STUDYING THE IN-VIVO ASSEMBLY OF THE BACTERIAL RIBOSOME

ELUCIDATING THE ROLE OF THE YJEQ AND RBGA GTPASES
IN THE ASSEMBLY OF THE BACTERIAL RIBOSOME

By AHMAD JOMAA, B.SC.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (2013) McMaster University

(Biochemistry and Biomedical Sciences) Hamilton, Ontario

TITLE: Elucidating the Role of the YjeQ and RbgA GTPases in the Assembly of the Bacterial Ribosome

AUTHOR: Ahmad Jomaa, B.Sc. (McMaster University)

SUPERVISOR: Dr. Joaquin Ortega

NUMBER OF PAGES: xvii, 188

Abstract

Ribosome assembly is a complex process, facilitated by more than 20 protein factors in bacteria. GTPases and ATPases represent the energy driving force of these factors. In my research as a PhD student, I studied the function of two GTPases, YjeQ and RbgA, involved in the assembly of the small and the large ribosomal subunits, respectively.

We isolated and characterized *in-vivo* assembled immature small (30S) and large (50S) subunits using a perturbation in the genes coding for these proteins. We observed that both subunits contained an incomplete ribosomal protein content, mainly lacking late-binding r-proteins. Additionally, we observed distortions in the functional core of the immature ribosomal subunit, particularly in the mRNA decoding center of the 30S subunit, the peptidyltransferase center of the 50S subunit, and tRNA binding sites.

Additionally, we have determined that the YjeQ protein interacts with the 30S subunit through its N-terminal OB-fold domain, and C-terminal Zn-finger motif. The binding site of YjeQ on the 30S subunit prevents the interaction with tRNAs, translation factors, and the 50S subunit.

Finally, we uncovered a novel functional interplay between RbgA and the ribosomal protein L16 during late stages of ribosomal assembly. We proposed that recruitment of L16 to the assembling 50S subunit would induce a conformational rearrangement that would ultimately promote the GTP-dependent release of RbgA.

The function of the assembly factors associated with the process of *in-vivo* ribosome assembly is not known, and thus a framework on how ribosomes are built is still elusive. I believe the research presented in this thesis provides novel insights into the role of YjeQ and RbgA in the assembly of ribosomes.

Acknowledgements

I would like to acknowledge my supervisor, Dr. Joaquin Ortega, for all his support, guidance, and patience throughout the years. It has truly been an everyday learning journey for me and I appreciate all the time he invested in me inside and outside of the lab environment. I will really miss our long conversations while driving to conferences together.

I would like to thank my committee members, Dr. Eric Brown and Dr. Justin Nodwell for their guidance and valuable advice. I would like to thank them for all the challenging committee meetings and encouraging me always to learn. I would also like to thank Dr. Alba Guarné for providing suggestions and for sharing lab equipment.

I am pleased to acknowledge our collaborators: Dr. Eric Brown, Dr. Janine Maddock, Dr. Jaime Martín-Benito, Dr. Jason Mears, Dr. Robert Britton, and Dr. Jamie Williamson.

I would like to acknowledge the past and present members of the Ortega group. In particular, I would like to extend my thanks to Vivian Leong and John Alexopoulos for all the scientific discussions and providing advice.

I wish to express my deepest gratitude to my family members. I like to thank my mother and father for all their support and understanding throughout my years away for my education. Finally, I would like to thank my wife, Nicole. Truly, without her support and patience, I would not be here.

Table of Contents

List of Figures

Chapter 1.

Chapter 2.

Chapter 3.

Chapter 4.

Chapter 5.

List of Tables

Chapter 1.

List of Abbreviations

2D/3D	two/three dimensional
30S	small bacterial ribosomal subunit
50S	large bacterial ribosomal subunit
Å	Angstroms
A	Adenine
A-site	aminoacyl-site
A-tRNA	Aminoacyl-tRNA
ACL	anti-codon loop
ASF	A-site finger
ASL	accepter stem loop
B50S	Bacillus subtilis 50S subunit
C	Cytosine
C-terminus	carboxyl-terminus
CC	cross correlation
CCF	cross correlation function
CP	central protuberance
cryo-EM	cryo-electron microscopy
CTF	contrast transfer function
D	Aspartic acid
DC	decoding center
DMS	dimethyl sulfate
E50S	E. coli 50S subunit
E	Glutamic acid
E-site	exit-site

E-tRNA	exit-tRNA
EF	elongation factor
ESI-TOF	electro-spray ionization time-of-flight
fMet	formyl-methionine
G-domain	GTPase-domain
GAC	GTPase activating center
GDP	Guanosine diphosphate
GDPNP	5'-Guanylyl imidodiphosphate
GTP	Guanosine triphosphate
h/H	small or large subunit rRNA helix
H-domain	Helical-domain
IF	initiation factor
IPTG	isopropyl-beta-D-thiogalactopyranoside
iTRAQ	isobaric tag for relative and absolute quantitation
L	large ribosomal subunit protein
lp	Large precursor
MALDI	Matrix-assisted laser desorption/ionization
ML	maximum-likelihood
mRNA	messenger RNA
MS	mass spectrometry
N-terminus	amino-terminus
OB	oligonucleotide/oligosaccharide binding
P-site	peptidyl-site
P-tRNA	peptidyl-tRNA
Pre	precursor
PTC	peptidyltransferase center

QMS	quantitative mass spectrometry
r-protein	ribosomal-protein
RAAFS	ribosome assembly-associated factors
RF	release factor
RNA	ribonucleic acid
RRF	ribosome recycling factor
rRNA	ribosomal RNA
S	small ribosomal subunit protein
sp	short precursor
TOF	time-of-flight
TRAFAC	translation factor class
tRNA	transfer RNA
Zn-finger	Zinc-finger

Declaration of Academic Achievement

All the experiments in this thesis are designed and conducted by Dr. Joaquin Ortega and myself, unless stated otherwise in the thesis preface. Experiments in Chapter 2 and Chapter 3 were designed in collaboration with Dr. Eric Brown. Experiments in Chapter 4 were designed in collaboration with Dr. Rob Britton.

CHAPTER 1. General Introduction

Ribosomes are macromolecular assemblies responsible for protein synthesis in all organisms. In actively growing bacteria, ribosomes compose up to 30% of the total cell mass where a total of 40% of energy turnover is invested in the process of protein synthesis (Wilson & Nierhaus, 2007). A large fraction of this energy is employed in producing the ribosomal components and building ribosomes. For example, the cell has to synthesize and assemble more than 4500 ribosomal RNA (rRNA) nucleotides and over 50 ribosomal proteins (r-proteins). Consequently, ribosome biogenesis is considered to be one of the most demanding tasks in the cell, where several factors have evolved to tightly and efficiently regulate this process in response to growth rates and environmental conditions.

Trying to understand how the ribosome assembles was a central question soon after its discovery. Vintage work by Nomura and coworkers (Traub & Nomura, 1968a; Mizushima & Nomura, 1970) and Nierhaus and coworkers (Nierhaus & Dohme, 1974; Herold et al., 1986; Herold & Nierhaus, 1987) established that the ribosome assembles from purified components (r-proteins and rRNA). These findings were fascinating as they indicated that the information required to assemble the ribosome is encoded in the primary structure of rRNA. A caveat in this study was that the *in-vitro* reconstitution experiments of the ribosomal subunits were carried out under non-physiological conditions, particularly using high ionic concentrations, varying temperatures, and long incubations hours. In contrast, *in-vivo* ribosome assembly involves: (i) the synthesis,

processing, and modification of the rRNA; (ii) the synthesis, and modification of the r-proteins; as well as (iii) the binding of these proteins to the folding rRNA in the correct manner to ensure the formation of an active ribosome particle.

In fact, a plethora of non-ribosomal factors have been indentified to regulate ribosome assembly *in-vivo* (Wilson & Nierhaus, 2007; Connolly & Culver, 2009; Shajani et al., 2011). A precise role for these factors in ribosome assembly is still elusive, and thus a framework for ribosome biogenesis in the cell is still unknown. In this thesis, I utilized biochemical and structural approaches to study the role of ribosome-associated factors implicated in the process of assembly. I believe that the results in this thesis provide novel information on the contribution of these factors to the process of building the ribosome in cells. The ribosome is a paradigm for the processes of RNA-protein recognition, and RNA and protein folding. Hence, I believe the work presented in this thesis can also be utilized to understand how multi-component RNA-protein complexes are generally assembled.

1.1 Ribosome Structure and Function

The bacterial ribosome sediments by ultra-centrifugation as a 70S particle, and is composed of two subunits, the large (50S) and the small (30S) subunits (see Figure 1.1). The 50S subunit is composed of 2 rRNA molecules: 23S (~3000 nucleotides) and 5S rRNA (~100 nucleotides), as well as 31-34 r-proteins designated as L-proteins (L1-L36, L for large subunit; Figure 1.2). The 30S subunit is composed of one rRNA molecule

known as 16S (~1500 nucleotides), and 21 r-proteins designated as S-proteins (S1-S21, S for small subunit; Figure 1.2).

1.1.1 Historic Events in Solving the Structure of the Bacterial Ribosome

Efforts to determine the structure of the ribosome started in the 70s of the past century, when the laboratory of Harry Noller determined the primary structure of the 16S and 23S rRNA from *Escherichia coli* (Brosius et al., 1978; Brosius et al., 1980). Secondary structure models of the rRNA followed soon after, which brought the structure of the ribosome into two dimensions (Noller et al., 1981; Noller & Woese, 1981). The first three-dimensional (3D) images of the ribosomes however, were determined by electron microscopy (Lake, 1976) and produced the first physical models that defined a simple low-resolution surface topology of the two ribosomal subunits.

The laboratory of Ada Yonath obtained the first x-ray crystals of the ribosome in 1980 (Yonath et al., 1980), but it was not until 20 years later when the first high-resolution structure of the ribosome was solved utilizing this technique. Meanwhile, efforts in electron microscopy from several groups, aided with developments in the software used for image processing (Frank et al., 1996; van Heel et al., 1996), were successful in solving the structure of the bacterial ribosome at 25Å resolution. An increased level of complexity emerged at this point, and structural domains of the two subunits were visualized (Frank et al., 1995a; Frank et al., 1995b; Stark et al., 1995).

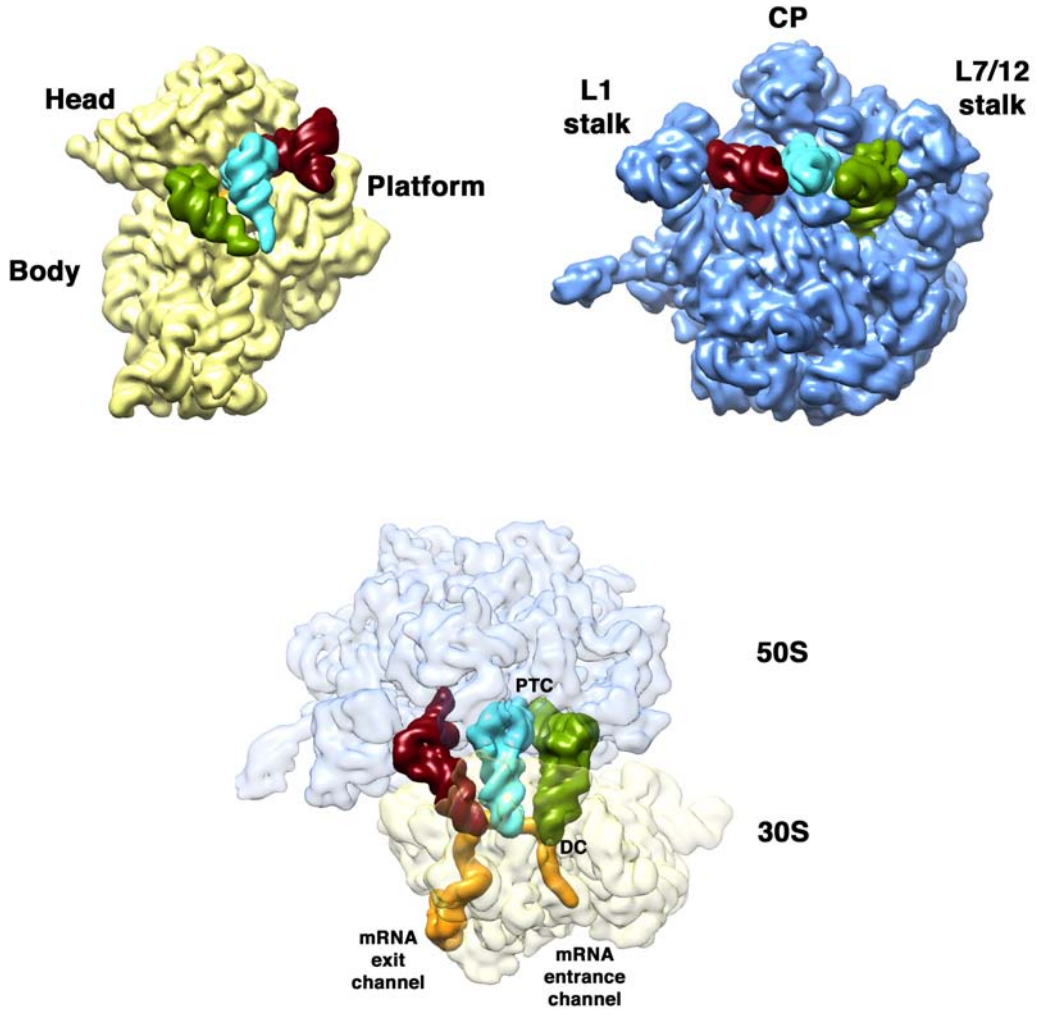


Figure 1.1 Structure of the bacterial ribosome

Structure of the bacterial ribosome in complex with A-site (green), P-site (cyan), and E-site (red) tRNAs; and mRNA (orange). Small (30S) subunit is colored in yellow and large (50S) subunit is colored in blue. Abbreviations: CP, central protuberance; DC, decoding center; and PTC, peptidyl transferase center. Pdb ID: 2HGP and 2HGQ.

Concurrent efforts in the field of x-ray crystallography were finally successful in producing the first complete atomic structures of the archeal 50S subunit from

Haloarcula marismortui (Ban et al., 2000), the bacterial 50S subunit from *Dinococcus radioduarans* (Harms et al., 2001), and the bacterial 30S subunit from *Thermus thermophilus* (Schluenzen et al., 2000; Wimberly et al., 2000). In addition, the complete crystal structure of the bacterial 70S ribosome was solved from *T. thermophilus* soon after (Yusupov et al., 2001). Since then, more than 50 structures of the ribosome in complex with tRNAs, mRNAs, translation factors and antibiotics have been solved, providing us with a wealth of information on the complex structure and function of this intricate cellular machinery.

1.1.2 Description of the Structure of the Ribosome

The bacterial ribosome is composed of two subunits, the small (30S) and the large (50S) subunits. The overall shape of the 30S subunit can be divided into 4 structural domains (see Figure 1.1 & 1.2), which also represent the classical features of this subunit and is largely based on the secondary structure of the 16S rRNA (Wimberly et al., 2000). The domains are as follows: the body (5' domain), the platform (central domain), the head (3' major domain), and h44 and h45 (3' minor domain). The nexus of these domains represents the neck region, which also contains the messenger RNA decoding center (DC; Figure 1.1), the functional core of the 30S subunit. The ribosomal proteins decorate the exterior of the 30S subunit, where the interface region is devoid of proteins except for the ribosomal protein S12 (see Figure 1.2), which also lies in close proximity of the 30S subunit functional core.

The structure of the 50S subunit is composed of a monolithic compact particle that cannot be divided structurally into separate domains (see Figure 1.1 & 1.2). The secondary structure of the 23S rRNA however, is divided into 6 asymmetrical domains that fit together like a 'jigsaw puzzle' (Ban et al., 2000). The interface view of the 50S subunit, also known as the crown view, displays one central protuberance (CP, see figure 1.1), and two lateral protuberances: L1 stalk, and L7/L12 stalk. Both lateral protuberances display inherent flexibilities that contribute to the functional aspect of this ribosomal subunit. Similar to the 30S subunit, the L-proteins are concentrated on the surface (see Figure 1.2), where the N-termini of these proteins that could reach the vicinity of the functional core, also known as the peptidyltransferase center (PTC; see Figure 1.1), are disordered in the x-ray structure.

The complete structure of the 70S ribosome was first determined from *T. thermopilus* at 5Å (Yusupov et al., 2001) and then from *E. coli* at 3.5Å (Schuwirth et al., 2005), where it provided insights on the interaction of the 30S subunit with the 50S subunit, and the interaction of the ribosome with its two main ligands: messenger RNA (mRNA) and transfer RNA (tRNA; see Figure 1.1). In particular, the 30S subunit undergoes large conformational rearrangements upon the association with the 50S subunit (Schuwirth et al., 2005). The penultimate stem (h44), which runs along the body of the 30S subunit, is laterally displaced in the structure of the 50S-bound 30S subunit when compared to the free 30S subunit structure (Schuwirth et al., 2005). High degree of flexibility was also observed in the head region, which had functional implications in the ratchet-like motion of the ribosome during the process of protein synthesis.

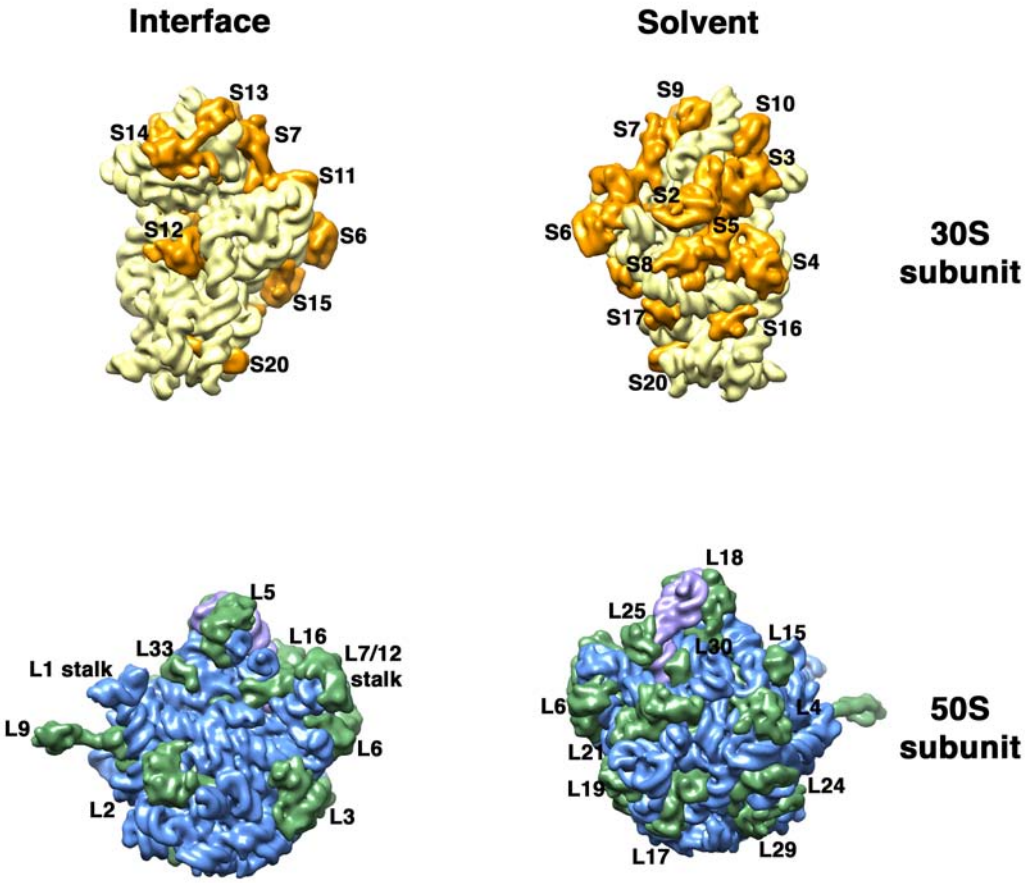


Figure 1.2 Protein and rRNA components of the 30S and 50S subunits

Proteins of the 30S subunit are colored in orange, and 16S rRNA is colored in yellow. Proteins of the 50S subunit are colored in green, 23S rRNA is colored in blue, and 5S rRNA is colored in purple. PDB ID 2AVY and 2AW4.

The structure of the 70S ribosome also indentified several contact points on the 30S and 50S subunits, some of which were essential for the association of the two subunits. These contact points were recognized as 'inter-subunit' bridges that were also visible in the cryo-EM structures of the ribosome and are dispersed throughout the interface regions of the two subunits (Lata et al., 1996). These inter-subunit bridges are

mainly composed of rRNA contacts, albeit few protein-protein interactions also exist (Yusupov et al., 2001). Additional information we learned from the structure of the ribosome was pertaining to the interactions with its substrates, mRNA and tRNA (Yusupov et al., 2001). The mRNA solely interacts with the 30S subunit, while tRNAs contact both the 30S and the 50S subunits at three distinct sites: Aminoacyl(A)-site, peptidyl(P)-site, and exit(E)-site tRNAs (Rheinberger et al., 1981) (Figure 1.1). Several interactions also occur between mRNA and tRNAs through base pairing between the codon and anti-codon nucleotides, respectively (Crick, 1966). Specifically, the anti-codon ends, also known as the anti-codon stem loops (ASL) of tRNA interact with mRNA in the RNA-rich groove at the nexus of head, body, and the platform of the 30S subunit. The other end of the tRNA, which includes the elbows and the acceptor arms, make contacts with the CP and the central regions in the interface of the 50S subunit (Yusupov et al., 2001).

1.1.3 Description of the Function of the Ribosome

The ribosome is responsible for protein biosynthesis in all organisms (Doudna & Rath, 2002), a process also identified as 'translation'. This process is assisted by several non-ribosomal accessory factors, which comprise tRNAs and mRNA, in addition to several protein factors that play a catalytic and regulatory role in the cell. These factors are classified into 4 sub-groups that are composed of initiation factors (IF1, IF2, and IF3), elongation factors (EF-Tu and EF-G), termination or release factors (RF1, RF2, and RF3), and lastly ribosome recycling factors (RRF1 and EF-G) (Noller, 1991; Kozak, 1999; Ramakrishnan, 2002; Steitz, 2008). All of these factors interact with the ribosome in co-

ordination with its two main functional cores, the DC and the PTC located on the 30S subunit and the 50S subunit respectively.

The anti-codon of the tRNA and codon of the mRNA interaction first occur at the A-site of the 30S subunit (see Figure 1.1) in the mRNA DC (Rodnina & Wintermeyer, 2001b, a). This interaction is a proofreading step by the 30S subunit where it determines whether the tRNA will be accepted (cognate state) or rejected (non-cognate state) (Crick, 1966; Hopfield, 1974; Ehrenberg & Blomberg, 1980). The DC of the 30S subunit then recognizes the geometry of the canonical codon-anticodon base pairing, and a conformational change is induced in the universally conserved nucleotides A1492, A1493, and G530 (Ogle et al., 2001). Mainly, the sugars of these bases will flip out and interact intimately with the minor groove of the first two base pairs between the codon and anti-codon. The third base pair generally assumes a wobble position where it is less monitored by the ribosome in a phenomenon known as the degeneracy of the genetic code (Crick, 1966).

After the proofreading step of the anti-codon loop (ACL) of the cognate tRNA by the DC, the other end of the tRNA (CCA end) carrying the amino acid will be accommodated into the PTC, a pocket-like structure on the 50S subunit (see Figure 1.1)(Traut & Monro, 1964; Monro, 1967). The PTC will then undergo a conformational change to properly position the ∝-amino group of the amino acid on the A-tRNA, with the carbonyl group of the amino acid on the P-tRNA (Nissen et al., 2000; Hansen et al.,

2002) and the catalysis of the new peptide bond formation will be carried out (Monro, 1967). This process leaves the P-tRNA in a 'de-acylated' or 'uncharged' state.

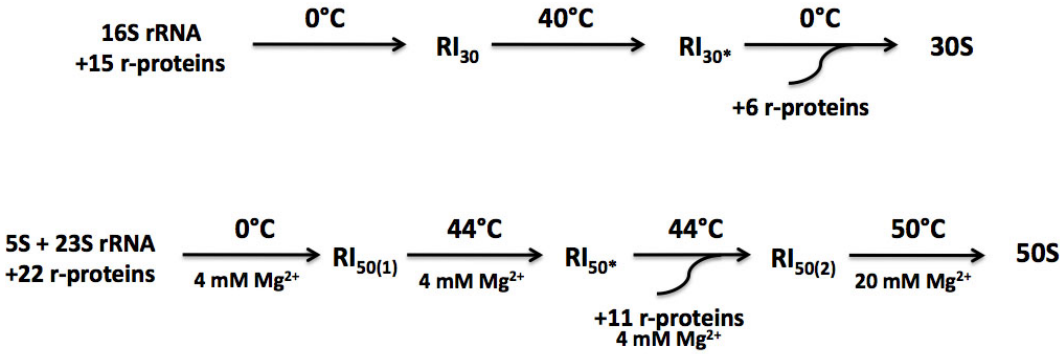


Figure 1.3 Schematic of the in-vitro ribosome assembly process.

Ribosomal subunits can be reconstituted in-vitro from purified components (rRNA and and r-proteins). RI stands for reconstitution intermediate.

Crystal structures of the 70S ribosome manifesting the 'pre-attack' conformation placed the two nucleotides A2451 of the 23S rRNA and A76 of the P-tRNA as the two candidates that play a chemical role in the catalysis (Nissen et al., 2000; Bashan et al., 2003). However, the contribution of A2451 in peptide bond catalysis was completely eliminated through genetic and biochemical studies (Youngman et al., 2004). In contrast, A76 was illustrated to play an important role as a proton shuttle catalyzing peptide bond formation (Dorner et al., 2003; Weinger et al., 2004; Schmeing et al., 2005a; Schmeing et al., 2005b). The PTC is exclusively composed of rRNA, albeit more recently L-proteins have been suggested to take part in translation (Maguire et al., 2005; Schuwirth et al.,

2005). The catalytic role of rRNA was considered to be a major breakthrough as it confirmed previous studies (Noller et al., 1992) that the ribosome is in fact a 'ribozyme' (Nissen et al., 2000).

1.2 Assembly of the Ribosome

1.2.1 *In-vitro* ribosomal assembly

It has been established that the 30S subunit can be reconstituted *in-vitro*, using purified (Traub & Nomura, 1969a; Held et al., 1973; Held et al., 1974) or recombinant r-proteins (Culver & Noller, 2000) (see Figure 1.3). In these experiments, a strict hierarchal and cooperative mode in the binding of the r-proteins to the rRNA was observed. Consequently, the binding of the ribosomal proteins has been classified under three main classes (see Figure 1.4) : (i) primary r-proteins can bind directly to naked 16S rRNA; (ii) secondary proteins require at least one bound primary protein to be able to bind the rRNA; and (iii) tertiary proteins require at least one primary and one secondary protein bound in order to bind the rRNA (reviewed in Culver, 2003). The process of RNA-protein recognition was also shown to follow a 5' to 3' directionality (see Figure 1.4), and r-proteins were classified as early, mid, or late binding proteins (Powers et al., 1993; Adilakshmi et al., 2005; Sykes & Williamson, 2009). Early-binding proteins bind to the 5' region (Body region), mid-binding proteins bind to the central domain (Platform region), and late-binding domains bind to the 3' domain (Head region). A similar, rather more complicated, assembly map (Figure 1.4) was established for the 50S subunit *in-vitro* (Herold et al., 1986; Herold & Nierhaus, 1987; Nierhaus, 1991), due to its higher complexity, larger rRNA content (~3000 nucleotides for the 50S subunit versus ~1500

nucleotides for the 30S subunit), and larger r-protein content (32-34 proteins in the 50S subunit versus 21 proteins in the 30S subunit). Consequently, the 50S assembly has been much less studied then the 30S subunit.

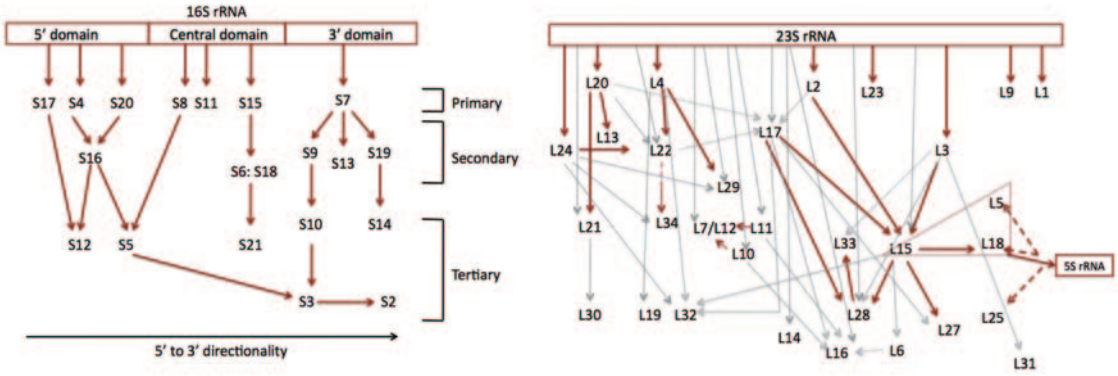


Figure 1.4 Nomura and Nierhaus assembly maps of the 30S and 50S subunits.

Nomura map (left) and Nierhaus map (right) depict r-protein binding to the 16S and 23S rRNA in a strict hierarchical and cooperative mode following a 5' to 3' directionality.

Ribosome assembly was initially thought to proceed through a single step process with the existence of rate limiting intermediates (Lindahl, 1975; Hayes & Vasseur, 1976; Sieber & Nierhaus, 1978) (see Figure 1.3). However, recent studies utilizing time resolved X-ray hydroxyl radical foot-printing experiments established that the assembly of the 30S subunit assembly nucleates concurrently at multiple sites (Adilakshmi et al., 2008). Additionally, pulse-chase labeling experiments followed by quantitative mass spectrometry analysis indicated that the 30S assembly proceeds via a landscape pathway dotted with various local conformational changes similar to that observed for protein folding (Talkington et al., 2005). Altogether, these findings combined with recent time-

resolved cryo-EM studies (Mulder et al., 2010), indicated that the assembly of the ribosome proceeds through parallel pathways rather than a single pathway (Woodson, 2008; Shajani et al., 2011).

1.2.2 *In-vivo* ribosomal assembly

The process of ribosomal assembly in the cell is divided into three basic components: (i) synthesis, processing, modification, and folding of rRNA; (ii) synthesis, modification, and folding of r-proteins; and (iii) binding of the r-proteins to the rRNA. In *E. coli*, there are seven transcriptional units of rRNA (Ellwood & Nomura, 1982). Initially, the rRNA is synthesized as a single transcript containing 16S rRNA, 23SrRNA, and 5S rRNA (see Figure 1.5). The spacer region also contains one or two tRNAs (Lund et al., 1976). The folding of the RNA transcript happens concomitantly before transcription is completed (Kaczanowska & Ryden-Aulin, 2007). The first endonuclease to cleave the RNA transcript is RNAse III, separating the precursor (pre)-16S-23S rRNA and the pre-tRNA-5S rRNA (Figure 1.5). Subsequently, RNAse III will recognize and cleave the 16S-23S rRNA transcript into three parts: a pre-16S rRNA, also known as 17S rRNA, a pre-23S rRNA, and a pre-tRNA-5S rRNA (Young & Steitz, 1978). The 5' terminus of the pre-tRNA-5S rRNA is cleaved by RNAse P, which will then release the pre-5S rRNA (Young & Steitz, 1978).

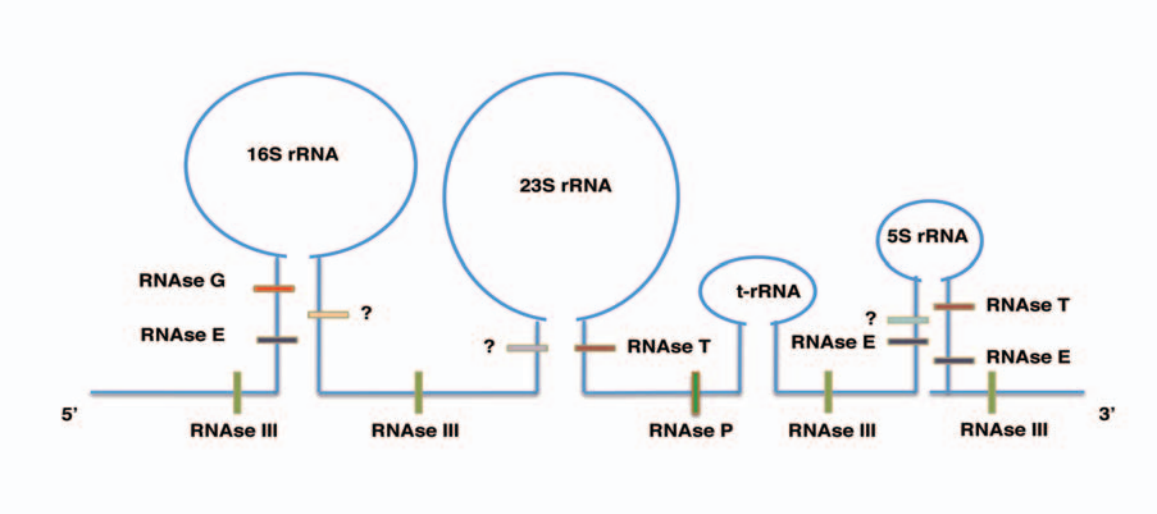


Figure 1.5 Processing of the rRNA transcript by endonucleases

The rRNA transcript is depicted from 5' to 3'. Colored dashes indicate the cleavage by the RNAses where known. An unkown RNAses cleavage is indicated by (?).

The 5' terminus of the 17S rRNA contains 115 nucleotides on the 5' terminus, which is recognized by two homologous enzymes, RNAse E and RNAse G (Li et al., 1999b). RNAse E will cleave the first 49 nucleotides, while RNAse G will cleave the final 66 nucleotides (Figure 1.5). An unknown RNAse cleaves the remaining 33 nucleotides on the 3' terminus to produce the mature 16S rRNA. The 3' terminus of the pre-23S rRNA contains 7 or 9 nucleotides and is processed by RNAse T (Li et al., 1999a), while an unknown RNAse cleaves the 3 or 7 extra nucleotides on the 5' terminus (Figure 1.5). The pre-5S rRNA contains additional 84 nucleotides on the 5' and 42 nucleotides on the 3' terminus. Both ends are recognized and partially cleaved by RNAse E (Roy & Apirion, 1983). The remaining nucleotides on the 3' terminus are cleaved by

RNAse T (Li et al., 1999a), while the RNAse that completes the processing of the 5' terminus of the 5S rRNA is still unknown.

In addition to the endonucleatic cleavage, rRNA undergoes extensive covalent modifications (Decatur & Fournier, 2002). There are two types of rRNA modifications in the cell: methylation and pseudo-uridinylation. Both occur predominantly post-transcriptionally and the modification sites are clustered in functionally important regions of the ribosome, albeit these sites are usually not conserved among different organisms (Brimacombe et al., 1988; Smith et al., 1992). In *E. coli*, there are 24 methylated nucleotides (Wilson & Nierhaus, 2007) and 10 pseudo-uridines sites on the rRNA (Del Campo et al., 2001; Ofengand et al., 2001a; Ofengand et al., 2001b). Although the exact role of these modifications is not known, some are related to antibiotic drug resistance (Doi & Arakawa, 2007). Consequently, studying these events will contribute toward the design of new drug targets. One interesting modification is the methylation of the A2058 of the 23S rRNA in the exit tunnel of the 23S rRNA, and in close proximity to the functional core of the 50S, the PTC. The Erm methyltransferase perform the methylation at this site and confer high level of resistance to macrolides such erythromycin (Poehlsgaard & Douthwaite, 2003). Another example of rRNA modification is the di-methylation of the A1518 and A1519 on the 3' end of the 16S rRNA of the 30S, also close to the its functional core, the DC. The KsgA enzyme carries out this modification, while the lack of this methylation confers resistance to the antibiotic ksaugamycin (Helser et al., 1972).

Similar to rRNA, r-proteins can also be modified (Kaczanowska & Ryden-Aulin, 2007). The most common r-protein modifications in bacteria are: methylation and acetylation. The exact role of these modifications is not known, although it is possible that these modifications are involved in ribosomal assembly or the regulation of degradation and turnover of ribosomes. An interesting modification is the methylation of the ribosomal protein L3 by the PrmB protein (Lhoest & Colson, 1977). A mutation in the *prmB* gene displays a slow growth phenotype and accumulates precursor ribosomal particles (Lhoest & Colson, 1981). Another interesting modification is the acetylation of the ribosomal protein S5 on the 30S subunit. The N-acetyltransferase enzyme, RimJ, performs the catalysis of this modification (Cumberlidge & Isono, 1979). It was later demonstrated that *in-vitro* binding of the aceylated form of S5 binds at a slower rate than the unaceylated form to the 16S rRNA in *in-vitro* 30S subunit reconstitution experiments (Talkington et al., 2005), indicating that the acetylation of S5 by RimJ occurs on mature ribosomes and is not involved in the 30S subunit assembly process.

1.3 Ribosome assembly is a factor-mediated process

In the cell, the process of ribosome assembly is a highly coordinated process, where it progresses through alternating events of conformational changes triggered by RNA-folding. This process is also coupled to r-protein recognition and binding, rRNA processing, and ion-binding events. These steps are believed to occur both co-transcriptionally and concomitantly (Powers et al., 1993; Talkington et al., 2005; Adilakshmi et al., 2008).

The folding of long rRNA helices can give rise to 'kinetic trap' intermediates due to folding errors during assembly (Noller & Nomura, 1987). Evidence that these misfolded states can be stable *in-vitro* was demonstrated by assembly experiments of the 30S and 50S subunits accumulating rate-limiting reconstitution intermediates (RI) (Traub et al., 1967; Herold & Nierhaus, 1987) (see Figure 1.3). Subsequently, the completion of the reconstitution process would require an activation energy step (i.e. heating step; see Figure 1.3) that induces a conformational rearrangement in the rRNA. This process is coupled by the binding of the r-proteins along with the conversion of these intermediate states into mature ribosomal subunits (Culver, 2003).

The *in-vitro* ribosome reconstitution experiments were also associated with several differences in comparison to the *in-vivo* ribosome assembly conditions. First, reconstitution experiments were performed with processed rRNA, and purified r-proteins. Second, *in-vitro* assembly of 30S and 50S subunits proceeded at a much slower rate (hours versus minutes) than *in-vivo* ribosome assembly and requires non-physiological ionic and salt concentrations (see Figure 1.3)(Culver, 2003). As mentioned earlier, the completion of the assembly also requires a heat activation step to convert the intermediates to the mature state. This led to the belief that the process of ribosomal assembly in the cell is more complex and would require the participation of auxiliary protein factors. These factors could limit the mis-folding events of rRNA, as well as lower the activation energy that is required for assembly intermediates to be converted to the mature subunits.

RAAF	30S or 50S assembly	Type
CsdA	50S	Helicase
RhlE	50S	Helicase
DbpA	50S	Helicase
DnaK, DnaJ	30S	Chaperone
GroEL-GroES	50S	Chaperone
RbfA	30S	Maturation Factor
RimM	30S	Maturation Factor
Era	30S	GTPase
YjeQ/RsgA	30S	GTPase
Obg/CgtA	50S	GTPase
RbgA/YlqF	50S	GTPase
YphC/EngA	50S	GTPase
YsxC	50S	GTPase

Table 1.1 List of non-ribosomal factors associated with ribosome assembly

Indeed more than 20 protein factors have been identified to be associated with the *in-vivo* process of ribosomal assembly in bacteria (Wilson & Nierhaus, 2007; Connolly & Culver, 2009; Shajani et al., 2011). In contrast, more than 200 factors have been identified to be involved in the process of ribosomal assembly factors in eukaryotes, particularly in yeast (Dez & Tollervey, 2004; Dlakic, 2005). In bacteria, ribosome-associated assembly factors (RAAFs) have been also classified into three major classes: DEAD Box-Proteins, Chaperones and Maturation Factors, and GTPases (Shajani et al., 2011). In the following sections, I will outline the most significant contributions of some the RAAFs present in bacteria (see Table 1.1).

1.3.1 DEAD-box Helicases

Helicases play an important role in ribosome assembly due to their ability to unwind rRNA and promote its proper folding (Bleichert & Baserga, 2007). Most of the

helicases identified to be associated with the ribosome assembly *in-vivo* belong to a larger family of DEAD-box ATPases, containing conserved D-E-A-D amino acids in their ATPase domain (Linder et al., 1989; Tanner et al., 2003). Although these ATPases are essential in eukaryotes, they are dispensable for cell survival under normal growth conditions in *E. coli*. Consequently, helicases associated with ribosome assembly were mostly identified from gene-knockout bacterial strains displaying cold shock sensitivities. One of these factors was the cold shock helicase activity CsdA (Charollais et al., 2004). The deletion of the *csdA* gene leads to the accumulation of precursor (pre)-50S subunits. Biochemical studies demonstrated that CsdA interacts with pre-50S particles, and that it is essential for ribosome biogenesis at low temperatures (Charollais et al., 2004). Recently, complementation studies utilizing Δ*csdA* strains identified RhlE as a putative helicase that could facilitate 50S assembly at low temperatures (Jain, 2008), which alludes that these two proteins are functionally related. The DEAD-box DpbA ATPase has been also implicated in the assembly of the 50S subunit, where an inactive mutant of DpbA over-expressed in *E. coli*, blocks ribosome assembly and accumulates pre-50S subunits, also known as 45S subunits (Sharpe Elles et al., 2009). DbpA can interact with these pre-50S particles *in-vitro*, and a stimulation in its ATPase activity was observed (Sharpe Elles et al., 2009).

1.3.2 Chaperones and Maturation Factors

Chaperones were also implicated in the *in-vivo* assembly of the ribosome. These small molecular machines can assist in the assembly process at either low temperatures such the DnaK-DnaJ-GrpE co-chaperones or at elevated temperatures such as the GroEL-

GroES heat shock proteins (El Hage et al., 2001; Al Refaii & Alix, 2009). The deletion of *dnaK* or *dnaJ* genes in bacteria exhibits a slow growth phenotype at low temperatures (below 30°C) and accumulates pre-30S subunits at elevated temperatures (above 40°C). These immature subunits were able to convert slowly to mature subunits (Al Refaii & Alix, 2009), which implicates the DnaJ-DnaK complex in catalyzing the assembly of the 30S subunit (Maki et al., 2002). The ability of the GroEL chaperone to convert pre-50S subunits to mature 50S subunits however, implicated this protein in catalyzing the process of 50S subunit assembly (El Hage et al., 2001). At elevated temperatures, the over-expression of GroEL-GroES was indentified as a partial suppressor for phenotype of the *dnaJ* or *dnaK* knockout strains (El Hage et al., 2001; Al Refaii & Alix, 2009), demonstrating a functional interplay between the two chaperones.

The cold shock protein RbfA was identified to be associated with the assembly of the 30S subunit (Jones & Inouye, 1996). Although its role is still elusive, the depletion of this protein accumulates immature 30S subunits that contains 17S rRNA and cannot associate with the 50S subunits (Bylund et al., 1998; Xia et al., 2003). Cryo-electron microscopy (cryo-EM) studies on the co-structure of the RbfA and the mature 30S subunit mapped this protein to the DC in the 50S-interface region (Datta et al., 2007). Interaction of RbfA with the 30S subunit induces a conformational change in the 3'minor domain of the 16S rRNA, displacing h45 and h44. Interestingly, RbfA is a known suppressor of another putative assembly factor RimM. *E. coli* strains containing a knockout of the *rimM* gene also accumulate pre-30S, containing 17S rRNA, and exhibit a slow growth phenotype (Bylund et al., 1998).

1.3.3 GTPases, putative regulators of ribosome assembly

GTPases play a role in regulating diverse cellular processes, particularly with the process of protein synthesis (Caldon et al., 2001), and have also been involved in the assembly of ribosome (Britton, 2009; Connolly & Culver, 2009). Although it is not clear how they contribute to the building of the ribosomal subunits, the role of GTPases could be exemplified as molecular switches that turn the process of assembly on or off by sensing the variation in GTP levels (Britton, 2009).

***GTPases and the assembly of the 30S subunit*:** Several GTPases have been implicated in the assembly of the 30S subunit (Brown, 2005; Britton, 2009). The contribution of two well characterized GTPases, Era and YjeQ/RsgA, will be highlighted below.

The Era protein is highly conserved GTPase that is involved in the assembly of the 30S subunit (Sharma et al., 2005). Era is essential in *E. coli* and its depletion leads to a phenotype reminiscent of the *ΔrimM* and *ΔrbfA* strains by accumulating pre-30S particles containing unprocessed rRNA (Inoue et al., 2003). Recently, genetic studies linked Era to the methyltransferase KsgA and the maturation factor RbfA. For instance, the over-expression of Era compliments *ΔrbfA* strains at high temperatures (Inoue et al., 2006), while the overexpression of KsgA compliments a cold sensitive Era mutants (Xu et al., 2008). The cryo-EM structure of Era in complex with the 30S subunit from *T. thermophilus*, placed Era on the cleft formed by the platform and the head of the 30S subunit, in close proximity to the DC (Sharma et al., 2005). The contact points formed by

the KH domain of Era and the 3' terminus of 16S rRNA overlaps with the KsgA methylation sites, A1518 and A1519, on h45. More recently, biochemical studies demonstrated that the overexpression of the RbfA protein potentiated a slow growth phenotype of *ΔksgA* strains from *E. coli* (Connolly & Culver, 2013). These strains accumulated 70S-like ribosomes that contained unprocessed 16S rRNA. Altogether, these findings indicate a yet uncharacterized functional interplay between Era, RbfA, and KsgA during the *in-vivo* 30S assembly.

The YjeQ protein, also known as RsgA, is involved in the assembly of the 30S subunit in *E. coli* and *B. subtilis* (Daigle & Brown, 2004; Himeno et al., 2004; Campbell et al., 2005). The GTPase activity of YjeQ is stimulated in the presence of mature 30S subunits and 70S ribosomes (Daigle & Brown, 2004), where it can bind to the 30S subunit in the presence of 5'-Guanylyl imidodiphosphate (GMPPNP), a non-hydrolysable form of GTP (Daigle & Brown, 2004; Himeno et al., 2004). *ΔyjeQ* strains exhibit a slow growth phenotype and accumulate immature 30S subunits containing 17S rRNA instead of processed 16S rRNA (Himeno et al., 2004; Campbell et al., 2005). Recently, studies employing chemical modification experiments demonstrated a change in the reactivity of nucleotides in the interface region of the 30S subunit, in close proximity to the mRNA DC (Kimura et al., 2008). YjeQ was also shown to compete for tRNA binding sites (Kimura et al., 2008), whereas antibiotics that bind to the decoding center inhibit the stimulation of its GTPase activity (Campbell et al., 2005; Kimura et al., 2007).

***GTPases and assembly of the 50S subunit*:** There are several GTPases that are involved in the assembly of the 50S subunit (Brown, 2005; Britton, 2009). Below, I will highlight the contributions of the ObgE/CgtA, RbgA/YlqF, YphC/EngA, and YsxC proteins in the assembly of the ribosome.

ObgE is an essential and highly conserved GTPase in all organisms, and has been implicated in several cellular processes including the assembly of the 50S subunit (reviewed in Brown, 2005). The depletion of ObgE leads to the accumulation of pre-50S particles and free 30S subunit and a decrease in 70S ribosomes (Jiang et al., 2006). Biochemical studies have suggested that the ObgE protein has a tendency to interact with the ribosome, particularly with the 50S subunit (Lin et al., 1999; Wout et al., 2004; Zhang & Haldenwang, 2004; Sato et al., 2005). ObgE and RrmJ, the methyltransferase of U2552 on 23S rRNA, are genetically linked, where the over-expression of ObgE compliments a slow growth phenotype of the *ΔrrmJ* strains and restore the defective ribosome profile (Tan et al., 2002). Pull-down experiments also identified the helicase CsdA, a putative 50S subunit assembly factor, as an interaction partner with ObgE.

RbgA, previously known as YlqF, is an essential protein for the growth of *B. subtilis*. This GTPase is widely conserved from bacteria to eukaryotes, but not present in *E. coli* (reviewed in Britton, 2009). The depletion of the RbgA protein accumulates free 30S subunits and pre-50S (45S) subunits that migrate slower on sucrose gradients during ultra-centrifugation (Matsuo et al., 2006; Uicker et al., 2006). RbgA interacts with the 45S subunits in the presence of GTP, but it can only form a stable interaction with 50S

subunit in the presence of GDPNP (Matsuo et al., 2006; Uicker et al., 2006; Achila et al., 2012).

These 45S subunits also contain an incomplete protein complement lacking late binding ribosomal protein L16, L27, and L36 (Uicker et al., 2006; Matsuo et al., 2007). All three proteins cluster in the interface region of the 50S close to the PTC, and make contacts with A-site and P-site tRNAs (Selmer et al., 2006). The absence of L16 is particularly interesting, as it was demonstrated that the binding of L16 to L16-deficient pre-50S particles induces a large conformational change in the rRNA (Teraoka & Nierhaus, 1978). It is thus believed that L16 could act as a "molecular glue" to stabilize the compact structure of the 50S subunit (Nishimura et al., 2004).

The two GTPases YsxC and YphC/EngA are also involved in the assembly of the 50S subunit (Bharat et al., 2006; Schaefer et al., 2006; Wicker-Planquart et al., 2008). Both GTPases are widely conserved and essential for cell growth in bacteria (Galperin & Koonin, 2004; Ruzheinikov et al., 2004; Hwang & Inouye, 2006) The depletion of either YsxC or YphC from *B. subtilis* accumulates pre-50S particles that lack the ribosomal proteins L16, L27 and L36 (Schaefer et al., 2006), similar to the immature 45S subunit that accumulate in RbgA-depleted strains. Interestingly, all three GTPases have been demonstrated to interact with the ribosome in close proximity to the PTC and tRNA binding sites (Matsuo et al., 2006; Wicker-Planquart et al., 2008). These findings suggest some sort of coordination between these GTPases during 50S subunit assembly, possibly

by acting as checkpoint proteins monitoring the final fold of the functionally important regions and preventing immature subunits from participating in translation.

1.4 Thesis Objective

Several GTPases have been implicated in the assembly of the ribosomes. These factors belong to the TRAFAC (translation factor class) family of GTPases (Leipe et al., 2002; Brown, 2005; Britton, 2009). However, no known function on how these GTPases contribute and regulate ribosomal assembly in the cell has been determined. It is fascinating how mutations affecting the functions of these GTPases confer pleiotropic phenotypes linking the process of ribosome assembly to cell cycle and cell growth (Brown, 2005). Accordingly, GTPases could act as molecular switches that turn ON/OFF ribosome assembly in response to nutrients availability (Britton, 2009), which would couple intrinsic cellular GTP levels to ribosome assembly.

In this thesis, I will focus on the structural and biochemical characterization of the function of the two enigmatic GTPases, YjeQ and RbgA, in the assembly process of the 30S and 50S subunits, respectively. The subsequent work in this thesis will be addressing the answers to three central questions:

1. What is the structural and biochemical state of the immature 30S subunit that accumulates in the Δ*yjeQ* strains?
2. How does the YjeQ GTPase interact with the 30S subunit during its maturation?
3. How do the recruitment of the ribosomal protein L16 and late-binding L-proteins contribute to the role RbgA protein during late stages of 50S assembly?

1.5 Thesis Organization

This thesis is divided into three research chapters (2-4), with additional introduction and discussion chapters, Chapter 1 and 5 respectively. Chapter 2 is a published manuscript in the *RNA journal* and gives insights on late stages of 30S assembly by using a perturbation in the function of the YjeQ protein. The aim of Chapter 3 is to develop an understanding of the role of YjeQ in ribosomal assembly by determining its binding site on the 30S subunit and describing the conformational rearrangements imposed by this interaction. This work was also published in the *RNA journal*. Chapter 3 uncovered a novel functional interplay between the ribosomal protein L16 and the RbgA protein during late stages 50s assembly. This work is currently in preparation for submission in April. Chapter 1 is aimed toward providing appropriate background information on the Ribosome: structure, function, and assembly, while Chapter 5 is intended as a summary and a general discussion linking the findings in each research chapter. This final chapter also depicts the significance of this work and its contribution to the overall understanding of the ribosome biogenesis in bacteria. Overall, this thesis was written and formatted as a "sandwich thesis".

CHAPTER 2. Understanding Ribosome Assembly: the Structure of *in- vivo* Assembled Immature 30S Subunits Revealed by Cryo-electron Microscopy

2.1 Author's Preface

The data presented in chapter 2 has been published in the peer-reviewed *RNA Journal*, and appears in its published format. This work characterized an *in-vivo* assembled 30S subunit isolated from Δ*yjeQ* strains. This work was carried in collaboration with Dr. Eric Brown, Dr. Janine Maddock and Dr. Jaime Martín-Benito. I have performed all cryo-EM microscopy experiments including sample preparation, data collection, and image processing. Geordie Stewart from the Brown group performed the Northern Blot experiments. Dr. Ryzard Zielke with Dr. Janine Maddock performed the quantitative mass spectrometry experiments. Dr. Jaime Martín-Benito performed part of the image classification approach. Dr. Joaquin Ortega, Dr. Eric Brown, Geordie Stewart and I analyzed the data and wrote the manuscript.

The full citations is as follows:

Jomaa, A., Stewart, G., Martín-Benito, J., Zielke, R. Campbell, T.L., Maddock, J.R., Brown, E.D., Ortega, J. (2011) Understanding Ribosome Assembly: the Structure of *in vivo* Assembled Immature 30S Subunits Revealed by Cryo-electron Microscopy. RNA. 2011 Apr;17(4):697-709.

2.2 ABSTRACT

Four decades after early *in vitro* assembly studies demonstrated that ribosome assembly is a controlled process, our understanding of ribosome assembly is still incomplete. Just as structure determination has been so important to understanding ribosome function so too will it be critical to sorting out the assembly process. Here, we used a viable deletion in the *yjeQ* gene, a recognized ribosome assembly factor, to isolate and structurally characterize immature 30S subunits assembled *in vivo*. These small ribosome subunits contained unprocessed 17S rRNA and lacked some late ribosomal proteins. Cryo-electron microscopy reconstructions revealed that the presence of precursor sequences in the rRNA induces a severe distortion in the 3' minor domain of the subunit involved in the decoding of mRNA and interaction with the large ribosome subunit. These findings suggest that rRNA processing events induce key local conformational changes directing the structure toward the mature assembly. We concluded that the rRNA processing, folding, and entry of tertiary r-proteins are interdependent events in the late stages of 30S subunit assembly. In addition, we demonstrate how studies of emerging assembly factors in ribosome biogenesis can help to elucidate the path of subunit assembly *in-vivo*.

2.3 INTRODUCTION

The ribosome constitutes perhaps one of the most intricate macromolecular complexes in the cell and assembles with remarkable reliability. The high-resolution atomic models of the ribosome (Ban et al., 2000; Schluenzen et al., 2000; Wimberly et al., 2000; Harms et al., 2001; Yusupov et al., 2001; Schuwirth et al., 2005) provide the detailed structure of the end product of the assembly line. However, very little information about the assembly process itself can be interpreted from the inspection of these structures.

Ribosomal subunit assembly has been the subject of more than 40 years of investigation that began with seminal studies by the Nomura laboratory in the 1960s to map the reconstitution of the 30S subunit *in vitro* (Hosokawa et al., 1966; Traub & Nomura, 1968b, a, 1969b). Interest in this field has resurfaced in the last several years, sparked by the high-resolution X-ray structures of the ribosome and by the development of new techniques for the study of the kinetics and mechanism of rRNA folding in subunit assembly *in vitro* (Talkington et al., 2005; Adilakshmi et al., 2008). Together these studies have considerably advanced our understanding of the fundamental driving forces of assembly, including RNA folding, protein binding order and the possible existence of intermediates. However, just as structures of the intermediates in protein synthesis have transformed our understanding of ribosome function (Frank, 2003; Mitra & Frank, 2006), so too will structural studies constitute the next frontier in ribosome assembly.

The 70S ribosome in *Escherichia coli* is composed of a large 50S and a small 30S

subunit. The 30S subunit consists of one 16S ribosomal RNA (rRNA) molecule and 21 ribosomal proteins (r-proteins) and the 50S subunit is made of two RNA molecules, the 23S and 5S rRNAs and 34 proteins. Bacterial ribosome assembly commences with the transcription of rRNA as a single precursor transcript containing the three rRNAs (Srivastava & Schlessinger, 1990). Processing of the primary rRNA transcript occurs very rapidly and begins before transcription is completed. RNase III performs the primary processing that separates the three rRNAs. The resulting fragments are called precursor rRNAs and contain additional nucleotides at both their 5' and 3' ends (Srivastava & Schlessinger, 1990). Subsequent secondary processing events cleave the extra nucleotides and produce the mature rRNA molecules. In particular, the precursor rRNA for the 30S ribosomal subunit (called 17S rRNA), contains an additional 115 and 33 nucleotides at the 5' and 3' ends, respectively. The coordinated action of RNase E and RNase G removes the precursor sequence at the 5' terminus (Li et al., 1999b). Processing of the 3'end is less well characterized but may occur in a single cleavage step that removes the 33 extra nucleotides at this end of the molecule (Hayes & Vasseur, 1976).

After transcription, the 17S rRNA is thought to form local secondary structures rapidly prior to trimming of the 5' and 3' precursor sequences. The initially folded precursor rRNA is quickly recognized and bound by the r-proteins. Vintage experiments by Nomura (Hosokawa et al., 1966; Traub & Nomura, 1968b, a, 1969b) and more recent experiments from the Williamson and Woodson laboratories (Talkington et al., 2005; Adilakshmi et al., 2008) have defined the hierarchy and kinetic pathway of binding of the 21 r-proteins of the 30S subunit to mature 16S rRNA *in vitro*. The r-proteins are

designated as primary (bind directly to the rRNA), secondary (binding is dependent on primary binding r-proteins) or tertiary (binding is dependent on secondary binding r-proteins). Using pulse-chase labeling quantified by mass spectrometry, Williamson and colleagues (Talkington et al., 2005) have presented compelling evidence to suggest protein binding to the rRNA drives conformational rearrangements that stabilize the native fold of the 30S subunit. This paradigm has been strengthened by Woodson and coworkers (Adilakshmi et al., 2008) who have used time-resolved X-ray hydroxyl radical foot-printing to show multiple early folding nucleation events and induced fit of protein-rRNA complexes. Binding of primary proteins appears to stabilize the local RNA structure and induces conformational changes that create new binding sites for secondary proteins. Thus *in vitro* studies using mature 16S rRNA suggest that assembly proceeds by an alternating series of RNA conformational changes and protein binding events until the mature structure is produced. A long-standing question, however, is the predictive strength of these *in vitro* experiments to the assembly processes occurring *in vivo* with freshly transcribed precursor rRNA and newly expressed ribosomal proteins.

As described in early studies by Nomura, the *in vitro* reconstitution of intact ribosomal subunits (and even functional ribosomes) is possible by providing the processed rRNA and r-proteins (Mizushima & Nomura, 1970). However, *in vitro* ribosomal assembly requires conditions that are far from physiological and occurs much more slowly than *in vivo*. Indeed, trans-acting protein factors, including ribonucleases, helicases, rRNA (and r-protein) modification enzymes and GTPases are increasingly recognized as having a role in subunit assembly in the cell (Wilson & Nierhaus, 2007).

While the details of their mechanism of action are largely unknown, the deletion of genes encoding many of these factors causes accumulation of precursor rRNAs (Wilson & Nierhaus, 2007; Connolly & Culver, 2009). Perhaps the most notable and enigmatic of these are the GTPases (Brown, 2005) including the YjeQ protein from *E. coli*. YjeQ is a slow GTPase that is highly stimulated by purified 30S subunits (Daigle et al., 2002; Daigle & Brown, 2004; Himeno et al., 2004). Disruption of the *yjeQ* gene in *E. coli* severely affects growth, likely due to the accumulation of ribosomal subunits (Daigle & Brown, 2004; Himeno et al., 2004; Campbell et al., 2005; Campbell & Brown, 2008). In addition, a large proportion of the 30S ribosomal subunits in a *yjeQ* deletion strain of *E. coli* (Δ*yjeQ*) are immature as they contain a precursor of the 16S rRNA molecule (Himeno et al., 2004). In vitro, recombinant YjeQ interact strongly with the 30S ribosomal subunit (Daigle & Brown, 2004; Himeno et al., 2004) and this interaction results in a 160-fold stimulation of the YjeQ GTPase activity (Daigle & Brown, 2004). The binding location of YjeQ to the 30S subunit is still unclear and existing studies do not agree. Two studies suggested that YjeQ binds in close contact with the A site and its binding induces conformational changes around the decoding center (Himeno et al., 2004; Kimura et al., 2008). Conversely, a chemical genetic based study (Campbell et al., 2005) concluded that YjeQ binds at a distal site on the ribosome, which is different from the A site.

In the present work, we used the Δ*yjeQ* strain of *E. coli* to purify immature 30S subunits that were assembled *in vivo*. We showed by Northern blotting that the 16S rRNA of these small subunits was incompletely processed and using isotope tagging proteomic

methodology, we demonstrated that 30S isolated from the Δ*yjeQ* strain had an incomplete protein complement. The three-dimensional reconstruction obtained by cryo-electron microscopy (cryo-EM) of these 30S subunits revealed a profound distortion of helix 44 that is located in the proximity (~25Å) of the 3' mature terminus of its 16S rRNA. Helix 44 is a structural element in the 30S subunit crucial to the decoding of mRNA and interaction with the large ribosome subunit. This finding indicates that processing of the 3' precursor sequence of the 17S rRNA is critical to the development of a conformation compatible with 50S binding and protein synthesis. Our results highlight the importance of the rRNA processing events in the assembly process and confirm that the YjeQ protein is a *bona fide* ribosome assembly factor assisting in late stages of assembly and link its function to processing of the 17S rRNA.

2.4 RESULTS

2.4.1 Maturation of both the 5' and 3' terminus of the 17S rRNA is delayed in an *E. coli* Δ*yjeQ* strain

We isolated total cellular RNA from an *E. coli* deletion strain lacking the gene encoding the ribosome assembly factor YjeQ. The purified RNA was resolved and visualized in 0.7% modified agarose gels (Figure 2.1A). In addition to the species corresponding to the mature 16S rRNA molecules seen in wild type cells, we also observed a band of slower mobility corresponding to a precursor form of this molecule. A third band with faster mobility was also visible. The existence of this latter band is consistent with previous studies in Δ*yjeQ* cells (Himeno et al., 2004) and it may constitute either a 16S rRNA degradation product or an aberrantly processed rRNA produced in the absence of YjeQ.

We used Northern hybridization to determine if the 5' and 3' precursor sequences were present in the immature 16S rRNA molecule accumulated in Δ*yjeQ* cells. Transferred membranes were complemented with end fragments of the 5' and 3' precursor sequences of 17S rRNA and, indeed, both the 5' and 3' precursor sequences were found (Figure 2.1B). These results indicate that in cells lacking the *yjeQ* gene the secondary processing reactions leading to the production of mature 16S rRNA are delayed. Consequently, these bacteria accumulate unprocessed 17S rRNA molecules with precursor sequences at both, the 5' and 3', ends. Furthermore, our findings confirm that the YjeQ protein is a *bona fide* ribosome assembly factor with a role in processing the precursor 17S rRNA.

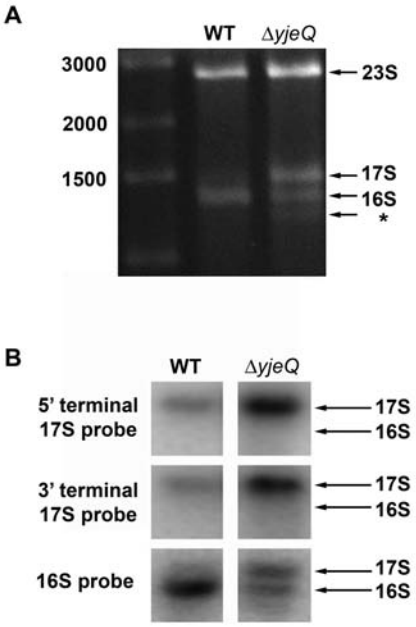


Figure 2.1 Characterization of the rRNA from *E. coli* Δ*yjeQ* strain.

(A) Electrophoretic analysis of rRNA derived from the wild type and yjeQ null strain cell extracts. The wild type gel pattern exhibited only two bands with migration corresponding to mature 16S and 23S rRNA. Two additional bands were observed in the yjeQ deletion strain corresponding to a precursor form of 16S rRNA (17S label) and a previously identified species suspected to be a degradation product or an aberrantly processed rRNA (asterisk). The sizes of the fragments of the DNA ladder in the left lane of the gel are indicated. (B) The identity of the precursor 16S rRNA found in the yjeQ null strain was confirmed by Northern blot analysis using radiolabeled DNA sequence-specific probes directed at the 5' and 3' terminal sequences of 17S rRNA. Hybridization of these probes with the rRNA precursor derived from the yjeQ deletion cell extract indicated that the species was indeed unprocessed 17S rRNA with intact 5' and 3' terminal sequences. Mobility of the 16S rRNA and 17S rRNA is indicated.

2.4.2 Protein complement of the immature 30S ribosomal subunits purified from *E. coli* Δ*yjeQ* cells

Next, we aimed to identify the protein complement present in the immature 17S rRNA-containing 30S subunits from Δ*yjeQ* cells and compared their relative protein levels with the mature, wild type, 16S rRNA-containing 30S subunits using the iTRAQ (isobaric tag for relative and absolute quantification) method (Ross et al., 2004). Considering that Δ*yjeQ* cells have approximately 30% of mature 16S rRNA-containing 30S subunits (Figure 2.1A) and that they are mostly associated to large ribosomal subunits forming 70S ribosomes (Himeno et al., 2004), we purified the 30S fraction from the Δ*yjeQ* cells under non-dissociating conditions in order to eliminate most of the mature 16S rRNA-containing 30S subunits and produce a preparation enriched in immature 17S rRNA-containing small subunits. Similarly, mature 16S rRNA-containing 30S subunits were prepared from wild type *E. coli* for comparative purposes.

We performed a double duplex isobaric labeling experiment (Jiang et al., 2007) to find the proteins present in our two 30S subunit preparations. We identified 1207 peptides to give us 41 unique proteins identified by at least 2 different peptides at 95% confidence The identification was comprehensive and included the 21 r-proteins of the 30S subunit (Figure 2.2 and Supplemental Table 2.1). We also identified eleven r-proteins from the 50S subunit and nine non-ribosomal proteins in these purified preparations (Supplemental Figure S2.1 and Supplemental Table 2.1). Identification of these proteins suggests a minor contamination of our 30S subunit preparations with 50S subunits and other proteins, most of them unrelated to the ribosome.

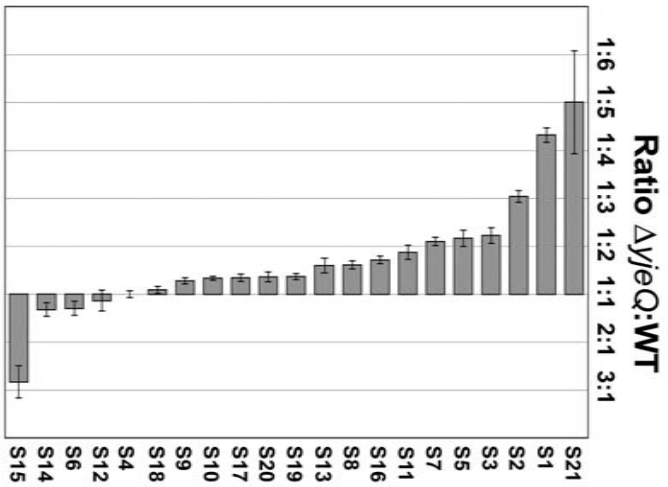


Figure 2.2 iTRAQ analysis of the 30S subunits purified from Δ*yjeQ* Cells.

Small subunit r-proteins. The plot shows the relative levels of the small subunit r-proteins between the 30S subunit purified from ΔyjeQ and wild type (WT) cells. Relative levels for each protein are expressed as the average ratio ΔyjeQ:WT obtained from two replicas of the experiment. Only peptides that identify proteins with ≥95% confidence were used for the calculation of these ratios. The error bars indicate the standard error of the mean for each ratio.

The iTRAQ double duplex isobaric labeling was also used to compare the relative amounts of the proteins in the 30S subunits purified from Δ*yjeQ* (largely immature 30S subunits) and wild type cells (largely mature 30S subunits). Results from this analysis were expressed as the Δ*yjeQ*:WT and WT:Δ*yjeQ* ratios for the 30S subunit r-proteins and for the other identified proteins (Figure 2.2, Supplemental Figure S2.1 and Supplemental Table 2.1). These figures and table provided two levels of analysis of the data. Only peptides that identified proteins with ≥95% confidence were used for the calculation of the ratios shown in Figure 2.1 and Supplemental Figure S2.1. Additional ratios are presented in Supplemental Table 2.1 that include some peptide identification data that lie

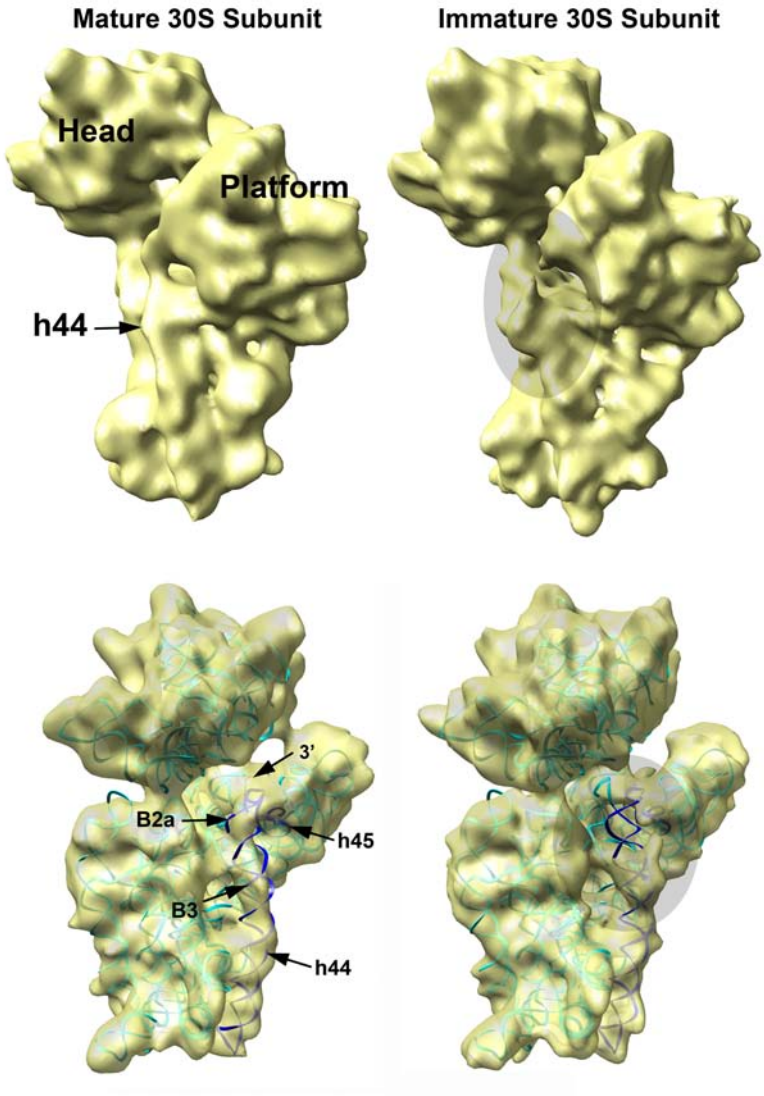
outside of this strict confidence interval. Both analyses concluded that the 30S subunit r-protein complement was largely the same in the immature and mature 30S subunits purified from Δ*yjeQ* and wild type cells, respectively. However, three r-proteins, S21, S1 and S2, were found to be significantly underrepresented in the immature 30S subunits (Figure 2.1) with a high degree of confidence (p-values for these proteins in two biological replicas was ≤ 0.05; see Supplemental Table 2.1). In addition, r-proteins S3, S5, S7, S11, S16, S8 and S13 were found slightly underrepresented in the immature 30S subunits (Figure 2.2). Interestingly, S15 was slightly underrepresented in the mature 30S subunit (Figure 2.2). This was a curious result that may well be consistent with a primary role for S15 in ribosome assembly rather than ribosome function (Bubunenko et al., 2006; Sykes & Williamson, 2009) and thus, binding to the immature structure could conceivable be tighter than to the mature structure. Accordingly, S15 may have dissociated during the purification process from some of the mature 30S subunits prepared from wild type cells but not from the immature particles from Δ*yjeQ* cells, resulting in the underrepresentation of this protein in the former. Alternatively, it is possible that S15 was bound non-specifically to the immature particles in superstoichiometric amounts and generated this result. Our data cannot rule out any of these two possibilities.

In conclusion, this analysis indicated that the protein complement of the immature 30S subunits purified from Δ*yjeQ* cells was similar to the mature small subunits except for the r-proteins S21, S1 and S2, which were significantly underrepresented. These findings suggest that in the absence of YjeQ, binding of these proteins is weaker and/or

their entry in the assembling 30S subunit is delayed. Interestingly, these r-proteins are tertiary proteins that bind to the 30S subunit only at the latest stages of the assembly process, which indicates that while immature, these 30S particles are at a late stage in the subunit assembly process (Mizushima & Nomura, 1970; Culver, 2003).

2.4.3 The immature 30S subunit from Δ*yjeQ* cells features a distorted helix 44 and decoding center

The structure of the immature 30S subunit purified from Δ*yjeQ* cells was obtained (Figure 2.3, upper right panel) using cryo-EM, image classification and three-dimensional (3D) reconstructions techniques. To this end, 30S subunits projections were selected from the electron micrographs. The contamination of the immature 30S subunit preparation with 50S subunits revealed by the iTRAQ technique (Supplemental Figure S2.1 and Supplemental Table 2.1) was very minor and the electron micrographs rarely showed any 50S subunit projections (Supplemental Figure S2.2). These particles were easily excluded during the picking process. The supervised classification approach (Valle et al., 2002; Gao et al., 2004) described in Supplemental Figure S2.3 was first used to eliminate the projections representing mature 30S subunits (27.4%) from the data set. The structure of the immature 30S subunit (Figure 2.3, upper right panel) was then obtained from the remaining projections in the data set and the structure was refined to 11.6 Å resolution (Supplemental Figure S2.4). In addition, a homogeneous preparation of mature 16S rRNA-containing 30S subunits purified from wild type *E. coli* cells was used to produce a control cryo-EM reconstruction that we used to evaluate the structural defects in the immature structure (Figure 2.3, upper left panel).


Mature 30S Subunit
Immature 30S Subunit
Head
Platform
h44
3'
B2a
h45
B3
h44

continued

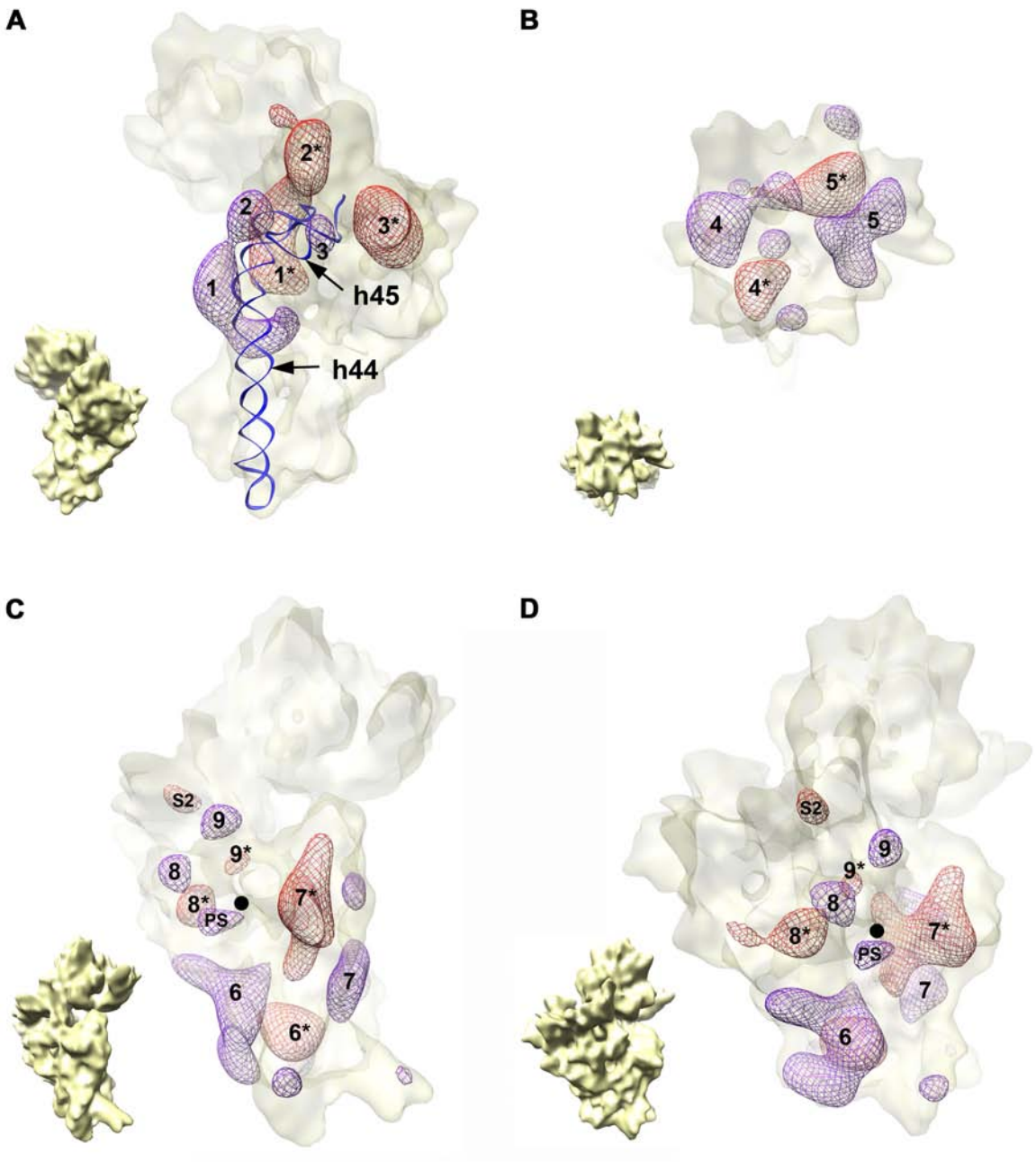
Figure 2.3 3D Reconstruction of the immature 30S ribosomal subunits purified from Δ*yjeQ E.coli* cells.

Platform-view of the cryo-EM maps of the immature (upper right panel) and mature (upper left panel) 30S subunits. The immature structure was produced from 42,873 projections of 30S subunits purified from ΔyjeQ cells after 16,180 projections (27.4% of the total) representing mature small subunits were removed from the data set using a supervised classification approach. The mature structure (left panel) was produced from 19,924 projection images obtained from a homogeneous sample of mature 16S rRNA-containing 30S subunits purified from wild type cells. Main landmarks of the 30S subunit are indicated. In the panels below, the X-ray structure of the wild type 30S subunit, shown as a ribbon representation, was fitted into the cryo-EM maps of both the mature (left) and immature (right) 30S subunits to illustrate their divergences. In the X-ray structure the rRNA is shown in cyan except helix 44 and 45 that are colored in dark blue. The r-proteins are not shown. Main landmarks of the 30S subunit indicated include helix 44 (h44), helix 45 (h45), the inter-subunit bridges B2a (B2a) and B3 (B3) important for the interaction with the 50S subunit and the 3' terminus of the 16S rRNA molecule (3'). The shadowed ovals in the right panels show the area in the helix 44 and decoding center distorted in the EM map of the immature 30S subunit.

When compared to the control reconstruction, the structure of the immature 30S subunit showed the main landmarks existing in the mature subunit suggesting that these particles represent a late stage of assembly where the main structural domains of the 30S subunit have been assembled (Figure 2.4, upper panels). However, the reconstruction of the immature subunit revealed that the conformation of the 3' minor domain, including helix 44, was clearly distorted. In particular, the upper segment of helix 44, responsible for tRNA binding, was displaced outwardly, protruding from the interface of the 30S subunit that contacts the large 50S subunit (Figure 2.3, lower right panel). In the mature 30S subunit, this structural motif forms part of the decoding center and then further meanders towards helix 45 and the 3' terminus of the 16S rRNA (Figure 2.3, lower left panel). In the immature structure, the distortion in helix 44 is likely also displacing helix

45 and the unprocessed 3' terminus to an outward flexible position. These changes indicate that the decoding center and two intersubunit bridges with the 50S subunit (B3 and B2a) are still not fully structured (Figure 2.3, lower panels).

The very significant distortions observed in helix 44 and the decoding center of immature 30S subunits isolated from the *yjeQ* deletion strain suggests that the presence of the precursor sequences (in particular the one at the 3' end) imposes a non-functional conformation in the 30S subunit and that rRNA processing is required to gain a conformation compatible with 50S binding and protein synthesis.



continued

Figure 2.4 Difference map analysis of the immature 30S subunit cryo-EM structure.

Overlay of the forward and reverse difference maps. Densities from the forward and reverse difference maps are shown as a violet and red mesh, respectively contained in a semitransparent surface representation of the immature 30S subunit structure from ΔyjeQ cells. Matching densities are identified with the same number and the density in the reverse map is identified by an asterisk ''. Panel (A) shows the densities present in the 3' minor (helix 44 & 45) and central domains of the two difference maps. Similarly, panel (B) displays only the densities present in the 3' major domain of these difference maps and panel (C) and (D) show two views of the densities existing on the 5' domain. The small cryo-EM map in the bottom left corner of each panel is an aid for orientation and represents the immature 30S subunit in the same view as the difference maps in the corresponding panel. Helix 44 (h44) and helix 45 (h45) are also shown as a ribbon representation in panel (A) to illustrate the displacement of helix 44 in the immature 30S subunit structure. The 5' end of the mature 16S rRNA is indicated with a black dot in panel (C) and (D). The density labeled as "S2" in panel (C) and (D) indicates the amount of density corresponding to this protein missing in the immature structure. The density labeled as "PS" in the same panels represent the additional density existing in the immature structure due to the presence of the 5' precursor sequence of the rRNA.*

2.4.4 Conformational differences in the immature 30S subunit relative to the mature structure

To obtain further details on the displacement of helix 44 as well as to analyze the conformational differences between the immature and mature 30S subunit structure, we calculated the "forward" difference map by subtracting the control map of the mature 30S subunit from wild type cells (Figure 2.3, upper left panel) from that of the immature 30S subunit from Δ*yjeQ* cells (Figure 2.3, upper right panel). In addition, to differentiate those regions in the EM map of the immature 30S subunit in a different conformation from areas lacking or containing an additional density, we also calculated the "reverse" difference map (i.e. mature 30S subunit map minus immature 30S map). Overlaying these two maps allowed us to study the differences between the two structures (Figure 2.4).

Most of the densities observed in the forward map had a nearby density in the reverse map (matching densities are identified with the same number in Figure 2.4; the density in the reverse map is identified by an asterisk '*') highlighting movement and/or conformational variations in the immature subunit with respect to the mature structure. Consistent with the large distortion observed in the upper segment of helix 44 in the structure of the immature 30S subunit, the overlaid difference maps of the 3' minor (helix 44 and 45) and central (platform) domains showed two matching densities (Figure 2.4A; densities 1 and 1*) that describe the outward displacement of the helix in the immature subunit. Regarding the other three major domains of the small subunit (5' (body), central (platform), 3' major (head)), the surface rendering representation of the immature 30S subunit EM map showed that the structure of these domains in the immature subunit was very similar to the mature structure (Figure 2.3, upper panels). However, we found nearby densities in the overlaid difference maps of all three domains (Figure 2.4) indicating that the relative orientation of the major structural domains in the immature subunit differs slightly from the mature subunit. This is consistent with a recently proposed model suggesting that folding of the different subunit domains occurs independently (Sykes & Williamson, 2009) and that they are properly oriented at late stages of assembly by the formation of the central pseudoknot involving helix 2 and connecting the three major structural elements of the small subunit (Poot et al., 1998; Holmes & Culver, 2004).

It is important to note that the nearby densities in the overlaid difference maps did not always match in size and shape, since parts of these densities were canceled out in the difference maps due to partial overlapping of the moved domains between the two

structures. A proximity criterion was used to pair the densities between the forward and reverse difference maps (Figure 2.4). The assignment of the pair 1 and 1* (Figure 2.4, upper left panel) defining the movement of helix 44 between both structures was unambiguous corroborating the distortion of this helix in the immature 30S subunit structure. Conversely, an alternative assignment of pairs was possible for some of the other densities. Nevertheless, a different assignment does not change our conclusion that the relative orientation of the major structural domains in the immature subunit differs slightly from the mature subunit.

The forward map failed to reveal any large densities that could be assigned to the 115 and 33 nucleotides of the 5' and 3' precursor sequences, respectively. Interestingly, we only observed a small density very close to the 5' mature end of the rRNA that probably represents only the first few residues of the 5' precursor sequence (Figure 2.4C and D; density labeled as PS). Concerning the 3' precursor sequence, the outward displacement of helix 44 probably displaced the helix 45 and the extra 33 nucleotides of the precursor sequence into the interface region where it most likely adopted a flexible conformation. These results suggest that the precursor sequences are flexible and occupy a variable position in the structure, resulting in a smearing of their densities during the three-dimensional reconstruction.

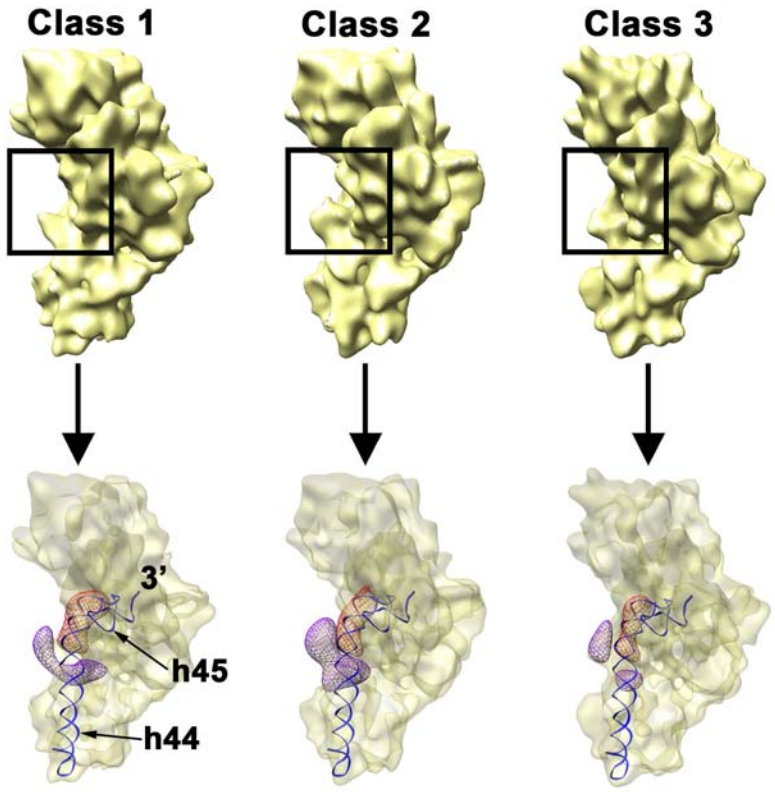


Figure 2.5 Conformational variability of the immature 30S subunit from Δ*yjeQ* cells.

Platform-view of the three cryo-EM structures (labeled from class 1 to class 3) representing three subpopulations of immature 30S subunits present in ΔyjeQ cells (upper panel). These 3D reconstructions were produced from 13,447 (class 1), 12,570 (class2) and 11,931 (class 3) projections and refined to 14.1 Å (class 1), 15.4 Å (class2) and 14.7 Å (class 3) resolution. The framed region in each map highlights the area of helix 44 that is distorted in the three subpopulations. The bottom panel illustrates the degree of distortion of helix 44 in each cryo-EM map. The density corresponding to the upper segment of helix 44 in each conformational state is shown as a violet mesh. The displacement of helix 44 in each conformation is shown by comparison to the position of the helix 44 in the mature 30S subunit shown as a red mesh. Helix 44 (h44) and helix 45 (h45) in the mature structure are also shown for reference.

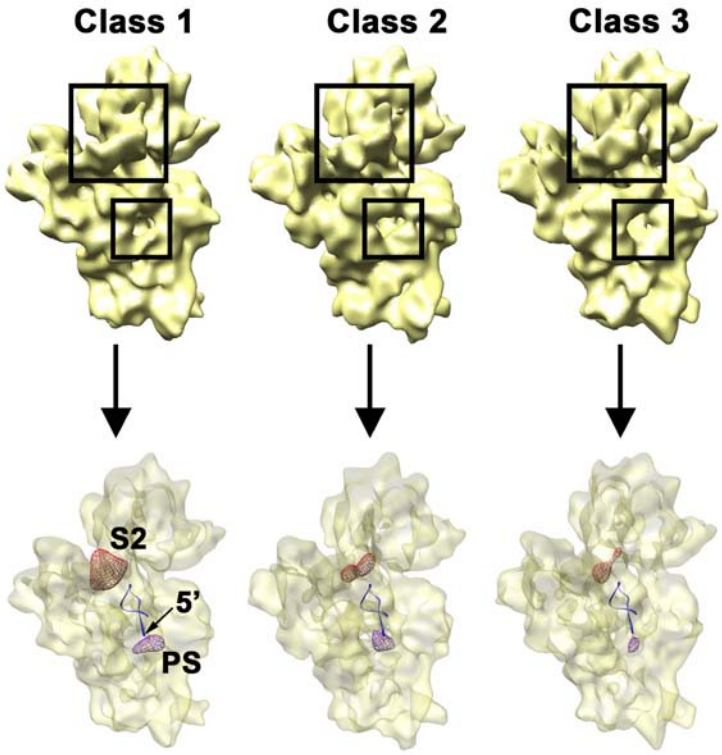


Figure 2.6 Variability of the density representing the S2 r-protein and 5' precursor sequences in the immature 30S subunit from Δ*yjeQ* cells.

Back-view of the three cryo-EM structures (labeled from class 1 to class 3) representing three subpopulations of immature 30S subunits present in ΔyjeQ cells (upper panel). The framed regions highlight the location of the S2 r-protein and the 5' mature end of the 16S rRNA. The dissimilar amount of density in these two locations in the cryo-EM maps is shown at the bottom panel. The density corresponding to the S2 r-protein (red mesh) from the reverse difference map (i.e. mature 30S subunit map minus immature 30S map) and the one corresponding to the 5' precursor sequence (violet mesh) from the forward difference map (i.e. immature 30S subunit map minus mature 30S map) for each cryo-EM map are displayed within a semitransparent surface of the 30S subunit structure. The first one indicates the amount of density corresponding to the S2 r-protein missing in each immature structure. Conversely, the second one shows the additional density existing in each structure due to the presence of the 5' precursor sequence. The first twenty nucleotides of the 5' end of the mature 16S rRNA are shown as a ribbon representation.

In addition, we expected to observe a density co-localizing with the S1, S21 and S2 r-proteins in the reverse map since they were found to be significantly underrepresented in the immature 30S subunits (Figure 2.2). However, only S2 showed a corresponding density in the difference map (Figure 2.4C and D; density labeled as S2). The size of this density was small compared to the size of the S2 protein suggesting that there were still a significant number of immature subunits that contain this r-protein. A matching density corresponding to S21 and S1 were not observed in the reverse map. S21 is a r-protein located between the head and platform and probably not observed because the small size of this protein. S1, when present in the 30S subunit, fills the cleft region between the head and platform (Sengupta et al., 2001). However, a corresponding density was not observed in the EM maps of the immature and mature 30S subunits. This result was expected because the high salt conditions we used to purify the 30S subunits lead to a loss of the S1 protein. The small Δ*yjeQ*:WT ratio found for S1 in the iTRAQ experiment (Figure 2.2 and Supplemental Table 2.1) may reflect the fact that our purification conditions eliminated the S1 protein even more efficiently in the immature 30S subunits.

In summary, the difference map analysis presented here indicated that the relative orientation of the structural domains in the immature 30S subunit diverged somewhat from that of the mature subunit. Most importantly, they revealed that the immature 30S subunits did not constitute a structurally homogeneous population.

2.4.5 Helix 44 in the immature 30S subunit moves inwards into its correct position during late stages of the maturation process

In order to structurally characterize the subpopulations of immature 30S subunits purified from Δ*yjeQ* cells, we further classified the projections that produced the 3D reconstruction of the immature 30S subunit (Figure 2.3, upper right panel) using a maximum likelihood-based classification technique (Scheres et al., 2005b; Scheres et al., 2005c; Scheres et al., 2007) (Supplemental Figure S2.5).

This procedure yielded three cryo-EM structures (Figure 2.5 and 2.6). The three EM maps were similar to the previously calculated 3D reconstruction of the immature 30S subunit (Figure 2.3, upper right panel). Difference map analysis of each individual map with respect to the mature structure from wild type cells showed that the relative orientation of the major structural domains in these three structures also diverged somewhat from that of the mature subunit (data not shown). However, they contained varying degrees of distortion in helix 44 (Figure 2.5) and differed in the amount of density associated with the S2 r-protein in the neck of the subunit (Figure 2.6). Conversely, a control classification of all the projections collected from the 30S subunit fraction purified from wild type cells was performed using the same approach and rendered three very similar 3D reconstructions, all resembling the mature 30S subunit structure (Supplemental Figure S2.6).

Regarding the 3D reconstructions produced by the classification of the immature particles from Δ*yjeQ* cells, one of the maps presented the most severe distortion with

respect to helix 44 (Figure 2.5, upper panel, class 1 map, framed area) and showed this structural element protruding outward from the interfacial surface of the subunit. A second map presented a less severe distortion (Figure 2.5, upper panel, class 3 map, framed area). The helix was still displaced outward but closer to the location in the mature structure. Finally, the third map showed an intermediate degree of distortion between the two previous structures (Figure 2.5, upper panel, class 3 map, framed area). To observe the degree of displacement of helix 44 in these three structures, the forward and reverse difference maps from each structure were overlaid and the corresponding densities for helix 44 were displayed (Figure 2.5, bottom panel). This analysis revealed a progressive movement of helix 44 into correct position indicated by the decrease in size of the corresponding densities in the difference maps.

The density representing the S2 r-protein in the three maps of the immature 30S subunit was also variable (Figure 2.6, large framed area). Interestingly, the map showing the largest distortion on helix 44 contained the least amount of density in this area (Figure 2.6, class 1 map) indicating that the occupancy of S2 in this subpopulation was significantly decreased. This group of particles constitutes in fact, the subpopulation causing the relative amount of S2 r-protein to be significantly underrepresented in the iTRAQ experiments (see above). Conversely, the EM map with helix 44 closer to the mature position showed comparable S2 density to the mature structure (Figure 2.6, class 3 map). This result is consistent with our interpretation that this last structure represents an immature 30S subunit but it is closer to the mature state. In agreement with this statement is the finding that the amount of density that we have assigned to the 5' precursor

sequence in these structures (Figure 2.5, small framed area) was largest in the EM map with the most severe distortion in helix 44 and lowest S2 density (class 1 map) and smaller in the structure with helix 44 and S2 densities similar to the mature particle (class 3 map).

In addition, one finding of the rRNA analysis performed on the Δ*yjeQ* cells was that apart from the 30S particles containing 17S rRNA, these cells also possessed degraded or aberrantly processed 16S rRNA, equivalent to approximately 10% of the total rRNA (Figure 2.1). We considered that these species may have been incorporated into 30S subunits. Interestingly, even when additional maximum likelihood-based classifications were performed assuming the existence of more than three classes in the data set, we consistently observed duplication in the obtained structures and it was possible to merge them into three, closely resembling the structures described above (data not shown). However, none of these classification experiments found a subpopulation that could be assigned to the particles containing the degraded rRNA. This result led us to conclude first, that mainly three subpopulations of immature 30S subunit accumulate in Δ*yjeQ* cells and second, that particles containing degraded rRNA may have diverged largely from the wild type 30S subunit conformation. It is possible that these particles might have been excluded during the initial step of particle image selection or alternatively, they may not represent a structurally congruent class and thus, they were eliminated during the 3D reconstruction process.

In sum, through our particle classification approach we have characterized three structures that highlight changing structural features spanning the immature and mature 30S subunits derived from a strain of *E. coli* that lacks the biogenesis factor YjeQ. Further, the observed correlations between the distortion in helix 44 and the amount of density in the EM maps associated with S2 and the 5' precursor sequence suggest that the processing of the rRNA precursor sequences, proper folding of the 3' minor domain of the subunit (helix 44 & 45) and the entry of tertiary proteins, are interdependent events in the late stages of 30S subunit assembly.

2.5 DISCUSSION

In the work reported here, we have made use of a viable deletion in the *yjeQ* gene in order to isolate and structurally characterize for the first time an *in vivo* assembled immature 30S subunit. Previous studies of 30S subunit assembly have largely been systematic *in vitro* investigations of rRNA folding and protein binding using processed 16S rRNA and r-proteins (Mizushima & Nomura, 1970; Talkington et al., 2005; Adilakshmi et al., 2008). These works are responsible for the remarkable understanding that we have today of 30S assembly, not the least of which is the hierarchy of protein binding described by the Nomura assembly map. Perhaps the most important principle to emerge from recent studies is that r-proteins drive conformational rearrangements in the RNA and stabilize the native fold. The *in vivo* approach followed here has raised some additional guiding principles. These include the idea that a specific lesion in ribosome assembly such as that achieved with a deletion in *yjeQ* can lead to an immature 30S subunit of profoundly distorted 3D structure, incapable of participating in either protein

synthesis or 50S binding. Further our results imply that rRNA processing events are key contributors to the final structure of the 30S subunit by defining a fold that is sterically capable of binding late subunit proteins.

We further conclude that ongoing structural studies will be critical to achieving precise understanding of the assembly process in three dimensions. To this end, as we demonstrate herein, emerging trans-acting factors in ribosomal subunit assembly represent an extraordinary opportunity to manipulate the assembly process *in vivo* in order to study subunit maturation in three dimensions using cryo-electron microscopy.

There have been a series of studies suggesting that the final rRNA processing events occur in polysomes (Mangiarotti et al., 1974; Hayes & Vasseur, 1976). In contrast, the structure of the immature 30S subunit presented here suggests that processing of the 3' precursor sequence of the 17S rRNA is a prerequisite to the formation of functional 70S ribosomes. Unfortunately, the precursor 16S rRNAs used in these early studies (Mangiarotti et al., 1974; Hayes & Vasseur, 1976) were not fully characterized and thus, it is unclear which of the precursor sequences where present in these precursor rRNA molecules. More recently, other studies have characterized precursor 16S rRNAs produced in *E. coli* strains containing mutations in the *rne* (encoding for RNase E) and *rng* genes (encoding for RNase G) and the incorporation of these precursor molecules into 70S ribosomes. One of the studies (Li et al., 1999b) found that a strain lacking *rng* combined with a temperature-sensitive allele of *rne* produces a precursor 16S rRNA with 115 extra nucleotides in the 5' end but its 3' end is fully processed. Interestingly, this

precursor 16S rRNA is incorporated into 70S ribosomes. A second report (Roy-Chaudhuri et al., 2010) found that a *rng* deletion strain produces two precursor 16S rRNA forms. These two forms contain 115 (lp 16S rRNA) and 66 (sp 16S rRNA) extra nucleotides in the 5' end, respectively but only the sp 16S rRNA is incorporated into the 70S ribosomes. Unfortunately, in this later study (Roy-Chaudhuri et al., 2010) the presence of the precursor sequence in the 3' end of lp and sp 16S rRNA was not investigated. Therefore, in light of the available evidence, our cryo-EM structure of the immature 30S subunit is not in disagreement with the idea that 5' rRNA processing may occur in the context of the 70S ribosome or even later in the polysomes. However, we cannot rule out whether the 3' rRNA processing occurs in the 70S ribosome. Nevertheless, recent work in our laboratory showed that while the sp 16S rRNA in the *rng* deletion strain is processed at the 3' terminus, the lp 16S rRNA is not and still contains its precursor sequence in the 3'end (Vivian Leong; unpublished results). Consequently, we favor the model that processing of the 3' precursor sequence occurs in the 30S subunit prior to association with the 50S subunit.

The data presented here also suggest that YjeQ has an indirect role in processing the precursor 16S rRNA, either by directly facilitating the interaction of the 30S subunit with these RNAases or by stabilizing the 30S subunit in a specific conformation that ensures efficient processing. Most likely YjeQ is mediating this process in conjunction with other trans-acting factors. A recent publication (Goto et al., 2011) has reported a functional interplay between YjeQ and RbfA during maturation of the 30S subunit. RbfA is a small protein that binds the 30S subunit at the junction of the head and body and its

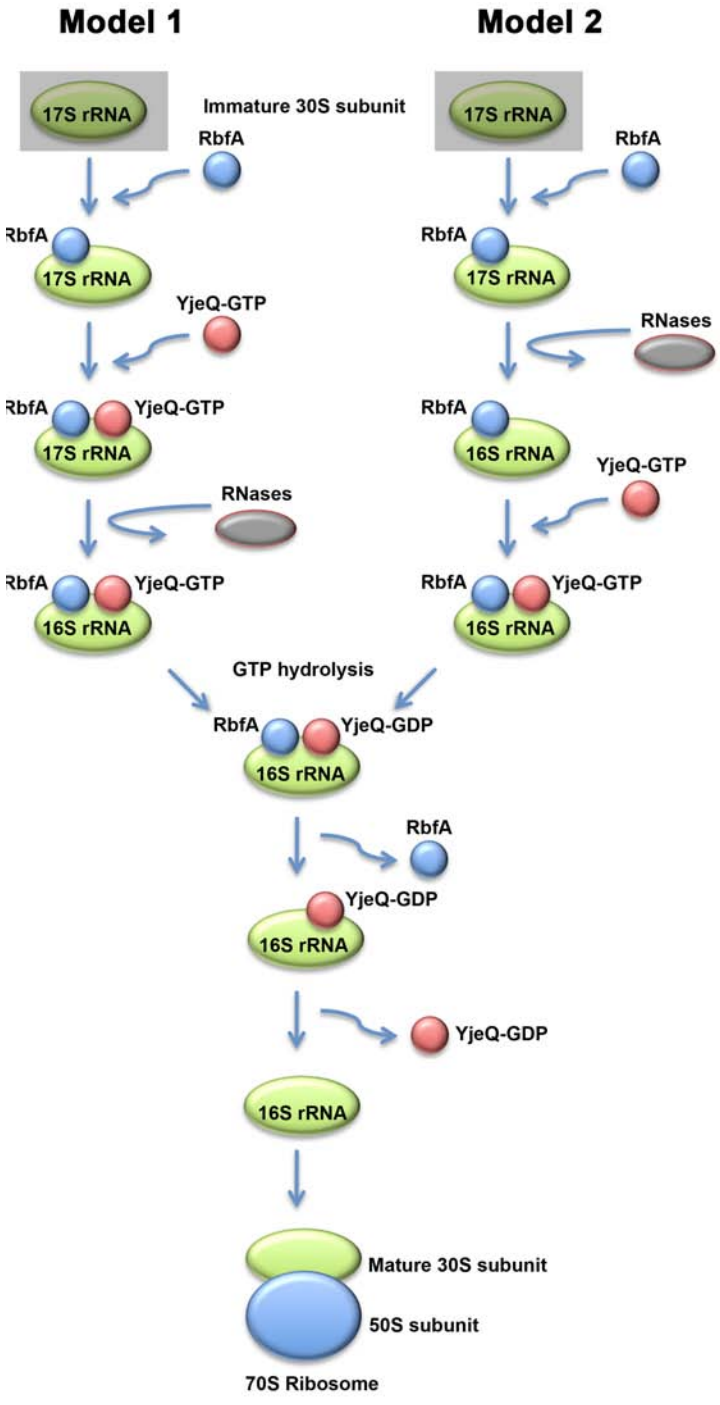
binding alters the position of helix 44 and 45 (Datta et al., 2007). Interestingly, RbfA seems to promote the association of YjeQ to the 30S subunits and binding of YjeQ stimulates the dissociation of RbfA from the small subunit (Goto et al., 2011).

Taken together, this and previous studies of YjeQ appear consistent with the following events in the late stages of 30S assembly (Figure 7, model 1): 1) RbfA binds first to an immature form of the 30S subunit; 2) this facilitates the entry of YjeQ-GTP into the complex either directly or after inducing a conformation that is then recognized by YjeQ; 3) this binding causes and additional conformational change in the immature particle; 4) the 30S subunit is recognized by RNases that process the 17S rRNA; 5) subsequent hydrolysis of GTP facilitates dissociation of RbfA and 6) YjeQ-GDP dissociates from the mature 30S subunit that is now in a conformation compatible for binding to the 50S subunit. Under this model, both RbfA and YjeQ would trigger specific conformations recognized by the RNases or directly recruit these enzymes to the maturing subunit. This order of events is somewhat different from that proposed recently by Goto and colleagues (Goto et al., 2011) where it was suggested that processing of the 17S rRNA is triggered by RbfA before YjeQ binding. In our view, the latter model seems inconsistent with the accumulation of 17S rRNA containing 30S subunits in the Δyj*eQ* strain, a finding that is reinforced by the data presented herein. Nevertheless, this accumulation could be explained as a consequence of impaired protein synthesis due to a shortage of mature ribosomes in these cells. Indeed, it has been reported that disturbances in protein synthesis arising from exposure to a 50S subunit-targeting antibiotic, such as puromycin, erythromycin or chloramphenicol also induce accumulation of 17S rRNA

(Hosokawa & Nomura, 1965; Himeno et al., 2004; Siibak et al., 2009). Thus it is conceivable that the structure of the immature 30S subunit purified from the Δ*yjeQ* strain may not represent the binding substrate for YjeQ but rather for RbfA. In any case, this cryo-EM reconstruction constitutes the first detailed structure of an immature 30S subunit prior to rRNA processing and containing 17S rRNA.

In addition to *yjeQ* and *rbfA*, there are several other genes encoding trans-acting factors that appear to have a role in assembly of the 30S subunit including *era*, *rimM*, and *rimN* (Wilson & Nierhaus, 2007; Connolly & Culver, 2009) with which YjeQ may also exert a functional interplay. *era* is of particular interest because of the genetic interaction of this gene with *yjeQ*: overexpression of *era* suppresses defects in growth and ribosome maturation of a yjeQ-null mutant (Campbell & Brown, 2008). Interestingly, a cryo-EM reconstruction of *Thermus thermophilus* Era protein bound to the mature 30S subunit reveals that this protein binds between the head and platform in a cavity formed by r-proteins S2, S7, S11 and S18 (Sharma et al., 2005). This binding site overlaps with the most severely distorted domain in the 30S subunit from cells lacking YjeQ. Perhaps Era performs its function in conjunction with YjeQ or YjeQ binds the 30S subunit in a location overlapping with the Era binding site. Additional structural, genetic and biochemical studies of YjeQ and other assembly factors will ultimately describe the existing functional relationships between these factors. The characterization of immature subunits generated with genetic lesions in these factors will be an important approach in on-going measures to understand bacterial ribosome biogenesis.

Finally, an important question in our present study is whether the obtained structures of immature 30S subunits constitute genuine intermediates of the ribosome assembly process or off pathway assembly species generated in cells lacking the assembly factor YjeQ. In this context it is worth considering a recent *in vitro* study (Talkington et al., 2005) suggesting that 30S subunit assembly does not proceed through a single pathway. Instead, there may be many parallel folding pathways that converge in the mature 30S subunit and thus, it is also unclear whether the characterized structures may represent intermediates within the same or different assembly pathway. The *yjeQ* gene is dispensable to cells and its loss leads not to cell death but to slow growth (Campbell et al., 2005). Perhaps, the slow growth phenotype originates from a slower uncatalyzed transformation of maturing species in the late stages of ribosome assembly. In this instance, we may well have characterized *bona fide* intermediates that accumulate in cells lacking YjeQ. Conversely, it is conceivable that YjeQ is a critical catalyst in a particular assembly pathway and that we have characterized off-pathway end products accumulated by the cell. Nevertheless, structural characterization of such species in ribosome assembly will surely also prove to be ultimately informative. Indeed, enzymologists have long endeavored to understand the structure of short-lived reaction intermediates by isolating and characterizing the by-products thereof. Thus while we cannot be definitive about the nature of the immature 30S subunits isolated and characterized in this work, these structures provide new and exciting insights into the process of ribosomal subunit assembly.


Model 1
Model 2
17S rRNA
Immature 30S subunit
RbfA
YjeQ-GTP
RNases
16S rRNA
GTP hydrolysis
YjeQ-GDP
Mature 30S subunit
50S subunit
70S Ribosome

continued

Figure 2.7 Placement of the structurally characterized immature 30S subunits in the current biogenesis model of the 30S subunit.

The diagram describes the likely order of events involving YjeQ in the late stages of assembly of the 30S subunit and processing of the 17S rRNA. A recent publication (Goto et al., 2011) has suggested that YjeQ mediate this process in conjunction with RbfA and thus, this protein is included in the diagram. RbfA seems to bind first to an immature form of the 30S subunit facilitating the subsequent entry of YjeQ-GTP to the maturing subunit. At this moment, there is not enough experimental evidence to establish whether the processing of the 17S rRNA occurs before (model 2) or after (model 1) YjeQ binding to the immature 30S subunit. The presented cryo-EM reconstruction of the immature 30S subunit purified from the ΔyjeQ strain constitutes the structure of a 17S rRNA-containing 30S subunit prior to rRNA processing and could possibly represent the structure of the immature 30S subunit bound by RbfA (shadowed squares in Model 1 and 2).

2.6 MATERIAL AND METHODS

2.6.1 Isolation and analysis of cellular rRNA

Overnight cultures of *E. coli* Δ*yjeQ* (EB1140) and the parental strain (EB334) were sub-cultured in LB media (1:100 dilution) and grown with shaking at 37 °C to an optical OD_{600} of 0.2. Subsequently, 1 ml of culture was pelleted by centrifugation at 5,000 x *g* for 10 minutes in an Eppendorf 5424 microcentrifuge. Cell pellets were resuspended in 100 µl of lysis buffer (Tris-HCl pH 8.0, 10 mg/ml lysozyme, 1 mg/ml proteinase K, 0.15 mM calcium acetate). Lysis mixtures were incubated at 22°C for 10 minutes with shaking. Ribosomal RNA was then purified from cells using RNeasy® Mini Kit (Qiagen) according to manufacturer's protocols. Purified RNA was incubated with an equal volume of RNA loading buffer (50% w/v urea, 10% w/v sucrose, 1x TBE) for 10 minutes at 75°C. Samples were loaded onto a modified agarose gel comprised of 0.7% agarose and 0.9% Synergel™ (Diversified Biotech, Boston, MA). RNA was separated by

electrophoresis for 4 hours at 4.5 V/cm. Gels were stained with ethidium bromide and visualized under UV light.

Northern hybridization analysis was performed for both the 5' and 3' precursor sequences of the 17S rRNA to determine its identity. Samples were transferred to Hybond-N^+ charged nylon membranes (GE Healthcare) under alkaline conditions (0.1 M NaOH, 3M NaCl) using the standard capillary method. Membranes were incubated at 32 °C in 50 ml of prehybridization buffer (0.5 M soidum phosphate ph 7.2, 7% (w/v) SDS, 1 mM EDTA pH7.0) for two hours. Single stranded DNA probes complimentary to the 5' or 3' terminal precursor sequences (Sigma Canada) were labeled with ^{32}P using T4 polynucleotide kinase (Fermentas). The sequence of the 5' probe was 5'-TTAAGAATCCGTATCTTCGAGTGCCCACA-3' and the sequence of the 3' probe was 5'-TGTGTGAGCACTACAAAGTACGCTTCTTTAAGGTAAGG-3.'

Labeled probes (20μg) were added to the prehybridization solutions and incubated with rotation for 20 hours at 32 °C. Subsequent washes with 1 x SSC, 0.5 x SSC and 0.1 x SSC in 1%SDS were performed at 50 °C for 15 minutes each. Blots were visualized using a Typhoon™ Trio+ phosphorimager (GE Healthcare). Northern blot analysis was also performed with a single stranded DNA probe complimentary to an internal fragment of 16S RNA as a control. The sequence of this probe was as follows: 5'-GGGCCATGATGACTTGACGTCATCCCCACC-3'. The method employed was identical to that used for the precursor probes.

2.6.2 Purification of 30S ribosomal subunits

Purified 30S subunits from *Escherichia coli* Δ*yjeQ* (EB1140) and BW25113 parental (EB334) strains were prepared using centrifugations over sucrose cushions and gradients as previously described (Daigle & Brown, 2004). For each strain, one liter of Luria broth (LB) was inoculated with 10 ml of a saturated overnight culture. Cultures were grown at 37 ºC to an OD_{600} of 0.2. Cells were cooled down to 4 ºC and all subsequent steps were conducted at this temperature.

Cultures were harvested by centrifugation at 8,500 x *g* for 15 min and the cell pellet resuspended in buffer A (20 mM Tris-HCl pH 7.5, 10.5 mM magnesium acetate, 100 mM NH_4Cl, 0.5 mM EDTA and 3 mM 2-mercaptoethanol). Cell lysis was performed by three consecutive passes of the cell suspension through a French pressure cell at 20,000 lb/in^2. The cell lysate was spun at 30,000 x *g* for 45 min to clear cell debris. Recovered supernatant (S30 fraction) was overlaid on an equal volume of 1.1 M sucrose cushion made up in buffer B (20 mM Tris-HCl pH 7.5, 10.5 mM magnesium acetate, 500 mM NH_4Cl, 0.5 mM EDTA and 3 mM 2-mercaptoethanol) and centrifuged at 100,000 x *g* for 16 hours. The ribosomal pellet was gently washed and sucrose removed by resuspension in buffer C (10 mM Tris-HCl pH 7.5, 10.5 mM magnesium acetate, 100 mM NH_4Cl, 0.5 mM EDTA and 7 mM 2-mercaptoethanol). The crude ribosomes were pelleted at 100,000 x *g* for 16 hours.

To obtain the 30S fraction from the *E. coli* Δ*yjeQ* strain, the crude ribosome pellet was resuspended in buffer E (10 mM Tris-HCl pH 7.5, 10 mM magnesium acetate, 60

mM NH_4Cl, 3 mM 2-mercaptoethanol) (non-dissociating conditions). In the case of the 30S fraction from the parental strain, the crude ribosome pellet was resuspended in buffer F (10 mM Tris-Hcl pH 7.5, 1.1 mM magnesium acetate, 60 mM NH_4Cl, 0.5 mM EDTA, and 2 mM 2-mercaptoethanol) (dissociating conditions). In both cases, a portion of the subunit suspension (50-60 A_{260} units) was layered onto a 32 ml 10-30% (wt/vol) sucrose gradient made up in buffer E (for non-dissociating conditions) or F (for dissociating conditions) and centrifuged at 43,000 x *g* for 16 hours using a Beckman SW32 Ti rotor. Gradients were fractionated using an AKTAprime™ purification system (GE Healthcare) and the elution peaks corresponding to 30S, 50S and 70S (in the case of non-dissociating conditions) particles peaks were monitored by absorbance at A_{260}. The 30S ribosomal subunits were then recovered by centrifugation at 100,000x *g* for 16 hours and the pellet was resuspended in Buffer E (10 mM Tris-HCl pH 7.5, 10 mM magnesium acetate, 60 mM NH_4Cl, 3 mM 2-mercaptoethanol), and stored at -80°C until further use. Quantification of the 30S subunits was accomplished by absorbance at 260 nm (1 A260 unit is equivalent to 69 pmol of 30S).

2.6.3 iTRAQ analysis

Biological replicates of 30S particles from *Escherichia coli* Δ*yjeQ* (EB1140) and BW25113 parental (EB334) strains were purified separately using non-dissociating conditions and analyzed independently. Total protein was measured using a 2D Quant kit (GE Healthcare). A total of 25 mg from each 30S subunit preparation was precipitated with 80% ice-cold acetone and washed twice with 100% acetone. The pellet was resuspended in 21 ml 0.5 M triethylammonium bicarbonate (TEAB), 0.01% SDS and

proteins were reduced with tris(2-carboxyethyl)phosphine (TCEP), cysteines were blocked with methylmethanethiosulfate (MMTS) and digested with trypsin as described previously (Jiang et al., 2006). Labeling with the iTRAQ reagents was according to the manufacturer's recommendations (Applied Biosystems). For each biological replicate, the peptides derived from the parental and Δ*yjeQ* 30S subunits were labeled with two different iTRAQ reagents. The iTRAQ reagents 115 and 117 were used to label tryptic peptides from independently isolated Δ*yjeQ* 30S subunits preparations and 114 and 116 were used to label tryptic peptides from independently isolated parental 30S subunits. The iTRAQ reagent was reconstituted in ethanol and added directly to the protein digest (70% ethanol, final concentration). The mixture was incubated at room temperature for 1 hour and the reaction was stopped by adding 250 ml of 0.1% trifluoroacetic acid (TFA, Fisher Sci, USA). Reaction mixtures from both labeling experiments were combined, dried by lyophilization and stored at -20°C.

iTRAQ-labeled peptides were resuspended in buffer A (5 mM KH_2PO_4 pH 2.7, 30% Acetonitrile) and applied to a strong cation exchange chromatography (SCX) using a Paradigm MG4 HPLC system (Michrom BioResources, Inc. Auburn, CA, USA) on a packed Polysulfoethyl aspartaminde column; 200 × 2.1 mm; 5 μm; 200 Å (Poly LC, Inc., Columbia, MD, USA). Peptides were eluted using a linear gradient with B (buffer A with 500 mM KCl) using an flow rate of 200 μl/min over 50 min. Two-min fractions were collected throughout and pooled into twelve fractions. For RP-HPLC-MS analysis, fractions were dried and redissolved in 45 ml of 0.1% TFA, desalted and concentrated on a reversed-phase cartridge (Zorbax C18) calibrated with 0.1% TFA as described

previously (Chen et al., 2006). Peptides were eluted with a linear gradient using 0.1% TFA with 90% acetonitrile using a flow rate of 4 ml/min for 90 min and spotted at 20 seconds intervals onto a stainless steel MALDI target plate (192 wells; Applied Biosystems).

The MS and MS/MS spectra were acquired on an Applied Biosystems 4800 Proteomics Analyzer (TOF/TOF) (Applied Biosystems/MDX Sciex, Foster City, CA) in positive ion reflection mode (selected mass range of 800-3000 *m/z*) with a 200 Hz Nd:YAG laser operating at 355 nm. Accelerating voltage was 20 kV with 450 ns delay. For MS/MS spectra, the collision energy was 6 keV and the collision gas was air. The instrument was calibrated using a default mixture of 7 peptides. Typical MS spectra were obtained with the minimum possible laser energy in order to maintain the best resolution. MS/MS spectra were collected for up to 12 of most intense MS peaks (minimum signal-to-noise ratio of 60) from each well.

The peak lists were generated from the acquired MS/MS spectra with the Protein Pilot software (Applied Biosystem, version 3.0) using default parameters. Acquired MS/MS spectra were searched against a concatenated (forward and reverse) *E. coli* K12 protein sequence database downloaded as a FASTA file from UniProtKB/Swiss-Prot (http://www.uniprot.org/taxonomy/complete-proteomes) on April 10th, 2008. The Protein Pilot software with default settings and the Paragon search engine (Applied Biosystems, version 3.0) was the software used to perform all the database searches. Proteins were identified with 1% local false discovery rate calculated by a module built-in the software.

Peptides for which one or more of the iTRAQ values were absent and proteins identified by only a single distinct peptide were discarded. Peak areas for all of the reporter ions (m/z 114.1, 115.1, 116.1, and 117.1) were extracted and automatically corrected using Protein Pilot (Applied Biosystems, version 3.0). The MS/MS data obtained were deposited into TRANCHE (https://proteomecommons.org/tranche/) with the Tranche Hash:

dDZes+9slItFdCPmF8YepUC7GZEGcVekFr284mjTK4lsoEtAC1HZThAr5WcXUfce+J
eZ1BYon2HV3tLR42W0aBiwCFsAAAAAAGmUKQ==.

The Δ*yjeQ*:WT (wild type) and WT: Δ*yjeQ* protein ratios as well as the p-values in Supplemental Table 2.1 were generated with the Protein Pilot Software. The ratios in Figure 2 and Supplemental Figure S2.1 were generated manually, as we have done previously (Jiang et al., 2006; Jiang et al., 2007). In this case, only peptides that identify proteins with ≥95% confidence were used for the calculation. Peptide ratios for each replicate (115/114 and 117/116) were combined and then averaged. To obtain protein ratios, the average from all peptide ratios was obtained after removal of outliers using a Grubbs test (http://www.graphpad.com/quickcalcs/Grubbs1.cfm). The entire dataset was normalized at the level of the proteins by adjusting the median of the protein ratio to 1:1.

2.6.4 Cryo-electron microscopy, image classification and three-dimensional reconstruction

For cryo-EM, holey carbon grids (400 mesh copper) containing an additional continuous thin (5-10 nm) layer of carbon were previously washed with acetone vapor for

15 min and glow discharge in air for 30 seconds (Aebi & Pollard, 1987). Then, 3.5 µL aliquots containing 30S subunits purified from Δ*yjeQ* or wild type cells were applied to the grid for one min. The 30S subunits were in Buffer E and at a concentration of ~60 nM. Grids were blotted for 7 s and vitrified by rapidly plunging into liquid ethane at -180 ºC with a Vitrobot™ (FEI). Data acquisition was performed under low dose conditions (~10 $e^-/Å^2$) on a JEOL 2010F FEG microscope operated at 200 kV with a Gatan 914 side-entry cryo-holder and at a nominal magnification of 50,000x. A total number of 72 and 29 electron micrographs were collected from the samples containing 30S subunits purified from Δ*yjeQ* and wild type cells, respectively. The defocus range of the images was from -0.65 to -3.9 mm. The micrographs were digitized with a step size of 12.7 µm in a Nikon Supercool Scan 9000 producing images with a sampling value of 2.54 Å/pixel.

Projections representing 30S subunits purified from Δ*yjeQ* and wild type cells were picked using Boxer (Ludtke et al., 1999). Contrast transfer function of the micrographs was estimated using CTFFIND software (Mindell & Grigorieff, 2003) and corrected using Xmipp software package (Scheres et al., 2008). Image classification was performed using supervised (Valle et al., 2002; Gao et al., 2004) and 3D maximum likelihood-based (Scheres et al., 2005c; Scheres et al., 2007) methods. Detailed procedures regarding these methods are included in Supplemental Figure S2.3 and S2.5.

The 3D reconstructions were calculated using 3D projection alignment procedures as implemented in the Xmipp software package (Scheres et al., 2008). The reference map used to refine the 3D reconstruction of both the mature and immature 30S subunit was the

X-ray structure of the *E. coli* 30S subunit (PDB ID: 2Z4K) (Borovinskaya et al., 2007) low-pass filtered to 25Å resolution. In each refinement, sets of projections were calculated from the reference map using an angular spacing that ranged from 15º to 2º during the multiple cycles of refinement. The correct handedness of the structures was imposed by the X-ray crystal structure of the 30S subunit from *E. coli* (PDB ID: 2Z4K). Resolution of the cryo-EM maps was estimated by calculating two maps following the last cycle of refinement from the even and odd numbered particles. Then Fourier shell correlation was calculated between both maps and the resolution was estimated using a FSC value of 0.5. These values were used to low-pass filter the refined cryo-EM maps.

Difference maps between the immature and mature 30S subunit structures were calculated using Xmipp software package (Scheres et al., 2008). To this end, maps were subtracted after alignment of the two structures and normalization to a comparable grayscale.

2.6.5 Docking of X-ray crystallographic structures and structure visualization

Fitting of the X-ray crystallographic structure of the *E. coli* 30S subunit (PDB ID: 2Z4K) (Borovinskaya et al., 2007) was performed using rigid-body fitting of the entire X-ray structure of the 30S subunit onto the EM maps as implemented in the Situs software (Wriggers et al., 1999). Visualization of the fitted atomic structure and the cryo-EM density maps and difference maps was done with UCSF Chimera software

(Pettersen et al., 2004).

2.6.6 Accession Numbers

The EM maps of the immature 30S subunit purified from Δ*yjeQ E. coli* cells and the mature 30S subunit purified from wild type *E. coli* cells have been deposited in the Electron Microscopy Data Bank (EMDB IDs: 1774 and 1775).

2.7 ACKNOWLEDGEMENTS

We are grateful to the staff at the Canadian Centre for Electron Microscopy. We also acknowledge Vivian Leong for technical support, Sjors Scheres for assistance with the maximum likelihood-based classification programs implemented in Xmipp and the Supercomputing Center of Galicia (CESGA) for the use of their high-performance computers. JO is a Canadian Institutes of Health Research (CIHR) New Investigator and also acknowledges support from an Early Researcher Award from the Ministry of Research and Innovation. This work was supported by a grant from the National Institutes of Health (Grant # GM077628) to J.R.M. E.D.B is a Canada Research Chair and supported for this work by an operating grant from the Canadian Institutes of Health Research (MOP-6492). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

CHAPTER 3. Cryo-electron Microscopy Structure of the 30S Subunit in Complex with the YjeQ Biogenesis Factor

3.1 AUTHOR'S PREFACE

The data presented in chapter 2 has been published in the peer-reviewed *RNA Journal*, and appears in its published format. This work characterized the co-structure formed between the YjeQ protein and the 30S subunit. This work was carried in collaboration with Dr. Eric Brown, and Dr. Jason Mears. I have performed all cryo-EM microscopy experiments including sample preparation, data collection, and image processing. Geordie Stewart from the Brown group performed the binding assays. Dr. Jason Mears Performed the flexible docking of the x-ray coordinates into the cryo-EM maps. Dr. Inga Kireeva performed the mass spectrometry analysis on the S1 protein sample. Dr. Joaquin Ortega, Dr. Eric Brown, Geordie Stewart and I analyzed the data and wrote the manuscript.

The full citation is as follows:

Jomaa A, Stewart G, Mears JA, Kireeva I, Brown ED, Ortega J. (2011) Cryo-electron Microscopy Structure of the 30S Subunit in Complex with the YjeQ Biogenesis Factor. RNA. 2011 Nov;17(11):2026-38.

3.2 ABSTRACT

YjeQ is a protein broadly conserved in bacteria containing an N-terminal oligonucleotide/oligosaccharide fold (OB-fold) domain, a central GTPase domain and a C-terminal zinc-finger domain. YjeQ binds tightly and stoichiometrically to the 30S subunit, which stimulates its GTPase activity by160-fold. Despite growing evidence for the involvement of the YjeQ protein in bacterial 30S subunit assembly, the specific function and mechanism of this protein remains unclear. Here, we report the co-structure of YjeQ with the 30S subunit obtained by cryo-electron microscopy. The co-structure revealed that YjeQ interacts simultaneously with helix 44, the head and the platform of the 30S subunit. This binding location of YjeQ in the 30S subunit suggests a chaperone role in processing of the 3' end of the rRNA as well as in mediating the correct orientation of the main domains of the 30S subunit. In addition, YjeQ binding site partially overlaps with the interaction site of initiation factor 2 and 3 and upon binding, YjeQ covers three inter-subunit bridges that are important for the association of the 30S and 50S subunit. Hence, our structure suggests that YjeQ may assist in ribosome maturation by preventing premature formation of the translation initiation complex and association with the 50S subunit. Together, these results support a role for YjeQ in the late stages of 30S maturation.

3.3 INTRODUCTION

The 70S ribosome is comprised of two distinct subunits of unequal size (Ramakrishnan, 2002). The smaller subunit, termed 30S is formed by 21 ribosomal proteins (designated S1 to S21) and one RNA molecule, the 16S rRNA. The 50S, or large subunit is made of two RNA molecules, the 23S and 5S rRNAs and 34 proteins (Moazed et al., 1995). Despite the size and complexity of this macromolecule, *Escherichia coli* efficiently assembles the ribosome assisted by more than 50 trans-acting factors (Wilson & Nierhaus, 2007). These include RNases that perform the cleavage of the precursor rRNAs (Misra & Apirion, 1979; Li et al., 1999b, a) and factors executing rRNA and protein modifications (Connolly & Culver, 2009). In addition, there is a subset of factors for which identification of a particular functional role has been problematic. This list of 'enigmatic factors' includes the putative maturation factors RimM, RimP and RbfA as well as the highly conserved GTPases, Era and YjeQ (also known as RsgA) (Brown, 2005; Karbstein, 2007). All of these proteins have been shown to interact with the 30S subunit suggesting that they participate in late steps of 30S subunit assembly, facilitating the action of RNAses or the binding of late r-proteins (Shajani et al., 2011).

The YjeQ protein represents a subfamily of GTPases whose defining structural feature is a circular permutation of the GTPase domain (Leipe et al., 2002). The protein is broadly conserved in bacteria but absent in eukaryotes and has a unique architecture with an N-terminal oligonucleotide/oligosaccharide fold (OB-fold), a central GTPase domain with a classical P-loop fold, and a C-terminal TAZ2-like zinc-finger domain (Levdikov et al., 2004; Shin et al., 2004; Nichols et al., 2007). YjeQ is dispensable for growth but

critically important for the overall fitness of *E. coli*, *Bacillus subtilis*, and *Staphylococcus aureus* (Himeno et al., 2004; Campbell et al., 2005; Campbell et al., 2006).

YjeQ catalyzes the rapid hydrolysis of GTP (100 s^{-1}) with a low steady-state turnover rate, 9.4 h^{-1} (Daigle et al., 2002). Recombinant YjeQ binds tightly and stoichiometrically to the 30S subunit in the presence of a non-hydrolyzable GTP analogue, GMP-PNP. Binding to the 30S subunit stimulates the YjeQ GTPase activity by 160-fold (Daigle & Brown, 2004). Previous studies have suggested that YjeQ interacts with the 30S subunit through the A and P site in the decoding center of the ribosome. This is evidenced by two separate studies demonstrating that A-site targeting antibiotics such as neomycin, inhibit the ribosome stimulation of YjeQ GTPase activity (Himeno et al., 2004; Campbell et al., 2005). More recently, chemical footprinting studies (Kimura et al., 2008) suggested that binding of YjeQ to the 30S subunit induces conformational changes around the A site, P site and helix 44. Accordingly, it was further shown that YjeQ promotes dissociation of tRNA but has the capacity to coexist with mRNA in the 30S subunit (Kimura et al., 2008).

There is a growing body of evidence implicating a role for YjeQ in the late stages of 30S biogenesis. Fractionation of ribosomal subunits from strains lacking *yjeQ* revealed accumulation of free 30S and 50S subunits (Himeno et al., 2004; Campbell et al., 2005). Accumulated 30S subunits are nonfunctional and contain immature 17S rRNA species (Himeno et al., 2004; Jomaa et al., 2011a). A three-dimensional reconstruction by cryo-electron microscopy (cryo-EM) of these 30S subunits (Jomaa et al., 2011a) revealed that the presence of precursor sequences in the rRNA induces a severe distortion in the 3'

minor domain of the subunit, which is involved in the decoding of mRNA and interaction with the large ribosome subunit. These findings suggest that YjeQ has an indirect role in processing the precursor 16S rRNA, either by directly facilitating the recruitment of RNAses to the processing site or by stabilizing the 30S subunit in a specific conformation that ensures efficient processing. YjeQ is presumably not mediating this process by itself but in conjunction with other trans-acting factors including Era, RimM, RimP and RbfA (Bylund et al., 1998; Inoue et al., 2003; Inoue et al., 2006; Goto et al., 2011).

To further understand the involvement and mechanism of action of YjeQ in the late stages of the 30S subunit assembly, we obtained the co-structure of the 30S subunit in complex with YjeQ by cryo-EM. The 3D reconstruction shows the small ribosomal subunit with an additional density covering part of the platform domain and the upper region of helix 44. The atomic structure of the YjeQ protein fit unambiguously into this density. Contacts between the protein OB-fold domain and the 16S rRNA in the platform mediate the interaction between YjeQ and the 30S subunit. At the C-terminal end of the protein, the zinc-finger domain sits in the upper region of helix 44 inducing a displacement of this helix in the 30S subunit. In this location, YjeQ might be optimally placed to assist in late maturation events, including processing of the 3' end of the 17S rRNA and mediating the correct orientation of the main domains of the small subunit. In addition, YjeQ binding also covers three inter-subunit bridges in the 30S subunit that are important for its association with the 50S subunit. This finding structurally explains biochemical data indicating that YjeQ induces the dissociation of the 70S ribosome (Himeno et al., 2004).

3.4 RESULTS

3.4.1 Cryo-EM structure of the 30S subunit in complex with YjeQ

Attempting to maximize the occupancy level of YjeQ in the 30S:YjeQ complex, we found that a 5-fold excess of freshly purified YjeQ protein in the presence of GMP-PNP, incubated with a 1 μM preparation of 30S subunits purified from *Escherichia coli* produced an occupancy of ~70% (Figure 3.1). Higher excesses of YjeQ in the assembly reaction did not significantly increase the occupancy of the protein in the complex.

A diluted sample from this mixture was imaged using cryo-EM (Supplemental Figure S3.1). Dilution of the specimen was necessary to obtain a suitable concentration of 30S subunits for cryo-EM imaging. To guard against a decrease in the occupancy level of YjeQ in the 30S:YjeQ complex on dilution we experimented by including free YjeQ in the dilution buffer. A 5-fold molar excess of YjeQ protein with respect to the 30S subunits maximized the proportion of 30S:YjeQ complexes and minimized excessive background that came with higher excesses of YjeQ.

Cryo-EM images were collected and the two-dimensional projections in the electron micrographs were selected and subjected to a supervised classification approach with the aim of eliminating particles representing 30S subunits that were not bound to YjeQ (Figure 3.2). This classification approach produced four classes of particles from which corresponding 3D reconstructions were calculated. The SC1 class (Figure 3.2; bottom panel) produced a reconstruction with all of the recognizable landmarks of the 30S subunit and represents the particles that were not bound to YjeQ. The other three

classes, SC2-SC4, (Figure 3.2; bottom panel) produced a similar 3D reconstruction but with different amounts of additional density attached to the upper domain of helix 44 and the nearby region of the platform. The reconstruction from SC4 particles (Figure 3.2; bottom panel) showed the largest density and was produced from a subpopulation with the highest proportion of 30S subunits bound to YjeQ. This reconstruction represents the structure of the 30S:YjeQ complex (Figure 3.3A; left panel), was refined to 16.5 Å resolution (Supplemental Figure S3.2) and used in our subsequent analysis. Reconstructions SC2 and SC3 with smaller amounts of additional density represented mixed subpopulation containing both free 30S subunits and 30S:YjeQ complexes. These two classes suggested that the classification approach was effective in separating only a fraction of the 30S:YjeQ complexes present in our images. A second round of supervised classification using SC1 (as a free 30S subunit reference) and SC4 (as a reference for the 30S:YjeQ complex) did not improve the efficiency of the classification. Similarly, we also attempted to sort out our particles using a maximum likelihood-based classification technique (Scheres et al., 2005a; Scheres et al., 2005b; Scheres et al., 2007) but we did not obtain an improved separation of the two populations existing in the sample (data not shown).

In addition to a 3D reconstruction of the 30S:YjeQ complex, we also imaged free 30S subunits and a control cryo-EM map was calculated. This structure was used as one of the references in the supervised classification (Figure 3.2) and for the difference map analysis of the 30S:YjeQ structure (see below). In addition, the control cryo-EM map from free 30S subunits was a necessary aid for unequivocally assigning the extra density

in the cryo-EM map of the complex to the YjeQ protein. In particular, it was important to rule out that initiation factors 1 (IF1), 2 (IF2) and 3 (IF3) that bind in an area partially overlapping with the location of the additional density observed in the 30S:YjeQ complex (Moazed et al., 1995; Dallas & Noller, 2001; Allen et al., 2005) were not contaminating the purified 30S subunits used to assemble the complex. The presence of IF1 and IF2 in the 30S subunit preparation was discarded. iTRAQ (isobaric tag for relative and absolute quantification) analysis (Ross et al., 2004) of a sample of 30S subunits purified with the same procedure in a previous study from our laboratory (Jomaa et al., 2011a) revealed that IF1 and IF2 were not present in the 30S subunit preparation. However, the sample did contain initiation factor 3 (IF3). The interaction site of N-terminal domain of IF3 with the 30S subunit has been mapped near the E site and proteins S7 and S11 (Dallas & Noller, 2001), which is a different location than the one proposed here for binding of YjeQ. Differently, the C-terminal domain of IF3 interacts with helices 23, 24 and 45 (Dallas & Noller, 2001). This is the area of the platform covered by additional density in the cryo-EM map of the 30S:YjeQ complex. Nevertheless, the control cryo-EM map from free 30S subunits showed all the recognizable landmarks of the 30S subunit but did not show any additional density, including in the area occupied by the C-terminus of IF3. Therefore, the control reconstruction indicates that this factor in our preparation is either dissociated from the 30S subunits or in substoichiometric amounts.

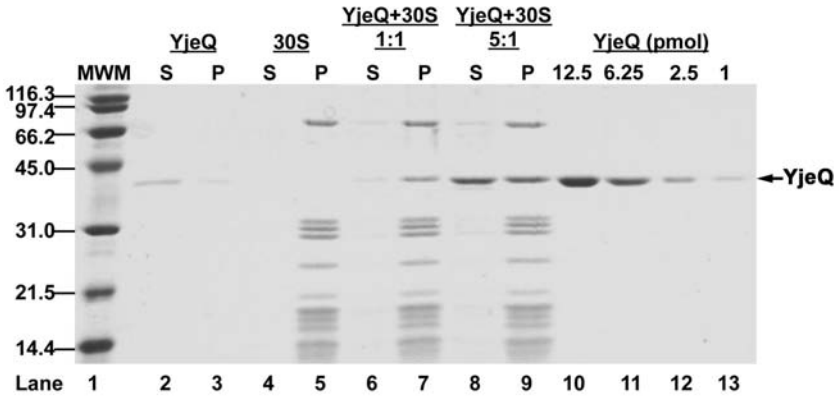


Figure 3.1 Assembly of the 30S:YjeQ complex.

Purified YjeQ was incubated with 30S subunits in ratios of 1:1 (lanes 6-7) and 5:1 (lanes 8-9) in the presence of 1 mM GMP-PNP. The concentration of 30S subunits in these reactions was 1 μM. Following incubation, samples were overlaid onto a sucrose cushion and pelleted by ultracentrifugation. Pelleted material was resuspended and both the supernatant (S) and pellet (P) were resolved by SDS-PAGE and stained with Coomassie brilliant blue. The lanes are as follows 1: molecular weight marker, 2-3: YjeQ + GMP-PNP, 4-5: 30S + GMP-PNP, 6-7: YjeQ + GMP-PNP + 30S (1:1), 8-9: YjeQ + GMP-PNP + 30S (5:1), 10-13: known quantities of YjeQ (1-12.5 pmol). The position of YjeQ in the gel is indicated. The occupancy of the YjeQ-30S complex was assessed by quantifying the amount of YjeQ that co-pelleted with ribosomes (2.5 pMol) in each reaction using the lanes with known quantities of YjeQ (lanes 10-13). For this calculation the fraction of YjeQ that pellets in the absence of ribosomes (25%) (lane 3) was subtracted from that quantified in the reaction pellet (lanes 7 and 9). Using the previously determined binding ratio of 1:1, occupancy for both 1:1 and 5:1 incubations was found to be approximately 70%.

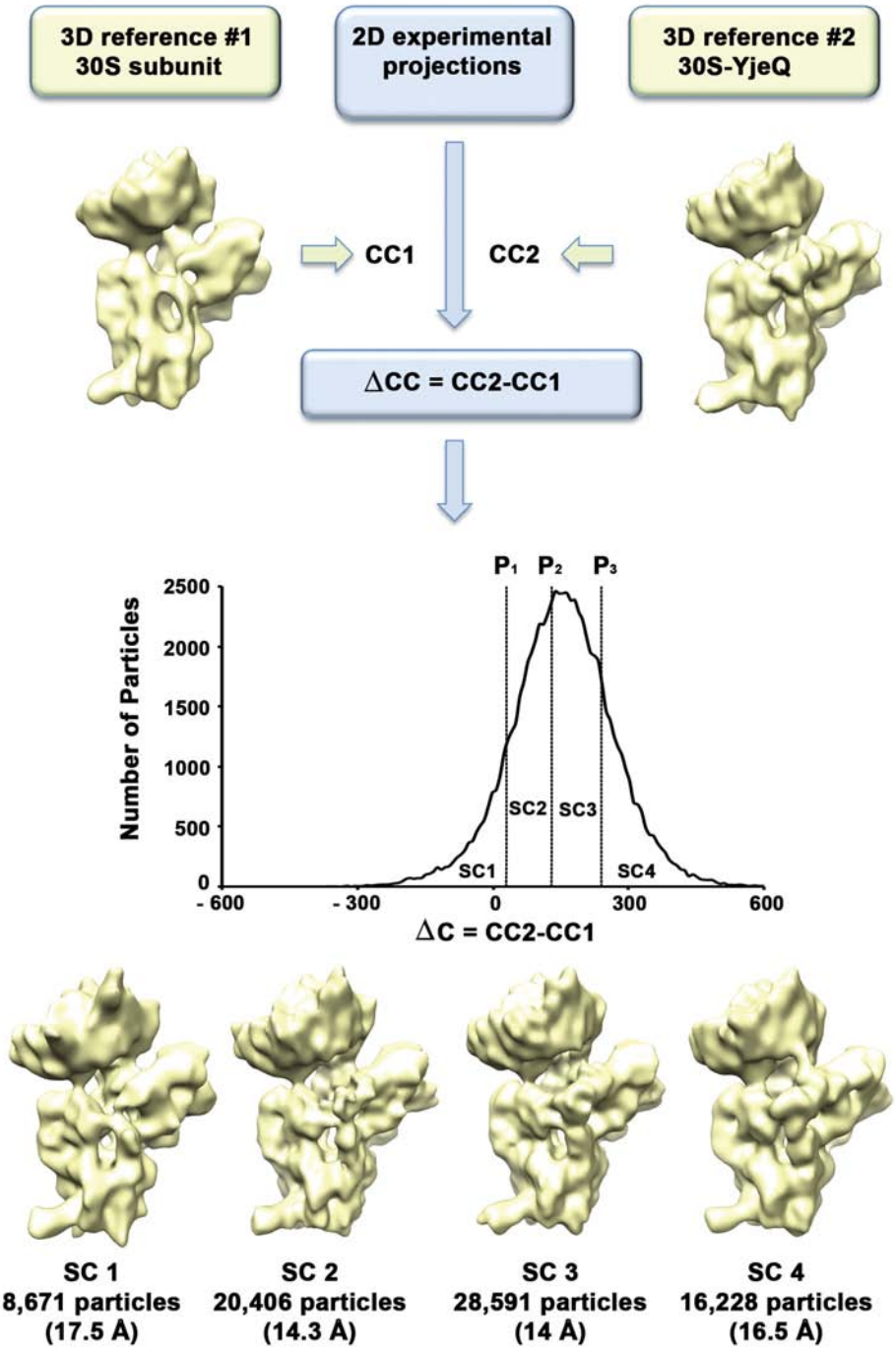
Under the high salt conditions used to purify the 30S subunits in our study, only small amounts of S1 protein were expected to remain associated to rRNA (Spedding, 1990). However, SDS-PAGE analysis of the purified 30S subunits used to assemble the 30S:YjeQ complex revealed a band at ~70 kDa (Figure 3.1) that we identified as the S1 protein by mass spectrometric analysis (see Material and Methods). When bound to the 30S subunit, S1 fills the cleft region between the head and platform (Sengupta et al., 2001) but its binding site does not overlap with the area occupied by the additional

density observed in the 30S:YjeQ complex. In any case, the control cryo-EM map obtained from free 30S subunits did not show any density in the area occupied by S1 (Figure 3.3A; right panel) suggesting that even though S1 is present in the preparation it probably dissociated from the 30S subunits during the cryo-EM sample preparation. This was not surprising, as it has been described that association of S1 to the small subunit is weak and reversible (Subramanian & van Duin, 1977). Taken together, these data enabled us to confidently attribute the additional density observed in the cryo-EM structure to the YjeQ protein.

3.4.2 Location of YjeQ in the structure of the 30S:YjeQ complex

The X-ray structures of YjeQ from *Salmonella typhimurium* (PDB ID 2RCN), *Thermotoga maritime* (PDB ID 1UOL) and *Aquifex aeolicus* (PDB ID 2YV5) have been determined (Shin et al., 2004; Nichols et al., 2007). These proteins share different degrees of sequence homology with the E. *coli* YjeQ used to form the 30S:YjeQ complex in our study. In particular, *S. typhymurium* YjeQ has a 92% sequence identity and a 94% sequence similarity with *E. coli* YjeQ (Supplemental Figure S3.3), whereas *A. aeolicus* and *T. maritime* only have 50% sequence similarity. In spite of their sequence diversity the structure of these three proteins is remarkably conserved (Supplemental Figure S3.4). The superimposition of the X-ray structure of *A. aeolicus* and *T. maritime* onto that of *S. typhymurium* returned root mean square deviations of 1.79 (244 residues superimposed) and 1.51 Å (247 residues superimposed), respectively. Nevertheless, due to the high degree of sequence homology the X-ray structure of *S. typhymurium* YjeQ was used for our docking experiments.

To locate the YjeQ protein in the complex structure, the cryo-EM map obtained as a control from free 30S subunits (Figure 3.3A; right panel) was subtracted from the cryo-EM structure of the 30S:YjeQ complex (Figure 3.3A; left panel). The resulting difference map showed a density clearly resembling the X-ray structure of YjeQ covering the upper domain of helix 44 and nearby region of the platform in the 30S subunit (Figure 3.3B). The *S. typhymurium* YjeQ atomic coordinates (Nichols et al., 2007) fit remarkably well and unambiguously into this density with a cross-correlation coefficient (CCC) value of 0.90. This was consistent with previous biochemical data indicating that only one molecule of YjeQ binds to the 30S subunit at a time (Daigle & Brown, 2004). Only small translational and rotational movements of each domain independently were necessary (Supplemental Figure S3.5) to optimize the fitting (Figure 3.3B). A CCC value of 0.91 was obtained after optimization. The finding that only small adjustments in the relative position of the YjeQ domains were necessary to optimize the fitting suggests that the three domains may not undergo a large rearrangement upon binding to the 30S subunit.


3D reference #1
30S subunit
2D experimental
projections
3D reference #2
30S-YjeQ
CC1
CC2
ΔCC = CC2-CC1
P1 P2 P3
Number of Particles
2500
2000
1500
1000
500
0
SC1
SC2
SC3
SC4
- 600
- 300
0
300
600
ΔC = CC2-CC1
SC 1
8,671 particles
(17.5 Å)
SC 2
20,406 particles
(14.3 Å)
SC 3
28,591 particles
(14 Å)
SC 4
16,228 particles
(16.5 Å)

continued

Figure 3.2 Supervised-classification of the projections representing the 30S:YjeQ complex.

Projections of 30S:YjeQ complexes were classified using two references: the control cryo-EM structure obtained from free 30S subunit low-pass filtered to 25Å resolution (reference #1) and the 3D reconstruction obtained from all the particle images in the data set (reference #2). Reference #2 already showed an extra mass of density attached to the upper domain of helix 44 and the nearby region of the platform. All images in the data set were aligned to the projections generated from each of the references and each image was assigned to the projection view from each reference that yielded the highest cross-correlation coefficient obtaining two cross-correlation coefficients per particle image (CC1 and CC2). Then, we plotted the distribution of particles against the difference between the two cross-correlation coefficients (DCC=CC2-CC1). The particles classified following a unimodal distribution. This histogram was divided into four classes (From SC1 to SC 4) using the P_x values indicated by the dash lines. The 3D map for each subset of particles was calculated using projection matching. The number of particles classified in each group is indicated in the figure. The reference map used to refine the four structures was the X-ray structure of the E. coli 30S subunit (PDB ID: 2AVY) (Schuwirth et al., 2005) low-pass filtered to 25Å resolution. The estimated resolution for each map is noted in the figure below the number of particles assigned to each class. The group of particles most similar to reference 2 (SC 4) formed the 3D cryo-EM map of the 30S:YjeQ complex.

The difference map analysis to locate the YjeQ protein in the 30S:YjeQ cryo-EM map was repeated using two other X-ray structures of the *E. coli* 30S subunit (PDB IDs: 2AW7 and 2AVY) (Schuwirth et al., 2005) filtered to match the resolution of our 30S:YjeQ cryo-EM map. These two X-ray structures only differ by a 6º rotation of the head around the neck region, however helix 44 and the platform are in the same conformation. The additional density isolated using these two difference maps was very similar to the one obtained when the cryo-EM map of the free 30S subunit (Figure 3.3; right panel) was used for the difference map analysis (data not shown). These results

suggested that our method to locate the YjeQ protein in the cryo-EM map of the complex was not biased by the 30S structure used for the difference map analysis.

Some small parts of the additional density covering helix 44 and platform of the 30S:YjeQ complex were not filled by the crystal structure of YjeQ. This may be due to the fact that some sequence motifs in YjeQ are not represented in the *S. typhymurium* YjeQ X-ray structure (Nichols et al., 2007). In particular, YjeQ in *S. typhymurium* contains a 42-amino acid N-terminal stretch that is not described by the atomic coordinates (Supplemental Figure S3.4; upper panel). This sequence is also present in *E. coli*, although it is a slightly shorter sequence of 34-amino acids (Supplemental Figure S3.3). Interestingly, the additional density in the 30S:YjeQ complex extends beyond the docked N-terminal OB-fold domain of YjeQ towards the tip of the platform (Figure 3.3B; density labeled as 1). This additional density may well encompass these residues. Another void in the density attributed to YjeQ was proximal to the GTPase domain and extended upward, contacting helix 31 and the C-terminus of S13 in the head of the subunit (Figure 3.3B; density labeled as 2). Here again, residues Asp 246 to Thr 258 (Asp 238 to Thr 250 in *E. coli* YjeQ) are not described in the X-ray structure of *S. typhymurium* YjeQ (Supplemental Figure S3.3 and S3.4) and may be filling this void (Figure 3.3B; density labeled as 2).

We noticed that the first and last a helices of the C-terminal zinc-finger domain of YjeQ protruded partially from the cryo-EM density of the 30S:YjeQ structure. However, in both cases we found a nearby void space in the additional mass attached to the

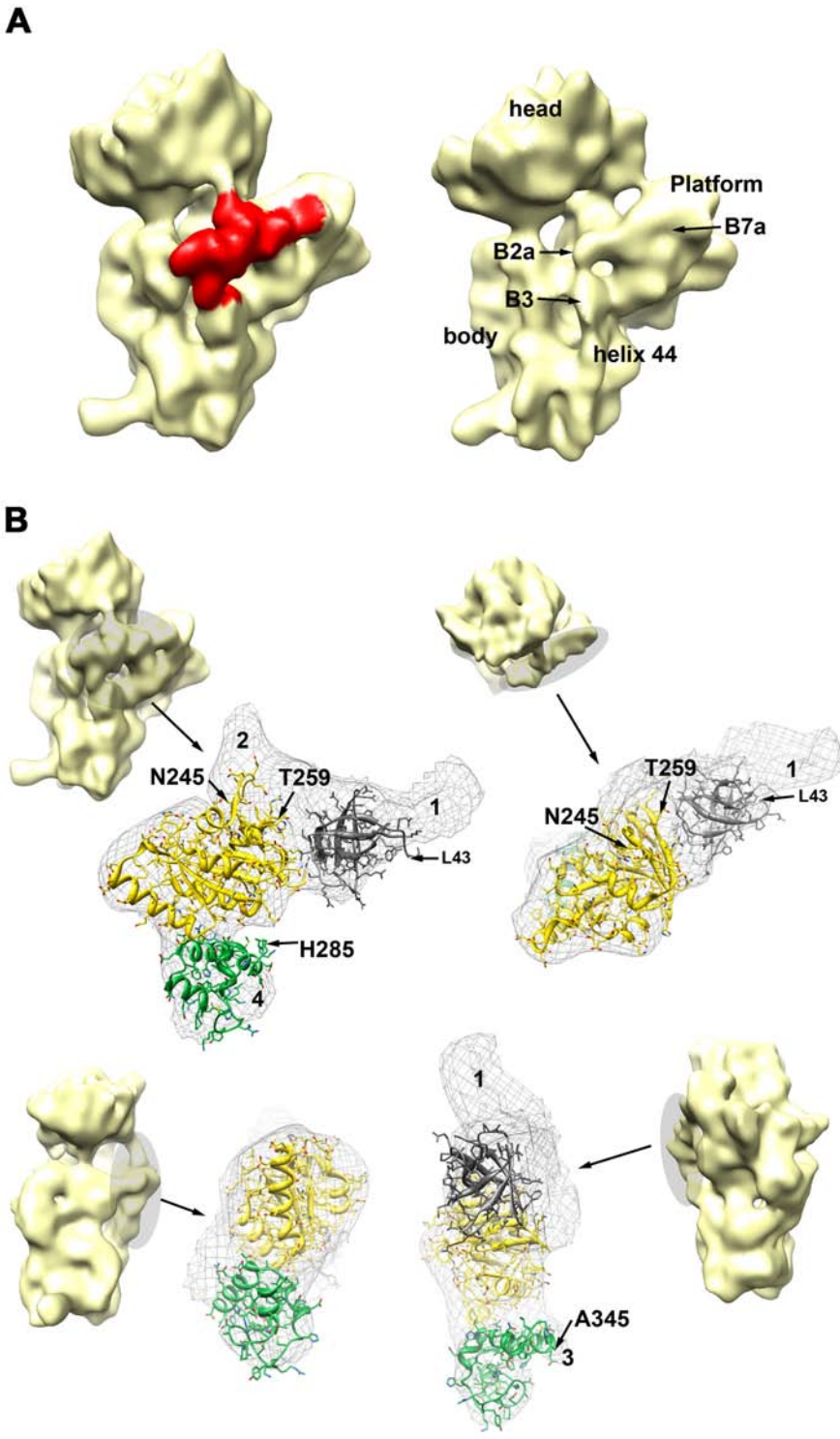
30S:YjeQ complex that could easily accommodate these two helices (Figure 3.3B; densities labeled as 3 and 4) suggesting that the conformation of these two motifs in YjeQ may be slightly different in the unbound and bound state of the protein. Indeed, the distal part of the C-terminal helix in the homologous *A. aeolicus* structure is oriented differently than in the *S. typhymurium* (Nichols et al., 2007) and *T. maritime* (Shin et al., 2004) structures (Supplemental Figure S3.4), indicating that this helix is flexible and capable of adopting alternate conformations.

Next, we superimposed the obtained difference map containing the docked X-ray structure of YjeQ into the cryo-EM structure of the 30S:YjeQ complex to locate YjeQ in the map. Subsequent docking of the X-ray structure of the 30S subunit (PDB ID: 2AVY) (Schuwirth et al., 2005) in the remaining density of the cryo-EM map allowed us to produce a pseudo-atomic model of the 30S:YjeQ complex (Figure 3.4A). In this model, we noticed that part of the density corresponding to the upper domain of helix 44 and helix 45 in the cryo-EM map of the 30S:YjeQ complex was very weak or completely absent suggesting that these two regions became displaced upon YjeQ binding (Figure 3.4B). Interestingly, the cryo-EM map of the 30S:YjeQ complex exhibited a density covering the r-protein S12 and adjacent to the segment of helix 44 missing in the cryo-EM map. It is plausible that these two regions move from their native location in the 30S subunit to this new site upon YjeQ binding (Figure 3.4C). However, this density is large enough to house the missing segment of helix 44 but not helix 45 suggesting that parts of these two helices became flexible or unfolded and thus, invisible in the cryo-EM structure. The displacement of parts of helix 44 and 45 in the structure of the 30S:YjeQ

complex is consistent with these two helices establishing a limited number of contacts with r-proteins and being stabilized by non-sequence specific backbone interactions with the rest of the 16S rRNA (Whitby et al., 2000).

3.4.3 Refinement and validation of the 30S:YjeQ structure

The pseudo-atomic model of the 30S:YjeQ complex obtained by manual placement of the X-ray structure of the 30S subunit and YjeQ protein into the cryo-EM map of the 30S:YjeQ complex was refined with flexible fitting using the YUP.scx refinement method within the YUP program (Tan et al., 2008). The model was initially placed into the cryo-EM map in a position equivalent to our manual placement and was then divided into 26 distinct units. Five of the fragments were the 16S rRNA structural domains (head, body, platform and the 3' minor domain that was broken up into helix 44 and helix 45). YjeQ and the ribosomal proteins constituted the remaining 21 fragments. The model was then refined to the cryo-EM density map of the 30S:YjeQ complex by simulated annealing with molecular dynamics. The energy function contains terms for scoring the quality of the fit of the model to the density map, restraint energies for a Gaussian Network Model that represents the all-atom structure, and volume exclusion terms.

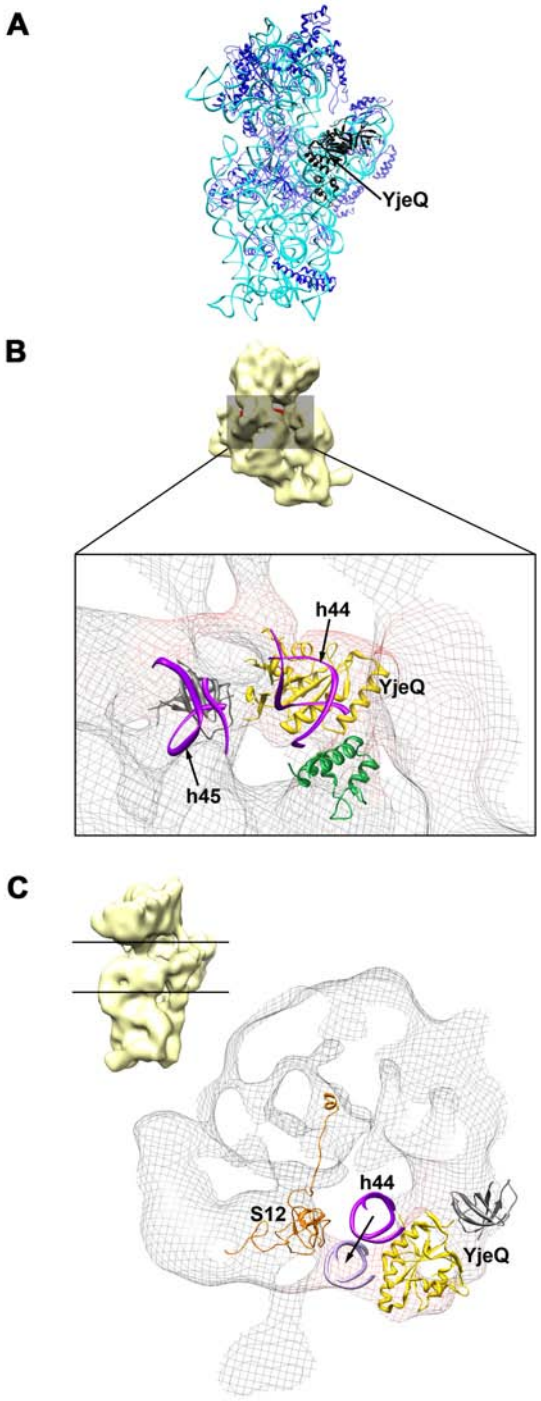

A
head
Platform
B7a
B2a
B3
body
helix 44
B
2
N245
T259
1
L43
H285
4
T259
1
L43
N245
1
A345
3

continued

Figure 3.3 Cryo-EM structure of the 30-YjeQ complex and docking of the YjeQ structure.

(A) Cryo-EM map of the 30S:YjeQ complex side by side with a control reconstruction obtained from free 30S subunits. The head, platform, body and helix 44, which are main landmarks in the 30S subunit are indicated as well as the inter-subunit bridges B2a, B3 and B7A covered upon YjeQ binding. The additional density representing the YjeQ protein has been colored in red. (B) The four panels provide different views of the additional density covering the upper domain of helix 44 and nearby region in the platform in the cryo-EM map of the 30S:YjeQ complex. The additional density is represented as a mesh with the docked X-ray structure of the YjeQ protein. The X-ray structure of the YjeQ protein from S. typhimurium (PDB ID 2RCN) is shown with the OB-fold domain colored in dark grey, the GTPase domain in yellow and the zinc-finger domain in green. As an aid for orientation, each panel shows the cryo-EM map of the 30S:YjeQ complex in a matching view. The first residue described in the YjeQ structure (L43) and residues N245 and T259 flanking a sequence not described by the atomic coordinates are labeled. Residues H285 and A345 are indicated to mark the first and last a helices at the C-terminal zinc-finger domain of YjeQ. Areas of the density not filled by the X-ray structure of YjeQ are labeled with numbers from 1 to 4.

The modeled structure fit the density well and the placement of the 30S subunit and YjeQ differed minimally from the placement obtained by manual rigid-body docking (Supplemental Figure S3.6A and Figure 3.4A), except for the OB-fold domain. This domain was rotated with respect to the manual placement and towards the strip of density connecting the head and platform (Supplemental Figure S3.6B). However, we favor the orientation predicted by the manual placement because it is consistent with other OB-fold containing proteins that interact with RNA (Supplemental Figure S3.7) (see below). In addition, the domains of the 16S rRNA were in slightly different orientations when compared with the X-ray structure, which suggests that the relative orientation of these 30S subunit domains change slightly upon YjeQ binding (Supplemental Figure S3.6C).



continued

Figure 3.4 Pseudo-atomic model of the 30S:YjeQ complex and displacement of the upper domain of helix 44 in the cryo-EM map of the 30S:YjeQ complex.

(A) Pseudo-atomic model of the 30S:YjeQ complex. The backbone of the 16S rRNA is represented as a ribbon in cyan. Ribosomal proteins are colored in dark blue. The YjeQ protein is colored in black. (B) Back view of the 30S:YjeQ complex (see orientation aid in panel above) after a clipping plane orthogonal to the direction of view was applied to the cryo-EM map removing the blocking densities on the back. The panel shows the density map of the 30S:YjeQ complex as a mesh. The OB-fold, GTPase and zinc-finger domains of the YjeQ structure are shown fitted into this density and colored in grey, yellow and green, respectively. The nucleotides in helix 44 and helix 45 for which a corresponding density in the 30S:YjeQ complex EM map does not exist are colored in blue. (C) Top view of a slab of the cryo-EM map of the 30S:YjeQ complex. The slab was cut orthogonal to the longitudinal axis of helix 44 (see orientation aid in panel above). Binding of YjeQ to the 30S subunit causes a displacement of the upper domain of helix 44. The location of this motif in the free 30S subunit and in the 30S:YjeQ complex is shown in blue and cyan, respectively. The black arrow indicates the direction of the shift.

To determine whether refinements using different starting models converge to the same solution, the process was repeated using the same manually constructed pseudo-atomic model but having removed the helix 45 of the 16S rRNA, which lacks a corresponding density in the cryo-EM map of the 30S:YjeQ complex. The refinement converged to an identical solution. Subsequently, we used a starting model with YjeQ bound in the same orientation but displaced laterally towards the platform. During refinement of this model, YjeQ drifted towards the previous solution and reached the same location. Starting models with YjeQ bound in drastically different orientation from our manual placement diverged to dissimilar solutions. However, these models were then subjected to subsequent refinements, and divergent orientations were often found, which suggests that these placements were far from optimal. Any of these solutions were consistent with our manually constructed pseudo-atomic model of the 30S:YjeQ complex (data not shown). In addition, they left large regions of the density assigned to YjeQ

unoccupied and regions of the YjeQ structure protruding from the cryo-EM map. Therefore, all these unstable solutions were discarded.

In conclusion, the convergence of refinements with multiple starting models to a solution very similar to our manually constructed pseudo-atomic model of the 30S:YjeQ complex validate our placement of the YjeQ and 30S subunit X-ray structures in the cryo-EM map.

3.4.4 Interactions of YjeQ with the 30S subunit

The manually constructed pseudo-atomic model of the 30S:YjeQ complex was used to study the interactions of the YjeQ protein with the 30S subunit. Bound YjeQ interacts with the 3' minor, the 3' major (head) and the central (platform) domains in the 30S subunit, (Figure 3.3A; left panel). In this location, YjeQ covers the inter-subunit bridge B2a at the top of helix 44 (Figure 3.3A; right panel), which is an essential region to maintain the association of the small and large subunit in the 70S ribosome (Maivali & Remme, 2004). In addition, the inter-subunit bridges B3 and B7a are also covered. The bridge B3, located in the upper domain of helix 44, serves as the pivot point during the ratcheting motion of the ribosome and the bridge B7a is at the apex of the platform (Figure 3.3A; right panel). The steric blockage of these inter-subunit bridges in the 30S:YjeQ complex provides a structural explanation to previous biochemical data indicating that YjeQ induces the dissociation of the 70S ribosome (Himeno et al., 2004).

In the region of the 30S subunit to which YjeQ binds, rRNA occupies most of the contact areas, except for the point of interaction in the head that is near the S13 r-protein

(Figure 3.5). The most prominent interaction of YjeQ with the 30S subunit involves the N-terminal OB-fold domain that contacts helix 23b and 24a in the platform of the subunit (Figure 3.5A), which correlates with a previous biochemical study of truncated YjeQ revealing that the N-terminal OB-fold region is essential for ribosome binding (Daigle & Brown, 2004). The interaction between the OB-fold domain and these two helices occurs through its binding face and encompasses loop 1 and β strands 2 and 3. The binding mechanism is very similar to r-proteins S12 or S17 that also have an OB-fold domain and bind RNA through this conserved structural motif (Supplemental Figure S3.7). Consistently, the electrostatic surface potential of YjeQ shows a continuous surface with positive potential covering the binding face of the OB-fold that interact with helix 23b and 24a (Figure 3.5C).

The C-terminal zinc-finger domain at the other end of YjeQ sits in the upper domain of helix 44 in its ascending way to the decoding center (Figure 3.5B). It is interesting that negatively charged residues in the surface of the loop region between the 3_{10}-helix and H5 a-helix in the zinc-finger domain face the negatively charged phosphate-oxygen backbone of nucleotides of helix 44 in the native 30S subunit (Figure 3.5B and 5C; right panel). This charge distribution supports the necessity of a shift in the helix 44 motif to avoid electrostatic repulsion. In addition, we noted a steric clash of a stretch of helix 44 (nucleotides U1481 to G1486) with the C-terminal α-helix of the zinc-finger domain in the manual placement of the YjeQ protein (Figure 3.5B). Indeed, during the flexible fitting refinement this domain slightly move outward and partially came out of the cryo-EM density (Supplemental Figure S3.6D; asterisk) minimizing the steric clash

with helix 44. Therefore, we hypothesized that YjeQ dislocates helix 44 from its mature conformation upon binding to the 30S subunit through the negatively charged residues in the surface of zinc-finger domain loop. Consistently with this, the obtained cryo-EM map of the 30S:YjeQ complex lacked any density matching the upper domain of helix 44 in the mature 30S subunit structure (Figure 3.4B). As a consequence of the displacement of helix 44, the surface of the GTPase domain facing the 30S subunit does not make significant contact with the rRNA. However, a small number of interactions with helix 45, involving residues from the strands b6 and b7 that connect this domain with the OB-fold domain, were observed (Figure 3.5A and 3.5B).

3.5 DISCUSSION

3.5.1 Binding site of YjeQ in the 30S subunit

The cryo-EM structure of 30S:YjeQ complex presented here shows that YjeQ binds in the platform domain and in the upper region of helix 44 in the 30S subunit. Interestingly, the area of helix 44 bound by YjeQ is distorted in immature particles purified from an *E. coli* strain lacking the *yjeQ* gene (Jomaa et al., 2011a). The cryo-EM reconstruction obtained from these immature 30S subunits revealed that the upper segment of helix 44 was displaced outwardly protruding from the interface of the 30S subunit that contacts the large 50S subunit. The fact the that immature particles accumulated in a strain lacking the *yjeQ* gene contain a distortion in the same area where YjeQ binds suggest that YjeQ may be involved in the correct folding of the upper domain of helix 44.

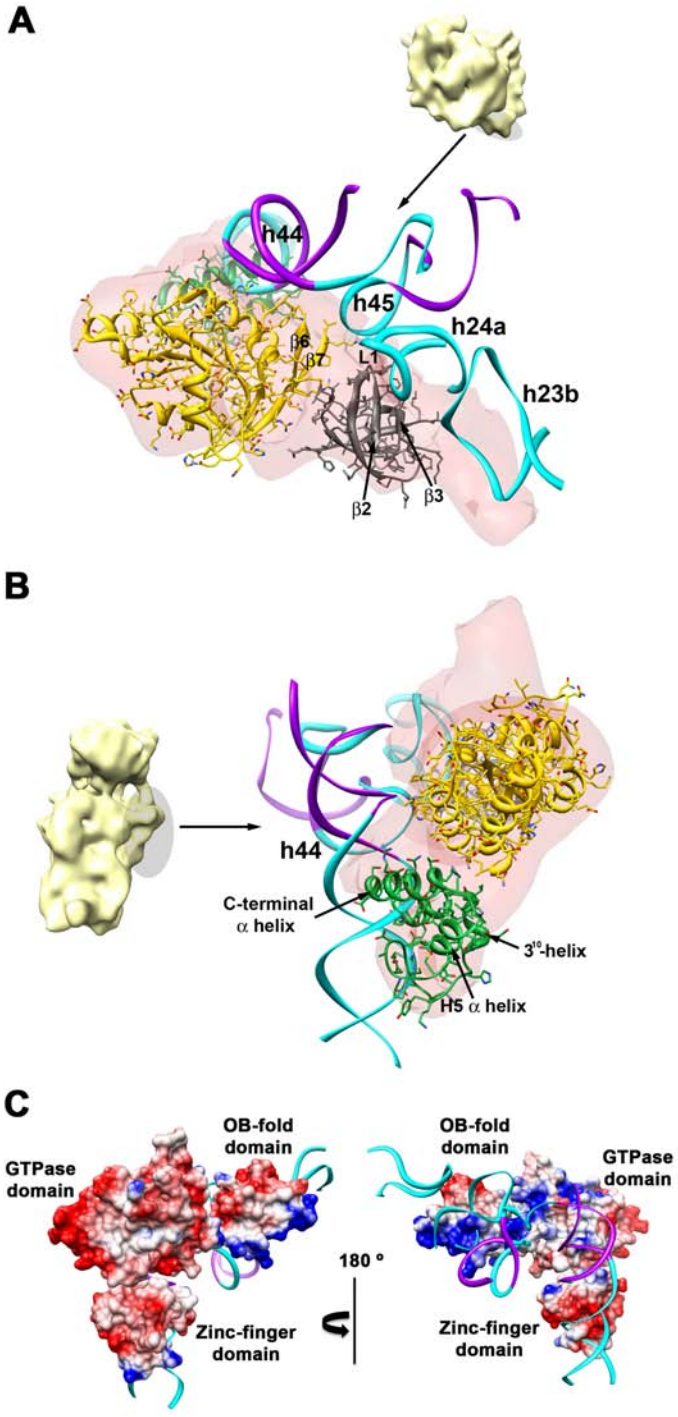
The binding location of YjeQ established by our cryo-EM structure is consistent with previous footprinting studies (Kimura et al., 2008) indicating that the chemical reactivity of residues around the A and P site (G530, A790, G925, G926, G966, C1054, G1339 and G1405) were reduced in the presence of YjeQ (bound to a non-hydrolyzable GTP analog) (Figure 3.6). The footprinting data suggested that YjeQ binding protects this region of the 30S subunit, which is near the YjeQ binding site established by cryo-EM here. In addition, these experiments found increased reactivity of nucleotides in helix 44 (A1408, A1468 and A1483) (Figure 3.6) suggesting a conformational change on YjeQ binding that increases the accessibility of these nucleotides to footprinting reagents. This finding is in accordance with the distortion of the upper domain in helix 44 observed in our 30S:YjeQ structure (Figure 3.4B and 3.4C). Furthermore, the footprinting experiments revealed that GDP-bound YjeQ produced significantly less changes or no changes in the reactivity of most of the residues mentioned above consistent with the lower affinity of YjeQ for the 30S subunit in its GDP-bound state (Daigle & Brown, 2004; Kimura et al., 2008).

3.5.2 Conformational changes in YjeQ

GDP-bound *S. typhymurium* YjeQ structure (Nichols et al., 2007) fits remarkably well in the additional density observed in our cryo-EM map of the 30S:YjeQ complex that was assembled in the presence of GMP-PNP. This nucleotide mimics the GTP state of YjeQ. This observation suggests that the protein does not undergo a large inter-domain rearrangement upon binding and GTP hydrolysis. Only small translational and rotational movements of each domain independently were necessary to optimize the fitting

(Supplemental Figure S3.5). The displacement of the N-terminal OB-fold with respect to the GTPase domain was larger than the one applied to the C-terminal zinc-finger domain. This is consistent with the higher flexibility predicted for the OB-fold domain with respect to the GTPase and zinc-finger domains (Shin et al., 2004). These small conformational differences observed between free and bound YjeQ contrast with the large interdomain rearrangements that Era undergoes upon binding to the 30S subunit (Sharma et al., 2005).

Association of YjeQ with the 30S subunit results in a 160-fold stimulation of YjeQ GTPase activity (Daigle & Brown, 2004). The most prominent interactions between the protein and the 30S subunit appear to occur through the N-terminal OB-fold domain, whereas the GTPase domain establishes comparatively very few contacts. Thus, the OB-fold domain may have a central role in this stimulation by contacting the 30S subunit and propagating conformational changes through the large area of interaction between the OB-fold and GTPase domains (~700 $Å^2$) (Shin et al., 2004). Likewise, conformational changes in the GTPase domain caused by GTP hydrolysis can also be easily transmitted to the OB-fold domain to induce dissociation from the 30S subunit. This interpretation of the structure is consistent with previous finding indicating that the OB-fold region is essential for ribosome binding and GTPase stimulation (Daigle & Brown, 2004).


A
h44
h45
h24a
h23b
β6
β7
L1
β2
β3
B
h44
C-terminal
α helix
3¹⁰-helix
H5 α helix
C
OB-fold
domain
GTPase
domain
Zinc-finger
domain
180 °
OB-fold
domain
GTPase
domain
Zinc-finger
domain

continued

Figure 3.5 Interactions of YjeQ with the 30S subunit.
(A) Interactions of the OB-fold and GTPase domain of YjeQ with the 16S rRNA helices 24a and 23b in the platform and helix 45 in the 3' minor domain of the 30S subunit. Important loops and b strands in YjeQ for the interaction with the 30S subunit are labeled. The OB-fold, GTPase and zinc-finger domains of the YjeQ structure are colored in grey, yellow and green, respectively. The extra density attributed to YjeQ is displayed as a red semitransparent surface. The rRNA is displayed in cyan except for the nucleotides for which a corresponding density in the 30S:YjeQ complex EM map does not exist that are colored in blue. (B) Interactions of the zinc-finger domain of YjeQ with the 16S rRNA helix 44. As an aid for orientation, each panel shows the cryo-EM map of the 30S:YjeQ complex in a matching view. (C) Electrostatic surface potential of S. typhimurium (PDB ID 2RCN) YjeQ. Zones in red, blue and white are negatively charged, positively charged and uncharged, respectively. The left panel shows the surface of YjeQ that is not in contact with the 30S subunit whereas the right panel shows the opposite face of the protein contacting the 16S rRNA.

3.5.3 Functional interplay of YjeQ with other ribosomal assembly factors

The KsgA protein is a universally conserved methyltransferase that dimethylates both A1518 and A1519 in helix 45 of the 16S rRNA (Sparling, 1970; Helser et al., 1972). It has been proposed that KsgA also binds the 30S subunit at late stages of assembly and these two dimethylations could act as a 'mark' for fully assembled 30S subunits (Xu et al., 2008). Therefore, it is plausible that KsgA and YjeQ may both be bound to the 30S subunit at some point during assembly. However, concurrent binding of these two factors has not been described. Footprinting experiments mapped the interaction of KsgA with the 30S subunit mainly in the lower domain of helix 44 and helices 11, 27, 45 (Xu et al., 2008), which is near but not overlapping with the YjeQ binding site. KsgA also contacts helix 24, which is the only interaction motif overlapping with the YjeQ binding site. Therefore, concurrent interaction of KsgA and YjeQ with the 30S subunit is not inconsistent with available structural data.

In addition, previous studies suggest a functional interplay between YjeQ and other biogenesis factors, especially Era, RbfA and RimM (Inoue et al., 2003; Inoue et al., 2006; Campbell & Brown, 2008; Goto et al., 2011). The principal known function of both YjeQ and Era is to assist the processing of the 3' end of the precursor 17S rRNA. The binding sites of Era and YjeQ to the 30S subunit do not overlap, thus simultaneous binding of the two proteins to the 30S subunit to perform this function is possible. Era sits in the cleft region between the head and platform and directly contacts the 3'end of the rRNA through its C-terminal domain (Sharma et al., 2005). The OB-fold domain in YjeQ binds in the immediate vicinity of Era's binding position but it does not contact the 3' end directly. In these locations, both proteins may mediate the processing of the 3' end by either recruiting the RNase or by inducing a specific conformation on this end that is recognized by the RNase. Alternatively, YjeQ may assist the processing of the 3' end by covering inter-subunit bridges and simply preventing premature association with 50S subunits, which in turn could block access of the RNases to the cleavage site. YjeQ also promotes the release of RbfA from the mature 30S subunit in a GTP-dependent manner (Goto et al., 2011). RbfA binds in the neck region of the 30S subunit. Binding of RbfA in this position induces a shift in the upper domain of helix 44 and 45 moving these motifs over RbfA and locking the protein deeply within the cleft between the head and body. YjeQ also produces a distortion of helices 44 and 45 upon binding to the 30S subunit. Therefore, it is possible that movement of these two helices induces the release of RbfA from the 30S subunit. In any case, plausibly the conformational changes induced by YjeQ in the context of these multi-complexes, where several biogenesis factors bind

simultaneously to the 30S subunit, may be different from the ones described here. Structures of these multi-complexes are necessary to establish how the functional interplay between these factors occurs.

The 30S:YjeQ complex in this work was assembled with mature 30S subunits, which bind YjeQ with much greater affinity than immature 30S subunits (Goto et al., 2011). However, It is possible that the rRNA precursor sequences in immature 30S subunits may modify the binding mechanism of YjeQ and the conformational changes induced upon binding. Therefore, it is plausible that the binding mode described herein and associated conformational states induced by YjeQ with mature subunits may be different from those associated with YjeQ upon binding to immature subunits.

3.5.4 Mechanisms of YjeQ to assist maturation of the 30S subunit

Each of the 30S subunit large rRNA domains (body, head and platform) can fold separately *in vitro* (Agalarov et al., 1998; Agalarov et al., 1999) driven by binding of r-proteins (Sykes & Williamson, 2009; Shajani et al., 2011). Interdomain interactions occur late in the assembly process to properly orient the domains in the mature structure. Our structure reveals that YjeQ binds the 30S subunit simultaneously contacting the head, platform and helix 44. This binding places YjeQ in a position to assist in the process of domain orientation during the final stages of the 30S subunit assembly (Poot et al., 1998). Interestingly, RbfA and Era biogenesis factors also bind at the junction between the 30S subunit domains and may also assist in this process.

YjeQ distorts the binding site for IF1, which binds the 30S subunit in the cleft between helix 44 and protein S12 (Carter et al., 2001). In addition, the YjeQ binding site partially overlaps with the interaction site of IF2 (Simonetti et al., 2008) and the C-terminal domain of IF3 (Moazed et al., 1995; Dallas & Noller, 2001). Hence, our structure suggests that YjeQ might assist in ribosome maturation by preventing premature formation of the translation initiation complex. Curiously, Era also inhibits formation of the translation initiation complex by contacting the conserved anti-Shine-Dalgarno nucleotides of the 30S subunit (Sharma et al., 2005), providing an additional example of partially overlapping functions between ribosome biogenesis factors.

3.6 MATERIAL AND METHODS

3.6.1 Construction of YjeQ overexpression clone

Gateway Recombination Cloning technology (Invitrogen Canada, Burlington, Ontario) was used to construct full-length YjeQ featuring an N-terminal His_6 affinity purification tag cleavable by TEV protease. Briefly, the *yjeQ* gene was amplified from *E. coli* MG1655 genomic DNA using the polymerase chain reaction. PCR was performed using VENT DNA polymerase (New England Biolabs, Beverly, MA) and the following primers: 5' G G GGG **ACA AGT TTG TAC AAA AAA GCA GGC TTA** GAT TAC GAT ATC CCA ACG ACC GAA AAC CTG TAT TTT CAG * GGC AGT AAA AAT AAA CTC TCC AAA GGC-3 and 5'-CGC GGA TCC TCA GTC ATC CGT ATC AGA AAA G-3, where the recombination sites, coding sequence and cleavage sites are denoted by bolded letters, underlining and an asterisk respectively. The amplified product was inserted into the pDONR201 (Invitrogen) vector as per the manufactures protocol and

subsequently cloned into the pDEST17 (Invitrogen) destination vector that encodes six histidine residues at the N-terminus of the insert. The resulting construct pDEST17-YjeQ was validated by sequencing (MOBIX, McMaster University).

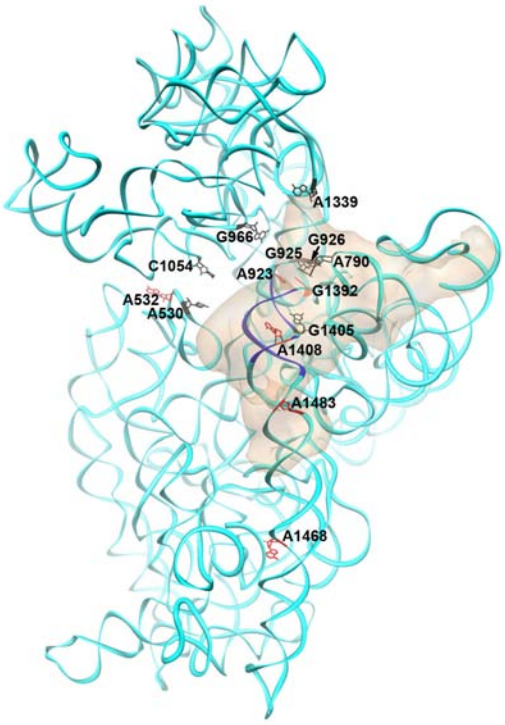


Figure 3.6 Comparison of the interaction model of YjeQ with the 30S subunit from cryo-EM and footprinting studies.

The backbone of the 16S rRNA of the 30S subunit is represented as a ribbon in cyan. Nucleotides whose chemical reactivity was decreased or increased by the addition of YjeQ with a non-hydrolyzable GTP analog (Kimura et al., 2008), are highlighted in black and red, respectively. The area covered by YjeQ according to the 30S:YjeQ cryo-EM structure presented here is shown as a pink semitransparent density. The 30S:YjeQ map suggests that a region of helix 44 is displaced upon YjeQ binding. This region is colored in dark blue.

3.6.2 Purification of YjeQ protein

YjeQ protein was expressed as a N-terminal His_6-tagged protein in *E. coli* BL21-AI competent cells transformed with the expression vector pDEST17-YjeQ (see above). One liter of cells was grown in LB medium at 37ºC to OD_{600} = 0.6 and expression was induced with 0.2% L-arabinose. Cells were induced for 3 hours at 37 ºC and harvested by centrifugation at 3,700 X *g* for 10 min. The cell pellet was washed with 1X PBS buffer (137 mM NaCl, 2.7 mM KCl, 8.1mM Na_2HPO_4, 1.76 mM KH_2PO_4, pH 7.4) and resuspended in 20 mL of lysis buffer (50 mM Tris-HCl, pH 8.0, 10% (wt/vol) sucrose, 100 mM NaCl). The cell suspension was passed through a French pressure cell at 20,000 lb/in^2 three consecutive times to lyse the cells. The lysate was spun at 30,000 x *g* for 45 min to clear cell debris. NaCl was added to bring the concentration to 0.5 M, and the lysate was filtered with a 0.45 mm filter and added to a HiTrap Metal Chelating column (GE Healthcare Life Sciences) equilibrated with 50 mM Tris-HCl, pH 8.0, 0.5 M NaCl, 5% (v/v) glycerol. Unspecifically bound proteins were washed with increasing concentrations of imidazole up to 90 mM. YjeQ was eluted with 240 mM imidazole. Purity of the fractions was monitored by SDS-PAGE. Fractions containing pure protein were pooled and dialyzed against 50 mM Tris-HCl, pH 8.0, 10% (v/v) glycerol.

The N-terminal His_6-tag was removed by digestion with purified tobacco etch virus (TEV) protease. An amount of 0.74 mg of TEV was added to the pooled and dialyzed fractions containing YjeQ (40 mg). The total volume of the reaction was 10 mL and the reaction mixture was incubated at 16 ºC for 4 hours. The mixture was added to 1 mL of Ni-NTA agarose resin (Qiagen) previously equilibrated with 50 mM Tris-HCl, pH

8.0, 0.5 M NaCl, 5% (v/v) glycerol and incubated in a rocking platform at 4 ºC for one hour. The mixture was spun at 1,400 x *g* for 5 min and the supernatant containing the untagged YjeQ protein was recovered, dialyzed against 50 mM Tris-HCl, pH 8.0, 10% (v/v) glycerol and concentrated using a 10 kDa-cutoff filter (Amicon). The protein was frozen in liquid nitrogen and stored at -80 ºC.

3.6.3 Purification of the 30S ribosomal subunits and preparation of the 30S:YjeQ complex

Purified 30S subunits from BW25113 of *Escherichia coli* were prepared using centrifugations over sucrose cushions and gradients (Daigle & Brown, 2004). One liter of Luria broth (LB) was inoculated with 10 ml of a saturated overnight culture. Cultures were grown at 37 ºC to an OD_{600} of 0.4 and cooled down to 4 ºC with all subsequent steps performed at this temperature. Harvesting of the cultures was done by centrifugation at 8,500 x *g* for 15 min and the cell pellet resuspended in buffer A (20 mM Tris-HCl pH 7.5, 10.5 mM magnesium acetate, 100 mM NH_4Cl, 0.5 mM EDTA and 3 mM 2-mercaptoethanol). Cell lysis was performed by three consecutive passes of the cell suspension through a French pressure cell at 20,000 lb/in^2. The cell lysate was spun at 30,000 x *g* for 45 min to clear cell debris. Recovered supernatant (S30 fraction) was overlaid on an equal volume of 1.1 M sucrose cushion made up in buffer B (20 mM Tris-HCl pH 7.5, 10.5 mM magnesium acetate, 500 mM NH_4Cl, 0.5 mM EDTA and 3 mM 2-mercaptoethanol) and centrifuged at 100,000 x *g* for 16 hours. The ribosomal pellet was gently washed and sucrose removed by resuspension in buffer C (10 mM Tris-HCl pH 7.5, 10.5 mM magnesium acetate, 100 mM NH_4Cl, 0.5 mM EDTA and 7 mM 2-

mercaptoethanol). The crude ribosomes were pelleted at 100,000 x *g* for 16 hours. To obtain the 30S fraction the crude ribosome pellet was resuspended in buffer F (10 mM Tris-Hcl pH 7.5, 1.1 mM magnesium acetate, 60 mM NH_4Cl, 0.5 mM EDTA, and 2 mM 2-mercaptoethanol) (dissociating conditions). A portion of the subunit suspension (50-60 A_{260} units) was layered onto a 32 ml 10-30% (wt/vol) sucrose gradient made up in buffer F and centrifuged at 43,000 x *g* for 16 hours using a Beckman SW32 Ti rotor. Gradients were fractionated using an AKTAprime™ purification system (GE Healthcare) and the elution peaks corresponding to 30S and 50S particles peaks were monitored by absorbance at A_{260}. The 30S ribosomal subunits were then recovered by centrifugation at 100,000 x *g* for 16 hours and the pellet was resuspended in Buffer E (10 mM Tris-HCl pH 7.5, 10 mM magnesium acetate, 60 mM NH_4Cl, 3 mM 2-mercaptoethanol), and stored at -80°C until further use. Quantification of the 30S subunits was accomplished by absorbance at 260 nm (1 A260 unit is equivalent to 69 pmol of 30S).

To determine the occupancy level of YjeQ in the 30S:YjeQ complex a 5-fold excess of freshly purified YjeQ protein was incubated with a 1 μM preparation of purified 30S subunits for one hour at 30 ºC in the presence of 1 mM GMP-PNP in buffer E. The volume of the reaction was 50 μL. The reaction mixture was layered onto a 1.1 M sucrose cushion (volume 150 μL) made up in buffer E. The ribosomal subunits were pelleted by ultracentrifugation at 330,000 x *g* for 4 hours using a Beckman TLA-120.1 rotor and the pellet was resuspended in 200 μL of buffer E. The supernatant and resuspended pellet of each sample (12 μL load volume) were resolved by SDS-PAGE and stained with Coomassie brilliant blue. The stained gel was subsequently photographed using a

FluorChem FC3 system (Alpha Innotech). The occupancy level of YjeQ in the 30S:YjeQ complex was estimated by quantifying the amount of YjeQ that co-pelleted with ribosomes (2.5 pMol) in each reaction. The integrated pixel density of each YjeQ band was calculated using Imagequant, version 5.2 (Molecular Dynamics), with the local average of the region surrounding the band used for background correction. This yielded an optical density for each band that was used to estimate the proportion of YjeQ found in the supernatant and pellet of each reaction. YjeQ protein (25%) was seen to pellet in the absence of ribosomes. Thus, to determine the amount of YjeQ that co-pelleted with 30S subunits it was necessary to subtract this fraction from the YjeQ observed in the reaction pellets (lanes 7 and 9). The previously established 1:1 stoichiometry for the 30S:YjeQ complex (Daigle & Brown, 2004) was assumed to determine the occupancy level.

To assemble the complex for cryo-EM a 50 mL reaction mixture in buffer E with 2 mM GMP-PNP was prepared by adding 30S subunits to a concentration of 1 mM and a 5-fold excess of freshly purified YjeQ protein. The reaction was incubated for one hour at 30 ºC and then diluted 100 times in buffer E containing 2 mM GMP-PNP before applying it to the EM grids.

3.6.4 Nanoscale microcapillary liquid chromatography electrospray ionization tandem mass-spectrometry (LC-MS/MS)

The protein gel band were distained with 50 mM ammonium bicarbonate and water until clear, then rinsed with water several times to remove all color. The band was reduced with 10 mM DTT at 56 ºC for 30 min and alkylated with 100 mM iodoacetamide

at room temperature for 15 min in the dark. Protein in the band was digested in-situ with 30 mL (13ng/mL) trypsin (Promega Corporation, Madison WI) in 50 mM ammonium bicarbonate at 37 ºC overnight followed by peptide extraction with 30 mL of 5% formic acid, then 30 mL of acetonitrile. The pooled extracts were concentrated to less than 5 mL on a SpeedVac spinning concentrator and then brought up in 0.1% formic acid and 5% acetonitrile for protein identification by micro-flow liquid chromatography electrospray tandem mass spectrometry (microLC-ESI-MS/MS) using a ThermoFisher LTQ-XL-Orbitrap Hybrid Mass Spectrometer (ThermoFisher, Bremen, Germany) coupled to an Eksigent nanoLC-2D HPLC system (Eksigent, Dublin, CA). Chromatography was performed using 0.1 % formic acid in both the A solvent (98%water, 2% acetonitrile) and B solvent (10% water, 80% acetonitrile, 10% isopropanol), and a 5% B to 95% B gradient over 30 min at 5 mL /min through an Eksigent capillary (CSP-3 C18 -100, 0.3m x 100mm) column.

The instrument method consisted of one MS full scan (200–2000 *m*/*z*) in the Orbitrap mass analyzer, an automatic gain control target of 500,000 with a maximum ion injection of 500 ms, one microscan, and a resolution of 60,000. Three data-dependent MS/MS scans were performed in the linear ion trap using the three most intense ions at 35% normalized collision energy. The MS and MS/MS scans were obtained in parallel fashion. In MS/MS mode automatic gain control targets were 10,000 with a maximum ion injection time of 100 ms. A minimum ion intensity of 1000 was required to trigger an MS/MS spectrum. The dynamic exclusion was applied using a maximum exclusion list of 500 with one repeat count with a repeat duration of 30 seconds and exclusion duration of

45 seconds. The lock-mass option was enabled for the FT full scans using the ambient air polydimethylcyclosiloxane (PCM) ion of m/z = 445.120024 or a common phthalate ion m/z = 391.284286 for real time internal calibration.

For database searching, tandem mass spectra were extracted by Proteome Discoverer 1.2. All MS/MS samples were analyzed using SEQUEST (ThermoFinnigan, San Jose, CA; version 27, revision 12) and X! Tandem (The Global Proteome Machine Organization; version 2006.04.01.2). Both search engines were set up to search the *E. coli* database, assuming trypsin digestion, allowing two missed cleavages, and using a fragment ion mass tolerance of 0.5 Da and a parent ion tolerance of 0.02 Da. The iodoacetamide derivative of cysteine was specified as a fixed modification. Deamidation of asparagine and glutamine and oxidation of methionine were specified in X! Tandem as variable modifications. Oxidation of methionine was specified in SEQUEST as a variable modification.

3.6.5 Cryo-electron microscopy, image classification and three-dimensional reconstruction

For cryo-EM, holey carbon grids (400 mesh copper) containing an additional continuous thin (5-10 nm) layer of carbon were previously washed with acetone vapor for 15 min and glow discharge in air for 30 seconds (Aebi & Pollard, 1987). Then, 3.5 μL aliquots of sample were applied to the grid for one min. Grids were blotted for 7 s and vitrified by rapidly plunging into liquid ethane at -180 ºC with a Vitrobot™ (FEI). Data acquisition was performed under low dose conditions (~15 $e^-/Å^2$) on a JEOL 2010F FEG

microscope operated at 200 kV with a Gatan 914 side-entry cryo-holder and at a nominal magnification of 50,000x. A total number of 121 electron micrographs were collected. The defocus range of the images was from -0.65 to -3.9 mm. The micrographs were digitized with a step size of 12.7 μm in a Nikon Supercool Scan 9000 producing images with a sampling value of 2.54 Å/pixel.

Particle projections from the electron micrographs were picked using Boxer (Ludtke et al., 1999). Contrast transfer function of the micrographs was estimated using CTFFIND software (Mindell & Grigorieff, 2003) and corrected using Xmipp software package (Scheres et al., 2008). Image classification was performed using supervised methods (Valle et al., 2002; Gao et al., 2004). Detailed procedures regarding the classification method are included in the figure legend for Figure 3.2. From the total 73,896 projections collected, 16,228 showed a clear YjeQ density in our image classification procedure and only those particles were subsequently used for the reconstruction of 30S:YjeQ complex. The 3D reconstruction of the 30S:YjeQ complex was calculated using 3D projection alignment procedures as implemented in the Xmipp software package (Scheres et al., 2008). The reference map used to refine the 3D reconstruction was the X-ray structure of the *E. coli* 30S subunit (PDB ID: 2AVY) (Schuwirth et al., 2005) low-pass filtered to 25Å resolution. In each refinement, sets of projections were calculated from the reference map using an angular spacing that ranged from 15º to 2º during the multiple cycles of refinement.

The correct handedness of the structures was imposed by the X-ray crystal structure of the 30S subunit from *E. coli* (PDB ID: 2AVY). Resolution of the cryo-EM maps was estimated by calculating two maps following the last cycle of refinement from the even and odd numbered particles. Then Fourier shell correlation was calculated between both maps and the resolution was estimated using a FSC value of 0.5. These values were used to low-pass filter the refined cryo-EM maps.

3.6.6 Construction of a pseudo-atomic model of the 30S:YjeQ complex

To locate the YjeQ protein in the cryo-EM map of the 30S:YjeQ complex, a difference map between the control cryo-EM structure of the free 30S subunit (Figure 3.3A; left panel) (low pass filtered at the same resolution) and the YjeQ-bound 30S subunit was calculated. To this end, we scaled the two maps relative to each other by normalizing their average density and standard deviation values (σ) to zero and one, respectively. Then, the two maps were superimposed by maximizing the correlation coefficient and subtracted. The resulting difference map had an average density equal to zero and a σ of 0.6 and in the rendering only significant values (>2σ) were displayed. The atomic structure of *S. typhymurium* YjeQ (PDB ID 2RCN) (Nichols et al., 2007) was then docked as a single rigid body into the density of the obtained difference map covering the upper domain of helix 44 and platform in the 30S subunit. Subsequently, we applied small translational and rotational movements to each domain independently to optimize the fitting (Supplemental Figure S3.5). Overlapping the obtained difference map containing the docked X-ray structure of YjeQ into the cryo-EM structure of the 30S:YjeQ complex and subsequent docking of the X-ray structure of the 30S subunit

(PDB ID: 2AVY) (Schuwirth et al., 2005) in the remaining density of the cryo-EM map allowed us to produce a pseudoatomic model of the 30S:YjeQ complex. The same difference map analysis was repeated using two *E. coli* 30S subunit X-ray structures (PDB IDs: 2AW7 and 2AVY) (Schuwirth et al., 2005). In these two cases, the structures were first low-pass filtered to match the resolution of our 30S:YjeQ cryo-EM map and then fitted into the cryo-EM map of the 30S:YjeQ complex to perform the map subtraction and calculate the difference map. The CCC values to measure the fitting of the X-ray coordinates of YjeQ into the cryo-EM density were determined after conversion of the fitted atomic coordinates into a density map.

3.6.7 Flexible fitting refinement

The starting model for the flexible fitting refinement was the atomic structures of the 30S subunit (PDB ID: 2AVY) and YjeQ (PDB ID: 2RCN) manually placed in the cryo-EM density of the 30S:YjeQ complex (see above). This assembly was divided into 26 distinct units (Five 16S rRNA structural domains including the head, body, platform, helix 44 and helix 45 along with 20 ribosomal proteins and YjeQ) that were refined with a flexible fitting algorithm using the YUP.scx module (Tan et al., 2008) of YUP (Tan et al., 2006). The potential energy function included terms for the all-atom structures of the 30S:YjeQ complex that were represented as a Gaussian Network Model, terms for scoring the quality of the fit of the model to the density map, and volume exclusion terms. Specifically, the Gaussian Network is created from the initial structure, taking in any atom pairs that are within the cutoff distance of 4 Å. This prevents motions that would otherwise distort the starting structure, while allowing for flexibility between more

isolated domains. The molecular dynamics simulation is carried out with a time step of 5 fs. After an initial energy minimization, the model is heated from a low temperature to 10K over 5,000 steps and then held at 10K over 20,000 steps. After equilibration, the annealing process is performed by reducing the temperature from 10K to 1K over ~50,000 steps, and the model is subjected to a final round of energy minimization. The final placement provides a pseudo-atomic model with stereochemical restraints that fits into the electron density map optimally.

3.6.8 Visualization of structures

Visualization of the fitted YjeQ and 30S atomic structures and cryo-EM maps was done with UCSF Chimera software (Pettersen et al., 2004). The electrostatic surface potential shown for YjeQ was also created using this software.

3.6.9 Accession Numbers

The EM map of the 30S:YjeQ complex has been deposited in the Electron Microscopy Data Bank (EMDB IDs: 1895)

3.7 ACKNOWLEDGEMENTS

We are grateful to the staff at the Canadian Centre for Electron Microscopy. We also thank Alba Guarne for advice in the docking experiments. JO is a Canadian Institutes of Health Research (CIHR) New Investigator and also acknowledges support from an Early Researcher Award from the Ministry of Research and Innovation. E.D.B is a Canada Research Chair and supported for this work by an operating grant from the Canadian Institutes of Health Research (MOP-64292).

CHAPTER 4. Structural Characterization of the 45S and L16-deficient 50S Ribosomal Subunits: A Functional Interplay between L16 and RbgA/YlqF GTPase during Ribosome Assembly

4.1 AUTHOR'S PREFACE

The data presented in this chapter is currently in preparation for submission to *Nature Structure and Molecular Biology* in April. In this work, we studied the role of the RbgA protein in the maturation of the large ribosomal subunit from *Bacillus subtilis*. We structurally and biochemically characterized the immature 45S subunits that accumulates in RbgA-depleted cells. We observed that interface region of the 50S subunit was distorted; particularly, its functional core and the 30S-50S contact points. Furthermore, we analyzed the cryo-EM structure of the L16-defficient 50S subunit. This subunit depicted distorted structural features similar to the 45S subunit. Our findings propose a novel functional interplay between RbgA and L16 during 50S subunit assembly. This work was done in collaboration with laboratories of Dr. Robert Britton at Michigan State University and Dr. Jamie Williamson at the Scripps Research Institute. Dr. David Achilla from the Britton lab isolated the L16-deficient 50S subunits. Dr. Nikhil Jain from the Britton group performed the chemical probing of the 45S subunit (not presented in this chapter). Dr. Joey Davis from the Williamson lab preformed the quantitative mass spectrometry analysis of the 45S subunits. I purified the 45S and 50S subunits under high magnesium conditions, as well as solved the cryo-EM structure of the 50S, 45S, and L16-defficient 50S subunits. Dr. Joaquin Ortega, Dr. Robert Britton, Dr. Nikhil Jain, Dr. Joey

Davis, and myself designed and analyzed the results. Dr. Ortega and I wrote the current version of the manuscript.

The citation is as follows:

Jomaa[1], A., Jain[1], N., Davis, J., Achila, D., Williamson, J., Britton, R., & Ortega J. Structural Characterization of the 45S and L16-deficient 50S Ribosomal Subunits from Bacillus Subtilis Using Cryo-electron Microscopy. (Manuscript in preparation, expected to submit in April 2013).

[1] Authors contributed equally to this work

4.2 ABSTRACT

The process of ribosome assembly is facilitated by multiple protein factors; however, their precise function and how they contribute to the assembly of ribosomes is not known. *The Bacillus subtilis* GTPase RbgA has been implicated in the assembly of the large ribosomal (50S) subunit. Depletion of the RbgA protein from *B. subtilis* cells leads to the accumulation of immature 50S subunits. Previous biochemical work has shown that these subunits have a slower sedimentation rate (~45S) and are missing three late binding ribosomal proteins including the ribosomal protein L16, L27, and L36. In this work, we analyzed the protein content of the immature 45S subunit using stable isotope labeling and quantitative mass spectrometry. We observed that L-proteins, which cluster at the base of the central protuberance, particularly L28, L33, and L35, are depleted from the 45S sample, in addition to the previously characterized L16, L27 and L36. The structure of the 45S subunit solved by cryo-electron microscopy (cryo-EM) depicted a distorted 30S-interface particularly in the central protuberance (CP), the peptidyltransferase center (PTC), and tRNA binding sites. Furthermore, we obtained a cryo-EM structure of the mature L16-deficient 50S subunit, which depicted similar features to the 45S subunit. Our findings propose a novel functional interplay between the RbgA protein and L16 during late stages of 50S subunit assembly.

4.3 INTRODUCTION

The assembly maps for both the 30S and 50S bacterial ribosomal subunits were elucidated several decades ago by the groups of Nomura (Traub & Nomura, 1968a, 1969b) and Nierhaus (Rohl & Nierhaus, 1982; Herold & Nierhaus, 1987). Recently, work from the Williamson (Talkington et al., 2005) and Woodson (Adilakshmi et al., 2008) laboratories established that the assembly of the 30S subunit does not flow through a single pathway but through multiple alternative pathways with no single rate-limiting step. Similar studies have not been done for the 50S subunit but the existence of multiple pathways has also been suggested from *in-vivo* studies (Sykes & Williamson, 2009). Overall, the assembly of the 50S subunit is still much less understood than the assembly of the 30S subunit due to its larger number of components and more intricate structure.

The assembly of the 50S ribosomal subunits is assisted in the cell by a number of protein factors. The precise role of many of these factors in the building of the 50S subunits is still unknown. An important class within these factors is a class of highly specific GTPases (Britton, 2009). Mutation studies suggest that they act late in the assembly; and due to their ability to respond quickly to rapid changes in guanidine nucleotide concentrations, these enzymes provide bacteria with a direct mechanism to shut down ribosome biogenesis in response to a drop in energy levels. It is also plausible that during ribosomal assembly these GTPases facilitate protein-RNA interactions, and rRNA processing, or otherwise ensure that the assembly process progresses quickly and efficiently by limiting the folding landscape of the rRNA and preventing the assembling

subunit from falling into the kinetic traps that often arise in the folding of large RNA molecules (Williamson, 2003).

The focus of this study is to gain insights into the function of RbgA (also known as YlqF), a GTPase widely conserved in all three kingdoms of life. Homologues of this protein are found in Gram-positive and Gram-negative bacteria as well as in all eukaryotes, but is not present in *Escherichia coli* (Uicker et al., 2006). This GTPase plays a non-elucidated, yet essential function during assembly of the 50S subunit. Cells depleted of this protein have a significant decrease in growth rate, a dramatic reduction of the amount of 70S ribosomes and lack 50S subunits. Instead, there is an accumulation of a large subunit intermediate that migrates as a 45S particle in sucrose gradients (Uicker et al., 2006).

The stoichiometry of the RbgA:50S subunit has been estimated to be two RbgA molecules per ribosomal subunit (Matsuo et al., 2006). Their binding site has been mapped at the interface of the 50S subunit with the 30S subunit around the A and P sites, in particular on helices 38, 81 and 85 of the 23S rRNA (Matsuo et al., 2006). The only r-protein that shows direct interaction with RbgA is L25 (named Ctc in *B. subtilis*)(Matsuo et al., 2006). This finding is surprising because in *B. subtilis* L25 is a general stress protein and is not present in 50S subunits from exponentially growing cells (Schmalisch et al., 2002).

RbgA stably interacts with the 45S immature particle in the GTP- and GMPPNP-bound forms but binding is severely reduced in the presence of GDP. With the mature

50S subunit, stable binding occurs only in the presence of GMPPNP (Matsuo et al., 2006; Uicker et al., 2006). The GTPase activity of RbgA is maximally stimulated by mature 50S subunits but not immature 45S particles (Achila et al., 2012). The binding of RbgA to the immature 45S particle however, does increase the affinity of the protein for GTP (Achila et al., 2012). Hydrolysis of the GTP molecule bound to RbgA changes the relative orientation of its GTPase and C-terminal domains (Kim do et al., 2008); this enzymatic activity is also necessary for the release of the protein factor from the 50S subunit (Matsuo et al., 2007). These findings led to the proposal of a model that attributes RbgA to the role of a checkpoint protein. Under this model, RbgA binds to the 45S particle and catalyzes a structural rearrangement that will transform the ribosomal particle into the mature state, which in turn triggers the hydrolysis of GTP and the subsequent dissociation of RbgA from the subunit.

The 45S particle contains an incomplete protein complement with L16, L27 and L36 r-proteins severely underrepresented (Matsuo et al., 2006; Schaefer et al., 2006; Uicker et al., 2006). These are late assembly proteins (Nierhaus, 1991), which suggest that the 45S particle is a late assembly intermediate of the 50S subunit. L27 and L36 sandwich L16 in their binding site in the 50S subunit close to the A site and the PTC. L36 is dispensable for growth but is an important protein in the structural organization of the mature 50S subunit. Deletion of this protein causes alterations in the 23S rRNA distant from its direct sites of interaction with the RNA (Maeder & Draper, 2005). L16 is an essential protein for the functionality of the 50S subunit that binds in close proximity to helices of the 23S rRNA, which are necessary for the peptidyltransferase activity (helix

89) (Ban et al., 2000; Harms et al., 2001) and IF2 binding site (La Teana et al., 2001). In addition, L16 also makes important contacts with the aminoacylated tRNA in the A site (Bashan et al., 2003). The absence of L16 is consistent with the altered migration as 45S particles in the sucrose gradients, since it has been noted that binding of L16 to the assembling 50S subunit causes a large, although still uncharacterized, conformational rearrangement in the subunit (Teraoka & Nierhaus, 1978). This finding posits the question of whether the structural rearrangement that RbgA induces in the 45S particle is a direct consequence of the binding of the GTPase or whether it may be caused indirectly by mediating the entry of L16, which in turn triggers the conformational transition to the mature structure. A combined model in which both the RbgA and L16 bindings produce some degree of structural rearrangement in the 45S particle is also possible.

To gain insights into the role of RbgA in ribosomal assembly, we carried out a comprehensive structural and biochemical characterization of the 45S particle. In this work, we analyzed the protein complement of the 45S particle with respect to the 50S subunit using stable isoptope labeling and quantitative mass spectrometry. We determined that the L-proteins that cluster in the base of the central protuberance (L16, L27, L28, L33, L35, and L36) were all depleted from the 45S sample. In addition, we solved the structure of the immature 45S subunit obtained by cryo-electron microscopy (cryo-EM). The structure of this subunit depicted distorted features in the 30S-interface region, particularly in the CP, the PTC, and tRNA binding sites. Finally, we obtained a cryo-EM structure of an L16-deficient 50S subunit, which also depicted distorted features similar to the 45S subunit. Most notably, the rRNA helices that contact the L16 binding site were

displaced. These results indicate that the 50S subunit lacking L16 reverts into an immature structure adopting a non-functional conformational state similar to the structure of the 45S subunit. Overall, our work indicates interdependency between the recruitment event of L16 and the role of RbgA in the *in-vivo* assembly process of the 50S subunit.

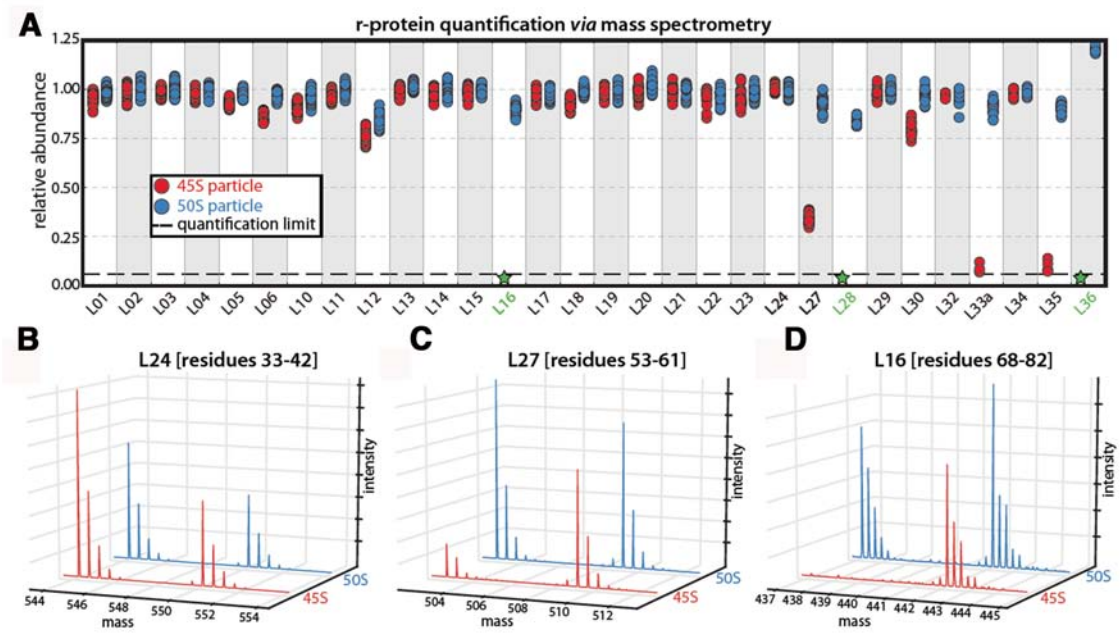


Figure 4.1 L-protein content of the 45S and 50S subunits.

A. Relative protein abundance. R-protein composition was quantified for the 45S (red) and 50S (blue) particles by mass spectrometry. Each marker represents a unique measurement of a peptide resulting from a tryptic digest of the parent protein. Each sample was analyzed in quadruplicate and datasets were merged to improve proteomic coverage. Green stars indicate proteins whose abundance was too low to quantify (see methods). B-D. Peptide mass spectra. Representative mass chromatograms of proteins present stochiometrically (b), depleted from the 45S (c), or completely absent from the 45S (d) are illustrated for the 45S (red) and 50S (blue) sample.

4.4 RESULTS

4.4.1 Isolation of 45S and 50S subunits from RbgA-depleted and IF2-depleted *Bacillus subtilis* cells, respectively, under non-dissociating conditions

In ultra-centrifugation sucrose gradients, large ribosomal subunits isolated from RbgA-depleted strains under non-dissociating conditions (15mM Mg^{2+}) migrate slower as ~45S instead of 50S as observed in mature ribosomes (Matsuo et al., 2006; Uicker et al., 2006). Under these conditions, mature 50S ribosomal subunits are associated with 30S subunits, and migrate as 70S ribosomes.

In order to isolate free 50S subunits under non-dissociating conditions, we used *B. subtilis* cell strains depleted of the initiation factor 2 (IF2). These strains accumulate free 50S subunits that migrate similar to the mature 50S subunits during ultra-centrifugation (Uicker et al., 2006). IF2 is an essential GTPase for the interaction of the 30S subunit with the 50S subunit in the presence of tRNA during the initiation of translation (Antoun et al., 2003; Antoun et al., 2004) and thus should not have an effect on the assembly of the 50S subunit. The isolated 50S subunits along with the 45S subunits will be used in this study for the biochemical and structural characterization of the role of RbgA in the assembly of the ribosome.

4.4.2 Large ribosomal subunits isolated from IF2-depeleted strains contain L-protein content similar to 70S ribosomes isolated from wild type cells

Using quantitative mass spectrometry of stable isotopically-labeled samples (see Material and Methods), we directly measured the protein composition of the 50S particle

purified from IF-2 depleted cells (Figure 4.1A). Each sample was purified from ^{14}N-labeled media and was then spiked with a fixed concentration of 70S particles purified from cells grown in ^{15}N-labeled media. The mixed samples were then analyzed by time-of-flight quantitative mass spectrometry (QMS). To control for concentration differences between the experimental sample and the ^{15}N-labeled 70S spike, the observed r-protein abundance was normalized to that of the primary-binding protein, L24. The vast majority of the r-proteins (*e.g.* L1-L36) were stoichiometrically-bound to the 50S particles, confirming the fidelity of the quantification method (Figure 4.1B).

The ribosomal protein L25 was lacking from both 50S and 70S samples. In *E. coli*, L25 forms contacts with the 5S rRNA, H38, and the ribosomal protein L16 and is constitutively bound to the ribosome (Yusupov et al., 2001). In contrast, the ribosomal protein L25 from *B. subtilis*, known as the stress protein Ctc, is composed of two domains: an N-terminal L25-homolgue domain, and C-terminal unique to the Ctc protein (Schmalisch et al., 2002). The x-ray structures of the 50S subunit from *Thermus thermophilus* (Petry et al., 2005) and *Dinococcus radiodurans* (Harms et al., 2001) depicts the Ctc domain extending and wrapping around the ribosomal protein L16 in the cleft between the CP and the L7/12 stalk. It was demonstrated previously that the Ctc/L25 protein interacts with the ribosome only under stress conditions in *B. subtilis* (Schmalisch et al., 2002; Nanamiya et al., 2004), while it is constitutively bound to the ribosome in *T. thermophilus* and *D. radiodurans*. The purified 50S subunits used in this study were purified from cells during the exponential growth phase (see Material and Methods), thus the absence of the L25 from the 50S subunits and 70S ribosome was anticipated.

In addition, ribosomal proteins L9 and L31 could not be identified in the 50S sample (data not shown). Consistently, previous studies could not identify the presence of the ribosomal protein L9 in purified 50S subunits from *B. subtilis* (Nanamiya et al., 2004). It is possible that L31 binds weakly to the 50S subunit and dissociates after the high salt washes during the purification process (see Materials and Methods).

Taken together, these results confirm that the protein content of the 50S subunits purified from the IF2-depleted cells is similar to the L-protein of the 70S ribosome purified from wild type cells.

4.4.3 Cryo-EM structure of the 50S subunit from *Bacillus subtilis* display similar classical features to the *Escherichia coli* 50S subunit

Using, cryo-EM we solved the structure of the 50S subunit purified from IF2-depleted cells. As an additional control, we also solved the cryo-EM structure of the mature 50S subunit isolated from *Escherichia coli* wild type cells for comparison.

The cryo-EM structure of the 50S subunit (B50S) isolated from IF2-depleted cells subunit depicted the classical landmarks of the large ribosomal subunit including the L1 stalk, L7/12 stalk, the central protuberance (CP), and the A-site finger (H38) (See Figure 4.2). Overall, the conformational state of the B50S is similar to the 50S subunit (E50S) isolated from the *E. coli* cells (see Figure 4.2A). We observed subtle differences in the two structures that could be attributed to short expansions in the rRNA of *E. coli* or *B. subtilis* 23S rRNA. Consistent with the QMS experiments, we could not identify any EM-

densities that correspond to the ribosomal protein L9, L25, and L31 in the B50S cryo-EM map (see Figure 4.2B).

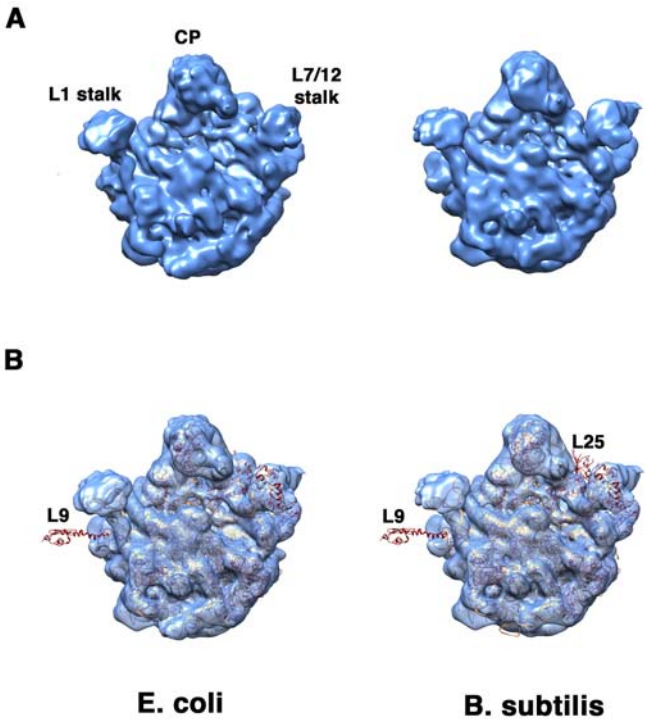


Figure 4.2 Reconstruction of the 50S subunit from B. subtilis IF2-depleted cells

Panel A is the cryo-EM reconstructions of the 50S subunit from E. coli and B. subtilis (IF2-depleted) are shown. Structural features labeled are the L1 stalk, the CP, and the L7/12 stalk. Panel B depicts the X-ray coordinates of the 50S subunit from E. coli (2AW4) are fitted as a rigid body into the cryo-EM maps.

Taken together, the structural analysis of the B50S in combination with the QMS experiments, demonstrate that this particle contains similar L-protein content and adopt a similar conformation to the mature wild-type 50S subunit. These findings are consistent with a role for the IF2 protein in the process of translation initiation and not ribosome assembly.

4.4.4 Immature 45S subunits lacks late-binding ribosomal proteins

Using QMS analysis, we comprehensively analyzed the L-protein content of the 45S subunits from RbgA-depleted cells. Similar to the 50S subunits, we did not identify the presence of the L-proteins L7, L9, L25, and L31 in the 45S sample. Hence, the lack of these proteins could be intrinsic to the purification approach of the large ribosomal subunit from *B. subtilis* and could not be due to the depletion of RbgA.

Three r-proteins (L16, L28, and L36) could not be identified in the 45S sample but were present in stoichiometric quantities in the 50S sample (Figure 1A). Because our quantification approach utilizes the exact mass as well as the offset between the ^{14}N and ^{15}N monoisotopic peaks to identify each peptide, we reasoned that these proteins were likely not identified because they were completely absent from the 45S sample. To confirm this hypothesis, the samples were analyzed in MS/MS mode, which allows for positive identification of each ^{14}N or ^{15}N species individually. In each instance, careful inspection of the mass chromatogram confirmed that the protein was present in the ^{15}N-labeled spike, but absent in the ^{14}N-lableled 45S sample (Figure 4.1D). Furthermore, QMS analysis of the 45S subunit revealed that the L-proteins L33a (rpmGa), and L35 were specifically depleted from this particle (Figure 4.1; relative abundance <0.25). The ribosomal protein L27 was also under-represented but to a lesser extent (relative abundance <0.5).

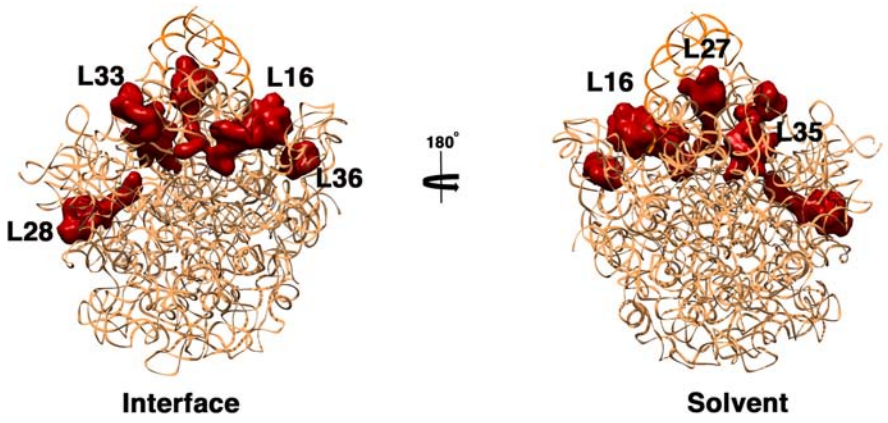


Figure 4.3 Lacking L-proteins cluster at the base of the central protuberance of the 50S subunit.

Interface and solvent views of the 50S subunit depicting lacking ribosomal proteins cluster at the base of the CP region. rRNA is shown in orange (PDB ID 2AW4). L-proteins are displayed as red surface.

The depletion of the ribosomal protein L16, L27, and L36 were also demonstrated by previous biochemical studies (Matsuo et al., 2006; Schaefer et al., 2006) (Uicker et al., 2006). With the higher sensitivity of our QMS experiments, we were also able to indentify the depletion additional proteins L28, L33, and L35. Except for L28, all these proteins cluster to the base of the central (CP) region (see Figure 4.3). In addition, all of the lacking proteins in the 45S sample bind at late stages of *in-vitro* 50S reconstitution experiments (Rohl & Nierhaus, 1982; Herold & Nierhaus, 1987). This indicates that 45S subunits isolated from RbgA-depleted strains represents a late *in-vivo* assembly intermediate.

4.4.5 Characterization of the structure of the 45S subunit from RbgA-depleted cells

The 45S subunit migrates slower than the 50S subunits on a sucrose gradient. The combined molecular weights of all under-represented and missing ribosomal proteins is estimated to be less than 10% of the total molecular weight of the 50S subunit. As a result, the slower migration cannot be attributed solely to a lower molecular weight of the 45S subunit. One possible explanation is that this subunit adopts a less compact or a more extended conformational state that would lead to a slower migration on the sucrose gradients.

To investigate a possible conformational change in the 45S subunit, we analyzed the structure of this subunit by cryo-EM. Initially, we obtained a reconstruction of all the particles combined; we then exposed the dataset to a Maximum-Likelihood image classification (ML3D) approach to minimize sample heterogeneity (see Materials and Methods). This process rendered four distinct cryo-EM maps (14-17Å resolution) with different distortions in the CP and the interface regions (see Figure 4.4). Classifying the particles into five or six classes produced one or more classes with similar features, which indicates the existence of only four homogeneous subsets of the 45S sample. All four classes displayed overall topology similar to the B50S and E50S subunits, which suggest that they represent late intermediates during 50S subunit assembly (see Figure 4.4)

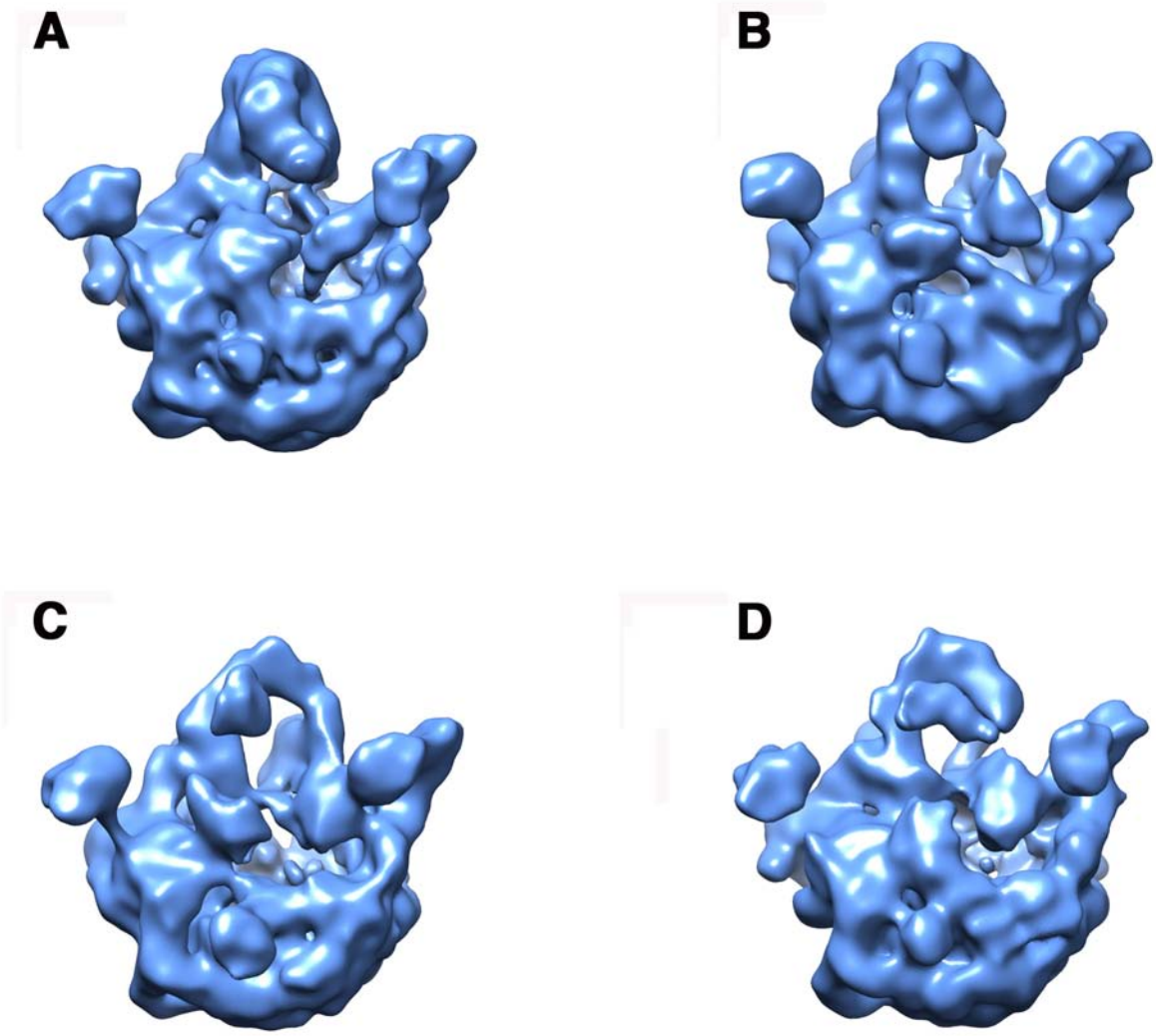


Figure 4.4A-D Cryo-EM Reconstruction of the 45S subunit from RbgA-depleted cells

Panels A-D depicts 45S intermediates one through four, respectively, after exposing the 45S dataset to maximum likelihood classification.

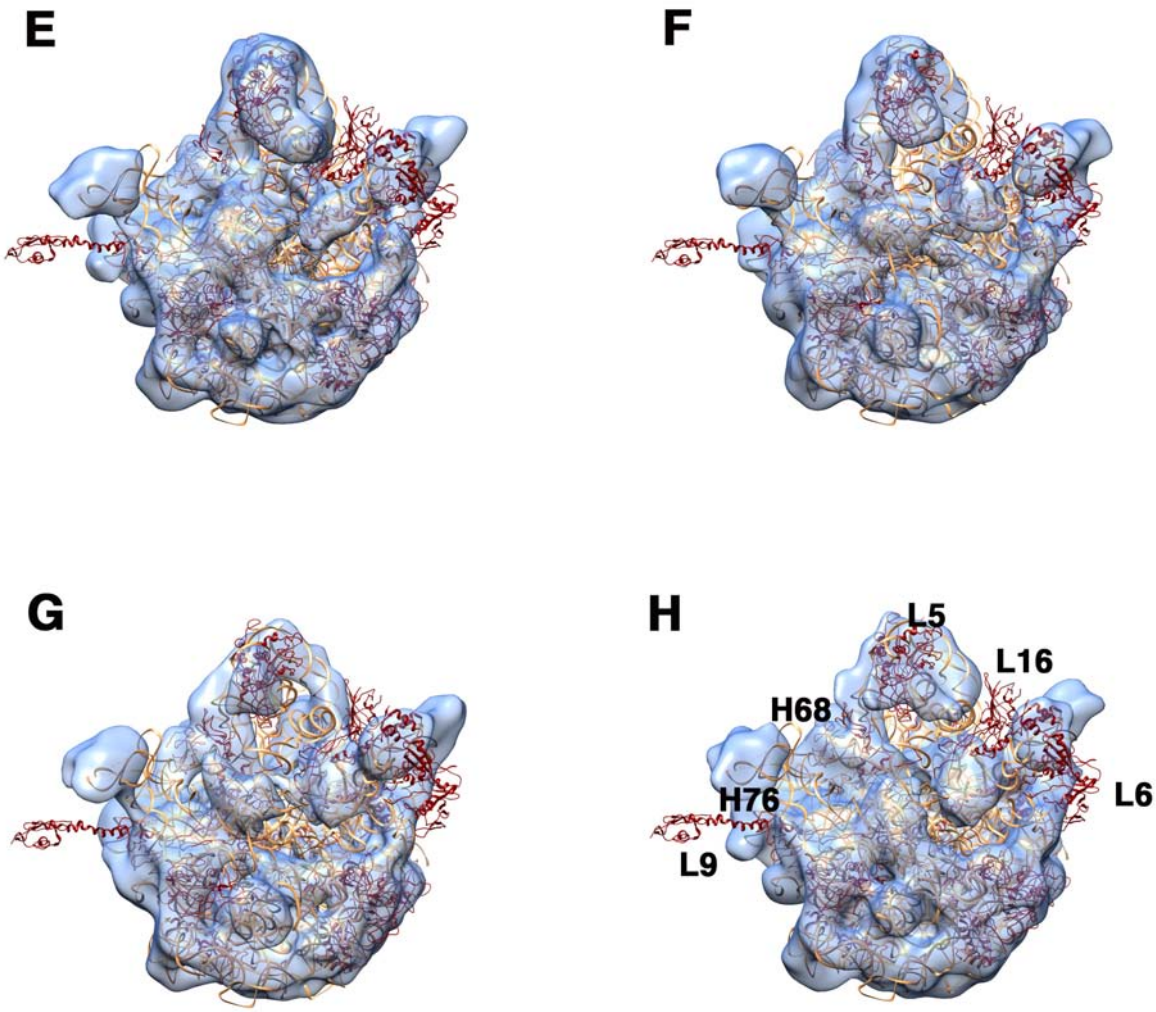


Figure 4.4E-F Cryo-EM Reconstruction of the 45S subunit from RbgA-depleted cells

Panels E-F depicts fitted x-ray coordinates of the 50S subunit fitted into the cryo-EM maps of the 45S intermediates (PDB ID 2AW4).

The cryo-EM structures of the 45S intermediates displayed several distortions throughout the 30S-interface region including the CP, the PTC, and the exit tunnel regions (Figure 4.4). Surprisingly, they cryo-EM map of the 45S intermediate 3 and 4

were lacking large densities in the CP, where only the density corresponding for the 5S rRNA can be visualized (see Figure 4.4C and D). These two intermediates (intermediate 3 and 4) represented 27.7% and 31.45%, respectively, of the 45S dataset. The two remaining 45S intermediates (intermediate 1 and 2) also displayed distortions in the CP but to a lesser degree, and represented 26.5% and 14.36%, respectively, of the total dataset (Figure 4.4E and F). Overlaying the cryo-EM map of the B50S and E50S subunit with the 45S intermediates, we observed that the latter had larger volumes than 50S subunit (data not shown). These findings indicate that the 45S subunit adopted a more extended conformational state, which would explain the slower migration of the 45S sample on ultra-centrifugation sucrose gradients when compared to the 50S subunit.

4.4.6 Ribosomal protein L5 and L6 adopt a flexible conformation in the cryo-EM map of the 45S subunit

Consistent with the QMS experiments, we could not identify EM-densities corresponding to the L-proteins L16, L27, L28, L33, L35 and L36 in all cryo-EM maps of the 45S intermediates (see Figure 4.5). Additionally, the binding sites for L5 and L6 were partially distorted (see Figure 4.4E-H), albeit these two proteins were present in stoichiometric quantities in the 45S sample (see Figure 1A).

The EM-density for the ribosomal protein L6 was partially segmented, and the fitted x-ray coordinates of the protein (PDB ID: 2AW4), as a rigid body, did not fit completely in all cryo-EM maps of the 45S intermediates (see Figure 4.4E-H). In

contrast, the density attributed to L5 could not be visualized only in the two cryo-EM maps of the 45S intermediates with large distortions in the CP (see Figure 4.4G and H).

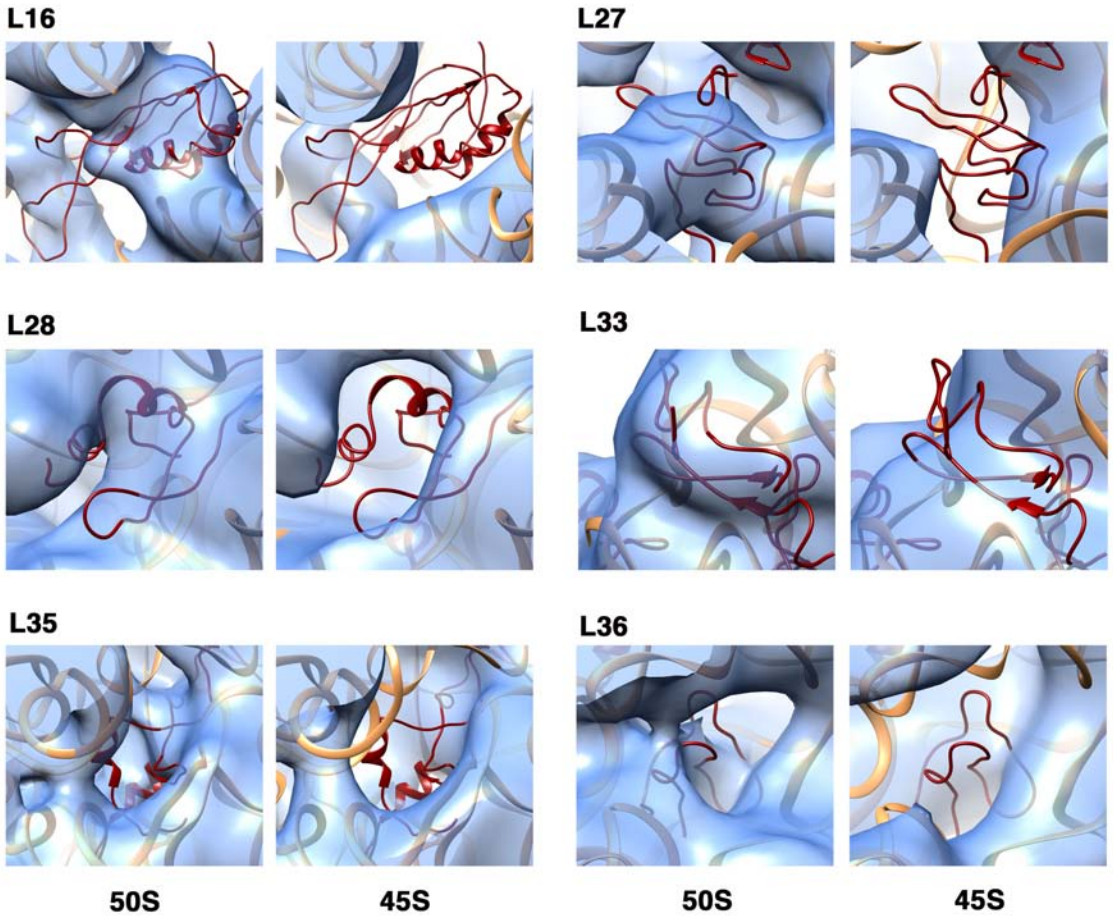


Figure 4.5 Lacking L-proteins portrayed from the cryo-EM map of the 45S subunit and are compared corresponding sites in the 50S subunit.

Cryo-EM densities are shown in blue, L-proteins are shown in red, and 23S rRNA is shown in orange. 45S intermediate 1 is only shown for demonstration.

Both ribosomal proteins are located within functionally important regions of the 50S subunit. The ribosomal protein L6 wraps like a clamp around H97 the L7/12 stalk, also forming contacts with H95 (Yusupov et al., 2001). The lobe formed by the ribosomal protein L6 and the H95 is known as the sarcin-ricin loop (SRL) (Gutell et al., 1992; Ban et al., 1998). The SRL contacts the G-domain of the translation factors EF-G and EF-Tu and is required for their ribosome-dependent GTPase activity (Bertina et al., 1973; Moazed et al., 1988; Wriggers et al., 2000). In contrast, the ribosomal protein L5 binding site in the CP region contacts H3 of the 5S rRNA and H84 of the 23S rRNA (Yusupov et al., 2001). This ribosomal protein interacts with the P-site tRNA (Yusupov et al., 2001) and was also demonstrated to be essential for the incorporation of 5S rRNA to the 50S subunit and the assembly of the CP region subunit (Rohl & Nierhaus, 1982; Korepanov et al., 2012).

The QMS experiments indicate full occupancy of L-proteins, L5 and L6, on the 45S subunit. Hence, we anticipate that the lack of the density of these two proteins in our cryo-EM maps is due to the flexibility in their binding sites. It is also possible that these proteins bind weakly to the immature 45S subunit and dissociate during the cryo-EM sample preparation procedure.

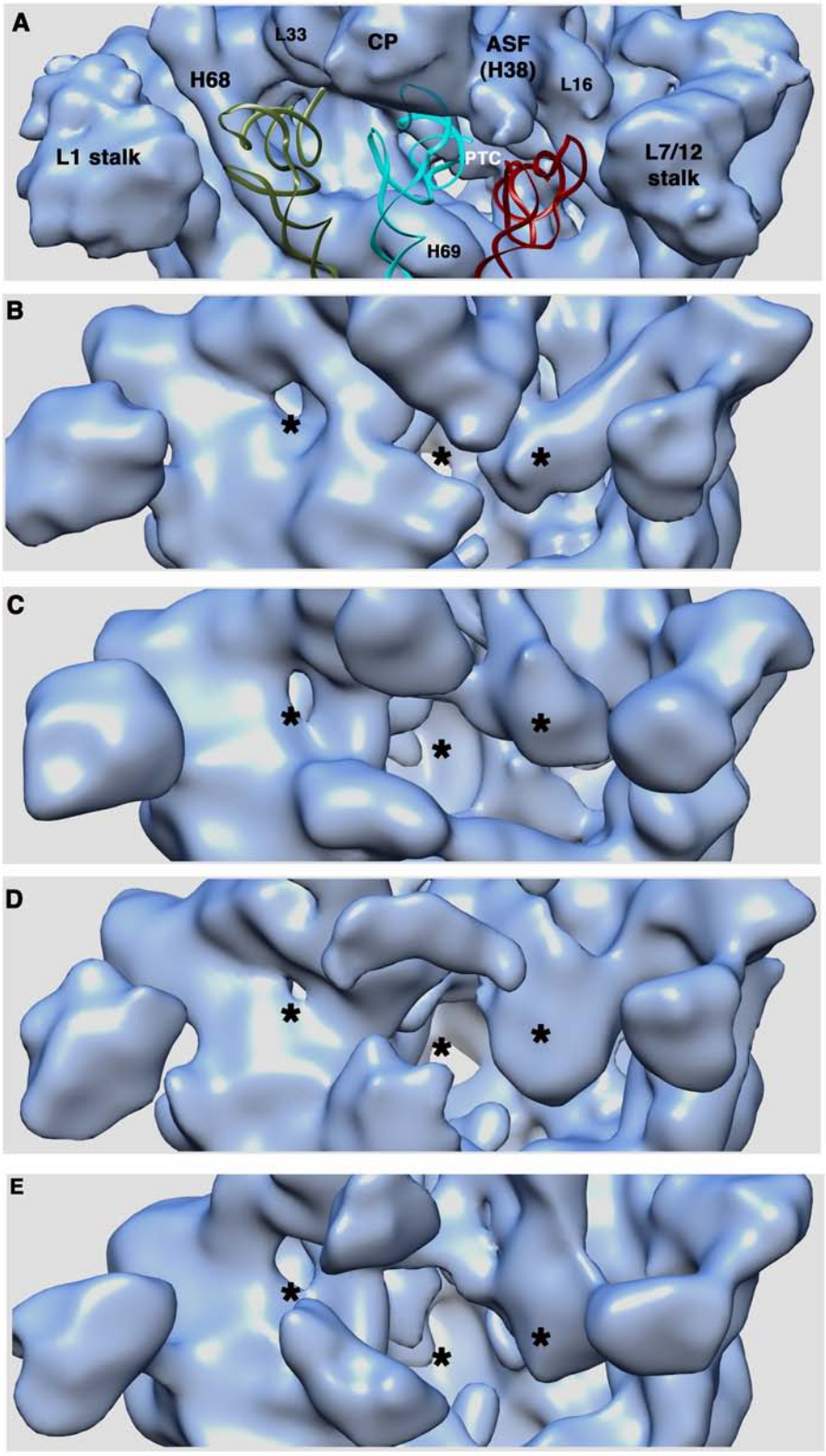
4.4.7 The tRNA binding sites, the PTC, and the exit tunnel are distorted in the 45S structure

Key rRNA helices in the CP, particularly H38 (A-site finger), H69, H81, H84, H85, H89, and H92 that also form part of the PTC and interact with A-site, P-site tRNAs,

and contacts 30S subunit are disordered in the cryo-EM maps 45S intermediates (Figure 4.6, 4.7C and F). In the structure of the mature 50S subunit, H81 and H85 form extensive contacts with the ribosomal L27, whereas the cleft formed by helix 38 and 89 forms the binding site for the ribosomal protein L16 (Yusupov et al., 2001). Furthermore, H89 make contacts with the L-protein L36 at the base of the L7/12 stalk (Yusupov et al., 2001). In addition, the displacement in H69, which would displace the inter-subunit bridge B2a, that contacts h44 of the 30S subunit and is essential for the association of the two subunits.

The acceptor stem of the E-site tRNA binds in a pocket formed by H68, H74, H75, H76, and H88, and the two ribosomal proteins, L28 and L33 (reviewed by Wilson & Nierhaus, 2006). The rRNA helices in this region are either partially or fully disordered in the cryo-EM maps of the 45S intermediates (Figure 4.6). Mainly, the EM-density corresponding to H68, a ~50Å long helix at the cleft between the L1 stalk and the CP, is completely disordered (Figure 4.6).

The exit tunnel was also not fully structured in all cryo-EM maps of the 45S intermediates (see Figure 4.6A-E). In particular, the EM-densities of the extended loops of L4 and L22, and segments of H71 and H74, were disordered, and thus the x-ray coordinates of these helices did not fit completely into the respective cryo-EM densities. Interestingly, the extended loops of L4 and L22 comprise part of the exit tunnel wall, which is predominantly composed of rRNA, forming a constriction site in the exit tunnel 30Å away from the PTC (Ban et al., 2000).



Continued

Figure 4.6 Immature 45S subunits contain distorted tRNA binding sites

Panel A depicts the cryo-EM map of the 50S subunit from IF2-depleted cells where the x-ray coordinates of the A (red), P (cyan), and E (green) tRNAs are overlaid on their respective binding site. Panels B-E depict the cryo-EM map of the 45S subunit classes 1 to 4. The asterisks are aligned with tRNAs acceptor stem binding sites on the 50S subunit.

Overall, the distortions observed in the functional core of the cryo-EM structures of the 45S subunit intermediate indicate that the 45S subunit adopt a non-functional conformational state. Mainly, this subunit cannot interact with tRNAs, translation factors, or the 30S subunit.

4.4.8 Cryo-EM structure of the L16-difficient 50S subunits is similar to the structure of the 45S subunit

We unexpectedly observed that the ribosomal protein L16 dissociated from the 50S subunit isolated under low magnesium concentrations (see Materials and Methods). To further analyze the effect of the dissociation of this protein on the conformational state of the 50S subunit, we analyzed the structure of the L16-deficient 50S subunit by cryo-EM.

Surprisingly, the structure of L16-deficient 50S subunit displayed a remarkable similarity to the cryo-EM structure of the 45S subunit, particularly the 45S intermediate with the lowest distortion in the CP (see Figure 4.4A, and 4.7). To further investigate the homogeneity of the L16-deficienct 50S sample, we exposed that dataset to ML3D classification. However, this approach did not render any new classes with distinct structural features.

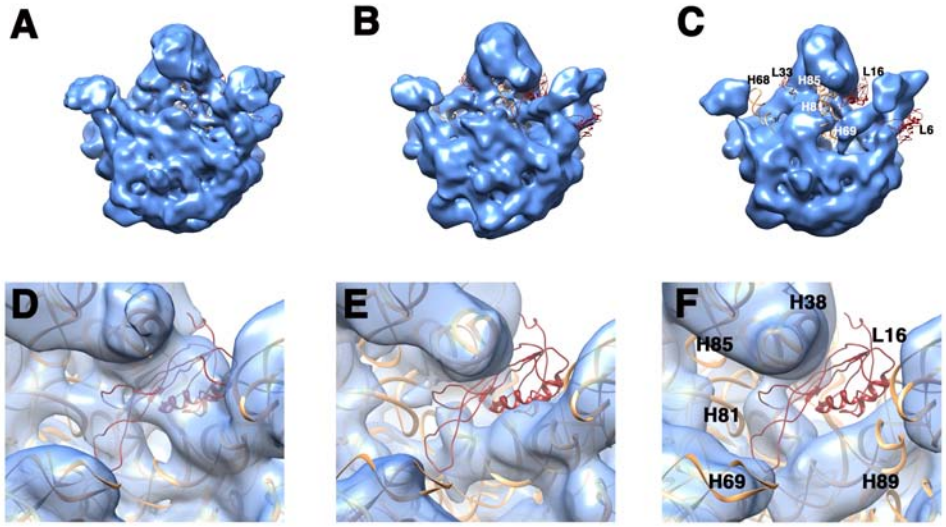


Figure 4.7 Cryo-EM reconstruction of the L16-deficient 50S subunit

The cryo-EM map of the 50S subunit from IF2-depleted cells (Panels A and D), the cryo-EM map of the L16-deficient 50S subunit (Panels B and E), and cryo-EM map of the 45S intermediate (Panels C and F) are displayed with the fitted coordinates of the rRNA helices (orange) and L-proteins (red, PDB ID 2AW4).

Consistent with the biochemical characterization of the L16-deficient 50S subunit, we could not identify a density attributed to ribosomal protein L16 in the cryo-EM structure. The rRNA helices, H89 and H38, which form the binding pocket of the ribosomal protein L16, were also displaced (see Figure 4.7 D-F). Furthermore, we could not indentify EM-densities corresponding to the ribosomal proteins L33, and L36. It is possible that these proteins also dissociated under low magnesium conditions. Interestingly, L33 and L36 were both late-binding L-proteins that were also lacking from the 45S subunits (see Figure 4.1A and Figure 4.5). In addition, a distortion in the binding site of the ribosomal protein L6 was also observed similar to the distortion identified in the cryo-EM structures of the 45S intermediates (see Figure 4.7A and B). This result

indicates that the binding site of L6 is dependent on the presence of L16 or on the concentration of magnesium ions.

Overall, our findings suggest that 50S subunit lacking the ribosomal protein L16 adopts a structure similar to that of the immature 45S subunit. This is consistent with the previously proposed role of L16 in stabilizing the final conformational state of the 50S subunit (Teraoka & Nierhaus, 1978; Nishimura et al., 2004). It is worth mentioning that the cryo-EM map and the image classification analysis of the L16-deficient 50S sample did not show any classes with distortions in the CP, which indicates that the distortion in this region is inherent to the 45S intermediates. Furthermore, this result suggests that the role of the RbgA protein during the assembly of the 50S subunit extends beyond monitoring the binding site of L16.

4.5 DISCUSSION

4.5.1 Immature 45S subunits lack late-binding proteins

In this study, we biochemically and structurally characterized an immature 45S subunit isolated from *B. subtilis* depleted of RbgA, We observed that this subunit contained incomplete protein content lacking L16, L27, L28, L33, L35 and L36. Interestingly, these proteins were demonstrated to bind at late stage of the *in-vitro* 50S reconstitution experiments (Rohl & Nierhaus, 1982; Herold & Nierhaus, 1987), and cluster to the base of the CP region. The L-protein content of the 45S subunit is similar to previously characterized 45S subunits isolated with either *E. coli* or *B. subitilis* with a mutation in the genes coding for 50S subunit assembly factors including DbpA (Sharpe

Elles et al., 2009), RrmJ (Bugl et al., 2000; Caldas et al., 2000), YphC, and YsxC (Schaefer et al., 2006). These findings indicate that these 45S subunits represent a common assembly intermediate that is recognized by multiple ribosome factors during late stages of ribosome assembly.

The structural analysis of the 45S subunit was consistent with QMS experiments, demonstrating that these late binding proteins were lacking. Interestingly, a distortion in the binding site of L5 was observed in two of the cryo-EM structures of the 45S intermediates, where QMS indicated stoichiometric quantities of this protein in the 45S sample. These intermediates also contained a large distortion in the CP region. The incorporation of the L5 protein to the assembling 50S subunit was implicated in the entry of 5S rRNA both *in-vitro* (Herold & Nierhaus, 1987) and *in-vivo* (Korepanov et al., 2012). The 5S rRNA, along with H81-88 of the 23SRNA, and L-proteins L5 and L18 composes the CP region of the 50S subunit (Ban et al., 2000). Thus, the distortion in the CP could indicate that the binding site of this protein is not yet structured and is adopting a flexible conformation. This observation also suggests a role for the RbgA protein in monitoring the folding of the CP including the binding site of L5.

In addition, the binding site of the ribosomal protein L6 was partially distorted in all cryo-EM structures of the 45S intermediates, particularly the EM-density attributed to its C-terminal domain. This domain forms intimate interactions with H97, in addition to H95, which contains the highly conserved SRL (Gutell et al., 1992). Recent biochemical and chemical probing studies demonstrated that the SRL is essential for the assembly of

the functional core of the 50S subunit (Lancaster et al., 2008). In the mature 50S subunit, the C-terminal end of L6 is also approached by H89 and L36 (Schuwirth et al., 2005). As explained earlier, H89 was displaced, while L36 was lacking from the cryo-EM map of the 45S subunit. This could impose an upstream or downstream effect on the binding site of L6 and the formation of the SRL during 50S assembly, and the folding of the functional core.

4.5.2 Immature 45S subunit adopts a non-functional conformational state which is similar to previously characterized immature 30S subunits

The cryo-EM structure of the 45S subunit contained distorted features throughout its 30S-interface region. The most striking feature was a distortion in the functional core of the 50S subunit, including its PTC, the exit tunnel, and the tRNA interactions sites. Additionally, our findings indicate that this immature 45S subunit is not capable of interacting with the 30S subunit due to the displacement of H69, which contains the essential inter-subunit bridge B2a. This is consistent with biochemical studies demonstrating that 45S subunits isolated from the RbgA-depleted strains from ultra-centrifugation sucrose gradients, migrate as free subunits and not in 70S ribosomes (Matsuo et al., 2006; Uicker et al., 2006).

The distorted functional regions of the immature 50S subunit portray common features with previously characterized immature 30S subunits. Recent cryo-EM studies of the immature 30S subunit that accumulated from *ΔyjeQ* (Jomaa et al., 2011a) and *ΔrimM* (Guo et al., 2013; leong et al., 2013), depicted a distorted mRNA decoding site and inter-subunit bridges. These results provide compelling evidence that the assembly of the 30S

subunit and 50S subunit progresses in a similar manner. In particular, the functional cores and the contact points of the ribosomal subunits will fold during late stages of assembly. Consistently, previous chemical foot-printing studies also indicated that rRNA helices in the functional regions fold at late stages during the *in-vitro* reconstitution of the 30S subunit (Holmes & Culver, 2004; Adilakshmi et al., 2008).

4.5.3 The L16-deficient 50S subunit adopts a structure with similar features to the structure of the 45S subunit

The cryo-EM structure of the L16-deficient 50S subunit displayed distortions in the PTC and tRNA binding sites. These distortions were similar to the cryo-EM maps of the 45S intermediates but to a lesser extent. Mainly, the rRNA helices (H38, H89) around the binding pocket of L16 were displaced. Previous biochemical studies have demonstrated that 50S subunit core particles treated by LiCl were deficient of the ribosomal protein L16 and lack peptidyltransferase activity (Teraoka & Nierhaus, 1978). The addition of L16 can reactivate these core particles under partial reconstitution conditions (Nierhaus & Montejo, 1973; Moore et al., 1975; de Bethune & Nierhaus, 1978), by inducing a large conformational change in their structure (Teraoka & Nierhaus, 1978).

Using cryo-EM, we were able visualize that the L16-deicient 50S subunit contained distortions similar to the structure of the 45S subunit. The L16-deficient 50S subunit however, did not contain a distorted CP similar to the one observed in the cryo-EM map of the 45S intermediate 3 and 4 (see Figure 4.4C and D). Hence, these results imply a functional role for RbgA as a checkpoint protein to monitor the correct folding of

the CP region, which includes the recruitment of L16 to the assembling 50S subunit. It is also possible that the conversion of the 45S subunit to a mature 50S subunit is induced by a combination of the binding event of RbgA and the recruitment of the ribosomal protein L16. The conformational change induced by this interaction could then signal the GTP-dependent release of RbgA.

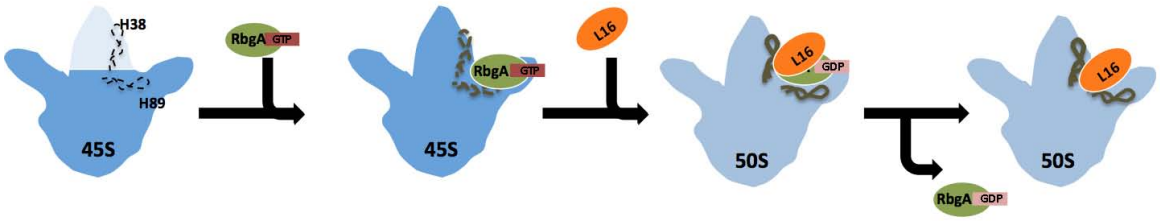


Figure 4.8 A model for the functional interplay between RbgA and L16

GTP-bound RbgA recognizes and binds to the 45S subunit. This event induces a conformational change in the CP and the binding site for L16 (H38 and H89). The ribosomal protein is then recruited to the 45S subunit. This induces the final confirmation state that would convert the 45S to mature 50S subunit and signals the GTP-dependent release of RbgA.

Based on these results we propose a novel functional interplay between RbgA and L16 during late stages of 50S assembly. Initially, the binding of RbgA to the assembling subunit is required for the structuring of the CP region. This event will be followed by the recruitment of L16, which induces a conformational change in the functional core of the 50S subunit. RbgA will then dissociate from the mature 50S subunit in a GTP dependent manner, after the maturation of the 50S subunit (see Figure 4.8 for proposed model).

4.6 MATERIAL AND METHODS

4.6.1 Isolation of large ribosomal subunits

Purification of the 50S ribosomes was carried out as previously described (Jomaa et al., 2011a). For 50S subunits isolated from *E. coli* cells (EB334), crude ribosomal pellets were washed with buffer F (10 mM Tris-Hcl pH 7.5, 1.1 mM magnesium acetate, 60 mM NH_4Cl, 0.5 mM EDTA, and 2 mM 2-mercaptoethanol) under dissociating conditions. A portion of the subunit suspension (50-60 A_{260} units) was layered onto a 32 ml 10-30% (wt/vol) sucrose gradient made up in buffer F and centrifuged at 43,000 x *g* for 16 hours using a Beckman SW32 Ti rotor. Gradients were fractionated using an AKTAprime™ purification system (GE Healthcare) and the elution peaks corresponding to 50S particles peaks were monitored at A_{260} and collected.

Isolation of the 50S and 45S ribosomal subunits from *B. subtilis* was carried out as previously described (Uicker et al., 2006), from IF2-depleted (RB419) and RbgA-depleted (RB301) strains, respectively. Bacterial strain RB301 (P_{spank}-*ylqF*) or RB419 (P_{spank}-*infB*) were grown on LB-agar plates O/N in the presence IPTG at 37°C. Cells were then scraped off the plate, re-suspended in 5 mL of LB media, and diluted in 1L media to a final OD_{600} of 0.05. Subsequently, cells were gown in the absence of IPTG at 37°C. OD_{600} readings were recorded every 30 minutes, and cells were diluted in LB media until a significant change in doubling time was observed. Cells were ultimately harvested at doubling time of 120 minutes at OD_{600} of 0.4 to obtain full depletion of the RbgA or IF2 proteins.

Cells were then lysed and spun down at 30,000 x g to remove cell debris as previously described (Jomaa et al., 2011a). Crude ribosomal pellets were washed with

Buffer E (10 mM Tris-Hcl pH 7.5, 15 mM magnesium acetate, 60 mM NH_4Cl, 0.5 mM EDTA, and 2 mM 2-mercaptoethanol) under non-dissociating conditions. Ribosome samples were layered onto a 32 ml 10-30% (wt/vol) sucrose gradient made up in buffer E and centrifuged at 43,000 x *g* for 16 hours using a Beckman SW32 Ti rotor. All isolated ribosomal subunits were stored in storage buffer (10 mM Tris-HCl at pH 7.5, 10 mM magnesium acetate, 60 mM NH_4Cl, 3 mM 2-mercaptoethanol) and kept frozen at -80°C until further use.

To dissociate the ribosomal protein L16, 50S subunits from IF2-depleted cells were lysed in buffer F (10 mM Tris-Hcl pH 7.5, 1.1 mM magnesium acetate, 60 mM KCl, 0.5 mM EDTA, and 2 mM 2-mercaptoethanol) under dissociating conditions. Samples were fractionated and isolated as described earlier. After the fractionation step, purified 50S ribosomal subunits were diluted to 1nM concentration in storage buffer (10mM Tris, 10mM $MgCl_2$ and 60mM KCl) and buffer exchanged by multiple rounds of centrifugation at 5000 x g using Amicon ultra 15 (100 KDa) cutoff filter tubes (Millipore). L16 dissociation from the ribosomal subunits was monitored on 12% Bis-Tris gels. Samples were then stored in storage buffer and kept frozen at -80°C until further use.

4.6.2 Quantitative Mass Spectrometry

^{15}N-labeled 70S ribosomes were prepared by growing *B. subtilis* strain ATCC 6051 aerobically in 1 L MSpitz9 medium (8.0 mM K_2HPO_4, 4.4 mM KH_2PO_4, 0.39 mM Na_3-citrate, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 5.6 mM glucose, 50 μM Na_3·EDTA, 25 mM $CaCl_2$, 50 μM $FeCl_3$, 0.5 μM $ZnSO_4$, 0.5 μM $CuSO_4$, 0.5 μM $MnSO_4$, 0.5 μM $CoCl_2$, 0.04

μM d-biotin, 0.02 μM folic acid, 0.08 μM vitamin B1, 0.11 μM calcium pantothenate, 0.4 nM vitamin B12, 0.2 μM nicotinamide, 0.07 μM riboflavin, **7.6 mM ($^{15}NH_4)_2SO_4$**) at 37 ºC. Cells were harvested at OD_{600} ~ 0.4 by directly adding the culture to 1 L of ice. After centrifugation at 6,000 x g for 20 minutes, cells were resuspended in buffer A (10 mM Tris (pH 7.8), 20 mM $MgCl_2$, 100 mM NH_4Cl, 6 mM 2-mercaptoethanol) and lysed using a BioSpec mini bead beater. The lysate was clarified *via* centrifugation (30,000 x g for 40 minutes) and loaded onto a 10%-40% (w/v) non-dissociating linear sucrose gradient containing 20 mM $MgCl_2$. The gradient was then centrifuged at 4 ºC in a Beckman SW32 rotor at 26,000 rpm for 28 hours. Fractions were collected using a Brandel gradient fractionator and those containing 70S particles were pooled and stored at 4 ºC.

^{14}N-labeled 70S ribosomes were purified as described above, substituting ^{14}N-ammonium sulfate in the growth medium. Samples were prepared for LC/MS analysis as described (Chen & Williamson, 2012). Briefly, 20 pmols of ^{15}N-labeled 70S spike was mixed with each experimental sample before the addition of trichloroacetic acid (TCA) to a final concentration of 13 % (v/v). Samples were incubated on ice overnight and precipitated material was isolated *via* centrifugation, washed with 10% TCA and finally resupended in 40 μL buffer B (100 mM NH_4CO_3, 5% acetonitrile, 5 mM dithiotreitol). After incubation for 10 minutes at 65 ºC, 10 mM iodoacetamide was added, and samples were incubated at 30 ºC for 30 minutes. Samples were then digested using 0.2 μg trypsin at 37 ºC overnight. Peptides were injected onto a C18 column and were eluted over 100 minutes across a concave 5-50% acetonitrile gradient and detected on an Agilent G1969A ESI-TOF mass spectrometer using a detection range set at 250-1,300 m/z. Unique

ribosomal peptides were identified in the raw data using a custom suite of software designed to identify $^{14}N/^{15}N$ peak pairs given a theoretical digest (Bunner et al., 2008). For each peptide, the quantities of ^{14}N or ^{15}N species were calculated by fitting the isotope distribution using a Least Square Fourier Transform Convolution algorithm (Sperling et al., 2008). The relative protein level for each r-protein was calculated as the ^{14}N fitted isotope distribution amplitude divided by the sum of the ^{14}N and ^{15}N fitted isotope distribution amplitudes ($^{14}N/(^{14}N+^{15}N)$). Each isotope distribution and its local chromatographic contour map was carefully examined and fits with low signal-to-noise ratios were excluded. To account for differences in the total amount of each experimental sample analyzed, each relative protein level was normalized to that of L24, a primary binding r-protein shown to be present in stoichiometric quantities in each sample analyzed (Figure 4.1A, 4.1B).

To improve protein coverage and quantification accuracy, samples were run multiple times and averaged across runs. Because proteins that are completely absent from an experimental sample cannot be identified by the $^{14}N/^{15}N$ peak pair, each experimental sample was also analyzed on a AB/Sciex Triple-TOF run in MS^2 mode. Peptides were eluted from analytical C18 nano-column across a 120 minute linear 5-40% acetonitrile gradient at 300 nL/min. Precursor ions were detected in a 400 to 1250 m/z mass window and product ions were detected with a 100-1800 m/z mass window. Precursor peptides were identified from the fragmentation data, using Mascot (Matrix Science) and relevant MS^1 scans were inspected to quantify the ^{14}N species (Figure 4.1D).

4.6.3 Cryo-electron microscopy

Ribosome samples were diluted to 16nM concentration and then 3.5 μL of the diluted sample was applied on 400 square/inch mesh copper grids. Grids were blotted and rapidly frozen by plunging in liquid ethane at liquid nitrogen temperature (~ -180°C) using FEI Virtobot™. Data acquisition was performed under low dose conditions (~15 e^- /Å^2) on a JEOL 2010F FEG microscope operated at 200 kV with a Gatan 914 side-entry cryo-holder and at a nominal magnification of 50,000x. The defocus range of the images was from -1 to -3.9 mm. The micrographs were digitized with a step size of 12.7 μm using a Nikon Supercool Scan 9000 producing images with a sampling value of 2.54 Å/pixel on the object scale.

4.6.4 Image processing

Projection images of the large ribosomal subunits were picked manually using Boxer (Ludtke et al., 1999). The total of number of particles picked for the *E. coli* 50S subunit (E50S), and the B. subtilis IF2-depleted 50S (B50S), RbgA-depleted 45S subunit were: 39,744; 53,053; and 76,499 particles, respectively. Additionally, 51,016 particles were picked for the mature 50S subunit with dissociated ribosomal protein L16 (F50S). Contrast transfer function of the micrographs was estimated using CTFFIND software (Mindell & Grigorieff, 2003) and corrected using the Xmipp software package (Scheres et al., 2008). The 3D reconstruction of the E50S and B50S subunits complex was calculated using 3D projection alignment procedures as implemented in the Xmipp software package (Scheres et al., 2008). The reference map used to refine the 3D reconstruction was the X-ray structure of the *E. coli* 50S subunit (PDB ID: 2AW4)

(Schuwirth et al., 2005) low-pass filtered to 25Å resolution. The resolution of the cryo-EM maps of the E50S and B50S was 12.7Å and 10.6Å using the FCS criteria at 0.5.

The dataset of immature 45S subunits was exposed to the Maximum-likelihood image classification (ML3D) approach for a total of 30 iterations, using the B50S map as a reference, low-pass filtered to 70Å. This approach rendered four distinct classes with particles (14-16Å resolution). Increasing the number of classes in the ML3D classifications produced classes with similar features. The total number of particles in each class was: 20,269; 10,990; 21,178; and 24,059. Each class was then refined using Maximum-likelihood refinement (Scheres et al., 2005c) approach followed projection matching approach using the cryo-EM map of the B50S low-pass filtered to 25Å.

The dataset obtained for the L16-deficient 50S subunit was refined following the projection matching approach using first Maximum-likelihood refinements to obtain an initial model then refinements using cross-correlation algorithms were used. The cryo-EM map of the B50S, low-pass filtered to 50Å, was used a reference. Further classification of this sample using ML3D did not produce classes with distinct structural features, indicating that the reconstruction was produced from a homogenous population. The final resolution of this map was 12.6Å. The X-ray coordinates of the *E. coli* 50S (PDB ID: 2AW4) were docked as a rigid body in all cryo-EM maps shown in this study. All figure illustrations were prepared using either Chimera (Pettersen et al., 2004).

CHAPTER 5. Discussion

5.1 Commonalties between immature 30S and 50S subunits

In this thesis, we structurally and biochemically characterized immature 30S and 50S subunits isolated from Δ*yjeQ* and RbgA-depleted cells, respectively. The overall structures of the immature 30S or 50S subunits depicted the classical structural features of the mature subunits. Hence, our findings indicate that the immature 30S and 50S subunit represent late stage assembly intermediates. A striking structural feature between the two subunits is the distortion in their functional regions. In the cryo-EM structure of the immature 30S subunit, the upper segment of helix 44, which forms part of the DC, and the entire helix 45, were displaced from their original sites and adopted a flexible conformation. In contrast, the cryo-EM structure of the immature 45S subunit contained distortions in the PTC, the CP, and its tRNA binding sites.

Our results are consistent with previous chemical probing studies analyzing *in-vitro* reconstruction experiments of the 30S subunit (Adilakshmi et al., 2008). In particular, helices dispersed in the body (5'domain) and the platform regions (central domain) and contacted by primary r- proteins were folded and protected from chemical reactivity early during assembly with 20-50 ms (rate constants > or = $20s^{-1}$). However, nucleotides close to the mRNA decoding site required the longest time to fold at around 1 min ($0.01\text{-}0.2s^{-1}$)(Adilakshmi et al., 2008). Changes in chemical reactivity of residues that are contained within functionally important regions in the 30S subunit were also monitored in the RI-to-RI* transition in 30S *in-vitro* reconstitution experiments (Holmes

& Culver, 2004). Most notably, nucleotides that are involved in tRNA and translation factors binding sites are folded and protected from chemical modifications during late stages of assembly (Holmes & Culver, 2004).

Several distortions were also observed in H68, 69, 81, 85, and 38 of the immature 50S subunit. These helices, particularly with the A-site finger (H38), make extensive contacts with the A- , P-, and E-site tRNAs. H69 contacts the penultimate stem h44 on the 30S subunit forming the essential inter-subunit bridge B2a (Yusupov et al., 2001; Schuwirth et al., 2005; Ali et al., 2006). Interestingly, both helices are displaced in the structure of the immature 50S and 30S subunits respectively. Furthermore, the distortion in helix 89 of the immature 50S subunit could affect key contacts with IF2, which also plays an essential role for 30S-50S subunit association during the initiation of translation (La Teana et al., 2001).

5.1.1 GTPases, key regulators of late stages of ribosomal assembly

Our findings present a novel role for the two GTPases, YjeQ and RbgA, acting at late stages during ribosomal assembly. Under this model, the immature 30S and 50S subunits, with the distorted DC and PTC respectively, are adopting non-functional conformational states. The two trans-acting GTPases, YjeQ and RbgA proteins, would recognize and bind the assembling ribosomal intermediates to monitor the folding of the functional regions. The final conformational state of the mature ribosomal subunits will then promote the GTP-dependent release of these factors once the assembly is completed.

GTPases have been proposed to act as intrinsic checkpoint proteins, where the coupling of ribosomal assembly to the intracellular GTP levels can be established (Britton, 2009). A decrease in the GTP/GDP ratios due to nutrient depletion would result in the arrest of ribosomal assembly probably by preventing these GTPases from interacting with the immature subunits. The production of the secondary messenger (p)ppGpp has also been proposed to partly control ribosome assembly by tightly binding GTPases under harsh environmental conditions (deLivron & Robinson, 2008; Potrykus & Cashel, 2008). Intriguingly, RbgA demonstrated a strong interaction with both mature and immature 50S subunit in the presence of the alarmone molecule, pppGpp (Achila et al., 2012).

Our results, in combination with previous biochemical data, envision a regulatory mechanism during late stages of ribosome assembly by GTPases. This would represent an energy efficient process to avoid any detrimental effects of premature initiation of protein synthesis by the assembling ribosomal subunits.

5.2 The parallel pathways of ribosomal assembly

The assembly of the ribosome was initially thought to proceed through a single pathway, which accumulates rate-limiting intermediates, similar to the ones that are observed *in-vitro* (Lindahl, 1975; Hayes & Vasseur, 1976; Sieber & Nierhaus, 1978). RAAFs would act on these intermediates by lowering the energy level required to transition to the mature state. Under non-optimal conditions in the cell, an assembly process governed by a global rate-limiting step could completely stall ribosome assembly.

The prevailing view that ribosome assembly proceeds through a single pathway was given away by more recent biochemical studies using *in-vitro* 30S reconstitution experiments. In these studies, it was demonstrated that 30S assembly proceeds through multiple local conformational transitions, where r-proteins bind to rRNA at similar rates independent of incubation temperatures (Talkington et al., 2005). Chemical-probing experiments later confirmed that multiple regions of naked rRNA can fold concomitantly and hence 30S assembly can nucleate from several places at once (Adilakshmi et al., 2008).

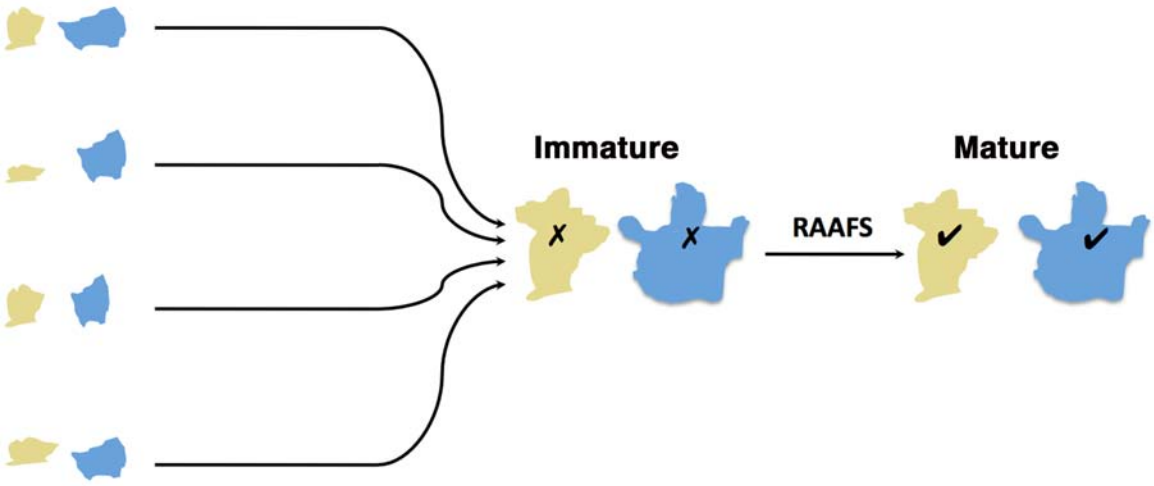


Figure 5.1 Schematic of the parallel assembly pathway of ribosome converging to a common intermediate recognized by RAAFs

Small and large ribosomal subunits are displayed in yellow and blue, respectively. DC and PTC indicated by the inactive and active state by the symbols (✗) and (✓), respectively.

In the light of these results, it was proposed that ribosomal assembly proceeds in parallel pathways rather than a single pathway (Culver, 2003; Woodson, 2008; Sykes & Williamson, 2009). This assembly pathway is also expected to proceed through the accumulation of multiple RI and multi-stage assembly kinetics (Noller & Nomura, 1987). Under this model, the role of RAAFs is predicted as checkpoint proteins that would monitor late stages of ribosomal assembly.

5.2.1 The observation of common intermediates before maturation

It has been generally observed that mutations in the RAAFs display similar defective phenotypes (Wilson & Nierhaus, 2007). For instance, immature 30S subunits isolated from mutant bacterial strain carrying a mutation in genes encoding for putative RAAFs accumulate immature 30S subunits, containing unprocessed rRNA (Bylund et al., 1998; Inoue et al., 2003; Himeno et al., 2004; Connolly et al., 2008; Goto et al., 2011) and incomplete S-protein content (Jomaa et al., 2011a; Guo et al., 2013; leong et al., 2013). A similar commonality aspect was also observed for the 50S assembly pathway, with immature 45S subunits travelling slower on ultra-centrifugation sucrose gradients and containing similar L-protein content (Schaefer et al., 2006; Uicker et al., 2006; Al Refaii & Alix, 2009; Sharpe Elles et al., 2009).

Two recent structural studies of immature 30S subunits isolated from *ΔrimM E. coli* strains demonstrated similar structural distorted features as the ones observed in our structure of the immature 30S subunit isolated from *ΔyjeQ* cells. Particularly the upper segment of h44, including the DC, was distorted (Guo et al., 2013; leong et al., 2013).

The S-protein content of the immature subunits was also incomplete, lacking predominantly late binding proteins such as S2, S3, and S21 (Guo et al., 2013; leong et al., 2013).

In sum, these findings suggest that the parallel pathways of ribosomal assembly converge at late stages of assembly where common 30S or 50S intermediates would accumulate (see Figure 5.1). These common assembly intermediates could represent as recognition templates for RAAFs that would act as checkpoint proteins to monitor the folding of functional regions of the ribosome and the recruitment of late-binding r-proteins.

5.3 Functional interplay between RAAFs during ribosomal assembly

The model we proposed for ribosomal assembly implicates the existence of a common immature state during late stages of assembly. There are two central questions under this model: How would different RAAFs act together to catalyze the immature intermediate to a mature stage? Do these factors bind this intermediate in a 'sequential' or 'simultaneous' model?

Under the sequential model, each factor would recognize a 'distinct' intermediate and would catalyze its maturation (see Figure 5.2). Under the simultaneous model however, several functionally related RAAFs would recognize a 'common' intermediate (see Figure 5.2), and thus factors would bind simultaneously with a defined hierarchy. In our opinion, the latter model emerges as a plausible energy-efficient process. The existence of a GTPase, within a "cluster" of RAAFS, could act as the energy driving

force of the complex. A GTP-bound state of the GTPase for instance, would signal the assembly of the RAAFs as a cluster on the assembling intermediate, where as a GDP-bound state of the GTPase would signal the dissociation of the cluster from the mature ribosomal subunit.

5.3.1 Functional interplay between YjeQ and other assembly factors

It has been demonstrated that YjeQ displaces RbfA on the mature but not the immature 30S subunit (Goto et al., 2011). Furthermore, the GTPase activity of YjeQ is stimulated by the mature 30S but not by the immature 30S subunit using *in-vitro* studies (Himeno et al., 2004). It is then possible that RbfA and YjeQ are locked during assembly, and YjeQ catalyzes RbfA release upon 30S maturation, in a GTP-dependent manner.

There are two available cryo-EM structures of 30S subunit solved in complex with YjeQ and GMPNP (Guo et al., 2011; Jomaa et al., 2011b). One of these structures was discussed in detail in Chapter 3 of this thesis. The cryo-EM structure solved by Guo *et al*. displayed a 360° rotation of the YjeQ protein in comparison to the structure discussed earlier. Particularly, the Zn finger domain interacts with the head domain of the 30S subunit, while the OB-fold domain binds the body in a similar fashion to initiation factor 1 (IF1) (Guo et al., 2011). Despite the two different orientations of YjeQ, the GTPase domain covers the upper segment of h44 in both cryo-EM structures, occluding the binding of tRNAs and translation factors (Guo et al., 2011; Jomaa et al., 2011b). The cryo-EM structure of the 30S subunit in complex with RbfA has also been determined (Datta et al., 2007). Although both YjeQ and RbfA bind close to the mRNA decoding

site, both factors do not share common contact points. In our cryo-EM structure of the YjeQ-30S complex, we observed a large conformational change in h44, where the upper segment of this helix adopts a flexible conformation. We attributed this movement to an electrostatic repulsion between the negatively charged residues of the C-terminal Zinc-motif of YjeQ and the negatively charged rRNA in the central segment of h44. Interestingly, the interaction of RbfA with the 30S subunit also induces a displacement of the upper segment of h44 and h45 (Datta et al., 2007). These two distinct conformational rearrangements in the mRNA DC site suggest that the C-terminal Zinc-motif of YjeQ plays a role in regulating the interaction of the RbfA-YjeQ interaction with the 30S subunit during its assembly or after maturation.

The maturation factor RimM, and the GTPase factor Era have also been linked to the function of YjeQ during ribosome assembly (Campbell & Brown, 2008). The cryo-EM structure of the Era-30S subunit placed Era on the cleft formed by the head and platform domains of the 30S subunit (Sharma et al., 2005). In contrast, biochemical studies indicated that RimM binds to the head of the 30S subunit (Lovgren et al., 2004). The binding sites of these two factors denote that they can also co-exist along with RbfA and YjeQ on the immature 30S subunit.

The methyltransferase KsgA binding site on the 30S subunit was also determined using cryo-EM. Interestingly, the binding site of KsgA overlaps with that of YjeQ, making similar contact points on h24a and h45 on the platform of the 30S subunit (Boehringer et al., 2012). While KsgA was identified as multi-copy suppressor of YjeQ

(Campbell & Brown, 2008), biochemical studies have demonstrated that KsgA interacts only with 'translationally-inactive' 30S subunits isolated under depleted Mg^{2+} ions concentration (Desai & Rife, 2006). Under this concentration, the penultimate stem h44 adopt a flexible conformation (Boehringer et al., 2012), and thus this conformation might not be favorable for the binding of YjeQ.

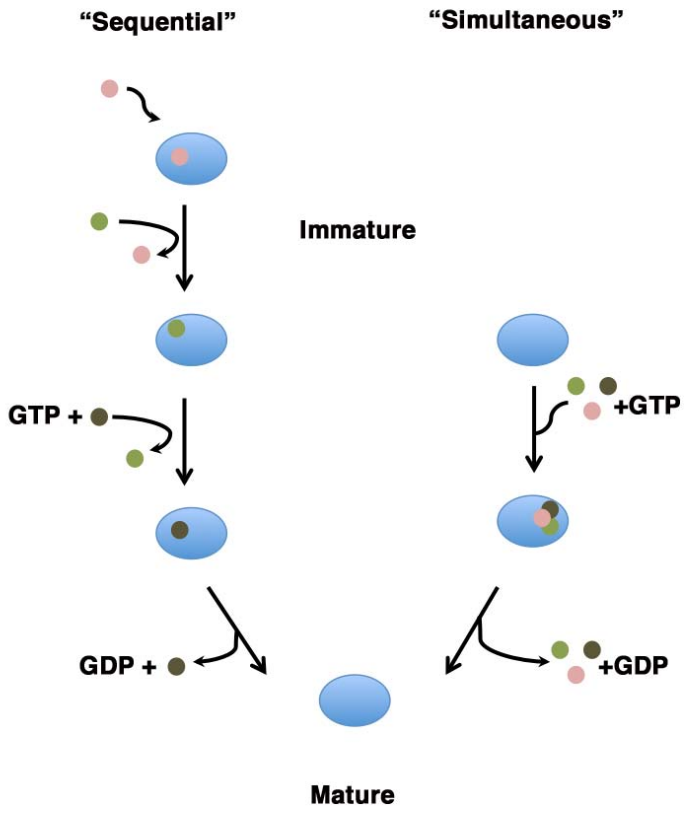


Figure 5.2 Schematic of the "Sequential" versus "Simultaneous" mode of RAAF recognition of the common maturation intermediate.

Immature and mature (30 or 50S) subunits are denoted by a blue circle. Small circles represent RAAFS.

Taken together, these results indicate that YjeQ and KsgA cannot bind simultaneously during ribosome assembly. It is then possible that the KsgA protein binds earlier during assembly to di-methylate the two nucleotides A1518 and A1519 on h45, and then YjeQ will replace it during later stages of assembly. More recently, a functional interplay was also demonstrated between KsgA and RbfA (Connolly & Culver, 2013). This study demonstrated that the over-expression of RbfA in Δ*ksgA* mutants accumulates 70S-like particles that are impaired in translation and lack the r-protein S21. It would be appealing to determine the hierarchy of the interaction of these factors on the immature 30S subunit.

In conclusion, our results in combination with previous structural and biochemical studies support a model where clusters of functionally related RAAFS could act simultaneously on assembling the 30S subunit. An attractive role the two GTPases (Era and YjeQ) play can be incorporated into this model, where they can act as 'ON/OFF' switches to release the remaining assembly factors (RimM and RbfA) from the 30S subunit once its maturation step is completed.

5.4 True intermediates or kinetic-traps

A perturbation in the role of RAAFs could affect cell growth, translational efficiency, and could lead to accumulation of either immature 30S subunits or 50S subunits with a concomitant drop in the number of 70S ribosomes. These immature subunits generally contain precursor rRNA and lack late-binding r-proteins. We demonstrated that immature subunits also contain distorted structural features in their

functional regions, which indicate that these regions are not yet folded. It is likely that the accumulated immature 30S or 50S subunits represent true assembly intermediates. Consequently, they would be able to interact with RAAFs as well as transition to the mature state. Another possibility is that the immature 30S or 50S subunits represent off-path kinetic traps. In this case, these immature subunits are in a misfolded state that cannot interact with RAAFs and hence cannot progress to the mature state.

Several biochemical studies have confirmed the interaction of RAAFs with immature 30S or 50S subunits. Immature 30S subunits that accumulate from *ΔyjeQΔrbfA* were verified to form a stable interaction *in-vitro* with RbfA but not with YjeQ (Goto et al., 2011). In contrast, immature 50S subunits isolated RbgA-depleted cells were confirmed to interact with RbgA in the presence of both GTP and GMPPNP (Matsuo et al., 2006; Uicker et al., 2006; Achila et al., 2012). The DEAD-box proteins SrmB and CsdA selectively interact with immature 50S particles isolated from *ΔsrmB* and *ΔcsdA* mutant stains, respectively, over mature 50S particles *in-vitro* (Charollais et al., 2003; Charollais et al., 2004). Additionally, an *in-vivo* interaction of immature 50S particles, isolated from *E. coli* strains expressing an inactive mutant of the $CtgA_E$ protein, was also demonstrated with the assembly factors: CsdA, RluB, and Yhb (Jiang et al., 2006).

There have been contradicting results addressing the ability of the immature 30S subunit or 50S subunit to transition to the mature state. Degradation of immature 30S subunits, containing 17S rRNA, was observed from YqeH-depleted cells (Loh et al., 2007). Additionally, supplementing the DbpA protein to cell lysates containing immature

45S subunits showed no recovery in ribosome profiles, which indicated that these subunits did not progress to the mature state (Sharpe Elles et al., 2009). Nevertheless, immature 50S subunits that accumulate Δ*csdA* mutant strains were demonstrated to progress to 70S particles *in-vivo* (Peil et al., 2008). Similarly, immature 50S subunits from Δ*dnaK* or Δ*dnaJ* mutant strains can be converted to mature 50S particles by the GRoEL-GroES complex (El Hage et al., 2001).

The different conformational states observed in our cryo-EM studies of the immature 30S subunits denoted the existence of several intermediates with a different degree of distortion in the DC. The two intermediates with a lesser distortion in the DC center also exhibited a higher occupancy in the S2 protein density. These results are in line with our hypothesis that the isolated immature 30S subunits form an ensemble of immature particles that are slowly progressing to the mature state. Likewise, the image classification analysis of the immature 45S subunits resulted in four intermediates with different distortions in the CP region. One of the classes contained similar features to the mature 50S structure except for a small distortion in H89, which forms extensive contacts with the ribosomal protein L16 (Yusupov et al., 2001). This result also indicates that 45S subunits represent a group of assembly intermediates progressing to the mature state.

In the line of these results, we believe it is crucial to establish biochemical assays to monitor the maturation of these intermediates. One approach of investigating this further is by performing pulse-chase experiments in order to analyze precursor rRNA processing, which accumulates in assembly intermediates. This would be ideal for tracing

unprocessed 16S rRNA that accumulate in immature 30S subunits. In fact, pulse-chase experiments were employed previously to characterize the processing of the 17S rRNA by the RNAses E and G (Li et al., 1999b).

A second approach would be to utilize stable isotope labeling of r-proteins followed by quantitative mass spectrometry (QMS) over time intervals to monitor the maturation of the pre-50S and 30S subunits (Chen et al., 2012; Chen & Williamson, 2012). If the immature subunits represent a bona fide intermediate on the pathway to the mature state, they will act as a sink for the label and the 70S ribosomes will show a delay in labeling during the QMS experiments.

In conclusion, these two future experiments will be essential to conclusively confirm whether immature ribosomal subunits represent true intermediates in the ribosome assembly pathway.

5.5 Significance

The work presented in this thesis provides novel biochemical and structural evidence on the *in-vivo* ribosome assembly pathway. We were able to determine for the first time the cryo-EM structures of immature 30S and 50S subunits isolated from bacterial mutant strains. We observed that both structures contained a distorted functional core. Our results suggest that these regions fold in late stages of ribosome assembly, probably to prevent immature initiation of protein synthesis.

We established novel roles for the GTPases YjeQ and RbgA in monitoring the folding of the DC and the PTC, respectively. Furthermore, we determined the binding site

of the YjeQ protein to the mRNA DC region of the 30S subunit, occluding tRNA and translation factors binding sites. We believe that YjeQ will act as a checkpoint protein during assembly locking the immature 30S subunit in an inactive state. Additionally, the conformational rearrangement of the upper segment of h44 induced by the interaction of the C-terminal Zn finger domain of YjeQ proposes a mechanism of displacing RbfA.

We also established a novel functional interplay between the RbgA protein and the ribosomal protein L16 during late stages of 50S subunit assembly. We observed that both proteins contribute to the conformational change observed in immature 45S subunits. Moreover, we believe that the conformational change induced by binding of L16 would lead to the GTP-dependent release of RbgA.

This thesis presents tantalizing clues to the mechanism of the RAAFs role in the *in-vivo* ribosome assembly. Nevertheless, structural and biochemical analysis, on the interaction of both immature small and large subunits with multiple RAAFs, will be required to provide a molecular-level understanding of how these factors contribute to building ribosomes in the cell.

REFERENCES

Achila D, Gulati M, Jain N, Britton RA. 2012. Biochemical characterization of ribosome assembly GTPase RbgA in Bacillus subtilis. *J Biol Chem 287*:8417-8423.

Adilakshmi T, Bellur DL, Woodson SA. 2008. Concurrent nucleation of 16S folding and induced fit in 30S ribosome assembly. *Nature 455*:1268-1272.

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

Aebi U, Pollard TD. 1987. A glow discharge unit to render electron microscope grids and other surfaces hydrophilic. *J Electron Microsc Tech 7*:29-33.

Agalarov SC, Selivanova OM, Zheleznyakova EN, Zheleznaya LA, Matvienko NI, Spirin AS. 1999. Independent in vitro assembly of all three major morphological parts of the 30S ribosomal subunit of Thermus thermophilus. *Eur J Biochem 266*:533-537.

Agalarov SC, Zheleznyakova EN, Selivanova OM, Zheleznaya LA, Matvienko NI, Vasiliev VD, Spirin AS. 1998. In vitro assembly of a ribonucleoprotein particle corresponding to the platform domain of the 30S ribosomal subunit. *Proc Natl Acad Sci U S A 95*:999-1003.

Al Refaii A, Alix JH. 2009. Ribosome biogenesis is temperature-dependent and delayed in Escherichia coli lacking the chaperones DnaK or DnaJ. *Mol Microbiol 71*:748-762.

Ali IK, Lancaster L, Feinberg J, Joseph S, Noller HF. 2006. Deletion of a conserved, central ribosomal intersubunit RNA bridge. *Mol Cell 23*:865-874.

Allen GS, Zavialov A, Gursky R, Ehrenberg M, Frank J. 2005. The cryo-EM structure of a translation initiation complex from Escherichia coli. *Cell 121*:703-712.

Antoun A, Pavlov MY, Andersson K, Tenson T, Ehrenberg M. 2003. The roles of initiation factor 2 and guanosine triphosphate in initiation of protein synthesis. *EMBO J 22*:5593-5601.

Antoun A, Pavlov MY, Tenson T, Ehrenberg MM. 2004. Ribosome formation from subunits studied by stopped-flow and Rayleigh light scattering. *Biol Proced Online 6*:35-54.

Ban N, Freeborn B, Nissen P, Penczek P, Grassucci RA, Sweet R, Frank J, Moore PB, Steitz TA. 1998. A 9 A resolution X-ray crystallographic map of the large ribosomal subunit. *Cell 93*:1105-1115.

Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science 289*:905-920.

Bashan A, Agmon I, Zarivach R, Schluenzen F, Harms J, Berisio R, Bartels H, Franceschi F, Auerbach T, Hansen HA, Kossoy E, Kessler M, Yonath A. 2003. Structural basis of the ribosomal machinery for peptide bond formation, translocation, and nascent chain progression. *Mol Cell 11*:91-102.

Bertina RM, Schrier PI, Slater EC. 1973. The binding of aurovertin to mitochondria, and its effect on mitochondrial respiration. *Biochim Biophys Acta 305*:503-518.

Bharat A, Jiang M, Sullivan SM, Maddock JR, Brown ED. 2006. Cooperative and critical roles for both G domains in the GTPase activity and cellular function of ribosome-associated Escherichia coli EngA. *J Bacteriol 188*:7992-7996.

Bleichert F, Baserga SJ. 2007. The long unwinding road of RNA helicases. *Mol Cell 27*:339-352.

Boehringer D, O'Farrell HC, Rife JP, Ban N. 2012. Structural insights into methyltransferase KsgA function in 30S ribosomal subunit biogenesis. *J Biol Chem 287*:10453-10459.

Borovinskaya MA, Pai RD, Zhang W, Schuwirth BS, Holton JM, Hirokawa G, Kaji H, Kaji A, Cate JH. 2007. Structural basis for aminoglycoside inhibition of bacterial ribosome recycling. *Nat Struct Mol Biol 14*:727-732.

Brimacombe R, Atmadja J, Stiege W, Schuler D. 1988. A detailed model of the three-dimensional structure of Escherichia coli 16 S ribosomal RNA in situ in the 30 S subunit. *J Mol Biol 199*:115-136.

Britton RA. 2009. Role of GTPases in bacterial ribosome assembly. *Annu Rev Microbiol 63*:155-176.

Brosius J, Dull TJ, Noller HF. 1980. Complete nucleotide sequence of a 23S ribosomal RNA gene from Escherichia coli. *Proc Natl Acad Sci U S A 77*:201-204.

Brosius J, Palmer ML, Kennedy PJ, Noller HF. 1978. Complete nucleotide sequence of a 16S ribosomal RNA gene from Escherichia coli. *Proc Natl Acad Sci U S A 75*:4801-4805.

Brown ED. 2005. Conserved P-loop GTPases of unknown function in bacteria: an emerging and vital ensemble in bacterial physiology. *Biochem Cell Biol 83*:738-746.

Bubunenko M, Korepanov A, Court DL, Jagannathan I, Dickinson D, Chaudhuri BR, Garber MB, Culver GM. 2006. 30S ribosomal subunits can be assembled in vivo without primary binding ribosomal protein S15. *RNA 12*:1229-1239.

Bugl H, Fauman EB, Staker BL, Zheng F, Kushner SR, Saper MA, Bardwell JC, Jakob U. 2000. RNA methylation under heat shock control. *Mol Cell 6*:349-360.

Bunner AE, Trauger SA, Siuzdak G, Williamson JR. 2008. Quantitative ESI-TOF analysis of macromolecular assembly kinetics. *Anal Chem 80*:9379-9386.

Bylund GO, Wipemo LC, Lundberg LA, Wikstrom PM. 1998. RimM and RbfA are essential for efficient processing of 16S rRNA in Escherichia coli. *J Bacteriol 180*:73-82.

Caldas T, Binet E, Bouloc P, Richarme G. 2000. Translational defects of Escherichia coli mutants deficient in the Um(2552) 23S ribosomal RNA methyltransferase RrmJ/FTSJ. *Biochem Biophys Res Commun 271*:714-718.

Caldon CE, Yoong P, March PE. 2001. Evolution of a molecular switch: universal bacterial GTPases regulate ribosome function. *Mol Microbiol 41*:289-297.

Campbell TL, Brown ED. 2008. Genetic interaction screens with ordered overexpression and deletion clone sets implicate the Escherichia coli GTPase YjeQ in late ribosome biogenesis. *J Bacteriol 190*:2537-2545.

Campbell TL, Daigle DM, Brown ED. 2005. Characterization of the Bacillus subtilis GTPase YloQ and its role in ribosome function. *Biochem J 389*:843-852.

Campbell TL, Henderson J, Heinrichs DE, Brown ED. 2006. The yjeQ gene is required for virulence of Staphylococcus aureus. *Infect Immun 74*:4918-4921.

Carter AP, Clemons WM, Jr., Brodersen DE, Morgan-Warren RJ, Hartsch T, Wimberly BT, Ramakrishnan V. 2001. Crystal structure of an initiation factor bound to the 30S ribosomal subunit. *Science 291*:498-501.

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

Charollais J, Pflieger D, Vinh J, Dreyfus M, Iost I. 2003. The DEAD-box RNA helicase SrmB is involved in the assembly of 50S ribosomal subunits in Escherichia coli. *Mol Microbiol 48*:1253-1265.

Chen SS, Sperling E, Silverman JM, Davis JH, Williamson JR. 2012. Measuring the dynamics of E. coli ribosome biogenesis using pulse-labeling and quantitative mass spectrometry. *Mol Biosyst 8*:3325-3334.

Chen SS, Williamson JR. 2012. Characterization of the Ribosome Biogenesis Landscape in E. coli Using Quantitative Mass Spectrometry. *J Mol Biol*.

Chen X, Walker AK, Strahler JR, Simon ES, Tomanicek-Volk SL, Nelson BB, Hurley MC, Ernst SA, Williams JA, Andrews PC. 2006. Organellar proteomics: analysis of pancreatic zymogen granule membranes. *Mol Cell Proteomics 5*:306-312.

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD. 2003. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res 31*:3497-3500.

Connolly K, Culver G. 2009. Deconstructing ribosome construction. *Trends Biochem Sci 34*:256-263.

Connolly K, Culver G. 2013. Overexpression of RbfA in the absence of the KsgA checkpoint results in impaired translation initiation. *Mol Microbiol 87*:968-981.

Connolly K, Rife JP, Culver G. 2008. Mechanistic insight into the ribosome biogenesis functions of the ancient protein KsgA. *Mol Microbiol 70*:1062-1075.

Crick FH. 1966. Codon--anticodon pairing: the wobble hypothesis. *J Mol Biol 19*:548-555.

Culver GM. 2003. Assembly of the 30S ribosomal subunit. *Biopolymers 68*:234-249.

Culver GM, Noller HF. 2000. In vitro reconstitution of 30S ribosomal subunits using complete set of recombinant proteins. *Methods Enzymol 318*:446-460.

Cumberlidge AG, Isono K. 1979. Ribosomal protein modification in Escherichia coli. I. A mutant lacking the N-terminal acetylation of protein S5 exhibits thermosensitivity. *J Mol Biol 131*:169-189.

Daigle DM, Brown ED. 2004. Studies of the interaction of Escherichia coli YjeQ with the ribosome in vitro. *J Bacteriol 186*:1381-1387.

Daigle DM, Rossi L, Berghuis AM, Aravind L, Koonin EV, Brown ED. 2002. YjeQ, an essential, conserved, uncharacterized protein from Escherichia coli, is an unusual GTPase with circularly permuted G-motifs and marked burst kinetics. *Biochemistry 41*:11109-11117.

Dallas A, Noller HF. 2001. Interaction of translation initiation factor 3 with the 30S ribosomal subunit. *Mol Cell 8*:855-864.

Datta PP, Wilson DN, Kawazoe M, Swami NK, Kaminishi T, Sharma MR, Booth TM, Takemoto C, Fucini P, Yokoyama S, Agrawal RK. 2007. Structural aspects of RbfA action during small ribosomal subunit assembly. *Mol Cell 28*:434-445.

de Bethune MP, Nierhaus KH. 1978. Characterisation of the binding of virginiamycin S to Escherichia coli ribosomes. *Eur J Biochem 86*:187-191.

Decatur WA, Fournier MJ. 2002. rRNA modifications and ribosome function. *Trends Biochem Sci 27*:344-351.

Del Campo M, Kaya Y, Ofengand J. 2001. Identification and site of action of the remaining four putative pseudouridine synthases in Escherichia coli. *RNA 7*:1603-1615.

deLivron MA, Robinson VL. 2008. Salmonella enterica serovar Typhimurium BipA exhibits two distinct ribosome binding modes. *J Bacteriol 190*:5944-5952.

Desai PM, Rife JP. 2006. The adenosine dimethyltransferase KsgA recognizes a specific conformational state of the 30S ribosomal subunit. *Arch Biochem Biophys 449*:57-63.

Dez C, Tollervey D. 2004. Ribosome synthesis meets the cell cycle. *Curr Opin Microbiol 7*:631-637.

Dlakic M. 2005. The ribosomal subunit assembly line. *Genome Biol 6*:234.

Doi Y, Arakawa Y. 2007. 16S ribosomal RNA methylation: emerging resistance mechanism against aminoglycosides. *Clin Infect Dis 45*:88-94.

Dorner S, Panuschka C, Schmid W, Barta A. 2003. Mononucleotide derivatives as ribosomal P-site substrates reveal an important contribution of the 2'-OH to activity. *Nucleic Acids Res 31*:6536-6542.

Doudna JA, Rath VL. 2002. Structure and function of the eukaryotic ribosome: the next frontier. *Cell 109*:153-156.

Ehrenberg M, Blomberg C. 1980. Thermodynamic constraints on kinetic proofreading in biosynthetic pathways. *Biophys J 31*:333-358.

El Hage A, Sbai M, Alix JH. 2001. The chaperonin GroEL and other heat-shock proteins, besides DnaK, participate in ribosome biogenesis in Escherichia coli. *Mol Gen Genet 264*:796-808.

Ellwood M, Nomura M. 1982. Chromosomal locations of the genes for rRNA in Escherichia coli K-12. *J Bacteriol 149*:458-468.

Frank J. 2003. Toward an understanding of the structural basis of translation. *Genome Biol 4*:237.

Frank J, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadj M, Leith A. 1996. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol 116*:190-199.

Frank J, Verschoor A, Li Y, Zhu J, Lata RK, Radermacher M, Penczek P, Grassucci R, Agrawal RK, Srivastava S. 1995a. A model of the translational apparatus based on a three-dimensional reconstruction of the Escherichia coli ribosome. *Biochem Cell Biol 73*:757-765.

Frank J, Zhu J, Penczek P, Li Y, Srivastava S, Verschoor A, Radermacher M, Grassucci R, Lata RK, Agrawal RK. 1995b. A model of protein synthesis based on cryo-electron microscopy of the E. coli ribosome. *Nature 376*:441-444.

Galperin MY, Koonin EV. 2004. 'Conserved hypothetical' proteins: prioritization of targets for experimental study. *Nucleic Acids Res 32*:5452-5463.

Gao H, Valle M, Ehrenberg M, Frank J. 2004. Dynamics of EF-G interaction with the ribosome explored by classification of a heterogeneous cryo-EM dataset. *J Struct Biol 147*:283-290.

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

Guo Q, Goto S, Chen Y, Feng B, Xu Y, Muto A, Himeno H, Deng H, Lei J, Gao N. 2013. Dissecting the in vivo assembly of the 30S ribosomal subunit reveals the role of RimM and general features of the assembly process. *Nucleic Acids Res*.

Guo Q, Yuan Y, Xu Y, Feng B, Liu L, Chen K, Sun M, Yang Z, Lei J, Gao N. 2011. Structural basis for the function of a small GTPase RsgA on the 30S ribosomal subunit maturation revealed by cryoelectron microscopy. *Proc Natl Acad Sci U S A 108*:13100-13105.

Gutell RR, Schnare MN, Gray MW. 1992. A compilation of large subunit (23S- and 23S-like) ribosomal RNA structures. *Nucleic Acids Res 20 Suppl*:2095-2109.

Hansen JL, Schmeing TM, Moore PB, Steitz TA. 2002. Structural insights into peptide bond formation. *Proc Natl Acad Sci U S A 99*:11670-11675.

Harms J, Schluenzen F, Zarivach R, Bashan A, Gat S, Agmon I, Bartels H, Franceschi F, Yonath A. 2001. High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell 107*:679-688.

Hayes F, Vasseur M. 1976. Processing of the 17-S Escherichia coli precursor RNA in the 27-S pre-ribosomal particle. *Eur J Biochem 61*:433-442.

Held WA, Ballou B, Mizushima S, Nomura M. 1974. Assembly mapping of 30 S ribosomal proteins from Escherichia coli. Further studies. *J Biol Chem 249*:3103-3111.

Held WA, Mizushima S, Nomura M. 1973. Reconstitution of Escherichia coli 30 S ribosomal subunits from purified molecular components. *J Biol Chem 248*:5720-5730.

Helser TL, Davies JE, Dahlberg JE. 1972. Mechanism of kasugamycin resistance in Escherichia coli. *Nat New Biol 235*:6-9.

Herold M, Nierhaus KH. 1987. Incorporation of six additional proteins to complete the assembly map of the 50 S subunit from Escherichia coli ribosomes. *J Biol Chem 262*:8826-8833.

Herold M, Nowotny V, Dabbs ER, Nierhaus KH. 1986. Assembly analysis of ribosomes from a mutant lacking the assembly-initiator protein L24: lack of L24 induces temperature sensitivity. *Mol Gen Genet 203*:281-287.

Himeno H, Hanawa-Suetsugu K, Kimura T, Takagi K, Sugiyama W, Shirata S, Mikami T, Odagiri F, Osanai Y, Watanabe D, Goto S, Kalachnyuk L, Ushida C, Muto A. 2004. A novel GTPase activated by the small subunit of ribosome. *Nucleic Acids Res 32*:5303-5309.

Holmes KL, Culver GM. 2004. Mapping structural differences between 30S ribosomal subunit assembly intermediates. *Nat Struct Mol Biol 11*:179-186.

Hopfield JJ. 1974. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc Natl Acad Sci U S A 71*:4135-4139.

Hosokawa K, Fujimura RK, Nomura M. 1966. Reconstitution of functionally active ribosomes from inactive subparticles and proteins. *Proc Natl Acad Sci U S A 55*:198-204.

Hosokawa K, Nomura M. 1965. Incomplete Ribosomes Produced in Chloramphenicol- and Puromycin-Inhibited Escherichia Coli. *J Mol Biol 12*:225-241.

Hwang J, Inouye M. 2006. The tandem GTPase, Der, is essential for the biogenesis of 50S ribosomal subunits in Escherichia coli. *Mol Microbiol 61*:1660-1672.

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

Inoue K, Chen J, Tan Q, Inouye M. 2006. Era and RbfA have overlapping function in ribosome biogenesis in Escherichia coli. *J Mol Microbiol Biotechnol 11*:41-52.

Jain C. 2008. The E. coli RhlE RNA helicase regulates the function of related RNA helicases during ribosome assembly. *RNA 14*:381-389.

Jiang M, Datta K, Walker A, Strahler J, Bagamasbad P, Andrews PC, Maddock JR. 2006. The Escherichia coli GTPase CgtAE is involved in late steps of large ribosome assembly. *J Bacteriol 188*:6757-6770.

Jiang M, Sullivan SM, Walker AK, Strahler JR, Andrews PC, Maddock JR. 2007. Identification of novel Escherichia coli ribosome-associated proteins using isobaric tags and multidimensional protein identification techniques. *J Bacteriol 189*:3434-3444.

Jomaa A, Stewart G, Martin-Benito J, Zielke R, Campbell TL, Maddock JR, Brown ED, Ortega J. 2011a. Understanding ribosome assembly: the structure of in vivo assembled immature 30S subunits revealed by cryo-electron microscopy. *RNA 17*:697-709.

Jomaa A, Stewart G, Mears JA, Kireeva I, Brown ED, Ortega J. 2011b. Cryo-electron microscopy structure of the 30S subunit in complex with the YjeQ biogenesis factor. *RNA 17*:2026-2038.

Jones PG, Inouye M. 1996. RbfA, a 30S ribosomal binding factor, is a cold-shock protein whose absence triggers the cold-shock response. *Mol Microbiol 21*:1207-1218.

Kaczanowska M, Ryden-Aulin M. 2007. Ribosome biogenesis and the translation process in Escherichia coli. *Microbiol Mol Biol Rev 71*:477-494.

Karbstein K. 2007. Role of GTPases in ribosome assembly. *Biopolymers 87*:1-11.

Kim do J, Jang JY, Yoon HJ, Suh SW. 2008. Crystal structure of YlqF, a circularly permuted GTPase: implications for its GTPase activation in 50 S ribosomal subunit assembly. *Proteins 72*:1363-1370.

Kimura T, Takagi K, Hirata Y, Hase Y, Muto A, Himeno H. 2008. Ribosome-small-subunit-dependent GTPase interacts with tRNA-binding sites on the ribosome. *J Mol Biol 381*:467-477.

Kimura T, Takagi K, Kyoko HS, Kalachnyuk L, Muto A, Himeno H. 2007. Interaction between RsgA and the ribosome. *Nucleic Acids Symp Ser (Oxf)*:375-376.

Korepanov AP, Korobeinikova AV, Shestakov SA, Garber MB, Gongadze GM. 2012. Protein L5 is crucial for in vivo assembly of the bacterial 50S ribosomal subunit central protuberance. *Nucleic Acids Res 40*:9153-9159.

Kozak M. 1999. Initiation of translation in prokaryotes and eukaryotes. *Gene 234*:187-208.

La Teana A, Gualerzi CO, Dahlberg AE. 2001. Initiation factor IF 2 binds to the alpha-sarcin loop and helix 89 of Escherichia coli 23S ribosomal RNA. *RNA 7*:1173-1179.

Lake JA. 1976. Ribosome structure determined by electron microscopy of Escherichia coli small subunits, large subunits and monomeric ribosomes. *J Mol Biol 105*:131-139.

Lancaster L, Lambert NJ, Maklan EJ, Horan LH, Noller HF. 2008. The sarcin-ricin loop of 23S rRNA is essential for assembly of the functional core of the 50S ribosomal subunit. *RNA 14*:1999-2012.

Lata KR, Agrawal RK, Penczek P, Grassucci R, Zhu J, Frank J. 1996. Three-dimensional reconstruction of the Escherichia coli 30 S ribosomal subunit in ice. *J Mol Biol 262*:43-52.

Leipe DD, Wolf YI, Koonin EV, Aravind L. 2002. Classification and evolution of P-loop GTPases and related ATPases. *J Mol Biol 317*:41-72.

Leong V, Kent M, Jomaa A, Ortega J. 2013. Escherichia coli rimM and yjeQ null strains accumulate 30S subunits of similar structure and protein complement. *RNA, In press*.

Levdikov VM, Blagova EV, Brannigan JA, Cladiere L, Antson AA, Isupov MN, Seror SJ, Wilkinson AJ. 2004. The crystal structure of YloQ, a circularly permuted GTPase essential for Bacillus subtilis viability. *J Mol Biol 340*:767-782.

Lhoest J, Colson C. 1977. Genetics of ribosomal protein methylation in Escherichia coli. II. A mutant lacking a new type of methylated amino acid, N5-methylglutamine, in protein L3. *Mol Gen Genet 154*:175-180.

Lhoest J, Colson C. 1981. Cold-sensitive ribosome assembly in an Escherichia coli mutant lacking a single methyl group in ribosomal protein L3. *Eur J Biochem 121*:33-37.

Li Z, Pandit S, Deutscher MP. 1999a. Maturation of 23S ribosomal RNA requires the exoribonuclease RNase T. *RNA 5*:139-146.

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

Lin B, Covalle KL, Maddock JR. 1999. The Caulobacter crescentus CgtA protein displays unusual guanine nucleotide binding and exchange properties. *J Bacteriol 181*:5825-5832.

Lindahl L. 1975. Intermediates and time kinetics of the in vivo assembly of Escherichia coli ribosomes. *J Mol Biol 92*:15-37.

Linder P, Lasko PF, Ashburner M, Leroy P, Nielsen PJ, Nishi K, Schnier J, Slonimski PP. 1989. Birth of the D-E-A-D box. *Nature 337*:121-122.

Loh PC, Morimoto T, Matsuo Y, Oshima T, Ogasawara N. 2007. The GTP-binding protein YqeH participates in biogenesis of the 30S ribosome subunit in Bacillus subtilis. *Genes Genet Syst 82*:281-289.

Lovgren JM, Bylund GO, Srivastava MK, Lundberg LA, Persson OP, Wingsle G, Wikstrom PM. 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.

Ludtke SJ, Baldwin PR, Chiu W. 1999. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol 128*:82-97.

Lund E, Dahlberg JE, Lindahl L, Jaskunas SR, Dennis PP, Nomura M. 1976. Transfer RNA genes between 16S and 23S rRNA genes in rRNA transcription units of E. coli. *Cell 7*:165-177.

Maeder C, Draper DE. 2005. A small protein unique to bacteria organizes rRNA tertiary structure over an extensive region of the 50 S ribosomal subunit. *J Mol Biol 354*:436-446.

Maguire BA, Beniaminov AD, Ramu H, Mankin AS, Zimmermann RA. 2005. A protein component at the heart of an RNA machine: the importance of protein l27 for the function of the bacterial ribosome. *Mol Cell 20*:427-435.

Maivali U, Remme J. 2004. Definition of bases in 23S rRNA essential for ribosomal subunit association. *RNA 10*:600-604.

Maki JA, Schnobrich DJ, Culver GM. 2002. The DnaK chaperone system facilitates 30S ribosomal subunit assembly. *Mol Cell 10*:129-138.

Mangiarotti G, Turco E, Ponzetto A, Altruda F. 1974. Precursor 16S RNA in active 30S ribosomes. *Nature 247*:147-148.

Matsuo Y, Morimoto T, Kuwano M, Loh PC, Oshima T, Ogasawara N. 2006. The GTP-binding protein YlqF participates in the late step of 50 S ribosomal subunit assembly in Bacillus subtilis. *J Biol Chem 281*:8110-8117.

Matsuo Y, Oshima T, Loh PC, Morimoto T, Ogasawara N. 2007. Isolation and characterization of a dominant negative mutant of Bacillus subtilis GTP-binding protein, YlqF, essential for biogenesis and maintenance of the 50 S ribosomal subunit. *J Biol Chem 282*:25270-25277.

Mindell JA, Grigorieff N. 2003. Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol 142*:334-347.

Misra TK, Apirion D. 1979. RNase E, an RNA processing enzyme from Escherichia coli. *J Biol Chem 254*:11154-11159.

Mitra K, Frank J. 2006. Ribosome dynamics: insights from atomic structure modeling into cryo-electron microscopy maps. *Annu Rev Biophys Biomol Struct 35*:299-317.

Mizushima S, Nomura M. 1970. Assembly mapping of 30S ribosomal proteins from E. coli. *Nature 226*:1214.

Moazed D, Robertson JM, Noller HF. 1988. Interaction of elongation factors EF-G and EF-Tu with a conserved loop in 23S RNA. *Nature 334*:362-364.

Moazed D, Samaha RR, Gualerzi C, Noller HF. 1995. Specific protection of 16 S rRNA by translational initiation factors. *J Mol Biol 248*:207-210.

Monro RE. 1967. Catalysis of peptide bond formation by 50 S ribosomal subunits from Escherichia coli. *J Mol Biol 26*:147-151.

Moore VG, Atchison RE, Thomas G, Moran M, Noller HF. 1975. Identification of a ribosomal protein essential for peptidyl transferase activity. *Proc Natl Acad Sci U S A 72*:844-848.

Mulder AM, Yoshioka C, Beck AH, Bunner AE, Milligan RA, Potter CS, Carragher B, Williamson JR. 2010. Visualizing ribosome biogenesis: parallel assembly pathways for the 30S subunit. *Science 330*:673-677.

Nanamiya H, Akanuma G, Natori Y, Murayama R, Kosono S, Kudo T, Kobayashi K, Ogasawara N, Park SM, Ochi K, Kawamura F. 2004. Zinc is a key factor in controlling alternation of two types of L31 protein in the Bacillus subtilis ribosome. *Mol Microbiol 52*:273-283.

Nichols CE, Johnson C, Lamb HK, Lockyer M, Charles IG, Hawkins AR, Stammers DK. 2007. Structure of the ribosomal interacting GTPase YjeQ from the enterobacterial species Salmonella typhimurium. *Acta Crystallogr Sect F Struct Biol Cryst Commun 63*:922-928.

Nierhaus KH. 1991. The assembly of prokaryotic ribosomes. *Biochimie 73*:739-755.

Nierhaus KH, Dohme F. 1974. Total reconstitution of functionally active 50S ribosomal subunits from Escherichia coli. *Proc Natl Acad Sci U S A 71*:4713-4717.

Nierhaus KH, Montejo V. 1973. A protein involved in the peptidyltransferase activity of Escherichia coli ribosomes. *Proc Natl Acad Sci U S A 70*:1931-1935.

Nishimura M, Yoshida T, Shirouzu M, Terada T, Kuramitsu S, Yokoyama S, Ohkubo T, Kobayashi Y. 2004. Solution structure of ribosomal protein L16 from Thermus thermophilus HB8. *J Mol Biol 344*:1369-1383.

Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. 2000. The structural basis of ribosome activity in peptide bond synthesis. *Science 289*:920-930.

Noller HF. 1991. Ribosomal RNA and translation. *Annu Rev Biochem 60*:191-227.

Noller HF, Hoffarth V, Zimniak L. 1992. Unusual resistance of peptidyl transferase to protein extraction procedures. *Science 256*:1416-1419.

Noller HF, Kop J, Wheaton V, Brosius J, Gutell RR, Kopylov AM, Dohme F, Herr W, Stahl DA, Gupta R, Waese CR. 1981. Secondary structure model for 23S ribosomal RNA. *Nucleic Acids Res 9*:6167-6189.

Noller HF, Nomura M. 1987. Ribosomes. In: Neidhardt FC, ed. *Escherichia coli and Salmonella typhimurium*. Washington, D.C.: American Society for Microbiology. pp 104-125.

Noller HF, Woese CR. 1981. Secondary structure of 16S ribosomal RNA. *Science 212*:403-411.

Ofengand J, Del Campo M, Kaya Y. 2001a. Mapping pseudouridines in RNA molecules. *Methods 25*:365-373.

Ofengand J, Malhotra A, Remme J, Gutgsell NS, Del Campo M, Jean-Charles S, Peil L, Kaya Y. 2001b. Pseudouridines and pseudouridine synthases of the ribosome. *Cold Spring Harb Symp Quant Biol 66*:147-159.

Ogle JM, Brodersen DE, Clemons WM, Jr., Tarry MJ, Carter AP, Ramakrishnan V. 2001. Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science 292*:897-902.

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

Petry S, Brodersen DE, Murphy FVt, Dunham CM, Selmer M, Tarry MJ, Kelley AC, Ramakrishnan V. 2005. Crystal structures of the ribosome in complex with release factors RF1 and RF2 bound to a cognate stop codon. *Cell 123*:1255-1266.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem 25*:1605-1612.

Poehlsgaard J, Douthwaite S. 2003. Macrolide antibiotic interaction and resistance on the bacterial ribosome. *Curr Opin Investig Drugs 4*:140-148.

Poot RA, van den Worm SH, Pleij CW, van Duin J. 1998. Base complementarity in helix 2 of the central pseudoknot in 16S rRNA is essential for ribosome functioning. *Nucleic Acids Res 26*:549-553.

Potrykus K, Cashel M. 2008. (p)ppGpp: still magical? *Annu Rev Microbiol 62*:35-51.

Powers T, Daubresse G, Noller HF. 1993. Dynamics of in vitro assembly of 16 S rRNA into 30 S ribosomal subunits. *J Mol Biol 232*:362-374.

Ramakrishnan V. 2002. Ribosome structure and the mechanism of translation. *Cell 108*:557-572.

Rheinberger HJ, Sternbach H, Nierhaus KH. 1981. Three tRNA binding sites on Escherichia coli ribosomes. *Proc Natl Acad Sci U S A 78*:5310-5314.

Rodnina MV, Wintermeyer W. 2001a. Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms. *Annu Rev Biochem 70*:415-435.

Rodnina MV, Wintermeyer W. 2001b. Ribosome fidelity: tRNA discrimination, proofreading and induced fit. *Trends Biochem Sci 26*:124-130.

Rohl R, Nierhaus KH. 1982. Assembly map of the large subunit (50S) of Escherichia coli ribosomes. *Proc Natl Acad Sci U S A 79*:729-733.

Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ. 2004. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics 3*:1154-1169.

Roy MK, Apirion D. 1983. Purification and properties of ribonuclease E, an RNA-processing enzyme from Escherichia coli. *Biochim Biophys Acta 747*:200-208.

Roy-Chaudhuri B, Kirthi N, Culver GM. 2010. Appropriate maturation and folding of 16S rRNA during 30S subunit biogenesis are critical for translational fidelity. *Proc Natl Acad Sci U S A 107*:4567-4572.

Ruzheinikov SN, Das SK, Sedelnikova SE, Baker PJ, Artymiuk PJ, Garcia-Lara J, Foster SJ, Rice DW. 2004. Analysis of the open and closed conformations of the GTP-binding protein YsxC from Bacillus subtilis. *J Mol Biol 339*:265-278.

Sato A, Kobayashi G, Hayashi H, Yoshida H, Wada A, Maeda M, Hiraga S, Takeyasu K, Wada C. 2005. The GTP binding protein Obg homolog ObgE is involved in ribosome maturation. *Genes Cells 10*:393-408.

Schaefer L, Uicker WC, Wicker-Planquart C, Foucher AE, Jault JM, Britton RA. 2006. Multiple GTPases participate in the assembly of the large ribosomal subunit in Bacillus subtilis. *J Bacteriol 188*:8252-8258.

Scheres SH, Gao H, Valle M, Herman GT, Eggermont PP, Frank J, Carazo JM. 2007. Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat Methods 4*:27-29.

Scheres SH, Marabini R, Lanzavecchia S, Cantele F, Rutten T, Fuller SD, Carazo JM, Burnett RM, San Martin C. 2005a. Classification of single-projection reconstructions for cryo-electron microscopy data of icosahedral viruses. *J Struct Biol 151*:79-91.

Scheres SH, Nunez-Ramirez R, Sorzano CO, Carazo JM, Marabini R. 2008. Image processing for electron microscopy single-particle analysis using XMIPP. *Nat Protoc 3*:977-990.

Scheres SH, Valle M, Carazo JM. 2005b. Fast maximum-likelihood refinement of electron microscopy images. *Bioinformatics 21 Suppl 2*:ii243-244.

Scheres SH, Valle M, Nunez R, Sorzano CO, Marabini R, Herman GT, Carazo JM. 2005c. Maximum-likelihood multi-reference refinement for electron microscopy images. *J Mol Biol 348*:139-149.

Schluenzen F, Tocilj A, Zarivach R, Harms J, Gluehmann M, Janell D, Bashan A, Bartels H, Agmon I, Franceschi F, Yonath A. 2000. Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell 102*:615-623.

Schmalisch M, Langbein I, Stulke J. 2002. The general stress protein Ctc of Bacillus subtilis is a ribosomal protein. *J Mol Microbiol Biotechnol 4*:495-501.

Schmeing TM, Huang KS, Kitchen DE, Strobel SA, Steitz TA. 2005a. Structural insights into the roles of water and the 2' hydroxyl of the P site tRNA in the peptidyl transferase reaction. *Mol Cell 20*:437-448.

Schmeing TM, Huang KS, Strobel SA, Steitz TA. 2005b. An induced-fit mechanism to promote peptide bond formation and exclude hydrolysis of peptidyl-tRNA. *Nature 438*:520-524.

Schuwirth BS, Borovinskaya MA, Hau CW, Zhang W, Vila-Sanjurjo A, Holton JM, Cate JH. 2005. Structures of the bacterial ribosome at 3.5 A resolution. *Science 310*:827-834.

Selmer M, Dunham CM, Murphy FVt, Weixlbaumer A, Petry S, Kelley AC, Weir JR, Ramakrishnan V. 2006. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science 313*:1935-1942.

Sengupta J, Agrawal RK, Frank J. 2001. Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. *Proc Natl Acad Sci U S A 98*:11991-11996.

Shajani Z, Sykes MT, Williamson JR. 2011. Assembly of bacterial ribosomes. *Annu Rev Biochem 80*:501-526.

Sharma MR, Barat C, Wilson DN, Booth TM, Kawazoe M, Hori-Takemoto C, Shirouzu M, Yokoyama S, Fucini P, Agrawal RK. 2005. Interaction of Era with the 30S ribosomal subunit implications for 30S subunit assembly. *Mol Cell 18*:319-329.

Sharpe Elles LM, Sykes MT, Williamson JR, Uhlenbeck OC. 2009. A dominant negative mutant of the E. coli RNA helicase DbpA blocks assembly of the 50S ribosomal subunit. *Nucleic Acids Res 37*:6503-6514.

Shin DH, Lou Y, Jancarik J, Yokota H, Kim R, Kim SH. 2004. Crystal structure of YjeQ from Thermotoga maritima contains a circularly permuted GTPase domain. *Proc Natl Acad Sci U S A 101*:13198-13203.

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

Siibak T, Peil L, Xiong L, Mankin A, Remme J, Tenson T. 2009. Erythromycin- and chloramphenicol-induced ribosomal assembly defects are secondary effects of protein synthesis inhibition. *Antimicrob Agents Chemother 53*:563-571.

Simonetti A, Marzi S, Myasnikov AG, Fabbretti A, Yusupov M, Gualerzi CO, Klaholz BP. 2008. Structure of the 30S translation initiation complex. *Nature 455*:416-420.

Smith JE, Cooperman BS, Mitchell P. 1992. Methylation sites in Escherichia coli ribosomal RNA: localization and identification of four new sites of methylation in 23S rRNA. *Biochemistry 31*:10825-10834.

Sparling PF. 1970. Kasugamycin resistance: 30S ribosomal mutation with an unusual location on the Escherichia coli chromosome. *Science 167*:56-58.

Spedding G. 1990. *Isolation and analysisi of ribosomes from prokaryotes, eukaryotes and organelles*. New York: IRL Press.

Sperling E, Bunner AE, Sykes MT, Williamson JR. 2008. Quantitative analysis of isotope distributions in proteomic mass spectrometry using least-squares Fourier transform convolution. *Anal Chem 80*:4906-4917.

Srivastava AK, Schlessinger D. 1990. Mechanism and regulation of bacterial ribosomal RNA processing. *Annu Rev Microbiol 44*:105-129.

Stark H, Mueller F, Orlova EV, Schatz M, Dube P, Erdemir T, Zemlin F, Brimacombe R, van Heel M. 1995. The 70S Escherichia coli ribosome at 23 A resolution: fitting the ribosomal RNA. *Structure 3*:815-821.

Steitz TA. 2008. A structural understanding of the dynamic ribosome machine. *Nat Rev Mol Cell Biol 9*:242-253.

Subramanian AR, van Duin J. 1977. Exchange of individual ribosomal proteins between ribosomes as studied by heavy isotope-transfer experiments. *Mol Gen Genet 158*:1-9.

Sykes MT, Williamson JR. 2009. A Complex Assembly Landscape for the 30S Ribosomal Subunit. *Annu Rev Biophys 38*:197-215.

Talkington MW, Siuzdak G, Williamson JR. 2005. An assembly landscape for the 30S ribosomal subunit. *Nature 438*:628-632.

Tan J, Jakob U, Bardwell JC. 2002. Overexpression of two different GTPases rescues a null mutation in a heat-induced rRNA methyltransferase. *J Bacteriol 184*:2692-2698.

Tan RK, Devkota B, Harvey SC. 2008. YUP.SCX: coaxing atomic models into medium resolution electron density maps. *J Struct Biol 163*:163-174.

Tan RK-Z, Petrov AS, Harvey SC. 2006. YUP: A molecular simulation program for coarse-grained and multi-scale models. . *J Chem Theory Comput 2*:529-540.

Tanner NK, Cordin O, Banroques J, Doere M, Linder P. 2003. The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol Cell 11*:127-138.

Teraoka H, Nierhaus KH. 1978. Protein L16 induces a conformational change when incorporated into a L16-deficient core derived from Escherichia coli ribosomes. *FEBS Lett 88*:223-226.

Traub P, Hosokawa K, Craven GR, Nomura M. 1967. Structure and function of E. coli ribosomes, IV. Isolation and characterization of functionally active ribosomal proteins. *Proc Natl Acad Sci U S A 58*:2430-2436.

Traub P, Nomura M. 1968a. Structure and function of E. coli ribosomes. V. Reconstitution of functionally active 30S ribosomal particles from RNA and proteins. *Proc Natl Acad Sci U S A 59*:777-784.

Traub P, Nomura M. 1968b. Structure and function of Escherichia coli ribosomes. I. Partial fractionation of the functionally active ribosomal proteins and reconstitution of artificial subribosomal particles. *J Mol Biol 34*:575-593.

Traub P, Nomura M. 1969a. Structure and function of Escherichia coli ribosomes. VI. Mechanism of assembly of 30 s ribosomes studied in vitro. *J Mol Biol 40*:391-413.

Traub P, Nomura M. 1969b. Studies on the assembly of ribosomes in vitro. *Cold Spring Harb Symp Quant Biol 34*:63-67.

Traut RR, Monro RE. 1964. The Puromycin Reaction and Its Relation to Protein Synthesis. *J Mol Biol 10*:63-72.

Uicker WC, Schaefer L, Britton RA. 2006. The essential GTPase RbgA (YlqF) is required for 50S ribosome assembly in Bacillus subtilis. *Mol Microbiol 59*:528-540.

Valle M, Sengupta J, Swami NK, Grassucci RA, Burkhardt N, Nierhaus KH, Agrawal RK, Frank J. 2002. Cryo-EM reveals an active role for aminoacyl-tRNA in the accommodation process. *Embo J 21*:3557-3567.

van Heel M, Harauz G, Orlova EV, Schmidt R, Schatz M. 1996. A new generation of the IMAGIC image processing system. *J Struct Biol 116*:17-24.

Weinger JS, Parnell KM, Dorner S, Green R, Strobel SA. 2004. Substrate-assisted catalysis of peptide bond formation by the ribosome. *Nat Struct Mol Biol 11*:1101-1106.

Whitby FG, Masters EI, Kramer L, Knowlton JR, Yao Y, Wang CC, Hill CP. 2000. Structural basis for the activation of 20S proteasomes by 11S regulators. *Nature 408*:115-120.

Wicker-Planquart C, Foucher AE, Louwagie M, Britton RA, Jault JM. 2008. Interactions of an essential Bacillus subtilis GTPase, YsxC, with ribosomes. *J Bacteriol 190*:681-690.

Williamson JR. 2003. After the ribosome structures: how are the subunits assembled? *RNA 9*:165-167.

Wilson DN, Nierhaus KH. 2006. The E-site story: the importance of maintaining two tRNAs on the ribosome during protein synthesis. *Cell Mol Life Sci 63*:2725-2737.

Wilson DN, Nierhaus KH. 2007. The weird and wonderful world of bacterial ribosome regulation. *Crit Rev Biochem Mol Biol 42*:187-219.

Wimberly BT, Brodersen DE, Clemons WM, Jr., Morgan-Warren RJ, Carter AP, Vonrhein C, Hartsch T, Ramakrishnan V. 2000. Structure of the 30S ribosomal subunit. *Nature 407*:327-339.

Woodson SA. 2008. RNA folding and ribosome assembly. *Curr Opin Chem Biol 12*:667-673.

Wout P, Pu K, Sullivan SM, Reese V, Zhou S, Lin B, Maddock JR. 2004. The Escherichia coli GTPase CgtAE cofractionates with the 50S ribosomal subunit and interacts with SpoT, a ppGpp synthetase/hydrolase. *J Bacteriol 186*:5249-5257.

Wriggers W, Agrawal RK, Drew DL, McCammon A, Frank J. 2000. Domain motions of EF-G bound to the 70S ribosome: insights from a hand-shaking between multi-resolution structures. *Biophys J 79*:1670-1678.

Wriggers W, Milligan RA, McCammon JA. 1999. Situs: A package for docking crystal structures into low-resolution maps from electron microscopy. *J Struct Biol 125*:185-195.

Xia B, Ke H, Shinde U, Inouye M. 2003. The role of RbfA in 16S rRNA processing and cell growth at low temperature in Escherichia coli. *J Mol Biol 332*:575-584.

Xu Z, O'Farrell HC, Rife JP, Culver GM. 2008. A conserved rRNA methyltransferase regulates ribosome biogenesis. *Nat Struct Mol Biol 15*:534-536.

Yonath A, Mussig J, Tesche B, Lorenz S, Erdman VA, Wittman HG. 1980. Crystaliization of the large ribosomal subunit from B. stearothermophilus. *Biochem Int 1*:315-428.

Young RA, Steitz JA. 1978. Complementary sequences 1700 nucleotides apart form a ribonuclease III cleavage site in Escherichia coli ribosomal precursor RNA. *Proc Natl Acad Sci U S A 75*:3593-3597.

Youngman EM, Brunelle JL, Kochaniak AB, Green R. 2004. The active site of the ribosome is composed of two layers of conserved nucleotides with distinct roles in peptide bond formation and peptide release. *Cell 117*:589-599.

Yusupov MM, Yusupova GZ, Baucom A, Lieberman K, Earnest TN, Cate JH, Noller HF. 2001. Crystal structure of the ribosome at 5.5 A resolution. *Science 292*:883-896.

Zhang S, Haldenwang WG. 2004. Guanine nucleotides stabilize the binding of Bacillus subtilis Obg to ribosomes. *Biochem Biophys Res Commun 322*:565-569.

Appendix

Appendix to Chapter 2

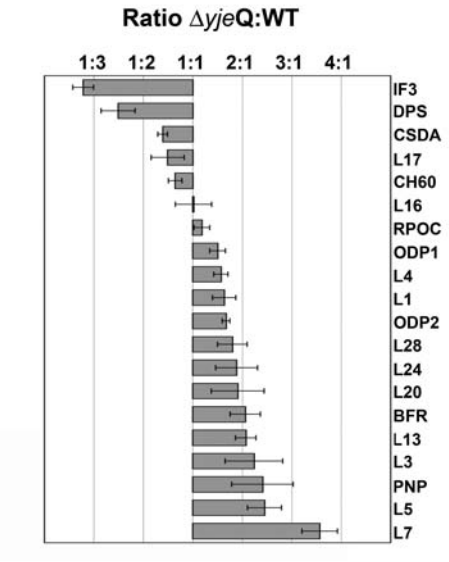


Figure S2.1 iTRAQ analysis of the 30S subunits purified from D*yjeQ* cells. Large subunit proteins and non-ribosomal proteins.
The plot shows the relative levels of the 50S subunit r-proteins (L-proteins) and other non-ribosomal proteins between the 30S subunit purified from DyjeQ and wild type (WT) cells. Relative levels for each protein are expressed as the average ratio DyjeQ:WT obtained from two replicas of the experiment. Only peptides that identify proteins with ≥95% confidence were used for the calculation of these ratios. The error bars indicate the standard error of the mean for each ratio. The abbreviations used to name the non-ribosomal proteins are as follow: Translation initiation factor IF-3 (IF3), DNA protection during starvation protein (DPS), Cold-shock DEAD box protein A (CSDA), 60 kDa chaperonin (CH60), DNA-directed RNA polymerase subunit beta' (RPOC), Pyruvate dehydrogenase E1 component (ODEP1), Pyruvate dehydrogenase E2 component (ODP2), Bacterioferritin (BFR) and Polyribonucleotide nucleotidyltransferase (PNP).

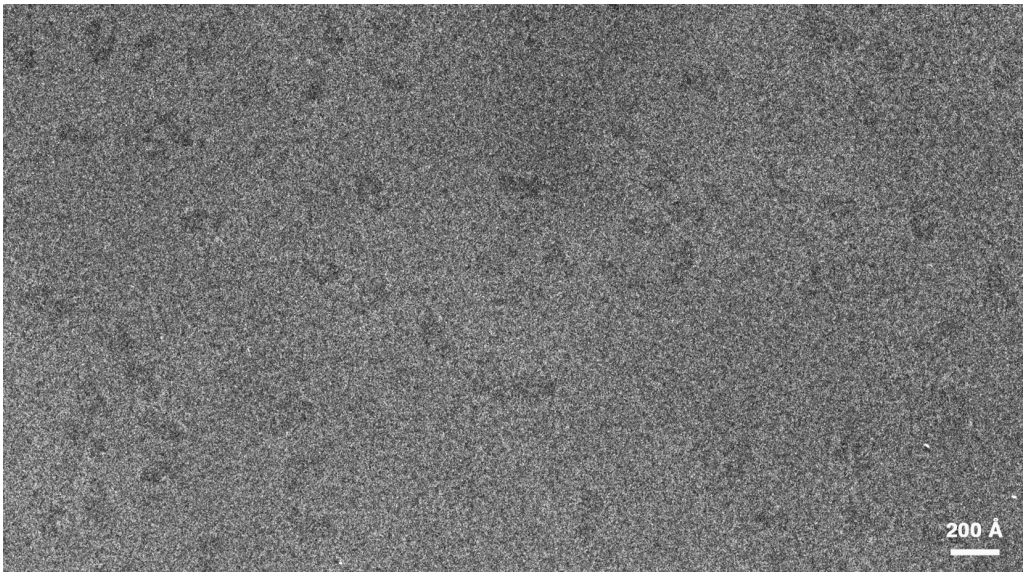
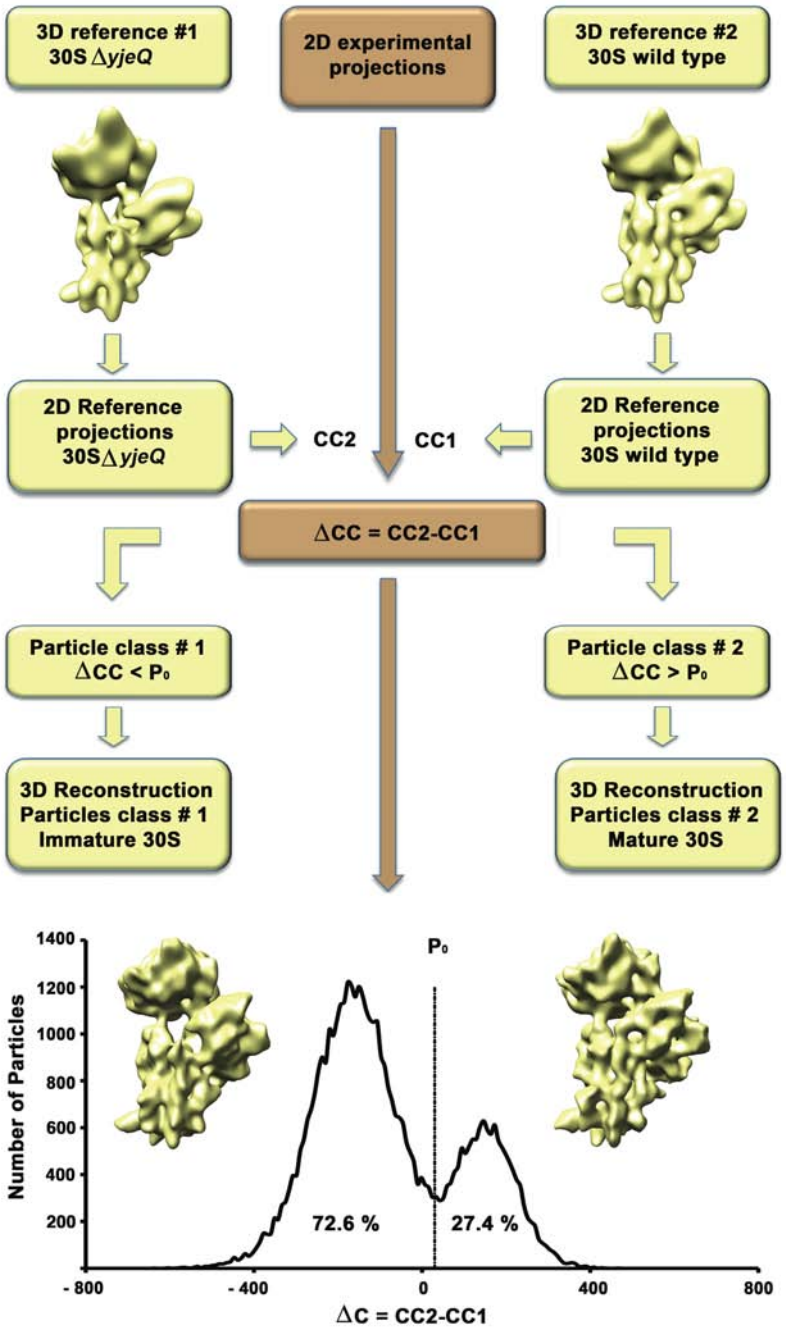


Figure S2.2 Cryo-electron microscopy of immature 30S subunits.
Representative cryo-electron micrograph of 30S subunits purified from the E. coli DyjeQ strain.


3D reference #1
30S ΔyjeQ
2D experimental projections
3D reference #2
30S wild type
2D Reference projections
30SΔyjeQ
CC2
CC1
2D Reference projections
30S wild type
ΔCC = CC2-CC1
Particle class # 1
ΔCC < P0
Particle class # 2
ΔCC > P0
3D Reconstruction
Particles class # 1
Immature 30S
3D Reconstruction
Particles class # 2
Mature 30S
P0
Number of Particles
1400
1200
1000
800
600
400
200
72.6 %
27.4 %
- 800
- 400
0
400
800
ΔC = CC2-CC1

continued

Figure S3.4 Supervised classification of 30S subunits purified from D*yjeQ* cells.
Supervised classification strategy to separate the projections representing mature from immature 30S subunits in the cryo-EM data set collected from the 30S fraction purified from DyjeQ cells. The two references used in this classification were the 3D reconstruction obtained from all the particle images in the data set (3D reference #1 30S DyjeQ) and the structure of the mature 30S subunit obtained by low-pass filtering the X-ray structure of the E. coli 30S subunit (PDB ID: 2Z4K) (Borovinskaya et al., 2007) to 25Å resolution (3D reference #2 30S wild type). Interestingly, reference #1 readily showed a distortion in the upper segment of helix 44 (upper panel), implying that the immature 30S subunits are structurally different from the mature 30S subunits and that most of the projections in our cryo-EM data set represented immature particles. All images in the data set were aligned to the projections generated from each of the references and ach image was assigned to the projection view from each reference that yielded the highest cross-correlation coefficient and consequently obtaining two cross-correlation coefficients per particle image (CC1 and CC2). The similarity of the particle images to the two references was derived by plotting the distribution of particles against the difference between the two cross-correlation coefficients (DCC=CC2-CC1). The particles classified following a bimodal distribution with a large (72.6%) and small (27.4%) peaks representing the population of immature and mature 30S subunits, respectively. The dash line indicates the P_0 value used to divide the data into two subsets. The 3D maps of the mature and the immature 30S subunit were calculated from both homogeneous subpopulations of images using a standard projection matching approach as implemented in the Xmipp program (Scheres et al., 2008). The reference map used to refine both maps was the X-ray structure of the E. coli 30S subunit (PDB ID: 2Z4K) (Borovinskaya et al., 2007) low-pass filtered to 25Å resolution.

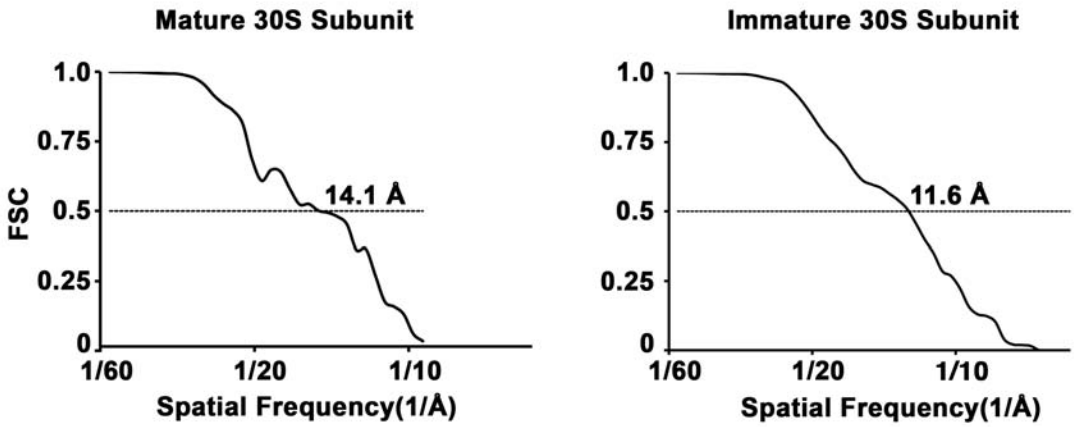
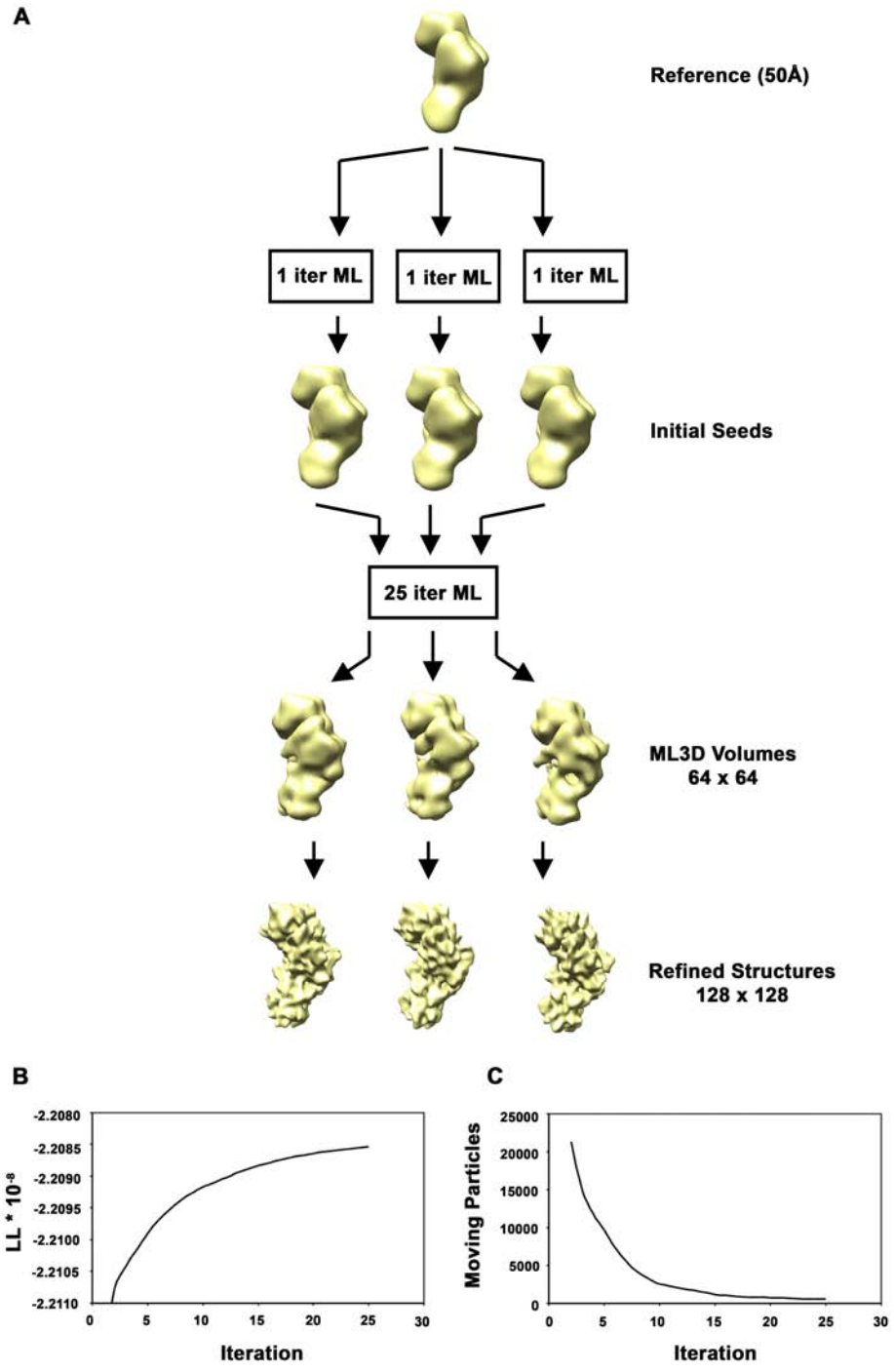


Figure S2.4 Estimation of the resolution of the 3D reconstructions of the mature and immature 30S subunits.

Fourier Shell Correlation (FSC) plots corresponding to the EM maps of the mature (left panel) and immature 30S subunit (right panel). The estimated resolution for both maps using a FSC value of 0.5 is indicated. These plots relate to the 3D reconstructions of the immature and mature 30S subunit structures shown in Fig. 3.


A
Reference (50Å)
1 iter ML
1 iter ML
1 iter ML
Initial Seeds
25 iter ML
ML3D Volumes
64 x 64
Refined Structures
128 x 128
B
LL * 10-8
-2.2080
-2.2085
-2.2090
-2.2095
-2.2100
-2.2105
-2.2110
0
5
10
15
20
25
30
Iteration
C
Moving Particles
25000
20000
15000
10000
5000
0
0
5
10
15
20
25
30
Iteration

continued

Figure S2.5 3D maximum likelihood-based classification of immature 30S subunits. (*A) Diagram representing the strategy followed for the maximum likelihood-based classification (Scheres et al., 2005b; Scheres et al., 2005c; Scheres et al., 2007). Projections under the large peak resulting from the supervised classification approach representing immature 30S subunits purified from DyjeQ cells (Supplemental Fig. S2) were classified by this method. The volume at the top of the diagram is the low-resolution structure of the mature 30S subunit obtained from the X-ray structure of the E. coli 30S subunit (PDB ID: 2Z4K) that was used as the initial reference to start the classification process. The three maps below constitute the seeds that were generated from random subsets of the experimental images after performing one iteration of maximum likelihood optimization. The seeds served as initial references for 25 iterations of maximum likelihood-based classification using downscaled images of 64 x 64 pixels that resulted in three low-resolution reconstructions shown below the seeds. The volumes at the very bottom are the three refined structures of the immature 30S subunit found in the data set and were constructed from the original 128x128 projections in each class using a standard projection matching approach (Scheres et al., 2008). The reference map used to refine the three subsets of particles was the X-ray structure of the E. coli 30S subunit (PDB ID: 2Z4K) (Borovinskaya et al., 2007) low-pass filtered to 25Å resolution. The resulting map with the most, intermediate and less severe distortion in the helix 44 represented the 35.6 %, 33.1% and 31.3% of the particle population, respectively. (B) The plot shows the values of the log-likelihood target function versus iterations. (C) Graph showing the number of particles for which the class membership changes over subsequent iterations.*

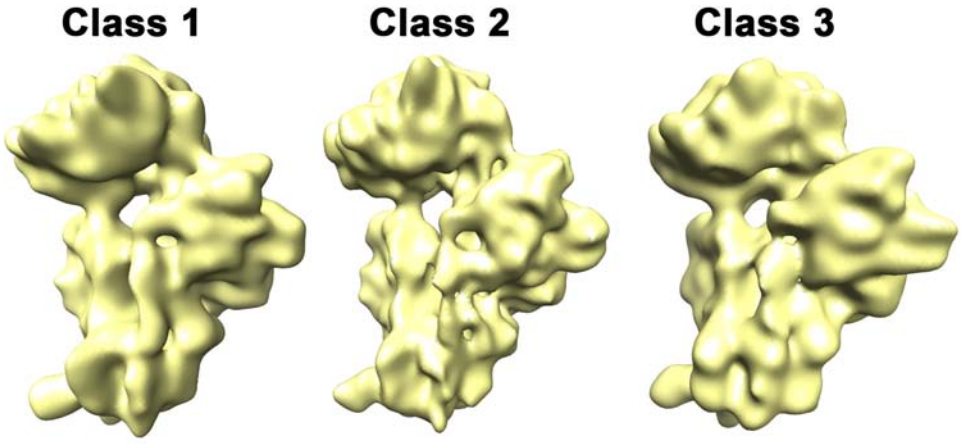


Figure S2.6 Maximum likelihood-based classification of the mature 30S subunits from wild type E. coli cells.

Three cryo-EM structures (labeled from class 1 to class 3) obtained from classifying projections obtained from the 30S subunit fraction purified from wild type E. coli cells. Particles were classified following an identical maximum likelihood-based classification approach to the one used with the immature 30S subunits from ΔyjeQ cells (Supplemental Fig. S3). The three reconstructions were all very similar and resembled the mature 30S subunit structure. The reconstructions were produced from 8, 487 (class 1), 9,208 (class2) and 8,093 (class 3) projections.

							First replicate				Second replicate			
	Accession	Protein Name	Molecular Weight (kDa)	Sequence Coverage (%)[a]	Unique peptides[b]	Total peptides[c]	Ratio ΔyjeQ:WT	Pvalue ΔyjeQ:WT	Ratio WT:ΔyjeQ	Pvalue WT:ΔyjeQ	Ratio ΔyjeQ:WT	Pvalue ΔyjeQ:WT	Ratio WT:ΔyjeQ	Pvalue WT:ΔyjeQ
Small subunit proteins														
	P0AG67	S1	60.1	55.8	125	30			4.41	<0.01			2.61	<0.01
	P0A7V0	S2	26.6	51.0	72	11			2.12	0.05			2.55	<0.01
	P0A7V3	S3	25.8	59.7	61	15			1.29	0.62			2.13	<0.01
	P0A7V8	S4	23.3	51.9	100	21	1.12	0.82					1.61	<0.01
	P0A7W1	S5	17.5	74.8	52	16			1.26	0.66			1.96	<0.01
	P02358	S6	15	79.3	16	81			1.33	0.75			1.31	0.45
	P02359	S7	19.9	54.8	14	99			1.62	0.42			1.55	<0.01
	P0A7W7	S8	14	63.8	41	10			1.19	0.77			1.55	<0.01
	P0A7X3	S9	14.7	55.4	56	11	1.08	0.86					1.53	<0.01
	P0A7R5	S10	11.6	68.0	34	6			1.21	0.45			1.63	<0.01
	P0A7R9	S11	13.7	61.2	30	8			3.77	0.12			1.22	0.18
	P0A7S3	S12	13.6	37.9	13	3	1.08	0.51					1.86	0.22
	P0A7S9	S13	13	67.0	40	9			1.26	0.64			1.42	0.06
	P0AG59	S14	11.4	34.6	22	4			1.16	0.80			1.75	0.01
	P0ADZ4	S15	10.1	56.2	31	7	1.31	0.74					2.00	0.05
	P0A7T3	S16	9	52.4	29	6			1.32	0.50			1.41	0.11
	P0AG63	S17	9.6	39.3	30	4			1.11	0.71			1.54	0.02
	P0A7T7	S18	8.8	40.0	26	4	1.27	0.68					1.66	0.11
	P0A7U3	S19	10.3	68.5	18	5			1.07	0.90			1.52	<0.01
	P0A7U7	S20	9.5	35.6	33	4			1.23	0.73			1.44	0.02
	P68679	S21	8.4	39.4	9	3			5.06	0.01			4.48	0.01
Large subunit proteins														
	P0A7L0	L1	24.6	23.5	4	10	1.52	0.46			2.43	0.02		
	P60438	L3	22.1	10.0	5	4	2.22	0.37			1.59	0.03		
	P60723	L4	21.9	34.3	10	5	1.79	0.27			1.67	0.06		
	P62399	L5	10.1	27.4	10	6	2.32	0.04			1.96	0.01		
	P0A7K2	L7/L12	12.1	54.5	4	8	1.35	0.71			6.88	0.03		
	P0AA10	L13	15.9	12.0	2	2			2.12	0.37	1.29	0.76		
	P0ADY7	L16	15.1	22.1	2	2	1.89				1.02			
	P0AG44	L17	14.2	19.7	3	2	1.70	0.60			1.48	0.33		
	P0A7L3	L20	13.4	16.1	5	2	1.81	0.66			1.55	0.08		
	P60624	L24	11.2	26.9	5	2	2.04	0.64			2.15	0.31		
	P0A7M2	L28	8.9	24.4	5	2	1.78	<0.01			1.57	0.32		
Other proteins														
	P0ABD3	Bfr	18.4	39.2	11	5	1.36	0.48			3.04	<0.01		
	P0A6F5	Ch60	57.3	23.5	7	17			1.58	0.28	1.13	0.27		
	P0A9P6	CsdA	70.5	24.9	14	11			1.91	0.03			1.34	0.04
	P0ABT2	Dps	18.7	16.8	5	2			1.23	0.71	1.21	0.82		
	P0A707	IF-3	20.6	27.8	3	6			7.92	0.09			1.29	0.36
	P0AFG8	Odp1	99.7	11.5	11	7			1.19	0.42	2.37	<0.01		
	P06959	Odp2	66.1	7.9	2	2			1.50	0.61	4.24	<0.01		
	P05055	Pnp	77.1	6.6	3	3			2.18	0.52	5.67	0.14		
	P0A8T7	RpoC	155.2	12.8	13	10			1.45	0.35	1.25	0.11		

Supplemental Table 2.1: Relative iTRAQ Ratios for Detected Proteins

[a] The number of amino acids spanned by the assigned peptides divided by the sequence length.[b] The total number of non-overlapping peptides with at least 95% confidence.[c] Total peptides with at least 95% confidence.

Appendix to Chapter 3

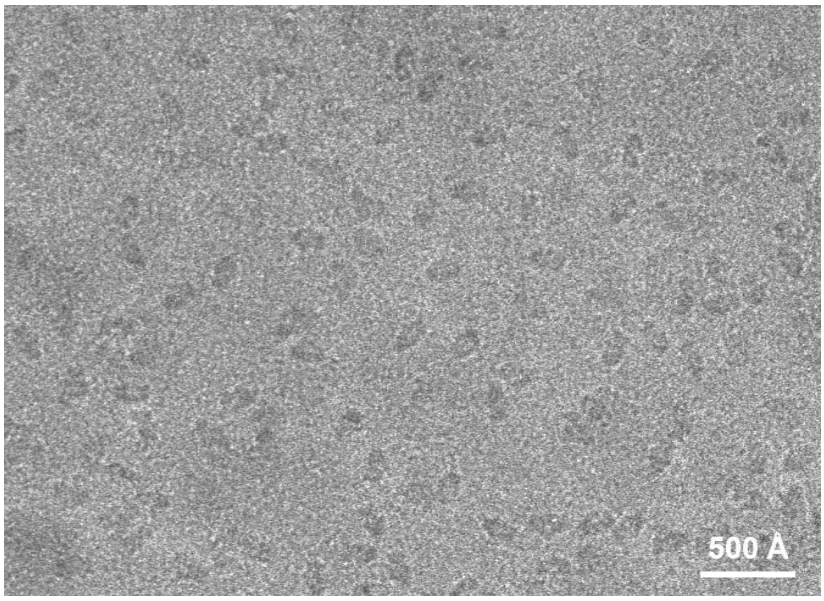


Figure S3.1 Cryo-electron microscopy of 30S-YjeQ complexes.
Representative cryo-electron micrograph of a preparation of 30S-YjeQ complexes.

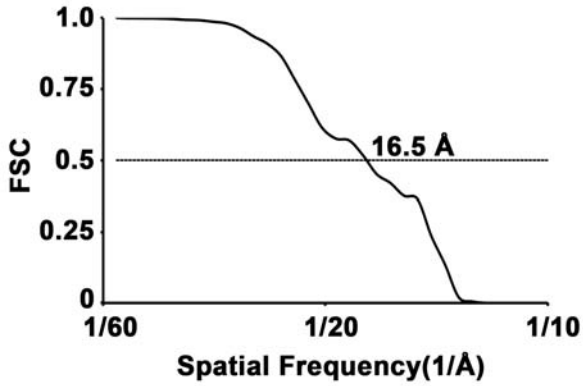


Figure S3.2 Resolution of the cryo-EM structure of the 30S-YjeQ complex.
Fourier Shell Correlation (FSC) plot estimating the resolution of the 3D reconstruction o the 30S subunit in complex with the YjeQ protein. A FSC value of 0.5 was used to provide the resolution value indicated in the plot.

```
E.coli          --------MSKNKLSKGQQRRVNANHQRRLKTSKEKPDYDDNLFGEPDEG 42
S.typhimurium   MGDQEPVRLSKNKLSKGQQRRVNANHQRRLKTSAEKADYDDNLFGEPAEG 50
                         :************************ **.********** **

E.coli          IVISRFGMHADVESADGDVHRCNIRRTIRSLVTGDRVVWRPGKPAAEGVN 92
S.typhimurium   IVISRFGMHADVESADGEVHRCNIRRTIRSLVTGDRVVWRPGKAAAEGVN 100
                ****************:*************************.******

E.coli          VKGIVEAVHERTSVLTRPDFYDGVKPIAANIDQIVIVSAILPELSLNIID 142
S.typhimurium   VKGIVEAVHERTSVLTRPDFYDGVKPIAANIDQIVIVSAILPELSLNIID 150
                **************************************************

E.coli          RYLVACETLQIEPIIVLNKIDLLDDEGMAFVNEQMDIYRNIGYRVLMVSS 192
S.typhimurium   RYLVGCETLQVEPLIVLNKIDLLDDEGMDFVNEQMDIYRNIGYRVLMVSS 200
                ****.*****:**:************** ********************

E.coli          HTQDGLKPLEEALTGRISIFAGQSGVGKSSLLNALLGLQKEILTNDISDN 242
S.typhimurium   HTQDGLKPLEEALTGRISIFAGQSGVGKSSLLNALLGLQNEILTNDVSNV 250
                ****************************************:******:*:

E.coli          SGLGQHTTTAARLYHFPHGGDVIDSPGVREFGLWHLEPEQITQGFVEFHD 292
S.typhimurium   SGLGQHTTTAARLYHFPHGGDVIDSPGVREFGLWHLEPEQITQGFVEFHD 300
                **************************************************

E.coli          YLGLCKYRDCKHDTDPGCAIREAVEEGKIAETRFENYHRILESMAQVKTR 342
S.typhimurium   YLGHCKYRDCKHDADPGCAIREAVENGAIAETRFENYHRILESMAQVKTR 350
                *** *********:***********:* **********************

E.coli          KNFSDTDD 350
S.typhimurium   KNFSDTDD 358   92% identity / 94% similarity
                ********
```

Figure S3.3 Sequence alignment of *Escherichia coli* and *Salmonella typhimurium* YjeQ.

Sequences of E. coli and S. typhimurium YjeQ were aligned using the sequence alignment tool ClustalW2 (www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool= clustalw2). Percent identities and similarities were generated by pairwise alignment using the EMBOSS water pairwise alignment algorithm on the European Bioinformatics Institute Web site (www.ebi.ac.uk/Tools/emboss/align/index.html).

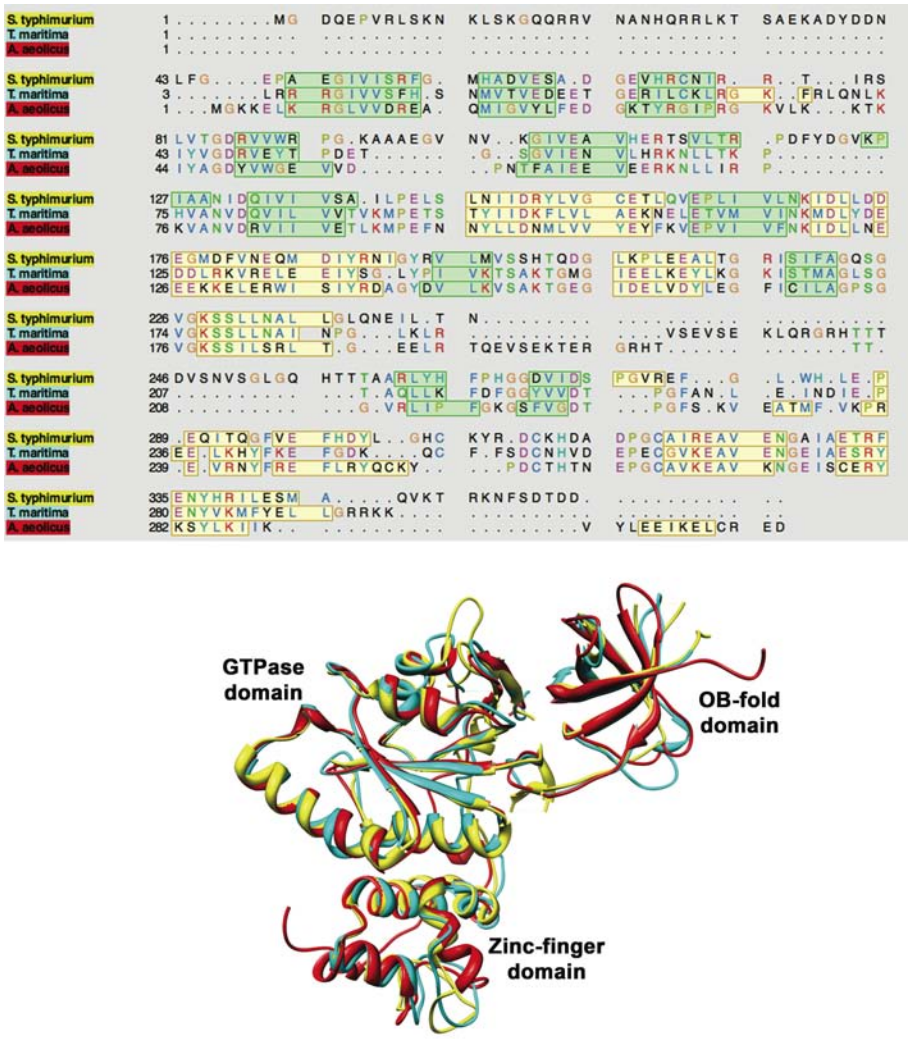


FIGURE S3.4 Structure based alignment of YjeQ from different organisms.

Structure-based sequence alignment (top panel) of YjeQ from Salmonella typhimurium (yellow), Thermotoga maritime (cyan) and Aquifex aeolicus (red). Residues were colored using the Clustal X coloring scheme (Chenna et al., 2003), which depends on both residue type and pattern of conservation. Residues boxed in green and yellow form the β-strands and α-helices in the structure, respectively. Residues outside the box form loop or unstructured regions. In the bottom panel the three structures are shown superimposed. The three domains of YjeQ are labeled.

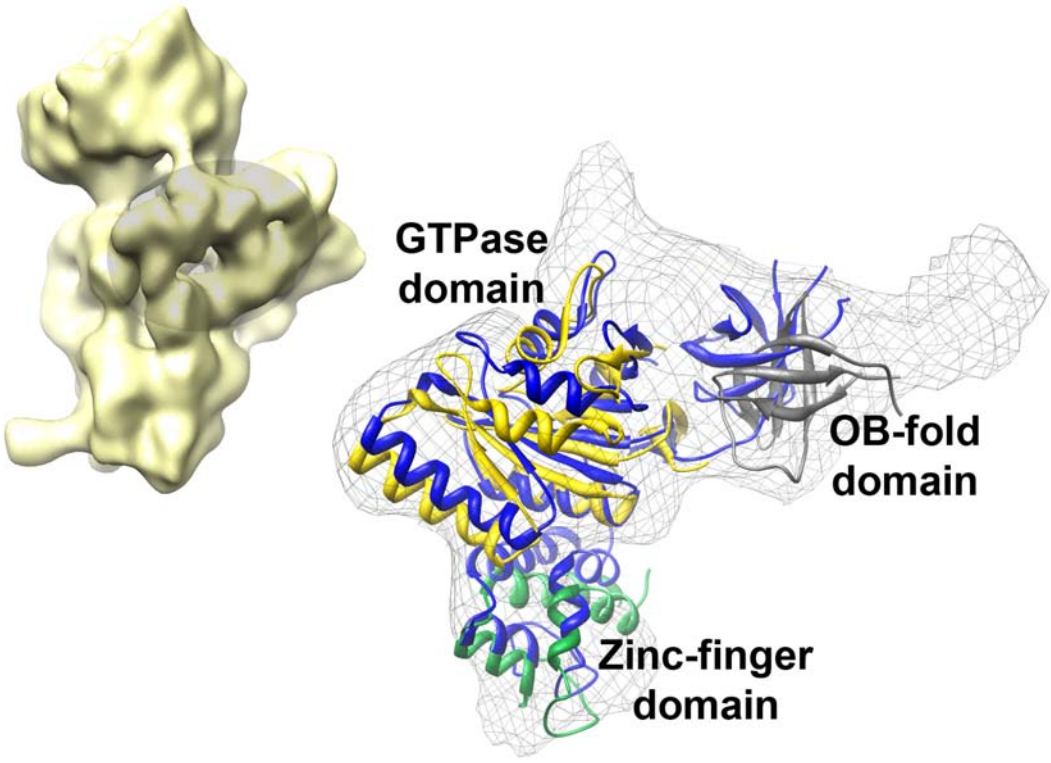


Figure S3.5 Optimizing the fitting of the YjeQ X-ray structure into the cryo-EM map of the 30S-YjeQ complex .

Rigid body fitting of the entire YjeQ X-ray structure into the additional mass of density covering in the cryo-EM map of the 30S-YjeQ complex is shown in blue. Small translational and rotational movements of each domain independently were performed to optimize the fitting. The YjeQ structure with optimized fitting is shown with the OB-fold, GTPase and zinc-finger domains colored in grey, yellow and green, respectively. The cryo-EM map of the 30S-YjeQ complex in a matching view with the additional density is shown as an aid for orientation.

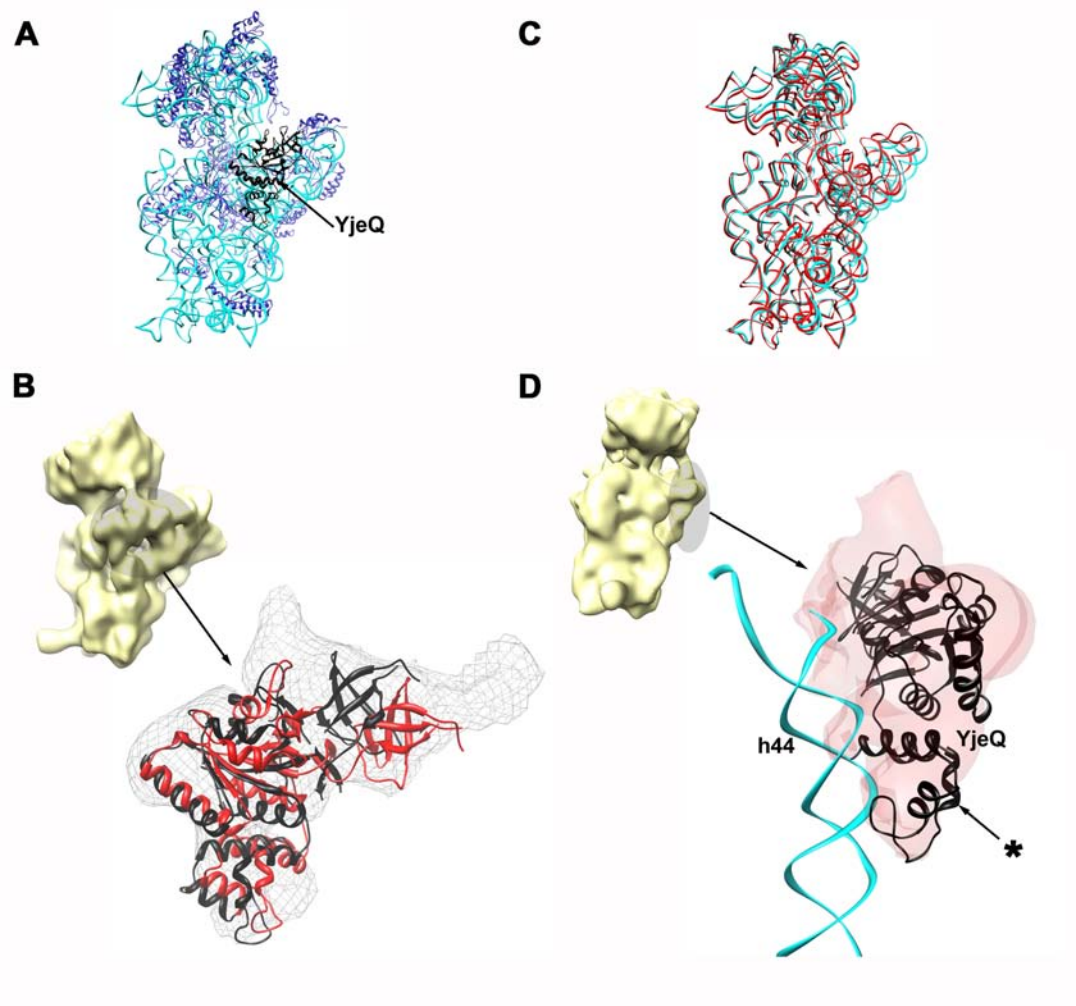


FIGURE S6. Refinement of the 30S-YjeQ by flexible fitting.

(A) Pseudo-atomic model of the 30S-YjeQ complex obtained by flexible fitting. (B) Placement of the X-ray structure of YjeQ in the additional density of the cryo-EM map by manual (red) and flexible (black) fitting. (C) Overlapping of the 16S rRNA from the pseudo-atomic model of the 30S-YjeQ complex obtained by manual (red) and flexible (cyan) fitting. (D) Interactions of the zinc-finger domain of YjeQ with the 16S rRNA helix 44 according to the pseudo-atomic model of the 30S-YjeQ complex obtained by flexible fitting. An asterisk indicates the region of the YjeQ zinc-finger domain protruding outside the cryo-EM density. As an aid for orientation, panels B and D shows the cryo-EM map of the 30S-YjeQ complex in a matching view.

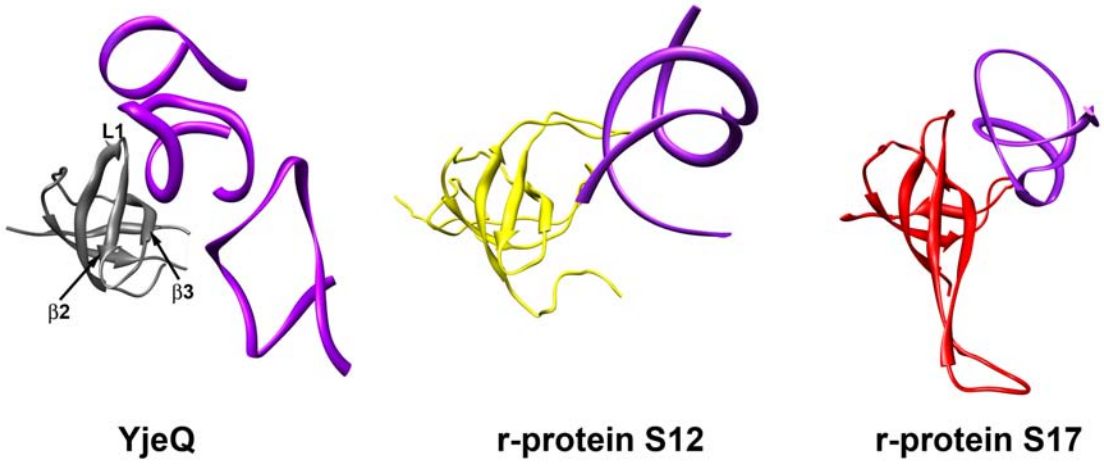


Figure S3.7 Binding of OB-fold domains to RNA.

Comparison of the mode of binding to RNA of the OB-fold domains (left) of YjeQ and r-proteins S12 and S17. The OB-fold domains for the three proteins are shown in equivalent orientations after a structure based alignment was performed. RNA is colored in purple.