli using a rhamnose-based tightly regulated and tunable promoter system. BioTechniques *40*, 355-364.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., *et al.* (2002). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387.

Gibbs, M.R., Moon, K.-M., Chen, M., Balakrishnan, R., Foster, L.J., and Fredrick, K. (2017). Conserved GTPase LepA (Elongation Factor 4) functions in biogenesis of the 30S

subunit of the 70S ribosome. Proceedings of the National Academy of Sciences *114*, 980-985.

Gingras, A.-C., Gstaiger, M., Raught, B., and Aebersold, R. (2007). Analysis of protein complexes using mass spectrometry. Nature Reviews Molecular Cell Biology *8*, 645-654.

Goto, S., Kato, S., Kimura, T., Muto, A., and Himeno, H. (2011a). RsgA releases RbfA from 30S ribosome during a late stage of ribosome biosynthesis. The EMBO journal *30*, 104-114.

Goto, S., Kato, S., Kimura, T., Muto, A., and Himeno, H. (2011b). RsgA releases RbfA from 30S ribosome during a late stage of ribosome biosynthesis. The EMBO journal *30*, 104-114.

Greber, B.J., Boehringer, D., Godinic-Mikulcic, V., Crnkovic, A., Ibba, M., Weygand-Durasevic, I., and Ban, N. (2012). Cryo-EM structure of the archaeal 50S ribosomal subunit in complex with initiation factor 6 and implications for ribosome evolution. J Mol Biol *418*, 145-160.

Green, R., and Noller, H.F. (1996). In vitro complementation analysis localizes 23S rRNA posttranscriptional modifications that are required for Escherichia coli 50S ribosomal subunit assembly and function. Rna *2*, 1011-1021.

Gualerzi, C., Risuleo, G., and Pon, C.L. (1977). Initial rate kinetic analysis of the mechanism of initiation complex formation and the role of initiation factor IF-3. Biochemistry *16*, 1684-1689.

Gualerzi, C.O., and Pon, C.L. (1990). Initiation of mRNA translation in prokaryotes. Biochemistry *29*, 5881-5889.

Guo, Q., Yuan, Y., Xu, Y., Feng, B., Liu, L., Chen, K., Sun, M., and Yang, Z. (2011). Structural basis for the function of a small GTPase RsgA on the 30S ribosomal subunit maturation revealed by cryoelectron microscopy. Proceedings of the National Academy of Sciences *108*, 13100-13105.

Gur, E., and Sauer, R.T. (2008a). Evolution of the ssrA degradation tag in Mycoplasma: Specificity switch to a different protease. Proceedings of the National Academy of Sciences *105*, 16113.

Gur, E., and Sauer, R.T. (2008b). Evolution of the ssrA degradation tag in Mycoplasma: Specificity switch to a different protease. Proceedings of the National Academy of Sciences.

Gustafsson, C., and Persson, B.C. (1998). Identification of the rrmA gene encoding the 23S rRNA m1G745 methyltransferase in Escherichia coli and characterization of an m1G745-deficient mutant. Journal of bacteriology *180*, 359-365.

Gutgsell, N.S., and Jain, C. (2012). Gateway role for rRNA precursors in ribosome assembly. Journal of bacteriology *194*, 6875-6882.

Guthrie, C., Nashimoto, H., and Nomura, M. (1969). Structure and function of E. coli ribosomes, VIII. Cold-sensitive mutants defective in ribosome assembly. Proc. Natl. Acad. Sci. U.S.A. *63*, 384-391.

Hadjiolov, A.A., and Nikolaev, N. (1978). Maturation of ribosomal ribonucleic acids and the biogenesis of ribosomes. Progress in Biophysics and Molecular Biology *31*, 95-144.

Hagerman, P.J. (1997). Flexibility of RNA. Annu Rev Biophys Biomol Struct *26*, 139-156.

Haggerty, T.J., and Lovett, S.T. (1997). IF3-mediated suppression of a GUA initiation codon mutation in the recJ gene of Escherichia coli. J Bacteriol *179*, 6705-6713.

Hansen, J.L., Schmeing, T.M., Moore, P.B., and Steitz, T.A. (2002). Structural insights into peptide bond formation. Proceedings of the National Academy of Sciences *99*, 11670.

Hassouna, N., Michot, B., and Bachellerie, J.P. (1984). The complete nucleotide sequence of mouse 28S rRNA gene. Implications for the process of size increase of the large subunit rRNA in higher eukaryotes. Nucleic Acids Research *12*, 3563-3583.

Hayes, F., and Hayes, D.H. (1971). Biosynthesis of ribosomes in E. coli. I. Properties of ribosomal precursor particles and their RNA components. Biochimie *53*, 369-382.

Hecht, A., Glasgow, J., Jaschke, P.R., Bawazer, L.A., Munson, M.S., Cochran, J.R., Endy, D., and Salit, M. (2017). Measurements of translation initiation from all 64 codons in E. coli. Nucleic Acids Research *45*, 3615-3626.

Helser, T.L., Davies, J.E., and Dahlberg, J.E. (1972). Mechanism of Kasugamycin Resistance in Escherichia coli. Nature New Biology *235*, 6.

Herries, D.G. (1985). Enzyme structure and mechanism Biochemical Education *13*, 146-146.

Herschlag, D. (1995). RNA Chaperones and the RNA Folding Problem. Journal of Biological Chemistry *270*, 20871-20874.

Himeno, H., Hanawa-Suetsugu, K., Kimura, T., Takagi, K., Sugiyama, W., Shirata, S., Mikami, T., Odagiri, F., Osanai, Y., Watanabe, D., *et al.* (2004). A novel GTPase activated by the small subunit of ribosome. Nucleic Acids Research *32*, 5303-5309.

Hirashima, A., and Kaji, A. (1972). Factor-dependent release of ribosomes from messenger RNA. Requirement for two heat-stable factors. J Mol Biol *65*, 43-58.

Hirvonen, C.A., Ross, W., Wozniak, C.E., Marasco, E., Anthony, J.R., Aiyar, S.E., Newburn, V.H., and Gourse, R.L. (2001). Contributions of UP elements and the transcription factor FIS to expression from the seven rrn P1 promoters in Escherichia coli. Journal of bacteriology *183*, 6305-6314.

Holthusen, K., Talaty, P., and Everly, D.N. (2015). Regulation of Latent Membrane Protein 1 Signaling through Interaction with Cytoskeletal Proteins. Journal of Virology *89*, 7277.

Hsiao, C., Mohan, S., Kalahar, B.K., and Williams, L.D. (2009). Peeling the Onion: Ribosomes Are Ancient Molecular Fossils. Molecular Biology and Evolution *26*, 2415-2425.

Hsu, S.-T.D., Blaser, G., Behrens, C., Cabrita, L.D., Dobson, C.M., and Jackson, S.E. (2010). Folding Study of Venus Reveals a Strong Ion Dependence of Its Yellow Fluorescence under Mildly Acidic Conditions. Journal of Biological Chemistry *285*, 4859-4869.

Hury, J., Nagaswamy, U., Larios-Sanz, M., and Fox, G.E. (2006). Ribosome origins: the relative age of 23S rRNA Domains. Origins of life and evolution of the biosphere : the journal of the International Society for the Study of the Origin of Life *36*, 421-429.

Hussain, T., Llácer, J.L., Wimberly, B.T., Kieft, J.S., and Ramakrishnan, V. (2016). Large-Scale Movements of IF3 and tRNA during Bacterial Translation Initiation. Cell *167*, 133-144.e113.

Hutchison, C.A., Peterson, S.N., Gill, S.R., Cline, R.T., White, O., Fraser, C.M., Smith, H.O., and Craig Venter, J. (1999). Global Transposon Mutagenesis and a Minimal Mycoplasma Genome. Science *286*, 2165.

Hwang, J., and Inouye, M. (2006a). The tandem GTPase, Der, is essential for the biogenesis of 50S ribosomal subunits in Escherichia coli. Molecular microbiology *61*, 1660-1672.

Hwang, J., and Inouye, M. (2006b). The tandem GTPase, Der, is essential for the biogenesis of 50S ribosomal subunits in Escherichia coli. Molecular microbiology *61*, 1660-1672.

Inoue, K., Alsina, J., Chen, J., and Inouye, M. (2003). Suppression of defective ribosome assembly in a rbfA deletion mutant by overexpression of Era, an essential GTPase in Escherichia coli. Molecular microbiology *48*, 1005-1016.

Iost, I., and Dreyfus, M. (2006). DEAD-box RNA helicases in Escherichia coli. Nucleic Acids Res *34*, 4189-4197.

Jacob, Asha I., Köhrer, C., Davies, Bryan W., RajBhandary, Uttam L., and Walker, Graham C. (2013). Conserved Bacterial RNase YbeY Plays Key Roles in 70S Ribosome Quality Control and 16S rRNA Maturation. Molecular Cell *49*, 427-438.

Jagessar, K.L., and Jain, C. (2010a). Functional and molecular analysis of Escherichia coli strains lacking multiple DEAD-box helicases. RNA *16*, 1386-1392.

Jagessar, K.L., and Jain, C. (2010b). Functional and molecular analysis of Escherichia coli strains lacking multiple DEAD-box helicases. RNA *16*, 1386-1392.

Janssen, B.D., and Hayes, C.S. (2012). The tmRNA ribosome-rescue system. In Advances in Protein Chemistry and Structural Biology, A. Marintchev, ed. (Academic Press), pp. 151-191.

Jomaa, A., Stewart, G., Martín-Benito, J., Zielke, R., Campbell, T.L., Maddock, J.R., Brown, E.D., and Ortega, J. (2011). Understanding ribosome assembly: the structure of in vivo assembled immature 30S subunits revealed by cryo-electron microscopy. RNA *17*, 697-709.

Jones, P.G., and Inouye, M. (1996). RbfA, a 30S ribosomal binding factor, is a cold-shock protein whose absence triggers the cold-shock response. Molecular microbiology *21*, 1207-1218.

Julián, P., Milon, P., Agirrezabala, X., Lasso, G., Gil, D., Rodnina, M.V., and Valle, M. (2011). The Cryo-EM Structure of a Complete 30S Translation Initiation Complex from Escherichia coli. PLOS Biology *9*, e1001095.

Jungbluth, M., Renicke, C., and Taxis, C. (2010). Targeted protein depletion in Saccharomyces cerevisiae by activation of a bidirectional degron. BMC Systems Biology *4*, 176.

Kamath, R.S., Fraser, A.G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., *et al.* (2003). Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature *421*, 231.

Karbstein, K. (2013). Quality control mechanisms during ribosome maturation. Trends Cell Biol *23*, 242-250.

Karimi, R., Pavlov, M.Y., Buckingham, R.H., and Ehrenberg, M. (1999). Novel Roles for Classical Factors at the Interface between Translation Termination and Initiation. Molecular Cell *3*, 601-609.

Katunin, V.I., Muth, G.W., Strobel, S.A., Wintermeyer, W., and Rodnina, M.V. (2002). Important Contribution to Catalysis of Peptide Bond Formation by a Single Ionizing Group within the Ribosome. Molecular Cell *10*, 339-346.

Kaufmann, A., and Knop, M. (2011). Genomic Promoter Replacement Cassettes to Alter Gene Expression in the Yeast Saccharomyces cerevisiae. In Strain Engineering: Methods and Protocols, J.A. Williams, ed. (Totowa, NJ: Humana Press), pp. 275-294.

Kavaliauskas, D., Chen, C., Liu, W., Cooperman, B.S., Goldman, Y.E., and Knudsen, C.R. (2018). Structural dynamics of translation elongation factor Tu during aa-tRNA delivery to the ribosome. Nucleic Acids Res *46*, 8651-8661.

Kavaliauskas, D., Nissen, P., and Knudsen, C.R. (2012). The busiest of all ribosomal assistants: elongation factor Tu. Biochemistry *51*, 2642-2651.

Kelly, C.L., Taylor, G.M., Hitchcock, A., Torres-Méndez, A., and Heap, J.T. (2018). A Rhamnose-Inducible System for Precise and Temporal Control of Gene Expression in Cyanobacteria. ACS Synthetic Biology *7*, 1056-1066.

Kerppola, T.K. (2008). Bimolecular Fluorescence Complementation (BiFC) Analysis as a Probe of Protein Interactions in Living Cells. Annual Review of Biophysics *37*, 465-487.

Khatter, H., Myasnikov, A.G., Natchiar, S.K., and Klaholz, B.P. (2015). Structure of the human 80S ribosome. Nature *520*, 640.

Kim, D.I., Jensen, S.C., Noble, K.A., Kc, B., Roux, K.H., Motamedchaboki, K., and Roux, K.J. (2016). An improved smaller biotin ligase for BioID proximity labeling. Mol Biol Cell *27*, 1188-1196.

Kim, D.I., Kc, B., Zhu, W., Motamedchaboki, K., Doye, V., and Roux, K.J. (2014). Probing nuclear pore complex architecture with proximity-dependent biotinylation. Proceedings of the National Academy of Sciences *111*, E2453.

Kirillov, S.V., Wower, J., Hixson, S.S., and Zimmermann, R.A. (2002). Transit of tRNA through the Escherichia coli ribosome: cross-linking of the 3′ end of tRNA to ribosomal proteins at the P and E sites. FEBS Letters *514*, 60-66.

Klein, D.J., Moore, P.B., and Steitz, T.A. (2004). The Roles of Ribosomal Proteins in the Structure Assembly, and Evolution of the Large Ribosomal Subunit. Journal of Molecular Biology *340*, 141-177.

Koch, A.L., and Deppe, C.S. (1971). In vivo assay of protein synthesizing capacity of Escherichia coli from slowly growing chemostat cultures. Journal of Molecular Biology *55*, 549-562.

Korepanov, A.P., Gongadze, G.M., Garber, M.B., Court, D.L., and Bubunenko, M.G. (2007). Importance of the 5 S rRNA-binding Ribosomal Proteins for Cell Viability and Translation in Escherichia coli. Journal of Molecular Biology *366*, 1199-1208.

Korobeinikova, A.V., Garber, M.B., and Gongadze, G.M. (2012). Ribosomal proteins: structure, function, and evolution. Biochemistry. Biokhimiia *77*, 562-574.

Korostelev, A.A. (2011). Structural aspects of translation termination on the ribosome. RNA *17*, 1409-1421.

Lake, J.A. (1976). Ribosome structure determined by electron microscopy of Escherichia coli small subunits, large subunits and monomeric ribosomes. Journal of Molecular Biology *105*, 131-159.

Lang, K., Erlacher, M., Wilson, D.N., Micura, R., and Polacek, N. (2008). The Role of 23S Ribosomal RNA Residue A2451 in Peptide Bond Synthesis Revealed by Atomic Mutagenesis. Chemistry & Biology *15*, 485-492.

Lareau, L.F., Hite, D.H., Hogan, G.J., and Brown, P.O. Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments.

Laursen, B.S., Sørensen, H.P., Mortensen, K.K., and Sperling-Petersen, H.U. (2005). Initiation of Protein Synthesis in Bacteria. Microbiology and Molecular Biology Reviews *69*, 101.

Lee, J.W., Beebe, K., Nangle, L.A., Jang, J., Longo-Guess, C.M., Cook, S.A., Davisson, M.T., Sundberg, J.P., Schimmel, P., and Ackerman, S.L. (2006). Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration. Nature *443*, 50-55.

Leive, L. (1965). RNA degradation and the assembly of ribosomes in actinomycin-treated Escherichia coli. Journal of Molecular Biology *13*, 862-875.

Lemke, J.J., Sanchez-Vazquez, P., Burgos, H.L., Hedberg, G., Ross, W., and Gourse, R.L. (2011). Direct regulation of Escherichia coli ribosomal protein promoters by the transcription factors ppGpp and DksA. Proceedings of the National Academy of Sciences of the United States of America *108*, 5712-5717.

Leong, V., Kent, M., Jomaa, A., and Ortega, J. (2013). Escherichia coli rimM and yjeQ null strains accumulate immature 30S subunits of similar structure and protein complement. Rna *19*, 789-802.

Lerner, C.G., and Inouye, M. (1991). Pleiotropic changes resulting from depletion of Era, an essential GTP-binding protein in Escherichia coli. Molecular Microbiology *5*, 951-957.

Lewandowski, L.J., and Brownstein, B.L. (1966). An altered pattern of ribosome synthesis in a mutant of E. coli. Biochem. Biophys. Res. Commun. *25*, 554-561.

Li, G.-W., Burkhardt, D., Gross, C., and Weissman, Jonathan S. (2014). Quantifying Absolute Protein Synthesis Rates Reveals Principles Underlying Allocation of Cellular Resources. Cell *157*, 624-635.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Li, Z., and Deutscher, M.P. (1995). The tRNA processing enzyme RNase T is essential for maturation of 5S RNA. Proceedings of the National Academy of Sciences of the United States of America *92*, 6883-6886.

Li, Z., Pandit, S., and Deutscher, M.P. (1999a). Maturation of 23S ribosomal RNA requires the exoribonuclease RNase T. RNA *5*, 139-146.

Li, Z., Pandit, S., and Deutscher, M.P. (1999b). RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. The EMBO Journal *18*, 2878-2885.

Lin, T., Troy, E.B., Hu, L.T., Gao, L., and Norris, S.J. (2014). Transposon mutagenesis as an approach to improved understanding of Borrelia pathogenesis and biology. Frontiers in Cellular and Infection Microbiology *4*, 63.

Lindahl, L. (1973a). Two New Ribosomal Precursor Particles in E. coli. Nature New Biology *243*, 170.

Lindahl, L. (1973b). Two new ribosomal precursor particles in E. coli. Nature: New biology *243*, 170-172.

Lindahl, L. (1975). Intermediates and time kinetics of the in vivo assembly of Escherichia coli ribosomes. Journal of Molecular Biology *92*, 15-37.

Liu, X., Salokas, K., Tamene, F., Jiu, Y., Weldatsadik, R.G., Ohman, T., and Varjosalo, M. (2018). An AP-MS- and BioID-compatible MAC-tag enables comprehensive mapping of protein interactions and subcellular localizations. Nature communications *9*, 1188.

Loftfield, R.B. (1957). The biosynthesis of protein. Progress in biophysics and biophysical chemistry *8*, 347-386.

López-Alonso, J.P., Kaminishi, T., Kikuchi, T., Hirata, Y., Iturrioz, I., Dhimole, N., Schedlbauer, A., Hase, Y., Goto, S., Kurita, D., *et al.* (2017). RsgA couples the maturation state of the 30S ribosomal decoding center to activation of its GTPase pocket. Nucleic Acids Research *45*, 6945-6959.

Lovgren, J.M., Bylund, G.O., Srivastava, M.K., Lundberg, L.A., Persson, O.P., Wingsle, G., and Wikstrom, P.M. (2004). The PRC-barrel domain of the ribosome maturation protein RimM mediates binding to ribosomal protein S19 in the 30S ribosomal subunits. Rna *10*, 1798-1812.

Lu, Q., and Inouye, M. (1998). The gene for 16S rRNA methyltransferase (ksgA) functions as a multicopy suppressor for a cold-sensitive mutant of era, an essential RAS-like GTP-binding protein in Escherichia coli. Journal of bacteriology *180*, 5243-5246.

Maguire, B.A., Beniaminov, A.D., Ramu, H., Mankin, A.S., and Zimmermann, R.A. (2005). A Protein Component at the Heart of an RNA Machine: The Importance of Protein L27 for the Function of the Bacterial Ribosome. Molecular Cell *20*, 427-435.

Maitra, A., and Dill, K.A. (2015). Bacterial growth laws reflect the evolutionary importance of energy efficiency. Proceedings of the National Academy of Sciences *112*, 406.

Mangiarotti, G., Apirion, D., Schlessinger, D., and Silengo, L. (1968). Biosynthetic precursors of 30S and 50S ribosomal particles in Escherichia coli. Biochemistry *7*, 456-472.

Marks, J., Kannan, K., Roncase, E.J., Klepacki, D., Kefi, A., Orelle, C., Vázquez-Laslop, N., and Mankin, A.S. (2016). Context-specific inhibition of translation by ribosomal antibiotics targeting the peptidyl transferase center. Proceedings of the National Academy of Sciences *113*, 12150-12155.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal; Vol 17, No 1: Next Generation Sequencing Data Analysis.

Martin, R., Straub, A.U., Doebele, C., and Bohnsack, M.T. (2013). DExD/H-box RNA helicases in ribosome biogenesis. RNA Biology *10*, 4-18.

Mašek, T., Valášek, L., and Pospíšek, M. (2011). Polysome Analysis and RNA Purification from Sucrose Gradients. In RNA: Methods and Protocols, H. Nielsen, ed. (Totowa, NJ: Humana Press), pp. 293-309.

McCarthy, B.J., Britten, R.J., and Roberts, R.B. (1962). The Synthesis of Ribosomes in E. coli: III. Synthesis of Ribosomal RNA. Biophys J *2*, 57-82.

Mehmood, S., Allison, T.M., and Robinson, C.V. (2015). Mass Spectrometry of Protein Complexes: From Origins to Applications. Annual Review of Physical Chemistry *66*, 453-474.

Mehta, P., Woo, P., Venkataraman, K., and Karzai, A.W. (2012). Ribosome Purification Approaches for Studying Interactions of Regulatory Proteins and RNAs with the Ribosome. In Bacterial Regulatory RNA: Methods and Protocols, K.C. Keiler, ed. (Totowa, NJ: Humana Press), pp. 273-289.

Mehus, A.A., Anderson, R.H., and Roux, K.J. (2016). BioID Identification of Lamin-Associated Proteins. Methods in enzymology *569*, 3-22.

Miller, J.H. (1992). Experiments in molecular genetics, 11. print edn (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory).

Milon, P., Konevega, A.L., Gualerzi, C.O., and Rodnina, M.V. (2008). Kinetic Checkpoint at a Late Step in Translation Initiation. Molecular Cell *30*, 712-720.

Milon, P., Maracci, C., Filonava, L., Gualerzi, C.O., and Rodnina, M.V. (2012). Real-time assembly landscape of bacterial 30S translation initiation complex. Nature structural & molecular biology *19*, 609-615.

Milon, P., Tischenko, E., Tomsic, J., Caserta, E., Folkers, G., La Teana, A., Rodnina, M.V., Pon, C.L., Boelens, R., and Gualerzi, C.O. (2006). The nucleotide-binding site of bacterial translation initiation factor 2 (IF2) as a metabolic sensor. Proceedings of the National Academy of Sciences *103*, 13962-13967.

Minden, J., Namba, R., Mergliano, J., and Cambridge, S. (2000). Photoactivated Gene Expression for Cell Fate Mapping and Cell Manipulation. Science *2000*, pl1.

Misra, T.K., and Apirion, D. (1979). RNase E, an RNA processing enzyme from Escherichia coli. The Journal of biological chemistry *254*, 11154-11159.

Mitra, K., and Frank, J. (2006). RIBOSOME DYNAMICS: Insights from Atomic Structure Modeling into Cryo-Electron Microscopy Maps. Annual Review of Biophysics and Biomolecular Structure *35*, 299-317.

Mizushima, S., and Nomura, M. (1970). Assembly Mapping of 30S Ribosomal Proteins from E. coli. Nature *226*, 1214.

Moore, P.B., and Steitz, T.A. (2003). After the ribosome structures: How does peptidyl transferase work? RNA *9*, 155-159.

Murray, H.D., Appleman, J.A., and Gourse, R.L. (2003). Regulation of the Escherichia coli rrnB P2 Promoter. Journal of Bacteriology *185*, 28-34.

Mushegian, A.R., and Koonin, E.V. (1996). A minimal gene set for cellular life derived by comparison of complete bacterial genomes. Proceedings of the National Academy of Sciences *93*, 10268.

Neidhardt, F.C. (1999). Bacterial growth: constant obsession with dN/dt. Journal of bacteriology *181*, 7405-7408.

Neidhardt, F.C., and Magasanik, B. (1960). Studies on the role of ribonucleic acid in the growth of bacteria. Biochim. Biophys. Acta *42*, 99-116.

Nierhaus, D., and Nierhaus, K.H. (1973). Identification of the Chloramphenicol-Binding Protein in Escherichia coli Ribosomes by Partial Reconstitution. Proceedings of the National Academy of Sciences *70*, 2224.

Nierhaus, K.H. (1991). The assembly of prokaryotic ribosomes. Biochimie *73*, 739-755.

Nierhaus, K.H., and Dohme, F. (1974). Total Reconstitution of Functionally Active 50S Ribosomal Subunits from Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America *71*, 4713-4717.

Nierhaus, K.H., and Montejo, V. (1973). A Protein Involved in the Peptidyltransferase Activity of Escherichia coli Ribosomes. Proceedings of the National Academy of Sciences *70*, 1931.

Nikolaev, N., Schlessinger, D., and Wellauer, P.K. (1974). 30 S pre-ribosomal RNA of Escherichia coli and products of cleavage by ribonuclease III: length and molecular weight. J Mol Biol *86*, 741-747.

Nikolaev, N., Silengo, L., and Schlessinger, D. (1973). A Role for Ribonuclease III in Processing of Ribosomal Ribonucleic Acid and Messenger Ribonucleic Acid Precursors in Escherichia coli. Journal of Biological Chemistry *248*, 7967-7969.

Nikolay, R., Schloemer, R., Mueller, S., and Deuerling, E. (2015). Fluorescence-based monitoring of ribosome assembly landscapes. BMC Molecular Biology *16*, 3.

Nikolay, R., Schloemer, R., Schmidt, S., Mueller, S., Heubach, A., and Deuerling, E. (2014). Validation of a fluorescence-based screening concept to identify ribosome assembly defects in Escherichia coli. Nucleic Acids Research *42*, e100.

Nikolay, R., Schmidt, S., Schlomer, R., Deuerling, E., and Nierhaus, K.H. (2016). Ribosome Assembly as Antimicrobial Target. Antibiotics *5*, 2.

Nishi, K., Morel-Deville, F., Hershey, J.W., Leighton, T., and Schnier, J. (1988). An eIF-4A-like protein is a suppressor of an Escherichia coli mutant defective in 50S ribosomal subunit assembly. Nature *336*, 496-498.

Nissen, P., Hansen, J., Ban, N., Moore, P.B., and Steitz, T.A. (2000). The Structural Basis of Ribosome Activity in Peptide Bond Synthesis. Science *289*, 920.

Noller, H.F., Hoffarth, V., and Zimniak, L. (1992). Unusual resistance of peptidyl transferase to protein extraction procedures. Science *256*, 1416.

Nomura, M. (1973). Assembly of bacterial ribosomes. Science *179*, 864-873.

Nomura, M., Gourse, R., and Baughman, G. (1984). Regulation of the synthesis of ribosomes and ribosomal components. Annu Rev Biochem. *53*, 75-117.

Nomura, M., and Traub, P. (1966). Organizational biosynthesis : a symposium. In Symposium on Organizational Biosynthesis (1st : 1966 : Rutgers University), H.J. Vogel, Lampen, Jay Oliver, Bryson, Vernon, ed. (Rutgers University, Academic Press).

Nomura, M., Yates, J.L., Dean, D., and Post, L.E. (1980). Feedback regulation of ribosomal protein gene expression in Escherichia coli: structural homology of ribosomal RNA and ribosomal protein MRNA. Proceedings of the National Academy of Sciences *77*, 7084.

Nord, S., Bylund, G.O., Lovgren, J.M., and Wikstrom, P.M. (2009). The RimP protein is important for maturation of the 30S ribosomal subunit. J Mol Biol *386*, 742-753.

Nowotny, V., May, R.P., and Nierhaus, K.H. (1986). Neutron-Scattering Analysis of Structural and Functional Aspects of the Ribosome: The Strategy of the Glassy Ribosome. In Structure, Function, and Genetics of Ribosomes, B. Hardesty, and G. Kramer, eds. (New York, NY: Springer New York), pp. 101-111.

O'Connor, M. (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 Research *25*, 1185-1193.

O'Connor, M., Gregory, S.T., and Dahlberg, A.E. (2004). Multiple defects in translation associated with altered ribosomal protein L4. Nucleic Acids Res *32*, 5750-5756.

O'Connor, M., Thomas, C.L., Zimmermann, R.A., and 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*, 1185-1193.

Ooi, H.S., Schneider,G., Chan, Ying L, Lim, Chan T., Eisenhaber, Brigit., Eisenhaber, Frank (2010). Databases of protein-protein interactions and complexes (Humana Press).

Ortega, J.P., and Hill, W.E. (1973). A molecular weight determination of the 16S ribosomal ribonucleic acid from Escherichia coli. Biochemistry *12*, 3241-3243.

Ottmann, C., Weyand, M., Wolf, A., Kuhlmann, J., and Ottmann, C. (2008). Applicability of superfolder YFP bimolecular fluorescence complementation *in vitro*. Biological Chemistry *390*, 81-90.

Palade, G.E. (1955). A Small Particulate Component of the Cytoplasm. The Journal of Biophysical and Biochemical Cytology *1*, 59-68.

Peil, L., Virumae, K., and Remme, J. (2008). Ribosome assembly in Escherichia coli strains lacking the RNA helicase DeaD/CsdA or DbpA. The FEBS journal *275*, 3772-3782.

Penczek, P.A., Grassucci, R.A., and Frank, J. (1994). The ribosome at improved resolution: New techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. Ultramicroscopy *53*, 251-270.

Pennisi, E. (1999). The Race to the Ribosome Structure. Science *285*, 2048.

Penzo, M., and Montanaro, L. (2018). Turning Uridines around: Role of rRNA Pseudouridylation in Ribosome Biogenesis and Ribosomal Function. *8*.

Peters, J.M., Colavin, A., Shi, H., Czarny, T.L., Larson, M.H., Wong, S., Hawkins, J.S., Lu, C.H.S., Koo, B.-M., Marta, E., *et al.* (2016). A Comprehensive, CRISPR-based Functional Analysis of Essential Genes in Bacteria. Cell *165*, 1493-1506.

Petrelli, D., La Teana, A., Garofalo, C., Spurio, R., Pon, C.L., and Gualerzi, C.O. (2001). Translation initiation factor IF3: two domains, five functions, one mechanism? The EMBO Journal *20*, 4560.

Petrillo, E., Godoy Herz, M.A., Barta, A., Kalyna, M., and Kornblihtt, A.R. (2014). Let there be light: Regulation of gene expression in plants. RNA Biology *11*, 1215-1220.

Polikanov, Y.S., Melnikov, S.V., Soll, D., and Steitz, T.A. (2015). Structural insights into the role of rRNA modifications in protein synthesis and ribosome assembly. Nature structural & molecular biology *22*, 342-344.

Powers, T., and Noller, H.F. (1995). Hydroxyl radical footprinting of ribosomal proteins on 16S rRNA. RNA *1*, 194-209.

Proux, F., Dreyfus, M., and Iost, I. (2011). Identification of the sites of action of SrmB, a DEAD-box RNA helicase involved in Escherichia coli ribosome assembly. Molecular Microbiology *82*, 300-311.

Pulk, A., and Cate, J.H.D. (2013). Control of Ribosomal Subunit Rotation by Elongation Factor G. Science *340*.

Qin, D., and Fredrick, K. (2009). Control of translation initiation involves a factor-induced rearrangement of helix 44 of 16S ribosomal RNA. Mol Microbiol *71*, 1239-1249.

Qin, D., Liu, Q., Devaraj, A., and Fredrick, K. (2012). Role of helix 44 of 16S rRNA in the fidelity of translation initiation. RNA *18*, 485-495.

Rajagopala, S.V., Sikorski, P., Kumar, A., Mosca, R., Vlasblom, J., Arnold, R., Franca-Koh, J., Pakala, S.B., Phanse, S., Ceol, A., *et al.* (2014). The binary protein-protein interaction landscape of Escherichia coli. Nature Biotechnology *32*, 285-290.

Ramakrishnan, V. (1986). Distribution of protein and RNA in the 30S ribosomal subunit. Science *231*, 1562.

Ramaswamy, P., and Woodson, S.A. (2009). S16 throws a conformational switch during assembly of 30S 5' domain. Nature structural & molecular biology *16*, 438-445.

Ramesh, V., Gite, S., Li, Y., and RajBhandary, U.L. (1997). Suppressor mutations in Escherichia coli methionyl-tRNA formyltransferase: Role of a 16-amino acid insertion module in initiator tRNA recognition. Proceedings of the National Academy of Sciences *94*, 13524.

Rancati, G., Moffat, J., Typas, A., and Pavelka, N. (2017). Emerging and evolving concepts in gene essentiality. Nature Reviews Genetics *19*, 34.

Rauniyar, N., and Yates, J.R. (2014). Isobaric Labeling-Based Relative Quantification in Shotgun Proteomics. Journal of Proteome Research *13*, 5293-5309.

Razi, A., Guarné, A., and 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. Proceedings of the National Academy of Sciences *114*, E3396-E3403.

Régnier, P., and Grunberg-Manago, M. (1989). Cleavage by RNase III in the transcripts of the met Y-nus-A-infB operon of Escherichia coli releases the tRNA and initiates the decay of the downstream mRNA. J Mol Biol *210*, 293-302.

Régnier, P., and Grunberg-Manago, M. (1990). RNase III cleavages in non-coding leaders of Escherichia coli transcripts control mRNA stability and genetic expression. Biochimie *72*, 825-834.

Reuber, T.L., Long, S., and Walker, G.C. (1991). Regulation of Rhizobium meliloti exo genes in free-living cells and in planta examined by using TnphoA fusions. Journal of Bacteriology *173*, 426.

Reuber, T.L., and Walker, G.C. (1993). Biosynthesis of succinoglycan, a symbiotically important exopolysaccharide of Rhizobium meliloti. Cell *74*, 269-280.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Séraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nature Biotechnology *17*, 1030.

Roberts, R.B. (1958). Microsomal particles and protein synthesis (Pergamon Press).

Robertson, M.P., and Joyce, G.F. (2012). The Origins of the RNA World. Cold Spring Harbor Perspectives in Biology *4*.

Robida, A.M., and Kerppola, T.K. (2009). Bimolecular Fluorescence Complementation Analysis of Inducible Protein Interactions: Effects of Factors Affecting Protein Folding on Fluorescent Protein Fragment Association. Journal of Molecular Biology *394*, 391-409.

Ross, W., Thompson, J.F., Newlands, J.T., and Gourse, R.L. (1990). E.coli Fis protein activates ribosomal RNA transcription in vitro and in vivo. The EMBO journal *9*, 3733-3742.

Roux, K.J., Kim, D.I., Burke, B., and May, D.G. (2018). BioID: A Screen for Protein-Protein Interactions. Current Protocols in Protein Science *91*, 19.23.11-19.23.15.

Roux, K.J., Kim, D.I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. The Journal of Cell Biology *196*, 801.

Roy-Chaudhuri, B., Kirthi, N., Kelley, T., and Culver, G.M. (2008). Suppression of a cold-sensitive mutation in ribosomal protein S5 reveals a role for RimJ in ribosome biogenesis. Molecular microbiology *68*, 1547-1559.

Ryals, J., Little, R., and Bremer, H. (1982). Control of rRNA and tRNA syntheses in Escherichia coli by guanosine tetraphosphate. Journal of bacteriology *151*, 1261-1268.

Samaha, R.R., O'Brien, B., O'Brien, T.W., and Noller, H.F. (1994). Independent in vitro assembly of a ribonucleoprotein particle containing the 3' domain of 16S rRNA. Proceedings of the National Academy of Sciences of the United States of America *91*, 7884-7888.

Sashital, D.G., Greeman, C.A., Lyumkis, D., Potter, C.S., Carragher, B., and Williamson, J.R. (2014). A combined quantitative mass spectrometry and electron microscopy analysis of ribosomal 30S subunit assembly in E. coli. eLife *3*, e04491.

Satterthwait, A.C., and Jencks, W.P. (1974). The mechanism of the aminolysis of acetate esters. Journal of the American Chemical Society *96*, 7018-7031.

Schaechter, M., Maaloe, O., and Kjeldgaard, N.O. (1958). Dependency on medium and temperature of cell size and chemical composition during balanced grown of Salmonella typhimurium. Journal of general microbiology *19*, 592-606.

Schleif, R., Hess, W., Finkelstein, S., and Ellis, D. (1973). Induction kinetics of the L-arabinose operon of Escherichia coli. Journal of Bacteriology *115*, 9-14.

Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., *et al.* (2000). Structure of Functionally Activated Small Ribosomal Subunit at 3.3 Å Resolution. Cell *102*, 615-623.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nature Methods *9*, 671.

Schnider, C.B., Bausch-Fluck, D., Brühlmann, F., Heussler, V.T., and Burda, P.-C. (2018). BioID Reveals Novel Proteins of the Plasmodium Parasitophorous Vacuole Membrane. mSphere *3*, e00522-00517.

Schuwirth, B.S., Borovinskaya, M.A., Hau, C.W., Zhang, W., Vila-Sanjurjo, A., Holton, J.M., and Cate, J.H.D. (2005). Structures of the Bacterial Ribosome at 3.5 Å Resolution. Science *310*, 827.

Scott, M., and Hwa, T. (2011). Bacterial growth laws and their applications. Current opinion in biotechnology *22*, 559-565.

Seeliger, J.C., Topp, S., Sogi, K.M., Previti, M.L., Gallivan, J.P., and Bertozzi, C.R. (2012). A Riboswitch-Based Inducible Gene Expression System for Mycobacteria. PLOS ONE *7*, e29266.

Selmer, M., Dunham, C.M., Murphy, F.V., Weixlbaumer, A., Petry, S., Kelley, A.C., Weir, J.R., and Ramakrishnan, V. (2006). Structure of the 70S Ribosome Complexed with mRNA and tRNA. Science *313*, 1935.

Semrad, K., and Green, R. (2002). Osmolytes stimulate the reconstitution of functional 50S ribosomes from in vitro transcripts of Escherichia coli 23S rRNA. RNA *8*, 401-411.

Serdyuk, I.N., Agalarov, S.C., Sedelnikova, S.E., Spirin, A.S., and May, R.P. (1983). Shape and compactness of the isolated ribosomal 16 S RNA and its complexes with ribosomal proteins. J Mol Biol *169*, 409-425.

Sergiev, P.V., Aleksashin, N.A., Chugunova, A.A., Polikanov, Y.S., and Dontsova, O.A. (2018). Structural and evolutionary insights into ribosomal RNA methylation. Nature Chemical Biology *14*, 226.

Shajani, Z., Sykes, M.T., and Williamson, J.R. (2011). Assembly of Bacterial Ribosomes. Annual Review of Biochemistry *80*, 501-526.

Shan, G. (2010). RNA interference as a gene knockdown technique. The International Journal of Biochemistry & Cell Biology *42*, 1243-1251.

Sharma, M.R., Barat, C., Wilson, D.N., Booth, T.M., Kawazoe, M., Hori-Takemoto, C., Shirouzu, M., Yokoyama, S., Fucini, P., and Agrawal, R.K. (2005). Interaction of Era with the 30S Ribosomal Subunit: Implications for 30S Subunit Assembly. Molecular Cell *18*, 319-329.

Sharrock, R.A., Gourse, R.L., and Nomura, M. (1985). Defective antitermination of rRNA transcription and derepression of rRNA and tRNA synthesis in the nusB5 mutant of Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America *82*, 5275-5279.

Shcherbakova, I., Mitra, S., Laederach, A., and Brenowitz, M. (2008). Energy barriers, pathways and dynamics during folding of large, multi-domain RNAs. Current opinion in chemical biology *12*, 655-666.

Shetty, S., and Varshney, U. (2016a). An evolutionarily conserved element in initiator tRNAs prompts ultimate steps in ribosome maturation. Proceedings of the National Academy of Sciences *113*, E6126-E6134.

Shetty, S., and Varshney, U. (2016b). An evolutionarily conserved element in initiator tRNAs prompts ultimate steps in ribosome maturation. Proceedings of the National Academy of Sciences of the United States of America *113*, E6126-E6134.

Shi, X., Chiu, K., Ghosh, S., and Joseph, S. (2009). Bases in 16S rRNA Important for Subunit Association, tRNA Binding, and Translocation. Biochemistry *48*, 6772-6782.

Shimamoto, T., and Inouye, M. (1996). Mutational analysis of Era, an essential GTP-binding protein of Escherichia coli. FEMS Microbiology Letters *136*, 57-62.

Shirokikh, N.E., Archer, S.K., Beilharz, T.H., Powell, D., and Preiss, T. (2017). Translation complex profile sequencing to study the in vivo dynamics of mRNA–ribosome interactions during translation initiation, elongation and termination. Nature Protocols *12*, 697-731.

Shoji, S., Dambacher, C.M., Shajani, Z., Williamson, J.R., and Schultz, P.G. (2011). Systematic Chromosomal Deletion of Bacterial Ribosomal Protein Genes. Journal of Molecular Biology *413*, 751-761.

Sieber, G., and Nierhaus, K.H. (1978). Kinetic and thermodynamic parameters of the assembly in vitro of the large subunit from Escherichia coli ribosomes. Biochemistry *17*, 3505-3511.

Sievers, A., Beringer, M., Rodnina, M.V., and Wolfenden, R. (2004). The ribosome as an entropy trap. Proceedings of the National Academy of Sciences of the United States of America *101*, 7897.

Siibak, T., Peil, L., Xiong, L., Mankin, A., Remme, J., and Tenson, T. (2009). Erythromycin- and Chloramphenicol-Induced Ribosomal Assembly Defects Are Secondary Effects of Protein Synthesis Inhibition. Antimicrobial Agents and Chemotherapy *53*, 563.

Simonović, M., and Steitz, T.A. (2008). Peptidyl-CCA deacylation on the ribosome promoted by induced fit and the O3′-hydroxyl group of A76 of the unacylated A-site tRNA. RNA *14*, 2372-2378.

Sirdeshmukh, R., and Schlessinger, D. (1985). Ordered processing of Escherichia coli 23S rRNA in vitro. Nucleic Acids Research *13*, 5041-5054.

Sousa, F.L., Nelson-Sathi, S., and Martin, W.F. (2016). One step beyond a ribosome: The ancient anaerobic core. Biochimica et Biophysica Acta (BBA) - Bioenergetics *1857*, 1027-1038.

Spahn, C.M.T., Beckmann, R., Eswar, N., Penczek, P.A., Sali, A., Blobel, G., and Frank, J. (2001). Structure of the 80S Ribosome from Saccharomyces cerevisiae—tRNA-Ribosome and Subunit-Subunit Interactions. Cell *107*, 373-386.

Spillmann, S., Dohme, F., and Nierhaus, K.H. (1977). Assembly in vitro of the 50 S subunit from Escherichia coli ribosomes: proteins essential for the first heat-dependent conformational change. J Mol Biol *115*, 513-523.

Srivastava, A.K., and Schlessinger, D. (1988). Coregulation of processing and translation: mature 5' termini of Escherichia coli 23S ribosomal RNA form in polysomes. Proceedings of the National Academy of Sciences *85*, 7144.

Srivatsan, A., and Wang, J.D. (2008). Control of bacterial transcription, translation and replication by (p)ppGpp. Current Opinion in Microbiology *11*, 100-105.

St-Pierre, F., Cui, L., Priest, D.G., Endy, D., Dodd, I.B., and Shearwin, K.E. (2013). One-step cloning and chromosomal integration of DNA. ACS Synthetic Biology *2*, 537-541.

Staehelin, T., and Meselson, M. (1966). In vitro recovery o ribosomes and of synthetic activity from synthetically inactive ribosomal subunits. J Mol Biol *16*, 245-249.

Starosta, A.L., Lassak, J., Jung, K., and Wilson, D.N. (2014). The bacterial translation stress response. FEMS microbiology reviews *38*, 1172-1201.

Steitz, J.A., and 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. Proceedings of the National Academy of Sciences of the United States of America *72*, 4734-4738.

Steitz, T.A. (2008). A structural understanding of the dynamic ribosome machine. Nature reviews. Molecular cell biology *9*, 242-253.

Stelling, J., Sauer, U., Szallasi, Z., Doyle, F.J., 3rd, and Doyle, J. (2004). Robustness of cellular functions. Cell *118*, 675-685.

Stern, S., Powers, T., Changchien, L.M., and Noller, H.F. (1989). RNA-protein interactions in 30S ribosomal subunits: folding and function of 16S rRNA. Science *244*, 783-790.

Stoffler, G., and Stoffler-Meilicke, M. (1984). Immunoelectron microscopy of ribosomes. Annual review of biophysics and bioengineering *13*, 303-330.

Stokes, J.M., Davis, J.H., Mangat, C.S., Williamson, J.R., and Brown, E.D. (2014a). Discovery of a small molecule that inhibits bacterial ribosome biogenesis. Elife *3*, e03574.

Stokes, J.M., Davis, J.H., Mangat, C.S., Williamson, J.R., and Brown, E.D. (2014b). Discovery of a small molecule that inhibits bacterial ribosome biogenesis. Elife *3*, e03574.

Strunk, Bethany S., Novak, Megan N., Young, Crystal L., and Karbstein, K. (2012). A Translation-Like Cycle Is a Quality Control Checkpoint for Maturing 40S Ribosome Subunits. Cell *150*, 111-121.

Sulthana, S., and Deutscher, M.P. (2013). Multiple Exoribonucleases Catalyze Maturation of the 3′ Terminus of 16S Ribosomal RNA (rRNA). The Journal of Biological Chemistry *288*, 12574-12579.

Sun, C., and Woolford, J.L., Jr. (1994). The yeast NOP4 gene product is an essential nucleolar protein required for pre-rRNA processing and accumulation of 60S ribosomal subunits. Embo j *13*, 3127-3135.

Sussman, J.K., Simons, E.L., and Simons, R.W. (1996). Escherichia coli translation initiation factor 3 discriminates the initiation codon in vivo. Mol Microbiol *21*, 347-360.

Sykes, M.T., and Williamson, J.R. (2009). A Complex Assembly Landscape for the 30S Ribosomal Subunit. Annual review of biophysics *38*, 197-215.

Szewczak, A.A., Ortoleva-Donnelly, L., Ryder, S.P., Moncoeur, E., and Strobel, S.A. (1998). A minor groove RNA triple helix within the catalytic core of a group I intron. Nature structural biology *5*, 1037-1042.

Talkington, M.W., Siuzdak, G., and Williamson, J.R. (2005). An assembly landscape for the 30S ribosomal subunit. Nature *438*, 628-632.

Tan, Z.-J., and Chen, S.-J. (2011). Salt Contribution to RNA Tertiary Structure Folding Stability. Biophysical Journal *101*, 176-187.

Thurlow, B., Davis, J.H., Leong, V., Moraes, T.F., Williamson, J.R., and Ortega, J. (2016). Binding properties of YjeQ (RsgA), RbfA, RimM and Era to assembly intermediates of the 30S subunit. Nucleic Acids Research *44*, 9918-9932.

Tissieres, A., and Watson, J.D. (1958). Ribonucleoprotein particles from Escherichia coli. Nature *182*, 778-780.

Toone, W.M., Rudd, K.E., and Friesen, J.D. (1991). deaD, a new Escherichia coli gene encoding a presumed ATP-dependent RNA helicase, can suppress a mutation in rpsB, the gene encoding ribosomal protein S2. Journal of bacteriology *173*, 3291-3302.

Torres, M., Balada, J.-M., Zellars, M., Squires, C., and Squires, C.L. (2004). In Vivo Effect of NusB and NusG on rRNA Transcription Antitermination. Journal of Bacteriology *186*, 1304.

Tourigny, D.S., Fernández, I.S., Kelley, A.C., and Ramakrishnan, V. (2013). Elongation Factor G Bound to the Ribosome in an Intermediate State of Translocation. Science *340*.

Towbin, B.D., Korem, Y., Bren, A., Doron, S., Sorek, R., and Alon, U. (2017). Optimality and sub-optimality in a bacterial growth law. Nat Commun *8*, 14123.

Traub, P., and Nomura, M. (1968a). Structure and function of E. coli ribosomes. V. Reconstitution of functionally active 30S ribosomal particles from RNA and proteins. Proceedings of the National Academy of Sciences of the United States of America *59*, 777-784.

Traub, P., and Nomura, M. (1968b). Structure and function of E. coli ribosomes. V. Reconstitution of functionally active 30S ribosomal particles from RNA and proteins. Proceedings of the National Academy of Sciences *59*, 777.

Traub, P., and Nomura, M. (1969). Studies on the Assembly of Ribosomes in vitro. Cold Spring Harbor Symposia on Quantitative Biology *34*, 63-67.

Treiber, D.K., and Williamson, J.R. (1999). Exposing the kinetic traps in RNA folding. Current opinion in structural biology *9*, 339-345.

Trobro, S., and Åqvist, J. (2005). Mechanism of peptide bond synthesis on the ribosome. Proceedings of the National Academy of Sciences of the United States of America *102*, 12395.

Trubetskoy, D., Proux, F., Allemand, F., Dreyfus, M., and Iost, I. (2009). SrmB, a DEAD-box helicase involved in Escherichia coli ribosome assembly, is specifically targeted to 23S rRNA in vivo. Nucleic Acids Research *37*, 6540-6549.

Ts'o, P.O.P. (1962). The Ribosomes-Ribonucleoprotein Particles. Annual Review of Plant Physiology *13*, 45-80.

Tsai, A., Petrov, A., Marshall, R.A., Korlach, J., Uemura, S., and Puglisi, J.D. (2012). Heterogeneous pathways and timing of factor departure during translation initiation. Nature *487*, 390.

Turtola, M., and Belogurov, G.A. (2016). NusG inhibits RNA polymerase backtracking by stabilizing the minimal transcription bubble. eLife *5*, e18096.

Valášek, L., Szamecz, B., Hinnebusch, A.G., and Nielsen, K.H. (2007). Chapter Eight - In Vivo Stabilization of Preinitiation Complexes by Formaldehyde Cross-Linking. In Methods in Enzymology, J. Lorsch, ed. (Academic Press), pp. 163-183.

Verstraeten, N., Knapen, W.J., Kint, C.I., Liebens, V., Van den Bergh, B., Dewachter, L., Michiels, J.E., Fu, Q., David, C.C., Fierro, A.C., *et al.* (2015). Obg and Membrane Depolarization Are Part of a Microbial Bet-Hedging Strategy that Leads to Antibiotic Tolerance. Mol Cell *59*, 9-21.

von Mering, C., Jensen, L.J., Snel, B., Hooper, S.D., Krupp, M., Foglierini, M., Jouffre, N., Huynen, M.A., and Bork, P. (2005). STRING: known and predicted protein-protein associations, integrated and transferred across organisms. Nucleic acids research *33*, D433-437.

Wachi, M., Umitsuki, G., Shimizu, M., Takada, A., and Nagai, K. (1999). Escherichia coli cafA Gene Encodes a Novel RNase, Designated as RNase G, Involved in Processing of the 5′ End of 16S rRNA. Biochemical and Biophysical Research Communications *259*, 483-488.

Ware, V.C., Tague, B.W., Clark, C.G., Gourse, R.L., Brand, R.C., and Gerbi, S.A. (1983). Sequence analysis of 28S ribosomal DNA from the amphibian Xenopus laevis. Nucleic Acids Research *11*, 7795-7817.

Watson, J.D., Baker, Tania. A, Bell, Stephen P. , Gann Alexander ,Levine Michael, and Losick Richard M. (2004). Molecular Biology of the Gene, 7th edn.

Weber, P.C., Ohlendorf, D.H., Wendoloski, J.J., and Salemme, F.R. (1989). Structural origins of high-affinity biotin binding to streptavidin. Science *243*, 85.

Wilson, D.N. (2014). Ribosome-targeting antibiotics and mechanisms of bacterial resistance. Nature reviews. Microbiology *12*, 35-48.

Wilson, D.N., and Nierhaus, K.H. (2003). The Ribosome through the Looking Glass. Angewandte Chemie International Edition *42*, 3464-3486.

Wilusz, C.J., and Wilusz, J. (2013). Lsm proteins and Hfq: Life at the 3' end. RNA Biology *10*, 592-601.

Wimberly, B.T., Brodersen, D.E., Clemons Jr, W.M., Morgan-Warren, R.J., Carter, A.P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. Nature *407*, 327.

Wollman, R. (2018). Robustness, Accuracy, and Cell State Heterogeneity in Biological Systems. Current opinion in systems biology *8*, 46-50.

Woodson, S.A. (2000). Recent insights on RNA folding mechanisms from catalytic RNA. Cellular and molecular life sciences : CMLS *57*, 796-808.

Wuchty, S., and Uetz, P. (2014). Protein-protein Interaction Networks of E. coli and S. cerevisiae are similar. Scientific Reports *4*, 7187.

Xu, Z., O'Farrell, H.C., Rife, J.P., and Culver, G.M. (2008). A conserved rRNA methyltransferase regulates ribosome biogenesis. Nat Struct Mol Biol *15*, 534-536.

Young, R.A., and Steitz, J.A. (1978). Complementary sequences 1700 nucleotides apart form a ribonuclease III cleavage site in Escherichia coli ribosomal precursor RNA. Proceedings of the National Academy of Sciences of the United States of America *75*, 3593-3597.

Youngman, E.M., McDonald, M.E., and Green, R. (2008). Peptide release on the ribosome: mechanism and implications for translational control. Annual review of microbiology *62*, 353-373.

Yu, H., Braun, P., Yıldırım, M.A., Lemmens, I., Venkatesan, K., Sahalie, J., Hirozane-Kishikawa, T., Gebreab, F., Li, N., Simonis, N., *et al.* (2008). High-Quality Binary Protein Interaction Map of the Yeast Interactome Network. Science *322*, 104.

Yusupova, G.Z., Yusupov, M.M., Cate, J.H.D., and Noller, H.F. (2001). The Path of Messenger RNA through the Ribosome. Cell *106*, 233-241.

Zeghouf, M., Li, J., Butland, G., Borkowska, A., Canadien, V., Richards, D., Beattie, B., Emili, A., and Greenblatt, J.F. (2004). Sequential Peptide Affinity (SPA) System for the Identification of Mammalian and Bacterial Protein Complexes. Journal of Proteome Research *3*, 463-468.

Zengel, J.M., and Lindahl, L. (1994). Diverse mechanisms for regulating ribosomal protein synthesis in Escherichia coli. Progress in nucleic acid research and molecular biology *47*, 331-370.

Zheng, H., Ho, P.-Y., Jiang, M., Tang, B., Liu, W., Li, D., Yu, X., Kleckner, N.E., Amir, A., and Liu, C. (2016). Interrogating the Escherichia coli cell cycle by cell dimension perturbations. Proceedings of the National Academy of Sciences *113*, 15000.

Appendix

The mass spectrometric data presented in this part of the thesis corresponds to the section 2.3.5 and Figure 2.8

Table A1: The BioID interactome of Era.BirA*. Proteins interacting with Era.BirA* were identified by mass spectrometry.

Accession	Description	Score	Coverage	#Unique Peptides	# Peptides	# PSMs
P0CE47	Elongation factor Tu 1 OS=Escherichia coli (strain K12) GN=tufA PE=1 SV=1 - [EFTU1_ECOLI]	4982.244	89.09	42	42	1455
P0AG30	Transcription termination factor Rho OS=Escherichia coli (strain K12) GN=rho PE=1 SV=1 - [RHO_ECOLI]	500.8343	70.64	37	37	169
C4ZSH9	D-tagatose-1,6-bisphosphate aldolase subunit GatZ OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=gatZ PE=3 SV=1 - [GATZ_ECOBW]	483.5135	75.71	23	23	112
B1XGY0	Translation initiation factor IF-2 OS=Escherichia coli (strain K12 / DH10B) GN=infB PE=3 SV=1 - [IF2_ECODH]	445.0075	51.35	42	42	109
P0A9P6	ATP-dependent RNA helicase DeaD OS=Escherichia coli (strain K12) GN=deaD PE=1 SV=2 - [DEAD_ECOLI]	352.5647	49.13	29	29	113
B1XBZ0	DNA-directed RNA polymerase subunit beta' OS=Escherichia coli (strain K12 / DH10B) GN=rpoC PE=3 SV=1 - [RPOC_ECODH]	343.8494	52.24	58	58	97
P16659	Proline--tRNA ligase OS=Escherichia coli (strain K12) GN=proS PE=1 SV=4 - [SYP_ECOLI]	317.7734	69.76	32	32	82
B1XBY9	DNA-directed RNA polymerase subunit beta OS=Escherichia coli (strain K12 / DH10B) GN=rpoB PE=3 SV=1 - [RPOB_ECODH]	316.0587	48.51	56	56	90
C4ZR79	30S ribosomal protein S18 OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=rpsR PE=3 SV=1 - [RS18_ECOBW]	301.9047	40	3	3	79

P06709	Bifunctional ligase/repressor BirA OS=Escherichia coli (strain K12) GN=birA PE=1 SV=1 - [BIRA_ECOLI]	263.764	75.08	20	20	70
P0A7V3	30S ribosomal protein S3 OS=Escherichia coli (strain K12) GN=rpsC PE=1 SV=2 - [RS3_ECOLI]	257.1406	60.09	16	16	71
P09373	Formate acetyltransferase 1 OS=Escherichia coli (strain K12) GN=pflB PE=1 SV=2 - [PFLB_ECOLI]	253.2653	43.03	29	29	67
P0AAI5	3-oxoacyl-[acyl-carrier-protein] synthase 2 OS=Escherichia coli (strain K12) GN=fabF PE=1 SV=2 - [FABF_ECOLI]	242.1975	52.78	16	16	82
P0A953	3-oxoacyl-[acyl-carrier-protein] synthase 1 OS=Escherichia coli (strain K12) GN=fabB PE=1 SV=1 - [FABB_ECOLI]	240.9725	62.32	16	16	58
C5A0L1	S-adenosylmethionine synthase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=metK PE=3 SV=1 - [METK_ECOBW]	236.7753	63.8	21	21	64
P0AG67	30S ribosomal protein S1 OS=Escherichia coli (strain K12) GN=rpsA PE=1 SV=1 - [RS1_ECOLI]	231.651	45.6	26	26	69
P0AES6	DNA gyrase subunit B OS=Escherichia coli (strain K12) GN=gyrB PE=1 SV=2 - [GYRB_ECOLI]	197.5915	51	29	29	48
B1X6J0	Elongation factor G OS=Escherichia coli (strain K12 / DH10B) GN=fusA PE=3 SV=1 - [EFG_ECODH]	190.7647	51.14	28	28	53
P0A9M0	Lon protease OS=Escherichia coli (strain K12) GN=lon PE=1 SV=1 - [LON_ECOLI]	181.4423	39.67	29	29	50
P24182	Biotin carboxylase OS=Escherichia coli (strain K12) GN=accC PE=1 SV=2 - [ACCC_ECOLI]	180.4462	56.12	26	26	50
C4ZUH2	50S ribosomal protein L2 OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=rplB PE=3 SV=1 - [RL2_ECOBW]	170.5628	59.34	18	18	49
P32132	GTP-binding protein TypA/BipA OS=Escherichia coli (strain K12) GN=typA PE=1 SV=2 - [TYPA_ECOLI]	161.6074	45.47	21	21	43

P09152	Respiratory nitrate reductase 1 alpha chain OS=Escherichia coli (strain K12) GN=narG PE=1 SV=4 - [NARG_ECOLI]	157.5415	34.48	27	30	51
C4ZZ10	ATP synthase subunit beta OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=atpD PE=3 SV=1 - [ATPB_ECOBW]	154.1785	62.39	20	20	39
C4ZRJ3	Protein translocase subunit SecA OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=secA PE=3 SV=1 - [SECA_ECOBW]	146.0395	35.85	23	23	41
C4ZQ19	Serine--tRNA ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=serS PE=3 SV=1 - [SYS_ECOBW]	138.6499	58.84	20	20	40
P0A9Q7	Aldehyde-alcohol dehydrogenase OS=Escherichia coli (strain K12) GN=adhE PE=1 SV=2 - [ADHE_ECOLI]	138.6237	41.64	28	28	39
P0AEI1	tRNA-2-methylthio-N(6)-dimethylallyladenosine synthase OS=Escherichia coli (strain K12) GN=miaB PE=1 SV=1 - [MIAB_ECOLI]	132.4479	56.96	23	23	36
P0A6A3	Acetate kinase OS=Escherichia coli (strain K12) GN=ackA PE=1 SV=1 - [ACKA_ECOLI]	128.2275	50.75	14	14	33
P21507	ATP-dependent RNA helicase SrmB OS=Escherichia coli (strain K12) GN=srmB PE=1 SV=1 - [SRMB_ECOLI]	124.6454	49.77	19	19	35
P0ABD8	Biotin carboxyl carrier protein of acetyl-CoA carboxylase OS=Escherichia coli (strain K12) GN=accB PE=1 SV=1 - [BCCP_ECOLI]	122.3235	89.1	7	7	35
P00561	Bifunctional aspartokinase/homoserine dehydrogenase 1 OS=Escherichia coli (strain K12) GN=thrA PE=1 SV=2 - [AK1H_ECOLI]	117.9488	41.34	24	24	33
B1XF11	tRNA sulfurtransferase OS=Escherichia coli (strain K12 / DH10B) GN=thiI PE=3 SV=1 - [THII_ECODH]	110.9351	50.62	20	20	32
P08201	Nitrite reductase (NADH) large subunit OS=Escherichia coli (strain K12) GN=nirB PE=3 SV=4 - [NIRB_ECOLI]	105.4942	28.45	19	19	27
P0AFG8	Pyruvate dehydrogenase E1 component OS=Escherichia coli (strain K12) GN=aceE PE=1 SV=2 - [ODP1_ECOLI]	101.0202	24.92	19	19	30

P0A9X4	Rod shape-determining protein MreB OS=Escherichia coli (strain K12) GN=mreB PE=1 SV=1 - [MREB_ECOLI]	99.87098	48.41	14	14	25
P28903	Anaerobic ribonucleoside-triphosphate reductase OS=Escherichia coli (strain K12) GN=nrdD PE=1 SV=2 - [NRDD_ECOLI]	99.19083	32.16	18	18	30
P21170	Biosynthetic arginine decarboxylase OS=Escherichia coli (strain K12) GN=speA PE=1 SV=2 - [SPEA_ECOLI]	99.02528	26.44	16	16	27
C4ZRP7	Glutamate-1-semialdehyde 2,1-aminomutase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=hemL PE=3 SV=1 - [GSA_ECOBW]	95.25827	42.96	12	12	28
B1X7X6	Ribosomal protein S12 methylthiotransferase RimO OS=Escherichia coli (strain K12 / DH10B) GN=rimO PE=3 SV=1 - [RIMO_ECODH]	93.24936	38.32	12	12	29
B1XD38	30S ribosomal protein S2 OS=Escherichia coli (strain K12 / DH10B) GN=rpsB PE=3 SV=1 - [RS2_ECODH]	93.18001	74.69	15	15	28
P00562	Bifunctional aspartokinase/homoserine dehydrogenase 2 OS=Escherichia coli (strain K12) GN=metL PE=1 SV=3 - [AK2H_ECOLI]	92.44514	33.21	26	26	30
C4ZX90	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (flavodoxin) OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=ispG PE=3 SV=1 - [ISPG_ECOBW]	90.51271	55.11	19	19	27
P07913	L-threonine 3-dehydrogenase OS=Escherichia coli (strain K12) GN=tdh PE=1 SV=1 - [TDH_ECOLI]	89.18739	39.88	10	10	23
P76658	Bifunctional protein HldE OS=Escherichia coli (strain K12) GN=hldE PE=1 SV=1 - [HLDE_ECOLI]	87.9995	51.78	18	18	25
P33602	NADH-quinone oxidoreductase subunit G OS=Escherichia coli (strain K12) GN=nuoG PE=1 SV=4 - [NUOG_ECOLI]	87.06376	20.26	12	12	20
C4ZPY3	RNA polymerase-associated protein RapA OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=rapA PE=3 SV=1 - [RAPA_ECOBW]	86.54792	28.2	23	23	29
P22523	Chromosome partition protein MukB OS=Escherichia coli (strain K12) GN=mukB PE=1 SV=2 - [MUKB_ECOLI]	84.09475	21.47	23	23	25

B1XAY6	GTPase Der OS=Escherichia coli (strain K12 / DH10B) GN=der PE=3 SV=1 - [DER_ECODH]	82.61288	42.04	15	15	26
B1XFM6	ATP-dependent Clp protease ATP-binding subunit ClpX OS=Escherichia coli (strain K12 / DH10B) GN=clpX PE=3 SV=1 - [CLPX_ECODH]	77.18264	55.42	16	16	21
P00961	Glycine--tRNA ligase beta subunit OS=Escherichia coli (strain K12) GN=glyS PE=1 SV=4 - [SYGB_ECOLI]	76.57484	39.19	19	19	21
B1XBY6	50S ribosomal protein L1 OS=Escherichia coli (strain K12 / DH10B) GN=rplA PE=3 SV=1 - [RL1_ECODH]	73.76688	47.44	11	11	23
P0A9A6	Cell division protein FtsZ OS=Escherichia coli (strain K12) GN=ftsZ PE=1 SV=1 - [FTSZ_ECOLI]	73.53234	56.92	12	12	20
C5A098	ATP-dependent protease ATPase subunit HslU OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=hslU PE=3 SV=1 - [HSLU_ECOBW]	71.3951	40.63	16	16	23
C4ZXA5	Cysteine desulfurase IscS OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=iscS PE=3 SV=1 - [ISCS_ECOBW]	70.37174	36.63	12	12	22
P11349	Respiratory nitrate reductase 1 beta chain OS=Escherichia coli (strain K12) GN=narH PE=1 SV=3 - [NARH_ECOLI]	70.23798	32.23	12	12	20
P63389	Uncharacterized ABC transporter ATP-binding protein YheS OS=Escherichia coli (strain K12) GN=yheS PE=1 SV=1 - [YHES_ECOLI]	70.12244	29.2	16	16	21
P00960	Glycine--tRNA ligase alpha subunit OS=Escherichia coli (strain K12) GN=glyQ PE=1 SV=2 - [SYGA_ECOLI]	69.19585	38.94	7	7	21
C4ZWB5	Lipoyl synthase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=lipA PE=3 SV=1 - [LIPA_ECOBW]	67.85336	57.32	13	13	19
C5A093	Glycerol kinase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=glpK PE=3 SV=1 - [GLPK_ECOBW]	66.59585	35.86	15	15	20
P16456	Selenide, water dikinase OS=Escherichia coli (strain K12) GN=selD PE=1 SV=1 - [SELD_ECOLI]	66.04207	40.06	12	12	19

P0AF20	N-acetylglucosamine repressor OS=Escherichia coli (strain K12) GN=nagC PE=2 SV=1 - [NAGC_ECOLI]	64.12278	43.35	15	15	20
P0A749	UDP-N-acetylglucosamine 1-carboxyvinyltransferase OS=Escherichia coli (strain K12) GN=murA PE=1 SV=1 - [MURA_ECOLI]	63.61467	38.9	13	13	20
P23538	Phosphoenolpyruvate synthase OS=Escherichia coli (strain K12) GN=ppsA PE=1 SV=5 - [PPSA_ECOLI]	62.87668	20.96	13	13	18
P52647	Probable pyruvate-flavodoxin oxidoreductase OS=Escherichia coli (strain K12) GN=ydbK PE=3 SV=2 - [NIFJ_ECOLI]	62.74173	20.1	14	14	18
P00579	RNA polymerase sigma factor RpoD OS=Escherichia coli (strain K12) GN=rpoD PE=1 SV=2 - [RPOD_ECOLI]	62.37409	22.84	13	13	20
B1X8Z8	NADH-quinone oxidoreductase subunit C/D OS=Escherichia coli (strain K12 / DH10B) GN=nuoC PE=3 SV=1 - [NUOCD_ECODH]	62.29811	25.17	14	14	19
B1XCM5	Alanine--tRNA ligase OS=Escherichia coli (strain K12 / DH10B) GN=alaS PE=3 SV=1 - [SYA_ECODH]	61.65482	24.2	15	15	21
B1XB43	Elongation factor 4 OS=Escherichia coli (strain K12 / DH10B) GN=lepA PE=3 SV=1 - [LEPA_ECODH]	61.28494	26.04	14	14	19
P37177	Phosphoenolpyruvate-dependent phosphotransferase system OS=Escherichia coli (strain K12) GN=ptsP PE=1 SV=2 - [PT1P_ECOLI]	61.09615	29.01	15	15	17
C4ZZT2	Enolase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=eno PE=3 SV=1 - [ENO_ECOBW]	60.69667	37.5	13	13	15
B1XDJ0	CTP synthase OS=Escherichia coli (strain K12 / DH10B) GN=pyrG PE=3 SV=1 - [PYRG_ECODH]	58.51811	28.44	16	16	22
B1XAZ2	Dual-specificity RNA methyltransferase RlmN OS=Escherichia coli (strain K12 / DH10B) GN=rlmN PE=3 SV=1 - [RLMN_ECODH]	57.92763	31.77	12	12	19
P08839	Phosphoenolpyruvate-protein phosphotransferase OS=Escherichia coli (strain K12) GN=ptsI PE=1 SV=1 - [PT1_ECOLI]	57.34058	32	12	12	16

P0A7W1	30S ribosomal protein S5 OS=Escherichia coli (strain K12) GN=rpsE PE=1 SV=2 - [RS5_ECOLI]	55.79135	31.14	4	4	16
P0ABH0	Cell division protein FtsA OS=Escherichia coli (strain K12) GN=ftsA PE=1 SV=1 - [FTSA_ECOLI]	55.70038	42.38	9	9	12
C4ZYU4	Protein RecA OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=recA PE=3 SV=1 - [RECA_ECOBW]	54.59214	29.75	7	7	14
C4ZY45	Hydroxylamine reductase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=hcp PE=3 SV=1 - [HCP_ECOBW]	52.9942	22.36	9	9	13
P76536	Probable deferrochelatase/peroxidase YfeX OS=Escherichia coli (strain K12) GN=yfeX PE=1 SV=2 - [YFEX_ECOLI]	52.89946	35.79	6	6	15
P0AFK0	Metalloprotease PmbA OS=Escherichia coli (strain K12) GN=pmbA PE=1 SV=1 - [PMBA_ECOLI]	52.19734	29.56	11	11	16
P0A9W3	Energy-dependent translational throttle protein EttA OS=Escherichia coli (strain K12) GN=ettA PE=1 SV=2 - [ETTA_ECOLI]	52.0608	35.14	13	13	17
C4ZXC6	Serine hydroxymethyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=glyA PE=3 SV=1 - [GLYA_ECOBW]	51.57606	28.06	11	11	16
C4ZQF0	Aspartate--tRNA ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=aspS PE=3 SV=1 - [SYD_ECOBW]	51.26872	28.31	11	11	14
P0A9C0	Anaerobic glycerol-3-phosphate dehydrogenase subunit A OS=Escherichia coli (strain K12) GN=glpA PE=1 SV=1 - [GLPA_ECOLI]	50.24137	28.97	11	11	13
C4ZX89	Histidine--tRNA ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=hisS PE=3 SV=1 - [SYH_ECOBW]	49.51572	37.5	12	12	13
B1X9Z4	ATP-dependent RNA helicase RhlB OS=Escherichia coli (strain K12 / DH10B) GN=rhlB PE=3 SV=1 - [RHLB_ECODH]	49.41408	31.12	10	10	14
B1X736	3-dehydroquinate synthase OS=Escherichia coli (strain K12 / DH10B) GN=aroB PE=3 SV=1 - [AROB_ECODH]	47.98023	42.27	9	9	12

P14081	Selenocysteine-specific elongation factor OS=Escherichia coli (strain K12) GN=selB PE=1 SV=3 - [SELB_ECOLI]	46.87258	25.9	11	11	13
C4ZSI0	D-tagatose-1,6-bisphosphate aldolase subunit GatY OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=gatY PE=3 SV=1 - [GATY_ECOBW]	46.85755	34.86	10	10	15
P77596	Uncharacterized protein YagF OS=Escherichia coli (strain K12) GN=yagF PE=3 SV=1 - [YAGF_ECOLI]	46.05291	24.27	9	9	14
P75864	Ribosomal RNA large subunit methyltransferase K/L OS=Escherichia coli (strain K12) GN=rlmL PE=1 SV=1 - [RLMKL_ECOLI]	45.32896	21.94	12	12	13
P00934	Threonine synthase OS=Escherichia coli (strain K12) GN=thrC PE=1 SV=1 - [THRC_ECOLI]	44.08678	31.54	9	9	11
C4ZRS7	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=accA PE=3 SV=1 - [ACCA_ECOBW]	43.75425	44.2	11	11	13
P60723	50S ribosomal protein L4 OS=Escherichia coli (strain K12) GN=rplD PE=1 SV=1 - [RL4_ECOLI]	42.60365	39.3	8	8	13
P0ACE0	Hydrogenase-2 large chain OS=Escherichia coli (strain K12) GN=hybC PE=1 SV=2 - [MBHM_ECOLI]	41.96234	22.57	9	9	11
P0AAB6	UTP--glucose-1-phosphate uridylyltransferase OS=Escherichia coli (strain K12) GN=galF PE=1 SV=1 - [GALF_ECOLI]	41.92544	35.35	8	8	13
P0AGG8	Metalloprotease TldD OS=Escherichia coli (strain K12) GN=tldD PE=1 SV=1 - [TLDD_ECOLI]	40.8053	21.62	7	7	10
P00954	Tryptophan--tRNA ligase OS=Escherichia coli (strain K12) GN=trpS PE=1 SV=3 - [SYW_ECOLI]	40.64489	33.83	8	8	12
P0A6Y8	Chaperone protein DnaK OS=Escherichia coli (strain K12) GN=dnaK PE=1 SV=2 - [DNAK_ECOLI]	40.51747	21.63	10	10	10
C4ZTG3	Queuine tRNA-ribosyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=tgt PE=3 SV=1 - [TGT_ECOBW]	40.40235	29.33	9	9	13

B1X9T7	Tryptophanase OS=Escherichia coli (strain K12 / DH10B) GN=tnaA PE=3 SV=1 - [TNAA_ECODH]	40.15368	31	11	11	13
P00452	Ribonucleoside-diphosphate reductase 1 subunit alpha OS=Escherichia coli (strain K12) GN=nrdA PE=1 SV=2 - [RIR1_ECOLI]	39.83684	19.32	13	13	16
C4ZRI6	UDP-N-acetylmuramate--L-alanine ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=murC PE=3 SV=1 - [MURC_ECOBW]	39.80019	28.31	8	8	10
P0AG24	Bifunctional (p)ppGpp synthase/hydrolase SpoT OS=Escherichia coli (strain K12) GN=spoT PE=1 SV=1 - [SPOT_ECOLI]	39.38312	15.1	10	10	13
P0AFI2	DNA topoisomerase 4 subunit A OS=Escherichia coli (strain K12) GN=parC PE=1 SV=1 - [PARC_ECOLI]	39.30757	14.36	8	8	13
C4ZZN7	DNA mismatch repair protein MutS OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=mutS PE=3 SV=1 - [MUTS_ECOBW]	39.21143	15.94	11	11	13
P0AGD7	Signal recognition particle protein OS=Escherichia coli (strain K12) GN=ffh PE=1 SV=1 - [SRP54_ECOLI]	38.98482	22.74	10	10	13
C4ZY85	Adenosine deaminase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=add PE=3 SV=1 - [ADD_ECOBW]	38.66576	45.95	11	11	13
P0ABU2	Ribosome-binding ATPase YchF OS=Escherichia coli (strain K12) GN=ychF PE=1 SV=2 - [YCHF_ECOLI]	38.47233	28.65	8	8	11
P67910	ADP-L-glycero-D-manno-heptose-6-epimerase OS=Escherichia coli (strain K12) GN=hldD PE=1 SV=1 - [HLDD_ECOLI]	37.32431	31.61	8	8	11
P0A7Z4	DNA-directed RNA polymerase subunit alpha OS=Escherichia coli (strain K12) GN=rpoA PE=1 SV=1 - [RPOA_ECOLI]	36.91777	29.79	9	9	12
P00363	Fumarate reductase flavoprotein subunit OS=Escherichia coli (strain K12) GN=frdA PE=1 SV=3 - [FRDA_ECOLI]	36.35264	20.27	8	8	11
P77434	Glutamate-pyruvate aminotransferase AlaC OS=Escherichia coli (strain K12) GN=alaC PE=1 SV=1 - [ALAC_ECOLI]	36.23557	33.74	10	10	12

P0AG20	GTP pyrophosphokinase OS=Escherichia coli (strain K12) GN=relA PE=1 SV=1 - [RELA_ECOLI]	35.80544	22.98	9	9	10
P0A9U3	Uncharacterized ABC transporter ATP-binding protein YbiT OS=Escherichia coli (strain K12) GN=ybiT PE=1 SV=1 - [YBIT_ECOLI]	34.45424	20.38	7	7	10
P13009	Methionine synthase OS=Escherichia coli (strain K12) GN=metH PE=1 SV=5 - [METH_ECOLI]	34.19222	11.49	10	10	10
P23830	CDP-diacylglycerol--serine O-phosphatidyltransferase OS=Escherichia coli (strain K12) GN=pssA PE=1 SV=2 - [PSS_ECOLI]	33.71275	25.28	7	7	7
C4ZPV2	Isoleucine--tRNA ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=ileS PE=3 SV=1 - [SYI_ECOBW]	33.53104	10.02	9	9	13
P0A9M8	Phosphate acetyltransferase OS=Escherichia coli (strain K12) GN=pta PE=1 SV=2 - [PTA_ECOLI]	33.03977	15.97	8	8	10
P0ADI4	Enterobactin synthase component B OS=Escherichia coli (strain K12) GN=entB PE=1 SV=1 - [ENTB_ECOLI]	32.94193	23.86	6	6	10
P0AES4	DNA gyrase subunit A OS=Escherichia coli (strain K12) GN=gyrA PE=1 SV=1 - [GYRA_ECOLI]	32.90384	11.54	9	9	11
P06710	DNA polymerase III subunit tau OS=Escherichia coli (strain K12) GN=dnaX PE=1 SV=1 - [DPO3X_ECOLI]	32.05144	19.6	9	9	9
P00393	NADH dehydrogenase OS=Escherichia coli (strain K12) GN=ndh PE=1 SV=2 - [DHNA_ECOLI]	31.64116	26.27	7	7	8
P06996	Outer membrane protein C OS=Escherichia coli (strain K12) GN=ompC PE=1 SV=1 - [OMPC_ECOLI]	31.43143	25.89	6	7	9
C4ZTJ3	Trigger factor OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=tig PE=1 SV=1 - [TIG_ECOBW]	30.84747	21.99	10	10	11
P42641	GTPase ObgE/CgtA OS=Escherichia coli (strain K12) GN=obgE PE=1 SV=1 - [OBG_ECOLI]	30.81549	27.18	7	7	9

P21513	Ribonuclease E OS=Escherichia coli (strain K12) GN=rne PE=1 SV=6 - [RNE_ECOLI]	30.47954	11.03	8	8	9
C4ZZ12	ATP synthase subunit alpha OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=atpA PE=3 SV=1 - [ATPA_ECOBW]	30.27282	20.08	8	8	9
P0A959	Glutamate-pyruvate aminotransferase AlaA OS=Escherichia coli (strain K12) GN=alaA PE=1 SV=1 - [ALAA_ECOLI]	30.16953	25.68	7	7	10
C4ZYI9	GTPase Era OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=era PE=3 SV=1 - [ERA_ECOBW]	30.13866	20.27	4	4	9
P0AEP3	UTP--glucose-1-phosphate uridylyltransferase OS=Escherichia coli (strain K12) GN=galU PE=1 SV=2 - [GALU_ECOLI]	29.07194	28.81	5	5	7
P09980	ATP-dependent DNA helicase Rep OS=Escherichia coli (strain K12) GN=rep PE=1 SV=3 - [REP_ECOLI]	29.06052	17.53	8	8	9
C4ZUF1	30S ribosomal protein S4 OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=rpsD PE=3 SV=1 - [RS4_ECOBW]	28.69427	33.98	6	6	8
P37744	Glucose-1-phosphate thymidylyltransferase 1 OS=Escherichia coli (strain K12) GN=rfbA PE=1 SV=2 - [RMLA1_ECOLI]	27.82324	26.28	7	7	12
P02359	30S ribosomal protein S7 OS=Escherichia coli (strain K12) GN=rpsG PE=1 SV=3 - [RS7_ECOLI]	27.36906	37.99	6	6	7
P0A9B2	Glyceraldehyde-3-phosphate dehydrogenase A OS=Escherichia coli (strain K12) GN=gapA PE=1 SV=2 - [G3P1_ECOLI]	27.1737	31.72	6	6	8
B1X6G0	50S ribosomal protein L5 OS=Escherichia coli (strain K12 / DH10B) GN=rplE PE=3 SV=1 - [RL5_ECODH]	26.89174	48.04	6	6	8
P04825	Aminopeptidase N OS=Escherichia coli (strain K12) GN=pepN PE=1 SV=2 - [AMPN_ECOLI]	26.32168	9.66	7	7	8
P0ADF6	Phosphogluconate dehydratase OS=Escherichia coli (strain K12) GN=edd PE=2 SV=1 - [EDD_ECOLI]	26.32101	17.41	6	6	8

P0A799	Phosphoglycerate kinase OS=Escherichia coli (strain K12) GN=pgk PE=1 SV=2 - [PGK_ECOLI]	26.31899	27.65	6	6	7
C4ZSQ5	Polyribonucleotide nucleotidyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=pnp PE=3 SV=1 - [PNP_ECOBW]	26.21635	12.52	6	6	7
P25888	ATP-dependent RNA helicase RhlE OS=Escherichia coli (strain K12) GN=rhlE PE=1 SV=3 - [RHLE_ECOLI]	25.65669	19.6	6	6	7
P14175	Glycine betaine/proline betaine transport system ATP-binding protein ProV OS=Escherichia coli (strain K12) GN=proV PE=1 SV=1 - [PROV_ECOLI]	25.3899	26.25	8	8	10
P0AD61	Pyruvate kinase I OS=Escherichia coli (strain K12) GN=pykF PE=1 SV=1 - [KPYK1_ECOLI]	24.857	19.57	5	5	6
P17169	Glutamine--fructose-6-phosphate aminotransferase [isomerizing] OS=Escherichia coli (strain K12) GN=glmS PE=1 SV=4 - [GLMS_ECOLI]	24.75468	13.96	5	5	6
P13033	Anaerobic glycerol-3-phosphate dehydrogenase subunit B OS=Escherichia coli (strain K12) GN=glpB PE=1 SV=1 - [GLPB_ECOLI]	23.26812	14.32	3	3	5
P0AB91	Phospho-2-dehydro-3-deoxyheptonate aldolase, Phe-sensitive OS=Escherichia coli (strain K12) GN=aroG PE=1 SV=1 - [AROG_ECOLI]	22.46141	23.71	5	5	6
C4ZQJ1	UvrABC system protein C OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=uvrC PE=3 SV=1 - [UVRC_ECOBW]	21.91528	10.98	5	5	5
C4ZR80	50S ribosomal protein L9 OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=rplI PE=3 SV=1 - [RL9_ECOBW]	21.69076	45.64	6	6	8
P0ABH9	ATP-dependent Clp protease ATP-binding subunit ClpA OS=Escherichia coli (strain K12) GN=clpA PE=1 SV=1 - [CLPA_ECOLI]	21.60538	9.37	5	5	6
P03004	Chromosomal replication initiator protein DnaA OS=Escherichia coli (strain K12) GN=dnaA PE=1 SV=2 - [DNAA_ECOLI]	21.00631	14.56	5	5	6
B1X6F3	50S ribosomal protein L15 OS=Escherichia coli (strain K12 / DH10B) GN=rplO PE=3 SV=1 - [RL15_ECODH]	20.88318	43.06	5	5	9

P0A879	Tryptophan synthase beta chain OS=Escherichia coli (strain K12) GN=trpB PE=1 SV=2 - [TRPB_ECOLI]	20.55681	14.61	4	4	7
P21829	Pyridoxal phosphate phosphatase YbhA OS=Escherichia coli (strain K12) GN=ybhA PE=1 SV=3 - [YBHA_ECOLI]	20.41705	20.22	6	6	7
P0A8N5	Lysine--tRNA ligase, heat inducible OS=Escherichia coli (strain K12) GN=lysU PE=1 SV=2 - [SYK2_ECOLI]	20.32496	19.21	4	7	8
P37759	dTDP-glucose 4,6-dehydratase 1 OS=Escherichia coli (strain K12) GN=rfbB PE=3 SV=2 - [RMLB1_ECOLI]	20.29947	30.47	7	7	7
B1XDY5	Glutamate 5-kinase OS=Escherichia coli (strain K12 / DH10B) GN=proB PE=3 SV=1 - [PROB_ECODH]	20.2574	22.89	5	5	5
P06987	Histidine biosynthesis bifunctional protein HisB OS=Escherichia coli (strain K12) GN=hisB PE=1 SV=1 - [HIS7_ECOLI]	20.21633	20	6	6	7
P06993	HTH-type transcriptional regulator MalT OS=Escherichia coli (strain K12) GN=malT PE=1 SV=2 - [MALT_ECOLI]	20.2121	9.54	6	6	7
P07395	Phenylalanine--tRNA ligase beta subunit OS=Escherichia coli (strain K12) GN=pheT PE=1 SV=2 - [SYFB_ECOLI]	20.08888	15.6	7	7	7
B1XF08	1-deoxy-D-xylulose-5-phosphate synthase OS=Escherichia coli (strain K12 / DH10B) GN=dxs PE=3 SV=1 - [DXS_ECODH]	19.96712	12.42	6	6	7
P46889	DNA translocase FtsK OS=Escherichia coli (strain K12) GN=ftsK PE=1 SV=2 - [FTSK_ECOLI]	19.71168	6.17	8	8	8
P0ABF1	Poly(A) polymerase I OS=Escherichia coli (strain K12) GN=pcnB PE=1 SV=2 - [PCNB_ECOLI]	19.33031	12.26	6	6	7
P30178	Hydroxycarboxylate dehydrogenase B OS=Escherichia coli (strain K12) GN=hcxB PE=1 SV=2 - [HCXB_ECOLI]	19.15465	12.74	2	2	4
C4ZYG8	UPF0061 protein YdiU OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=ydiU PE=3 SV=1 - [YDIU_ECOBW]	19.1029	18.2	5	5	5

P0ABQ0	Coenzyme A biosynthesis bifunctional protein CoaBC OS=Escherichia coli (strain K12) GN=coaBC PE=1 SV=2 - [COABC_ECOLI]	18.89768	16.26	4	4	5
P08200	Isocitrate dehydrogenase [NADP] OS=Escherichia coli (strain K12) GN=icd PE=1 SV=1 - [IDH_ECOLI]	18.4397	12.74	4	4	5
P21599	Pyruvate kinase II OS=Escherichia coli (strain K12) GN=pykA PE=1 SV=3 - [KPYK2_ECOLI]	18.36489	16.04	5	5	5
P0A7B1	Polyphosphate kinase OS=Escherichia coli (strain K12) GN=ppk PE=1 SV=2 - [PPK_ECOLI]	18.22731	11.19	6	6	6
P37747	UDP-galactopyranose mutase OS=Escherichia coli (strain K12) GN=glf PE=1 SV=1 - [GLF_ECOLI]	18.08334	14.17	5	5	6
B1X9R9	Glutamate--tRNA ligase OS=Escherichia coli (strain K12 / DH10B) GN=gltX PE=3 SV=1 - [SYE_ECODH]	18.07866	15.29	5	5	5
P0A9S3	Galactitol 1-phosphate 5-dehydrogenase OS=Escherichia coli (strain K12) GN=gatD PE=1 SV=1 - [GATD_ECOLI]	18.06141	18.21	5	5	8
P76027	Oligopeptide transport ATP-binding protein OppD OS=Escherichia coli (strain K12) GN=oppD PE=3 SV=2 - [OPPD_ECOLI]	18.0212	10.09	3	3	5
C5A059	Protein FdhE OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=fdhE PE=3 SV=1 - [FDHE_ECOBW]	17.2186	21.68	5	5	5
P0A9C9	Fructose-1,6-bisphosphatase 1 class 2 OS=Escherichia coli (strain K12) GN=glpX PE=1 SV=1 - [GLPX_ECOLI]	17.18332	18.75	4	4	5
P0A8N3	Lysine--tRNA ligase OS=Escherichia coli (strain K12) GN=lysS PE=1 SV=2 - [SYK1_ECOLI]	17.01225	15.25	4	7	8
P0A698	UvrABC system protein A OS=Escherichia coli (strain K12) GN=uvrA PE=1 SV=1 - [UVRA_ECOLI]	16.98924	6.49	5	5	6
P08312	Phenylalanine--tRNA ligase alpha subunit OS=Escherichia coli (strain K12) GN=pheS PE=1 SV=2 - [SYFA_ECOLI]	16.91329	20.49	4	4	4

P0AEZ3	Septum site-determining protein MinD OS=Escherichia coli (strain K12) GN=minD PE=1 SV=2 - [MIND_ECOLI]	16.74127	24.07	6	6	6
Q46861	UPF0313 protein YgiQ OS=Escherichia coli (strain K12) GN=ygiQ PE=3 SV=3 - [YGIQ_ECOLI]	16.48777	10.01	6	6	7
B1X9T5	tRNA modification GTPase MnmE OS=Escherichia coli (strain K12 / DH10B) GN=mnmE PE=3 SV=1 - [MNME_ECODH]	16.4845	14.54	4	5	5
P0AEW9	1-phosphofructokinase OS=Escherichia coli (strain K12) GN=fruK PE=3 SV=1 - [K1PF_ECOLI]	16.42109	15.38	4	4	4
P0A993	Fructose-1,6-bisphosphatase class 1 OS=Escherichia coli (strain K12) GN=fbp PE=1 SV=1 - [F16PA_ECOLI]	16.38321	19.88	5	5	6
P43329	ATP-dependent RNA helicase HrpA OS=Escherichia coli (strain K12) GN=hrpA PE=3 SV=3 - [HRPA_ECOLI]	16.22785	5.77	5	5	5
P76104	Uncharacterized protease YdcP OS=Escherichia coli (strain K12) GN=ydcP PE=1 SV=2 - [YDCP_ECOLI]	16.17149	11.33	5	5	5
P0A9H7	Cyclopropane-fatty-acyl-phospholipid synthase OS=Escherichia coli (strain K12) GN=cfa PE=1 SV=2 - [CFA_ECOLI]	16.16218	13.87	4	4	5
B1XEU8	S-formylglutathione hydrolase FrmB OS=Escherichia coli (strain K12 / DH10B) GN=frmB PE=3 SV=1 - [SFGH1_ECODH]	16.12867	19.49	3	3	4
P27833	dTDP-4-amino-4,6-dideoxygalactose transaminase OS=Escherichia coli (strain K12) GN=wecE PE=1 SV=2 - [WECE_ECOLI]	16.07518	24.73	5	5	5
C4ZSA8	ATP phosphoribosyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=hisG PE=3 SV=1 - [HIS1_ECOBW]	16.02656	17.73	3	3	4
P75829	Uncharacterized protein YbjX OS=Escherichia coli (strain K12) GN=ybjX PE=4 SV=1 - [YBJX_ECOLI]	16.01727	13.64	5	5	5
C5A1D5	60 kDa chaperonin OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=groL PE=1 SV=1 - [CH60_ECOBW]	15.66032	11.13	4	4	4

P0A9N4	Pyruvate formate-lyase 1-activating enzyme OS=Escherichia coli (strain K12) GN=pflA PE=1 SV=2 - [PFLA_ECOLI]	15.05887	23.98	5	5	5
P37349	Protein-lysine deacetylase OS=Escherichia coli (strain K12) GN=dhaM PE=1 SV=3 - [DHAM_ECOLI]	15.04598	11.02	4	4	4
P19319	Respiratory nitrate reductase 2 alpha chain OS=Escherichia coli (strain K12) GN=narZ PE=1 SV=5 - [NARZ_ECOLI]	14.47923	2.97	1	4	6
P24203	Uncharacterized GTP-binding protein YjiA OS=Escherichia coli (strain K12) GN=yjiA PE=1 SV=3 - [YJIA_ECOLI]	14.25901	9.12	2	2	4
P0ADG4	Inositol-1-monophosphatase OS=Escherichia coli (strain K12) GN=suhB PE=1 SV=1 - [SUHB_ECOLI]	14.22729	24.72	4	4	4
P45527	Uncharacterized protease YhbU OS=Escherichia coli (strain K12) GN=yhbU PE=1 SV=1 - [YHBU_ECOLI]	14.21379	16.62	3	3	3
B1XDS8	Adenylosuccinate synthetase OS=Escherichia coli (strain K12 / DH10B) GN=purA PE=3 SV=1 - [PURA_ECODH]	14.04506	18.06	7	7	7
P0ADY7	50S ribosomal protein L16 OS=Escherichia coli (strain K12) GN=rplP PE=1 SV=1 - [RL16_ECOLI]	14.0317	10.29	1	1	5
P0A7X3	30S ribosomal protein S9 OS=Escherichia coli (strain K12) GN=rpsI PE=1 SV=2 - [RS9_ECOLI]	13.91612	20	3	3	6
P0A9B6	D-erythrose-4-phosphate dehydrogenase OS=Escherichia coli (strain K12) GN=epd PE=1 SV=2 - [E4PD_ECOLI]	13.78434	17.11	4	4	4
P37760	dTDP-4-dehydrorhamnose reductase OS=Escherichia coli (strain K12) GN=rfbD PE=3 SV=2 - [RMLD_ECOLI]	13.77983	20.4	3	3	3
P08957	Type I restriction enzyme EcoKI M protein OS=Escherichia coli (strain K12) GN=hsdM PE=1 SV=1 - [T1MK_ECOLI]	13.66226	10.02	4	4	4
B1XB82	ATP-dependent 6-phosphofructokinase isozyme 1 OS=Escherichia coli (strain K12 / DH10B) GN=pfkA PE=3 SV=1 - [PFKA_ECODH]	13.62015	10.63	3	3	5

P04079	GMP synthase [glutamine-hydrolyzing] OS=Escherichia coli (strain K12) GN=guaA PE=1 SV=1 - [GUAA_ECOLI]	13.59698	8.57	3	3	4
P0AB77	2-amino-3-ketobutyrate coenzyme A ligase OS=Escherichia coli (strain K12) GN=kbl PE=1 SV=1 - [KBL_ECOLI]	13.25349	11.31	3	3	4
C4ZUG5	50S ribosomal protein L14 OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=rplN PE=3 SV=1 - [RL14_ECOBW]	13.19013	26.83	3	3	4
B1XA82	DNA ligase OS=Escherichia coli (strain K12 / DH10B) GN=ligA PE=3 SV=1 - [DNLJ_ECODH]	13.05137	5.66	2	2	3
P25539	Riboflavin biosynthesis protein RibD OS=Escherichia coli (strain K12) GN=ribD PE=1 SV=1 - [RIBD_ECOLI]	12.99905	9.81	3	3	4
P07813	Leucine--tRNA ligase OS=Escherichia coli (strain K12) GN=leuS PE=1 SV=2 - [SYL_ECOLI]	12.95352	7.56	5	5	5
B1XG88	Ribosomal RNA large subunit methyltransferase G OS=Escherichia coli (strain K12 / DH10B) GN=rlmG PE=3 SV=1 - [RLMG_ECODH]	12.80645	11.38	3	3	4
P0A7S9	30S ribosomal protein S13 OS=Escherichia coli (strain K12) GN=rpsM PE=1 SV=2 - [RS13_ECOLI]	12.78905	38.14	3	3	3
P0ACP1	Catabolite repressor/activator OS=Escherichia coli (strain K12) GN=cra PE=1 SV=1 - [CRA_ECOLI]	12.74687	10.78	3	3	4
B1X9G8	UPF0176 protein YceA OS=Escherichia coli (strain K12 / DH10B) GN=yceA PE=3 SV=1 - [YCEA_ECODH]	12.53626	12.86	3	3	3
P06612	DNA topoisomerase 1 OS=Escherichia coli (strain K12) GN=topA PE=1 SV=2 - [TOP1_ECOLI]	12.37564	7.75	3	3	3
B1X8M6	Chromosome partition protein MukF OS=Escherichia coli (strain K12 / DH10B) GN=mukF PE=3 SV=1 - [MUKF_ECODH]	12.34992	10.23	3	3	3
P30958	Transcription-repair-coupling factor OS=Escherichia coli (strain K12) GN=mfd PE=1 SV=2 - [MFD_ECOLI]	12.34287	4.18	3	3	3

P24183	Formate dehydrogenase, nitrate-inducible, major subunit OS=Escherichia coli (strain K12) GN=fdnG PE=1 SV=3 - [FDNG_ECOLI]	12.24714	7.68	4	5	5
P06959	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex OS=Escherichia coli (strain K12) GN=aceF PE=1 SV=3 - [ODP2_ECOLI]	12.10945	7.3	3	3	4
P0ACB0	Replicative DNA helicase OS=Escherichia coli (strain K12) GN=dnaB PE=1 SV=1 - [DNAB_ECOLI]	12.05601	11.25	4	4	4
P03018	DNA helicase II OS=Escherichia coli (strain K12) GN=uvrD PE=1 SV=1 - [UVRD_ECOLI]	11.97727	5.97	3	3	4
P67087	Ribosomal RNA small subunit methyltransferase I OS=Escherichia coli (strain K12) GN=rsmI PE=1 SV=1 - [RSMI_ECOLI]	11.85595	21.68	4	4	4
P25534	2-octaprenyl-6-methoxyphenol hydroxylase OS=Escherichia coli (strain K12) GN=ubiH PE=1 SV=2 - [UBIH_ECOLI]	11.82563	11.99	3	3	3
P0A887	Ubiquinone/menaquinone biosynthesis C-methyltransferase UbiE OS=Escherichia coli (strain K12) GN=ubiE PE=1 SV=1 - [UBIE_ECOLI]	11.80817	17.93	3	3	3
P0AAZ4	Replication-associated recombination protein A OS=Escherichia coli (strain K12) GN=rarA PE=1 SV=1 - [RARA_ECOLI]	11.7682	10.51	4	4	5
P0AB89	Adenylosuccinate lyase OS=Escherichia coli (strain K12) GN=purB PE=1 SV=1 - [PUR8_ECOLI]	11.72042	10.75	4	4	4
C4ZZ73	Diaminopimelate epimerase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=dapF PE=3 SV=1 - [DAPF_ECOBW]	11.59955	18.25	3	3	3
P62623	4-hydroxy-3-methylbut-2-enyl diphosphate reductase OS=Escherichia coli (strain K12) GN=ispH PE=1 SV=1 - [ISPH_ECOLI]	11.40038	18.35	3	3	3
P00350	6-phosphogluconate dehydrogenase, decarboxylating OS=Escherichia coli (strain K12) GN=gnd PE=1 SV=2 - [6PGD_ECOLI]	11.30022	12.61	4	4	4
P52062	Oxygen-independent coproporphyrinogen-III oxidase-like protein YggW OS=Escherichia coli (strain K12) GN=yggW PE=3 SV=1 - [YGGW_ECOLI]	11.02946	10.85	4	4	4

P25437	S-(hydroxymethyl)glutathione dehydrogenase OS=Escherichia coli (strain K12) GN=frmA PE=1 SV=3 - [FRMA_ECOLI]	11.00297	11.65	3	3	3
B1XGY3	Argininosuccinate synthase OS=Escherichia coli (strain K12 / DH10B) GN=argG PE=3 SV=1 - [ASSY_ECODH]	10.851	12.53	3	3	3
P21693	ATP-dependent RNA helicase DbpA OS=Escherichia coli (strain K12) GN=dbpA PE=1 SV=2 - [DBPA_ECOLI]	10.83586	10.94	3	3	3
P08192	Dihydrofolate synthase/folylpolyglutamate synthase OS=Escherichia coli (strain K12) GN=folC PE=1 SV=2 - [FOLC_ECOLI]	10.66058	9.72	2	2	3
P0A910	Outer membrane protein A OS=Escherichia coli (strain K12) GN=ompA PE=1 SV=1 - [OMPA_ECOLI]	10.41462	15.03	3	3	3
P63284	Chaperone protein ClpB OS=Escherichia coli (strain K12) GN=clpB PE=1 SV=1 - [CLPB_ECOLI]	10.19746	5.02	3	3	3
P0A8M0	Asparagine--tRNA ligase OS=Escherichia coli (strain K12) GN=asnS PE=1 SV=2 - [SYN_ECOLI]	10.14215	12.02	3	3	3
P0AC53	Glucose-6-phosphate 1-dehydrogenase OS=Escherichia coli (strain K12) GN=zwf PE=1 SV=1 - [G6PD_ECOLI]	9.950286	7.74	3	3	3
P0A6E9	ATP-dependent dethiobiotin synthetase BioD 2 OS=Escherichia coli (strain K12) GN=bioD2 PE=3 SV=1 - [BIOD2_ECOLI]	9.82686	19.05	3	3	3
B1XFI0	Ribosomal RNA small subunit methyltransferase C OS=Escherichia coli (strain K12 / DH10B) GN=rsmC PE=3 SV=1 - [RSMC_ECODH]	9.723831	9.91	2	2	3
P0A9D4	Serine acetyltransferase OS=Escherichia coli (strain K12) GN=cysE PE=1 SV=1 - [CYSE_ECOLI]	9.606718	21.61	3	3	3
P00959	Methionine--tRNA ligase OS=Escherichia coli (strain K12) GN=metG PE=1 SV=2 - [SYM_ECOLI]	9.586238	9.16	4	4	4
P36683	Aconitate hydratase B OS=Escherichia coli (strain K12) GN=acnB PE=1 SV=3 - [ACNB_ECOLI]	9.565858	4.97	4	4	4

P37689	2,3-bisphosphoglycerate-independent phosphoglycerate mutase OS=Escherichia coli (strain K12) GN=gpmI PE=1 SV=1 - [GPMI_ECOLI]	9.400729	7.78	3	3	3
P33136	Glucans biosynthesis protein G OS=Escherichia coli (strain K12) GN=mdoG PE=1 SV=1 - [OPGG_ECOLI]	9.250042	5.68	2	2	3
P27829	UDP-N-acetyl-D-mannosamine dehydrogenase OS=Escherichia coli (strain K12) GN=wecC PE=1 SV=4 - [WECC_ECOLI]	9.032727	9.29	3	3	3
P0A9F3	HTH-type transcriptional regulator CysB OS=Escherichia coli (strain K12) GN=cysB PE=3 SV=1 - [CYSB_ECOLI]	8.958551	12.35	2	2	2
P36938	Phosphoglucomutase OS=Escherichia coli (strain K12) GN=pgm PE=1 SV=1 - [PGM_ECOLI]	8.86815	8.61	3	3	3
P69874	Spermidine/putrescine import ATP-binding protein PotA OS=Escherichia coli (strain K12) GN=potA PE=1 SV=1 - [POTA_ECOLI]	8.860117	5.56	2	2	3
P09546	Bifunctional protein PutA OS=Escherichia coli (strain K12) GN=putA PE=1 SV=3 - [PUTA_ECOLI]	8.515907	5.3	5	5	5
C4ZU48	Periplasmic nitrate reductase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=napA PE=3 SV=1 - [NAPA_ECOBW]	8.404147	3.62	2	2	2
P0AEK4	Enoyl-[acyl-carrier-protein] reductase [NADH] FabI OS=Escherichia coli (strain K12) GN=fabI PE=1 SV=2 - [FABI_ECOLI]	8.329305	9.16	2	2	3
P07118	Valine--tRNA ligase OS=Escherichia coli (strain K12) GN=valS PE=1 SV=2 - [SYV_ECOLI]	8.306784	3.58	3	3	3
P21420	Putative outer membrane porin protein NmpC OS=Escherichia coli (strain K12) GN=nmpC PE=5 SV=2 - [NMPC_ECOLI]	8.22787	6.03	1	2	3
P75745	Uncharacterized protein YbgK OS=Escherichia coli (strain K12) GN=ybgK PE=1 SV=1 - [YBGK_ECOLI]	8.081942	6.45	2	2	3
P08660	Lysine-sensitive aspartokinase 3 OS=Escherichia coli (strain K12) GN=lysC PE=1 SV=2 - [AK3_ECOLI]	7.964732	4.23	1	1	2

C5A0C2	Phosphoenolpyruvate carboxylase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=ppc PE=3 SV=1 - [CAPP_ECOBW]	7.908599	3.62	2	2	2
P43672	ABC transporter ATP-binding protein uup OS=Escherichia coli (strain K12) GN=uup PE=1 SV=2 - [UUP_ECOLI]	7.87323	5.83	2	2	2
P22106	Asparagine synthetase B [glutamine-hydrolyzing] OS=Escherichia coli (strain K12) GN=asnB PE=1 SV=3 - [ASNB_ECOLI]	7.858747	6.5	2	2	2
P77783	Probable dimethyl sulfoxide reductase chain YnfF OS=Escherichia coli (strain K12) GN=ynfF PE=1 SV=4 - [YNFF_ECOLI]	7.779649	4.34	3	3	3
B1X9W9	tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG OS=Escherichia coli (strain K12 / DH10B) GN=mnmG PE=3 SV=1 - [MNMG_ECODH]	7.659205	5.56	3	3	3
C4ZRJ1	UDP-3-O-acyl-N-acetylglucosamine deacetylase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=lpxC PE=3 SV=1 - [LPXC_ECOBW]	7.657356	12.13	3	3	3
P0A717	Ribose-phosphate pyrophosphokinase OS=Escherichia coli (strain K12) GN=prs PE=1 SV=2 - [KPRS_ECOLI]	7.645351	10.16	2	2	2
C4ZS02	Dihydroorotase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=pyrC PE=3 SV=1 - [PYRC_ECOBW]	7.464237	10.06	3	3	3
P10443	DNA polymerase III subunit alpha OS=Escherichia coli (strain K12) GN=dnaE PE=1 SV=1 - [DPO3A_ECOLI]	7.463567	3.19	3	3	3
P0AFF6	Transcription termination/antitermination protein NusA OS=Escherichia coli (strain K12) GN=nusA PE=1 SV=1 - [NUSA_ECOLI]	7.456522	5.86	2	2	2
P24192	Hydrogenase isoenzymes formation protein HypD OS=Escherichia coli (strain K12) GN=hypD PE=1 SV=1 - [HYPD_ECOLI]	7.411872	4.83	1	1	2
P60438	50S ribosomal protein L3 OS=Escherichia coli (strain K12) GN=rplC PE=1 SV=1 - [RL3_ECOLI]	7.382845	15.31	3	3	3
P10121	Signal recognition particle receptor FtsY OS=Escherichia coli (strain K12) GN=ftsY PE=1 SV=1 - [FTSY_ECOLI]	7.378947	7.24	2	2	2

C4ZUG0	50S ribosomal protein L6 OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=rplF PE=3 SV=1 - [RL6_ECOBW]	7.318969	18.08	3	3	3
P0AF24	Ribonucleotide monophosphatase NagD OS=Escherichia coli (strain K12) GN=nagD PE=1 SV=1 - [NAGD_ECOLI]	7.314887	12.8	2	2	2
P75824	NADH oxidoreductase HCR OS=Escherichia coli (strain K12) GN=hcr PE=1 SV=3 - [HCR_ECOLI]	7.295831	11.8	2	2	2
P0A6N8	Elongation factor P-like protein OS=Escherichia coli (strain K12) GN=yeiP PE=1 SV=1 - [EFPL_ECOLI]	7.271196	7.37	1	1	2
P37773	UDP-N-acetylmuramate--L-alanyl-gamma-D-glutamyl-meso-2,6-diaminoheptandioate ligase OS=Escherichia coli (strain K12) GN=mpl PE=1 SV=3 - [MPL_ECOLI]	7.269703	6.13	2	2	2
P77536	Uncharacterized electron transport protein YkgF OS=Escherichia coli (strain K12) GN=ykgF PE=3 SV=1 - [YKGF_ECOLI]	7.252494	6.11	2	2	2
B1XBN7	Exoribonuclease 2 OS=Escherichia coli (strain K12 / DH10B) GN=rnb PE=3 SV=1 - [RNB_ECODH]	7.244056	4.66	2	2	2
P0AB28	Large ribosomal RNA subunit accumulation protein YceD OS=Escherichia coli (strain K12) GN=yceD PE=2 SV=1 - [YCED_ECOLI]	7.240918	6.36	2	2	3
B1XHC8	Holliday junction ATP-dependent DNA helicase RuvB OS=Escherichia coli (strain K12 / DH10B) GN=ruvB PE=3 SV=1 - [RUVB_ECODH]	7.027968	8.04	2	2	2
C4ZWP8	NAD-dependent malic enzyme OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=maeA PE=3 SV=1 - [MAO1_ECOBW]	6.909779	5.49	2	2	2
P0A9Q1	Aerobic respiration control protein ArcA OS=Escherichia coli (strain K12) GN=arcA PE=1 SV=1 - [ARCA_ECOLI]	6.89185	14.71	2	2	2
P0A6R0	3-oxoacyl-[acyl-carrier-protein] synthase 3 OS=Escherichia coli (strain K12) GN=fabH PE=1 SV=1 - [FABH_ECOLI]	6.88405	6.62	1	1	1
P76403	Uncharacterized protease YegQ OS=Escherichia coli (strain K12) GN=yegQ PE=3 SV=1 - [YEGQ_ECOLI]	6.789061	11.04	3	3	3

P39356	Uncharacterized transcriptional regulator YjhU OS=Escherichia coli (strain K12) GN=yjhU PE=3 SV=2 - [YJHU_ECOLI]	6.665281	10.37	3	3	4
P0AGJ9	Tyrosine--tRNA ligase OS=Escherichia coli (strain K12) GN=tyrS PE=1 SV=2 - [SYY_ECOLI]	6.650552	6.37	2	2	2
P05852	tRNA N6-adenosine threonylcarbamoyltransferase OS=Escherichia coli (strain K12) GN=tsaD PE=1 SV=2 - [TSAD_ECOLI]	6.618659	8.01	2	2	3
B1X6E9	30S ribosomal protein S11 OS=Escherichia coli (strain K12 / DH10B) GN=rpsK PE=3 SV=1 - [RS11_ECODH]	6.551288	28.68	2	2	2
C4ZSR5	Phosphoglucosamine mutase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=glmM PE=3 SV=1 - [GLMM_ECOBW]	6.546573	6.29	2	2	2
C4ZUJ7	30S ribosomal protein S12 OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=rpsL PE=3 SV=1 - [RS12_ECOBW]	6.519565	10.48	1	1	2
B1XA17	Beta-hexosaminidase OS=Escherichia coli (strain K12 / DH10B) GN=nagZ PE=3 SV=1 - [NAGZ_ECODH]	6.299098	5.28	1	1	2
B1XC78	Cell division protein ZapD OS=Escherichia coli (strain K12 / DH10B) GN=zapD PE=3 SV=1 - [ZAPD_ECODH]	6.232635	12.55	2	2	2
P32176	Formate dehydrogenase-O major subunit OS=Escherichia coli (strain K12) GN=fdoG PE=1 SV=5 - [FDOG_ECOLI]	6.19939	2.95	1	2	2
P25535	2-octaprenylphenol hydroxylase OS=Escherichia coli (strain K12) GN=ubiI PE=1 SV=2 - [UBII_ECOLI]	6.130609	5.5	2	2	2
B1XAK9	Xaa-Pro dipeptidase OS=Escherichia coli (strain K12 / DH10B) GN=pepQ PE=3 SV=1 - [PEPQ_ECODH]	6.068495	4.97	2	2	2
P0A6K6	Phosphopentomutase OS=Escherichia coli (strain K12) GN=deoB PE=1 SV=1 - [DEOB_ECOLI]	6.058149	6.14	2	2	2
P0AAI3	ATP-dependent zinc metalloprotease FtsH OS=Escherichia coli (strain K12) GN=ftsH PE=1 SV=1 - [FTSH_ECOLI]	5.904557	5.9	2	2	2

P39452	Ribonucleoside-diphosphate reductase 2 subunit alpha OS=Escherichia coli (strain K12) GN=nrdE PE=2 SV=3 - [RIR3_ECOLI]	5.86711	3.92	1	1	1
B1XHI0	RNase adapter protein RapZ OS=Escherichia coli (strain K12 / DH10B) GN=rapZ PE=3 SV=1 - [RAPZ_ECODH]	5.809325	6.69	1	1	1
P00968	Carbamoyl-phosphate synthase large chain OS=Escherichia coli (strain K12) GN=carB PE=1 SV=2 - [CARB_ECOLI]	5.616251	2.33	2	2	2
P0ABS5	DNA primase OS=Escherichia coli (strain K12) GN=dnaG PE=1 SV=1 - [DNAG_ECOLI]	5.612765	4.3	2	2	2
B1XG23	50S ribosomal protein L20 OS=Escherichia coli (strain K12 / DH10B) GN=rplT PE=3 SV=1 - [RL20_ECODH]	5.436881	16.1	2	2	2
P0ADQ2	Uncharacterized protein YiiD OS=Escherichia coli (strain K12) GN=yiiD PE=4 SV=1 - [YIID_ECOLI]	5.434523	7.9	2	2	2
P0C0V0	Periplasmic serine endoprotease DegP OS=Escherichia coli (strain K12) GN=degP PE=1 SV=1 - [DEGP_ECOLI]	5.372044	4.01	2	2	2
P36879	Uncharacterized ABC transporter ATP-binding protein YadG OS=Escherichia coli (strain K12) GN=yadG PE=3 SV=1 - [YADG_ECOLI]	5.343012	8.44	2	2	3
P04993	RecBCD enzyme subunit RecD OS=Escherichia coli (strain K12) GN=recD PE=1 SV=2 - [RECD_ECOLI]	5.340537	2.96	2	2	2
P75838	Ribosomal protein S12 methylthiotransferase accessory factor YcaO OS=Escherichia coli (strain K12) GN=ycaO PE=1 SV=2 - [YCAO_ECOLI]	5.3401	3.24	2	2	2
P0A9G6	Isocitrate lyase OS=Escherichia coli (strain K12) GN=aceA PE=1 SV=1 - [ACEA_ECOLI]	5.317436	5.99	2	2	2
P27242	Lipopolysaccharide 1,2-N-acetylglucosaminetransferase OS=Escherichia coli (strain K12) GN=waaU PE=4 SV=2 - [WAAU_ECOLI]	5.280689	6.72	2	2	2
C4ZRQ9	Bifunctional uridylyltransferase/uridylyl-removing enzyme OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=glnD PE=3 SV=1 - [GLND_ECOBW]	5.258397	5.51	4	4	4

P37440	Oxidoreductase UcpA OS=Escherichia coli (strain K12) GN=ucpA PE=3 SV=3 - [UCPA_ECOLI]	5.069444	6.46	1	1	1
P64624	Uncharacterized protein YheO OS=Escherichia coli (strain K12) GN=yheO PE=4 SV=1 - [YHEO_ECOLI]	5.051862	11.25	1	1	1
Q46927	tRNA threonylcarbamoyladenosine dehydratase OS=Escherichia coli (strain K12) GN=tcdA PE=1 SV=1 - [TCDA_ECOLI]	5.037906	10.07	1	1	1
P0AGG0	Thiamine-monophosphate kinase OS=Escherichia coli (strain K12) GN=thiL PE=1 SV=1 - [THIL_ECOLI]	4.842215	7.08	2	2	2
P35340	Alkyl hydroperoxide reductase subunit F OS=Escherichia coli (strain K12) GN=ahpF PE=1 SV=2 - [AHPF_ECOLI]	4.762915	4.22	1	1	1
P21645	UDP-3-O-(3-hydroxymyristoyl)glucosamine N-acyltransferase OS=Escherichia coli (strain K12) GN=lpxD PE=1 SV=2 - [LPXD_ECOLI]	4.759872	7.62	1	1	1
B1XCC1	Sugar fermentation stimulation protein A OS=Escherichia coli (strain K12 / DH10B) GN=sfsA PE=3 SV=1 - [SFSA_ECODH]	4.758421	8.55	1	1	2
B1XA76	D-amino acid dehydrogenase OS=Escherichia coli (strain K12 / DH10B) GN=dadA PE=3 SV=1 - [DADA_ECODH]	4.758037	5.56	2	2	2
P37751	Putative glycosyltransferase WbbK OS=Escherichia coli (strain K12) GN=wbbK PE=4 SV=1 - [WBBK_ECOLI]	4.745873	6.18	2	2	2
C4ZS59	N-acetyl-D-glucosamine kinase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=nagK PE=3 SV=1 - [NAGK_ECOBW]	4.709603	5.61	1	1	1
C4ZX73	Uracil phosphoribosyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=upp PE=3 SV=1 - [UPP_ECOBW]	4.684175	10.1	2	2	2
C4ZRR2	Elongation factor Ts OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=tsf PE=3 SV=1 - [EFTS_ECOBW]	4.673696	6.71	1	1	1
B1X929	tRNA pseudouridine synthase A OS=Escherichia coli (strain K12 / DH10B) GN=truA PE=3 SV=1 - [TRUA_ECODH]	4.652658	6.3	1	1	1

P0AE12	AMP nucleosidase OS=Escherichia coli (strain K12) GN=amn PE=1 SV=1 - [AMN_ECOLI]	4.63329	5.17	1	1	1
P21893	Single-stranded-DNA-specific exonuclease RecJ OS=Escherichia coli (strain K12) GN=recJ PE=1 SV=2 - [RECJ_ECOLI]	4.622648	3.47	2	2	2
P0A9J0	Ribonuclease G OS=Escherichia coli (strain K12) GN=rng PE=1 SV=2 - [RNG_ECOLI]	4.542322	4.09	1	1	1
P77580	Acetaldehyde dehydrogenase OS=Escherichia coli (strain K12) GN=mhpF PE=1 SV=1 - [ACDH_ECOLI]	4.538018	7.59	1	1	1
P05719	Type-1 restriction enzyme EcoKI specificity protein OS=Escherichia coli (strain K12) GN=hsdS PE=1 SV=1 - [T1SK_ECOLI]	4.491043	4.09	1	1	1
P0A9D8	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase OS=Escherichia coli (strain K12) GN=dapD PE=1 SV=1 - [DAPD_ECOLI]	4.4323	5.84	1	1	1
B1XG10	Phosphoenolpyruvate synthase regulatory protein OS=Escherichia coli (strain K12 / DH10B) GN=ppsR PE=3 SV=1 - [PSRP_ECODH]	4.428607	6.14	1	1	1
P0A6T3	Galactokinase OS=Escherichia coli (strain K12) GN=galK PE=1 SV=2 - [GAL1_ECOLI]	4.396665	5.24	1	1	1
B1XBX8	Glutamate racemase OS=Escherichia coli (strain K12 / DH10B) GN=murI PE=3 SV=1 - [MURI_ECODH]	4.380487	6.67	1	1	1
P24554	DNA repair protein RadA OS=Escherichia coli (strain K12) GN=radA PE=1 SV=1 - [RADA_ECOLI]	4.365266	4.78	1	1	1
P0A9Q9	Aspartate-semialdehyde dehydrogenase OS=Escherichia coli (strain K12) GN=asd PE=1 SV=1 - [DHAS_ECOLI]	4.34383	4.63	1	1	1
P0AA39	Ribosomal large subunit pseudouridine synthase C OS=Escherichia coli (strain K12) GN=rluC PE=1 SV=1 - [RLUC_ECOLI]	4.304317	5.33	2	2	2
B1XAJ3	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase OS=Escherichia coli (strain K12 / DH10B) GN=metE PE=3 SV=1 - [METE_ECODH]	4.28505	2.12	1	1	1

C4ZZQ5	Sulfate adenylyltransferase subunit 1 OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=cysN PE=3 SV=1 - [CYSN_ECOBW]	4.280105	4	1	1	1
P30744	L-serine dehydratase 2 OS=Escherichia coli (strain K12) GN=sdaB PE=2 SV=2 - [SDHM_ECOLI]	4.188429	4.4	2	2	2
P0ADG7	Inosine-5'-monophosphate dehydrogenase OS=Escherichia coli (strain K12) GN=guaB PE=1 SV=1 - [IMDH_ECOLI]	4.180875	4.1	1	1	1
P0AFG3	2-oxoglutarate dehydrogenase E1 component OS=Escherichia coli (strain K12) GN=sucA PE=1 SV=1 - [ODO1_ECOLI]	4.162061	2.36	2	2	2
Q46925	Cysteine desulfurase CsdA OS=Escherichia coli (strain K12) GN=csdA PE=1 SV=1 - [CSDA_ECOLI]	4.120872	3.49	1	1	1
P04968	L-threonine dehydratase biosynthetic IlvA OS=Escherichia coli (strain K12) GN=ilvA PE=1 SV=1 - [ILVA_ECOLI]	4.111911	5.84	2	2	2
P0ACL0	Glycerol-3-phosphate regulon repressor OS=Escherichia coli (strain K12) GN=glpR PE=1 SV=1 - [GLPR_ECOLI]	4.038617	5.16	1	1	1
P75849	Hydroxyacylglutathione hydrolase GloC OS=Escherichia coli (strain K12) GN=gloC PE=1 SV=1 - [GLO22_ECOLI]	4.029521	11.16	1	1	1
P04693	Aromatic-amino-acid aminotransferase OS=Escherichia coli (strain K12) GN=tyrB PE=1 SV=1 - [TYRB_ECOLI]	3.977978	4.28	1	1	1
P68187	Maltose/maltodextrin import ATP-binding protein MalK OS=Escherichia coli (strain K12) GN=malK PE=1 SV=1 - [MALK_ECOLI]	3.945694	6.2	1	1	1
P37051	Formyltetrahydrofolate deformylase OS=Escherichia coli (strain K12) GN=purU PE=1 SV=1 - [PURU_ECOLI]	3.882563	5.36	1	1	1
P0A9P0	Dihydrolipoyl dehydrogenase OS=Escherichia coli (strain K12) GN=lpdA PE=1 SV=2 - [DLDH_ECOLI]	3.842908	4.43	1	1	1
P0AFP6	GTP cyclohydrolase 1 type 2 homolog OS=Escherichia coli (strain K12) GN=ybgI PE=1 SV=1 - [GCH1L_ECOLI]	3.710349	6.88	1	1	1

P36928	Uncharacterized chaperone protein YegD OS=Escherichia coli (strain K12) GN=yegD PE=3 SV=2 - [YEGD_ECOLI]	3.709621	3.78	1	1	1
C4ZZ41	Dihydroxy-acid dehydratase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=ilvD PE=3 SV=1 - [ILVD_ECOBW]	3.615542	2.27	1	1	1
P0A996	Anaerobic glycerol-3-phosphate dehydrogenase subunit C OS=Escherichia coli (strain K12) GN=glpC PE=1 SV=1 - [GLPC_ECOLI]	3.559338	3.79	1	1	1
B1XBY7	50S ribosomal protein L10 OS=Escherichia coli (strain K12 / DH10B) GN=rplJ PE=3 SV=1 - [RL10_ECODH]	3.546613	9.7	1	1	1
P75682	Probable 2-keto-3-deoxy-galactonate aldolase YagE OS=Escherichia coli (strain K12) GN=yagE PE=1 SV=2 - [YAGE_ECOLI]	3.544104	5.63	1	1	1
P00490	Maltodextrin phosphorylase OS=Escherichia coli (strain K12) GN=malP PE=1 SV=7 - [PHSM_ECOLI]	3.536541	2.01	1	1	1
C5A0R1	Soluble pyridine nucleotide transhydrogenase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=sthA PE=3 SV=1 - [STHA_ECOBW]	3.506781	4.08	1	1	1
P77252	Uncharacterized protein YkgE OS=Escherichia coli (strain K12) GN=ykgE PE=3 SV=1 - [YKGE_ECOLI]	3.473802	6.28	1	1	1
C4ZTK1	7-cyano-7-deazaguanine synthase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=queC PE=3 SV=1 - [QUEC_ECOBW]	3.427602	9.96	1	1	1
P39099	Periplasmic pH-dependent serine endoprotease DegQ OS=Escherichia coli (strain K12) GN=degQ PE=1 SV=1 - [DEGQ_ECOLI]	3.379516	3.08	1	1	1
P52129	mRNA endoribonuclease LS OS=Escherichia coli (strain K12) GN=rnlA PE=1 SV=2 - [RNLA_ECOLI]	3.355839	3.92	1	1	1
B1XAX4	Phosphoribosylformylglycinamidine cyclo-ligase OS=Escherichia coli (strain K12 / DH10B) GN=purM PE=3 SV=1 - [PUR5_ECODH]	3.344254	4.64	1	1	1
P0AGJ5	Uncharacterized tRNA/rRNA methyltransferase YfiF OS=Escherichia coli (strain K12) GN=yfiF PE=3 SV=1 - [YFIF_ECOLI]	3.337731	6.67	1	1	1

B1XEK1	HTH-type transcriptional regulator ArgP OS=Escherichia coli (strain K12 / DH10B) GN=argP PE=3 SV=1 - [ARGP_ECODH]	3.288699	5.05	1	1	1
P77737	Oligopeptide transport ATP-binding protein OppF OS=Escherichia coli (strain K12) GN=oppF PE=3 SV=1 - [OPPF_ECOLI]	3.267135	9.28	3	3	3
P0AFM6	Phage shock protein A OS=Escherichia coli (strain K12) GN=pspA PE=1 SV=2 - [PSPA_ECOLI]	3.256055	5.41	1	1	1
P69908	Glutamate decarboxylase alpha OS=Escherichia coli (strain K12) GN=gadA PE=1 SV=1 - [DCEA_ECOLI]	3.256007	3.65	1	1	1
C5A0W2	Glucose-6-phosphate isomerase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=pgi PE=3 SV=1 - [G6PI_ECOBW]	3.246464	7.1	2	2	2
P15038	DNA helicase IV OS=Escherichia coli (strain K12) GN=helD PE=1 SV=2 - [HELD_ECOLI]	3.245464	4.09	2	2	2
C4ZXM7	Formamidopyrimidine-DNA glycosylase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=mutM PE=3 SV=1 - [FPG_ECOBW]	3.19261	6.32	1	1	1
P0AFG6	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex OS=Escherichia coli (strain K12) GN=sucB PE=1 SV=2 - [ODO2_ECOLI]	3.167551	4.44	1	1	1
P0ACL2	Exu regulon transcriptional regulator OS=Escherichia coli (strain K12) GN=exuR PE=4 SV=1 - [EXUR_ECOLI]	3.155151	6.2	1	1	1
P75960	NAD-dependent protein deacylase OS=Escherichia coli (strain K12) GN=cobB PE=1 SV=3 - [NPD_ECOLI]	3.129424	7.89	1	1	1
P07003	Pyruvate dehydrogenase [ubiquinone] OS=Escherichia coli (strain K12) GN=poxB PE=1 SV=1 - [POXB_ECOLI]	3.111768	2.8	1	1	1
P0AF08	Iron-sulfur cluster carrier protein OS=Escherichia coli (strain K12) GN=mrp PE=3 SV=1 - [APBC_ECOLI]	3.106342	4.34	1	1	1
B1XBE0	Chaperone protein DnaJ OS=Escherichia coli (strain K12 / DH10B) GN=dnaJ PE=3 SV=1 - [DNAJ_ECODH]	3.104943	5.85	2	2	2

P76422	Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase OS=Escherichia coli (strain K12) GN=thiD PE=1 SV=1 - [THID_ECOLI]	3.097109	9.4	1	1	1
C4ZPZ5	3-isopropylmalate dehydratase large subunit OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=leuC PE=3 SV=1 - [LEUC_ECOBW]	3.082834	4.72	1	1	1
P46118	HTH-type transcriptional regulator HexR OS=Escherichia coli (strain K12) GN=hexR PE=3 SV=2 - [HEXR_ECOLI]	3.056759	5.19	1	1	1
P0AFB5	Nitrogen regulation protein NR(II) OS=Escherichia coli (strain K12) GN=glnL PE=1 SV=1 - [NTRB_ECOLI]	3.055393	4.3	1	1	1
P0AG44	50S ribosomal protein L17 OS=Escherichia coli (strain K12) GN=rplQ PE=1 SV=1 - [RL17_ECOLI]	3.046828	7.87	1	1	1
P76586	Uncharacterized protein YphH OS=Escherichia coli (strain K12) GN=yphH PE=3 SV=2 - [YPHH_ECOLI]	3.046201	2.77	1	1	1
P0AFL9	Paraquat-inducible protein A OS=Escherichia coli (strain K12) GN=pqiA PE=2 SV=1 - [PQIA_ECOLI]	3.032519	3.12	1	1	1
P27128	Lipopolysaccharide 1,3-galactosyltransferase OS=Escherichia coli (strain K12) GN=rfaI PE=3 SV=1 - [RFAI_ECOLI]	2.985376	3.54	1	1	1
C4ZUX7	Bifunctional protein FolD OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=folD PE=3 SV=1 - [FOLD_ECOBW]	2.959088	5.9	1	1	1
P0A9L3	FKBP-type 22 kDa peptidyl-prolyl cis-trans isomerase OS=Escherichia coli (strain K12) GN=fklB PE=1 SV=2 - [FKBB_ECOLI]	2.941667	6.31	1	1	1
P0AG76	Nuclease SbcCD subunit D OS=Escherichia coli (strain K12) GN=sbcD PE=1 SV=1 - [SBCD_ECOLI]	2.927418	3.25	1	1	1
C4ZUV3	tRNA 2-selenouridine synthase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=selU PE=3 SV=1 - [SELU_ECOBW]	2.916784	3.57	1	1	1
C4ZWK2	Succinate--CoA ligase [ADP-forming] subunit beta OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=sucC PE=3 SV=1 - [SUCC_ECOBW]	2.915723	4.12	1	1	1

P77806	Methionine aminotransferase OS=Escherichia coli (strain K12) GN=ybdL PE=1 SV=1 - [YBDL_ECOLI]	2.893929	3.63	1	1	1
P00888	Phospho-2-dehydro-3-deoxyheptonate aldolase, Tyr-sensitive OS=Escherichia coli (strain K12) GN=aroF PE=1 SV=1 - [AROF_ECOLI]	2.890631	3.65	1	1	1
P18775	Dimethyl sulfoxide reductase DmsA OS=Escherichia coli (strain K12) GN=dmsA PE=1 SV=2 - [DMSA_ECOLI]	2.874827	1.6	1	1	1
B1XDL5	Ribosomal RNA large subunit methyltransferase M OS=Escherichia coli (strain K12 / DH10B) GN=rlmM PE=3 SV=1 - [RLMM_ECODH]	2.870817	3.01	1	1	1
P0A8V0	Ribonuclease BN OS=Escherichia coli (strain K12) GN=rbn PE=1 SV=1 - [RBN_ECOLI]	2.85076	4.92	1	1	1
P33643	Ribosomal large subunit pseudouridine synthase D OS=Escherichia coli (strain K12) GN=rluD PE=1 SV=4 - [RLUD_ECOLI]	2.841599	4.6	1	1	1
P0A9K9	FKBP-type peptidyl-prolyl cis-trans isomerase SlyD OS=Escherichia coli (strain K12) GN=slyD PE=1 SV=1 - [SLYD_ECOLI]	2.816442	4.59	1	1	1
C4ZS67	Peptidase T OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=pepT PE=3 SV=1 - [PEPT_ECOBW]	2.814862	2.94	1	1	1
C4ZZ22	Aspartate--ammonia ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=asnA PE=3 SV=1 - [ASNA_ECOBW]	2.784961	4.55	1	1	1
P12282	Molybdopterin-synthase adenylyltransferase OS=Escherichia coli (strain K12) GN=moeB PE=1 SV=1 - [MOEB_ECOLI]	2.784947	4.02	1	1	1
P0A832	SsrA-binding protein OS=Escherichia coli (strain K12) GN=smpB PE=1 SV=2 - [SSRP_ECOLI]	2.783981	6.88	1	1	1
P37013	Anaerobic nitric oxide reductase transcription regulator NorR OS=Escherichia coli (strain K12) GN=norR PE=1 SV=2 - [NORR_ECOLI]	2.781419	1.98	1	1	1
P0AC38	Aspartate ammonia-lyase OS=Escherichia coli (strain K12) GN=aspA PE=1 SV=1 - [ASPA_ECOLI]	2.759288	1.88	1	1	1

B1XFJ2	Thymidine phosphorylase OS=Escherichia coli (strain K12 / DH10B) GN=deoA PE=3 SV=1 - [TYPH_ECODH]	2.750229	3.18	1	1	1
B1XC53	4-hydroxythreonine-4-phosphate dehydrogenase OS=Escherichia coli (strain K12 / DH10B) GN=pdxA PE=3 SV=1 - [PDXA_ECODH]	2.744393	6.08	1	1	1
P08390	USG-1 protein OS=Escherichia coli (strain K12) GN=usg PE=1 SV=1 - [USG_ECOLI]	2.744247	3.26	1	1	1
P0AC41	Succinate dehydrogenase flavoprotein subunit OS=Escherichia coli (strain K12) GN=sdhA PE=1 SV=1 - [SDHA_ECOLI]	2.741433	1.53	1	1	1
P27126	Lipopolysaccharide core biosynthesis protein RfaS OS=Escherichia coli (strain K12) GN=rfaS PE=4 SV=1 - [RFAS_ECOLI]	2.727968	3.86	1	1	1
P08142	Acetolactate synthase isozyme 1 large subunit OS=Escherichia coli (strain K12) GN=ilvB PE=1 SV=1 - [ILVB_ECOLI]	2.711188	2.49	1	1	1
P0AF26	Nitrate reductase molybdenum cofactor assembly chaperone NarJ OS=Escherichia coli (strain K12) GN=narJ PE=1 SV=1 - [NARJ_ECOLI]	2.709156	5.93	1	1	1
B1XFI4	Peptide chain release factor 3 OS=Escherichia coli (strain K12 / DH10B) GN=prfC PE=3 SV=1 - [RF3_ECODH]	2.681521	2.84	1	1	1
P77360	Uncharacterized zinc-type alcohol dehydrogenase-like protein YphC OS=Escherichia coli (strain K12) GN=yphC PE=3 SV=2 - [YPHC_ECOLI]	2.671574	4.25	1	1	1
P17115	Arabinose 5-phosphate isomerase GutQ OS=Escherichia coli (strain K12) GN=gutQ PE=1 SV=3 - [GUTQ_ECOLI]	2.662151	2.49	1	1	1
P67660	Probable HTH-type transcriptional regulator YhaJ OS=Escherichia coli (strain K12) GN=yhaJ PE=3 SV=1 - [YHAJ_ECOLI]	2.635306	3.36	1	1	1
P0AG16	Amidophosphoribosyltransferase OS=Escherichia coli (strain K12) GN=purF PE=1 SV=2 - [PUR1_ECOLI]	2.618547	2.38	1	1	1
P28904	Trehalose-6-phosphate hydrolase OS=Escherichia coli (strain K12) GN=treC PE=1 SV=3 - [TREC_ECOLI]	2.6067	2.54	1	1	1

P07604	Transcriptional regulatory protein TyrR OS=Escherichia coli (strain K12) GN=tyrR PE=1 SV=2 - [TYRR_ECOLI]	2.559558	1.75	1	1	1
P0ABH7	Citrate synthase OS=Escherichia coli (strain K12) GN=gltA PE=1 SV=1 - [CISY_ECOLI]	2.559061	3.28	1	1	1
P0AES0	Bifunctional glutathionylspermidine synthetase/amidase OS=Escherichia coli (strain K12) GN=gss PE=1 SV=1 - [GSP_ECOLI]	2.547998	2.42	1	1	1
P31436	Sugar efflux transporter C OS=Escherichia coli (strain K12) GN=setC PE=1 SV=1 - [SETC_ECOLI]	2.506188	6.35	1	1	1
B1XGC4	Cysteine--tRNA ligase OS=Escherichia coli (strain K12 / DH10B) GN=cysS PE=3 SV=1 - [SYC_ECODH]	2.49866	6.07	2	2	2
P0ABQ2	2-hydroxy-3-oxopropionate reductase OS=Escherichia coli (strain K12) GN=garR PE=1 SV=1 - [GARR_ECOLI]	2.444608	5.1	1	1	1
P21179	Catalase HPII OS=Escherichia coli (strain K12) GN=katE PE=1 SV=1 - [CATE_ECOLI]	2.442985	1.46	1	1	1
C4ZWF6	Glutamine--tRNA ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=glnS PE=3 SV=1 - [SYQ_ECOBW]	2.427482	1.99	1	1	1
P33595	HTH-type transcriptional regulator SgrR OS=Escherichia coli (strain K12) GN=sgrR PE=1 SV=2 - [SGRR_ECOLI]	2.420133	2.36	1	1	1
C4ZVX9	Glycogen synthase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=glgA PE=3 SV=1 - [GLGA_ECOBW]	2.402691	2.31	1	1	1
P0CG19	Inactive ribonuclease PH OS=Escherichia coli (strain K12) GN=rph PE=3 SV=1 - [RNPH_ECOLI]	2.373947	5.7	1	1	1
C5A1F5	Small ribosomal subunit biogenesis GTPase RsgA OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=rsgA PE=3 SV=1 - [RSGA_ECOBW]	2.372116	2.57	1	1	1
P37634	Ribosomal RNA large subunit methyltransferase J OS=Escherichia coli (strain K12) GN=rlmJ PE=1 SV=1 - [RLMJ_ECOLI]	2.369488	3.57	1	1	1

P33021	Putative nucleoside permease NupX OS=Escherichia coli (strain K12) GN=nupX PE=1 SV=1 - [NUPX_ECOLI]	2.358887	2.4	1	1	1
P19323	Formate hydrogenlyase transcriptional activator FhlA OS=Escherichia coli (strain K12) GN=fhlA PE=1 SV=1 - [FHLA_ECOLI]	2.349321	1.73	1	1	1
P0A870	Transaldolase B OS=Escherichia coli (strain K12) GN=talB PE=1 SV=2 - [TALB_ECOLI]	2.341436	2.84	1	1	1
P13035	Aerobic glycerol-3-phosphate dehydrogenase OS=Escherichia coli (strain K12) GN=glpD PE=1 SV=3 - [GLPD_ECOLI]	2.298097	3.59	2	2	2
P00936	Adenylate cyclase OS=Escherichia coli (strain K12) GN=cyaA PE=1 SV=5 - [CYAA_ECOLI]	2.261919	1.06	1	1	1
C5A0G3	Peptide chain release factor 2 OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=prfB PE=3 SV=1 - [RF2_ECOBW]	2.246346	3.84	1	1	1
P0AFQ7	Uncharacterized metal-dependent hydrolase YcfH OS=Escherichia coli (strain K12) GN=ycfH PE=1 SV=1 - [YCFH_ECOLI]	2.24573	3.77	1	1	1
P0A8A0	Probable transcriptional regulatory protein YebC OS=Escherichia coli (strain K12) GN=yebC PE=1 SV=1 - [YEBC_ECOLI]	2.213795	4.07	1	1	1
P0ACB2	Delta-aminolevulinic acid dehydratase OS=Escherichia coli (strain K12) GN=hemB PE=1 SV=2 - [HEM2_ECOLI]	2.209007	2.16	1	1	1
P0AEN1	NAD(P)H-flavin reductase OS=Escherichia coli (strain K12) GN=fre PE=1 SV=2 - [FRE_ECOLI]	2.189597	3.86	1	1	1
P09053	Valine--pyruvate aminotransferase OS=Escherichia coli (strain K12) GN=avtA PE=1 SV=3 - [AVTA_ECOLI]	2.188671	2.16	1	1	1
P33919	Putative DNA repair helicase RadD OS=Escherichia coli (strain K12) GN=radD PE=1 SV=2 - [RADD_ECOLI]	2.1846	1.88	1	1	1
P0ADR8	Pyrimidine/purine nucleotide 5'-monophosphate nucleosidase OS=Escherichia coli (strain K12) GN=ppnN PE=1 SV=1 - [PPNN_ECOLI]	2.178563	2.42	1	1	1

P24215	Mannonate dehydratase OS=Escherichia coli (strain K12) GN=uxuA PE=1 SV=2 - [UXUA_ECOLI]	2.169566	3.05	1	1	1
B1X7D7	Ribosomal RNA large subunit methyltransferase F OS=Escherichia coli (strain K12 / DH10B) GN=rlmF PE=3 SV=1 - [RLMF_ECODH]	2.158236	8.44	1	1	1
P05824	DNA repair protein RecN OS=Escherichia coli (strain K12) GN=recN PE=1 SV=2 - [RECN_ECOLI]	2.156606	2.53	1	1	1
C4ZQP1	Protein-methionine-sulfoxide reductase catalytic subunit MsrP OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=msrP PE=3 SV=1 - [MSRP_ECOBW]	2.1096	2.1	1	1	1
P0AEX9	Maltose-binding periplasmic protein OS=Escherichia coli (strain K12) GN=malE PE=1 SV=1 - [MALE_ECOLI]	2.079256	5.56	1	1	1
P0ACR7	Uncharacterized HTH-type transcriptional regulator YfeR OS=Escherichia coli (strain K12) GN=yfeR PE=4 SV=1 - [YFER_ECOLI]	1.95253	1.95	1	1	1
P27241	Lipopolysaccharide core biosynthesis protein RfaZ OS=Escherichia coli (strain K12) GN=rfaZ PE=4 SV=2 - [RFAZ_ECOLI]	1.930532	2.47	1	1	1
P31979	NADH-quinone oxidoreductase subunit F OS=Escherichia coli (strain K12) GN=nuoF PE=1 SV=3 - [NUOF_ECOLI]	1.879873	1.8	1	1	1
P09831	Glutamate synthase [NADPH] large chain OS=Escherichia coli (strain K12) GN=gltB PE=1 SV=3 - [GLTB_ECOLI]	1.876248	0.47	1	1	1
P76257	Probable ATP-dependent helicase YoaA OS=Escherichia coli (strain K12) GN=yoaA PE=3 SV=1 - [YOAA_ECOLI]	1.784524	1.26	1	1	1
P02931	Outer membrane protein F OS=Escherichia coli (strain K12) GN=ompF PE=1 SV=1 - [OMPF_ECOLI]	1.736053	3.04	1	1	1
P45807	Uncharacterized protein YbaM OS=Escherichia coli (strain K12) GN=ybaM PE=4 SV=1 - [YBAM_ECOLI]	1.71385	11.32	1	1	1
P0AB80	Branched-chain-amino-acid aminotransferase OS=Escherichia coli (strain K12) GN=ilvE PE=1 SV=2 - [ILVE_ECOLI]	1.686642	3.56	1	1	1

B1X9K4	UPF0115 protein YfcN OS=Escherichia coli (strain K12 / DH10B) GN=yfcN PE=3 SV=1 - [YFCN_ECODH]	0	3.83	1	1	1
P06149	D-lactate dehydrogenase OS=Escherichia coli (strain K12) GN=dld PE=1 SV=3 - [DLD_ECOLI]	0	1.23	1	1	1
P07364	Chemotaxis protein methyltransferase OS=Escherichia coli (strain K12) GN=cheR PE=4 SV=2 - [CHER_ECOLI]	0	4.55	1	1	1
P07762	1,4-alpha-glucan branching enzyme GlgB OS=Escherichia coli (strain K12) GN=glgB PE=1 SV=1 - [GLGB_ECOLI]	0	0.82	1	1	1
P0A715	2-dehydro-3-deoxyphosphooctonate aldolase OS=Escherichia coli (strain K12) GN=kdsA PE=1 SV=1 - [KDSA_ECOLI]	0	2.11	1	1	1
P0A9E5	Fumarate and nitrate reduction regulatory protein OS=Escherichia coli (strain K12) GN=fnr PE=1 SV=1 - [FNR_ECOLI]	0	2.4	1	1	1
P0AC33	Fumarate hydratase class I, aerobic OS=Escherichia coli (strain K12) GN=fumA PE=1 SV=2 - [FUMA_ECOLI]	0	3.1	1	1	1
P0ACX5	Fumarase D OS=Escherichia coli (strain K12) GN=fumD PE=1 SV=1 - [FUMD_ECOLI]	0	24.64	1	1	1
P0AEE8	DNA adenine methylase OS=Escherichia coli (strain K12) GN=dam PE=1 SV=1 - [DMA_ECOLI]	0	6.12	1	1	1
P0AGI8	Trk system potassium uptake protein TrkA OS=Escherichia coli (strain K12) GN=trkA PE=1 SV=1 - [TRKA_ECOLI]	0	1.97	1	1	1
P37692	ADP-heptose--LPS heptosyltransferase 2 OS=Escherichia coli (strain K12) GN=rfaF PE=1 SV=1 - [RFAF_ECOLI]	0	3.74	1	1	1
P46855	Uncharacterized protein YhhZ OS=Escherichia coli (strain K12) GN=yhhZ PE=3 SV=1 - [YHHZ_ECOLI]	0	3.83	1	1	1
P64515	Inner membrane protein YecN OS=Escherichia coli (strain K12) GN=yecN PE=1 SV=1 - [YECN_ECOLI]	0	15.27	1	1	1

P64612	Cell division protein ZapE OS=Escherichia coli (strain K12) GN=zapE PE=1 SV=1 - [ZAPE_ECOLI]	0	2.4	1	1	1
P77302	Probable diguanylate cyclase YdaM OS=Escherichia coli (strain K12) GN=ydaM PE=1 SV=2 - [YDAM_ECOLI]	0	1.95	1	1	1
P77732	Uncharacterized HTH-type transcriptional regulator RhmR OS=Escherichia coli (strain K12) GN=rhmR PE=1 SV=1 - [RHMR_ECOLI]	0	4.23	1	1	1
P77739	Putative kinase YniA OS=Escherichia coli (strain K12) GN=yniA PE=3 SV=1 - [YNIA_ECOLI]	0	5.94	1	1	1

Table A2: The BioID interactome of S18.BirA*. Proteins interacting with S18.BirA* were identified by mass spectrometry.

Accession	Description	Score	Coverage	#Unique Peptides	# Peptides	# PSMs
P0CE47	Elongation factor Tu 1 OS=Escherichia coli (strain K12) GN=tufA PE=1 SV=1 - [EFTU1_ECOLI]	4840.121	88.07	38	38	1330
C4ZSH9	D-tagatose-1,6-bisphosphate aldolase subunit GatZ OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=gatZ PE=3 SV=1 - [GATZ_ECOBW]	974.1245	84.05	26	26	212
P0AG30	Transcription termination factor Rho OS=Escherichia coli (strain K12) GN=rho PE=1 SV=1 - [RHO_ECOLI]	749.2824	66.59	33	33	198
P16659	Proline--tRNA ligase OS=Escherichia coli (strain K12) GN=proS PE=1 SV=4 - [SYP_ECOLI]	499.3476	81.29	35	35	131

B1XBZ0	DNA-directed RNA polymerase subunit beta' OS=Escherichia coli (strain K12 / DH10B) GN=rpoC PE=3 SV=1 - [RPOC_ECODH]	468.357	57	65	65	128
B1XGY0	Translation initiation factor IF-2 OS=Escherichia coli (strain K12 / DH10B) GN=infB PE=3 SV=1 - [IF2_ECODH]	463.1146	58.2	49	49	114
C5A0L1	S-adenosylmethionine synthase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=metK PE=3 SV=1 - [METK_ECOBW]	426.3784	69.79	21	21	107
B1XBY9	DNA-directed RNA polymerase subunit beta OS=Escherichia coli (strain K12 / DH10B) GN=rpoB PE=3 SV=1 - [RPOB_ECODH]	421.6949	56.86	67	67	115
P09373	Formate acetyltransferase 1 OS=Escherichia coli (strain K12) GN=pflB PE=1 SV=2 - [PFLB_ECOLI]	415.4801	59.47	36	38	104
P0A8M3	Threonine--tRNA ligase OS=Escherichia coli (strain K12) GN=thrS PE=1 SV=1 - [SYT_ECOLI]	334.4552	52.34	35	35	97
P32132	GTP-binding protein TypA/BipA OS=Escherichia coli (strain K12) GN=typA PE=1 SV=2 - [TYPA_ECOLI]	318.5668	50.58	27	27	81
C4ZQ19	Serine--tRNA ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=serS PE=3 SV=1 - [SYS_ECOBW]	284.1021	62.33	27	27	75
P0AES6	DNA gyrase subunit B OS=Escherichia coli (strain K12) GN=gyrB PE=1 SV=2 - [GYRB_ECOLI]	270.452	57.96	40	40	62
P0A6A3	Acetate kinase OS=Escherichia coli (strain K12) GN=ackA PE=1 SV=1 - [ACKA_ECOLI]	257.6806	83.5	22	22	57

P24182	Biotin carboxylase OS=Escherichia coli (strain K12) GN=accC PE=1 SV=2 - [ACCC_ECOLI]	229.1684	60.13	25	25	56
P0A953	3-oxoacyl-[acyl-carrier-protein] synthase 1 OS=Escherichia coli (strain K12) GN=fabB PE=1 SV=1 - [FABB_ECOLI]	223.8288	50.25	12	12	47
P00961	Glycine--tRNA ligase beta subunit OS=Escherichia coli (strain K12) GN=glyS PE=1 SV=4 - [SYGB_ECOLI]	210.2762	61.39	31	31	52
P09152	Respiratory nitrate reductase 1 alpha chain OS=Escherichia coli (strain K12) GN=narG PE=1 SV=4 - [NARG_ECOLI]	197.9239	35.93	28	33	56
P0AAI5	3-oxoacyl-[acyl-carrier-protein] synthase 2 OS=Escherichia coli (strain K12) GN=fabF PE=1 SV=2 - [FABF_ECOLI]	186.6945	50.12	13	13	50
P0A9Q7	Aldehyde-alcohol dehydrogenase OS=Escherichia coli (strain K12) GN=adhE PE=1 SV=2 - [ADHE_ECOLI]	175.5264	43.32	29	29	50
C4ZRJ3	Protein translocase subunit SecA OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=secA PE=3 SV=1 - [SECA_ECOBW]	171.9648	44.06	29	29	46
P0A9M0	Lon protease OS=Escherichia coli (strain K12) GN=lon PE=1 SV=1 - [LON_ECOLI]	166.4379	45.15	31	31	46
P0A9C0	Anaerobic glycerol-3-phosphate dehydrogenase subunit A OS=Escherichia coli (strain K12) GN=glpA PE=1 SV=1 - [GLPA_ECOLI]	164.1077	63.47	27	27	43
C5A098	ATP-dependent protease ATPase subunit HslU OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=hslU PE=3 SV=1 - [HSLU_ECOBW]	159.8542	55.53	24	24	46

B1X6J0	Elongation factor G OS=Escherichia coli (strain K12 / DH10B) GN=fusA PE=3 SV=1 - [EFG_ECODH]	158.6586	47.59	25	25	42
C4ZZ10	ATP synthase subunit beta OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=atpD PE=3 SV=1 - [ATPB_ECOBW]	156.2869	69.13	23	23	43
P0AEI1	tRNA-2-methylthio-N(6)-dimethylallyladenosine synthase OS=Escherichia coli (strain K12) GN=miaB PE=1 SV=1 - [MIAB_ECOLI]	153.3377	62.45	28	28	38
P0A9P6	ATP-dependent RNA helicase DeaD OS=Escherichia coli (strain K12) GN=deaD PE=1 SV=2 - [DEAD_ECOLI]	150.333	46.9	23	23	43
C4ZPY3	RNA polymerase-associated protein RapA OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=rapA PE=3 SV=1 - [RAPA_ECOBW]	149.5158	37.71	29	29	41
P76658	Bifunctional protein HldE OS=Escherichia coli (strain K12) GN=hldE PE=1 SV=1 - [HLDE_ECOLI]	147.7299	59.96	24	24	41
P21507	ATP-dependent RNA helicase SrmB OS=Escherichia coli (strain K12) GN=srmB PE=1 SV=1 - [SRMB_ECOLI]	138.009	51.58	19	19	39
C4ZSI0	D-tagatose-1,6-bisphosphate aldolase subunit GatY OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=gatY PE=3 SV=1 - [GATY_ECOBW]	137.6784	46.13	12	12	37
P08201	Nitrite reductase (NADH) large subunit OS=Escherichia coli (strain K12) GN=nirB PE=3 SV=4 - [NIRB_ECOLI]	136.9504	34.83	22	22	36

P28903	Anaerobic ribonucleoside-triphosphate reductase OS=Escherichia coli (strain K12) GN=nrdD PE=1 SV=2 - [NRDD_ECOLI]	133.968	37.08	22	22	36
P22523	Chromosome partition protein MukB OS=Escherichia coli (strain K12) GN=mukB PE=1 SV=2 - [MUKB_ECOLI]	129.4122	24.83	27	27	34
P00561	Bifunctional aspartokinase/homoserine dehydrogenase 1 OS=Escherichia coli (strain K12) GN=thrA PE=1 SV=2 - [AK1H_ECOLI]	127.1817	41.22	23	23	32
C5A093	Glycerol kinase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=glpK PE=3 SV=1 - [GLPK_ECOBW]	122.3348	47.01	17	17	36
B1XFM6	ATP-dependent Clp protease ATP-binding subunit ClpX OS=Escherichia coli (strain K12 / DH10B) GN=clpX PE=3 SV=1 - [CLPX_ECODH]	118.7388	56.37	16	16	28
P00562	Bifunctional aspartokinase/homoserine dehydrogenase 2 OS=Escherichia coli (strain K12) GN=metL PE=1 SV=3 - [AK2H_ECOLI]	115.3421	43.09	26	26	34
P00363	Fumarate reductase flavoprotein subunit OS=Escherichia coli (strain K12) GN=frdA PE=1 SV=3 - [FRDA_ECOLI]	114.7447	49.34	20	20	30
P0A9X4	Rod shape-determining protein MreB OS=Escherichia coli (strain K12) GN=mreB PE=1 SV=1 - [MREB_ECOLI]	110.0702	48.7	16	16	30
P21170	Biosynthetic arginine decarboxylase OS=Escherichia coli (strain K12) GN=speA PE=1 SV=2 - [SPEA_ECOLI]	109.7203	26.6	16	16	27

B1X7X6	Ribosomal protein S12 methylthiotransferase RimO OS=Escherichia coli (strain K12 / DH10B) GN=rimO PE=3 SV=1 - [RIMO_ECODH]	109.3789	30.39	10	10	30
C4ZXA5	Cysteine desulfurase IscS OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=iscS PE=3 SV=1 - [ISCS_ECOBW]	107.3517	43.81	15	15	31
C4ZY45	Hydroxylamine reductase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=hcp PE=3 SV=1 - [HCP_ECOBW]	106.8816	48.73	17	17	27
B1XDS8	Adenylosuccinate synthetase OS=Escherichia coli (strain K12 / DH10B) GN=purA PE=3 SV=1 - [PURA_ECODH]	106.834	49.07	19	19	30
C4ZQF0	Aspartate--tRNA ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=aspS PE=3 SV=1 - [SYD_ECOBW]	105.9647	46.61	22	22	29
P0AFG8	Pyruvate dehydrogenase E1 component OS=Escherichia coli (strain K12) GN=aceE PE=1 SV=2 - [ODP1_ECOLI]	105.9261	33.15	27	27	32
C4ZRP7	Glutamate-1-semialdehyde 2,1-aminomutase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=hemL PE=3 SV=1 - [GSA_ECOBW]	100.2793	49.3	11	11	27
P0A9A6	Cell division protein FtsZ OS=Escherichia coli (strain K12) GN=ftsZ PE=1 SV=1 - [FTSZ_ECOLI]	96.84785	54.83	15	15	24
C4ZX90	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (flavodoxin) OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=ispG PE=3 SV=1 - [ISPG_ECOBW]	96.61245	57.53	18	18	27
P0A9H3	Inducible lysine decarboxylase OS=Escherichia coli (strain K12) GN=cadA PE=1 SV=1 - [LDCI_ECOLI]	96.56312	38.18	18	18	27

C4ZUH2	50S ribosomal protein L2 OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=rplB PE=3 SV=1 - [RL2_ECOBW]	95.01873	52.75	14	14	34
P13033	Anaerobic glycerol-3-phosphate dehydrogenase subunit B OS=Escherichia coli (strain K12) GN=glpB PE=1 SV=1 - [GLPB_ECOLI]	94.44764	43.44	10	10	19
P07913	L-threonine 3-dehydrogenase OS=Escherichia coli (strain K12) GN=tdh PE=1 SV=1 - [TDH_ECOLI]	94.31844	32.26	9	9	25
P0A749	UDP-N-acetylglucosamine 1-carboxyvinyltransferase OS=Escherichia coli (strain K12) GN=murA PE=1 SV=1 - [MURA_ECOLI]	93.44086	60.62	17	17	27
P00934	Threonine synthase OS=Escherichia coli (strain K12) GN=thrC PE=1 SV=1 - [THRC_ECOLI]	92.85761	46.5	14	14	22
P0A9W3	Energy-dependent translational throttle protein EttA OS=Escherichia coli (strain K12) GN=ettA PE=1 SV=2 - [ETTA_ECOLI]	91.54861	44.32	17	17	25
P00960	Glycine--tRNA ligase alpha subunit OS=Escherichia coli (strain K12) GN=glyQ PE=1 SV=2 - [SYGA_ECOLI]	90.89137	43.23	10	10	27
B1XF11	tRNA sulfurtransferase OS=Escherichia coli (strain K12 / DH10B) GN=thiI PE=3 SV=1 - [THII_ECODH]	89.27102	47.3	18	18	28
P0AGD7	Signal recognition particle protein OS=Escherichia coli (strain K12) GN=ffh PE=1 SV=1 - [SRP54_ECOLI]	88.89584	49.01	18	18	24
B1XB43	Elongation factor 4 OS=Escherichia coli (strain K12 / DH10B) GN=lepA PE=3 SV=1 - [LEPA_ECODH]	88.46764	33.72	17	17	24

C4ZXC6	Serine hydroxymethyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=glyA PE=3 SV=1 - [GLYA_ECOBW]	88.45534	29.74	10	10	21
B1X8Z8	NADH-quinone oxidoreductase subunit C/D OS=Escherichia coli (strain K12 / DH10B) GN=nuoC PE=3 SV=1 - [NUOCD_ECODH]	88.27257	40	20	20	27
P76536	Probable deferrochelatase/peroxidase YfeX OS=Escherichia coli (strain K12) GN=yfeX PE=1 SV=2 - [YFEX_ECOLI]	88.1865	46.15	9	9	23
C4ZRI6	UDP-N-acetylmuramate--L-alanine ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=murC PE=3 SV=1 - [MURC_ECOBW]	86.19536	31.77	12	12	26
P0AGF6	L-threonine dehydratase catabolic TdcB OS=Escherichia coli (strain K12) GN=tdcB PE=1 SV=1 - [TDCB_ECOLI]	85.38713	53.19	13	13	20
P08142	Acetolactate synthase isozyme 1 large subunit OS=Escherichia coli (strain K12) GN=ilvB PE=1 SV=1 - [ILVB_ECOLI]	85.28022	39.15	12	12	21
B1XDJ0	CTP synthase OS=Escherichia coli (strain K12 / DH10B) GN=pyrG PE=3 SV=1 - [PYRG_ECODH]	84.07355	39.27	18	18	26
B1XBY6	50S ribosomal protein L1 OS=Escherichia coli (strain K12 / DH10B) GN=rplA PE=3 SV=1 - [RL1_ECODH]	82.82238	46.15	9	9	25
P11349	Respiratory nitrate reductase 1 beta chain OS=Escherichia coli (strain K12) GN=narH PE=1 SV=3 - [NARH_ECOLI]	81.74545	39.26	13	13	20
P0AG67	30S ribosomal protein S1 OS=Escherichia coli (strain K12) GN=rpsA PE=1 SV=1 - [RS1_ECOLI]	81.3339	38.6	14	14	19

C4ZYU4	Protein RecA OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=recA PE=3 SV=1 - [RECA_ECOBW]	81.25392	46.18	9	9	25
P0AAB6	UTP--glucose-1-phosphate uridylyltransferase OS=Escherichia coli (strain K12) GN=galF PE=1 SV=1 - [GALF_ECOLI]	81.14549	57.91	12	12	21
P52647	Probable pyruvate-flavodoxin oxidoreductase OS=Escherichia coli (strain K12) GN=ydbK PE=3 SV=2 - [NIFJ_ECOLI]	80.33979	22.57	17	17	21
P0AFK0	Metalloprotease PmbA OS=Escherichia coli (strain K12) GN=pmbA PE=1 SV=1 - [PMBA_ECOLI]	79.71022	44.67	14	14	22
C4ZWB5	Lipoyl synthase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=lipA PE=3 SV=1 - [LIPA_ECOBW]	78.95297	50.47	12	12	19
P00452	Ribonucleoside-diphosphate reductase 1 subunit alpha OS=Escherichia coli (strain K12) GN=nrdA PE=1 SV=2 - [RIR1_ECOLI]	77.84976	28.52	16	16	25
P08839	Phosphoenolpyruvate-protein phosphotransferase OS=Escherichia coli (strain K12) GN=ptsI PE=1 SV=1 - [PT1_ECOLI]	77.70262	38.78	14	14	18
P37177	Phosphoenolpyruvate-dependent phosphotransferase system OS=Escherichia coli (strain K12) GN=ptsP PE=1 SV=2 - [PT1P_ECOLI]	77.27352	33.42	18	18	21
P16456	Selenide, water dikinase OS=Escherichia coli (strain K12) GN=selD PE=1 SV=1 - [SELD_ECOLI]	76.02012	54.18	12	12	20

P23538	Phosphoenolpyruvate synthase OS=Escherichia coli (strain K12) GN=ppsA PE=1 SV=5 - [PPSA_ECOLI]	73.40331	23.61	16	16	21
P0ABU2	Ribosome-binding ATPase YchF OS=Escherichia coli (strain K12) GN=ychF PE=1 SV=2 - [YCHF_ECOLI]	73.04729	46.01	11	11	19
C4ZRS7	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=accA PE=3 SV=1 - [ACCA_ECOBW]	70.82464	47.65	16	16	21
P0AEP3	UTP--glucose-1-phosphate uridylyltransferase OS=Escherichia coli (strain K12) GN=galU PE=1 SV=2 - [GALU_ECOLI]	69.7752	35.43	5	5	15
B1XAY6	GTPase Der OS=Escherichia coli (strain K12 / DH10B) GN=der PE=3 SV=1 - [DER_ECODH]	67.49218	43.06	14	14	19
B1X9T7	Tryptophanase OS=Escherichia coli (strain K12 / DH10B) GN=tnaA PE=3 SV=1 - [TNAA_ECODH]	66.40315	45.44	15	15	19
P0A8N5	Lysine--tRNA ligase, heat inducible OS=Escherichia coli (strain K12) GN=lysU PE=1 SV=2 - [SYK2_ECOLI]	66.13599	38.61	12	15	25
P0ABH9	ATP-dependent Clp protease ATP-binding subunit ClpA OS=Escherichia coli (strain K12) GN=clpA PE=1 SV=1 - [CLPA_ECOLI]	65.40826	24.93	14	14	18
C4ZY85	Adenosine deaminase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=add PE=3 SV=1 - [ADD_ECOBW]	63.87362	64.86	13	13	18
P33602	NADH-quinone oxidoreductase subunit G OS=Escherichia coli (strain K12) GN=nuoG PE=1 SV=4 - [NUOG_ECOLI]	63.77475	13.66	9	9	14

P63284	Chaperone protein ClpB OS=Escherichia coli (strain K12) GN=clpB PE=1 SV=1 - [CLPB_ECOLI]	62.32342	20.54	12	12	16
B1X736	3-dehydroquinate synthase OS=Escherichia coli (strain K12 / DH10B) GN=aroB PE=3 SV=1 - [AROB_ECODH]	61.83392	39.23	9	9	18
P14081	Selenocysteine-specific elongation factor OS=Escherichia coli (strain K12) GN=selB PE=1 SV=3 - [SELB_ECOLI]	60.45265	29.8	13	13	16
P0ACE0	Hydrogenase-2 large chain OS=Escherichia coli (strain K12) GN=hybC PE=1 SV=2 - [MBHM_ECOLI]	58.65768	28.57	10	10	14
P0A7V3	30S ribosomal protein S3 OS=Escherichia coli (strain K12) GN=rpsC PE=1 SV=2 - [RS3_ECOLI]	57.28084	44.64	9	9	15
P0A9C9	Fructose-1,6-bisphosphatase 1 class 2 OS=Escherichia coli (strain K12) GN=glpX PE=1 SV=1 - [GLPX_ECOLI]	57.14899	36.61	10	10	16
C4ZYI9	GTPase Era OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=era PE=3 SV=1 - [ERA_ECOBW]	57.14589	41.86	9	9	15
P0ABH0	Cell division protein FtsA OS=Escherichia coli (strain K12) GN=ftsA PE=1 SV=1 - [FTSA_ECOLI]	56.78618	34.05	7	7	11
P0A9U3	Uncharacterized ABC transporter ATP-binding protein YbiT OS=Escherichia coli (strain K12) GN=ybiT PE=1 SV=1 - [YBIT_ECOLI]	56.17563	23.02	7	7	14
P00393	NADH dehydrogenase OS=Escherichia coli (strain K12) GN=ndh PE=1 SV=2 - [DHNA_ECOLI]	55.80994	35.71	11	11	16

P0A7Z4	DNA-directed RNA polymerase subunit alpha OS=Escherichia coli (strain K12) GN=rpoA PE=1 SV=1 - [RPOA_ECOLI]	55.08909	43.77	10	10	15
P76440	NAD-dependent dihydropyrimidine dehydrogenase subunit PreT OS=Escherichia coli (strain K12) GN=preT PE=1 SV=1 - [PRET_ECOLI]	53.50727	33.25	9	9	12
C4ZU48	Periplasmic nitrate reductase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=napA PE=3 SV=1 - [NAPA_ECOBW]	53.05487	20.17	11	11	14
C4ZZT2	Enolase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=eno PE=3 SV=1 - [ENO_ECOBW]	51.69733	34.49	10	10	18
P17169	Glutamine--fructose-6-phosphate aminotransferase [isomerizing] OS=Escherichia coli (strain K12) GN=glmS PE=1 SV=4 - [GLMS_ECOLI]	51.07871	30.21	12	12	14
P63389	Uncharacterized ABC transporter ATP-binding protein YheS OS=Escherichia coli (strain K12) GN=yheS PE=1 SV=1 - [YHES_ECOLI]	50.10505	23.08	13	13	15
B1X9Z4	ATP-dependent RNA helicase RhlB OS=Escherichia coli (strain K12 / DH10B) GN=rhlB PE=3 SV=1 - [RHLB_ECODH]	49.52261	30.64	9	9	13
P13009	Methionine synthase OS=Escherichia coli (strain K12) GN=metH PE=1 SV=5 - [METH_ECOLI]	49.1226	15.57	13	13	14

B1XD38	30S ribosomal protein S2 OS=Escherichia coli (strain K12 / DH10B) GN=rpsB PE=3 SV=1 - [RS2_ECODH]	48.41051	54.36	8	8	13
P0AG24	Bifunctional (p)ppGpp synthase/hydrolase SpoT OS=Escherichia coli (strain K12) GN=spoT PE=1 SV=1 - [SPOT_ECOLI]	47.71877	24.36	12	12	15
P0AF20	N-acetylglucosamine repressor OS=Escherichia coli (strain K12) GN=nagC PE=2 SV=1 - [NAGC_ECOLI]	47.57133	35.96	12	12	15
P0AB77	2-amino-3-ketobutyrate coenzyme A ligase OS=Escherichia coli (strain K12) GN=kbl PE=1 SV=1 - [KBL_ECOLI]	46.5321	25.38	10	10	12
P21599	Pyruvate kinase II OS=Escherichia coli (strain K12) GN=pykA PE=1 SV=3 - [KPYK2_ECOLI]	46.37231	32.92	9	9	11
P0AES4	DNA gyrase subunit A OS=Escherichia coli (strain K12) GN=gyrA PE=1 SV=1 - [GYRA_ECOLI]	46.25941	11.31	7	7	14
P0AD61	Pyruvate kinase I OS=Escherichia coli (strain K12) GN=pykF PE=1 SV=1 - [KPYK1_ECOLI]	46.1884	34.68	10	10	12
P27833	dTDP-4-amino-4,6-dideoxygalactose transaminase OS=Escherichia coli (strain K12) GN=wecE PE=1 SV=2 - [WECE_ECOLI]	46.03329	39.63	9	9	12
P06709	Bifunctional ligase/repressor BirA OS=Escherichia coli (strain K12) GN=birA PE=1 SV=1 - [BIRA_ECOLI]	44.6746	41.12	9	9	12
P0AEW9	1-phosphofructokinase OS=Escherichia coli (strain K12) GN=fruK PE=3 SV=1 - [K1PF_ECOLI]	44.20991	31.41	7	7	12

C5A1D5	60 kDa chaperonin OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=groL PE=1 SV=1 - [CH60_ECOBW]	43.64823	27.92	9	9	10
P36683	Aconitate hydratase B OS=Escherichia coli (strain K12) GN=acnB PE=1 SV=3 - [ACNB_ECOLI]	43.61881	13.64	9	9	15
C5A0R1	Soluble pyridine nucleotide transhydrogenase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=sthA PE=3 SV=1 - [STHA_ECOBW]	43.012	25.32	7	7	10
P0A799	Phosphoglycerate kinase OS=Escherichia coli (strain K12) GN=pgk PE=1 SV=2 - [PGK_ECOLI]	42.21397	36.18	9	9	12
P75829	Uncharacterized protein YbjX OS=Escherichia coli (strain K12) GN=ybjX PE=4 SV=1 - [YBJX_ECOLI]	42.07149	37.27	10	10	11
P00954	Tryptophan--tRNA ligase OS=Escherichia coli (strain K12) GN=trpS PE=1 SV=3 - [SYW_ECOLI]	41.62681	41.92	10	10	12
B1XF08	1-deoxy-D-xylulose-5-phosphate synthase OS=Escherichia coli (strain K12 / DH10B) GN=dxs PE=3 SV=1 - [DXS_ECODH]	41.56366	25.97	12	12	14
P06710	DNA polymerase III subunit tau OS=Escherichia coli (strain K12) GN=dnaX PE=1 SV=1 - [DPO3X_ECOLI]	41.44125	23.95	12	12	13
P0AFI2	DNA topoisomerase 4 subunit A OS=Escherichia coli (strain K12) GN=parC PE=1 SV=1 - [PARC_ECOLI]	40.8568	22.21	11	11	13
C4ZTJ3	Trigger factor OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=tig PE=1 SV=1 - [TIG_ECOBW]	40.8054	28.94	10	10	14

P07395	Phenylalanine--tRNA ligase beta subunit OS=Escherichia coli (strain K12) GN=pheT PE=1 SV=2 - [SYFB_ECOLI]	40.74784	21.01	9	9	11
P21513	Ribonuclease E OS=Escherichia coli (strain K12) GN=rne PE=1 SV=6 - [RNE_ECOLI]	40.32119	10.56	8	8	10
P0A993	Fructose-1,6-bisphosphatase class 1 OS=Escherichia coli (strain K12) GN=fbp PE=1 SV=1 - [F16PA_ECOLI]	39.22681	31.02	7	7	13
P09546	Bifunctional protein PutA OS=Escherichia coli (strain K12) GN=putA PE=1 SV=3 - [PUTA_ECOLI]	39.07376	12.65	11	11	12
P06993	HTH-type transcriptional regulator MalT OS=Escherichia coli (strain K12) GN=malT PE=1 SV=2 - [MALT_ECOLI]	38.70962	13.87	8	8	10
C4ZZN7	DNA mismatch repair protein MutS OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=mutS PE=3 SV=1 - [MUTS_ECOBW]	38.65205	15.01	9	9	9
P0A9M8	Phosphate acetyltransferase OS=Escherichia coli (strain K12) GN=pta PE=1 SV=2 - [PTA_ECOLI]	38.46428	21.15	10	10	11
P0A6K6	Phosphopentomutase OS=Escherichia coli (strain K12) GN=deoB PE=1 SV=1 - [DEOB_ECOLI]	38.26211	32.92	9	9	11
P0A9B6	D-erythrose-4-phosphate dehydrogenase OS=Escherichia coli (strain K12) GN=epd PE=1 SV=2 - [E4PD_ECOLI]	37.29955	34.81	9	9	12
P0ACB0	Replicative DNA helicase OS=Escherichia coli (strain K12) GN=dnaB PE=1 SV=1 - [DNAB_ECOLI]	37.25136	26.75	8	8	10

P0AB89	Adenylosuccinate lyase OS=Escherichia coli (strain K12) GN=purB PE=1 SV=1 - [PUR8_ECOLI]	37.19054	27.19	10	10	13
B1X9R9	Glutamate--tRNA ligase OS=Escherichia coli (strain K12 / DH10B) GN=gltX PE=3 SV=1 - [SYE_ECODH]	37.13408	27.6	9	9	11
B1X8M6	Chromosome partition protein MukF OS=Escherichia coli (strain K12 / DH10B) GN=mukF PE=3 SV=1 - [MUKF_ECODH]	36.93561	25.91	8	8	9
P09980	ATP-dependent DNA helicase Rep OS=Escherichia coli (strain K12) GN=rep PE=1 SV=3 - [REP_ECOLI]	36.76304	18.42	9	9	10
C4ZX89	Histidine--tRNA ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=hisS PE=3 SV=1 - [SYH_ECOBW]	36.33873	26.65	10	10	11
P0A879	Tryptophan synthase beta chain OS=Escherichia coli (strain K12) GN=trpB PE=1 SV=2 - [TRPB_ECOLI]	36.00715	26.95	6	6	11
P00959	Methionine--tRNA ligase OS=Escherichia coli (strain K12) GN=metG PE=1 SV=2 - [SYM_ECOLI]	35.3687	18.76	11	11	13
P67910	ADP-L-glycero-D-manno-heptose-6-epimerase OS=Escherichia coli (strain K12) GN=hldD PE=1 SV=1 - [HLDD_ECOLI]	34.89655	36.45	10	10	11
P75838	Ribosomal protein S12 methylthiotransferase accessory factor YcaO OS=Escherichia coli (strain K12) GN=ycaO PE=1 SV=2 - [YCAO_ECOLI]	34.85133	16.21	7	7	9
P08200	Isocitrate dehydrogenase [NADP] OS=Escherichia coli (strain K12) GN=icd PE=1 SV=1 - [IDH_ECOLI]	34.42677	23.8	7	7	10

P0A9B2	Glyceraldehyde-3-phosphate dehydrogenase A OS=Escherichia coli (strain K12) GN=gapA PE=1 SV=2 - [G3P1_ECOLI]	34.41343	23.56	6	6	10
P0AGG8	Metalloprotease TldD OS=Escherichia coli (strain K12) GN=tldD PE=1 SV=1 - [TLDD_ECOLI]	33.80753	20.79	7	7	10
B1XCM5	Alanine--tRNA ligase OS=Escherichia coli (strain K12 / DH10B) GN=alaS PE=3 SV=1 - [SYA_ECODH]	33.22897	15.98	9	9	10
P25889	NAD-dependent dihydropyrimidine dehydrogenase subunit PreA OS=Escherichia coli (strain K12) GN=preA PE=1 SV=3 - [PREA_ECOLI]	33.16756	27.49	8	8	14
P0ADG4	Inositol-1-monophosphatase OS=Escherichia coli (strain K12) GN=suhB PE=1 SV=1 - [SUHB_ECOLI]	32.81856	34.83	6	6	9
P0A959	Glutamate-pyruvate aminotransferase AlaA OS=Escherichia coli (strain K12) GN=alaA PE=1 SV=1 - [ALAA_ECOLI]	32.56038	22.47	7	7	10
C5A059	Protein FdhE OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=fdhE PE=3 SV=1 - [FDHE_ECOBW]	32.12168	26.21	7	7	9
P0A9P0	Dihydrolipoyl dehydrogenase OS=Escherichia coli (strain K12) GN=lpdA PE=1 SV=2 - [DLDH_ECOLI]	31.30335	21.31	5	5	7
P0AFF6	Transcription termination/antitermination protein NusA OS=Escherichia coli (strain K12) GN=nusA PE=1 SV=1 - [NUSA_ECOLI]	31.30192	28.28	7	7	7
P25534	2-octaprenyl-6-methoxyphenol hydroxylase OS=Escherichia coli (strain K12) GN=ubiH PE=1 SV=2 - [UBIH_ECOLI]	31.12849	36.22	8	8	8

P14407	Fumarate hydratase class I, anaerobic OS=Escherichia coli (strain K12) GN=fumB PE=1 SV=2 - [FUMB_ECOLI]	30.91753	20.07	3	8	8
P0AB91	Phospho-2-dehydro-3-deoxyheptonate aldolase, Phe-sensitive OS=Escherichia coli (strain K12) GN=aroG PE=1 SV=1 - [AROG_ECOLI]	30.30827	26	6	6	7
P08957	Type I restriction enzyme EcoKI M protein OS=Escherichia coli (strain K12) GN=hsdM PE=1 SV=1 - [T1MK_ECOLI]	29.94277	17.96	6	6	8
P37759	dTDP-glucose 4,6-dehydratase 1 OS=Escherichia coli (strain K12) GN=rfbB PE=3 SV=2 - [RMLB1_ECOLI]	29.8041	33.52	7	7	8
P15038	DNA helicase IV OS=Escherichia coli (strain K12) GN=helD PE=1 SV=2 - [HELD_ECOLI]	29.48468	14.77	6	6	7
P37744	Glucose-1-phosphate thymidylyltransferase 1 OS=Escherichia coli (strain K12) GN=rfbA PE=1 SV=2 - [RMLA1_ECOLI]	29.28786	31.4	8	8	10
P36938	Phosphoglucomutase OS=Escherichia coli (strain K12) GN=pgm PE=1 SV=1 - [PGM_ECOLI]	29.15889	17.58	6	6	8
P06996	Outer membrane protein C OS=Escherichia coli (strain K12) GN=ompC PE=1 SV=1 - [OMPC_ECOLI]	28.78824	28.07	8	8	10
P23830	CDP-diacylglycerol--serine O-phosphatidyltransferase OS=Escherichia coli (strain K12) GN=pssA PE=1 SV=2 - [PSS_ECOLI]	28.48962	25.06	6	6	6
P0ADF6	Phosphogluconate dehydratase OS=Escherichia coli (strain K12) GN=edd PE=2 SV=1 - [EDD_ECOLI]	28.42597	13.43	5	5	7

P05852	tRNA N6-adenosine threonylcarbamoyltransferase OS=Escherichia coli (strain K12) GN=tsaD PE=1 SV=2 - [TSAD_ECOLI]	28.02917	18.69	5	5	9
C4ZTG3	Queuine tRNA-ribosyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=tgt PE=3 SV=1 - [TGT_ECOBW]	28.01065	29.6	8	8	8
P0ADI4	Enterobactin synthase component B OS=Escherichia coli (strain K12) GN=entB PE=1 SV=1 - [ENTB_ECOLI]	27.90043	32.63	7	7	8
P46889	DNA translocase FtsK OS=Escherichia coli (strain K12) GN=ftsK PE=1 SV=2 - [FTSK_ECOLI]	27.33956	8.13	8	8	9
P0AG20	GTP pyrophosphokinase OS=Escherichia coli (strain K12) GN=relA PE=1 SV=1 - [RELA_ECOLI]	27.22874	9.27	4	4	7
P69910	Glutamate decarboxylase beta OS=Escherichia coli (strain K12) GN=gadB PE=1 SV=1 - [DCEB_ECOLI]	27.19629	20.39	1	7	8
P37760	dTDP-4-dehydrorhamnose reductase OS=Escherichia coli (strain K12) GN=rfbD PE=3 SV=2 - [RMLD_ECOLI]	27.07514	36.79	4	4	4
P0A7B1	Polyphosphate kinase OS=Escherichia coli (strain K12) GN=ppk PE=1 SV=2 - [PPK_ECOLI]	26.65815	15.41	9	9	9
P06959	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex OS=Escherichia coli (strain K12) GN=aceF PE=1 SV=3 - [ODP2_ECOLI]	26.17033	15.08	7	7	8
P69908	Glutamate decarboxylase alpha OS=Escherichia coli (strain K12) GN=gadA PE=1 SV=1 - [DCEA_ECOLI]	26.12576	18.88	1	7	8

P0A9S3	Galactitol 1-phosphate 5-dehydrogenase OS=Escherichia coli (strain K12) GN=gatD PE=1 SV=1 - [GATD_ECOLI]	26.05096	24.86	8	8	8
P24203	Uncharacterized GTP-binding protein YjiA OS=Escherichia coli (strain K12) GN=yjiA PE=1 SV=3 - [YJIA_ECOLI]	26.03752	15.41	4	4	7
C5A0C2	Phosphoenolpyruvate carboxylase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=ppc PE=3 SV=1 - [CAPP_ECOBW]	25.86696	9.97	6	6	7
P0A6Y8	Chaperone protein DnaK OS=Escherichia coli (strain K12) GN=dnaK PE=1 SV=2 - [DNAK_ECOLI]	25.74091	13.79	7	7	8
P03004	Chromosomal replication initiator protein DnaA OS=Escherichia coli (strain K12) GN=dnaA PE=1 SV=2 - [DNAA_ECOLI]	25.51401	19.27	6	6	7
P0A996	Anaerobic glycerol-3-phosphate dehydrogenase subunit C OS=Escherichia coli (strain K12) GN=glpC PE=1 SV=1 - [GLPC_ECOLI]	25.33152	26.77	8	8	9
P0A8M0	Asparagine--tRNA ligase OS=Escherichia coli (strain K12) GN=asnS PE=1 SV=2 - [SYN_ECOLI]	25.0729	15.45	5	5	7
P36928	Uncharacterized chaperone protein YegD OS=Escherichia coli (strain K12) GN=yegD PE=3 SV=2 - [YEGD_ECOLI]	25.01055	19.33	4	4	5
P0ACP1	Catabolite repressor/activator OS=Escherichia coli (strain K12) GN=cra PE=1 SV=1 - [CRA_ECOLI]	24.73777	21.86	5	5	7
P0A698	UvrABC system protein A OS=Escherichia coli (strain K12) GN=uvrA PE=1 SV=1 - [UVRA_ECOLI]	24.67319	9.68	8	8	8

P42632	PFL-like enzyme TdcE OS=Escherichia coli (strain K12) GN=tdcE PE=1 SV=2 - [TDCE_ECOLI]	24.349	11.39	5	7	9
P0A717	Ribose-phosphate pyrophosphokinase OS=Escherichia coli (strain K12) GN=prs PE=1 SV=2 - [KPRS_ECOLI]	24.18507	22.86	5	5	7
P04825	Aminopeptidase N OS=Escherichia coli (strain K12) GN=pepN PE=1 SV=2 - [AMPN_ECOLI]	24.17459	7.7	5	5	7
P14175	Glycine betaine/proline betaine transport system ATP-binding protein ProV OS=Escherichia coli (strain K12) GN=proV PE=1 SV=1 - [PROV_ECOLI]	24.15436	18	5	5	7
B1XDY5	Glutamate 5-kinase OS=Escherichia coli (strain K12 / DH10B) GN=proB PE=3 SV=1 - [PROB_ECODH]	24.04472	17.71	4	4	6
P28629	Biodegradative arginine decarboxylase OS=Escherichia coli (strain K12) GN=adiA PE=1 SV=1 - [ADIA_ECOLI]	23.81658	9.93	5	5	6
B1XGY3	Argininosuccinate synthase OS=Escherichia coli (strain K12 / DH10B) GN=argG PE=3 SV=1 - [ASSY_ECODH]	23.75124	19.69	7	7	9
P25539	Riboflavin biosynthesis protein RibD OS=Escherichia coli (strain K12) GN=ribD PE=1 SV=1 - [RIBD_ECOLI]	23.70704	13.35	3	3	5
P0AC33	Fumarate hydratase class I, aerobic OS=Escherichia coli (strain K12) GN=fumA PE=1 SV=2 - [FUMA_ECOLI]	23.66546	13.87	1	6	6
P03018	DNA helicase II OS=Escherichia coli (strain K12) GN=uvrD PE=1 SV=1 - [UVRD_ECOLI]	23.63937	11.53	6	6	6

P22106	Asparagine synthetase B [glutamine-hydrolyzing] OS=Escherichia coli (strain K12) GN=asnB PE=1 SV=3 - [ASNB_ECOLI]	23.38762	22.74	7	7	7
P0AEZ3	Septum site-determining protein MinD OS=Escherichia coli (strain K12) GN=minD PE=1 SV=2 - [MIND_ECOLI]	23.33584	21.85	5	5	7
P16431	Formate hydrogenlyase subunit 5 OS=Escherichia coli (strain K12) GN=hycE PE=1 SV=1 - [HYCE_ECOLI]	23.2741	13.88	5	5	6
B1XC53	4-hydroxythreonine-4-phosphate dehydrogenase OS=Escherichia coli (strain K12 / DH10B) GN=pdxA PE=3 SV=1 - [PDXA_ECODH]	23.24101	24.62	4	4	5
P39384	Putative dehydratase subunit YjiM OS=Escherichia coli (strain K12) GN=yjiM PE=3 SV=2 - [YJIM_ECOLI]	23.20026	15.67	5	5	6
B1XAZ2	Dual-specificity RNA methyltransferase RlmN OS=Escherichia coli (strain K12 / DH10B) GN=rlmN PE=3 SV=1 - [RLMN_ECODH]	22.97251	21.88	7	7	8
C4ZZ12	ATP synthase subunit alpha OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=atpA PE=3 SV=1 - [ATPA_ECOBW]	22.85562	18.52	7	7	8
P77536	Uncharacterized electron transport protein YkgF OS=Escherichia coli (strain K12) GN=ykgF PE=3 SV=1 - [YKGF_ECOLI]	22.7198	16.84	6	6	8
P37342	Uncharacterized protein YjjI OS=Escherichia coli (strain K12) GN=yjjI PE=4 SV=2 - [YJJI_ECOLI]	22.64631	14.73	5	5	6

B1XHI0	RNase adapter protein RapZ OS=Escherichia coli (strain K12 / DH10B) GN=rapZ PE=3 SV=1 - [RAPZ_ECODH]	22.18911	27.82	4	4	5
P0AEW6	Inosine-guanosine kinase OS=Escherichia coli (strain K12) GN=gsk PE=1 SV=1 - [INGK_ECOLI]	22.12861	22.12	5	5	5
B1XA82	DNA ligase OS=Escherichia coli (strain K12 / DH10B) GN=ligA PE=3 SV=1 - [DNLJ_ECODH]	21.52805	12.97	6	6	6
P0ADQ2	Uncharacterized protein YiiD OS=Escherichia coli (strain K12) GN=yiiD PE=4 SV=1 - [YIID_ECOLI]	21.48308	13.68	4	4	7
P0ABQ0	Coenzyme A biosynthesis bifunctional protein CoaBC OS=Escherichia coli (strain K12) GN=coaBC PE=1 SV=2 - [COABC_ECOLI]	21.40392	22.66	4	4	4
P77434	Glutamate-pyruvate aminotransferase AlaC OS=Escherichia coli (strain K12) GN=alaC PE=1 SV=1 - [ALAC_ECOLI]	21.24875	18.69	5	5	6
P0ACD8	Hydrogenase-1 large chain OS=Escherichia coli (strain K12) GN=hyaB PE=1 SV=1 - [MBHL_ECOLI]	21.0065	12.56	6	6	7
P30178	Hydroxycarboxylate dehydrogenase B OS=Escherichia coli (strain K12) GN=hcxB PE=1 SV=2 - [HCXB_ECOLI]	20.74386	12.74	2	2	4
C4ZSQ5	Polyribonucleotide nucleotidyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=pnp PE=3 SV=1 - [PNP_ECOBW]	20.21219	11.25	5	5	6
P0A8N3	Lysine--tRNA ligase OS=Escherichia coli (strain K12) GN=lysS PE=1 SV=2 - [SYK1_ECOLI]	20.00573	16.63	5	8	11

P37440	Oxidoreductase UcpA OS=Escherichia coli (strain K12) GN=ucpA PE=3 SV=3 - [UCPA_ECOLI]	19.90497	27.38	6	6	6
P07604	Transcriptional regulatory protein TyrR OS=Escherichia coli (strain K12) GN=tyrR PE=1 SV=2 - [TYRR_ECOLI]	19.65897	14.62	6	6	8
P09053	Valine--pyruvate aminotransferase OS=Escherichia coli (strain K12) GN=avtA PE=1 SV=3 - [AVTA_ECOLI]	19.62035	12.71	4	4	6
B1X9T5	tRNA modification GTPase MnmE OS=Escherichia coli (strain K12 / DH10B) GN=mnmE PE=3 SV=1 - [MNME_ECODH]	19.49439	18.72	6	6	6
P75864	Ribosomal RNA large subunit methyltransferase K/L OS=Escherichia coli (strain K12) GN=rlmL PE=1 SV=1 - [RLMKL_ECOLI]	19.41368	9.83	5	5	6
P24193	Hydrogenase isoenzymes formation protein HypE OS=Escherichia coli (strain K12) GN=hypE PE=1 SV=2 - [HYPE_ECOLI]	19.24475	19.05	4	4	5
P19319	Respiratory nitrate reductase 2 alpha chain OS=Escherichia coli (strain K12) GN=narZ PE=1 SV=5 - [NARZ_ECOLI]	19.16234	4.65	1	6	7
P04968	L-threonine dehydratase biosynthetic IlvA OS=Escherichia coli (strain K12) GN=ilvA PE=1 SV=1 - [ILVA_ECOLI]	19.11269	11.67	4	4	6
P24192	Hydrogenase isoenzymes formation protein HypD OS=Escherichia coli (strain K12) GN=hypD PE=1 SV=1 - [HYPD_ECOLI]	19.0295	13.4	4	4	5

C4ZS67	Peptidase T OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=pepT PE=3 SV=1 - [PEPT_ECOBW]	18.91257	17.89	4	4	4
P08660	Lysine-sensitive aspartokinase 3 OS=Escherichia coli (strain K12) GN=lysC PE=1 SV=2 - [AK3_ECOLI]	18.46995	11.8	3	3	4
P21829	Pyridoxal phosphate phosphatase YbhA OS=Escherichia coli (strain K12) GN=ybhA PE=1 SV=3 - [YBHA_ECOLI]	18.32168	28.31	6	6	6
C4ZWK2	Succinate--CoA ligase [ADP-forming] subunit beta OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=sucC PE=3 SV=1 - [SUCC_ECOBW]	18.19181	11.34	4	4	5
C4ZZ22	Aspartate--ammonia ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=asnA PE=3 SV=1 - [ASNA_ECOBW]	17.94015	21.21	5	5	6
P75824	NADH oxidoreductase HCR OS=Escherichia coli (strain K12) GN=hcr PE=1 SV=3 - [HCR_ECOLI]	17.68022	21.74	4	4	4
P37747	UDP-galactopyranose mutase OS=Escherichia coli (strain K12) GN=glf PE=1 SV=1 - [GLF_ECOLI]	17.32557	13.35	4	4	5
P0A6R0	3-oxoacyl-[acyl-carrier-protein] synthase 3 OS=Escherichia coli (strain K12) GN=fabH PE=1 SV=1 - [FABH_ECOLI]	17.06909	20.5	4	4	4
P0AFB8	Nitrogen regulation protein NR(I) OS=Escherichia coli (strain K12) GN=glnG PE=1 SV=1 - [NTRC_ECOLI]	16.65948	10.45	5	5	5
P76027	Oligopeptide transport ATP-binding protein OppD OS=Escherichia coli (strain K12) GN=oppD PE=3 SV=2 - [OPPD_ECOLI]	16.42166	22.55	5	5	5

P24554	DNA repair protein RadA OS=Escherichia coli (strain K12) GN=radA PE=1 SV=1 - [RADA_ECOLI]	16.38051	12.39	4	4	5
P21693	ATP-dependent RNA helicase DbpA OS=Escherichia coli (strain K12) GN=dbpA PE=1 SV=2 - [DBPA_ECOLI]	16.27059	14.88	4	4	4
P0A9D8	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase OS=Escherichia coli (strain K12) GN=dapD PE=1 SV=1 - [DAPD_ECOLI]	16.2394	16.06	3	3	4
P0ABD8	Biotin carboxyl carrier protein of acetyl-CoA carboxylase OS=Escherichia coli (strain K12) GN=accB PE=1 SV=1 - [BCCP_ECOLI]	16.0299	55.77	4	4	6
P0A9H7	Cyclopropane-fatty-acyl-phospholipid synthase OS=Escherichia coli (strain K12) GN=cfa PE=1 SV=2 - [CFA_ECOLI]	15.71766	13.87	4	4	5
C4ZWF6	Glutamine--tRNA ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=glnS PE=3 SV=1 - [SYQ_ECOBW]	15.68488	11.91	3	3	3
C4ZPV2	Isoleucine--tRNA ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=ileS PE=3 SV=1 - [SYI_ECOBW]	15.61592	5.97	6	6	6
P75728	2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol hydroxylase OS=Escherichia coli (strain K12) GN=ubiF PE=1 SV=1 - [UBIF_ECOLI]	15.41631	17.65	4	4	4
P08312	Phenylalanine--tRNA ligase alpha subunit OS=Escherichia coli (strain K12) GN=pheS PE=1 SV=2 - [SYFA_ECOLI]	15.41439	16.82	4	4	4

P00968	Carbamoyl-phosphate synthase large chain OS=Escherichia coli (strain K12) GN=carB PE=1 SV=2 - [CARB_ECOLI]	15.33612	6.62	4	4	4
P37749	Beta-1,6-galactofuranosyltransferase WbbI OS=Escherichia coli (strain K12) GN=wbbI PE=1 SV=1 - [WBBI_ECOLI]	15.31452	13.94	4	4	5
P0ADG7	Inosine-5'-monophosphate dehydrogenase OS=Escherichia coli (strain K12) GN=guaB PE=1 SV=1 - [IMDH_ECOLI]	15.19171	12.91	3	3	4
C4ZU73	Ubiquinone biosynthesis O-methyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=ubiG PE=3 SV=1 - [UBIG_ECOBW]	15.14121	17.5	2	2	3
P13035	Aerobic glycerol-3-phosphate dehydrogenase OS=Escherichia coli (strain K12) GN=glpD PE=1 SV=3 - [GLPD_ECOLI]	15.12227	16.37	6	6	6
B1XEU8	S-formylglutathione hydrolase FrmB OS=Escherichia coli (strain K12 / DH10B) GN=frmB PE=3 SV=1 - [SFGH1_ECODH]	15.09175	25.99	4	4	4
P62623	4-hydroxy-3-methylbut-2-enyl diphosphate reductase OS=Escherichia coli (strain K12) GN=ispH PE=1 SV=1 - [ISPH_ECOLI]	13.96222	18.04	4	4	4
P37773	UDP-N-acetylmuramate--L-alanyl-gamma-D-glutamyl-meso-2,6-diaminoheptandioate ligase OS=Escherichia coli (strain K12) GN=mpl PE=1 SV=3 - [MPL_ECOLI]	13.75597	8.53	2	2	3
P37349	Protein-lysine deacetylase OS=Escherichia coli (strain K12) GN=dhaM PE=1 SV=3 - [DHAM_ECOLI]	13.69625	13.77	4	4	4

P37689	2,3-bisphosphoglycerate-independent phosphoglycerate mutase OS=Escherichia coli (strain K12) GN=gpmI PE=1 SV=1 - [GPMI_ECOLI]	13.63611	11.67	3	3	3
P43329	ATP-dependent RNA helicase HrpA OS=Escherichia coli (strain K12) GN=hrpA PE=3 SV=3 - [HRPA_ECOLI]	13.31288	3.38	3	3	4
C4ZR48	tRNA dimethylallyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=miaA PE=3 SV=1 - [MIAA_ECOBW]	13.18106	19.94	4	4	4
P17115	Arabinose 5-phosphate isomerase GutQ OS=Escherichia coli (strain K12) GN=gutQ PE=1 SV=3 - [GUTQ_ECOLI]	13.05749	8.72	2	2	3
C4ZXM7	Formamidopyrimidine-DNA glycosylase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=mutM PE=3 SV=1 - [FPG_ECOBW]	12.98989	20.82	3	3	3
P0ADR8	Pyrimidine/purine nucleotide 5'-monophosphate nucleosidase OS=Escherichia coli (strain K12) GN=ppnN PE=1 SV=1 - [PPNN_ECOLI]	12.92481	12.56	4	4	4
B1XAJ3	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase OS=Escherichia coli (strain K12 / DH10B) GN=metE PE=3 SV=1 - [METE_ECODH]	12.84075	6.24	4	4	4
P08192	Dihydrofolate synthase/folylpolyglutamate synthase OS=Escherichia coli (strain K12) GN=folC PE=1 SV=2 - [FOLC_ECOLI]	12.76897	14.69	3	3	3

P0AGJ9	Tyrosine--tRNA ligase OS=Escherichia coli (strain K12) GN=tyrS PE=1 SV=2 - [SYY_ECOLI]	12.7097	12.74	3	3	3
P07813	Leucine--tRNA ligase OS=Escherichia coli (strain K12) GN=leuS PE=1 SV=2 - [SYL_ECOLI]	12.66971	4.65	3	3	4
P25535	2-octaprenylphenol hydroxylase OS=Escherichia coli (strain K12) GN=ubiI PE=1 SV=2 - [UBII_ECOLI]	12.61301	11.75	4	4	4
P43672	ABC transporter ATP-binding protein uup OS=Escherichia coli (strain K12) GN=uup PE=1 SV=2 - [UUP_ECOLI]	12.44291	8.19	3	3	3
P05719	Type-1 restriction enzyme EcoKI specificity protein OS=Escherichia coli (strain K12) GN=hsdS PE=1 SV=1 - [T1SK_ECOLI]	12.17018	11.64	3	3	3
P0A887	Ubiquinone/menaquinone biosynthesis C-methyltransferase UbiE OS=Escherichia coli (strain K12) GN=ubiE PE=1 SV=1 - [UBIE_ECOLI]	12.14107	17.93	3	3	3
P0A6L2	4-hydroxy-tetrahydrodipicolinate synthase OS=Escherichia coli (strain K12) GN=dapA PE=1 SV=1 - [DAPA_ECOLI]	11.99384	9.93	2	2	3
P00582	DNA polymerase I OS=Escherichia coli (strain K12) GN=polA PE=1 SV=1 - [DPO1_ECOLI]	11.71793	6.25	4	4	4
P76403	Uncharacterized protease YegQ OS=Escherichia coli (strain K12) GN=yegQ PE=3 SV=1 - [YEGQ_ECOLI]	11.63457	12.14	3	3	3
P0AC38	Aspartate ammonia-lyase OS=Escherichia coli (strain K12) GN=aspA PE=1 SV=1 - [ASPA_ECOLI]	11.45289	10.67	4	4	4

P00579	RNA polymerase sigma factor RpoD OS=Escherichia coli (strain K12) GN=rpoD PE=1 SV=2 - [RPOD_ECOLI]	11.42437	7.67	4	4	4
P07118	Valine--tRNA ligase OS=Escherichia coli (strain K12) GN=valS PE=1 SV=2 - [SYV_ECOLI]	11.19588	4.73	2	2	3
B1X9G8	UPF0176 protein YceA OS=Escherichia coli (strain K12 / DH10B) GN=yceA PE=3 SV=1 - [YCEA_ECODH]	11.17677	12.86	3	3	3
P0AG76	Nuclease SbcCD subunit D OS=Escherichia coli (strain K12) GN=sbcD PE=1 SV=1 - [SBCD_ECOLI]	11.1439	8	2	2	3
P68187	Maltose/maltodextrin import ATP-binding protein MalK OS=Escherichia coli (strain K12) GN=malK PE=1 SV=1 - [MALK_ECOLI]	11.01948	13.21	3	3	3
P0A6N8	Elongation factor P-like protein OS=Escherichia coli (strain K12) GN=yeiP PE=1 SV=1 - [EFPL_ECOLI]	11.00548	24.74	3	3	3
P04693	Aromatic-amino-acid aminotransferase OS=Escherichia coli (strain K12) GN=tyrB PE=1 SV=1 - [TYRB_ECOLI]	10.99619	10.33	3	3	3
P0A6Z3	Chaperone protein HtpG OS=Escherichia coli (strain K12) GN=htpG PE=1 SV=1 - [HTPG_ECOLI]	10.97528	9.78	4	4	4
P10443	DNA polymerase III subunit alpha OS=Escherichia coli (strain K12) GN=dnaE PE=1 SV=1 - [DPO3A_ECOLI]	10.93326	4.74	5	5	6
P00888	Phospho-2-dehydro-3-deoxyheptonate aldolase, Tyr-sensitive OS=Escherichia coli (strain K12) GN=aroF PE=1 SV=1 - [AROF_ECOLI]	10.91229	13.2	2	2	2

P52129	mRNA endoribonuclease LS OS=Escherichia coli (strain K12) GN=rnlA PE=1 SV=2 - [RNLA_ECOLI]	10.90753	14.01	4	4	4
P07658	Formate dehydrogenase H OS=Escherichia coli (strain K12) GN=fdhF PE=1 SV=2 - [FDHF_ECOLI]	10.70124	7.97	3	3	3
P00350	6-phosphogluconate dehydrogenase, decarboxylating OS=Escherichia coli (strain K12) GN=gnd PE=1 SV=2 - [6PGD_ECOLI]	10.59953	9.62	3	3	3
C4ZUX7	Bifunctional protein FolD OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=folD PE=3 SV=1 - [FOLD_ECOBW]	10.3777	7.64	1	1	2
P10121	Signal recognition particle receptor FtsY OS=Escherichia coli (strain K12) GN=ftsY PE=1 SV=1 - [FTSY_ECOLI]	10.34819	9.66	3	3	3
C4ZXK0	Glycerol-3-phosphate dehydrogenase [NAD(P)+] OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=gpsA PE=3 SV=1 - [GPDA_ECOBW]	10.22539	9.44	3	3	4
P25437	S-(hydroxymethyl)glutathione dehydrogenase OS=Escherichia coli (strain K12) GN=frmA PE=1 SV=3 - [FRMA_ECOLI]	10.13527	10.84	3	3	4
P0A9N4	Pyruvate formate-lyase 1-activating enzyme OS=Escherichia coli (strain K12) GN=pflA PE=1 SV=2 - [PFLA_ECOLI]	10.11999	17.48	2	2	3
B1X7N2	Cytidine deaminase OS=Escherichia coli (strain K12 / DH10B) GN=cdd PE=3 SV=1 - [CDD_ECODH]	10.06876	12.24	2	2	2

P25888	ATP-dependent RNA helicase RhlE OS=Escherichia coli (strain K12) GN=rhlE PE=1 SV=3 - [RHLE_ECOLI]	10.01375	7.93	2	2	2
P67087	Ribosomal RNA small subunit methyltransferase I OS=Escherichia coli (strain K12) GN=rsmI PE=1 SV=1 - [RSMI_ECOLI]	9.822261	15.03	2	2	2
C4ZSR5	Phosphoglucosamine mutase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=glmM PE=3 SV=1 - [GLMM_ECOBW]	9.815786	6.29	2	2	3
P15005	5-methylcytosine-specific restriction enzyme B OS=Escherichia coli (strain K12) GN=mcrB PE=1 SV=3 - [MCRB_ECOLI]	9.783869	8.28	3	3	3
B1X9W9	tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG OS=Escherichia coli (strain K12 / DH10B) GN=mnmG PE=3 SV=1 - [MNMG_ECODH]	9.294013	3.82	2	2	3
P30744	L-serine dehydratase 2 OS=Escherichia coli (strain K12) GN=sdaB PE=2 SV=2 - [SDHM_ECOLI]	9.265748	6.37	3	3	4
P27829	UDP-N-acetyl-D-mannosamine dehydrogenase OS=Escherichia coli (strain K12) GN=wecC PE=1 SV=4 - [WECC_ECOLI]	9.101173	8.81	3	3	3
Q46802	Uncharacterized sigma-54-dependent transcriptional regulator YgeV OS=Escherichia coli (strain K12) GN=ygeV PE=4 SV=1 - [YGEV_ECOLI]	9.044973	5.91	2	2	2

P0AAI3	ATP-dependent zinc metalloprotease FtsH OS=Escherichia coli (strain K12) GN=ftsH PE=1 SV=1 - [FTSH_ECOLI]	9.025153	5.9	2	2	2
P0A6T3	Galactokinase OS=Escherichia coli (strain K12) GN=galK PE=1 SV=2 - [GAL1_ECOLI]	9.005999	9.16	2	2	2
C4ZXI2	L-seryl-tRNA(Sec) selenium transferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=selA PE=3 SV=1 - [SELA_ECOBW]	8.868298	7.99	2	2	2
P00490	Maltodextrin phosphorylase OS=Escherichia coli (strain K12) GN=malP PE=1 SV=7 - [PHSM_ECOLI]	8.846161	2.89	2	2	3
P31060	Putative molybdenum transport ATP-binding protein ModF OS=Escherichia coli (strain K12) GN=modF PE=3 SV=2 - [MODF_ECOLI]	8.774251	9.39	3	3	3
P77596	Uncharacterized protein YagF OS=Escherichia coli (strain K12) GN=yagF PE=3 SV=1 - [YAGF_ECOLI]	8.628483	4.73	2	2	3
C5A134	Glycerol-3-phosphate acyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=plsB PE=3 SV=1 - [PLSB_ECOBW]	8.470447	4.46	3	3	3
B1XFJ2	Thymidine phosphorylase OS=Escherichia coli (strain K12 / DH10B) GN=deoA PE=3 SV=1 - [TYPH_ECODH]	8.427342	3.18	1	1	2
P77783	Probable dimethyl sulfoxide reductase chain YnfF OS=Escherichia coli (strain K12) GN=ynfF PE=1 SV=4 - [YNFF_ECOLI]	8.394686	3.72	2	2	2

C4ZYG8	UPF0061 protein YdiU OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=ydiU PE=3 SV=1 - [YDIU_ECOBW]	8.285548	7.74	2	2	2
B1XAK9	Xaa-Pro dipeptidase OS=Escherichia coli (strain K12 / DH10B) GN=pepQ PE=3 SV=1 - [PEPQ_ECODH]	8.199325	4.97	2	2	3
P22259	Phosphoenolpyruvate carboxykinase (ATP) OS=Escherichia coli (strain K12) GN=pckA PE=1 SV=2 - [PCKA_ECOLI]	8.145742	6.67	2	2	2
C4ZSA8	ATP phosphoribosyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=hisG PE=3 SV=1 - [HIS1_ECOBW]	8.113433	13.04	2	2	2
P30750	Methionine import ATP-binding protein MetN OS=Escherichia coli (strain K12) GN=metN PE=1 SV=4 - [METN_ECOLI]	8.09397	13.99	2	2	2
P76558	NADP-dependent malic enzyme OS=Escherichia coli (strain K12) GN=maeB PE=1 SV=1 - [MAO2_ECOLI]	8.068489	4.08	2	2	2
P04079	GMP synthase [glutamine-hydrolyzing] OS=Escherichia coli (strain K12) GN=guaA PE=1 SV=1 - [GUAA_ECOLI]	7.884047	8	3	3	3
P08390	USG-1 protein OS=Escherichia coli (strain K12) GN=usg PE=1 SV=1 - [USG_ECOLI]	7.830189	14.54	2	2	2
P45527	Uncharacterized protease YhbU OS=Escherichia coli (strain K12) GN=yhbU PE=1 SV=1 - [YHBU_ECOLI]	7.822476	13.6	3	3	3
P0ADY7	50S ribosomal protein L16 OS=Escherichia coli (strain K12) GN=rplP PE=1 SV=1 - [RL16_ECOLI]	7.80015	22.06	2	2	2

C4ZZ73	Diaminopimelate epimerase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=dapF PE=3 SV=1 - [DAPF_ECOBW]	7.79664	10.22	2	2	2
P06987	Histidine biosynthesis bifunctional protein HisB OS=Escherichia coli (strain K12) GN=hisB PE=1 SV=1 - [HIS7_ECOLI]	7.714962	12.68	3	3	3
P00887	Phospho-2-dehydro-3-deoxyheptonate aldolase, Trp-sensitive OS=Escherichia coli (strain K12) GN=aroH PE=1 SV=4 - [AROH_ECOLI]	7.549433	8.05	2	2	2
P0AF08	Iron-sulfur cluster carrier protein OS=Escherichia coli (strain K12) GN=mrp PE=3 SV=1 - [APBC_ECOLI]	7.387477	8.94	2	2	2
P21499	Ribonuclease R OS=Escherichia coli (strain K12) GN=rnr PE=1 SV=2 - [RNR_ECOLI]	7.366722	4.06	2	2	2
C4ZZ41	Dihydroxy-acid dehydratase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=ilvD PE=3 SV=1 - [ILVD_ECOBW]	7.365398	5.03	2	2	2
P00936	Adenylate cyclase OS=Escherichia coli (strain K12) GN=cyaA PE=1 SV=5 - [CYAA_ECOLI]	7.351512	3.77	2	2	2
P0ABT5	tRNA-dihydrouridine synthase B OS=Escherichia coli (strain K12) GN=dusB PE=3 SV=1 - [DUSB_ECOLI]	7.273063	9.97	2	2	2
C4ZTG2	S-adenosylmethionine:tRNA ribosyltransferase-isomerase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=queA PE=3 SV=1 - [QUEA_ECOBW]	7.232445	8.71	2	2	2

B1XFI4	Peptide chain release factor 3 OS=Escherichia coli (strain K12 / DH10B) GN=prfC PE=3 SV=1 - [RF3_ECODH]	7.22995	6.81	3	3	3
P0A6E9	ATP-dependent dethiobiotin synthetase BioD 2 OS=Escherichia coli (strain K12) GN=bioD2 PE=3 SV=1 - [BIOD2_ECOLI]	7.10843	12.55	2	2	2
B1XB82	ATP-dependent 6-phosphofructokinase isozyme 1 OS=Escherichia coli (strain K12 / DH10B) GN=pfkA PE=3 SV=1 - [PFKA_ECODH]	7.027476	8.44	3	3	3
P64624	Uncharacterized protein YheO OS=Escherichia coli (strain K12) GN=yheO PE=4 SV=1 - [YHEO_ECOLI]	6.974549	11.25	1	1	1
P0AFB5	Nitrogen regulation protein NR(II) OS=Escherichia coli (strain K12) GN=glnL PE=1 SV=1 - [NTRB_ECOLI]	6.949288	8.31	2	2	2
P30958	Transcription-repair-coupling factor OS=Escherichia coli (strain K12) GN=mfd PE=1 SV=2 - [MFD_ECOLI]	6.931969	2.7	3	3	3
C4ZRQ9	Bifunctional uridylyltransferase/uridylyl-removing enzyme OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=glnD PE=3 SV=1 - [GLND_ECOBW]	6.910604	4.49	3	3	3
P42630	L-serine dehydratase TdcG OS=Escherichia coli (strain K12) GN=tdcG PE=1 SV=4 - [TDCG_ECOLI]	6.865132	5.95	2	2	2
P39356	Uncharacterized transcriptional regulator YjhU OS=Escherichia coli (strain K12) GN=yjhU PE=3 SV=2 - [YJHU_ECOLI]	6.568667	7.93	2	2	2

P07023	T-protein OS=Escherichia coli (strain K12) GN=tyrA PE=1 SV=1 - [TYRA_ECOLI]	6.494095	7.24	2	2	2
B1XDL5	Ribosomal RNA large subunit methyltransferase M OS=Escherichia coli (strain K12 / DH10B) GN=rlmM PE=3 SV=1 - [RLMM_ECODH]	6.483147	6.83	2	2	2
P39410	Putative kinase YjjJ OS=Escherichia coli (strain K12) GN=yjjJ PE=3 SV=1 - [YJJJ_ECOLI]	6.447779	6.09	1	1	1
P0AAZ4	Replication-associated recombination protein A OS=Escherichia coli (strain K12) GN=rarA PE=1 SV=1 - [RARA_ECOLI]	6.39294	8.28	3	3	3
P0AC41	Succinate dehydrogenase flavoprotein subunit OS=Escherichia coli (strain K12) GN=sdhA PE=1 SV=1 - [SDHA_ECOLI]	6.253986	4.93	3	3	3
P0AE12	AMP nucleosidase OS=Escherichia coli (strain K12) GN=amn PE=1 SV=1 - [AMN_ECOLI]	6.246315	3.51	1	1	2
P76104	Uncharacterized protease YdcP OS=Escherichia coli (strain K12) GN=ydcP PE=1 SV=2 - [YDCP_ECOLI]	6.172837	4.13	2	2	2
P0A9D4	Serine acetyltransferase OS=Escherichia coli (strain K12) GN=cysE PE=1 SV=1 - [CYSE_ECOLI]	5.991639	8.79	2	2	2
B1XG88	Ribosomal RNA large subunit methyltransferase G OS=Escherichia coli (strain K12 / DH10B) GN=rlmG PE=3 SV=1 - [RLMG_ECODH]	5.916891	8.73	2	2	2

P0A9K9	FKBP-type peptidyl-prolyl cis-trans isomerase SlyD OS=Escherichia coli (strain K12) GN=slyD PE=1 SV=1 - [SLYD_ECOLI]	5.727189	18.37	1	1	1
P52062	Oxygen-independent coproporphyrinogen-III oxidase-like protein YggW OS=Escherichia coli (strain K12) GN=yggW PE=3 SV=1 - [YGGW_ECOLI]	5.717974	5.82	2	2	2
B1XBE0	Chaperone protein DnaJ OS=Escherichia coli (strain K12 / DH10B) GN=dnaJ PE=3 SV=1 - [DNAJ_ECODH]	5.563894	5.85	2	2	2
P21179	Catalase HPII OS=Escherichia coli (strain K12) GN=katE PE=1 SV=1 - [CATE_ECOLI]	5.547596	2.79	2	2	2
C4ZUV3	tRNA 2-selenouridine synthase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=selU PE=3 SV=1 - [SELU_ECOBW]	5.441372	6.59	1	1	1
P33919	Putative DNA repair helicase RadD OS=Escherichia coli (strain K12) GN=radD PE=1 SV=2 - [RADD_ECOLI]	5.322029	3.92	1	1	1
B1XA17	Beta-hexosaminidase OS=Escherichia coli (strain K12 / DH10B) GN=nagZ PE=3 SV=1 - [NAGZ_ECODH]	5.213741	8.8	2	2	2
B1XGC4	Cysteine--tRNA ligase OS=Escherichia coli (strain K12 / DH10B) GN=cysS PE=3 SV=1 - [SYC_ECODH]	5.185703	3.69	2	2	2
B1XA76	D-amino acid dehydrogenase OS=Escherichia coli (strain K12 / DH10B) GN=dadA PE=3 SV=1 - [DADA_ECODH]	5.063919	5.56	2	2	2

P24183	Formate dehydrogenase, nitrate-inducible, major subunit OS=Escherichia coli (strain K12) GN=fdnG PE=1 SV=3 - [FDNG_ECOLI]	5.046585	2.66	2	2	2
C4ZVX9	Glycogen synthase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=glgA PE=3 SV=1 - [GLGA_ECOBW]	4.87044	4.82	2	2	2
C4ZS59	N-acetyl-D-glucosamine kinase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=nagK PE=3 SV=1 - [NAGK_ECOBW]	4.854305	5.61	1	1	1
P27242	Lipopolysaccharide 1,2-N-acetylglucosaminetransferase OS=Escherichia coli (strain K12) GN=waaU PE=4 SV=2 - [WAAU_ECOLI]	4.75137	6.72	2	2	2
P69797	PTS system mannose-specific EIIAB component OS=Escherichia coli (strain K12) GN=manX PE=1 SV=2 - [PTNAB_ECOLI]	4.685439	6.5	1	1	1
B1XEK1	HTH-type transcriptional regulator ArgP OS=Escherichia coli (strain K12 / DH10B) GN=argP PE=3 SV=1 - [ARGP_ECODH]	4.68185	9.09	2	2	2
P0AA16	Transcriptional regulatory protein OmpR OS=Escherichia coli (strain K12) GN=ompR PE=1 SV=1 - [OMPR_ECOLI]	4.640294	10.04	2	2	2
P0AB71	Fructose-bisphosphate aldolase class 2 OS=Escherichia coli (strain K12) GN=fbaA PE=1 SV=2 - [ALF_ECOLI]	4.625154	7.52	1	1	1
C4ZUF1	30S ribosomal protein S4 OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=rpsD PE=3 SV=1 - [RS4_ECOBW]	4.612835	8.74	1	1	1

P77737	Oligopeptide transport ATP-binding protein OppF OS=Escherichia coli (strain K12) GN=oppF PE=3 SV=1 - [OPPF_ECOLI]	4.598031	5.09	2	2	2
P46118	HTH-type transcriptional regulator HexR OS=Escherichia coli (strain K12) GN=hexR PE=3 SV=2 - [HEXR_ECOLI]	4.588016	11.07	3	3	3
P0AGJ5	Uncharacterized tRNA/rRNA methyltransferase YfiF OS=Escherichia coli (strain K12) GN=yfiF PE=3 SV=1 - [YFIF_ECOLI]	4.547614	4.93	1	1	1
P33030	Zinc-binding GTPase YeiR OS=Escherichia coli (strain K12) GN=yeiR PE=1 SV=2 - [YEIR_ECOLI]	4.513877	6.4	2	2	2
P09832	Glutamate synthase [NADPH] small chain OS=Escherichia coli (strain K12) GN=gltD PE=1 SV=3 - [GLTD_ECOLI]	4.403302	4.03	1	1	1
P0ACP5	HTH-type transcriptional regulator GntR OS=Escherichia coli (strain K12) GN=gntR PE=1 SV=1 - [GNTR_ECOLI]	4.301704	4.83	1	1	1
B1XBX8	Glutamate racemase OS=Escherichia coli (strain K12 / DH10B) GN=murI PE=3 SV=1 - [MURI_ECODH]	4.266018	6.67	1	1	1
P07003	Pyruvate dehydrogenase [ubiquinone] OS=Escherichia coli (strain K12) GN=poxB PE=1 SV=1 - [POXB_ECOLI]	4.246956	3.32	2	2	2
P0AAN3	Hydrogenase isoenzymes nickel incorporation protein HypB OS=Escherichia coli (strain K12) GN=hypB PE=1 SV=1 - [HYPB_ECOLI]	4.221004	7.59	1	1	1
P33136	Glucans biosynthesis protein G OS=Escherichia coli (strain K12) GN=mdoG PE=1 SV=1 - [OPGG_ECOLI]	4.220096	3.13	1	1	1

P0A9J0	Ribonuclease G OS=Escherichia coli (strain K12) GN=rng PE=1 SV=2 - [RNG_ECOLI]	4.173019	3.27	1	1	1
P21893	Single-stranded-DNA-specific exonuclease RecJ OS=Escherichia coli (strain K12) GN=recJ PE=1 SV=2 - [RECJ_ECOLI]	4.162608	2.25	1	1	1
P0A988	DNA polymerase III subunit beta OS=Escherichia coli (strain K12) GN=dnaN PE=1 SV=1 - [DPO3B_ECOLI]	4.155298	4.1	1	1	1
P77739	Putative kinase YniA OS=Escherichia coli (strain K12) GN=yniA PE=3 SV=1 - [YNIA_ECOLI]	4.141877	5.94	1	1	1
B1XBN7	Exoribonuclease 2 OS=Escherichia coli (strain K12 / DH10B) GN=rnb PE=3 SV=1 - [RNB_ECODH]	4.113503	2.64	1	1	1
P28631	DNA polymerase III subunit delta' OS=Escherichia coli (strain K12) GN=holB PE=1 SV=2 - [HOLB_ECOLI]	4.105003	4.19	1	1	1
C4ZRR2	Elongation factor Ts OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=tsf PE=3 SV=1 - [EFTS_ECOBW]	4.066476	6.36	1	1	1
C4ZVY0	Glucose-1-phosphate adenylyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=glgC PE=3 SV=1 - [GLGC_ECOBW]	4.059067	3.25	1	1	1
C5A0G3	Peptide chain release factor 2 OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=prfB PE=3 SV=1 - [RF2_ECOBW]	4.013648	3.84	1	1	1
P37051	Formyltetrahydrofolate deformylase OS=Escherichia coli (strain K12) GN=purU PE=1 SV=1 - [PURU_ECOLI]	3.931808	5.36	1	1	1

P75863	Uncharacterized protein YcbX OS=Escherichia coli (strain K12) GN=ycbX PE=1 SV=1 - [YCBX_ECOLI]	3.921844	4.88	1	1	1
P28904	Trehalose-6-phosphate hydrolase OS=Escherichia coli (strain K12) GN=treC PE=1 SV=3 - [TREC_ECOLI]	3.846764	2.9	1	1	1
P77580	Acetaldehyde dehydrogenase OS=Escherichia coli (strain K12) GN=mhpF PE=1 SV=1 - [ACDH_ECOLI]	3.846711	7.59	1	1	1
B1X658	Trehalose-6-phosphate synthase OS=Escherichia coli (strain K12 / DH10B) GN=otsA PE=3 SV=1 - [OTSA_ECODH]	3.769569	4.85	1	1	1
B1XFI0	Ribosomal RNA small subunit methyltransferase C OS=Escherichia coli (strain K12 / DH10B) GN=rsmC PE=3 SV=1 - [RSMC_ECODH]	3.700012	4.37	1	1	1
P0A7W1	30S ribosomal protein S5 OS=Escherichia coli (strain K12) GN=rpsE PE=1 SV=2 - [RS5_ECOLI]	3.691953	10.78	1	1	1
B1XC78	Cell division protein ZapD OS=Escherichia coli (strain K12 / DH10B) GN=zapD PE=3 SV=1 - [ZAPD_ECODH]	3.58328	5.67	1	1	1
P13029	Catalase-peroxidase OS=Escherichia coli (strain K12) GN=katG PE=1 SV=2 - [KATG_ECOLI]	3.564599	1.93	1	1	1
P0ABH7	Citrate synthase OS=Escherichia coli (strain K12) GN=gltA PE=1 SV=1 - [CISY_ECOLI]	3.485211	3.28	1	1	1
P18775	Dimethyl sulfoxide reductase DmsA OS=Escherichia coli (strain K12) GN=dmsA PE=1 SV=2 - [DMSA_ECOLI]	3.451255	1.6	1	1	1

C4ZWP8	NAD-dependent malic enzyme OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=maeA PE=3 SV=1 - [MAO1_ECOBW]	3.402134	2.12	1	1	1
B1XC12	Homoserine O-succinyltransferase OS=Escherichia coli (strain K12 / DH10B) GN=metAS PE=3 SV=1 - [METAS_ECODH]	3.40045	8.09	1	1	1
P76316	D-cysteine desulfhydrase OS=Escherichia coli (strain K12) GN=dcyD PE=1 SV=4 - [DCYD_ECOLI]	3.381505	3.96	1	1	1
B1X6F3	50S ribosomal protein L15 OS=Escherichia coli (strain K12 / DH10B) GN=rplO PE=3 SV=1 - [RL15_ECODH]	3.375041	9.72	1	1	1
B1X998	Adenine deaminase OS=Escherichia coli (strain K12 / DH10B) GN=ade PE=3 SV=1 - [ADEC_ECODH]	3.362915	2.72	1	1	1
P77689	FeS cluster assembly protein SufD OS=Escherichia coli (strain K12) GN=sufD PE=1 SV=1 - [SUFD_ECOLI]	3.357718	3.55	1	1	1
P69924	Ribonucleoside-diphosphate reductase 1 subunit beta OS=Escherichia coli (strain K12) GN=nrdB PE=1 SV=2 - [RIR2_ECOLI]	3.347764	3.46	1	1	1
P69741	Hydrogenase-2 small chain OS=Escherichia coli (strain K12) GN=hybO PE=1 SV=1 - [MBHT_ECOLI]	3.344275	3.76	1	1	1
P27126	Lipopolysaccharide core biosynthesis protein RfaS OS=Escherichia coli (strain K12) GN=rfaS PE=4 SV=1 - [RFAS_ECOLI]	3.338925	7.07	2	2	2

P46837	Protein YhgF OS=Escherichia coli (strain K12) GN=yhgF PE=1 SV=3 - [YHGF_ECOLI]	3.310799	1.81	1	1	1
P54901	Stationary phase-inducible protein CsiE OS=Escherichia coli (strain K12) GN=csiE PE=4 SV=3 - [CSIE_ECOLI]	3.300247	3.52	1	1	1
P0ADW6	Protein YhcC OS=Escherichia coli (strain K12) GN=yhcC PE=3 SV=1 - [YHCC_ECOLI]	3.288904	4.85	1	1	1
C4ZQG0	Arginine--tRNA ligase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=argS PE=3 SV=1 - [SYR_ECOBW]	3.283226	1.91	1	1	1
P20083	DNA topoisomerase 4 subunit B OS=Escherichia coli (strain K12) GN=parE PE=1 SV=3 - [PARE_ECOLI]	3.278103	3.33	1	1	1
P69874	Spermidine/putrescine import ATP-binding protein PotA OS=Escherichia coli (strain K12) GN=potA PE=1 SV=1 - [POTA_ECOLI]	3.268765	3.17	1	1	1
P39288	Epoxyqueuosine reductase OS=Escherichia coli (strain K12) GN=queG PE=3 SV=1 - [QUEG_ECOLI]	3.188155	3.43	1	1	1
P0AA39	Ribosomal large subunit pseudouridine synthase C OS=Escherichia coli (strain K12) GN=rluC PE=1 SV=1 - [RLUC_ECOLI]	3.178523	3.45	1	1	1
P0ACY3	Uncharacterized protein YeaG OS=Escherichia coli (strain K12) GN=yeaG PE=3 SV=1 - [YEAG_ECOLI]	3.174146	2.64	1	1	1
P77252	Uncharacterized protein YkgE OS=Escherichia coli (strain K12) GN=ykgE PE=3 SV=1 - [YKGE_ECOLI]	3.172971	4.6	1	1	1

P27298	Oligopeptidase A OS=Escherichia coli (strain K12) GN=prlC PE=3 SV=3 - [OPDA_ECOLI]	3.172968	2.94	1	1	1
P31979	NADH-quinone oxidoreductase subunit F OS=Escherichia coli (strain K12) GN=nuoF PE=1 SV=3 - [NUOF_ECOLI]	3.16346	2.92	1	1	1
P0AG16	Amidophosphoribosyltransferase OS=Escherichia coli (strain K12) GN=purF PE=1 SV=2 - [PUR1_ECOLI]	3.129692	2.18	1	1	1
C4ZQJ1	UvrABC system protein C OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=uvrC PE=3 SV=1 - [UVRC_ECOBW]	3.124326	2.79	1	1	1
P0A7S9	30S ribosomal protein S13 OS=Escherichia coli (strain K12) GN=rpsM PE=1 SV=2 - [RS13_ECOLI]	3.122386	15.25	1	1	1
P67662	HTH-type transcriptional activator AaeR OS=Escherichia coli (strain K12) GN=aaeR PE=4 SV=1 - [AAER_ECOLI]	3.102667	3.88	1	1	1
P22188	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase OS=Escherichia coli (strain K12) GN=murE PE=1 SV=2 - [MURE_ECOLI]	3.096303	3.23	1	1	1
P0AGG2	Acyl-CoA thioesterase 2 OS=Escherichia coli (strain K12) GN=tesB PE=1 SV=2 - [TESB_ECOLI]	3.068325	3.5	1	1	1
P0A9Q9	Aspartate-semialdehyde dehydrogenase OS=Escherichia coli (strain K12) GN=asd PE=1 SV=1 - [DHAS_ECOLI]	3.063127	4.36	1	1	1
P31131	Protein YdeJ OS=Escherichia coli (strain K12) GN=ydeJ PE=3 SV=1 - [YDEJ_ECOLI]	3.032011	13.95	1	1	1

P0ABK5	Cysteine synthase A OS=Escherichia coli (strain K12) GN=cysK PE=1 SV=2 - [CYSK_ECOLI]	3.01721	8.67	1	1	1
P0AG44	50S ribosomal protein L17 OS=Escherichia coli (strain K12) GN=rplQ PE=1 SV=1 - [RL17_ECOLI]	3.008997	7.87	1	1	1
P0AFL6	Exopolyphosphatase OS=Escherichia coli (strain K12) GN=ppx PE=1 SV=2 - [PPX_ECOLI]	2.97432	2.53	1	1	1
P39407	Uncharacterized protein YjjU OS=Escherichia coli (strain K12) GN=yjjU PE=3 SV=1 - [YJJU_ECOLI]	2.961048	2.8	1	1	1
P0AGG0	Thiamine-monophosphate kinase OS=Escherichia coli (strain K12) GN=thiL PE=1 SV=1 - [THIL_ECOLI]	2.927783	4.31	1	1	1
P23721	Phosphoserine aminotransferase OS=Escherichia coli (strain K12) GN=serC PE=1 SV=4 - [SERC_ECOLI]	2.924736	3.59	1	1	1
P19323	Formate hydrogenlyase transcriptional activator FhlA OS=Escherichia coli (strain K12) GN=fhlA PE=1 SV=1 - [FHLA_ECOLI]	2.892398	1.73	1	1	1
P0AEJ2	Isochorismate synthase EntC OS=Escherichia coli (strain K12) GN=entC PE=1 SV=1 - [ENTC_ECOLI]	2.881617	2.81	1	1	1
P76141	Transcriptional regulator LsrR OS=Escherichia coli (strain K12) GN=lsrR PE=1 SV=1 - [LSRR_ECOLI]	2.845747	4.42	1	1	1
P16433	Formate hydrogenlyase subunit 7 OS=Escherichia coli (strain K12) GN=hycG PE=1 SV=2 - [HYCG_ECOLI]	2.837107	8.63	1	1	1

C5A0W2	Glucose-6-phosphate isomerase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=pgi PE=3 SV=1 - [G6PI_ECOBW]	2.814593	2.91	1	1	1
P67660	Probable HTH-type transcriptional regulator YhaJ OS=Escherichia coli (strain K12) GN=yhaJ PE=3 SV=1 - [YHAJ_ECOLI]	2.774523	6.71	2	2	2
P76586	Uncharacterized protein YphH OS=Escherichia coli (strain K12) GN=yphH PE=3 SV=2 - [YPHH_ECOLI]	2.772981	2.77	1	1	1
P04993	RecBCD enzyme subunit RecD OS=Escherichia coli (strain K12) GN=recD PE=1 SV=2 - [RECD_ECOLI]	2.772114	1.64	1	1	1
P0A9K3	PhoH-like protein OS=Escherichia coli (strain K12) GN=ybeZ PE=3 SV=2 - [PHOL_ECOLI]	2.756644	3.47	1	1	1
P27250	Aldehyde reductase Ahr OS=Escherichia coli (strain K12) GN=ahr PE=1 SV=2 - [AHR_ECOLI]	2.729346	3.24	1	1	1
C5A0H7	Aminomethyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=gcvT PE=3 SV=1 - [GCST_ECOBW]	2.704652	2.47	1	1	1
B1X6G0	50S ribosomal protein L5 OS=Escherichia coli (strain K12 / DH10B) GN=rplE PE=3 SV=1 - [RL5_ECODH]	2.700862	6.7	1	1	1
P0AFG3	2-oxoglutarate dehydrogenase E1 component OS=Escherichia coli (strain K12) GN=sucA PE=1 SV=1 - [ODO1_ECOLI]	2.700094	0.96	1	1	1
P0AF24	Ribonucleotide monophosphatase NagD OS=Escherichia coli (strain K12) GN=nagD PE=1 SV=1 - [NAGD_ECOLI]	2.696054	4.4	1	1	1

P02359	30S ribosomal protein S7 OS=Escherichia coli (strain K12) GN=rpsG PE=1 SV=3 - [RS7_ECOLI]	2.692768	9.5	1	1	1
P37013	Anaerobic nitric oxide reductase transcription regulator NorR OS=Escherichia coli (strain K12) GN=norR PE=1 SV=2 - [NORR_ECOLI]	2.652758	1.98	1	1	1
B1X971	50S ribosomal protein L28 OS=Escherichia coli (strain K12 / DH10B) GN=rpmB PE=3 SV=1 - [RL28_ECODH]	2.648528	12.82	1	1	1
P09831	Glutamate synthase [NADPH] large chain OS=Escherichia coli (strain K12) GN=gltB PE=1 SV=3 - [GLTB_ECOLI]	2.639892	1.08	1	1	1
P77374	Putative dimethyl sulfoxide reductase chain YnfE OS=Escherichia coli (strain K12) GN=ynfE PE=1 SV=1 - [YNFE_ECOLI]	2.631776	1.49	1	1	1
P0ACQ4	Hydrogen peroxide-inducible genes activator OS=Escherichia coli (strain K12) GN=oxyR PE=1 SV=1 - [OXYR_ECOLI]	2.612756	3.93	1	1	1
P16095	L-serine dehydratase 1 OS=Escherichia coli (strain K12) GN=sdaA PE=1 SV=3 - [SDHL_ECOLI]	2.525115	6.39	2	2	2
P0AFU4	Transcriptional regulatory protein GlrR OS=Escherichia coli (strain K12) GN=glrR PE=1 SV=1 - [GLRR_ECOLI]	2.509808	2.93	1	1	1
P0A7X3	30S ribosomal protein S9 OS=Escherichia coli (strain K12) GN=rpsI PE=1 SV=2 - [RS9_ECOLI]	2.500647	9.23	1	1	1
P24175	Phosphomannomutase OS=Escherichia coli (strain K12) GN=manB PE=3 SV=1 - [MANB_ECOLI]	2.449327	2.63	1	1	1

C4ZQ44	3-deoxy-manno-octulosonate cytidylyltransferase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=kdsB PE=3 SV=1 - [KDSB_ECOBW]	2.415252	4.84	1	1	1
P0A870	Transaldolase B OS=Escherichia coli (strain K12) GN=talB PE=1 SV=2 - [TALB_ECOLI]	2.398519	2.84	1	1	1
P36879	Uncharacterized ABC transporter ATP-binding protein YadG OS=Escherichia coli (strain K12) GN=yadG PE=3 SV=1 - [YADG_ECOLI]	2.392689	4.22	1	1	1
C4ZXV3	GTP 3',8-cyclase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=moaA PE=3 SV=1 - [MOAA_ECOBW]	2.391444	6.08	2	2	2
P0A9G6	Isocitrate lyase OS=Escherichia coli (strain K12) GN=aceA PE=1 SV=1 - [ACEA_ECOLI]	2.3755	3.69	1	1	1
P0ACB2	Delta-aminolevulinic acid dehydratase OS=Escherichia coli (strain K12) GN=hemB PE=1 SV=2 - [HEM2_ECOLI]	2.363438	2.16	1	1	1
P20966	PTS system fructose-specific EIIB'BC component OS=Escherichia coli (strain K12) GN=fruA PE=1 SV=1 - [PTFBC_ECOLI]	2.356761	2.31	1	1	1
B1XHC8	Holliday junction ATP-dependent DNA helicase RuvB OS=Escherichia coli (strain K12 / DH10B) GN=ruvB PE=3 SV=1 - [RUVB_ECODH]	2.347992	3.57	1	1	1
P0AFQ7	Uncharacterized metal-dependent hydrolase YcfH OS=Escherichia coli (strain K12) GN=ycfH PE=1 SV=1 - [YCFH_ECOLI]	2.296493	3.77	1	1	1

P0ACL2	Exu regulon transcriptional regulator OS=Escherichia coli (strain K12) GN=exuR PE=4 SV=1 - [EXUR_ECOLI]	2.282765	3.49	1	1	1
P46890	Uncharacterized protein YbaE OS=Escherichia coli (strain K12) GN=ybaE PE=4 SV=2 - [YBAE_ECOLI]	2.274684	1.59	1	1	1
C5A1F5	Small ribosomal subunit biogenesis GTPase RsgA OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=rsgA PE=3 SV=1 - [RSGA_ECOBW]	2.267208	2.57	1	1	1
P0ABF1	Poly(A) polymerase I OS=Escherichia coli (strain K12) GN=pcnB PE=1 SV=2 - [PCNB_ECOLI]	2.264895	1.94	1	1	1
P75745	Uncharacterized protein YbgK OS=Escherichia coli (strain K12) GN=ybgK PE=1 SV=1 - [YBGK_ECOLI]	2.217952	2.58	1	1	1
P0AB80	Branched-chain-amino-acid aminotransferase OS=Escherichia coli (strain K12) GN=ilvE PE=1 SV=2 - [ILVE_ECOLI]	2.217476	3.56	1	1	1
P76187	Oxidoreductase YdhF OS=Escherichia coli (strain K12) GN=ydhF PE=1 SV=2 - [YDHF_ECOLI]	2.216756	4.03	1	1	1
P45524	Putative esterase YheT OS=Escherichia coli (strain K12) GN=yheT PE=3 SV=1 - [YHET_ECOLI]	2.207857	3.53	1	1	1
Q46861	UPF0313 protein YgiQ OS=Escherichia coli (strain K12) GN=ygiQ PE=3 SV=3 - [YGIQ_ECOLI]	2.200274	1.76	1	1	1
P06612	DNA topoisomerase 1 OS=Escherichia coli (strain K12) GN=topA PE=1 SV=2 - [TOP1_ECOLI]	2.183221	1.04	1	1	1

P0AC53	Glucose-6-phosphate 1-dehydrogenase OS=Escherichia coli (strain K12) GN=zwf PE=1 SV=1 - [G6PD_ECOLI]	2.132419	2.04	1	1	1
P0ACH5	Multiple antibiotic resistance protein MarA OS=Escherichia coli (strain K12) GN=marA PE=1 SV=2 - [MARA_ECOLI]	2.130593	5.51	1	1	1
P27128	Lipopolysaccharide 1,3-galactosyltransferase OS=Escherichia coli (strain K12) GN=rfaI PE=3 SV=1 - [RFAI_ECOLI]	2.128326	3.83	1	1	1
C4ZRJ1	UDP-3-O-acyl-N-acetylglucosamine deacetylase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=lpxC PE=3 SV=1 - [LPXC_ECOBW]	2.092627	2.62	1	1	1
B1XDY2	Esterase FrsA OS=Escherichia coli (strain K12 / DH10B) GN=frsA PE=3 SV=1 - [FRSA_ECODH]	2.043153	2.9	1	1	1
P45568	1-deoxy-D-xylulose 5-phosphate reductoisomerase OS=Escherichia coli (strain K12) GN=dxr PE=1 SV=2 - [DXR_ECOLI]	2.006543	2.76	1	1	1
C4ZZE0	UPF0229 protein YeaH OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=yeaH PE=3 SV=1 - [YEAH_ECOBW]	1.96724	2.11	1	1	1
P77473	Putative cyclic-di-GMP phosphodiesterase YlaB OS=Escherichia coli (strain K12) GN=ylaB PE=2 SV=2 - [YLAB_ECOLI]	1.951022	1.16	1	1	1
P60723	50S ribosomal protein L4 OS=Escherichia coli (strain K12) GN=rplD PE=1 SV=1 - [RL4_ECOLI]	1.93743	3.48	1	1	1

C4ZS02	Dihydroorotase OS=Escherichia coli (strain K12 / MC4100 / BW2952) GN=pyrC PE=3 SV=1 - [PYRC_ECOBW]	1.935337	2.87	1	1	1
Q06065	Acetoacetate metabolism regulatory protein AtoC OS=Escherichia coli (strain K12) GN=atoC PE=3 SV=2 - [ATOC_ECOLI]	1.904739	2.17	1	1	1
P37751	Putative glycosyltransferase WbbK OS=Escherichia coli (strain K12) GN=wbbK PE=4 SV=1 - [WBBK_ECOLI]	1.890528	2.15	1	1	1
P0ABS5	DNA primase OS=Escherichia coli (strain K12) GN=dnaG PE=1 SV=1 - [DNAG_ECOLI]	1.876528	2.24	1	1	1
P42912	Putative galactosamine-6-phosphate isomerase OS=Escherichia coli (strain K12) GN=agaI PE=3 SV=1 - [AGAI_ECOLI]	1.829043	4.38	1	1	1
P0A9T0	D-3-phosphoglycerate dehydrogenase OS=Escherichia coli (strain K12) GN=serA PE=1 SV=2 - [SERA_ECOLI]	1.809803	3.17	1	1	1
P09148	Galactose-1-phosphate uridylyltransferase OS=Escherichia coli (strain K12) GN=galT PE=1 SV=2 - [GAL7_ECOLI]	1.79137	2.3	1	1	1
P24180	Multidrug export protein AcrE OS=Escherichia coli (strain K12) GN=acrE PE=1 SV=2 - [ACRE_ECOLI]	1.752631	1.82	1	1	1
P0ACX5	Fumarase D OS=Escherichia coli (strain K12) GN=fumD PE=1 SV=1 - [FUMD_ECOLI]	1.751139	10.14	1	1	1
P0AG80	sn-glycerol-3-phosphate-binding periplasmic protein UgpB OS=Escherichia coli (strain K12) GN=ugpB PE=1 SV=1 - [UGPB_ECOLI]	1.694535	1.83	1	1	1

P76612	Uncharacterized protein YpjB OS=Escherichia coli (strain K12) GN=ypjB PE=4 SV=1 - [YPJB_ECOLI]	1.655606	2.28	1	1	1
P76594	Protein lysine acetyltransferase Pka OS=Escherichia coli (strain K12) GN=pka PE=1 SV=1 - [PKA_ECOLI]	1.631728	1.24	1	1	1
B1XAS5	UPF0225 protein YchJ OS=Escherichia coli (strain K12 / DH10B) GN=ychJ PE=3 SV=1 - [YCHJ_ECODH]	0	7.89	1	1	1
B1XEM6	Aspartate carbamoyltransferase OS=Escherichia coli (strain K12 / DH10B) GN=pyrB PE=3 SV=1 - [PYRB_ECODH]	0	1.93	1	1	1
P07330	Chemotaxis response regulator protein-glutamate methylesterase OS=Escherichia coli (strain K12) GN=cheB PE=1 SV=3 - [CHEB_ECOLI]	0	7.16	1	1	1
P0A7U3	30S ribosomal protein S19 OS=Escherichia coli (strain K12) GN=rpsS PE=1 SV=2 - [RS19_ECOLI]	0	10.87	1	1	1
P0ABP6	Protein DedA OS=Escherichia coli (strain K12) GN=dedA PE=1 SV=1 - [DEDA_ECOLI]	0	4.11	1	1	1
P0ADY1	Peptidyl-prolyl cis-trans isomerase D OS=Escherichia coli (strain K12) GN=ppiD PE=1 SV=1 - [PPID_ECOLI]	0	0.96	1	1	1
P0AE30	Arginine ABC transporter permease protein ArtM OS=Escherichia coli (strain K12) GN=artM PE=1 SV=1 - [ARTM_ECOLI]	0	2.7	1	1	1
P0AEM6	RNA polymerase sigma factor FliA OS=Escherichia coli (strain K12) GN=fliA PE=1 SV=1 - [FLIA_ECOLI]	0	5.44	1	1	1

P0AEX9	Maltose-binding periplasmic protein OS=Escherichia coli (strain K12) GN=malE PE=1 SV=1 - [MALE_ECOLI]	0	5.56	1	1	1
P0AFU2	Uncharacterized transporter YfbS OS=Escherichia coli (strain K12) GN=yfbS PE=3 SV=1 - [YFBS_ECOLI]	0	1.15	1	1	1
P0AGH8	Phosphate transport system permease protein PstC OS=Escherichia coli (strain K12) GN=pstC PE=1 SV=1 - [PSTC_ECOLI]	0	8.46	1	1	1
P0AGI8	Trk system potassium uptake protein TrkA OS=Escherichia coli (strain K12) GN=trkA PE=1 SV=1 - [TRKA_ECOLI]	0	1.75	1	1	1
P0CF91	Putative transposase InsL for insertion sequence element IS186A OS=Escherichia coli (strain K12) GN=insL1 PE=3 SV=1 - [INSL1_ECOLI]	0	2.43	1	1	1
P16926	Cell shape-determining protein MreC OS=Escherichia coli (strain K12) GN=mreC PE=1 SV=1 - [MREC_ECOLI]	0	1.91	1	1	1
P24215	Mannonate dehydratase OS=Escherichia coli (strain K12) GN=uxuA PE=1 SV=2 - [UXUA_ECOLI]	0	2.54	1	1	1
P33570	Transketolase 2 OS=Escherichia coli (strain K12) GN=tktB PE=1 SV=1 - [TKT2_ECOLI]	0	6.9	1	1	1
P42626	UPF0597 protein YhaM OS=Escherichia coli (strain K12) GN=yhaM PE=2 SV=3 - [YHAM_ECOLI]	0	2.06	1	1	1
P52101	Sensor histidine kinase GlrK OS=Escherichia coli (strain K12) GN=glrK PE=1 SV=3 - [GLRK_ECOLI]	0	2.11	1	1	1

P61517	Carbonic anhydrase 2 OS=Escherichia coli (strain K12) GN=can PE=1 SV=1 - [CAN_ECOLI]	0	5.91	1	1	1
P69786	PTS system glucose-specific EIICB component OS=Escherichia coli (strain K12) GN=ptsG PE=1 SV=1 - [PTGCB_ECOLI]	0	9.64	1	1	1
P76518	Acetyl-CoA:oxalate CoA-transferase OS=Escherichia coli (strain K12) GN=yfdE PE=1 SV=2 - [ACOCT_ECOLI]	0	7.87	1	1	1
P77174	Uncharacterized protein YbdM OS=Escherichia coli (strain K12) GN=ybdM PE=1 SV=1 - [YBDM_ECOLI]	0	2.87	1	1	1
P77463	Probable D,D-dipeptide transport system permease protein DdpC OS=Escherichia coli (strain K12) GN=ddpC PE=1 SV=1 - [DDPC_ECOLI]	0	2.01	1	1	1

List of publications

1. **Himanshu Sharma,** B. Anand (2016). Fluorescence Bimolecular Complementation Enables Facile Detection of Ribosome Assembly Defects in *Escherichia coli***.** *RNA Biology*, 13, 872-882.
2. **Himanshu Sharma**, B. Anand "Ribosome quality control in *E. coli* and its implication in translation". *(Under preparation)*

List of posters presented in conferences

1. **Himanshu Sharma*** and B. Anand (2017). "Investigating the role of ribosome assembly factors in translation initiation" **9th RNA Group Meet**, Banaras Hindu University, Varanasi, India, October 26-28th, 2017.
2. **Himanshu Sharma**, B Anand***** (2015). "Fluorescence Bimolecular Complementation enables facile detection of Ribosome Assembly Defects in *Escherichia coli*." **RNA 2015**, ***The Twentieth Annual Meeting of the RNA Society***, University of Wisconsin, Madison, USA. May 26–31, 2015.
3. **Himanshu Sharma*,** B Anand. (2015). "A Bimolecular fluorescence complementation system to detect ribosome assembly defects in *Escherichia coli*." **MCB75*: from Molecules to Organisms***, Indian Institute of Science, Bengaluru, India. December 11-14, 2015.
4. **Himanshu Sharma***, B Anand. (2014). "Structural basis for K^+(potassium) dependent hydrolysis and domain crosstalk in a highly conserved GTPase Era." **PSI 6th Annual Meeting:** ***Proteomics from Discovery to Function***, International proteomics conference. Indian Institute of Technology Bombay, Mumbai, India. December 6 - 11, 2014.

***** *Presenting author*