STUDYING THE ASSEMBLY OF THE BACTERIAL

LARGE RIBOSOMAL SUBUNIT

ELUCIDATING THE FUNCTION OF ASSEMBLY FACTORS IN THE

MATURATION OF THE BACTERIAL LARGE RIBOSOMAL SUBUNIT

By XIAODAN NI, M.SC.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the

Requirement for the Degree Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (2016)　　　　　　　McMaster University

(Biochemistry and Biomedical Sciences)　　　　Hamilton, Ontario

TITLE:　　　　　　　　　Elucidating the Function of Assembly Factors in the

　　　　　　　　　　　　Maturation of the Bacterial Large Ribosomal Subunit

AUTHOR:　　　　　　　Xiaodan Ni, M.Sc.

SUPERVISOR:　　　　　Dr. Joaquin Ortega

NUMBER OF PAGES:　　xiix, 119

ABSTRACT

Antibiotic resistance in bacteria is becoming a major threat to public health. Many of the antibiotics used today in the clinic target the process of protein synthesis performed by the ribosome. Recent prospects for blocking ribosome function are increasingly focusing on preventing the assembly of bacterial ribosomes. A number of ribosome assembly factors are emerging as attractive targets for novel antibiotics that work in new ways.

YphC and YsxC are essential GTPases in *Bacillus subtilis* that facilitate the assembly of the 50S ribosomal subunit; however, their roles in this process are still uncharacterized. To explore their function, we biochemically and structurally characterized the $45S_{YphC}$ and $44.5S_{YsxC}$ precursor particles accumulated from strains depleted of YphC and YsxC, respectively. Quantitative mass spectrometry analysis and 5-6 Å resolution cryo-EM maps of the $45S_{YphC}$ and $44.5S_{YsxC}$ particles revealed that the two GTPases participate in maturation of functional sites of the 50S subunit. We also observed that YphC and YsxC bind specifically to the two immature particles. In addition, we characterized the structure of the 50S subunits in complex with the RbgA protein. The preliminary 3D structure shows that the RbgA protein binds to the P site of the 50S subunit and displaces h69. There are also missing densities in the structure for h68 and the uL16 ribosomal protein. We expect that the atomic resolution structure of the 50S.RbgA complex will reveal the function and molecular mechanisms of this assembly factor.

The deep structural understanding of protein synthesis process done by the ribosome led to the optimization of over a hundred antibiotics that are currently used in the

clinic. In the same manner, work described in this thesis provides novel insights into understanding the maturation of the large ribosomal subunit, and is paving the way to use the bacterial ribosome biogenesis pathway as a target for the development of new antimicrobials.

Acknowledgements

I would like to acknowledge my supervisor, Dr. Joaquin Ortega, for all his support, guidance and patience throughout all four years. It has been a long journey for me from the very difficult beginning to now defending my PhD Thesis. I enjoy very much the discussions and conversations with him inside and outside of the lab environment. It was extremely rewarding the challenge we overcame together to set up the image-processing environment in the lab. I will really miss those days and nights we worked together on the microscope, taking pictures to solve structures. He gave me enough freedom to be myself, inspired me to overcome the challenges that I encountered in my projects, encouraged me to find what I am really interested in with my research, and made me become an independent and more mature researcher.

I would like to thank my committee members, Dr. Yingfu Li and Dr. Russell Bishop for their guidance and valuable advice. I would like to thank Dr. Li for sharing his own experience and all those valuable encouragements that really guided my way during the qualification exam; I would like to thank Dr. Bishop for his kindness in my committee meetings and encouraging me to learn more and improve my critical thinking.

I would like to thank our close collaborator, Dr. Robert Britton. His open-mind and wisdom were always helpful when we discussed projects and he also gave us valuable suggestions. I would like to thank Dr. Nikhil Jain in Dr. Britton's lab for help to design experiments in a more efficient way. I would like to thank Dr. John Rubinstein for giving us access to his microscope at The Hospital for Sick Children in Toronto;

Table of Contents

List of Figures

Chapter 3.

Chapter 4.

Chapter 5.

List of Tables

Chapter 3.

List of Abbreviations

2D/3D	two/three dimensional
30S	small bacterial ribosomal subunit
50S	large bacterial ribosomal subunit
Å	Angstroms
ANTAR	AmiR-NasR transcription anti-termination regulator
A-site	Aminoacyl tRNA site
B	bound
CBP	calmodulin binding peptide
CCD	charge-coupled device
CP	central protuberance
Cryo-EM	cryo-electron microscopy
CTF	contrast transfer function
DC	decoding center
DDD	direct detection device
DQE	detective quantum efficiency
E-site	Exit-site
FSC	Fourier Shell Correlation
FT	flow through
GAR	GTPase associated region
IgG	immunoglobulin G
IPTG	isopropylbeta-D-thiogalactopyranoside
MDFF	molecular dynamics flexible fitting

MSLR	Mcmaster Service Lab and Repository
MRM	multiple reaction monitoring
P	pellet
PTC	peptidyl transferase center
Pr A	protein A
P-site	Peptidyl tRNA site
qMS	quantitative mass spectrometry
r-proteins	ribosomal proteins
rRNA	ribosomal RNA
RbgA	ribosome biogenesis GTPase A
S	supernatant
SDS	Shwachman Bodian Syndrome
tRNA	transfer RNA
TAP	tandem affinity purification

Declaration of Academic Achievement

All of the experiments in this thesis were designed and conducted by Dr. Joaquin Ortega and myself, unless stated otherwise in the chapter preface. Mass spectrometry analysis in Chapter 2 was performed by Dr. Joseph Davis.

CHAPTER 1. General Introduction

The bacterial ribosome is a ribonucleoprotein complex that is responsible for the process of protein synthesis. It constitutes up to 30% of the dry mass of the cell, making the synthesis of ribosomes one of the most energy demanding tasks in the cell (Wilson & Nierhaus 2007). In actively growing bacteria, one of the challenges for ribosome assembly is to coordinate the synthesis of ribosomal proteins (r-proteins) and ribosomal RNA (rRNA) as well as the binding of more than 50 r-proteins to the 4500 nucleotide long rRNA in a correct manner, following a defined hierarchy to ensure proper formation of ribosomes. Therefore, trying to understand the assembly process of the bacterial ribosome has been an intensive area of research for decades.

It was demonstrated during the 1970s that bacterial ribosomal subunits can be reconstituted *in vitro* from purified rRNA and r-protein components (Held et al. 1973; Held et al. 1974; Culver & Noller 1999). The conditions required for the reconstitution are far from physiological and included high temperature treatments and high salt concentrations to facilitate the assembly process (Herold & Nierhaus 1987). In contrast, several protein factors have been identified to regulate ribosome assembly *in vivo*. However, the exact functions of these factors in ribosome biogenesis remain largely elusive. In my thesis, I have used structural and biochemical approaches to study the role of ribosome biogenesis factors involved in the assembly process. Results in this thesis provide novel information for how these assembly factors contribute to the maturation of ribosome subunits. In addition, the high resolution structures provided in this thesis allow for direct visualization of the

assembly process. Hence, the work presented in this thesis help to better understand how the ribosome gets assembled.

1.1 Ribosome Structure and Function

The bacterial ribosome sediments as a 70S particle, and has a molecular weight of approximately 2.5 MDa, of which 60% is rRNA and the remaining 40% is r-proteins (Moore 1991). It consists of two subunits, the 30S and 50S subunit, which play different roles during protein synthesis.

1.1.1   Description of the Structure of the Ribosome

The bacterial ribosome is composed of two subunits, the large 50S and the small 30S subunits (Figure 1.1). The 50S subunit consists of 5S RNA, 23S RNA and more than 30 different r-proteins designated as L-proteins. The 30S subunit is comprised of the 16S RNA and 21 r-proteins designated as S-proteins. Except for r-protein L7/L12, all of these ribosomal components are present as only one copy per particle (Nomura et al. 1984). The primary structure of many ribosomal components and the secondary structures of the rRNAs from different bacterial species have been determined in detail. The 30S subunit can be divided into four structural domains based on the secondary structure of the 16S rRNA: the head (3' major domain), the body (5' domain), the platform (central domain) and the h44 and h45 located at the subunit interface (3' minor domain) (Figure 1.1). These four domains converge on the neck region of the subunit, which represents the functional core of 30S subunit, the decoding center (DC) (Wimberly et al. 2000). Different from the 30S subunit, only the secondary structure of the 23S rRNA of 50S subunit can be divided into six large asymmetric structural domains, which form together with 5S rRNA the main frame of

the 50S subunit (Ban 2000). The crown view of the 50S subunit displays a central protuberance (CP), and two lateral protuberances, with the L1 stalk to the left and the L7/L12 stalk to the right of CP (Figure 1.1). The large ribosome subunit is monolithic, except for the two stalks that are known to be flexible and not well ordered.
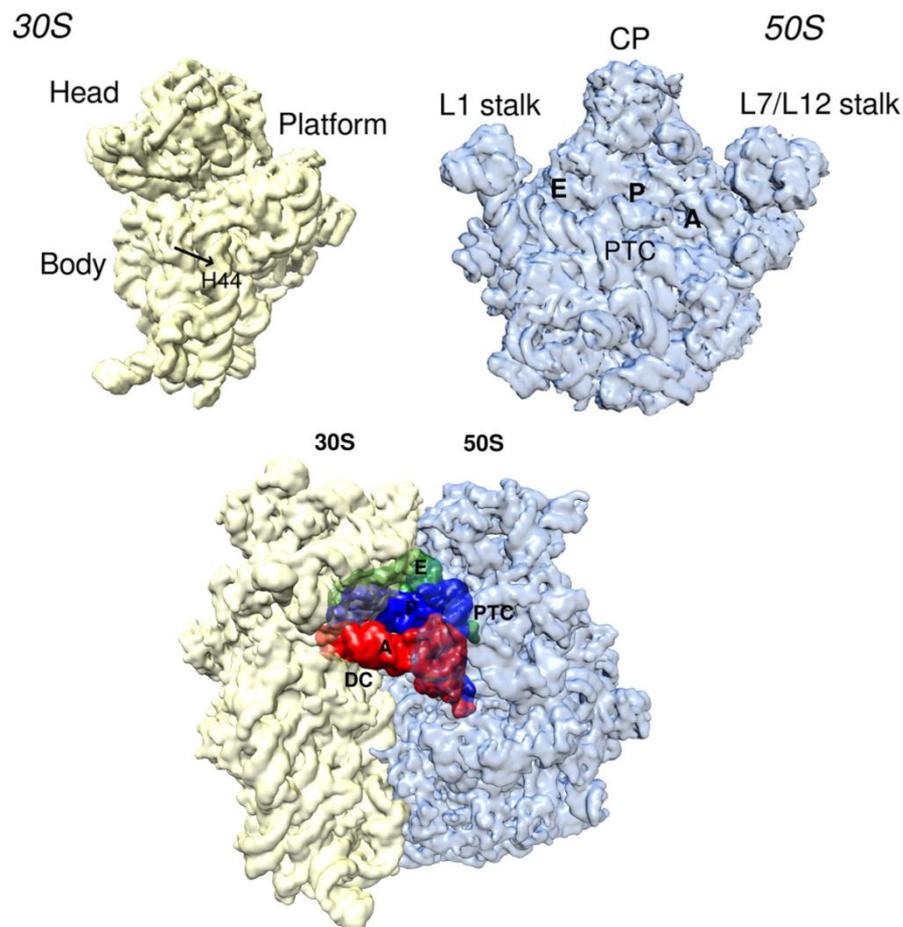
Figure 1.1 Structure of the bacterial ribosome

Structure of the bacterial ribosome including the small 30S subunit and the large 50S subunit. The 30S subunit is colored in yellow and the 50S subunit is colored in blue. The A-site (red), P-site (blue) and E-site (green) are the t-RNA binding sites that have special contacts with the DC (decoding center) in 30S subunit and the PTC (peptidyl transferase center) in the 50S subunit. PDB ID: 4V4Y and 3J9W.

The r-proteins of both subunits are dispersed throughout the entire structure, mostly on the surface. However, the proteins are largely absent from the interface between two subunits, the decoding center of 30S subunit and the peptidyl transferase center (PTC) of 50S subunit, which is of functional significance to protein synthesis (Figure 1.1).

The atomic-resolution structures of both the 30S and 50S ribosomal subunits have provided significant insights into the mechanism of protein synthesis. However, the complete structure of the intact the 70S ribosome reveals that the 30S subunit undergoes large conformational changes upon association with the 50S subunit. These conformational rearrangements include the ratchet-like motion between the small and the large subunits. The rotation of the head of the 30S subunit may related to the trigger of mRNA and transfer RNA (tRNA) movements during protein synthesis (Schuwirth 2005). The structure also reveals several inter-subunit contacts (inter-subunit bridges) that spread over the interface of both the 30S and the 50S subunits (Figure 1.2) (Yusupov et al. 2001). Most of the bridge contacts involve rRNA helices, with a few r-protein interactions. In the 30S subunit, the RNA-RNA contacts are centrally located around the decoding center (DC), directly abutting the tRNA binding sites, whereas the contacts involving proteins are located more distal from the functional sites (Figure 1.2, left panel). In the 50S subunit, the rRNA contacts are located centrally as well and around the peptidyl transferase center (PTC). These inter-subunit bridges are important interactions between the 30S and the 50S subunits when they associate during protein synthesis.
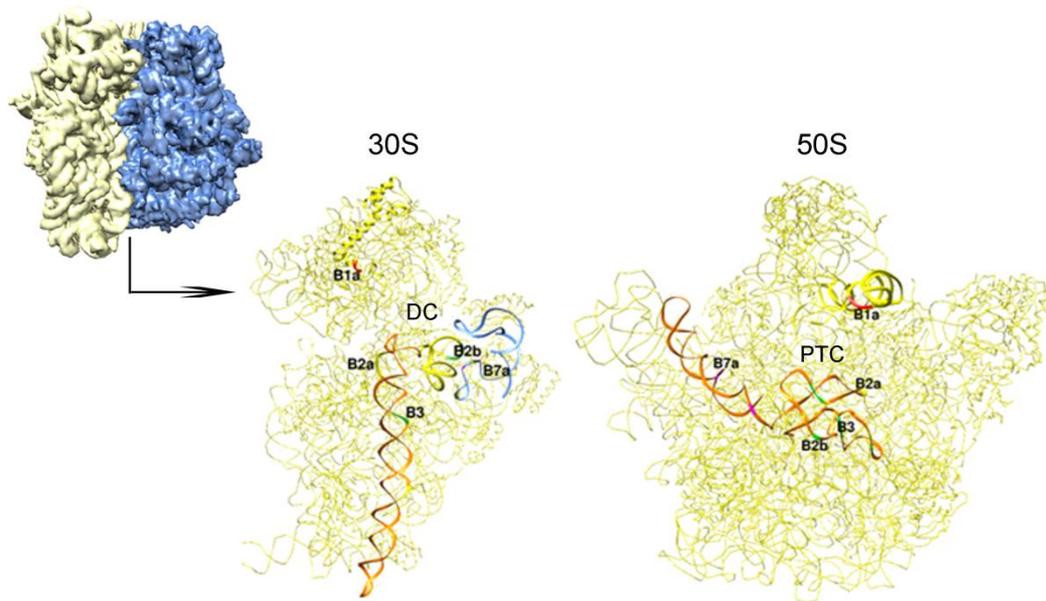
Figure 1.2 Interface views of the inter-subunit bridges

The intersubunit bridges between 30S and 50S subunit. Representative bridges are numbered as B1a, B2a, B2b, B3, B7a, which located around the PTC (peptidyl transferase center) on the 50S subunit and the DC (decoding center) on the 30S subunit.

1.1.2   Brief Description of the Function of the Ribosome

The correlation of ribosomal structure with specific ribosome function has been a major topic in the study of the ribosome for many years. The ribosome is responsible for protein synthesis (translation) (Bąkowska-Zywicka & Tyczewska 2009) in all forms of life (Schuwirth 2005). During this protein translation process, the mRNA interacts with the 30S subunit, while the tRNAs interact with both 30S and 50S in the decoding center (DC) and peptidyl transferase center (PTC) through three specific tRNA binding sites: the Aminoacyl (A)-site, the Peptidyl (P)-site, and Exit (E)-site (Steitz 2008; Rheinberger et al. 1981; Steitz & Moore 2003). Several GTP-hydrolyzing protein factors are involved during the stages of initiation, elongation and

termination of protein synthesis, such as the initiation factors (IF1, IF2, IF3), elongation factors (EF-Tu and EF-G), release factors (RF1, RF2 and RF3), and ribosome recycling factors (RRF1 and EF-G) (Noller 1991; Steitz 2008). Additionally, protein translation is a high fidelity process to ensure the correct transmission of genetic information from mRNA (Dnina & Wintermeyer 2001). For this purpose, the interaction between the anti-codon of the tRNA and codon of the mRNA is a kinetic proofreading process during Aminoacyl (A)-tRNAs (aa-tRNAs) selection by the 30S subunit (Dnina & Wintermeyer 2001; Hopfield 1974; Ehrenberg & Blomberg 1980), particularly, the codon-anticodon recognition assessed by the A1492, A1493 and G530 nucleosides of the 16S rRNA (Agris et al. 2007).

1.2 Ribosome Assembly and Assembly Factors

1.2.1 *In vitro* Ribosomal Assembly

It has been established that both 30S and 50S subunits can be reconstituted *in vitro* from purified rRNAs and r-proteins (Held et al. 1973; Held et al. 1974; Culver & Noller 1999). In the assembly of the 50S subunit *in vitro*, three reconstitution intermediates (RI) are formed consecutively from the association of r-proteins and rRNAs: RI_{50} (1), $RI_{50}*$ and RI_{50} (2), while two intermediates RI_{30} and $RI_{30}*$ are formed in the assembly of 30S subunit (Figure 1.3) (Herold & Nierhaus 1987; Culver & Noller 1999). The balance among different intermediates varies depending on the ionic conditions and the temperature of the reconstitutions (Shajani et al. 2011).
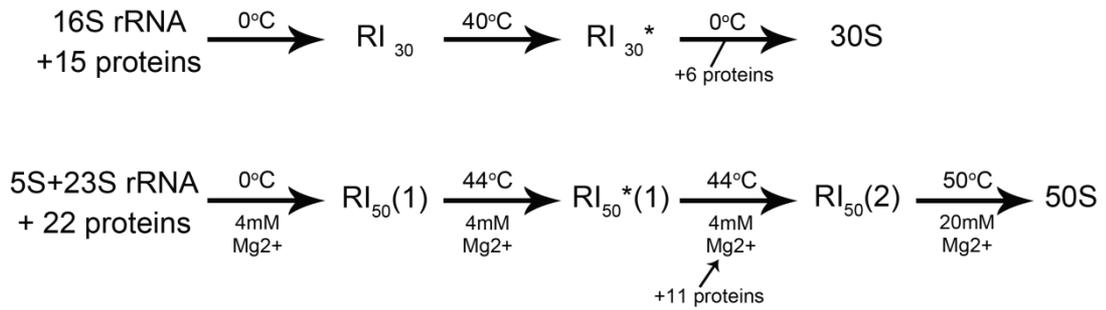
Figure 1.3 The *in vitro* ribosome assembly process

The ribosome subunits can be reconstituted in vitro from purified rRNA and r-protein components on specific ionic and temperature conditions.

During the assembly process, it was found that the binding of r-proteins to the rRNA follows a strict hierarchy. The r-proteins are defined as primary when they directly bind to the nascent 16S rRNA; secondary r-proteins require at least one bound primary protein to be able to bind to the rRNA; tertiary r-proteins require at least one bound primary and secondary protein to bind to the rRNA (Nomura et al. 1968; Shajani et al. 2011). The 50S subunit assembly map is significantly more complicated than the 30S subunit because of the number of components involved (Figure 1.4) (Nierhaus 1991; Herold & Nierhaus 1987).
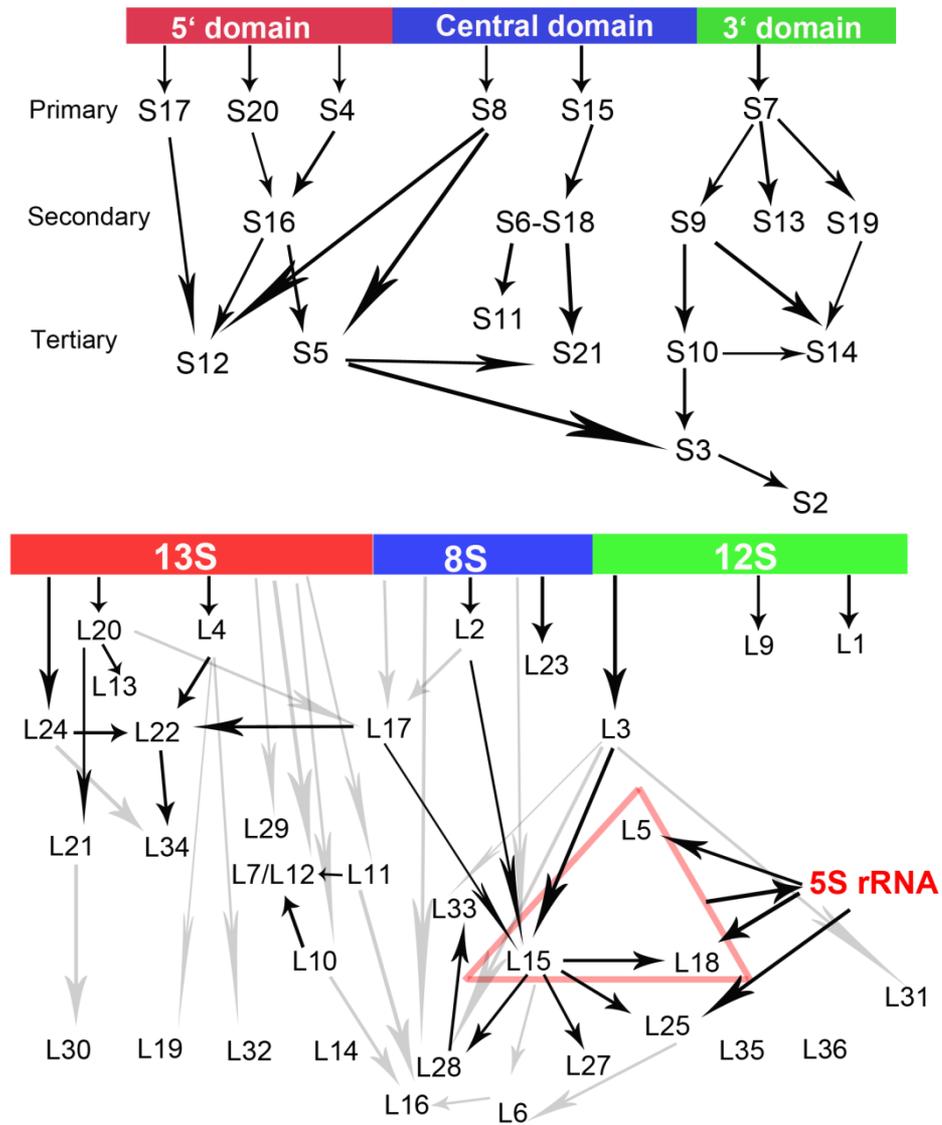
Figure 1.4 Nomura and Nierhaus assembly maps of the 30S and 50S subunits

The Nomura map (top) and Nierhaus map (bottom) show a strict hierarchal binding of the r-proteins to the 16S and 23S rRNAs.

Due to the existence of rate-limiting intermediates based on the studies of *in vitro* reconstitution of the 30S and 50S ribosomal subunits, ribosome assembly was initially thought to proceed through a single linear pathway (Shajani et al. 2011; Pan et al. 1997; Ralston et al. 2000). However, using time-resolved hydroxyl radical

footprinting methods, the Woodson lab (Adilakshmi et al. 2008) demonstrated that assembly is not cooperative at every stage. The multiphase rRNA folding kinetics strongly indicate that the 30S subunit takes different routes to the final structure. In addition, Williamson and coworkers utilized time-resolved electron microscopy method to identify and visualize several *in vitro* 30S intermediates. By the combination of pulse-chase quantitative mass spectrometry assay for measuring protein binding rates, they provided specific evidence for the parallel assembly pathways (Mulder et al. 2010). Hence, altogether these findings indicated that the assembly of the ribosome proceeds through multiple parallel pathways rather than a single pathway (Shajani et al. 2011).

1.2.2 *In vivo* Ribosomal Assembly

Different from *in vitro* ribosomal assembly, the process of assembly *in-vivo* is comprised of three main steps: the synthesis, processing, modification and folding of rRNA; the synthesis, modification and folding of r-proteins; and the binding of r-proteins to the rRNA. The 16S, 23S and 5S rRNA *in vivo* are initially synthesized as one primary transcript and processed by several RNases (Figure 1.5). The RNase III cleaves the initial transcript into precursor 16S rRNA (17S rRNA), precursor 23S rRNA, and precursor 5S rRNA (Ginsburg & Steitz 1975). The 17S rRNA contains an additional 115 nt at the 5' end of the 16S rRNA and 33 nt at the 3' end. The 5' end is further cleaved by RNase E and RNase G. However, many species lack RNase E and RNase G orthologues, thus in *B. subtilis,* the 16S rRNA is processed by ribonuclease RNase J1 (Britton et al. 2007). The 33nt extra residues at the 3' end are removed by any one of the four exoribonucleases, RNase II, R, PH or PNPase (Sulthana &

Deutscher 2013). The 7 to 9 nucleotides at the 3' terminus of the precursor 23S rRNA are processed by RNase T, but the enzyme for processing of the additional 3 or 7 nucleotides in the 5' end is still unknown (Li et al. 1999). The precursor 5S rRNA is composed of 84 additional nucleotides at the 5' end and 42 nucleotides at the 3' end. Precursor sequences at both ends are cleaved by RNase E. The same as the 23S rRNA, the final maturation of the 5S rRNA at the 5' end is processed by an unknown enzyme, while the 3' terminus is processed by RNase T (Li & Deutscher 1995).
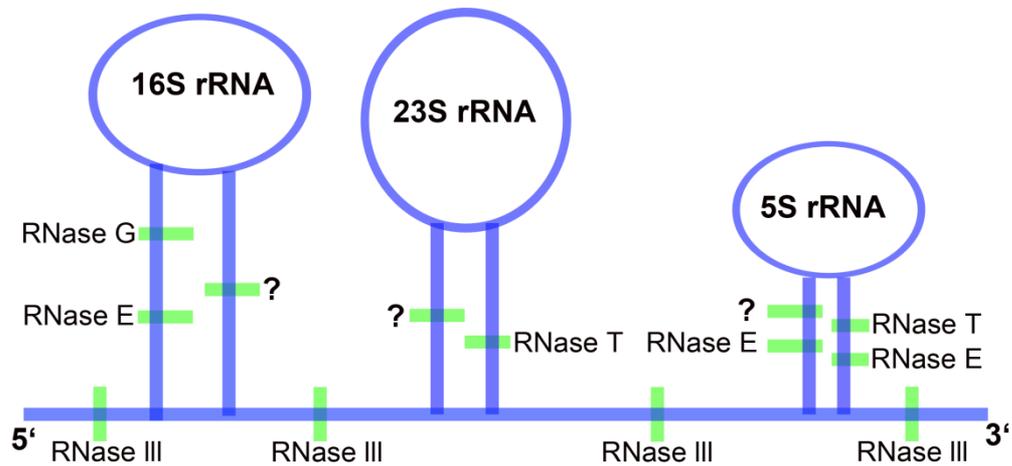
Figure 1.5 Maturation of ribosomal RNA (rRNA)

The 16S, 23S and 5S rRNA are originally transcribed as a single transcript and then are cleaved by a series of RNases to mature into the final products. Cleavage sites are indicated by green dashes. Unknown RNases are indicated by question marks.

The maturation of rRNA also includes extensive nucleotide modifications in addition to the rRNA cleavages. Most of the modification occurs in the important functional regions, such as the peptidyltransferase center, the A, P and E sites of tRNA and

mRNA binding, the polypeptide exit tunnel, and the inter-subunit interaction sites (Decatur & Fournier 2002). These modification-rich regions are also the sites where specific protein translation happens, which indicates that the modifications influence both the structure and function of the ribosome. The absence of modified nucleotides in the areas dominated by ribosomal proteins in the external surfaces of the ribosome indicates that most RNA-protein interactions are not influenced directly by rRNA modification (Ban 2000; Schluenzen et al. 2000). Hence, the clustered distribution of rRNA modification probably reflects its roles in properly 'building' the functional regions (Decatur & Fournier 2002). Some other functions of rRNA modification also related to antibiotic resistance, such as the methylation of specific nucleotides within the A-site of 16S rRNA have been identified as a new mechanism of resistance against aminoglycosides by hampering its binding to the 30S ribosomal subunits (Doi & Arakawa 2007).

Although the exact role of modification is not clearly understood, it has been suggested that methylation of ribosomal proteins might be important for the assembly of ribosomal particles as well. The methylation of ribosomal proteins mostly takes place on the proteins involved in peptidyltransferase or in polyphenylalanine-synthesizing activity, such as uL5, uL3, uL10, uL11, bL7/L12 and uL16 of *E. coli*, indicating the possible functional role in protein biosynthesis (Amaro & Jerez 1984). In addition, the methylation of ribosomal proteins is conserved in different bacterial organisms, and many of the proteins carry more than one modification, suggesting the potential general significance of these r-protein modifications for ribosome structure and function (Amaro & Jerez 1984; Kaczanowska & Rydén-Aulin 2007).

1.2.3 Ribosome Assembly Factors

Ribosome assembly is a tightly coordinated series of events that include synthesis, processing and modification of rRNA and r-proteins. One of the problems during RNA folding is that it is structurally promiscuous. With the growing length of the RNA, it is easy to create kinetic traps (Kaczanowska & Rydén-Aulin 2007; Williamson 2003). The rate-limiting steps of *in vitro* reconstitution of the 30S and 50S subunits are likely to be these kinetic traps that result in the accumulation of the reconstitution intermediates (RI) (Nierhaus 1991; Herold & Nierhaus 1987; Culver & Noller 1999). As previously described, certain non-physiological ionic conditions and high temperature are required to complete assembly. Coupled with the binding of r-proteins, the reconstitution intermediates finally convert into the mature ribosome subunits. Hence, the *in vitro* reconstitution of ribosome particles can be completed without exogenous factors. In comparison, the *in vivo* reconstitution for ribosome assembly occurs much faster than the assembly *in vitro* (Shajani et al. 2011), which indicates that there are protein factors in the cell that facilitate and speed up the assembly process (Kaczanowska & Rydén-Aulin 2007).

Over the past 30 years, a large number of protein factors (>200) have been identified to be involved in the different processes of ribosome assembly in eukaryotes, especially in yeast (Hage & Tollervey 2004; Kressler et al. 2010). In contrast, a very small number of assembly factors are involved in the synthesis of the bacterial ribosome (Shajani et al. 2011). The functions of these factors include promoting the processing and modification machinery, facilitating proper RNA folding and protein-RNA interactions during the ribosome assembly process (Hage & Tollervey 2004;

Shajani et al. 2011). They can be classified into several classes: DEAD-box proteins, RNA chaperones, ribosome-dependent GTPases and other maturation factors (Shajani et al. 2011; Connolly & Culver 2009; Kaczanowska & Rydén-Aulin 2007). Among all these assembly factors, the GTPases are an especially intriguing group. This thesis will focus on three of these GTPases (RbgA, YphC and YsxC) that are involved in the assembly of the large ribosomal subunit.

GTPases regulate diverse cellular processes including factors involved in protein translation, for example, IF2, EF-Tu or EF-G. They also play key roles in the assembly of ribosomes in bacteria and eukaryotes (Britton 2009). Although the precise roles of many of these GTPases are still unclear, it was predicted that the GTPase involved in the ribosome assembly might allow the cells to use them to mediate its own ribosome assembly process. The decreasing GTP levels in the cell under stress conditions would couple the ribosome assembly directly to the energy state of the cell (Blombach et al. 2011; Britton 2009). There are several GTPases that have been implicated in assembly of the large 50S subunit, such as RbgA (YlqF), YphC and YsxC proteins (Britton 2009; Wilson & Nierhaus 2007; Goto et al. 2013). RbgA (ribosome biogenesis GTPase A), also known as YlqF, is an essential GTPase involved in the late stages of assembly of the 50S subunits in *Bacillus subtilis* (Britton 2009). It contains two structural domains, an N-terminal circularly permuted G-domain, in which the order of the highly conserved G1-G4 motifs is altered to the G4-G1-G2-G3, and a C-terminal domain designated as the ANTAR domain due to its structural similarity to the AmiR-NasR Transcription Anti-Termination Regulator (ANTAR) domains. Mutational analysis in both the N- and C-terminal ends of the

protein suggested that the ANTAR domain of RbgA is essential for association with the ribosome, while mutations in the conserved G-domains abolished GTP hydrolysis (Gulati et al. 2013). Depletion of RbgA in *B. subtilis* results in a dramatic reduction of 70S ribosomes and the accumulation of immature 45S intermediates. Analysis of this immature particle reveals that it is physiologically relevant and is competent to mature into a complete 50S subunit (Jomaa, Jain, Joseph H Davis, et al. 2014). The 45S intermediate lacks several ribosomal proteins including uL16, bL28, bL27, bL33a, bL35 and bL36, which can be found at the base of the central protuberance (CP) region (Figure 1.6), suggesting that the functional core including the A-site, P-site and the peptidyl-transferase center (PTC) of the 50S subunit are not correctly structured. In addition, all of these underrepresented proteins are at the last branch of the *in vitro* 50S assembly map, which suggests that these 45S intermediates represent an assembly intermediate at late stages of assembly (Uicker et al. 2006). Consistently, the cryo-EM reconstructions of the 45S particle show a mature body domain, but exhibit significant conformational alterations in the CP region around the peptidyl-transferase center (PTC), the functional core of 50S subunits (Jomaa, Jain, Joseph H Davis, et al. 2014; Li et al. 2013). These results suggest the important activity of RbgA in facilitating the conformational changes in the functional core during the maturation of 50S subunits.

Homologs of RbgA protein can be found in all eukaryotes and are implicated in ribosome assembly in both the cytoplasm and mitochondria. In eukaryotes, the human Mtg1 protein likely plays a role in the assembly of mitochondrial large ribosome subunits, suggesting a similar and conserved functional activity of Mtg1 and RbgA (Britton, 2009; De Silva et al., 2015).
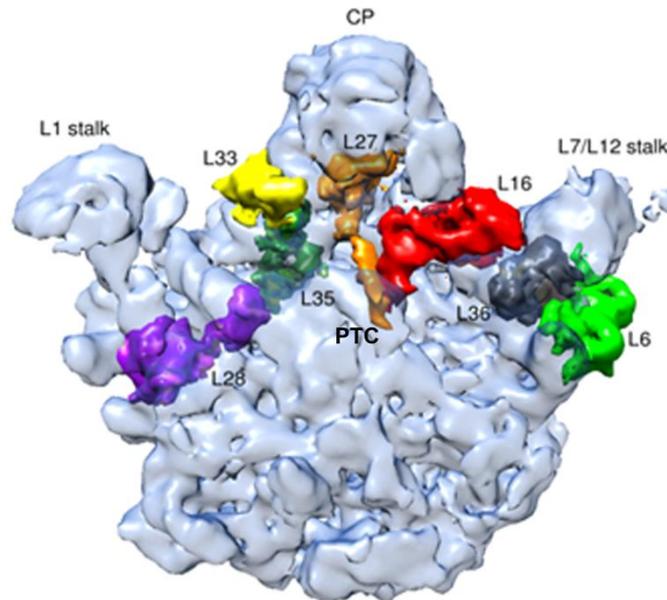
Figure 1.6 Location of missing proteins at the immature 45S intermediates

The 45S intermediate from RbgA-depletion lacks several ribosomal proteins including uL16, bL28, bL27, bL33a, bL35 and bL36, which are located at the base of the central protuberance (CP), around the peptidyl-transferase center (PTC).

The same as RbgA, the YphC and YsxC proteins are two GTPases that are essential for the growth of *B. subtilis* (Morimoto et al. 2002). YphC protein is widely conserved in both bacteria and eukarytic cells. It contains two GTP binding domains (GD1 and GD2) in the N-terminal region and a KH-like domain linked to the C-terminal domain (Goto et al. 2013). The cryo-EM reconstruction of 50S: EngA (Homolog of YphC in *E. coli*) complex showed that the EngA protein binds to the peptidyl transferase center (PTC), with its GD1 domain interacting with components of the E-site including H82, H88 and L1 protein, and the GD2 located underneath the H68 and largely displacing this helix, while the C-terminal KH domain shows strong interactions with the PTC functional core including H80, H89 and H93 (Zhang et al. 2014). These results suggest that the EngA protein is likely to facilitate conformational maturation of the

PTC of 50S subunits. And the two GTP domains, endowed with multiple conformations, might probably work as a sensor of the cellular GTP/GDP ratio to regulate the ribosome assembly process in the cell (Zhang et al. 2014).

YsxC (YihA in *E. coli*) is essential to the growth of *B. subtilis* and *E. coli* (Wicker-Planquart et al., 2008). It has several functions including participating in cell division and progression of the cell cycle, and particularly in the biogenesis of the large ribosomal subunit (Britton 2009; Schaefer et al. 2006; Wilson & Nierhaus 2007). YsxC binds to the 50S subunit in a GTP-dependent manner, and it also has weak binding to the 30S subunit. YsxC can directly interact with ribosomal proteins L1, L6 and L7/L12, however, the exact mechanism for interaction with the ribosome are still unknown (Wicker-Planquart et al. 2008).

1.2.4 Assembly Factors as Targets for Antimicrobials

Most of the antibiotics used in the clinic carry out their antimicrobial functions by blocking protein synthesis on the bacterial ribosome (Poehlsgaard & Douthwaite 2005). However, evolution drives resistance and thus, antibiotic resistance represents a continuous threat to public health (Brown & Wright 2016). Therefore, the discovery of new antimicrobials with new mechanisms is of great interest. In bacteria, many of the protein factors that are involved in ribosome assembly appear to be essential, highly conserved, and soluble to be able to generate high-resolution structural information. Hence, in recent years they are considered as new targets for antimicrobial drugs (Comartin & Brown 2006). Specifically, Era, YjeQ and RbgA are assembly factors that are involved in the late stages of assembly of the 30S or 50S subunit. Structural study of Era showed that it is overlapping the binding site of S1

protein and the conformation was exhibited to inhibit 30S association with 50S (Sharma et al. 2005). YjeQ, despite being a non-essential assembly factor, was identified as important for growth *in vivo* in an animal infection model. The deletion of *yjeQ* severely attenuated the virulence of *Staphylococcus aureus* (Campbell et al. 2006). RbgA, as previously described, is an essential factor present in a variety of gram-positive bacteria. The homologues of RbgA protein appear in eukaryotes, including human and yeast (Uicker et al. 2006). These factors are examples that appear to represent novel targets for new antibacterial drugs targeting ribosome subunit assembly. However, novel technologies and better understanding of the assembly process *in vivo* are required for the discovery of new inhibitors of ribosome assembly (Comartin & Brown 2006).

1.3 Recent Developments and Challenges of Single-Particle Cryo-electron Microscopy (cryo-EM)

Single-particle cryo-EM, a structural biology technique, is used to determine the structure of proteins and macromolecular complexes. It has come a long way, to a level of a quantum leap progression within only 20 years (Frank 2009). Structural studies of the 50S ribosomal subunit began with the milestone 25Å resolution cryo-EM map of the *E. coli* ribosome published in 1995 (Frank et al. 1995). Then a 5Å resolution map that can be fitted with protein and RNA structures was reported in 1999 (Ban et al. 1999). And one year later, a 2.4Å resolution map (Ban 2000) that allowed the building of a complete atomic model of the 50S subunit was available. It took only 5 years to realize the transition from a low-resolution map to atomic models. Additionally, with the recent advancements in electron detection and new image

processing tools, cryo-EM method is now a powerful alternative to X-ray crystallography and can be used to determine atomic resolution structures of macromolecules that are still difficult to crystallize in conformations representing functional states (Bai et al. 2014; Cheng 2015; Nogales & Scheres 2015).

1.3.1 Recent advances in Single-Particle Cryo-EM

Direct electron detectors

The potential of single-particle cryo-EM has dramatically improved with the development of direct electron detection (DDD) cameras in recent years. Traditionally, the frozen-hydrated biological samples were recorded with either photographic film or scintillator-based digital cameras, such as a charge-coupled device (CCD) (Wu et al. 2015). The attainable resolution using film or CCD camera is limited by the large point spread function of this recording media. In contrast, direct detection cameras capture the electrons directly onto the sensor and greatly improve image quality (Tem 2015). In addition, DDD cameras have a significant improvement in the detective quantum efficiency (DQE), a measurement of combined effects of the signal and noise performance to judge a high resolution EM camera (Cheng et al. 2015; Tem 2015). The dramatically improved signal-to-noise ratios in the experimental data by using DDD improves the contrast of the images that allows the use of a smaller defocus, ultimately contributing to the higher resolution information in the images and in the final three-dimensional reconstructions.

Another important breakthrough of the DDD camera is their fast frame readout rate that enables images to be acquired as a movie comprised of stacks of many frames instead of a single micrograph (Cheng 2015; Wu et al. 2015). The fast read out of the

DDD also allows splitting the total electron dose into single frames, where the beam-induced image blurring is minimized (Nogales & Scheres 2015). The computational frame alignment is then carried out after data collection to be able to correct the beam-induced image motions (Bai et al. 2013; Rubinstein & Brubaker 2015). The combination of dose splitting and motion correction significantly improves the image quality due to the preservation of high-resolution information, and ultimately can be reflected in the final 3D EM maps.

Improved image processing tools

Simultaneously with the developments in new detectors, improved image processing tools are now also available. Sample heterogeneity is a major factor that influences the final resolution of the obtained reconstructions. However, in a way, the major advantage of single-particle cryo-EM is that it does not require an absolutely homogeneous sample (Cheng 2015). Crystallization is prevented by sample heterogeneity, but in cryo-EM, particles can be computationally classified into different classes comprised of homogeneous populations. The "orientation problem" is available in single-particle analysis, in which the relative orientations of all particles can be considered as missing data (Nogales & Scheres 2015). A powerful image processing approach to integrate the distributions of all possible orientations and classify particles according to their conformations is maximum likelihood-based method for classification and refinement (Scheres 2010). Implementation of this method into user-friendly software packages, such as the FREALIGN (Lyumkis et al. 2013) and RELION programs (Scheres 2012), made this method easy to use in practice. Especially, the RELION program has quickly gained popularity due to its

powerful and user-friendly application for 3D classification and high resolution refinements (Bai et al. 2014).

1.3.2 Challenges and Perspectives in Cryo-EM

In recent years, single-particle cryo-EM has revolutionized structural biology. However, this technology is still facing challenges and further developments are still necessary. Direct electron detectors can still be improved. Certain complexes fall apart due to physical forces on the sample during blotting or interactions with the air-water interface during grid preparation (Nogales & Scheres 2015; Glaeser 2016). With a resolution higher than 3 Å is possible to visualize not only where but also how small ligands bind to target proteins, which is of particular interest for the pharmaceutical company to investigate a broad spectrum of drug-target interactions and to use it in structural-based drug design (Cheng 2015; Merk et al. 2016). However, for the complex itself with continuously functional states, the structure which is rich in flexible elements or complex that is very small, it is still challenging to separate the particles into structurally homogeneous subsets (Bai et al. 2014; Nogales 2016). Therefore, the continuing developments including the optimizations in sample preparation, the improved image processing techniques, as well as even better detectors and the commercial use of phase plates that will allow us to image smaller particles, are still required and under development in recent years and in the near future (Nogales & Scheres 2015; Nogales 2016).

In addition, with the development of single-particle cryo-EM, both academic and industry parties need to make efforts to make wider experimental access for the cryo-EM community to do high-throughput data acquisition. And the development of

enormous computational resources needed for data storage and image processing is also highly required (Cheng 2015; Nogales 2016; Bai et al. 2014; Glaeser 2016). These efforts are important steps to push the technology further in the single-particle cryo-EM perspectives.

1.4 Thesis Objective

As we previously described, there are multiple GTPases that have been implicated in the assembly of the large 50S subunit, such as RbgA (YlqF), YphC and YsxC proteins (Britton 2009; Wilson & Nierhaus 2007; Goto et al. 2013). Several genetic and biochemical experiments had been set up for the characterization of these factors (Uicker et al. 2006; Achila et al. 2012; Matsuo et al. 2006; Muench et al. 2006; Wicker-Planquart et al. 2008). However, we still do not know how these GTPases contribute to the ribosome assembly process. Ribosome assembly intermediates generated through depletion and knock-out bacterial strains constitute today an important tool to study the function of assembly factors. Thus, it would be interesting to use both biochemical and structural approaches to characterize these assembly intermediates and study the function of assembly factors during ribosome assembly.

The objective of this thesis is the identification of the function of three essential GTPases, RbgA, YphC and YsxC, in the assembly process of the large 50S ribosome subunit. Work in this thesis will be addressing these three specific aims:

1. Establish the function of YphC and YsxC based on the structural analysis of the immature ribosomal particles that accumulate in the YphC and YsxC depleted stains in *Bacillus Subtilis*.

2. Determine the nature and binding properties of the immature ribosomal

particles that accumulate in these depletion strains.

3. Identify the binding site and conformational changes induced by RbgA upon binding to the mature 50S subunit.

1.5 Thesis Organization

This thesis is divided into three research chapters (Chapter 2-4), with the additional general introduction in Chapter 1 and discussion in Chapter 5. Chapters 2 and 3 are modified from a published paper in the journal *Nucleic Acids Research*. I am the first author in this publication. The work in Chapter 2 studied the function of assembly factors YphC and YsxC by characterization of the *in vivo* assembly immature ribosome intermediates using the cryo-EM approach. The work in Chapter 3 describes the nature and binding properties of the immature ribosomal particles accumulating upon the RbgA, YphC or YsxC depletions. Chapter 4 is an ongoing project aiming to solve the structure of the 50S subunit in complex with RbgA and study the functional interplay between RbgA and ribosomal particles. Chapter 1 is aimed to provide appropriate background information on the ribosome and the development in single-particle cryo-EM, while Chapter 5 is intended as a summary and discussion about the findings and significance of our work in this thesis.

CHAPTER 2. Function of YphC and YsxC Revealed by Analysis of *in vivo* Assembled Immature Ribosomal Particles Using Cryo-electron Microscopy

2.1 Author's Preface

The work in chapter 2 is part of a paper published in the journal *Nucleic Acids Research* and is modified in this chapter. The work characterized the *in vivo* assembly immature ribosome intermediates by using cryo-EM. This work was carried in collaboration with Dr. Robert Britton's and Dr. James Williamson's laboratories. I have performed all of the cryo-electron microscope experiments including sample preparation, data collection, and image processing. Dr. Joseph Davis from the Williamson lab performed the quantitative mass spectrometry experiments. The YphC and YsxC depleted stains for us to purify the immature intermediates were provided from Dr. Britton's lab. Dr. Joaquin Ortega and I analyzed the data and wrote the manuscript.

The full citation is as follows:

Xiaodan Ni, Joseph H. Davis, Nikhil Jain, Aida Razi, Samir Benlekbir, Andrew G. McArthur, John L. Rubinstein, Robert A. Britton, James R. Williamson and Joaquin Ortega. (2016) YphC and YsxC GTPases assist the maturation of the central protuberance, GTPase associated region and functional core of the 50S ribosomal subunit. *Nucleic Acids Research*. 9140 (2016), 1–14.

2.2 ABSTRACT

YphC and YsxC are two GTPases in *Bacillus subtilis* that facilitate the assembly of the 50S ribosomal subunit, however their roles in this process are still uncharacterized. To explore their function, we used the strains in which the only copy of the *yphC* or *ysxC* genes were under the control of an IPTG inducible promoter. The incomplete ribosomal subunits accumulated under the depletion conditions were designated as $45S_{YphC}$ and $44.5S_{YsxC}$ particles. Quantitative mass spectrometry analysis and the cryo-EM maps of the $45S_{YphC}$ and $44.5S_{YsxC}$ particles revealed that the two GTPases participate in the maturation of the central protuberance, GTPase associated region (GAR) and the key RNA helices in the A, P and E functional sites of the 50S subunit. These results provide the first insights into the function of factors YphC and YsxC in 50S subunit assembly. The cryo-EM approach in combination with novel image processing methods to capture different states of maturation during the assembly process will also become a useful tool to study samples with flexibility and with different functional states.

2.3 INTRODUCTION

The ribosome is a ribonucleoprotein complex in all cells that is responsible for protein synthesis. Hence the biosynthesis of ribosomes becomes one of the major tasks in the cell (Wilson & Nierhaus 2007). A challenge in studying ribosome assembly in bacteria is that assembly *in vivo* is a highly efficient process and assembly intermediates are hard to accumulate. By using pulse labeling experiments, early studies characterized extremely small quantities of incomplete ribosomal particles that accumulate in normal cells (Lindahl L 1973; Lindahl 1975; Hayes et al. 1971). More recently, researchers have explored the use of small molecule screenings to find chemical inhibitors of the specific steps in the bacterial ribosome assembly process (Stokes & Brown 2015). Many of these small chemical molecules have been previously used as tools to dissect the translation process and capture the conformational changes that the mature ribosomes undertake during this process. However, there are only a few chemical probes that inhibit ribosome biogenesis in eukaryotes and only one in bacteria (Loibl et al. 2014; Pertschy et al. 2004; Stokes & Brown 2015; Drygin et al. 2011). Therefore, ribosome biogenesis has been studied almost exclusively by genetic and biochemical approaches nowadays.

Previous genetic studies on bacterial ribosomes found that depletion of several assembly factors results in the incomplete assembly of ribosome subunits and the accumulation of distinct immature intermediates (Jomaa et al. 2011; Leong et al. 2013; Guo et al. 2013), which can be isolated and characterized. Analysis of these particles provided important insights into how the protein factors assist the assembly process of the ribosomal subunits. The structural characterization of several 30S assembly

intermediates that accumulate in *Escherichia coli* cells lacking either YjeQ (Jomaa et al. 2011) or RimM (Leong et al. 2013; Guo et al. 2013) suggested that these assembly factors are assisting the maturation of the decoding center of the 30S subunit at the late stage of maturation.

Genetic approaches have also been used to establish the function of assembly factors assisting the maturation of the 50S subunit. RbgA (also known as YlqF), YphC and YsxC are three essential GTPases that are involved in the late stages of 50S ribosomal maturation (Uicker et al. 2006; Schaefer et al. 2006). Recently by using a *B. subtilis* strain in which RbgA was under the control of an IPTG inducible promoter, it was possible to purify the immature 50S particles ($45S_{RbgA}$) that accumulated in the cells under the depletion conditions (without IPTG) (Jomaa et al. 2014; Li et al. 2013). The quantitative mass spectrometry (qMS), cryo-electron microscopy (cryo-EM) and chemical probing revealed that RbgA plays a critical role in the maturation of the central protuberance and peptidyl transferase center of the 50S subunit. However, the mechanistic insights of how RbgA assists the maturation of these functional sites of the 50S subunit, and the exact functions of the other two GTPases (YphC and YsxC) remain largely unknown. Similarly to RbgA, cells depleted in YphC or YsxC also accumulate immature 50S subunits, designated as $45S_{YphC}$ and $44.5S_{YsxC}$ particles, respectively (Schaefer et al. 2006). Therefore, we use the structural analysis of these particles to demonstrate the function of YphC and YsxC in the assembly of the large 50S ribosomal subunit.

The Quantitative mass spectrometry (qMS) experiment indicated that the late-binding r-proteins including uL16, bL28, bL27, bL33a, bL35 and bL36 are dramatically

underrepresented in the $45S_{YphC}$ and $44.5S_{YsxC}$ particles, the same as the $45S_{RgbA}$ particles, which suggested they are late assembly intermediates of the 50S subunit. Additionally, the cryo-EM reconstructions showed that these particles exhibited significant structural differences from the mature 50S subunit in several important functional sites, including the A, P and E sites, central protuberance and GTPase associated region suggesting that YphC and YsxC, together with RbgA, play critical roles in the maturation of these functional regions.

2.4 RESULTS

2.4.1 The $45S_{YphC}$ and $44.5S_{YsxC}$ particles represent late assembly intermediates

The depletion of YphC and YsxC protein in *B. subtilis* results in the accumulation of defective large ribosome subunits that migrate in sucrose gradients slower than the mature 50S subunits and are called $45S_{YphC}$ and $44.5S_{YsxC}$ (Figure 2.1). To purify these immature particles, we used strains in which the only copy of the *yphC* or *ysxC* gene is under the control of an IPTG-inducible promoter (Schaefer et al. 2006). Hence, in the presence of IPTG, lysates from these strains fractionated by sucrose density gradient ultracentrifugation generated ribosome profiles that were indistinguishable from those of the wild type cells (Schaefer et al. 2006). However, the absence of IPTG produced a dramatic reduction in the level of 70S ribosome and the accumulation of $45S_{YphC}$ and $44.5S_{YsxC}$ particles (Figure 2.1). We also purified the mature 50S subunits and the RbgA-depleted 45S particles as previously described (Jomaa et al. 2014).
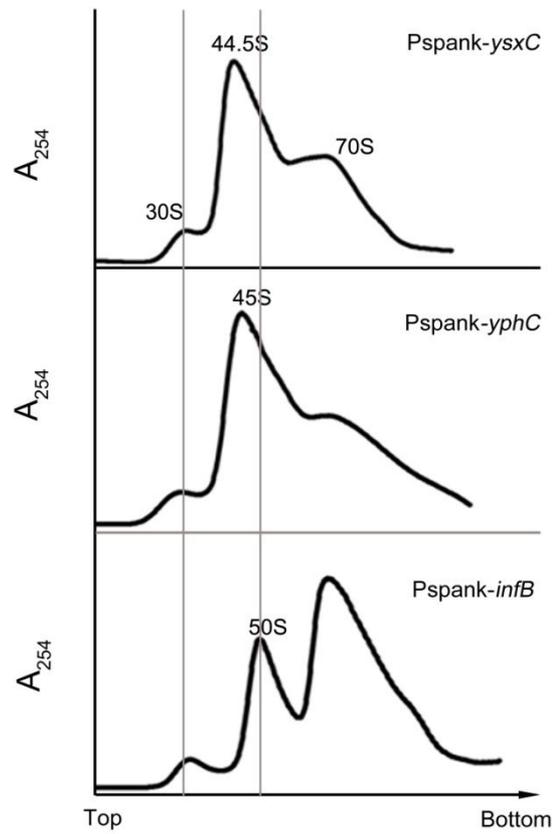
Figure 2.1 Purification of the immature 45S$_{YphC}$ and 44.5S$_{YsxC}$ particles and mature 50S subunits

The immature ribosomal particles were purified from cells where the only copy of the ysxC (RB260) or yphC (RB290) genes was under the control of the IPTG-inducible promoter P$_{spank}$. Mature 50S subunits were purified from a strain in which the only copy of infB (encoding for IF2) (RB419) was also under a P$_{spank}$ promoter. In the absence of IPTG these cells accumulate large amounts of immature and mature 50S subunits, respectively. The ribosomal profiles were generated by ultracentrifugation of the cell lysates through a 10–30% (wt/vol) sucrose gradient. Profiles were generated by monitoring the UV absorbance at 254nm through the sucrose gradient from the top to the bottom. Vertical lines indicate where the 30S and 50S mature particles migrate in the gradient.

To determine the protein complements of the three immature particles (45S$_{YphC}$, 44.5S$_{YsxC}$ and 44.5S$_{RbgA}$) and the mature 50S subunit, all cells were grown in the ^{14}N-labeled media. The purified particles from these cells were then mixed with a

reference spike containing a fixed concentration of ^{15}N-labeled 70S ribosomes and the protein levels of immature particles relative to this reference spike were measured using qMS. This analysis shows that r-proteins uL16, bL28 and bL35 (35) (r-protein nomenclature according to Ban et al) were severely depleted (occupancy < 0.25), and bL27, bL33 and bL36 were significantly reduced (occupancy > 0.25) from all three 45S particles (Figure 2.2A). The occupancy level for bL34 protein remained uncharacterized as no peptides were found for this protein. The mature 50S subunits exhibited a full complement and no signs of depletion for any of these r-proteins. To determine the effect of salt condition (NH$_4$Cl) on r-protein occupancy, these assays were performed in duplicate with particles purified either under low salt (150mM) or high salt (500mM) conditions. The purification under high salt conditions led to similar occupancy profiles, but with lower abundance of depleted proteins (Figure 2.2B). This control experiment confirmed that the buffer conditions used for particle purification and mass spectrometry analysis were not the cause of any of the depletions observed for these six r-proteins.

In addition, uL16, bL27, bL28, bL33, bL35 and bL36 are the late binding r-proteins when they bind to the 50S subunit in the reconstitution experiments (Herold & Nierhaus 1987; Chen & Williamson 2013). Therefore, this analysis suggested that the $45S_{YphC}$, $44.5S_{YsxC}$ particles, similar to the $45S_{RbgA}$ particles, are the late assembly intermediates.
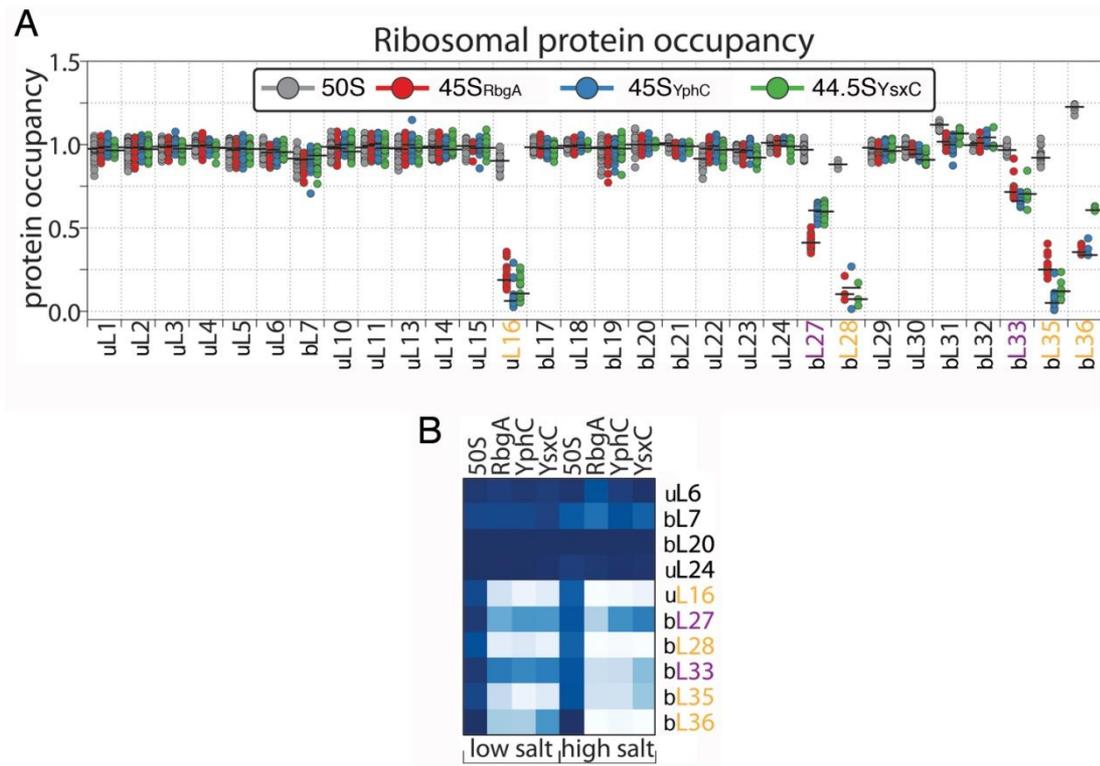
Figure 2.2 Ribosome protein occupancy measured by qMS

(A) Ribosomal protein occupancy in 50S (grey), $45S_{RbgA}$ (red), $45S_{YphC}$ (blue), and $44.5S_{YsxC}$ (green) particles purified in the presence of 150 mM NH_4Cl. MRM-like transitions were extracted for each product ion from the SWATH datasets (see methods) and $^{14}N/^{15}N$ abundance ratios were calculated and normalized to the median value determined for protein L20. Circles denote individual MRM-transition measurements, lines signify the median $^{14}N/^{15}N$ abundance ratio measured for each protein. (B) Protein occupancy for a subset of ribosomal proteins measured in particles purified in the presence of either 150 mM (low salt; left) or 500 mM (high salt; right) NH_4Cl. Occupancy from 0 to 1 scales from white to blue.

2.4.2 The immature particles accumulating in the YphC and YsxC-depleted cells exhibit multiple conformations

As the qMS experiment has demonstrated, the depletion pattern of protein complements in the $45S_{YphC}$, $44.5S_{YsxC}$ and $45S_{RbgA}$ particles is almost identical. This result led us to investigate the structural similarity between the three particles and

whether the function of YphC and YsxC could be inferred from the $45S_{YphC}$ and $44.5S_{YsxC}$ structures. To study the structures, the purified $45S_{YphC}$ and $44.5S_{YsxC}$ particles were imaged by cryo-EM (Figure 2.3, left panels) with a direct detector device camera allowing for the beam-induced motion correction of ribosomal particles in each frame during the image acquisition process (Figure 2.3, right panels).
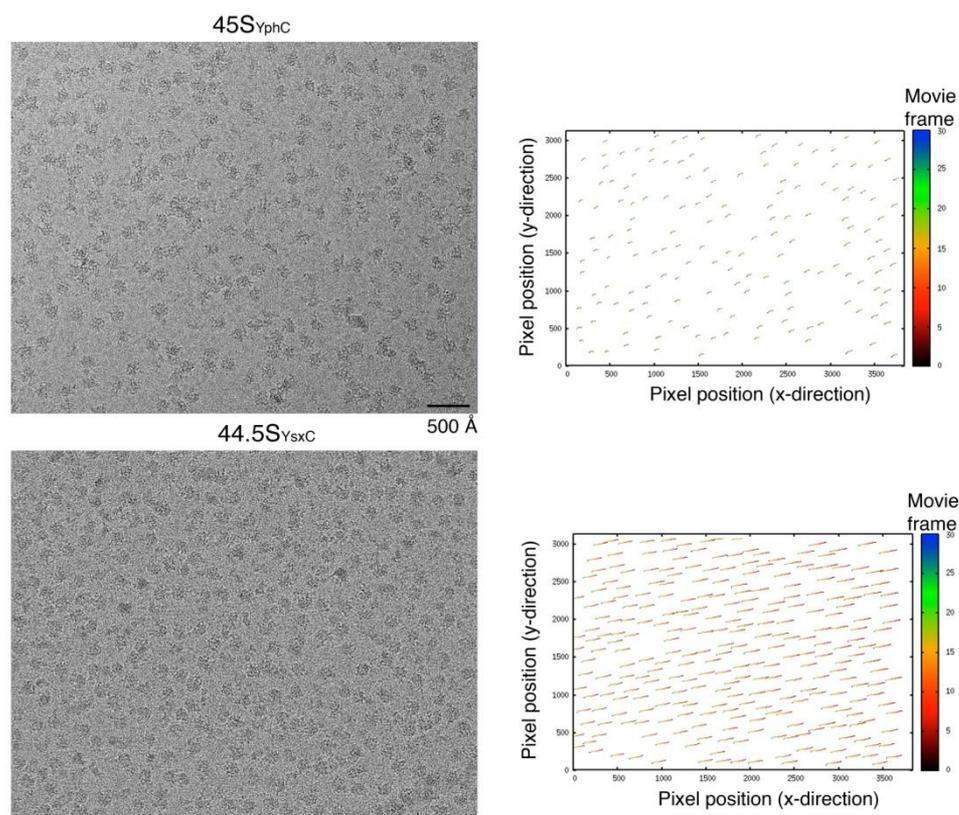
Figure 2.3 Cryo-EM images of the $45S_{YphC}$ and $44.5S_{YsxC}$ particles and the beam induced motion correction

The left panels are representative micrographs from $45S_{YphC}$ and $44.5S_{YsxC}$ particles. These micrographs were produced after frame alignment [Li, 2013 #3194] of the 30 frames in the original movies collected by the K2 Summit direct electron detector camera. Vector maps at the right panels display the particle trajectories during the 15 second exposures used to collect the movies. Trajectories are exaggerated by a factor of 5 to allow visualization. Beam induced motion correction was performed in individual particle images using the observed trajectories [Rubinstein, 2015 #3195].

Three-dimensional classification of the $45S_{YphC}$ data set using the entire signal in the particle images revealed one distinct three-dimensional class, while the $44.5S_{YsxC}$ particles exhibited two three-dimensional classes (Figure 2.4A and 2.4B). The percentage of the particle images assigned to the two classes is 44% and 56%, respectively. The most striking structural differences between the three immature particles and the mature 50S subunit structure were focused on the functional sites including A, P and E sites. The central protuberance (CP) and the region around the L7/L12 stalk also differed from the mature structure substantially (Figure 2.4A). These results suggested that the functional sites and the regions around the central protuberance are still in an immature state. In contrast, the core of these particles has already assembled into a close mature conformation (Figure 2.4C).

Consistent with the qMS results (Figure 2.2), the six r-proteins (uL16, bL27, bL28, bL33, bL35 and bL36) that were severely depleted in the immature particles were all missing their density in the cryo-EM maps of the $45S_{YphC}$ and $44.5S_{YsxC}$ particles. These depleted r-proteins are located at the base of the central protuberance (Figure 2.5A). In addition, interestingly, there are six more r-proteins that were presented as 100% occupancy according to the qMS results, however the density corresponding to these proteins was partially or completely missing from the cryo-EM maps (Figure 2.5B). These proteins are L6, L10, L11 in the L7/L12 stalk and L5, L31 and L18 in the central protuberance.
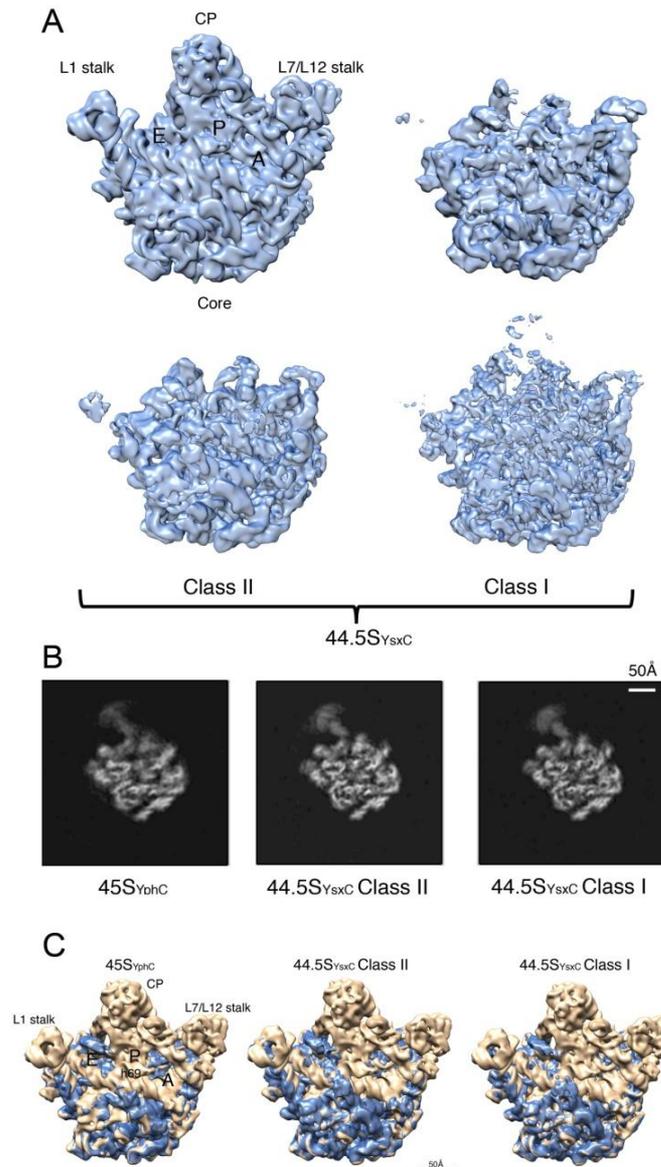
Figure 2.4 Cryo-EM maps of the 45S$_{YphC}$ and 44.5S$_{YsxC}$ ribosomal particles

(A) The cryo-EM maps show that the 45S$_{YphC}$ particle was present in one conformation, whereas the 44.5S$_{YsxC}$ particles exhibited two. The map for the mature 50S subunit was obtained by applying a low-pass filter at comparable resolution to the X-ray structure of the 50S subunit from B. subtilis (PDB ID: 3j9w) modeled from a 3.9 Å resolution cryo-EM map. Landmarks in the 50S subunit are labeled in the mature subunit. CP stands for central protuberance. (B) Cross-sections through the three-dimensional map of the 45S$_{YphC}$ particle and the two conformational states of the 44.5S$_{YsxC}$ particle. (C) Structural differences between the mature 50S subunit and the 45S$_{YphC}$ and 44.5S$_{YsxC}$ cryo-EM maps. The regions of missing density (in sand color) are shown in the figure by overlapping the map of the mature 50S subunit (from PDB ID: 3j9w) with the maps of the immature particles.
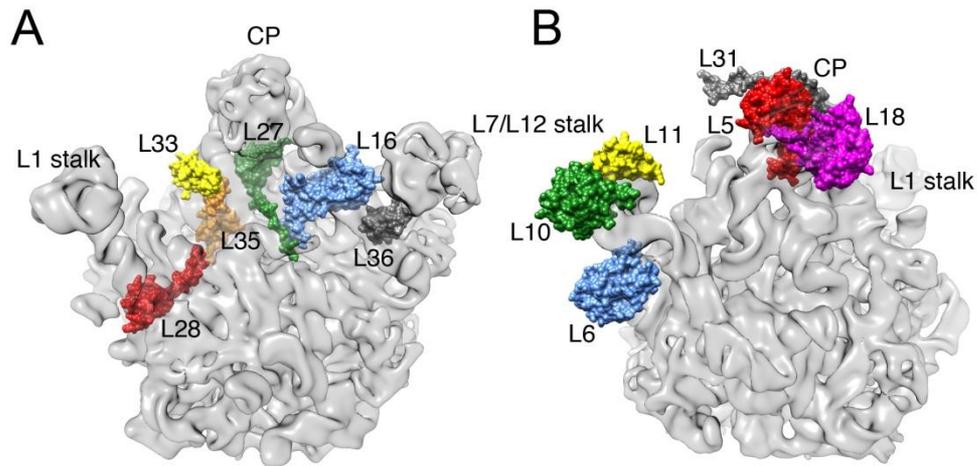
Figure 2.5 Depleted ribosomal proteins in the 45S$_{YphC}$ and 44.5S$_{YsxC}$ maps

(A) Location in the 50S subunit mature structure of the ribosomal proteins that were found severely depleted or absent in the 45S$_{YphC}$ and 44.5S$_{YsxC}$ particles. (B) The six ribosomal proteins displayed in the structure of the 50S subunit were found to be present at ~100% occupancy by qMS, but a corresponding density for these proteins was not observed in the 45S$_{YphC}$ and 44.5S$_{YsxC}$ maps.

The mean resolution of these cryo-EM classes is 6.5 Å (45S$_{YphC}$), 5.8 Å (44.5S$_{YsxC}$, class I) and 6.2 Å (44.5S$_{YsxC}$, class II), respectively (Figure 2.6A). The local resolution calculations, which estimate the different resolution in different areas within the structure, indicated that the resolution of the core of the immature particles is much higher than the mean resolution (Figure 2.6B and 2.6C). This is consistent with the features of the cryo-EM maps that show a close mature core with a clear fitting of α-helices and ß-sheets in the r-proteins (Figure 2.7A) and the pitch of the rRNA helices (Figure 2.7B). However, the central protuberance regions and the functional sites in the cryo-EM maps are still in an immature state, which refined to much lower resolutions compared to the mean resolution of the entire cryo-EM maps (Figure 2.6B and 2.6C). The variance in resolution within one cryo-EM map indicates

that the core of these particles is at a stable and already mature conformation, however, the relatively flexible nature of the central protuberance and functional sites indicate that they are still in an immature conformation.
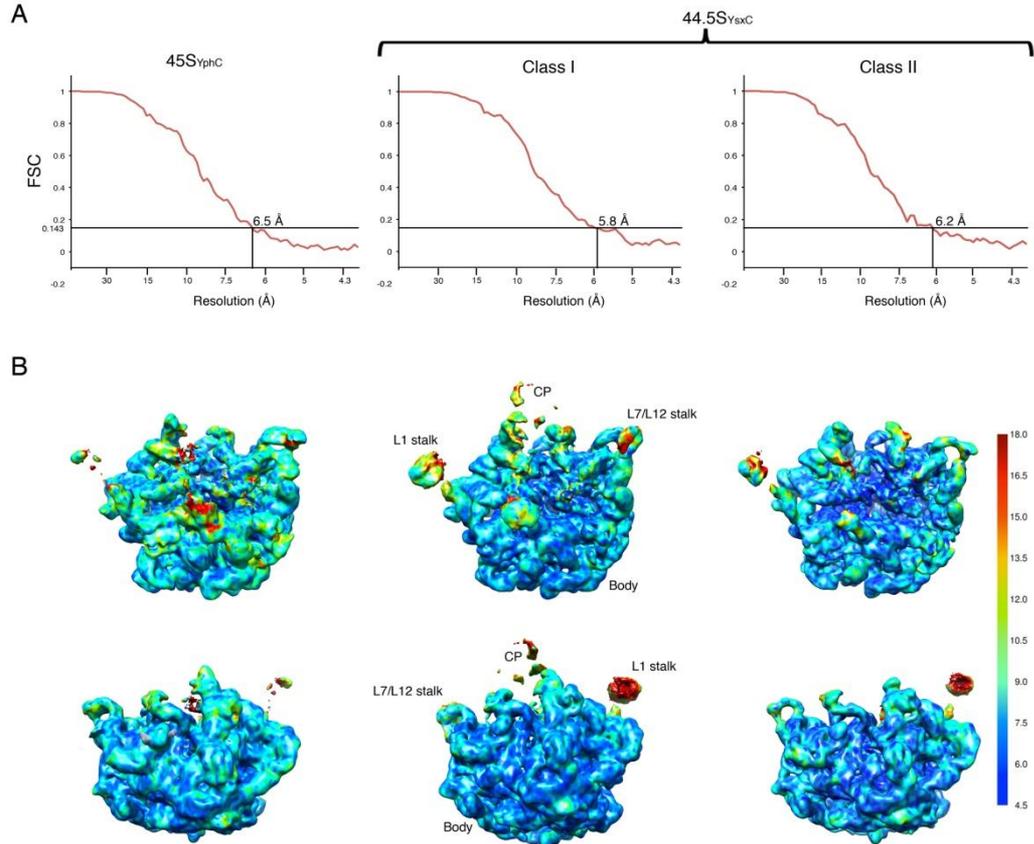
Figure 2.6 Local resolution of the cryo-EM structures of the 45S$_{YphC}$ and 44.5S$_{YsxC}$ immature particles

(A) Fourier Shell Correlation (FSC) curves of the cryo-EM structures of the 45S$_{YphC}$ and 44.5S$_{YsxC}$ immature particles obtained after a "gold standard" refinement and post-processing procedure as implemented in Relion 1.3. A FSC value of 0.143 was used to report the resolution. (B) Local resolution analysis of the maps shows that features in the body part are better resolved than in the CP and A, P and E site of the particles.

In summary, our structural study showed that the depletion of YphC or YsxC resulted in the accumulation of particles with the central protuberance, L7/L12 stalk and the

functional sites are still in an immature conformation, while the core of particles are formed to a mature conformation. Therefore, these results suggested that YphC and YsxC are involved in the maturation of the functional sites of 50S subunit, and this happens at the late stages of 50S assembly.
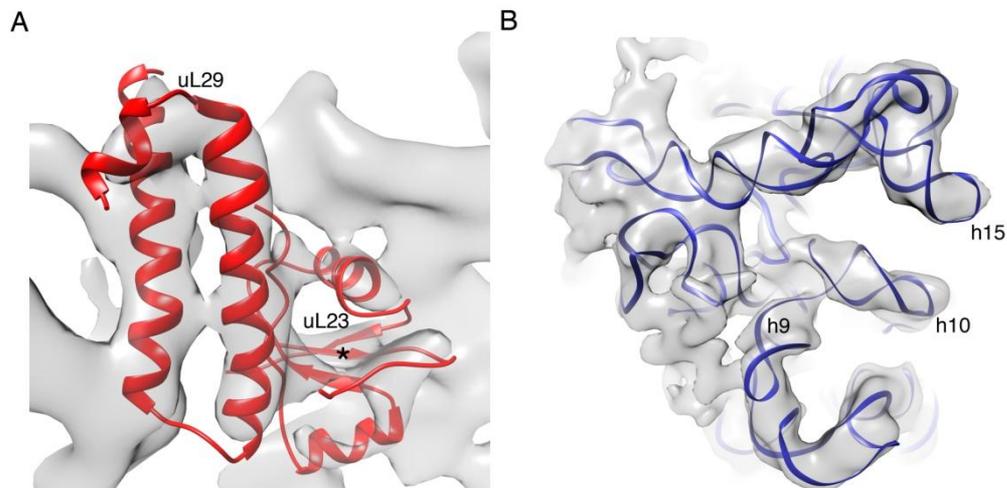
Figure 2.7 Density fitting of r-proteins and rRNA into the cryo-EM maps

Densities fitting of r-proteins and 23S rRNA into the 44.5S$_{YsxC}$ cryo-EM map (class 1) (A) α-helices in uL29 and β-sheets (asterisk) in uL23; (B) helices 9, 10 and 15 of the 23S rRNA.

2.4.3 Essential helices in the A, P and E sites of the 50S subunit adopt an immature state in the 45S$_{YphC}$ and 44.5S$_{YsxC}$ particles

The helices in the A, P and E site of the 50S subunit are a crucial helix group due to their location in the peptidyltransferase center (PTC), the functional sites of 50S subunit. The resolution we obtained from the cryo-EM maps for the 45S$_{YphC}$ and 44.5S$_{YsxC}$ particles by using the direct detector camera is sufficient to identify clearly the individual rRNA helices that differed from those of the mature 50S subunit.

One important group of helices that exhibited differently in the immature 45S$_{YphC}$ and

$44.5S_{YsxC}$ particles from the mature conformation were the helices located in the A, P and E sites. Densities for the helix 89 and helices 91-93, which are forming the A and P sites were not observed in the cryo-EM maps (Figure 2.8A). Similarly, helix 71 in the P site and helix 68, a major structural component of the E site, are also missing their correspondent density in the maps (Figure 2.8B). These helices are completely absent in the cryo-EM maps suggesting that they are still very flexible and may adopt multiple conformations within individual particle images.

Helix 69 is another functionally important motif located in the P site and structurally diverges from the mature 50S subunit in the immature particles (Figure 2.8B). This helix has special contacts to the decoding center of the 30S subunit and mediates the essential B2a intersubunit bridge between 30S and 50S subunits. A density corresponding to helix 69 in the $44.5S_{YsxC}$ class I and class II maps however was bent outward from the mature state by ~30 ° (Figure 2.8B, asterisks). This non-native conformation of the helix likely prevents the immature particles from association with the 30S subunit and joining into the translational pool.

Overall, these maps demonstrate that the $45S_{YphC}$ and $44.5S_{YsxC}$ particles have not structurally reached the mature state. The structural divergence in the A, P and E sites in the $45S_{YphC}$ and $44.5S_{YsxC}$ particles from the mature 50S subunits probably prevents these particles from prematurely engaging into translation. The structures also suggest that they are defective in tRNA binding and association with the 30S subunit.
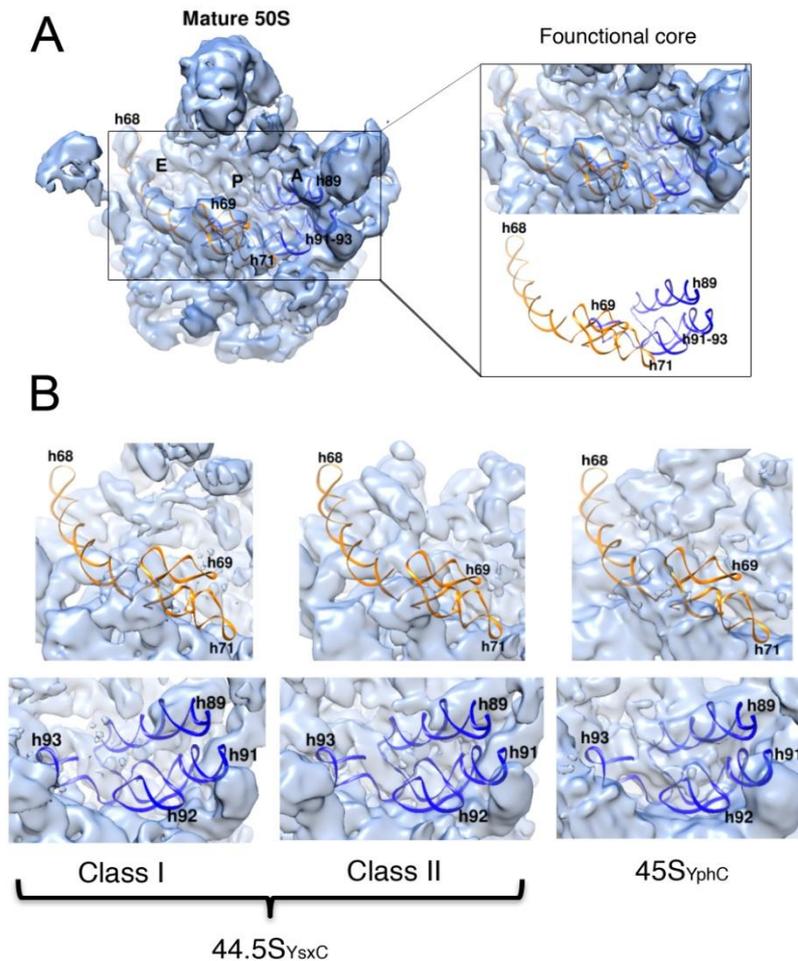
Figure 2.8 Structure of the functional core of the immature 45S$_{YphC}$ and 44.5S$_{YsxC}$ particles

(A) Zoomed view of functional core including helices 89 to 93 and helices 68 to 71 in the A, P and E sites of the 50S subunit. A ribbon representation of these helices (PDB ID: 3j9w) was fitted into the map of the mature 50S subunit. (B) A ribbon represents helices 68 to 71 and helices 89 to 93 was fitted into the immature 45S$_{YphC}$ and 44.5S$_{YsxC}$ density map. It shows the structural details of the A, P and E sites in the immature particles and how they are different from the mature 50S subunit.

2.4.4 Maturation dependencies between the central protuberance, helix 38 and the GTPase associated region

In addition to the RNA helices in the A, P and E sites, the 45S$_{YphC}$ and 44.5S$_{YsxC}$

particles also exhibited structural differences from the mature state in another functional motif including the central protuberance (5S rRNA, helices 80-88), helix 38 and the GTPase associated region.

The central protuberance region appeared at different assembly stages in the $45S_{YphC}$ and $44.5S_{YsxC}$ structures. However, none of the maps exhibited a fully assembled central protuberance (Figure 2.4A). This motif is composed of helices 80-88 from the domain V of the 23S rRNA. These helices form the bulk of the central protuberance, while the 5S rRNA forms its back. In the $45S_{YphC}$ and $44.5S_{YsxC}$ class II maps, densities for both helices 80-88 and 5S rRNA were missing. However, in the map of the $44.5S_{YsxC}$ class I, limited densities were apparent for these regions. Similarly, the amount of density representing helix 38 and helices 42-44 (GTPase associated region) were also featuring variable amounts of density among these structures (Figure 2.4A). To better understand the conformational changes of these important functional domains during the late stages of assembly, the focus classification was performed with three sets of cryo-EM maps from the $45S_{YphC}$, $44.5S_{YsxC}$ class I and $44.5S_{YsxC}$ class II, respectively. Briefly, we focused the signal only in the particle images corresponding to the central protuberance, helix 38, GTPase associated region and helices forming the A, P and E sites in the functional region during the 3D classification. Meanwhile, the signal from all ribosomal motifs that had already reached the mature state was masked out and subtracted from the entire particles images (Bai et al. 2015). Each data set generated three distinct classes with a resolution range of 8-10 Å for $44.5S_{YsxC}$ particle and 9-14 Å for the $45S_{YphC}$ particle (Figure 2.9).
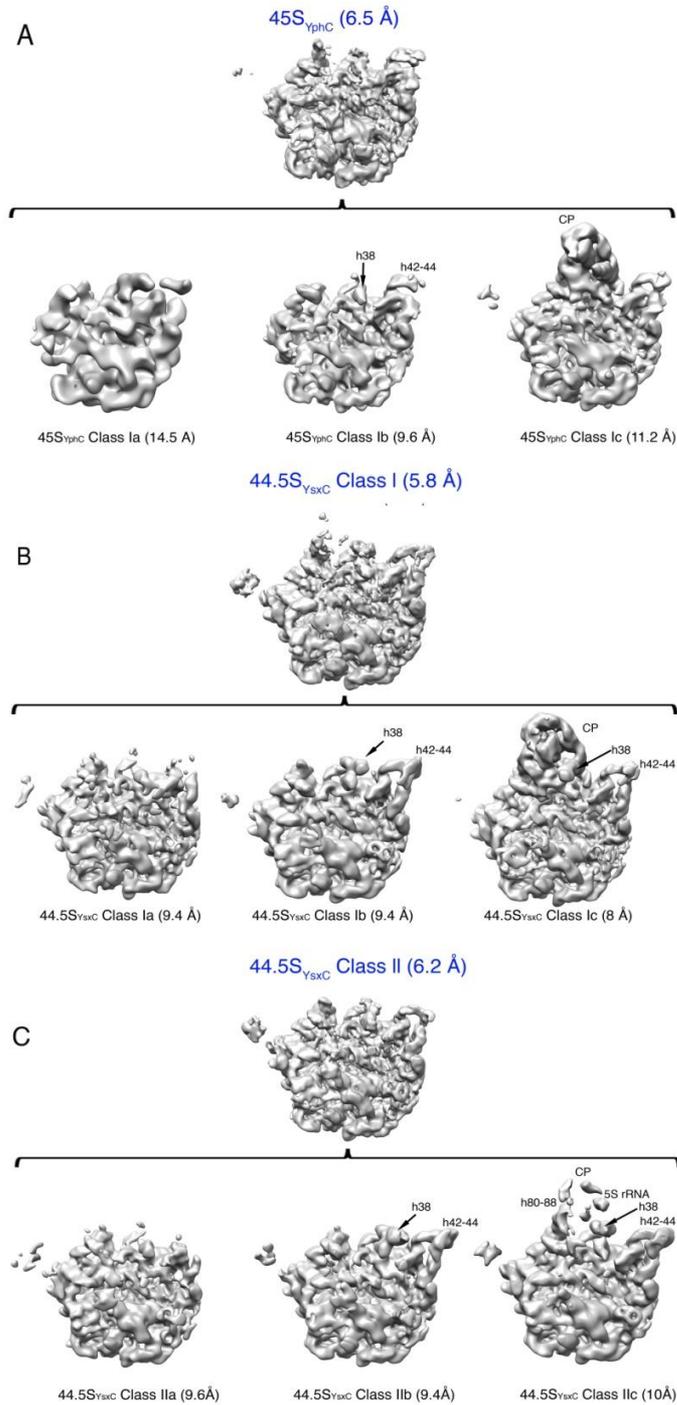
Figure 2.9 Three distinct sub-classes generated from each class of the 45S$_{YphC}$ and 44.5S$_{YsxC}$ cryo-EM maps

The focus classification rendered three sub-classes from each class of the 45S$_{YphC}$ and 44.5S$_{YsxC}$ cryo-EM maps. (A) The 9-14 Å classes from the 45S$_{YphC}$ map; (B) The 8-9 Å classes from the 44.5S$_{YsxC}$ class I; (C) The 9-10 Å classes from the 44.5S$_{YsxC}$ class II.

Comparison of the three sub-classes generated from the $45S_{YphC}$, $44.5S_{YsxC}$ class I and $44.5S_{YsxC}$ class II data sets revealed that the rRNA helices forming the A, P and E site were consistently not present in any of these derived maps. However, the other immature regions including the central protuberance, helix 38 and the GTPase associated region (helix 42-44) exhibited variable conformations (Figure 2.10). Among the three sub-classes from $44.5S_{YsxC}$ class I, the $44.5S_{YsxC}$ class Ia did not show density from any of these regions. The $44.5S_{YsxC}$ class Ib had density present for helix 42 and most of the helices 43-44. It also displayed density for the helix 38 with the direction of this helix deviated by ~30 ° from the mature conformation. Similarly, there was also no density for helices 80-88 or 5S rRNA suggesting that the central protuberance is still in an immature state. The most striking changes happened in the third class, which showed the central protuberance, helix 38 and the GTPase associated region adopting a conformation close to that of the mature subunit (Figure 2.10B).

The data set that generated the $44.5S_{YsxC}$ class II structure also produced three structures (Figure 2.10C). Two of them ($44.5S_{YsxC}$ class IIa and $44.5S_{YsxC}$ class IIb) were identical to the $44.5S_{YsxC}$ class Ia and $44.5S_{YsxC}$ class Ib described above. The third structure ($44.5S_{YsxC}$ class IIc) presents densities similar to the mature conformations for helix 38 and the helices 42-44 corresponding to the GTPase associated region. Additionally, it also featured with fragmented densities in the central protuberance region including helices 80-88 and the 5S rRNA. This sub-class probably represents an immature particle in the process of folding the central protuberance region. The RNA helices in the A, P and E site were consistently

41

unstructured in these three maps.
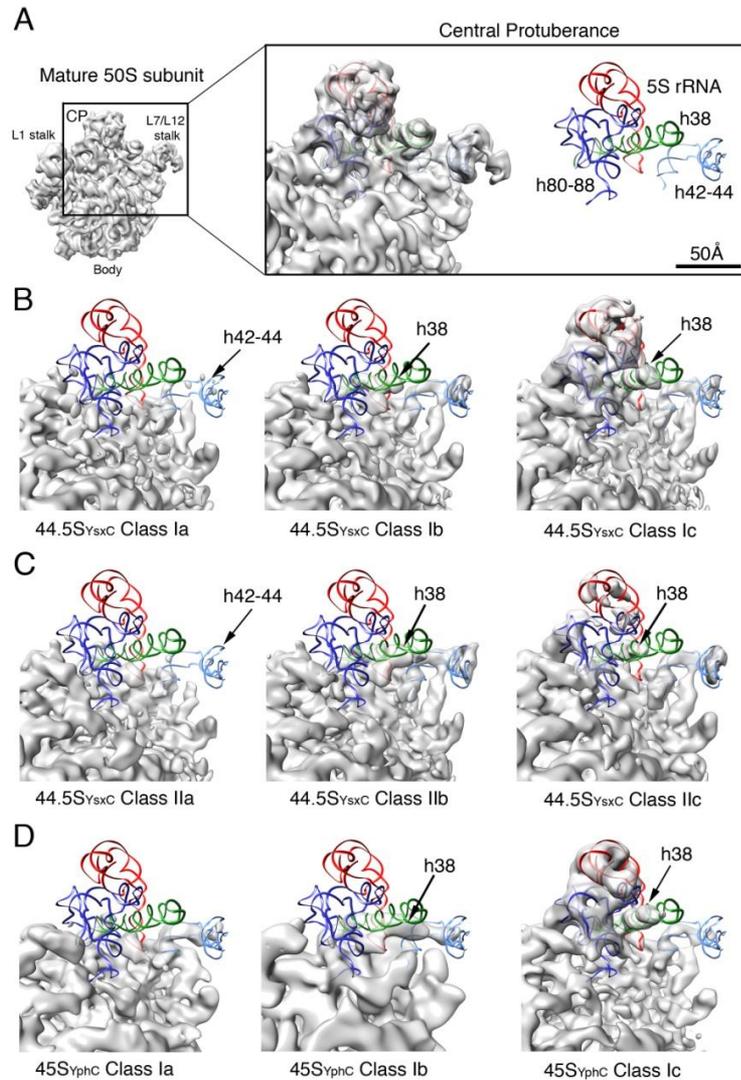
Figure 2.10 Structure of the central protuberance region of the 45S$_{YphC}$ and 44.5S$_{YsxC}$ immature particles

Zoomed views of the central protuberance, helix 38 and the GTPase associated region (helix 42-44) in the cryo-EM maps of the multiple classes obtained from focus classification for the 45S$_{YphC}$ (D) and 44.5S$_{YsxC}$ data sets (B and C). Panel (A) shows this region in the mature 50S subunit (PDB ID: 3j9w) and the three bottom panels (B), (C) and (D) in the immature particles. A ribbon representation of helices 80-88, helix 38 and helices 42-44 of the 23S rRNA and 5S rRNA were fitted to the cryo-EM maps. The densities representing helix 38 and helices 42-44 are indicated with black arrows in some of these classes.

Finally, the data set from the 45S$_{YphC}$ map also produced three sub-classes (Figure 2.10D). The first structure had no density at the central protuberance and helix 38, but with incipient densities for the GTPase associated region. The second structure was equivalent to the 44.5S$_{YsxC}$ class Ib with most of the density for the GTPase associated region and helix 38 density deviated by ~30° from the mature conformation. The third class presented a close mature conformation at these three regions. Same with all other structures, all of these sub-classes presented fully immature A, P and E sites.

In summary, these structural data obtained from further focus classification suggest that the central protuberance, helix 38, the GTPase associated region and the A, P and E functional sites of the 50S subunit fold sequentially and in a coordinated manner. It starts with the folding of helix 42 and followed by the maturation of helices 43-44 in the GTPase associated region. Simultaneously, helix 38 starts to extend but with an angle different from the mature structure. Subsequently, folding of helices 80-88 and 5S rRNA forming the central protuberance drags helix 38 towards its mature position. The densities corresponding to the RNA helices at the A, P and E sites were consistently missing in all the cryo-EM maps obtained from the 45S$_{YphC}$ and 44.5S$_{YsxC}$ particles, suggesting that these functional sites are the last regions to become mature.

2.5 DISCUSSION

2.5.1 High resolution cryo-EM structures provides precise testable models about YphC and YsxC function

The presented structures of immature 45S$_{YphC}$ and 44.5S$_{YsxC}$ particles that were obtained from a direct electron detector were refined to a resolution of around 5-6 Å. This resolution, to our knowledge so far, is the highest resolution cryo-EM structure

available for an immature bacterial ribosome intermediate. Compared with the previous moderate resolution cryo-EM maps on the assembly intermediates (Jomaa et al. 2011; Leong et al. 2013; Guo et al. 2013; Jomaa et al. 2014; Li et al. 2013), these high resolution structures allow the definition of individual rRNA helices in the $45S_{YphC}$ and $44.5S_{YsxC}$ particles that are still adopting an immature conformation. More importantly, these structures are making it possible to propose testable models regarding the function of YphC and YsxC protein in the assembly process. Previous study in the immature $45S_{RbgA}$ structure (Li et al. 2013) predicted that RbgA might play a role as an rRNA chaperone with the essential role of positioning helix 38 during the 50S subunit maturation. In our structures presented here, we found that the helix 38 is able to adopt its maturation in the absence of YphC or YsxC. In addition, our results also suggest that the assembly of the central protuberance and the GTPase associated center is not dependent on these two proteins. However, the RNA helices forming the A, P and E sites were consistently unstructured in the $45S_{YphC}$ and $44.5S_{YsxC}$ structures, suggesting that the essential role of YphC and YsxC might be more related to the remodeling of the RNA helices in the functional core of the particles. A recent high resolution structure of EngA, the *E. coli* homologue of YphC, in complex with the mature 50S subunit showed that the EngA protein binds to the peptidyltransferase center deeply into the tRNA passage at the P and E site (Zhang et al. 2014). Consistently, binding of EngA leads to significant rearrangement of the rRNA helices including helices 68-71 that we found is fully unstructured in our immature $45S_{YphC}$ and 44.5_{YsxC} particles.

The determinations of different sub-classes from the high resolution cryo-EM maps

using focus classification also provide us an important insight into the deficiencies found in the protein complement of these particles. In our structures, the cryo-EM maps for the 44.5S$_{YsxC}$ class Ib, 44.5S$_{YsxC}$ class IIb and the 45S$_{YphC}$ class Ib particles, the position of helix 38 deviated by ~30° from the mature conformation and completely blocks the binding site of uL16. Only when the central protuberance has formed, helix 38 is dragged to its mature position thus making the uL16 binding site accessible for the almost complete absence of uL16 found in the 44.5S$_{YsxC}$ and 45S$_{YphC}$ particles. This feature of helix 38 and the steric blockage of the uL16 binding site is shared with the previous described cryo-EM map of the 45S$_{RbgA}$ particles as well (Jomaa et al. 2014; Li et al. 2013).

In addition, certain features in the cryo-EM maps of 45S$_{YphC}$ and 44.5S$_{YsxC}$ particles are in agreement with the previous described model that the ribosome assembly is following multiple parallel pathways (Shajani et al. 2011; Woodson 2008; Mulder 2010). In particular, in our structures we found that there are no densities corresponding to the rRNA helices in the A, P and E sites of these immature particles, which suggests that the rRNA helices in this domain probably adopt a large number of conformational changes resulted from the multiple assembly pathways occurring in the cell. In addition, in our structures we found that the folding of the 23S rRNA is not proceeding in a 5' to 3' manner, as domain IV and V are still in an immature state, whereas all other domains including I, II, III and VI already reached the mature stage (Figure 2.11). This was also described in the previous 45S$_{RbgA}$ particles (Jomaa et al. 2014; Li et al. 2013). This 23S rRNA folding manner is different from the assembly of the 16S rRNA in the forming of 30S subunit where the rRNA folding follows a 5' to 3'

transcriptional order (Mulder 2010).
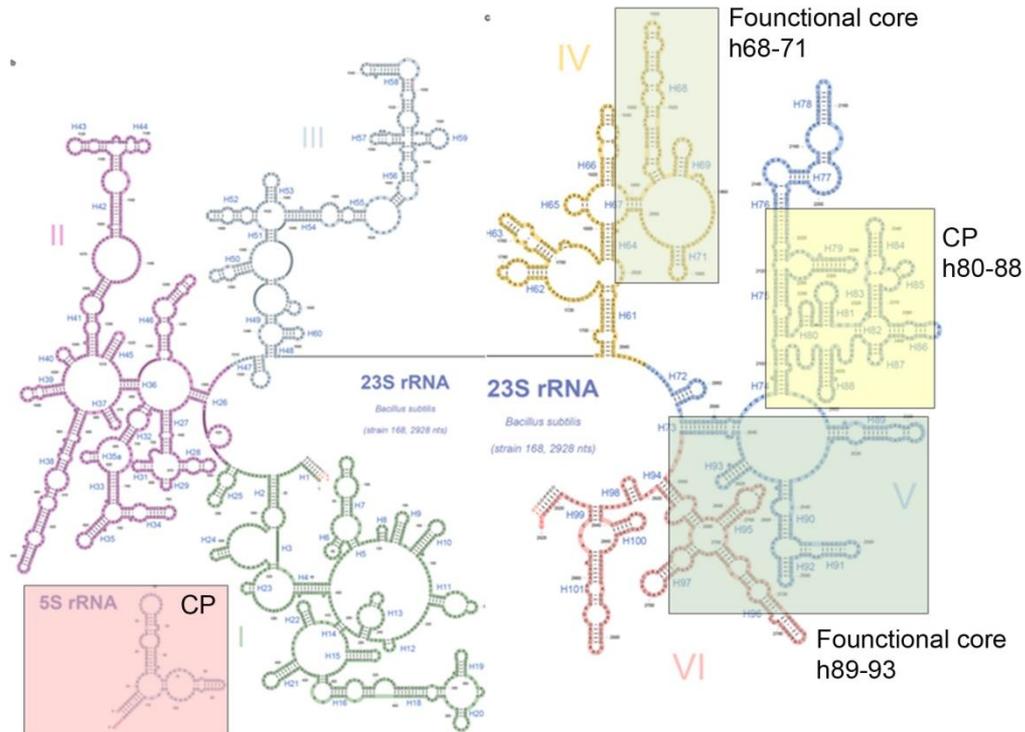
Figure 2.11 The overview of secondary structure of the 23S rRNA

The overview of the secondary structure shows that domains I, II, III and VI representing the core of the 50S subunit have formed into a mature conformation, whereas domains IV and V are still not fully structured.

2.5.2 Cryo-EM allows for direct visualization of the ribosomal assembly process

The cryo-EM maps of the $45S_{YphC}$ and $44.5S_{YsxC}$ particles also describe the discrete stages leading to the assembly of important functional sites of the 50S subunit, including the GTPase associated region, central protuberance and helix 38. It starts with the helix 42 reaching its mature conformation, followed by the folding of helices 43-44 to form the complete rRNA helices in GTPase associated region. Simultaneously, the helix 38 starts to grow but with a ~30° angle differed from the

mature conformation. As helices 80-88 and 5S rRNA start to form the central protuberance, helix 38 is dragged to its mature conformation. These results suggest that the conformational variability in this region among different particle images may not be as diverse as has been observed for the rRNA helices at the A, P and E sites.

These structures also suggest that the central protuberance in the bacterial ribosome assembled differently from the eukaryotic ribosome. A recent study of the energetically favored assembly intermediates of eukaryotic 60S ribosomal subunits indicated that there is a dramatic conformational rearrangement of the central protuberance (Leidig et al. 2014). Compared to its mature state, the 5S rRNA exhibits an essentially unchanged fold with the entire molecule rotated by 180°. This conformational change is stabilized by the assembly factors Rsa4 and Nog1, through which the 5S rRNA rotates to its native position during the assembly process. However, in the late stages of assembly of bacterial 50S subunit that we studied here, the central protuberance did not exhibit a conformational rearrangement in any of the assembly intermediates, which indicates the diverse assembly mechanisms between bacterial and eukaryotic ribosomes.

2.6 MATERIALS AND METHODS

2.6.1 Purification of mature 50S subunits and immature $45S_{YphC}$, $45S_{RbgA}$ and $44.5S_{YsxC}$ particles

The mature 50S subunits and immature $45S_{YphC}$, $44.5S_{YsxC}$ and $45S_{RbgA}$ particles were purified from IF2-depleted (RB419), YphC-depleted (RB290), YsxC-depleted (RB260) and RbgA-depleted (RB301) *B. subtilis* strains, respectively. As previously described, in these cells the transcription of *infB* (encoding IF2), *ylqF* (encoding

RbgA), *yphc* (encoding YphC) and *ysxc* (encoding YsxC) is under the control of an isopropylbeta-D-thiogalactopyranoside (IPTG) inducible P*spank* promoter (Schaefer et al. 2006; Uicker et al. 2006). The mature 50S subunit and immature $45S_{RbgA}$ particles were purified as described previously (Jomaa et al. 2014). The $44.5S_{YsxC}$ and $45S_{YphC}$ particles were purified in a similar procedure to the $45S_{RbgA}$ particles. Briefly, strains were first grown at 37 ℃ on LB plates containg 5mg/ml chloramphenicol and 1mM IPTG. Then cells were scraped off the plate and resuspended in 1ml of LB media, which was then inoculated in 200ml pre-warmed LB media without IPTG and cultures were grown from an OD_{600} of 0.1 to an OD_{600} of 0.5. The doubling time was calculated before the dilution into 1L of pre-warmed LB media and harvested when the doubling time reached 120-150 min. Cells were spun down through centrifugation at 8500g for 15 min, and the cell pellet was resuspended in buffer A (20mM Tris-HCl at pH 7.5, 15mM magnesium acetate, 100mM NH4Cl, 0.5mM EDTA, and 3 mM 2-mercaptoethanol). The cell suspension was passed through a French press at 20,000 lb/in^2 pressure three consecutive times to lyse the cells. The lysate was spun at 30,000g for 40min to clear cell debris using a Beckman 70Ti rotor. The crude ribosomes were obtained as previously described (Jomaa et al. 2014) except the buffers contained 15mM magnesium acetate. To obtain the 50S or intermediate fractions, the crude ribosome pellet was resuspended in buffer E (10mM Tris−HCl at pH 7.5, 15mM magnesium acetate, 60mM NH4Cl and 3mM 2-mercaptoethanol). About 50-60 A_{260} units of the subunit suspension were loaded onto a 32 ml 10-30% (wt/vol) sucrose gradient made in buffer E and centrifuged at 43000 g for 16h using a Beckman SW32 rotor. Gradients were fractionated using an AKTA prime purification

system (GE Healthcare) and a Brandel fractionator apparatus. The elution profiles were monitored by UV absorbance at A_{260}, and the fractions corresponding to the 50S subunits and intermediates were collected and pooled before being concentrated and stored at -80°C until further use.

2.6.2 Quantitative mass spectrometry

Samples were purified as described above with either 500 mM (high salt) or 150 mM (low salt) NH_4Cl present during the sucrose cushion centrifugation. For the low salt samples, 10 pmols of each sample (50S, $45S_{RbgA}$, $45S_{YphC}$, or $44.5S_{YxsC}$) were spiked with 10 pmols of 70S particles purified from wild-type cells grown in ^{15}N-labeled media. Samples were then precipitated, reduced, alkylated, and digested to tryptic peptides as described previously (Jomaa et al. 2014). Peptides were injected onto a C18 nanoflex column (Eksigent), and eluted using a 120 minute 5%-45% convex acetonitrile gradient.

Data were initially collected in a data-dependent acquisition mode with a cycle consisting of a 200 ms MS^1 scan followed by 30-100 ms MS^2 scans, selecting precursors exceeding 125 counts per second. Precursors were excluded for 12 seconds after their being selected twice. Datasets were searched against the *B. subtilis* proteome using Mascot. Search results were combined to generate a spectral library using Skyline (Brendan MacLean 2010).

Data were also collected in a data-independent SWATH acquisition mode (Gillet et al. 2012) using a 250 ms MS^1 scan followed by 32 MS^2 scans ranging 400-1200 Th, each 25 Th in width. Using the spectral library noted above, product ion chromatograms from the SWATH acquisition were extracted for ^{14}N and ^{15}N species using Skyline.

Peptides and transitions were filtered to eliminate spectral interference and poorly ionized precursors. $^{14}N/^{15}N$ abundance ratios were calculated for each transition. Protein abundance was calculated as the median value of this ratio, normalized to the median value observed for protein L20, which was expected to be bound stoichiometrically.

Samples undergoing the high salt wash were spiked with a mixture of ^{14}N- and ^{15}N- labeled 70S particles, and peptide abundances were determined from MS^1 data obtained on an Agilent G1969A ESI-TOF mass spectrometer as previous described (Gulati et al. 2014). Protein abundance was again calculated as the median $^{14}N/^{15}N$ ratio, normalized to that of protein L20.

2.6.3 Cryo-electron microscopy and image processing

The purified $45S_{YphC}$ and $44.5S_{YsxC}$ particles were diluted to a concentration of 40-50 nM in buffer E (10mM Tris–HCl at pH 7.5, 10mM magnesium acetate, 60mM NH_4Cl and 3mM ß-mercaptoethanol). Approximately 3.6 µl of the diluted sample was applied to the holey carbon grids (c-flat CF-2/2-2C-T) with an additional layer of continuous thin carbon (5-10nm). Before the sample was applied, grids were glow discharged in air at 5 mA for 15 seconds. Samples were then vitrificated in a Vitrobot (FEI) by blotting the grids twice, 15 seconds each time and with an offset of -1.5 before they were plunged into liquid ethane.

Grids were loaded in a Gatan 626 single tilt cryo-holder and introduced into a FEI Tecnai F20 electron microscope operated at 200 kV and equipped with a Gatan K2 Summit direct detector device camera. This detector was used in counting movie mode with five electrons per pixel per second for 15 seconds exposures and 0.5

seconds per frame. This method produced movies consisting of 30 frames with an exposure rate of ~1 $e^-/Å^2$. Movies were collected with a defocus range of 1 to 2.5 μ m and a nominal magnification of 25,000x, which produced images with a calibrated pixel size of 1.45 Å.

The 30 frames in each movie were aligned using the program alignframesleastsquares_list and averaged into one single micrograph with the shiftframes_list program (Rubinstein & Brubaker 2015). These programs are available from (https://sites.google.com/site/rubinsteingroup/home). These micrographs were used to estimate the parameters of the contrast transfer function using CTFFIND3 (Mindell & Grigorieff 2003) and also to determine the coordinates for particles in the frames of the movies. Then the particles from each averaged micrograph were picked using the autopicking procedure in Relion (Scheres 2012). The coordinates of particles in the frames were used to extract candidate particle images from the 30 unaligned frames in the movie. The motion of the individual particles in the frames was tracked and corrected using alignparts_lmbfgs algorithm (Rubinstein & Brubaker 2015). This procedure produced one stack of particle images fully corrected from beam-induced motion from the first 20 frames of each movie. Therefore, the total accumulated dose to produce these particle images was 20 electrons per square angstrom. The initial number of particles in the data sets for determining the $45S_{YphC}$ and $44.5S_{YsxC}$ structures was 105,302 and 91,724, respectively. These particles were then subjected to two- and three-dimensional classification using Relion (Scheres 2012). In the case of the $45S_{YphC}$ particle three-dimensional classes were built from 101,264 particle images and the 45,369 particles assigned to one class were used to

build the final consensus 3D map. A similar approach was followed with the 44.5S$_{YsxC}$ structure where three-dimensional classes were produced from 87,684 particle images. The final three-dimensional consensus map for the class I conformation was built from the 36,033 particle images that were assigned to one of the 3D classes obtained in the 3D classification and the class II conformation was produced from the 46,430 particle images from another of the 3D classes.

After the 3D refinement, the particles contributing to the final reconstructions were further subjected to the focus classification, an approach previously described to calculate the variance within one structure (Bai et al. 2015), with subtraction of the residual signal using Relion. The mask for focus classification on the central protuberance, helix 38, GTPase associated region and A, P, and E functional sites was generated by converting the atomic model of the 50S subunit (PDB ID 3j9w) into a density map after the following motifs were removed from the atomic model: 5S rRNA, helices from the 23S rRNA including h80-88 (nt 2280-2420), h38 (nt 890-980), h42-44 (nt 1080-1160), h89-93 (nt 2480-2630), h68-71 (nt 1870-2000), h76-78 (2140-2200), and ribosomal proteins uL16, bL27, uL6, bL33, bL35, bL28, bL36, bL31, uL18, uL5, uL30, uL15, uL10 and uL11. This density map was used to create a soft-edged mask and to also subtract the signal of the mature motifs in the experimental particles. The newly created stacks of particles after signal subtraction and the mask were used as input for the focus 3D classification run. During the classification step, we kept all orientations fixed at the values determined in the refinement of the consensus maps. Each data set rendered three distinct classes that were subjected to a

separate 3D auto-refinement using the cryo-EM structure of the 50S subunit from *B. subtilis* (PDB ID 3j9w) low pass filtered to 50 Å.

Prior to visualization, sharpening of the cryo-EM maps was done by applying a negative B-factor estimated using automated procedures (Rosenthal & Henderson 2003). Relion processes were calculated using the SciNet cluster (Loken et al. 2010) and a VMWare-based Ubuntu linux server with 32 processors / 256 GB RAM within the McMaster Service Lab and Repository (MSLR) computing cluster. We used the program ResMap (Kucukelbir et al. 2014) to estimate the local resolution of the structures. The UCSF Chimera program (Pettersen et al. 2004) was used for the visualization of cryo-EM maps and render figures. To identify the rRNA helices in the $45S_{YphC}$ and $44.5S_{YsxC}$ structures that were different from the mature 50S subunit, the atomic model of the *B. subtilis* 50S subunit (PDB ID 3J9W) was docked into the cryo-EM maps first as a rigid body using Chimera, and then the fitting was optimized by Molecular Dynamics Flexible Fitting (MDFF) (Leonardo G. Trabuco 2008).

The $44.5S_{YsxC}$ class I and class II cryo-EM maps have been assigned the EMDB IDs 8274 and 8275, respectively. The EMDB ID for the $45S_{YphC}$ cryo-EM map is 8276.

CHAPTER 3. Nature and binding properties of the immature ribosomal particles accumulating in bacterial strains depleted of RbgA, YphC or YsxC

3.1 Author's Preface

The work in Chapter 3 is part of a published paper in the journal *Nucleic Acids Research*, and is modified in this chapter. To determine whether the accumulated intermediates described in Chapter 2 are the actual substrates for the assembly factors, we characterized the nature and binding properties of the immature ribosomal particles accumulating in the RbgA, YphC or YsxC depletion *B. subtilis* strains. I have performed all of the binding assays and the GTPase assays, including protein and ribosome purifications for these assays. Dr. Nikhil Jain, a postdoctoral fellow working with Dr. Robert Britton, our close collaborator, purified part of the immature ribosome particles necessary for the GTPase assays, and also helped to design the experiments. Dr. Joaquin Ortega and I analyzed the data and wrote the manuscript.

The full citation is as follows:

Xiaodan Ni, Joseph H. Davis, Nikhil Jain, Aida Razi, Samir Benlekbir, Andrew G. McArthur, John L. Rubinstein, Robert A. Britton, James R. Williamson and Joaquin Ortega. (2016) YphC and YsxC GTPases assist the maturation of the central protuberance, GTPase associated region and functional core of the 50S ribosomal subunit. *Nucleic Acids Research*. 9140 (2016), 1–14.

3.2 ABSTRACT

The cryo-EM structures of the immature ribosomal particles that accumulate under depletion conditions of RbgA, YphC or YsxC revealed that these assembly factors are involved in facilitating the maturation of the main functional sites of the 50S subunit. However, whether these immature intermediates are the actual substrates for the assembly factors is still uncertain. To characterize these immature particles, we performed filtration assays to test the binding of assembly factors to the ribosomal particles. The RbgA, YphC and YsxC proteins can individually bind to both the mature 50S subunit and the immature $45S_{RbgA}$, $45S_{YphC}$ and $44.5S_{YsxC}$ particles. All three factors can also bind to both mature and immature 50S particles simultaneously. These results reconfirmed that the 50S subunit assembles following multiple parallel pathways. In addition, enzymatic assays indicated that the intrinsic GTPase activity of the three assembly factors can be maximally stimulated by the mature 50S subunit, and partially stimulated by the three immature particles. This specific functional interplay between assembly factors and ribosomal particles suggested that the binding of factors to the mature 50S subunit and the immature ribosomal particles is specific. Overall, the nature and binding properties of the immature ribosomal particles revealed that they either constitute the actual substrates for the assembly factors or their conformations have not diverged significantly from the actual substrates in which the assembly factors RbgA, YphC and YsxC still bind to them.

3.3 INTRODUCTION

RbgA, YphC and YsxC are three essential GTPases that participate in the late stages of the assembly of the 50S subunits. Depletion of each of them generates the immature ribosomal intermediates designated as $45S_{RbgA}$, $45S_{YphC}$ and $44.5S_{YsxC}$ respectively (Schaefer et al. 2006). Characterization of the three late 50S assembly intermediates revealed that the immature 50S particles that accumulated in these depleted strains are structurally similar and the main function of RbgA, YphC and YsxC is to facilitate the maturation of the functional sites of the subunit including central protuberance, GTPase associated region and A, P, E sites (Ni et al. 2016; Jomaa et al. 2014; Li et al. 2013). However, the nature of these immature 50S particles that accumulated upon depletion of RbgA, YphC or YsxC and whether they represent the true on-pathway assembly intermediates are still unknown.

Recent studies in 30S ribosome assembly found that assembly factors YjeQ, RbfA, Era and RimM exhibited weak binding affinity to the immature 30S particles accumulated. Therefore, it was proposed that these immature 30S particles probably represented the assembly intermediates at later stages with more energetically stability than the actual substrates for the assembly factors (Thurlow et al. 2016). Thus, it would be interesting to find out whether in the assembly of 50S subunit the same scenario happens. It is also unclear whether RbgA, YphC and YsxC can bind to these assembling immature 50S particles.

To address these questions, we purified the mature 50S subunits and the immature $45S_{RbgA}$, $45S_{YphC}$ and $44.5S_{YsxC}$ particles from RbgA, YphC or YsxC depleted *B. subtilis* strains. The pulse-labeling experiments in the previous study have determined

that the $45S_{RbgA}$ particles that accumulate in cells under RbgA depletion conditions are competent for maturation and progress into functional 70S particles (Jomaa et al. 2014). This finding was important as it provided reassurance that the immature particles we generated in these depletion conditions do not represent dead-end products of the reaction and thus, they render physiologically relevant information about the function of assembly factors. Subsequently, we tested the binding of each factor to the ribosome particles. We found that RbgA, YphC and YsxC can individually bind to each of the immature particles as well as to the mature 50S subunit. This binding is specific as it triggers a stimulation of the GTPase activity of the assembly factors. However, a hierarchy of binding similar to that found for bona fide r-proteins was not apparent for the binding of these factors. This finding is consistent with recent kinetic work revealing that assembly of the ribosome occurs through multiple parallel pathways, which introduce the necessary flexibility and redundancy to make ribosome assembly an extremely robust and efficient process. These results suggest that different from 30S assembly, the assembly intermediates that accumulate in the absence of RbgA, YphC or YsxC are thermodynamically stable. They either constitute the actual substrates for the assembly factors or their conformations have not diverged significantly from that present in the actual substrate, so that RbgA, YphC or YsxC still bind to them.

3.4 RESULTS

3.4.1 YphC, YsxC and RbgA are capable to bind to both the mature and immature ribosomal particles

As described in Chapter 2, the immature $45S_{YphC}$ and $44.5S_{YsxC}$ particles exhibit

similar conformations. In addition, the $45S_{RbgA}$ particle structurally resembles these two immature particles (Jomaa et al. 2014; Li et al. 2013). Therefore, we tested whether these assembly factors including RbgA, YphC and YsxC have the ability to bind to all these immature particles and to the mature 50S subunits as well.

We used filtration assays to test the binding of YphC, YsxC and RbgA to the mature 50S subunit and the three immature ribosomal particles. In these assays, a mixture of the assembly factor with the mature or immature ribosomal particles was incubated at 37℃ for 1hr in the presence of 1 mM GMPPNP. Reactions were then centrifuged through a membrane, the assembly factors bound to the ribosome particles and the free ribosomal particles are retained as bound fraction, while molecules not bound to the ribosome particles passed through the membrane as flow through. Both bound fraction and flow through were analyzed using SDS-PAGE (Figure 3.1).

The results showed that none of the three proteins were retained by the filter membrane in the absence of ribosome particles. Although the quantitative nature of these assays is limited, we could observe the YsxC protein exhibited similar binding density to both mature and immature particles when reacted with the ribosomal particles as shown in Figure 3.1C. However, the RbgA or YphC protein has a larger amount of bound fraction for the reactions containing the immature particles than the reactions containing the mature 50S subunit (Figure 3.1A and 3.1B).
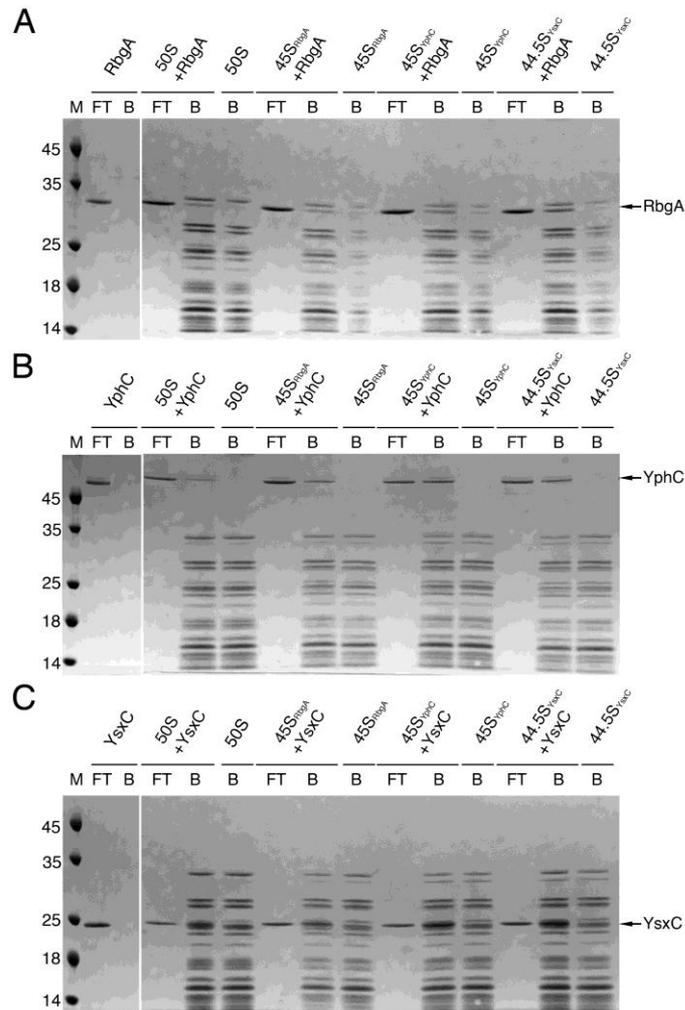
Figure 3.1 Binding of RbgA, YphC and YsxC to the mature 50S subunit and the immature 45S$_{RbgA}$, 45S$_{YphC}$ and 44.5S$_{YsxC}$ particles

(A) The binding of RbgA to the mature 50S subunit and the immature 45S$_{RbgA}$, 45S$_{YphC}$ and 44.5S$_{YsxC}$ particles. A Coomassie blue stained 4-12% bis-tris polyacrylamide gel shows the content of the flow-through (FT) and bound (B) fractions of the filtration assay. Reactions contained RbgA alone or a mixture of ribosomal particles with a five-fold molar excess of RbgA. The molecular weight (M) is in kDa. Similar filtration assays were performed to test the binding of YphC (B) or YsxC (C) to the mature 50S subunit and the immature particles.

In addition, in order to determine the effects of the nucleotides on the binding affinity

of RbgA, YphC and YsxC to the ribosomal particles, the identical reactions were

performed in the presence of 1mM GTP or 1mM GDP (Figure 3.2). In the case of

RbgA and YphC we found that the binding observed in the presence of these two nucleotides was weaker than in the presence of GMPPNP (Figure 3.2A and 3.2B). This observation for RbgA is consistent with previous studies (Uicker et al. 2006; Achila et al. 2012; Matsuo et al. 2006). However, the YsxC protein exhibited similar binding affinity to the ribosomal particles with the three nucleotides (Figure 3.2C).
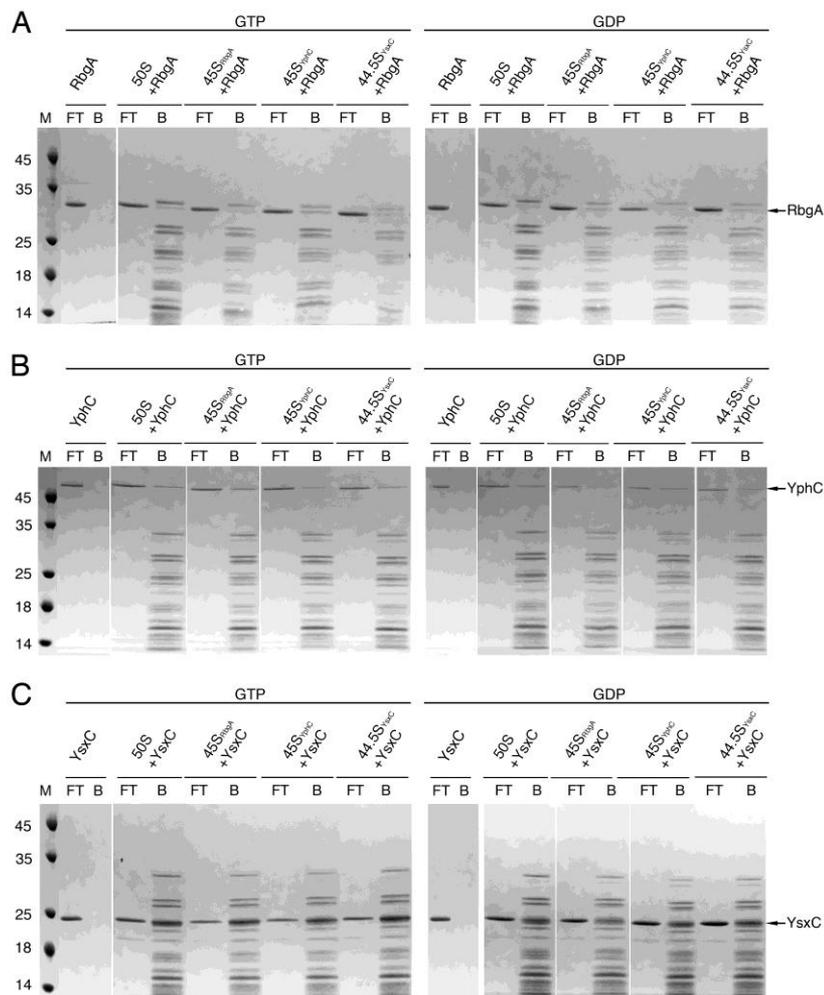
Figure 3.2 Binding of RbgA, YphC and YsxC to the mature 50S subunit and the immature particles in the presence of different nucleotides

Filtration assays were used to analyze the binding of RbgA (A), YphC (B) and YsxC (C) to the mature 50S subunits and the immature particles. The reactions were incubated in the presence of 1mM GTP or GDP instead of GMPPNP.

We also performed the binding assay to test whether multiple assembly factors could bind the ribosomal particles simultaneously. To perform this filtration assay, each of the ribosomal particles was incubated with 5-fold molar excess of all three assembly factors in the presence of 1mM GMPPNP. The assays suggested that multiple assembly factors bound to each of the immature particles and the mature 50S subunit in an approximately stoichiometric amount (Figure 3.3). Although we did not observe a strict hierarchy of binding of these factors to the ribosomal particles, the experiment suggested that the simultaneous binding of more than one of the assembly factors to the mature 50S subunit and the immature particles is possible.
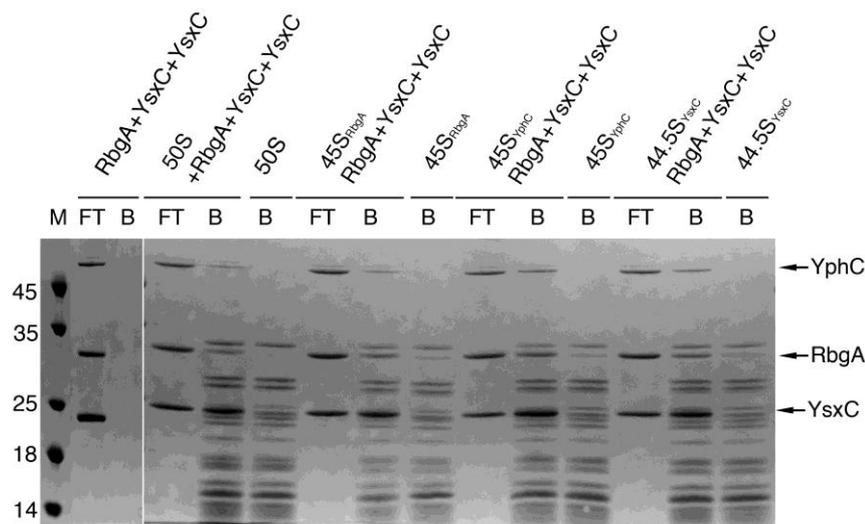
Figure 3.3 Binding of multiple assembly factors to the ribosomal particles

Filtration assay testing binding of multiple assembly factors including RbgA, YphC and YsxC to the mature 50S subunits and the immature $45S_{RbgA}$, $45S_{YphC}$ and $44.5S_{YsxC}$ particles. Five-fold molar excess of each factor with respect to the ribosomal particle was added to the reaction in this assay.

3.4.2 The GTPase activities of YphC, YsxC and RbgA are stimulated by the ribosomal particles

As previously described, RbgA exhibits low intrinsic GTPase activity. The activity can be dramatically stimulated upon the interaction with the mature 50S subunit (Matsuo et al. 2006; Achila et al. 2012). YphC and YsxC are the other two GTPase factors that also involved in the 50S ribosomal subunit assembly. We hypothesized that YphC and YsxC may exhibit similar behavior to RbgA, and thus we measured the GTPase activity of the factor itself and the stimulation of GTPase activities in the presence of ribosomal particles and also to test whether the binding we observed in the filtration assays was specific. RbgA was also included in these assays.

We first evaluated the intrinsic GTPase activity of YphC, YsxC and RbgA proteins by using the malachite green assay. The three proteins showed a low intrinsic GTPase activity. RbgA and YsxC had a k_{cat} of ~10 h^{-1} and 5 h^{-1} respectively (Figure 3.4A, 3.4C and Table 3.1), while YphC had a much higher GTP hydrolysis rate with a k_{cat} of ~82 h^{-1} (Figure 3.4B and Table 3.1) considering the YphC protein contains two GTPase domains. All three proteins exhibited an apparent K_M in the micromolar range indicating the low affinity for GTP binding (Table 3.1).

To test whether the GTP catalytic rate can be stimulated by the ribosomal particles, the GTPase activity of YphC, YsxC and RbgA was then tested in the presence of the mature 50S subunit and the three immature particles ($45S_{YphC}$, $44.5S_{YsxC}$ and $45S_{RbgA}$). Possible non-specific GTP hydrolysis by each ribosomal particle itself was subtracted at each GTP concentration. These assays showed that there was a significantly higher stimulation of the catalytic rate of the enzyme against GTP for all three GTPase proteins in the presence of the mature 50S subunit (Table 3.1, increase in k_{cat}). The three immature particles also stimulated the catalytic rate of assembly factors but to a

lesser extent. Overall, there was a much higher (~26-38 fold) stimulation in YsxC protein activity in the presence of ribosomal particles, while a lower (~2-11 fold) stimulation in the case of YphC and RbgA.
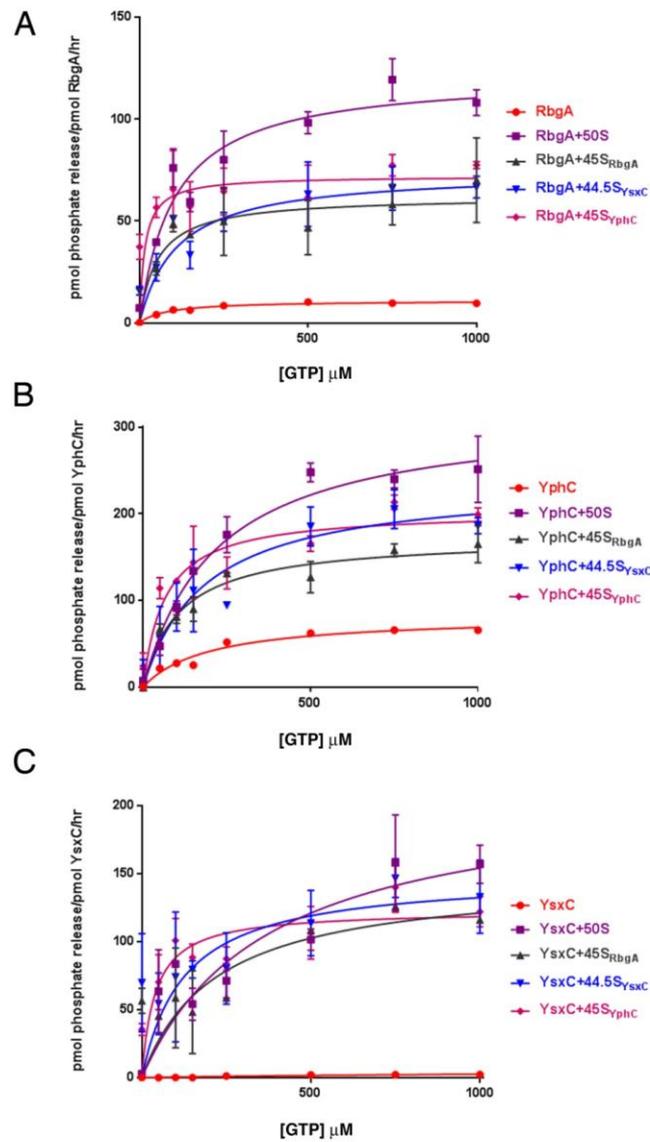
Figure 3.4 Stimulation of the GTPase activity of RbgA, YphC and YsxC by the mature 50S subunit and the immature 45S$_{RbgA}$, 45S$_{YphC}$ and 44.5S$_{YsxC}$ particles

(A) The GTP hydrolysis rates of RbgA itself and in the presence of the mature and immature ribosomal particles were measured at different concentrations of GTP. Kinetic parameters were determined in the equivalent experiments with YphC (B) and YsxC (C) as well.

In addition, the interaction with the ribosomal particles had an influence in the apparent K_M as well of the three factors (Table 3.1). YphC and RbgA exhibited a small increase in the apparent K_M value in the presence of the 50S subunits but with a decrease value with the immature particles (exception of RbgA in the presence of $44.5S_{YsxC}$ particle). This suggested that the YphC and RbgA exhibited a higher affinity to GTP when interacting with the immature particles compared with the mature 50S subunits. In contrast, YsxC showed a decrease in the apparent K_M value with both the mature and immature particles. To have an overview of the apparent K_M values in the three GTPases, we found that the biggest decrease in the apparent K_M values happened always in the presence of the $45S_{YphC}$ particle. Additionally, the enzyme efficiency (Table 3.1, increase in k_{cat}/apparent K_M) for all the reactions performed in the presence of ribosomal particles increased, comparing to the reactions with only assembly factors by themselves.

In summary, the GTPase activity of RbgA, YphC and YsxC can be stimulated by both the mature 50S subunit and the three immature ribosomal particles. These results suggest that each one of the three GTPases is capable to bind to both the mature and immature ribosomal particles in a specific manner. This finding is consistent with what we observed in the binding assays, in which each of these factors can bind to both mature and immature particles in a promiscuous manner, suggesting that the last steps of assembly of the 50S subunits probably follow multiple parallel pathways.

	Apparent K_M(mM)	kcat(h^{-1})	kcat/ Apparent K_M ($mM^{-1} h^{-1}$)	Increase in apparent K_M	Increase in kcat	Increase in kcat/ Apparent K_M
RbgA	82.8±9.4	10.9±0.3	0.1	1	1	1
RbgA +50S	102.8±24.4	121.7±6.9	1.2	1.2	11.1	8.9
RbgA +45S $_{YphC}$	15.5±16	72±7.5	4.6	0.2	6.6	35
RbgA +45S $_{RbgA}$	53.4±21.6	61.9±5.2	1.2	0.6	5.6	8.8
RbgA +44.5S $_{YsxC}$	110.5±36.7	74.2±6.3	0.7	1.3	6.8	5.1

	Apparent K_M(mM)	kcat(h^{-1})	kcat/ Apparent K_M ($mM^{-1} h^{-1}$)	Increase in apparent K_M	Increase in kcat	Increase in kcat/ Apparent K_M
YphC	199±31.4	82.2±4.3	0.41	1	1	1
YphC+50S	217.4±39.2	317.9±19.2	1.4	1.1	3.9	3.5
YphC+45S $_{YphC}$	64.5±18.9	203.3±13.1	3.1	0.3	2.5	7.7
YphC+45S $_{RbgA}$	108.3±21.2	172.4±9.3	1.5	0.5	2.1	3.9
YphC+44.5S $_{YsxC}$	185.3±57.4	236.5±24.6	1.2	0.9	2.9	3.1

	Apparent K_M(mM)	kcat(h^{-1})	kcat/ Apparent K_M ($mM^{-1} h^{-1}$)	Increase in apparent K_M	Increase in kcat	Increase in kcat/ Apparent K_M
YsxC	1268±116.6	5.5±3.3	0.004	1	1	1
YsxC +50S	377.1±166.7	212.3±40.9	0.5	0.3	38.1	128.1
YsxC +45S $_{YphC}$	39.5±18.4	123±9.6	3.1	0.03	22.1	707.9
YsxC +45S $_{RbgA}$	219.6±174.6	147.6±45.5	0.6	0.2	26.5	153
YsxC +44.5S $_{YsxC}$	126.3±72.5	149±25.3	1.1	0.1	26.7	268.5

Table 3.1 The kinetic parameters from the GTPase assay of RbgA, YphC and YsxC in the absence and presence of the mature 50S subunit and the immature 45S$_{RbgA}$, 45S$_{YphC}$ and 44.5S$_{YsxC}$ particles

3.5 DISCUSSION

3.5.1 Assembly factors bind to the ribosomal particles in a promiscuous manner

The structural study demonstrated that the main function of RbgA, YphC and YsxC is

to help to facilitate the maturation of the functional sites including the central

protuberance, GTPase associated region and A, P, E sites of the 50S subunit (Ni, Joseph H Davis, et al. 2016). This result introduced the question as to whether these assembly factors can bind to these ribosomal particles. We used filtration assays to test the binding of the assembly factors. The binding assay showed that both the mature 50S subunit and the immature $45S_{RbgA}$, $45S_{YphC}$ and $44.5S_{YsxC}$ particles bind individual assembly factors, as well as multiple assembly factors simultaneously. However, a hierarchy of binding of these factors similar to that found for ribosomal proteins was not apparent in these assays. Instead, the binding of these assembly factors to the ribosomal particles was highly promiscuous. This finding is consistent with recent kinetic and structural work revealing that assembly of the ribosome occurs through multiple parallel pathways (Mulder 2010; Shajani et al. 2011; Adilakshmi et al. 2008). Additionally, the structural study indicated that the immature $45S_{RbgA}$, $45S_{YphC}$ and $44.5S_{YsxC}$ particles exhibited similar conformations (Jomaa et al. 2014; Ni et al. 2016). All of these biochemical and structural features suggested that they introduce in the cell a necessary flexibility and redundancy to make ribosome assembly an extremely robust and efficient process.

3.5.2 Assembly factors exhibited specific functional interplays with the ribosomal particles

Previous studies of RbgA revealed that the GTPase activity of RbgA can be maximally stimulated by the mature 50S subunit, but not the immature $45S_{RbgA}$ intermediates, and provided evidence to support the model in which RbgA utilized GTPase activity to dissociate from the mature 50S subunit (Achila et al. 2012; Uicker et al. 2006). This kinetic analysis of RbgA also suggested the specific functional

interplay between RbgA and ribosome particles. Thus, we performed similar kinetic assays for the YphC and YsxC protein and took RbgA as a reference. The three proteins showed a low intrinsic GTPase activity except YphC that had a much higher GTP hydrolysis rate considering it contains two GTPase domains. More importantly, similarly to RbgA, there was a significantly higher stimulation of the catalytic rate of the enzyme representing as k_{cat} against GTP for YphC and YsxC proteins in the presence of the mature 50S subunit. The three immature particles also stimulated the catalytic rate of assembly factors but to a lesser extent. These results indicated that there are specific functional interplays between YphC or YsxC and the mature and immature ribosomal particles.

For all three assembly factors, the enzyme catalytic rate (k_{cat}) is always significantly higher in the presence of mature 50S subunit than in the presence of immature particles. In contrast, the apparent K_M is lower (with the exception of RbgA+ $44.5S_{YsxC}$) in the reaction with immature particles than with the mature 50S subunit. These kinetic parameters suggested a model similar to that proposed for RbgA in the previous study (Achila et al. 2012). This model suggests that the assembly factors utilized the GTPase activity to dissociate from the mature 50S subunit.

3.5.3 The intermediates are either the real substrates or not diverged significantly from the real substrates

As previously described (Jomaa et al. 2014; Li et al. 2013), by using a *B. subtilis* strain in which RbgA was under the control of an IPTG inducible promoter, it was possible to purify the immature $45S_{RbgA}$ particles that are competent for maturation (Jomaa et al. 2014). Therefore, we used the same genetic approach to generate the

YphC and YsxC depletion strains, and managed to purify the immature $45S_{YphC}$ and $44.5S_{YsxC}$ particles. These immature particles including $45S_{RbgA}$, $45S_{YphC}$ and $44.5S_{YsxC}$ provided us useful tools to study the ribosome assembly process.

As the binding assays showed us, the RbgA, YphC and YsxC proteins can individually bind to each of the immature particles as well as the mature 50S subunit. In addition, multiple assembly factors can simultaneously bind to both mature and immature ribosomal particles. These bindings are specific as they trigger stimulation of the intrinsic GTPase activity of the assembly factors. Therefore, these results suggest that the assembly intermediates that accumulate in the depletion conditions for RbgA, YphC or YsxC are thermodynamically stable. They either constitute the actual substrates for the assembly factors or their conformations have not diverged significantly from the actual substrates in which the assembly factors RbgA, YphC and YsxC still bind to them.

3.6 MATERIALS AND METHODS

3.6.1 Construction of strains for ribosome purification and protein overexpression

As previously described in Chapter 2, *B. subtilis* RB301 (Pspank-*rbgA*), RB419 (Pspank-*infB*), RB260 (Pspank-*ysxC*) and RB290 (Pspank-*yphC*) strains were grown in the absence of isopropylb-d-thiogalactopyranoside (IPTG) and used to purify the mature and immature ribosome particles.

The *E. coli* BL21 (DE3) cells transformed with the pET21b-*ylqF* plasmid containing the full-length of *rbgA* with a C-terminal His$_6$-tag was used to overexpress the RbgA protein as described previously (Uicker et al. 2006). The pET15b-*yphC* and pET15b-

ysxC plasmids were used to overexpress YphC and YsxC proteins with an N-terminal His$_6$ tag in the *E. coli* BL21 (DE3) cells and cleaved by thrombin protease. The sequence of the *yphC* gene (NCBI reference sequence: NC_000964.3) and *ysxC* gene (NCBI reference sequence: NC_016047.1) were optimized for overexpression in *E. coli* cells using the GeneOptimizer software® and subsequently synthesized (Life Technologies; Thermo Fisher Scientific) with a *NdeI* and a *BamHI* site in the 5' and 3' ends of the gene, respectively. The genes were originally cloned into the carrier pMA-RQ (ampR) plasmid using the *SfiI* and *SfiI* cloning sites and subsequently sub-cloned into the final expression vector pET15b using the *NdeI* and a *BamHI* restriction sites. The final products were verified by sequencing (MOBIX, McMaster University).

3.6.2 Protein over-expression and purification

YphC and YsxC were overexpressed as N-terminal His$_6$-tag proteins by transforming into the *E. coli* BL21 (DE3) cells with pET15b-*yphC* and pET15b-*ysxC* plasmids, respectively. For both proteins, one liter of LB medium containing 100µg/ml ampicillin was inoculated with 10 mL of saturated overnight culture and cells were grown to OD$_{600}$ = 0.6 by incubation at 37 °C and shaking at 225 rpm in an Excella E24 incubator (New Brunswick). Expression was induced by the addition of 1mM isopropyl-D-1-thiogalactopyranoside (IPTG) into the culture. Cells were then induced for 3 hours at 37 °C and harvested by centrifugation at 3,700g for 15 min. Cell pellets were washed with 1 X PBS buffer (137 mM NaCl, 2.7 mM KCl, 8.1 mM Na$_2$HPO$_4$ at pH 7.4) and re-suspended in 20 mL of binding buffer containing 20 mM NaPO$_4$ pH 7.5, 0.5 M NaCl and 20 mM imidazole with a protease inhibitor cocktail (Complete Protease Inhibitor Cocktail Tablets. Roche). The cell suspension was passed through a

French press at 20,000 lb/in^2 pressure three consecutive times and the lysate was spun at 30,000g for 45 min to clear cell debris, then filtered with a 0.45 µm filter and loaded into a HiTrap Metal Chelating Column (GE Healthcare Life Sciences) previously equilibrated with binding buffer. Nonspecifically bound proteins were washed with buffer containing 20 mM NaPO$_4$ pH 7.5, 0.5M NaCl and 60 mM imidazole. Elution of YphC and YsxC was done by increasing the concentration of imidazole to 250 mM. Purity of the fractions was assayed by SDS-PAGE and fractions containing each respective protein were collected, pooled together and dialyzed overnight against 20 mM NaPO$_4$ pH 7.5 and 5% glycerol. The N-terminal His6-tags of YphC and YsxC were removed by digestion with thrombin (Sigma) that was added in the amount of 25 U/mL to the pooled fractions containing the target proteins during dialysis. Dialyzed protein preparations were then centrifuged at 12,000g for ten minutes in a 50ml Falcon centrifuge tube to remove any precipitated protein. The reaction mixtures were then loaded onto a Hi-Trap Q HP anion exchange column (GE Healthcare Life Sciences) for YphC protein and Hi-Trap SP HP column (GE Healthcare Life Sciences) for YsxC protein. The columns were pre-equilibrated with Buffer A (20mM NaPO$_4$ pH 7.5, 5% glycerol). A linear gradient of NaCl concentration from 0 mM and 1 M was used to wash and elute the protein. YphC and YsxC were eluted at a concentration of 350 mM and 500 mM NaCl, respectively. Protein-containing fractions were verified by SDS-PAGE, concentrated, and NaCl was removed by exchanging the buffer to desalting buffer (50mM Tris-HCl (pH 7.5), 750mM KCl, 5mM MgCl$_2$, 20mM imidazole, 2 mM DTT, and 10% glycerol) using a 10 kDa-cutoff centrifuged concentrator (Amicon). In the final step of concentration,

the buffer was exchanged to storage buffer (50mM Tris-HCl at pH 7.5, 750mM KCl, 5mM MgCl$_2$, 2mM DTT, and 10% glycerol) also using a 10 kDa-cutoff centrifuged concentrator (Amicon). Pure proteins were frozen in liquid nitrogen and stored at -80 ℃.

The RbgA protein was overexpressed as a C-terminal His$_6$-tag protein by transforming into *E. coli* BL21 (DE3) cells with the pET21b-*ylqF* plasmid. The overexpression and purification of RbgA was similar with YphC and YsxC. However, because the C-terminal His$_6$-tag in RbgA was not removable, the purification of RbgA only required a HiTrap Metal Chelating Column (GE Healthcare Life Sciences) to be purified. After elution, protein-containing fractions were verified by SDS-PAGE, concentrated, and desalted by exchanging the buffer to desalting buffer (50mM Tris-HCl (pH 7.5), 750mM KCl, 5mM MgCl$_2$, 20mM imidazole, 2 mM DTT, and 10% glycerol) and then frozen in liquid nitrogen and stored at -80 ℃ with storage buffer (50mM Tris-HCl at pH 7.5, 750mM KCl, 5mM MgCl$_2$, 2mM DTT, and 10% glycerol).

3.6.3 Binding assay

To perform the binding assays, Nanosep Omega centrifugal devices (PALL) (100 kDa cut-off) were prepared in order to block for non-specific binding of proteins by incubating the filter membrane with 500 μL of 1% [w/v] bovine serum albumin (BSA) for 90 min. Filters were then washed by rinsing with 500 μl of RNase free water and then removing any residual blocking solution by adding 500 μL of RNase free water and spinning at 12000g for 10 min. The binding reactions were prepared by incubating 200 pmoles of each assembly factor with 40 pmoles of mature or immature ribosomal particles in a 100 μl reaction in the binding buffer (10 mM Tris-HCl at pH

7.5, 7 mM magnesium acetate, 150 mM NH$_4$Cl and 1 mM DTT). GTP, GDP and GMPPNP were added in the reactions at a final concentration of 1 mM. Reactions were incubated at 37 ℃ for 30 min followed by centrifugation in the 100 kDa centrifugal devices at 12000g for 10 min to separate ribosomal particles and bound factors (B) that were retained by the filter from unbound proteins in the flow-through (FT) fraction. The flow-through was collected and the filter was gently washed twice with 100 μL of binding buffer followed by a 5 min spin at 12,000g. Finally, the ribosomal particles and bound proteins retained by the filter were vigorously resuspended in 100 μL of binding buffer and collected as the bound fraction (B). To resolve the flow-through (FT) and bound (B) fractions, 30 μL of sample were mixed with 6X SDS-PAGE loading buffer and loaded into a 4-12% Criterion^{TM} XT Bis-tris gel (Bio-Rad). Samples were run in XT MOPS buffer (Bio-Rad). Gels were stained with Coomassie Brilliant Blue and visualized using a ChemiDoc MP system (Bio-Rad).

3.6.4 GTPase assays

To measure the intrinsic GTPase activity, RbgA and YsxC were incubated at a concentration of 2 μM with a range of GTP concentrations (0-1mM). YphC is at a concentration of 200 nM in the equivalent reactions. The background of the assay itself was measured by running control reactions with no enzyme at each GTP concentration. These background values were subtracted from the total GTPase activity exhibited by the reactions containing the assembly factor at each GTP concentration. To determine the stimulation of RbgA, YsxC and YphC GTPase activity by the ribosomal particles, we performed reactions containing 50 nM

concentration of assembly factor and an equal concentration of either mature 50S subunits or one of the immature particles with a range of GTP concentrations (0-1mM). All assays were performed by first calculating the background GTPase activity from each ribosomal particle (50S subunit, $45S_{YphC}$, $44.5S_{YsxC}$ and $45S_{RbgA}$ particles) at 50 nM incubated from 0 to 1mM of GTP. This background subtraction ensured accuracy in the calculations by removing all background phosphate production not due to the stimulation.

All reactions were incubated at 37 ℃ for 30 min before measuring the released free phosphate by the malachite green assay (BioAssays Systems). The assay showed a linear behaviour for this incubation time. Reactions were performed in the reaction buffer (50 mM Tris-HCl (pH 7.5), 200 mM KCl, 10 mM $MgCl_2$ and 1 mM DTT) and terminated by the addition of malachite green reagent. Released phosphate was detected by monitoring the color formation at 620 nm using a 96-well plate reader (Tecan Sunrise). The K_M and k_{cat} values were calculated by fitting the data to the Michaelis-Menten equation with non-linear regression using the GraphPad Prism software. All these assays were performed at least in triplicate and with a minimum of two different preparations of the assembly factors and ribosomal particles.

CHAPTER 4. Using Cryo-electron Microscopy to Solve the Structure of the 50S Subunit in Complex with the RbgA Ribosomal Biogenesis Factor

4.1 AUTHOR'S PREFACE

The previous two chapters revealed that the main function of assembly factors YphC and YsxC is to assist the maturation of the functional sites of the 50S ribosomal particles. Now, to identify the binding site and the functional interplay between RbgA and the ribosomal particle, I aim to solve the structure of the 50S subunit in complex with RbgA using cryo-EM. I have performed all the cryo-EM experiments including sample preparation, data collection, and image processing. The preliminary complex structure of $45S_{RbgA}$.RbgA discussed in this chapter is currently being done by Dr. Nikhil Jain, a postdoctoral fellow working in Dr. Robert Britton's laboratory at Baylor College of Medicine. The RbgA-F6A mutation strain and the suppressor stains generated from RbgA-F6A were provided from Dr. Robert Britton's lab. Dr. Joaquin Ortega and I designed the experiments, analyzed the results and set up the future directions.

4.2 ABSTRACT

Despite the biochemical and structural identification of multiple assembly factors that facilitate the maturation of the large ribosome subunit, the molecular mechanisms through which they drive ribosome assembly are still elusive. Here, we analyze the interaction between RbgA and the 50S ribosomal subunit by using cryo-EM. The obtained preliminary cryo-EM structure of the 50S.RbgA complex exhibited an additional density attached to the P site. Rigid body fitting experiments suggested that this density may represent the RbgA protein. This possible RbgA density is partially overlapping with the position of h69 in the mature 50S subunit. We found that upon binding of the RbgA protein, the mature 50S structure reversed back to the immature conformation and the density for h68 and L16 r-protein disappeared. These preliminary results suggest that the RbgA functions at the very latest steps of the assembly of the 50S subunit. The maturation of the functional sites of the 50S subunit, especially h69, may require the dissociation of RbgA from the mature 50S subunit. We expect that the once obtained the atomic resolution structure of the 50S subunit in complex with RbgA will reveal the function and molecular mechanisms of this assembly factor.

4.3 INTRODUCTION

RbgA (ribosome biogenesis GTPaseA), previously known as YlqF, is an essential GTPase required for the late stages of assembly of the 50S subunits in *Bacillus subtilis* (Uicker et al. 2006; Gulati et al. 2013; Jomaa et al. 2014). Analysis of ribosome profiles from cells depleted of RbgA demonstrated that the large subunit sedimented as a 45S instead of a 50S particle. The proportion of 70S ribosomes was also greatly reduced under RbgA depletion conditions (Uicker et al. 2006). The structural characterization of the $45S_{RbgA}$ particles demonstrated that the main function of RbgA is to facilitate the maturation of key functional sites of the large ribosomal subunit in *B. subtilis* (Jomaa et al. 2014). However, the precise function of this protein in ribosome assembly remains elusive. Previous biochemical characterization of RbgA has demonstrated that the GTPase activity of RbgA can be maximally stimulated by the mature 50S subunit and it associates with both the mature and immature 50S particles in a nucleotide-dependent manner (Achila et al. 2012; Ni et al. 2016). Hence, it was hypothesized that the GTPase activity of RbgA promotes the late steps of 50S maturation and dissociating RbgA from the mature 50S subunit.

RbgA-F6A is a mutant of RbgA that displays reduced GTPase activity upon association with the ribosome but still binds to the ribosome. Thus, it was taken as the substitution of RbgA and used in suppressor analysis (Gulati et al. 2014). Several suppressor strains have been identified by using the RbgA-F6A mutant. These suppressors can partially restore the growth defect caused by the RbgA-F6A mutant but with the mutations clustered in the L6 ribosomal protein (Gulati et al. 2014). This finding indicated the possibility of a direct or indirect functional interplay between

RbgA and L6 protein. Our preliminary structure of the 50S.RbgA complex will provide a structural context to explain the functional interplay between RbgA and L6 protein that genetics has suggested. The preliminary structure of 50S.RbgA exhibits an additional density at the P site of the 50S subunit. This finding agrees with a preliminary reconstruction of the $45S_{RbgA}$.RbgA complex structure that is currently being obtained by our collaborators at the Britton lab.

We found that upon binding of RbgA to the 50S subunit, the structure reversed back to an immature conformation with the density for h68 and the L16 protein missing. The N-terminal end of RbgA overlaps with the location of h69 in the mature structure of the 50S subunit. This finding indicates that the dissociation of the N-terminal end of RbgA is required for the maturation of h69. More experiments including optimizing the assembly condition, acquiring more electron microscopy images and the application of different image processing approaches are still required to obtain an atomic resolution structure of the 50S.RbgA complex that will allow to explore the molecular mechanisms of RbgA in the maturation of 50S ribosome.

4.4 PRELIMINARY RESULTS

4.4.1 The preliminary structure of the 50S.RbgA complex

Before we constructed the 50S.RbgA complex, pelleting assays were used to evaluate the binding of RbgA to the mature 50S subunit (Figure 4.1). Binding of RbgA to the immature $45S_{RbgA}$ particles was also assessed. The result agreed with the filtration assays described in Chapter 3 that the RbgA binds to both the mature 50S subunit and the immature $45S_{RbgA}$ particles in the presence of GMPPNP (Figrue 4.1). Thus, we performed cryo-EM analysis to study the structure of RbgA in complex with the 50S

subunit. The complex was formed by incubating the mature 50S subunit with RbgA protein in the presence of GMPPNP at 37 ℃ for 30min. Subsequently, this *in vitro* assembled 50S.RbgA complex was subjected to image processing after cryo-EM. Surprisingly, we found 75% of particles contributed to the final 3D reconstruction representing the mature 50S subunit, however, there was one class of reconstruction with only 7% particles had an additional density at the P site of the 50S subunit (Figure 4.2A, left panel). More interestingly, upon binding of RbgA to the 50S subunit, the structure reversed back to the immature conformation with the density for h68 and L16 protein missing in the map (Figure 4.2A, right panel).
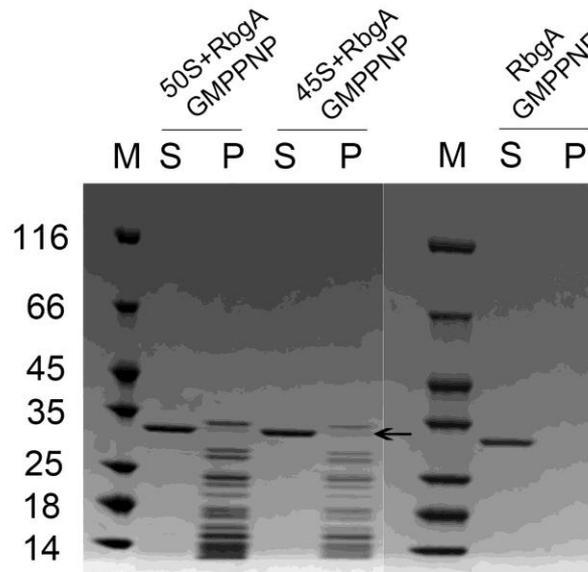
Figure 4.1 Pelleting assay to test the binding of RbgA to the mature 50S subunit or the immature 45S$_{RbgA}$ particles

The pelleting assay showed that RbgA binds to both mature 50S subunit and immature 45S$_{RbgA}$ particles in the presence of GMPPNP. The control showed that the protein itself was not aggregated in the pellet.

To better determine the percentage of 50S particles bound to RbgA, we applied a focus classification approach. This method considers only the region around the additional density for classification and excludes the influence of the other regions. We found that the occupancy of the additional density did not improve, however, the region around the additional density became more defined (Figure 4.2B). In addition, h69 and h89 exhibited an extended conformation (Figure 4.2B, right panel, labeled as star).
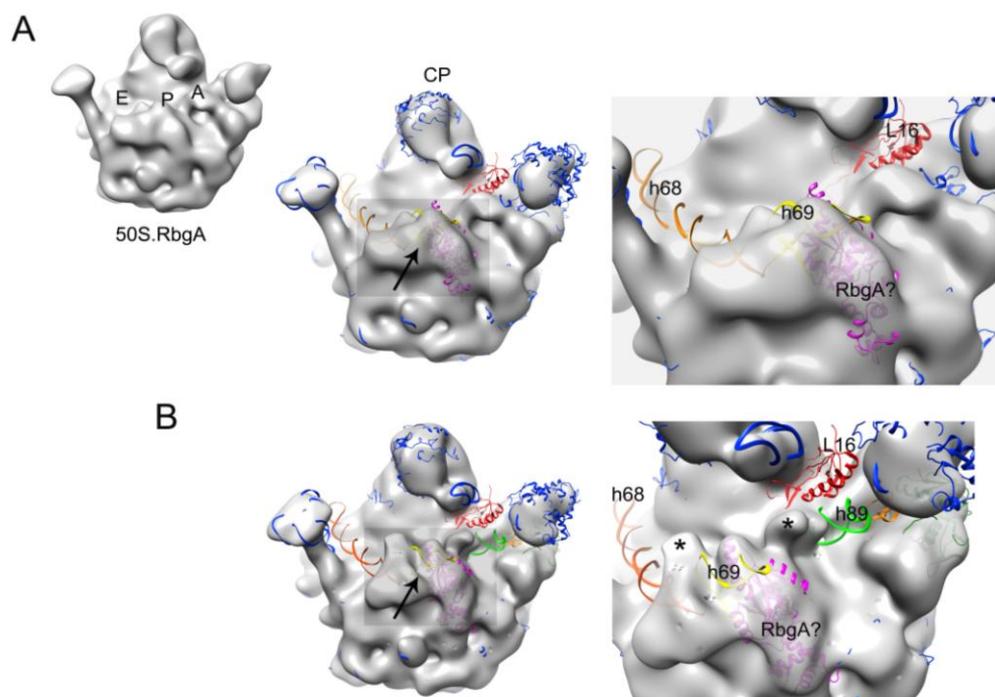
Figure 4.2 Conformation of 50S.RbgA complex with possible RbgA density

The 50S.RbgA complex had one class of conformation (7% occupancy) with an additional density at the P site of the 50S subunit (A, B, labeled as arrow). The whole reconstruction reverses back to the immature conformation with the L16 and h68 clearly missing in the density map, while the h69 and h89 exhibited a certain extent of conformational change (B, right panel, labeled as star).

Although the current reconstruction is still at low resolution, it is expected that with more datasets and particle images we will be able to refine this complex to a near-atomic or atomic resolution. At this level of resolution, we will be able to observe the interplay between the protein and the ribosome, and predict the function of RbgA during ribosome assembly. However, at the same time, it is also challenging that the low occupancy will dramatically increase the number of particles required to get to higher resolution structures. Thus, optimizing the assembly conditions and trying further classification methods are still required to study this complex.

4.5 DISCUSSION

4.5.1 Comparison with the 45S$_{RbgA}$.RbgA structure

Previously, genetic and biochemical approaches to analyze mutations of the RbgA protein provided insights into the interactions between RbgA and the 50S subunit (Gulati et al. 2013). The mutations of key residues in the C-terminal domain of RbgA altered the association with the 50S subunit, while mutations in certain residues in the N-terminal domain severely reduced the GTPase activity of RbgA. These results suggest that RbgA may influence ribosome assembly by interacting with the rRNA through its C-terminal domain. The binding of the C-terminal domain to rRNA is possible to couple with a conformational change in the N-terminal G-domain that could facilitate GTP hydrolysis (Gulati et al. 2013). To further identify the binding site of RbgA and the interactions between RbgA and the 50S subunit, we performed cryo-EM analysis and solved the structures of both 50S.RbgA and 45S$_{RbgA}$.RbgA complexes. As mentioned in the Results section, the low occupancy density map of 50S.RbgA complex showed an additional density around h69 at the P site of the 50S

subunit (Figure 4.2). In addition, the entire 50S structure reversed back to an immature conformation resembling the 50S immature intermediates (Ni et al. 2016; Jomaa et al. 2014) that are missing h68 and uL16 protein.

More interestingly, when our collaborator's laboratory tried to solve the structure of RbgA in complex with the immature $45S_{RbgA}$ particles, they found the same additional density in the immature particles to the one we observed in the 50S.RbgA complex (Figure 4.3A, left panel). However, the $45S_{RbgA}$.RbgA reconstruction with the additional density had around 40% occupancy among all the particles, which is much higher than the 7% in the 50S.RbgA complex. Thus, it allowed us to achieve higher resolution with a limited number of particles and perform a more accurate docking of the crystal structure of RbgA into the additional density that has higher resolution and with more details (Figure 4.3A, right panel). Similarly to the 50S.RbgA map, the RbgA protein binds to the P site of the subunit, with the missing density of h68 and L16 protein (Figure 4.3A, right panel). H89 also exhibits a certain extent of conformational change in the structure of $45S_{RbgA}$.RbgA complex (Figure 4.3A, right panel labeled as star). To further compare the reconstructions between 50S.RbgA and the $45S_{RbgA}$.RbgA complex, we fitted the immature $45S_{RbgA}$.RbgA structure into the mature 50S.RbgA density map. The two maps overlapped well with each other, including the additional density that represents RbgA protein and the missing density of h68 and L16 protein (Figure 4.3B). Compared with the previous structures of the immature $45S_{RbgA}$, $45S_{YphC}$ and $44.5S_{YsxC}$ particles (Ni et al. 2016; Jomaa et al. 2014), the $45S_{RbgA}$.RbgA complex exhibited a much more mature conformation that is similar to the 50S.RbgA complex. This result may indicate that the binding of RbgA

help to facilitate the folding of the immature particles into a close mature stage.

In summary, the two reconstructions suggested that the binding site of RbgA is likely at the P site of the large subunit around the h69 region. More work is still required to refine the structures to higher resolution and to have a more accurate fitting and an atomic model that describes the functional interplay between RbgA and the ribosome subunit at the molecular level.
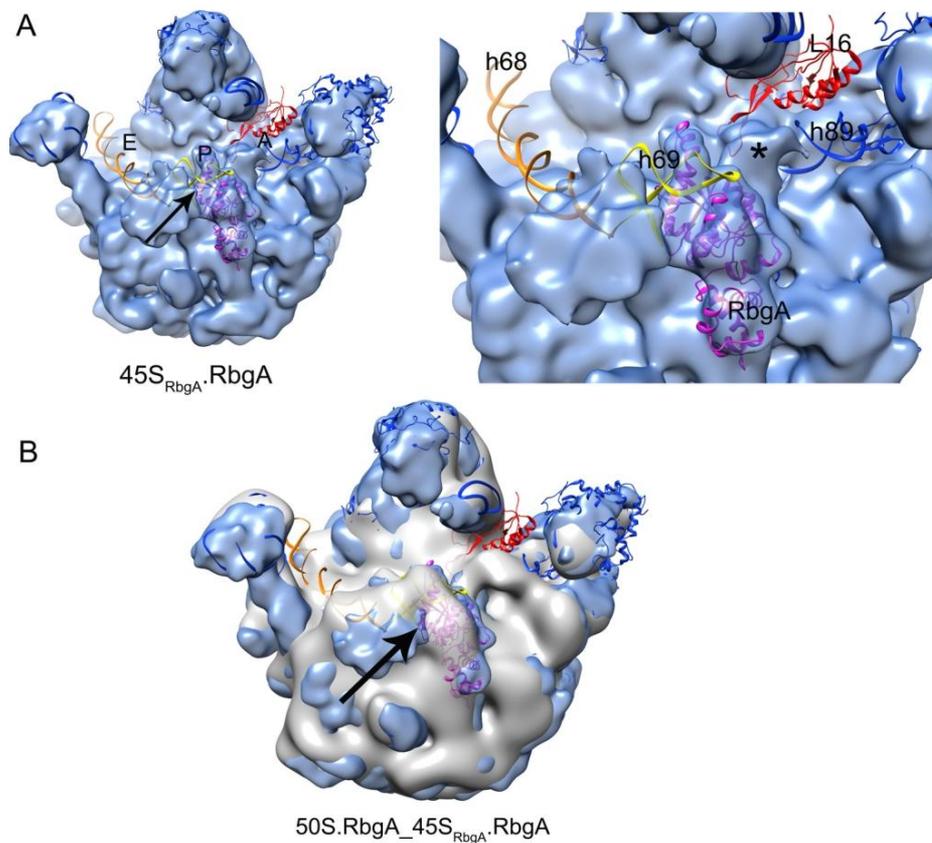
Figure 4.3 Cryo-EM structure of 45S$_{RbgA}$.RbgA complex

After rigid body fitting of the crystal structure of RbgA into the additional density in the 45S$_{RbgA}$.RbgA complex, it is exhibited that the RbgA protein binds to the P site of the subunit, with the missing density of h68 and L16 protein (A) and h89 also exhibits certain extent of conformation change (A, right panel labeled as star). (B) The immature 45S$_{RbgA}$.RbgA structure and the mature 50S.RbgA density map overlapped well with each other, including the additional density that represents RbgA protein and the missing density of h68 and L16 protein.

4.5.2 Hypothesis of the functions of RbgA in assisting the maturation of the 50S subunit

As previously described, RbgA help to facilitate the maturation of the key functional sites including the A, P and E sites of the 50S subunit. Our structural characterization of the 50S.RbgA and the 45S$_{RbgA}$.RbgA complex provides important insights of the binding sites of RbgA to the 50S subunit. The protein binds to the P site of the 50S subunit and may stimulate the conformational changes of this region.

To have a close view of the RbgA binding to the 45S$_{RbgA}$ or 50S particles, we found that the binding site of the N-terminal of RbgA protein overlaps with the position of h69 in the mature particle (Figure 4.4B, left panel). When overlapping the density map of 45S$_{RbgA}$.RbgA with the mature 50S map, we found that the density of the h69 (Figure 4.4B, right panel labeled in star) is superimposed onto the density representing the N-terminal of RbgA protein. This conformation suggested that the dissociation of the N-terminal of RbgA from the P site of 50S subunit is required for the maturation of h69. The GTPase activity of RbgA is maximally stimulated by the mature 50S subunit (Achila et al. 2012). Therefore, it is plausible that GTP hydrolysis promotes the complete dissociation of RbgA from the mature 50S subunit. However, the N-terminal of RbgA is required to dissociate first to allow the maturation of h69 and relevant helices such as h68 and h89 at the A, P and E site of the 50S subunit (Figure 4.4B). This might be the reason that we only capture a very low occupancy of additional density bound to the mature 50S subunit.

In both the 50S.RbgA and the 45S.RbgA density maps, we found that the density representing uL16 protein was missing (Figure 4.2 right panel, Figure 4.3A). This

protein is known to be crucial for the late steps of the maturation of the 50S subunit. When the mature 50S subunit was purified under low magnesium conditions, the mature 50S particles lost the uL16 protein and the structure of this uL16-depleted 50S particle reversed back to the immature conformation. The observed conformation was similar to that found here for the 50S.RbgA complex (Jomaa et al. 2014). So it is also possible that the binding of RbgA may facilitate the binding of uL16 protein, and the uL16 protein would work as a molecular glue to facilitate the maturation of the helices and r-proteins in functional sites of the 50S subunit, and promote the dissociation of the RbgA protein from the mature 50S ribosome through GTP hydrolysis.
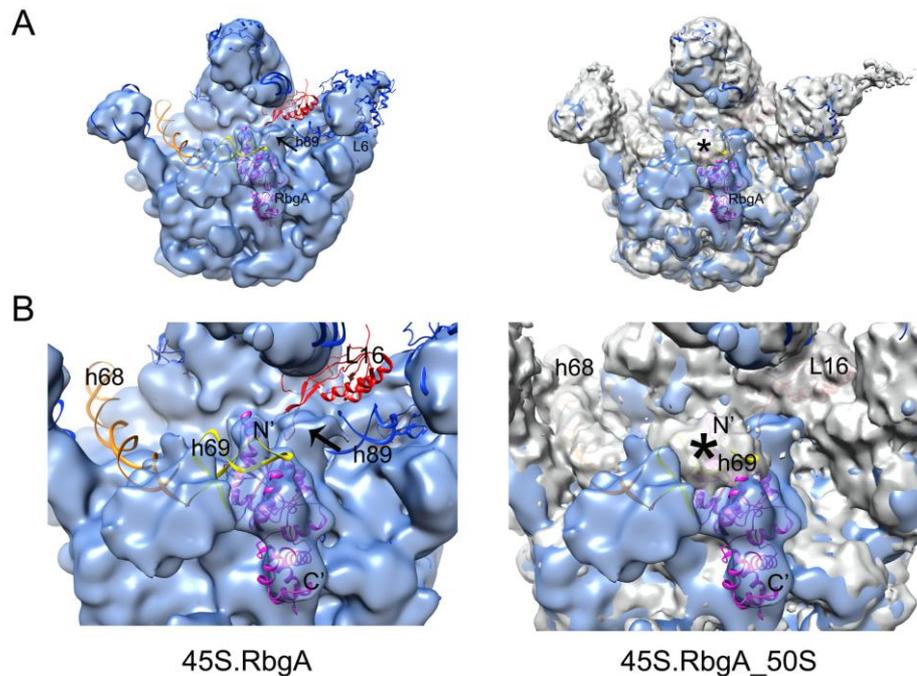
Figure 4.4 Zooming view of the binding site of the RbgA protein

The comparison of the structure of the mature 50S subunit (grey color) with the immature ribosome conformation bound with the RbgA protein (blue color). (A) The 45S.RbgA complex structure (left panel); the fitting map of 45S.RbgA into the mature 50S structure (right panel). (B) The zooming view of the interaction between h69 and the N-terminal of RbgA (left panel); the superimposed view of the density of h69 (labeled as star) and the N-terminal of RbgA (right panel).

4.5.3 The functional interplay between RbgA and uL6 protein

To gain insights into the function of RbgA, the Britton's group used genetic approaches and isolated suppressor mutations that could partially restore the growth defect from the RbgA-F6A mutant (Gulati et al. 2014). Analysis of these suppressors identified mutations clustered in the N-terminal region of the uL6 protein. Hence, this assay indicated that the uL6 mutations allow the defective RbgA-F6A protein to function more effectively in ribosome maturation. It suggested the possible direct or indirect functional interplay between RbgA and the uL6 protein. In the $45S_{RbgA}$.RbgA density map, we found the density representing the h89 adopting a 90-degree conformational change compared to its mature stage (Figure 4.5, right panel labeled as star). This shifted density exhibits special contacts with the N-terminal of RbgA protein. This conformation change likely also influence the conformation of the uL6 protein (Figure 4.5, right panel). The proper position of uL6 protein on the ribosome may help to set up the binding site for the uL16 protein, and stimulate the subsequent maturation of the functional sites of the 50S subunit as we previously described.

The low resolution of both 50S.RbgA and $45S_{RbgA}$.RbgA structures presently represent the limiting factor to gain more detailed information about these complexes. To explore the precise functions of RbgA in the maturation of the large ribosomal subunit, more experiments including optimizing the assembly condition, acquiring more particles and the application of different image processing approaches are still highly required.
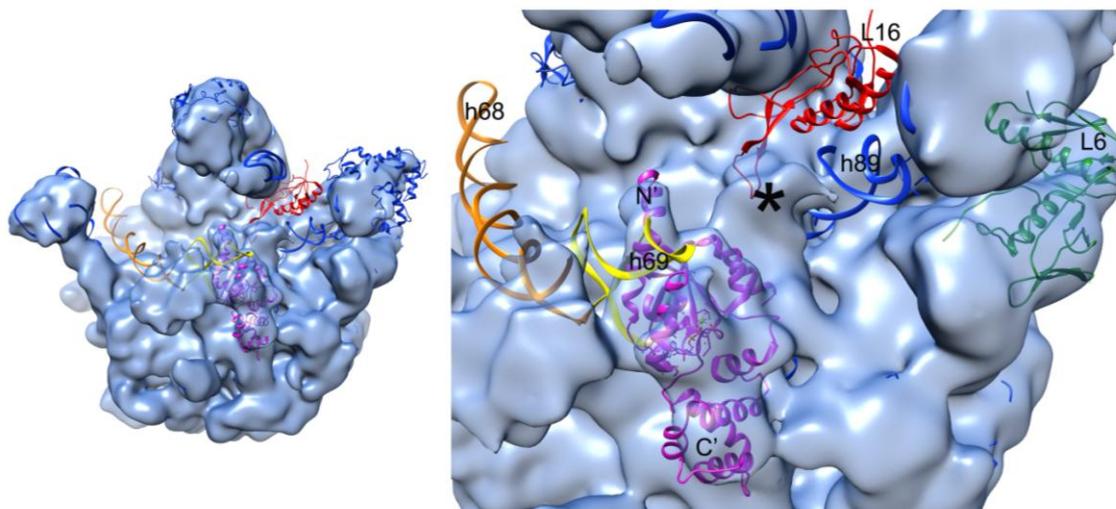
Figure 4.5 Zooming view of conformational change of h89

Upon binding of RbgA protein, the h89 had a 90-degree conformational change compared to its mature stage (right panel labeled as star), and had special contacts with the N-terminal of RbgA protein.

4.6 MATERIALS AND METHODS

4.6.1 Purification of wild-type RbgA protein

The *E. coli* BL21 (DE3) cells transformed with the pET21b-*ylqF* plasmid containing the full-length of *rbgA* with a C-terminal His$_6$-tag was used to overexpress the RbgA protein as described previously (Uicker et al. 2006). Briefly, one liter of LB medium containing 100µg/ml Ampicillin was inoculated with 10 mL of saturated overnight culture and cells were grown to OD$_{600}$ = 0.6 by incubation at 37 ℃ and shaking at 225 rpm in an Excella E24 incubator (New Brunswick). Expression was induced by the addition of 1mM isopropyl-D-1-thiogalactopyranoside (IPTG) into the culture. Cells were then induced for 3 hours at 37 ℃ and harvested by centrifugation at 3,700g for 15 min. Cell pellets were washed with 1 X PBS buffer (137 mM NaCl, 2.7 mM KCl, 8.1 mM Na$_2$HPO$_4$ at pH 7.4) and re-suspended in 20 mL of binding buffer containing

20 mM $NaPO_4$ pH 7.5, 0.5 M NaCl and 20 mM imidazole with a protease inhibitor cocktail (Complete Protease Inhibitor Cocktail Tablets. Roche). The cell suspension was passed through a French press at 20,000 lb/in^2 pressure three consecutive times and the lysate was spun at 30,000g for 45 min to clear cell debris, then filtered with a 0.45 μm filter and loaded into a HiTrap Metal Chelating Column (GE Healthcare Life Sciences) previously equilibrated with binding buffer. Nonspecifically bound proteins were washed with buffer containing 20 mM $NaPO_4$ pH 7.5, 0.5M NaCl and 60 mM imidazole. Elution of RbgA was done by increasing the concentration of imidazole to 250 mM. Because the C-terminal His_6-tag in RbgA was not removable, the purification of RbgA only required a HiTrap Metal Chelating Column (GE Healthcare Life Sciences) to be purified. After elution, protein-containing fractions were verified by SDS-PAGE, concentrated, and desalted by exchanging the buffer to desalting buffer (50mM Tris-HCl (pH 7.5), 750mM KCl, 5mM $MgCl_2$, 20mM imidazole, 2 mM DTT, and 10% glycerol) and then frozen in liquid nitrogen and stored at -80℃ with storage buffer (50mM Tris-HCl at pH 7.5, 750mM KCl, 5mM $MgCl_2$, 2mM DTT, and 10% glycerol).

4.6.2 Purification of the mature 50S subunit and the immature 45S$_{RbgA}$ particles

The mature 50S subunits and the immature 45S$_{RbgA}$ particles were purified from IF2-depleted (RB419) and RbgA-depleted (RB301) *B. subtilis* strains respectively. As previous described, in these cells the expression of IF2 and RbgA is under the control of an IPTG inducible promoter (Uicker et al. 2006). The mature 50S subunit and the immature 45S$_{RbgA}$ particles were purified as previous described (Jomaa et al. 2014; Ni

et al. 2016). Briefly, strains were first grown at 37 ℃ on LB plates containing 5mg/ml chloramphenicol and 1mM IPTG. Then cells were scraped off the plate and resuspended in 1ml of LB media, which was then inoculate in 200ml pre-warmed LB media without IPTG and cultures were grown from an OD_{600} of 0.1 to an OD_{600} of 0.5. The doubling time was calculated before the dilution into 1L of pre-warmed LB media and harvested when the doubling time reaches 120-150 min. Cells were spinned down through centrifugation at 8500g for 15 min, and the cell pellet was resuspended in buffer A (20mM Tris-HCl at pH 7.5, 15mM magnesium acetate, 100mM NH4Cl, 0.5mM EDTA, and 3 mM ß-mercaptoethanol). The cell suspension was passed through a French press at 20,000 lb/in^2 pressure three consecutive times to lyse the cells. The lysate was spun at 30,000g for 40min to clear cell debris using Beckman 70Ti rotor. The crude ribosomes were obtained as previous described (Jomaa et al. 2014) except the buffers contained 15mM magnesium acetate. To obtain the 50S or intermediates fractions, the crude ribosome pellet was resuspended in buffer E (10mM Tris–HCl at pH 7.5, 15mM magnesium acetate, 60mM NH4Cl and 3mM ß-mercaptoethanol) before loaded onto a 32 ml 10-30% (wt/vol) sucrose gradient made in buffer E and centrifuged at 43000 g for 16h using Beckman SW32 rotor. Gradients were fractionated using an AKTA prime purification system (GE Healthcare) and a Brandel fractionator apparatus. The elution profiles were monitored by UV absorbance at A_{260}, and the fractions corresponding to the 50S subunits or the $45S_{RbgA}$ particles were collected and pooled before concentrated and stored at -80℃ until further use.

4.6.3 Pelleting assay

The 50S subunit (50 pmol) and the WT RbgA (250 pmol) in the presence of nucleotides GMPPNP (1 mM) were incubated in 50µl binding buffer (10 mM Tris–HCl at pH 7.5, 15 mM magnesium acetate, 150 mM NH_4Cl and 1 mM DTT) at 37 ℃ for 1hr. The mixtures were then carefully loaded onto a 150µl sucrose cushion (10 mM Tris–HCl at pH 7.5, 15 mM magnesium acetate, 150 mM NH_4Cl, 1 mM DTT and 37.6% sucrose) and centrifuged using a TLA100 rotor (Beckman Coulter) at 85,000 rpm for 4hr. The supernatant (S) containing free protein that did not pellet with the 50S subunits was collected. The pellet (P) containing the free 50S particles and the 50S-bound proteins were resuspended in 200µl binding buffer. To resolve the supernatant and pellet fractions, 30µl of sample was mixed with 6x SDS-PAGE loading buffer and loaded into a 4%-12% Criterion XT Bis-tris gel (Bio-Rad). Samples were run in XT MOPS buffer (Bio-Rad). Gels were stained with Coomassie blue.

4.6.4 Assembly of the 50S.RbgA complex for cryo-EM

The mature 50S subunit (~1 µM) was incubated with His-RbgA (~40 µM) in the presence of GMPPNP (~1 mM) in the binding buffer (10mM Tris–HCl at pH 7.5, 10mM magnesium acetate, 60mM NH_4Cl and 3mM ß-mercaptoethanol) at 37 ℃ for 30min for forming the 50S.RbgA complex. After the incubation, the reactions were diluted for 20 times before applied to vitrification in Vitrobot. The dilution buffer was including 2 µM RbgA and 1 mM GMPPNP.

4.6.5 Cryo-electron microscopy and image processing

Approximately 3.6 µl of the diluted sample was applied in the holey carbon grids (c-flat CF-2/2-2C-T) with an additional layer of continuous thin carbon (5-10nm). Before

the sample was applied, grids were glow discharged in air at 5 mA for 15 seconds. Samples were then vitrificated in a Vitrobot (FEI) by blotting the grids once, 10 seconds each time and with an offset of 0 before they were plunged into liquid ethane. Grids were loaded in a Gatan 626 single tilt cryo-holder and introduced into a FEI Tecnai F20 electron microscope operated at 200 kV and equipped with a Gatan K2 Summit direct detector device camera. This detector was used in counting movie mode with five electrons per pixel per second for 15 seconds exposures and 0.5 seconds per frame. This method produced movies consisting of 30 frames with an exposure rate of ~1 e$^-$/Å^2. Movies were collected with a defocus range of 1 to 2.5µm and a nominal magnification of 25,000x, which produced images with a calibrated pixel size of 1.45 Å.

The 30 frames in each movie were aligned using the program alignframesleastsquares_list and averaged into one single micrograph with the shiftframes_list program (Rubinstein & Brubaker 2015). These programs are available from (https://sites.google.com/site/rubinsteingroup/home). These micrographs were used to estimate the parameters of the contrast transfer function using CTFFIND3 (Mindell & Grigorieff 2003) and also to determine the coordinates for particles in the frames of the movies. Then the particles from each averaged micrograph were picked using the autopicking procedure in Relion (Scheres 2012). The coordinates of particles in the frames were used to extract candidate particle images from the 30 unaligned frames in the movie. The motion of the individual particles in the frames was tracked and corrected using alignparts_lmbfgs algorithm (Rubinstein & Brubaker 2015). This procedure produced one stack of particle images fully corrected from

beam-induced motion from the first 20 frames of each movie. Therefore, the total accumulated dose to produce these particles images was 20 electrons per square angstrom.

The initial number of particles in the data sets for determining the preliminary 50S.RbgA was 140,519. These particles were then subjected to two- and three-dimensional classification using Relion (Scheres 2012). In the 50S.RbgA complex, the three-dimensional classes were produced from 101,077 particle images. The final three-dimensional consensus map with the binding of possible RbgA density was built from the 7,075 particle images that were assigned to one of the 3D classes obtained in the 3D classification.

CHAPTER 5. Discussion

5.1 A unified mechanism for the late stages of assembly of the 30S and the 50S subunits

5.1.1 The functional sites mature at the late stages of assembly in both the 30S and 50S subunits

Previously, in order to visualize the assembly intermediates of the 30S ribosome, the 16S rRNA of the 30S subunit was probed by using time-resolved hydroxyl radical footprinting (Adilakshmi et al. 2008). In general, nucleotides that were protected early from X-ray cleavage were helices contacted by the primary assembly proteins at the 5' domain (body) and the central domain (platform regions). In contrast, nucleotides close to the decoding center, which is the functional site of the 30S subunit required the longest time to fold (Adilakshmi et al. 2008). Therefore, it was concluded that the functional site of the subunit is the last region to fold into the mature conformation.

As described in this thesis, depletion of assembly factors RbgA, YphC and YsxC in *B. subtilis* results in the accumulation of late stage assembly intermediates of the 50S subunit. We structurally and biochemically characterized these immature particles and found that they exhibit distortions in the central protuberance, GTPase associated region (GAR) and key RNA helices in the A, P and E functional sites of the 50S subunits (Figure 5.1B) (Ni et al. 2016; Jomaa et al. 2014). Consistently, in a previous study deletion of YjeQ and RimM proteins in *E. coli*, cause the accumulation of 30S intermediates at late stages of assembly. These immature 30S particles exhibited distortions in the functional regions. Helix 44, which forms part of the decoding center (Figure 5.1A), and the entire helix 45, adopted a flexible conformation (Leong et al.

92

2013; Jomaa et al. 2011). However, in contrast to these unstructured functional sites, the body region of both the immature 30S and the 50S particles were folded into a mature conformation. Hence, the functional sites are the last regions to become mature in both the 30S and the 50S subunits. The assembly factors are likely to play their role in catalyzing the conformational changes within the functional sites and facilitating the final stages of ribosomal maturation.
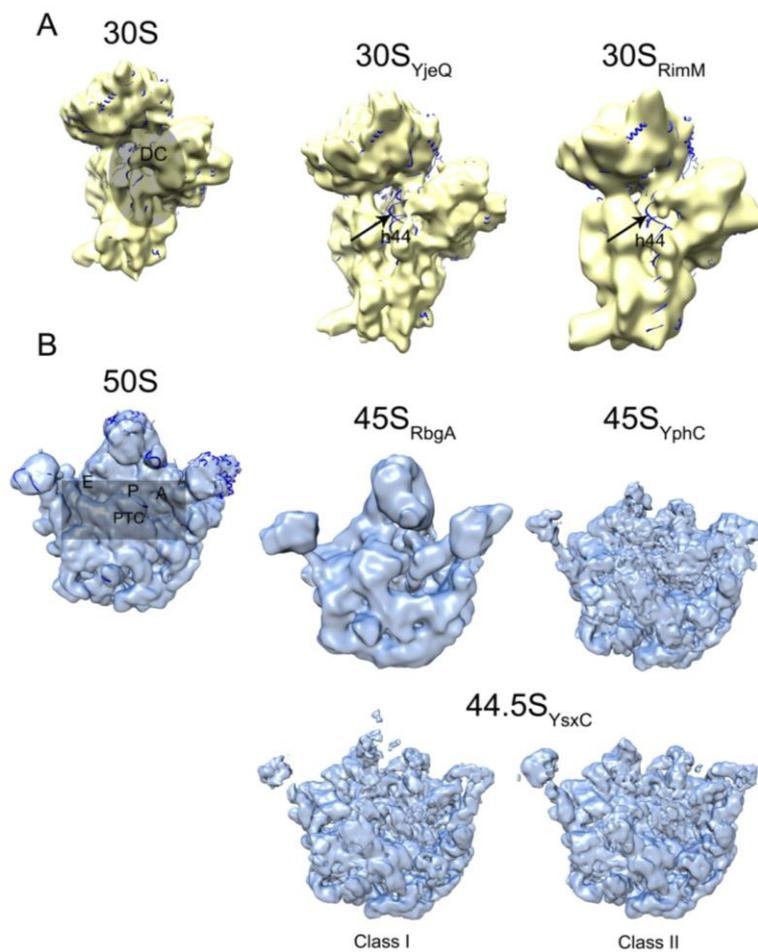
Figure 5.1 The immature intermediates of the 30S and 50S subunits (why in this figure things are not aligned? The 30S structures should be at the same scale.

The immature 30S and 50S particles exhibited different conformations. (A) Distortions in the decoding center of the immature 30S$_{YjeQ}$ and 30S$_{RimM}$. (B) The unstructured functional sites of the immature 45S$_{RbgA}$, 45S$_{YphC}$ and 44.5S$_{YsxC}$ particles.

5.1.2 The possible quality control checkpoint function of the assembly factors during ribosome assembly

Quality control in protein synthesis should be carefully controlled to ensure genetic information is transmitted from mRNA to proteins with fidelity. Thus, it is crucial to make sure the ribosome, the machine for protein translation, is properly assembled (Karbstein 2014).

In eukaryotic cells, it has been shown that releasing the misassembled ribosomes into the translational pool leads to disease. The Shwachman Bodian syndrome (SDS) is caused by the inability to release the assembly factor eIF6 and results in the impaired association of 40S and 60S subunits (Burwick et al. 2012). The alterations of cell cycles, cell proliferation and cell growth that are predisposed to cancer, has been associated with the dysregulation of ribosome biogenesis (Freed et al. 2010). These findings underscored the importance of a quality control system for ribosomes in eukaryotic cells. Only the fully assembled and functional ribosomes can proceed into the translating pool.

The assembly of ribosome subunits *in vivo* requires about 200 assembly factors in eukaryotes and about 40 factors in bacteria. It has been proven that the late cytoplasmic 40S ribosome assembly intermediates from *Saccharomyces cerevisiae* show a distorted decoding center that blocks 60S subunit association. Most importantly, the assembly factors located on the interface of the 30S subunit overlap the binding sites of translational initiation factors and thus prevent translation initiation steps (Strunk et al. 2012). These assembly factors may also play a role as quality control checkpoints that ensure the newly formed ribosomal subunits are fully

matured before engaging in the translation pathway.

A similar pattern has been found in bacteria. In the structures of immature 30S and 50S intermediates generated from assembly factor depletion or deletion strains, the functional sites are highly distorted. These functional sites include the decoding center of the 30S subunit and the peptidyl transferase center where peptide-bond formation is catalyzed during protein translation. Meanwhile, several of the helices that compose the inter-subunit bridges between 30S and 50S subunits exhibit conformational changes in the two immature subunits. In the 50S subunit, helix 38, in the A-site finger, forms the inter-subunit bridge B1a and reaches from the right side of the central protuberance of the 50S subunit to the middle of the head of the 30S subunit. Bridges B2a, B3, B5 and B6 are all involved in the interactions between the 50S subunit and the helix 44 of the 30S subunit (Yusupov et al. 2001). Therefore, the unstructured functional sites and the distortions of the inter-subunit bridges reveal that without these assembly factors, the 30S ribosomal subunits cannot associate with 50S subunits and form into mature 70S ribosomes that proceed into the translating ribosomal pool.

Additionally, structural studies have also reported that assembly factors RbfA and KsgA bind at positions that block the binding of translation factors IF1 and IF3 respectively. RbfA binds in a position overlapping the binding sites of the A- and P-site tRNAs, while KsgA overlaps the position of h38 (A-site finger) of 50S subunits (Datta et al. 2007; Boehringer et al. 2012). Similarly, Era binds at the 3' end of the 16S rRNA at the anti-Shine-Dalgarno sequence and thus, is likely to prevent mRNA recruitment (Tu et al. 2009). In the case of the 50S subunit, through our preliminary

work on the 50S or $45S_{RbgA}$ in complex with RbgA, we also found that the binding site of RbgA overlapped with the h69 of the 23S rRNA where the P- site tRNA binds. In addition, with the binding of RbgA the entire structure reverses back to the immature conformation. Taken together, although the precise functions of these assembly factors in bacteria are still elusive, it is likely that similarly to the eukaryotic assembly factors, assembly factors in bacteria may also have a role as checkpoint proteins and ensure only the fully matured ribosomes are functionally competent.

5.1.3 GTPase factors, main regulators of the late stages of ribosomal assembly

GTPases are the key players in the latest stages of ribosomal maturation. They include factors such as YjeQ and Era for the 30S subunit and RbgA, YphC (EngA in *E. coli*) and YsxC for the 50S subunit. In addition to a potential role as checkpoint proteins, it has been suggested that bacteria could use GTPase assembly factors to control the number of translational competent ribosomes based on the cellular GTP concentrations (Li et al. 2013; Jomaa et al. 2014; Britton 2009). In our previous biochemical assays of the 50S subunit, the assembly factors exhibited low affinities to the ribosomal particles in the presence of GTP, while the GTP/GDP exchange rates can be maximally stimulated by the mature 50S subunit and thus stimulating the dissociation of the assembly factors. Therefore, a decrease in the cellular GTP concentration could inhibit the GTPase activity of these assembly factors and stall the maturation of the ribosomal subunits.

A previous study in the 50S.EngA complex suggested that EngA might have a role as a sensor of the cellular GTP/GDP ratio. The EngA protein exhibited different conformational states to regulate the ribosome assembly based on this ratio (Zhang et

al. 2014). Similar to the 50S.EngA complex, binding of RbgA to the mature 50S subunit caused a reversion of the ribosomal particle to an immature state. This finding suggested that perhaps under nutrient deprivation, assembly factors might keep the ribosome subunits immature before becoming competent for translation in response to the GTP level in the cell.

5.2 Difference between the immature 30S and the immature 50S particles

We have been talking in the thesis about the immature intermediates that accumulated from the depletion of RbgA, YphC, YsxC, or deletion of YjeQ, RbfA, RimM. Pulse labeling experiments coupled with quantitative mass spectrometry, determined that these accumulated immature intermediates in both the 30S and the 50S subunits are competent for maturation (Jomaa et al. 2014; Thurlow et al. 2016). However, whether these immature particles are representing the real substrates for the assembly factors or they represent on-pathway intermediates remains unknown.

Previous work using time-resolved hydroxyl radical footprinting and time-resolved electron microscopy revealed that the assembly of the ribosome proceeds through multiple parallel pathways rather than a single linear pathway (Adilakshmi et al. 2008; Mulder 2010; Shajani et al. 2011). In addition, a convergence model was proposed that multiple parallel pathways of the subunits converge into a late assembly structurally similar common intermediate, in which the functional related assembly factors bind and catalyze the last steps of maturation (Figure 5.1A) (Leong et al. 2013; Guo et al. 2013; Jomaa et al. 2011; Jomaa et al. 2014). However, this scenario has changed in the recent studies of both the immature 30S and 50S particles.

For the 30S immature intermediates, the assembly factors seem to have very low

binding affinity to the immature 30S particles as determined by microscale thermophoresis (Thurlow et al. 2016). This result was consistent with mass spectrometry analysis revealing that the occupancy of these factors is actually below 10% in these immature 30S intermediates (Thurlow et al. 2016). Therefore, it was proposed that these structurally similar intermediates accumulated from the assembly factor deletion strains are probably local energy minimum intermediates that are thermodynamically more stable, but exhibited low affinity to the factors. In contrast, the actual on-pathway assembly intermediate that constitutes the real substrate for each factor likely precedes the ribosome particles accumulating in the deletion strains. In the case of the 50S subunit, our structural study of the immature 50S intermediates ($45S_{RbgA}$, $45S_{YphC}$ and $44.5S_{YsxC}$) revealed that the three immature particles present unstructured functional sites (Figure 5.1B) (Ni, Joseph H. Davis, et al. 2016). This result is consistent with what was found in the 30S immature particles (Figure 5.1A). However, with respect to the binding behavior, immature 50S particles were able to associate with each of the individual assembly factors, as well as multiple assembly factors at the same time. In addition, the GTPase activity of the assembly factors exhibited maximal stimulation of the catalytic rate in the reaction with the mature 50S subunit, and lower stimulation rate with the immature particles (Achila et al. 2012; Ni, Joseph H. Davis, et al. 2016). These results indicated the possible functional interplays between assembly factors and ribosome particles, and that the observed binding of assembly factors to the mature or immature 50S ribosomal particles is specific. Taken together, different from the immature 30S particles, the immature 50S particles are either the actual on-pathway intermediates or their conformations have

not significantly diverged from the actual substrate that the assembly factors still recognized.

5.3 Recent developments in studying ribosome assembly

5.3.1 Small chemical inhibitors as probes in the ribosome assembly process

Small chemical inhibitors of protein translation have been an extremely powerful tool to dissect the translation process and study the ribosome structure and function. A great variety of natural or synthetic antibiotics inhibit the proliferation of bacteria by binding to their ribosomes and interfering with protein translation (Poehlsgaard & Douthwaite 2005). However, with the continuing threat of drug resistance in bacteria, it is becoming increasingly clear that molecules of new chemical class and mechanism that could block ribosome function are necessary for future drug discovery (Comartin & Brown 2006).

Ribosome assembly, another largely enigmatic process, involves rRNA processing, binding of r-proteins, binding and release of the ribosomal assembly factors. Ribosome biogenesis and protein translation are inextricably linked. Thus, a challenge in using the chemical probes as inhibitors of ribosome assembly is that it is difficult to differentiate a phenotype is specifically caused by the inhibition of ribosome assembly process or the influence from the downstream protein translation, or from both (Stokes & Brown 2015). For example, the inhibition of protein translation results in the reduction of r-protein production and thus accumulating the immature ribosomal particles. Hence, previous studies in ribosome assembly were almost exclusively focusing on genetic and biochemical approaches. However, there is a paucity of chemical probes that have been identified to inhibit ribosome biogenesis in yeast and

mammals, such as the drug diazaborine that inhibits the maturation of rRNAs for the large ribosomal subunit (Loibl et al. 2014; Pertschy et al. 2004); or the small molecule CX-5461 that selectively inhibits the RNA polymerase I mediated rRNA synthesis in cancer cells (Drygin et al. 2011). In addition, there is one chemical inhibitor for studying the ribosome assembly in bacteria (Stokes & Brown 2015). Because mutations in many ribosome assembly factors result in a cold sensitivity phenotype, the small chemical inhibitor was found by screening diverse chemical compounds that induced cold sensitive growth inhibition in *E.coli* (Stokes et al. 2014). An anticonvulsant drug lamotrigine was finally found to have the most cold sensitive. It was shown that this small molecule was targeted to the initiation factor IF2 and was incapable of inhibiting protein synthesis and thus directly perturbed the ribosome assembly process (Stokes et al. 2014). The beneficial aspect for using this method is that it allows to kinetically monitor the immediate response after perturbation of ribosome biogenesis by the treatment of chemical inhibitors in the cell. This finding shed light on the usage of small chemical inhibitors as probes in the study of the ribosome assembly process. More importantly, the target specificity of ribosome biogenesis inhibitors will improve temporal resolution and might avoid the phenotype pleiotropy (Lerner & Inouye 1991) that is introduced by using genetic manipulations.

5.3.2 Using tandem affinity purification (TAP) to capture the true substrates for assembly factors

Ribosome assembly is a highly efficient process in the cell and intermediates do not accumulate. The genetic manipulation of ribosome assembly factors has been used to slow down this process and make it possible to isolate and characterize the immature

intermediates (Jomaa et al. 2014; Li et al. 2013; Ni et al. 2016; Leong et al. 2013). However, particles accumulating in genetically modified strains might be the thermodynamically stable off-pathway products that are not representing the true substrates for assembly factors (Thurlow et al. 2016). Thus, in addition to the utility of small chemical probes for studying ribosome assembly, there have been significant advances by using the tandem affinity purification (TAP) to purify the native complex with the protein of interest for structural studies (Puig et al. 2001; Wyler et al. 2011; van der Feltz & Pomeranz Krummel 2016).

The TAP method uses two tags, the calmodulin binding peptide (CBP) tag that can be recognized by protein calmodulin in the presence of calcium, and the protein A tag that is used to bind to the immunoglobulin G (IgG). Between these two tags a linker for the cleavage of TEV protease was incorporated (Figure 5.2). The TAP approach is performed in two steps, briefly, the complex of interest is genomically fused with a sequence comprised of a CBP tag (green), a cleavage site for the protease TEV (blue), and two IgG binding regions of protein A (Pr A, red) (Figure 5.2). Then the cell lysate is incubated with an IgG resin and the complex binds to the resin through its interaction with the protein A in the sequence. The complex is first released from the IgG resin by the cleavage of TEV protease. In the second step, the complex is incubated with a calmodulin resin before releasing via addition of the calcium chelator EGTA (van der Feltz & Pomeranz Krummel 2016). The purified complex with the protein of interest is a bona fide sample for cryo-EM structural determination due to its compositionally and structurally homogeneous state as well as high concentration. More importantly, the TAP method preserves the integrity of a complex and can be

used with different tagged proteins in a complex for purifying a native complex.

This method was initially developed in yeast but in recent years it has been successfully applied in various organisms especially in eukaryotic cells (Puig et al. 2001; Rigaut et al. 1999). In the study of ribosome assembly in eukaryotes, this method has been used for the isolation of pre-ribosomal complexes under physiological conditions (Bradatsch et al. 2013; Fromont-Racine et al. 2003; Wu et al. 2016). With the combination of mass spectrometry identification of the components and the cryo-EM structural determination, this approach became a powerful tool for the analysis of the ribosome assembly pathway.
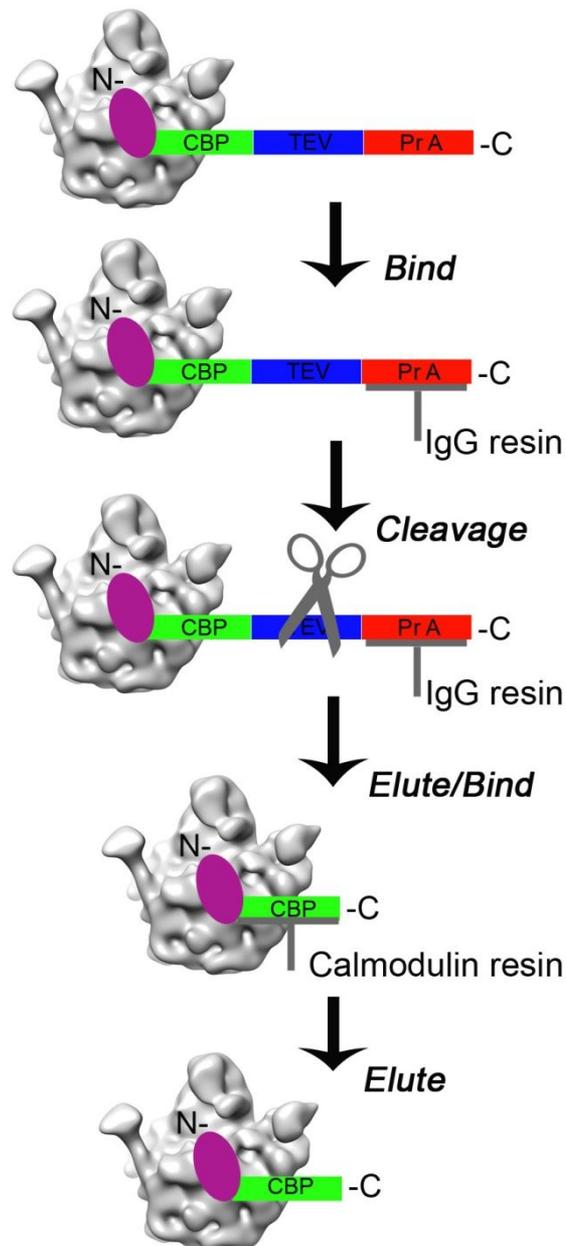
Figure 5.2 Schematic of the two steps tandem affinity purification (TAP)

Genomically fused a sequence comprised of a calmodulin binding peptide (CBP, green), linker site for the recognition of TEV protease (Blue), and two IgG binding regions of Protein A (Pr A, red) with the C-terminal of a protein of interest (Purple) that will be the target to pull out the complex of interest (Grey, representing the ribosome/protein complex).

5.4 Significance

The work in this thesis provides novel biochemical and structural insights on the functions of the assembly factors at the late stages of large ribosomal subunit assembly. Ribosome assembly intermediates generated through depletion or deletion bacterial strains still constitute today a significant tool for studying the function of assembly factors. We were able to determine for the first time the high resolution cryo-EM structure of the immature 50S subunits isolated from the YphC or YsxC depletion bacterial strains. By comparison to the mature 50S subunits, these structures exhibited distorted functional sites, however with an already mature conformation in the body region. The results suggest that the main function of assembly factors is to facilitate the maturation of functional sites. These functional sites fold in the last steps of ribosome assembly, probably to prevent immature initiation of protein synthesis as we previously described.

Future structural studies of the functions of assembly factors can be optimized by improving the sample preparation approach such as using tandem affinity purification (TAP) to capture different native complexes with the factors of interest. With the recent breakthroughs in both hardware and software in single particle cryo-EM, structural studies of these complexes to the atomic resolution will allow us to understand the function of assembly factors at the molecular level.

The atomic model of ribosome rationalizes over four decades of biochemical data and provides a large amount of information about RNA and protein structure, and protein-RNA interactions (Wimberly et al. 2000). The structure of the ribosome allows us to analyze the functions of the ribosomal subunits, such as decoding, peptide bond

formation and also for understanding the action of antibiotics. Our ongoing work aimed at understanding the function of assembly factors during the ribosome assembly process will also set up the biochemical and structural basis for future perspectives in the application, especially in drug discovery. As we described previously, considering the continuing threat of drug resistance in bacteria, new mechanisms of inhibition are necessary, as most of the antibiotic currently used are interfering with the protein translation process (Comartin & Brown 2006; Poehlsgaard & Douthwaite 2005; Champney 2006). Recent prospects for blocking ribosome function are increasingly focusing on preventing the assembly of the bacterial ribosome, and a number of assembly factors such as Era, Obg, YjeQ and RbgA represent promising targets for modern antimicrobial drug discovery (Comartin & Brown 2006). In addition, the assembly factor/ribosome complexes purified by TAP approaches could be expected to yield additional factors important for this process, and provide us with novel targets for drug discovery. This research may have a great impact on the development of future strategies to fight bacterial infection.

REFERENCES

Achila, D. et al., 2012. Biochemical characterization of ribosome assembly gtpase RbgA in Bacillus subtilis. Journal of Biological Chemistry, 287(11), pp.8417–8423.

Adilakshmi, T., Bellur, D.L. & Woodson, S.A., 2008. Concurrent nucleation of 16S folding and induced fit in 30S ribosome assembly. Nature, 455(7217), pp.1268–1272.

Agris, P.F., Vendeix, F.A.P. & Graham, W.D., 2007. tRNA's Wobble Decoding of the Genome: 40 Years of Modification. Journal of Molecular Biology, 366(1), pp.1–13.

Amaro, A.M. & Jerez, C.A., 1984. Methylation of ribosomal proteins in bacteria: evidence of conserved modification of the eubacterial 50S subunit. Journal of bacteriology, 158(1), pp.84–93.

Bai, X. chen, McMullan, G. & Scheres, S.H.W., 2014. How cryo-EM is revolutionizing structural biology. Trends in Biochemical Sciences, 40(1), pp.49–57.

Bai, X.C. et al., 2013. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. eLife, 2013(2), pp.2–4.

Bai, X.-C. et al., 2015. Sampling the conformational space of the catalytic subunit of human γ-secretase. eLife, 4(November), p.e11182.

Bąkowska-Zywicka, K. & Tyczewska, A., 2009. The structure of the ribosome -short history. Biotechnologia, 1(1), pp.14–23.

Ban, N. et al., 1999. Placement of protein and RNA structures into a 5 A-resolution map of the 50S ribosomal subunit. Nature, 400(6747), pp.841–847.

Ban, N., 2000. The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 A Resolution. Science, 289(5481), pp.905–920.

Blombach, F., Brouns, S.J.J. & van der Oost, J., 2011. Assembling the archaeal ribosome: roles for translation-factor-related GTPases. Biochemical Society transactions, 39(1), pp.45–50.

Boehringer, D. et al., 2012. Structural insights into methyltransferase KsgA function in 30S ribosomal subunit biogenesis. The Journal of biological chemistry, 287(13), pp.10453–9.

Bradatsch, B. et al., 2013. Structure of the pre-60S ribosomal subunit with nuclear export factor Arx1 bound at the exit tunnel. Nat Struct Mol Biol, 19(12), pp.1234–1241.

Brendan MacLean. et al., 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics, 26(7), p.966.

Britton, R.A. et al., 2007. Maturation of the 5' end of Bacillus subtilis 16S rRNA by the essential ribonuclease YkqC/RNase J1. Molecular Microbiology, 63(1), pp.127–138.

Britton, R.A., 2009. Role of GTPases in Bacterial Ribosome Assembly. Annu. Rev. Microbiol, 63, pp.155–76.

Brown, E.D. & Wright, G.D., 2016. Antibacterial drug discovery in the resistance era. Nature, 529, pp336-343.

Burwick, N. et al., 2012. Impaired ribosomal subunit association in Shwachman-Diamond syndrome. Blood, 120(26), pp.5143–5152.

Campbell, T.L. et al., 2006. The yjeQ gene is required for virulence of Staphylococcus aureus. Infection and immunity, 74(8), pp.4918–21.

Champney, W.S., 2006. The other target for ribosomal antibiotics: inhibition of bacterial ribosomal subunit formation. Infectious disorders drug targets, 6(4), pp.377–90.

Chen, S.S. & Williamson, J.R., 2013. Characterization of the ribosome biogenesis landscape in E. Coli using quantitative mass spectrometry. Journal of Molecular Biology, 425(4), pp.767–779.

Cheng, Y. et al., 2015. A primer to single-particle cryo-electron microscopy. Cell, 161(3), pp.439–449.

Cheng, Y., 2015. Single-particle Cryo-EM at crystallographic resolution. Cell, 161(3), pp.450–457.

Comartin, D.J. & Brown, E.D., 2006. Non-ribosomal factors in ribosome subunit assembly are emerging targets for new antibacterial drugs. Current Opinion in Pharmacology, 6(5), pp.453–458.

Connolly, K. & Culver, G., 2009. Deconstructing ribosome construction. Trends in Biochemical Sciences, 34(5), pp.256–263.

Culver, G.M. & Noller, H.F., 1999. Efficient reconstitution of functional Escherichia coli 30S ribosomal subunits from a complete set of recombinant small subunit ribosomal proteins. RNA, 5(6), pp.832–43.

Datta, P.P. et al., 2007. Structural aspects of RbfA action during small ribosomal subunit assembly. Molecular cell, 28(3), pp.434–45.

Decatur, W.A. & Fournier, M.J., 2002. rRNA modifications and ribosome function. Trends in Biochemical Sciences, 27(7), pp.344–351.

Dnina, M. V & Wintermeyer, W., 2001. Ribosome fidelity: tRNA discrimination, proofreading and induced fit. Trends in Biochemical Sciences, 26(2), pp.124–130.

Doi, Y. & Arakawa, Y., 2007. 16S Ribosomal RNA Methylation: Emerging Resistance Mechanism against Aminoglycosides. Clinical Infectious Diseases, 45(1), pp.88–94.

Drygin, D. et al., 2011. Targeting RNA polymerase I with an oral small molecule CX-5461 inhibits ribosomal RNA synthesis and solid tumor growth. Cancer research, 71(4), pp.1418–30.

Ehrenberg, M. & Blomberg, C., 1980. Thermodynamic constraints on kinetic proofreading in biosynthetic pathways. Biophysical Journal, 31(3), pp.333–358.

Van der Feltz, C. & Pomeranz Krummel, D., 2016. Purification of Native Complexes for Structural Study Using a Tandem Affinity Tag Method. Journal of Visualized Experiments, (113), pp.e54389–e54389.

Frank, J. et al., 1995. A model of the translational apparatus based on a three-dimensional reconstruction of the Escherichia coli ribosome. Biochemistry and cell biology, 73(11–12), pp.757–65.

Frank, J., 2009. Single-particle reconstruction of biological macromolecules in electron microscopy--30 years. Quarterly reviews of biophysics, 42(3), pp.139–58.

Freed, E.F. et al., 2010. When ribosomes go bad: diseases of ribosome biogenesis. Molecular bioSystems, 6(3), pp.481–93.

Fromont-Racine, M. et al., 2003. Ribosome assembly in eukaryotes. Gene, 313(1–2), pp.17–42.

Gillet, L.C. et al., 2012. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Molecular & cellular proteomics : MCP, 11(6), p.O111.016717.

Ginsburg, D. & Steitz, J.A., 1975. The 30 S ribosomal precursor RNA from Escherichia coli. A primary transcript containing 23 S, 16 S, and 5 S sequences. The Journal of biological chemistry, 250(14), pp.5647–54.

Glaeser, R.M., 2016. How good can cryo-EM become? Nat Methods, 13(1), pp.28–32.

Goto, S., Muto, A. & Himeno, H., 2013. GTPases involved in bacterial ribosome maturation. Journal of Biochemistry, 153(5), pp.403–414.

Gulati, M. et al., 2014. Functional Interaction between Ribosomal Protein L6 and RbgA during Ribosome Assembly. PLoS Genetics, 10(10).

Gulati, M. et al., 2013. Mutational analysis of the ribosome assembly GTPase RbgA provides insight into ribosome interaction and ribosome-stimulated GTPase activation. Nucleic Acids Research, 41(5), pp.3217–3227.

Guo, Q. et al., 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 Research, 41(4), pp.2609–2620.

Hage, A. El & Tollervey, D., 2004. A surfeit of factors: why is ribosome assembly so much more complicated in eukaryotes than bacteria? RNA biology, 1(1), pp.10–5.

Hayes, F. et al., 1971. Additional nucleotide sequences in precursor 16S ribosomal RNA from Escherichia coli. Nature: New biology, 232(28), pp.54–5.

Held, W.A. et al., 1974. Assembly Mapping of 30 S Ribosomal Proteins from Escherichia coli. Further studies. , 249(10), pp.3103–3111.

Held, W.A., Mizushima, S. & Nomura, M., 1973. Reconstitution of Escherichia coli 30 S ribosomal subunits from purified molecular components. J Biol Chem, 248(16), pp.5720–5730.

Herold, M. & Nierhaus, K.H., 1987. Incorporation of six additional proteins to complete the assembly map of the 50 S subunit from Escherichia coli ribosomes. Journal of Biological Chemistry, 262(18), pp.8826–8833.

Hopfield, J.J., 1974. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. Proceedings of the National Academy of Sciences of the United States of America, 71(10), pp.4135–9.

Jomaa, A., Jain, N., Davis, J.H., et al., 2014. Functional domains of the 50S subunit mature late in the assembly process. Nucleic acids research, 42(5), pp.3419–35.

Jomaa, A. et al., 2011. Understanding ribosome assembly: the structure of in vivo assembled immature 30S subunits revealed by cryo-electron microscopy. RNA (New York, N.Y.), 17(4), pp.697–709.

Kaczanowska, M. & Rydén-Aulin, M., 2007. Ribosome biogenesis and the translation process in Escherichia coli. Microbiology and molecular biology reviews : MMBR, 71(3), pp.477–94.

Karbstein, K., 2014. Quality Control Mechanisms During Ribosome Maturation. Trends Cell Biol, 23(5), pp.242–250.

Kressler, D., Hurt, E. & Ba??ler, J., 2010. Driving ribosome assembly. Biochimica et Biophysica Acta - Molecular Cell Research, 1803(6), pp.673–683.

Kucukelbir, A., Sigworth, F.J. & Tagare, H.D., 2014. Quantifying the local resolution of cryo-EM density maps. Nature methods, 11(1), pp.63–5.

Leidig, C. et al., 2014. 60S ribosome biogenesis requires rotation of the 5S ribonucleoprotein particle. Nature Communications, 5, pp 1-8.

Leonardo G. Trabuco, E.V.K.M.J.F.K.S., 2008. Flexible Fitting of Atomic Structures into Electron Microscopy Maps Using Molecular Dynamics. Structure (London, England : 1993), 16(5), p.673.

Leong, V. et al., 2013. Escherichia coli rimM and yjeQ null strains accumulate immature 30S subunits of similar structure and protein complement. RNA (New York, N.Y.), 19(6), pp.789–802.

Lerner, C.G. & Inouye, M., 1991. Pleiotropic changes resulting from depletion of Era, an essential GTP-binding protein in Escherichia coli. Molecular Microbiology, 5(4), pp.951–957.

Li, N. et al., 2013. Cryo-EM structures of the late-stage assembly intermediates of the bacterial 50S ribosomal subunit. Nucleic Acids Research, 41(14), pp.7073–7083.

Li, Z. & Deutscher, M.P., 1995. The tRNA processing enzyme RNase T is essential for maturation of 5S RNA. Proceedings of the National Academy of Sciences of the United States of America, 92(15), pp.6883–6.

Li, Z., Pandit, S. & Deutscher, M.P., 1999. RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. The EMBO journal, 18(10), pp.2878–85.

Lindahl, L., 1975. Intermediates and time kinetics of the in vivo assembly of Escherichia coli ribosomes. Journal of Molecular Biology, 92(1), pp.15–37.

Lindahl L, 1973. Two new ribosomal precursor particles in E. coli. Nat New Biol, 243(127), pp.170–2.

Loibl, M. et al., 2014. The Drug Diazaborine Blocks Ribosome Biogenesis by Inhibiting the AAA-ATPase Drg1. Journal of Biological Chemistry, 289(7), pp.3913–3922.

Loken, C. et al., 2010. SciNet: Lessons Learned from Building a Power-efficient Top-20 System and Data Centre. Journal of Physics: Conference Series, 256(1), p.12026.

Lyumkis, D. et al., 2013. Likelihood-based classification of cryo-EM images using FREALIGN. Journal of structural biology, 183(3), pp.377–88.

Matsuo, Y. et al., 2006. The GTP-binding protein YlqF participates in the late step of 50 S ribosomal subunit assembly in Bacillus subtilis. Journal of Biological Chemistry, 281(12), pp.8110–8117.

Merk, A. et al., 2016. Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. Cell, 165(7), pp.1698–1707.

Mindell, J.A. & Grigorieff, N., 2003. Accurate determination of local defocus and specimen tilt in electron microscopy. Journal of Structural Biology, 142(3), pp.334–347.

Moore, P.B., 1991. Ribosomes. Current Opinion in Structural Biology, 1(2), pp.258–263.

Morimoto, T. et al., 2002. Six GTP-binding proteins of the Era/Obg family are essential for cell growth in Bacillus subtilis. Microbiology (2002), 148, 3539–3552.

Muench, S.P. et al., 2006. The essential GTPase YphC displays a major domain rearrangement associated with nucleotide binding. Proceedings of the National Academy of Sciences of the United States of America, 103(33), pp.12359–64.

Mulder, A.M., 2010. Visualizing Ribosome Biogenesis. Science, 330(October), pp.673–677.

Ni, X., Davis, J.H., et al., 2016. YphC and YsxC GTPases assist the maturation of the central protuberance, GTPase associated region and functional core of the 50S ribosomal subunit, Nucleic Acids Research, 9140, pp.1–14.

Nierhaus, K.H., 1991. The assembly of prokaryotic ribosomes. Biochimie, 73(6), pp.739–55.

Nogales, E., 2016. The development of cryo-EM into a mainstream structural biology technique. Nat Meth, 13(1), pp.24–27.

Nogales, E. & Scheres, S.H.W., 2015. Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. Molecular Cell, 58(4), pp.677–689.

Noller, H.F., 1991. Ribosomes. Drug and the RNA world. Nature, 353, pp 302-303.

Nomura, M., Gourse, R. & Baughman, G., 1984. Regulation of the synthesis of ribosomes and ribosomal components. Annual review of biochemistry, 53(1997), pp.75–117.

Nomura, M., Traub, P. & Bechmann, H., 1968. Hybrid 30S ribosomal particles reconstituted from components of different bacterial origins. Nature, 219(5156), pp.793–9.

Pan, J., Thirumalai, D. & Woodson, S.A., 1997. Folding of RNA involves parallel pathways. Journal of Molecular Biology, 273(1), pp.7–13.

Pertschy, B. et al., 2004. Diazaborine treatment of yeast cells inhibits maturation of the 60S ribosomal subunit. Molecular and cellular biology, 24(14), pp.6476–87.

Pettersen, E.F. et al., 2004. UCSF Chimera-A visualization system for exploratory research and analysis. Journal of Computational Chemistry, 25(13), pp.1605–1612.

Poehlsgaard, J. & Douthwaite, S., 2005. The bacterial ribosome as a target for antibiotics. Nature reviews. Microbiology, 3(11), pp.870–81.

Puig, O. et al., 2001. The Tandem Affinity Purification (TAP) Method: A General Procedure of Protein Complex Purification. Methods, 24(3), pp.218–229.

Ralston, C.Y. et al., 2000. Time-resolved synchrotron X-ray footprinting and its application to RNA folding. Methods in Enzymology, 317, pp.353–368.

Rheinberger, H.J., Sternbach, H. & Nierhaus, K.H., 1981. Three tRNA binding sites on Escherichia coli ribosomes. Proceedings of the National Academy of Sciences of the United States of America, 78(9), pp.5310–4.

Rigaut, G. et al., 1999. A generic protein purification method for protein complex characterization and proteome exploration. NATURE BIOTECHNOLOGY, 17, pp.1030–1032.

Rosenthal, P.B. & Henderson, R., 2003. Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-particle Electron Cryomicroscopy. Journal of Molecular Biology, 333(4), pp.721–745.

Rubinstein, J.L. & Brubaker, M.A., 2015. Alignment of cryo-EM movies of individual particles by optimization of image translations. Journal of Structural Biology, 192(2), pp.188–195.

Schaefer, L. et al., 2006. Multiple GTPases Participate in the Assembly of the Large Ribosomal Subunit in Bacillus subtilis. , 188(23), pp.8252–8258.

Scheres, S.H.W., 2010. Maximum-likelihood methods in cryo-EM. Part II: application to experimental data. Methods Enzymol, 6879(10), pp.295–320.

Scheres, S.H.W., 2012. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. Journal of Structural Biology, 180(3), pp.519–530.

Schluenzen, F. et al., 2000. Structure of Functionally Activated Small Ribosomal Subunit at 3.3 Å Resolution. Cell, 102(5), pp.615–623.

Schuwirth, B.S., 2005. Structures of the Bacterial Ribosome at 3.5 A Resolution. Science, 310(5749), pp.827–834.

Shajani, Z., Sykes, M.T. & Williamson, J.R., 2011. Assembly of bacterial ribosomes. Annual review of biochemistry, 80, pp.501–26.

Sharma, M.R. et al., 2005. Interaction of Era with the 30S ribosomal subunit: Implications for 30S subunit assembly. Molecular Cell, 18(3), pp.319–329.

De Silva, D. et al., 2015. Mitochondrial ribosome assembly in health and disease. Cell cycle (Georgetown, Tex.), 14(14), pp.2226–50.

Steitz, T.A. & Moore, P.B., 2003. RNA, the first macromolecular catalyst: The ribosome is a ribozyme. Trends in Biochemical Sciences, 28(8), pp.411–418.

Steitz, T. a, 2008. A structural understanding of the dynamic ribosome machine. Nature reviews. Molecular cell biology, 9(3), pp.242–253.

Stokes, J.M. et al., 2014. Discovery of a small molecule that inhibits bacterial ribosome biogenesis. eLife, 3, p.e03574.

Stokes, J.M. & Brown, E.D., 2015. Chemical modulators of ribosome biogenesis as biological probes. Nature chemical biology, 11(12), pp.924–932.

Strunk, B.S. et al., 2012. A Translation-Like Cycle Is a Quality Control Checkpoint for Maturing 40S Ribosome Subunits. Cell, 150(1), pp.111–121.

Sulthana, S. & Deutscher, M.P., 2013. Multiple exoribonucleases catalyze maturation of the 3' terminus of 16S ribosomal RNA (rRNA). The Journal of biological chemistry, 288(18), pp.12574–9.

Tem, S., 2015. K2 Direct Detection Camera. Gatan, pp.3–4.

Thurlow, B. et al., 2016. Binding properties of YjeQ (RsgA), RbfA, RimM and Era to assembly intermediates of the 30S subunit. Nucleic Acids Research, p.gkw613.

Tu, C. et al., 2009. Structure of ERA in complex with the 3□ end of 16S rRNA: Implications for ribosome biogenesis. PNAS, 106(35), pp.14843–14848.

Uicker, W.C., Schaefer, L. & Britton, R.A., 2006. The essential GTPase RbgA (YlqF) is required for 50S ribosome assembly in Bacillus subtilis. Molecular Microbiology, 59(2), pp.528–540.

Wicker-Planquart, C. et al., 2008. Interactions of an essential Bacillus subtilis GTPase, YsxC, with ribosomes. Journal of Bacteriology, 190(2), pp.681–690.

Williamson, J.R., 2003. After the ribosome structures: how are the subunits assembled? RNA, 9(2), pp.165–167.

Wilson, D.N. & Nierhaus, K.H., 2007. The weird and wonderful world of bacterial ribosome regulation. Critical Reviews in Biochemistry and Molecular Biology, 42, pp. 187–219.

Wimberly, B.T. et al., 2000. Structure of the 30S ribosomal subunit. Nature, 407, pp.327–339.

Woodson, S.A., 2008. RNA folding and ribosome assembly. Current Opinion in Chemical Biology, 12(6), pp.667–673.

Wu, S. et al., 2016. Diverse roles of assembly factors revealed by structures of late nuclear pre-60S ribosomes. Nature, 534(7605), pp.133–7.

Wu, S., Armache, J.-P. & Cheng, Y., 2015. Single-particle cryo-EM data acquisition by using direct electron detection camera. Microscopy, pp.1–7.

Wyler, E. et al., 2011. Tandem affinity purification combined with inducible shRNA expression as a tool to study the maturation of macromolecular assemblies. RNA, 17(1), pp.189–200.

Yusupov, M.M. et al., 2001. Crystal structure of the ribosome at 5.5 A resolution. Science, 292(5518), pp.883–96.

Zhang, X. et al., 2014. Structural insights into the function of a unique tandem GTPase EngA in bacterial ribosome assembly. Nucleic Acids Research, 42(21), pp.13430–13439.