Published in final edited form as:

Curr Opin Struct Biol. 2008 June ; 18(3): 299–304.

Biophysical Studies of Bacterial Ribosome Assembly

James R. Williamson

*Departments of Molecular Biology and Chemistry and The Skaggs Institute for Chemical Biology,
The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037*

SUMMARY

The assembly of the bacterial ribosome involves association of over 50 proteins to three large RNA molecules, and it represents a major metabolic activity for rapidly growing bacteria. The availability of atomic structures of the ribosome and the application of biochemical and biophysical methods have led to rapid progress in understanding the mechanistic details of ribosome assembly. The basic steps required to assemble a ribosome are outlined, and the contributions of mass spectrometry, computational methods, and RNA folding studies in understanding these steps are detailed. This complex process takes place with both sequential and parallel processing that is coordinated to ensure efficient and complete of assembly of ribosomes to meet the demands of cell growth.

INTRODUCTION

The bacterial 70S ribosome is composed of the large 50S subunit that harbors the peptidyl transferase activity, and the small 30S subunit that is responsible for the mRNA decoding activity. In rapidly growing bacteria, assembly of new ribosomes for synthesis of cellular proteins is a major metabolic activity that is subject to extraordinary regulatory control to ensure the appropriate and timely synthesis of the many components. The 30S subunit is composed of 16S (~1500 nucleotides (nt) ribosomal RNA (rRNA) and 20 small proteins, and the 50S subunit is composed of 23S (~2300 nt), and 5S (~120 nt) rRNAs and 35 small proteins. To ensure the proper production of this very large and complex macromolecular machine, the assembly process must be extremely efficient.

The basic set of processes that are known to be required for ribosome assembly is outlined in Table I. First, the ribosomal RNAs are transcribed as a single transcript from the rRNA operon, a process that is subject to intricate regulation coupled to growth rate, recently reviewed[1]. In parallel, the ~55 ribosomal proteins are synthesized in a highly coordinated manner, involving both transcriptional and posttranscriptional gene regulation mechanisms[2]. The 16S and 23S rRNAs are modified at ~30–40 specific positions either by base methylation and/or pseudouridylation, recently reviewed[3–5]. Similarly, >10 specific ribosomal proteins are modified in *E. coli*, primarily by methylation and acetylation, recently reviewed [3,6]. These modifications require dozens of gene products, although the role of these modifications in ribosome function is not entirely understood, as many of them are nonessential. Potential roles include stabilizing RNA structure or RNA-protein interactions, mediating translation factor interactions, or roles as checkpoints in ribosome assembly. The processing of the primary

jrwill@scripps.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ribosomal RNA transcript to produce the mature rRNAs requires at least five nucleases, and several of the required processing steps have not yet been assigned to specific enzymes[3,7].

The formation of local secondary structures such as A-form helices and hairpin loops occurs very rapidly, and partial folding of ribosomal RNAs certainly occurs co-transcriptionally. The formation of the active conformation and many of the tertiary interactions between distant sequence elements probably occurs later in the overall assembly process. Protein binding is also thought to occur co-transcriptionally with a defined hierarchy and order of addition. There are 20–30 assembly cofactors that bind at various times, promoting the orderly and efficient process of assembly, recently reviewed[3,4].

The order of the events in Table I only corresponds roughly to the order of events for assembly in cells, although many of these processes are occurring in parallel. All of these events must occur for a ribosome to be completed, but the assembly pathway is stochastic and statistical in nature. Elements of RNA processing, RNA folding, protein binding, and cofactor mediation occur by an ensemble of pathways.

This review will focus on recent progress in using biophysical methods to study ribosome assembly. The majority of the studies in this area can be grouped into three areas that have not been extensively discussed elsewhere in the many recent reviews on the ribosome. First, there is a growing use of mass spectrometry to study ribosome composition and assembly. Second, computational methods have yielded important insights into the energetics of ribosome assembly. Third, biochemical approaches are providing insights into the nature of the series of conformational changes that occur during assembly.

Applications of Mass Spectrometry

There is rapid growth in the application of mass spectrometry (MS) techniques to study the structure and dynamics of the ribosome. Molecular ions for peptide or RNA fragments, intact ribosomal proteins, and even intact ribosomal particles have been observed. The three main types of applications of MS techniques to ribosomes are inventory of the components, identification of modifications of ribosomes, and study of ribosome dynamics.

The ribosome is abundant in bacteria and is readily isolated from cells, making it an excellent candidate for proteomic studies to identify ribosomal components and associated factors. The composition of the ribosomal proteome from *Caulobacter crescentus* was determined using a combination of MS analysis of intact proteins and of proteolytic fragments [8]. The masses of most of the ribosomal proteins were determined directly, and sequences were confirmed by MS-MS analysis of tryptic peptides. The quantitative iTRAQ MS technique was used to examine the inventory of ribosomal proteins and associated factors in an *E. coli* strain carrying mutations in the GTPase assembly cofactor CgtA(E) [9]. Incomplete precursor-50S particles are observed in the mutant strain that lack some late-binding proteins, and the precursor particles also contain a different complement of associated assembly factors. Extending the quantitative iTRAQ approach to analysis of 30S, 50S, and 70S ribosomes, the association of about a dozen proteins with the various ribosomal subunits was quantified [10]. Mass spectrometry has also been applied to quantitate the kinetics of 30S ribosome assembly using an isotope pulse-chase strategy, and the results from this method are discussed in the RNA folding section below [11].

MS has proven to be a powerful tool for the analysis of the protein and RNA modifications that occur during ribosome assembly, but there are only extensive characterizations of a few bacterial species. It is clear that each species has a unique complement of modifications, and despite the availability of many genomic sequences, MS analysis of isolated ribosomes remains the only way to completely characterize the spectrum of modifications for a new species. In

the proteomic study of *Caulobacter crescentus*, some modifications similar to those observed in *E. coli* were observed, as well as some novel species-specific modifications [8]. The ability to map modifications in the ribosomal proteome will permit the study of changes in modifications with growth conditions or genotype, which may elucidate the specific role of ribosomal protein modifications in assembly or translation.

The modifications of ribosomal RNAs are also varied among bacterial species, and an extensive catalog of such modifications is available for bacterial 16S rRNA [12]. MS analysis of RNAs and nucleolytic RNA fragments is not nearly as widely practiced as proteomic MS, and application of this powerful technology to RNA has lagged compared to its application to proteins. Recently, liquid chromatography coupled MS (LC-MS) has been applied to map all of the modifications in *Thermus thermophilus* 16S rRNA [13]. MS analysis has also been used to identify the site of 16S rRNA modification by the *E. coli* gene product YebU, which was previously unannotated [14].

The dynamics of ribosomal proteins has been investigated by a diverse set of approaches that use MS as a readout. First, it is remarkable that intact ribosomes and ribosomal subunits can be directly observed in the mass spectrometer, recently reviewed [15]. This technique can be used to analyze the association of translation factors, and the stability of the association of ribosomal proteins to the rRNA under various conditions. Hydrogen-deuterium exchange (HDX) experiments have been developed to monitor protein dynamics in solution by counting the number of exchangeable hydrogen atoms using MS. Application of HDX to the 70S ribosome revealed which regions of the bound ribosomal proteins were most flexible, and these regions may correspond to functionally important regions during translation [16].

Complementary information was obtained by determining the susceptibility of ribosomal proteins to proteolysis in the 70S ribosome or the isolated 50S and 30S subunits [17]. Surface accessibility of lysine residues in the ribosome was probed using the amine-reactive probe S-methyl thioacetimidate in conjunction with MS analysis of the proteins [18]. Taken together, these chemical approaches offer the opportunity to add a dynamic dimension to the atomic structures now available for the ribosome.

Applications of Computational Methods

With the atomic coordinates of the ribosome in hand, it is possible to apply computational methods to investigate the dynamics and energetics of the ribosome. The large number of atoms in the ribosome presents a significant computational challenge, although all-atom simulations of the 70S ribosome have been carried out [19]. Several recent studies have tackled the issue of 30S ribosome assembly using a range of computational approaches.

The first effort in this area was carried out using a reduced representation, where each residue in the RNA and protein was replaced with a single pseudo-atom for computational efficiency [20]. Using this simplified computational approach, the assembly pathway was investigated using Monte Carlo simulated annealing, beginning with a 16S rRNA model containing secondary structure but no tertiary structure. The secondary structure diagram for 16S rRNA is shown in Figure 1A, highlighting the three structural domains, shown in Figure 1B. The original Nomura assembly map defined the order of addition of the ribosomal proteins during *in vitro* reconstitution [21], and the revised map reorganized according to the structural domains is shown in Figure 1C. Proteins were added sequentially using restraints to guide formation of the native structure. The calculations showed that the flexibility of the protein binding sites roughly correlated with the stage at which the proteins bind, where primary protein binding sites are less flexible than secondary and tertiary protein binding sites. The assembly of the 3'-domain in the Nomura and reversed "anti-Nomura" direction was studied in more detail. Interestingly, the flexibility of the protein binding sites downstream of the primary binding protein S7 was decreased after S7 binding, but no such decrease was observed if the tertiary

binding protein S3 was bound first. This computational study supports the idea that one role of ribosomal proteins in assembly is to gradually reduce the flexibility of the 16S rRNA in a stepwise and defined way.

An implicit solvent model was used to facilitate calculation of the binding free energies of the 30S ribosomal proteins, using a continuum Poisson-Boltzmann model for calculation of the electrostatic interactions [22]. Protein binding energies were calculated for binding of each single protein to isolated 16S rRNA as well as to the 30S subunit lacking that protein. The electrostatic energy of binding of the primary binding proteins to 16S rRNA was much higher than for the later binding proteins. The binding energy calculations correlate well with both the observed 5' to 3' assembly direction and the observed order of primary, secondary, and tertiary protein binding. One caveat to this approach is that RNA folding processes and explicit counterion effects that occur during assembly could not be taken into account, and both these effects will certainly impact the energetics of the electrostatic interactions. Nevertheless, the basic picture that emerges from analysis of the interactions between the proteins and the rRNA during assembly is that the order of assembly is also guided in part by electrostatic binding energy.

An alternative approach to computation of protein binding energies was implemented using the self-consistent pair contact probability (SCPCP) approximation [23]. This computational approach also uses a reduced representation for quasi-molecular dynamics simulations that provide interaction energies as well as entropic information from fluctuations, and offers the advantage that conformational changes can be handled to some extent. In this study, the cooperativity of protein binding energy was investigated by removing each protein, or pairs of proteins, and computing the changes in binding energy for the remaining proteins. The computed energetic couplings between proteins agree well with the cooperative protein binding observed in the original Nomura map, and the energetic coupling between proteins is most often observed within the structural 5', central, or 3' domains, with fewer couplings between domains. The computed energies reflect changes in both the strong electrostatic interactions and the flexibility in the rRNA binding sites.

These three studies all provide informative insights into the energetics of protein binding during assembly, at equilibrium, as reflected by the Nomura assembly map constructed *in vitro*. Bacterial ribosome assembly occurs co-transcriptionally in cells, and the assembly process may also be controlled kinetically, rather than thermodynamically. Application of these powerful computational methods will certainly assist in understanding assembly kinetics once the nature of the intermediates in cells is elucidated.

RNA folding and Ribosomal Protein Binding, *in vitro*

One of the most challenging aspects of understanding ribosome assembly is understanding the many RNA folding processes and protein binding reactions required to produce mature ribosomal subunits. Because the efficiency of reconstitution of 30S ribosomes from purified components is very high, while reconstitution of 50S ribosomes remains challenging, *in vitro* studies have focused on 30S ribosome assembly. Again, these important *in vitro* assembly studies carried out with intact 16S rRNA must be interpreted with the understanding that assembly proceeds co-transcriptionally in cells. The transient nature of some structures formed in precursor 23S rRNAs stands in contrast to the gradual and continuous formation of native structure observed in the majority of *in vitro* studies [24].

Chemical footprinting of RNA is a powerful way to monitor both RNA folding and protein binding reactions. Changes in the accessibility of RNA to chemical probes have long been used to monitor these reactions, and recently, significant new insights have been obtained. Time-resolved hydroxyl radical footprinting was applied to study the magnesium ion dependent

folding of the isolated 5'-domain from *E. coli* 16S rRNA [25]. The domain folds with multiple kinetic phases, but nearly all of the tertiary interactions expected in the complete 30S subunit are formed in the absence of proteins, and the fastest folding regions correspond to the primary protein binding sites. The fast self-folding of the 5'-domain is consistent with its critical role as the first domain to be synthesized during cotranscriptional assembly.

Due to the slow reaction rate of many RNA chemical probes used for footprinting, it is difficult to directly monitor assembly reactions in real time. However, it is possible to assemble and examine the conformation of defined ribonucleoprotein particles that correspond to likely intermediates in assembly. A reconstitution intermediate (RI) with a subset of 15 ribosomal proteins was identified in the original work by Nomura when reconstitution reactions were carried out at low temperatures. The RI intermediate undergoes a conformational change upon heating to form a new structure (RI*) that is competent to bind the remaining 5 proteins and to form 30S subunits. The conformational changes in 16S rRNA have been monitored in this series of intermediates using chemical probes [26,27]. Both protections and enhancements are observed through the series of intermediates, and some of the changes that occur are due to local interactions of the proteins with the RNA, but many are distributed in nearby regions where there are no protein contacts. There is strong evidence for some misfolding in the 3'-domain, which bears the binding sites for most of the post-RI* binding proteins. Chemical probes were also used to monitor the temperature dependent conformational changes in 16S rRNA that occur upon binding of individual primary binding proteins [28]. Again, changes were both local consequences of protein binding and remote consequences of nearby RNA folding. Differences in the accessibility of the RNA to probes at different temperatures for some of the proteins reveals another level of complexity to protein binding reactions. Protein binding with different regions of the rRNA can occur at different times due to intervening RNA folding reactions.

Binding kinetics for most of the 30S ribosomal proteins can be measured simultaneously using an isotope pulse chase assay with an MS readout for quantitation [11]. The times for association of the proteins range from seconds to minutes, as shown in Figure 1C, D, with the order of binding in general agreement with the previously observed 5' to 3' directionality, and primary binding proteins binding faster than tertiary proteins in their respective domains. Most revealing was the temperature dependence of the binding rates which revealed that each ribosomal protein experiences a unique kinetic barrier to binding, and there is no global rate-limiting step or obligatory kinetic intermediate through which all assembly reactions must proceed.

CONCLUSIONS

Biophysical studies of ribosome assembly have been greatly stimulated by the availability of the atomic structure of the ribosome and the application of new and powerful biophysical techniques. In the next few years, we look forward to a deeper understanding of RNA folding and protein binding involved in ribosome assembly. Furthermore, the specific role of ribosome assembly cofactors is now a very active area of investigation, and the many interesting and exciting studies in this area could not be contained in this review. The complex process of ribosome assembly in bacteria has been fine tuned during evolution to be efficient under a wide variety of environmental circumstances. It is intriguing to speculate that the parallel nature of some parts of the assembly process provides the possibility for redundant assembly pathways that may be differentially used under diverse conditions. In an automobile assembly line, each car is assembled in a precisely predetermined order. However, it is apparent that on the ribosome assembly line, sometimes the doors are attached before the wheels, and sometimes after.

Acknowledgements

This work was supported by a grant from the National Institutes of Health, GM-53757 to JRW. The author thanks Anne Bunner, Stephen Chen, William Ridgeway, and Michael Sykes for critical comments on the manuscript.

References

1. Paul BJ, Ross W, Gaal T, Gourse RL. rRNA transcription in *Escherichia coli*. *Annual Review of Genetics* 2004;38:749–770.
2. Zengel, JM.; Lindahl, L. Progress in Nucleic Acid Research and Molecular Biology, Vol 47. 1994. Diverse Mechanisms for Regulating Ribosomal-Protein Synthesis in *Escherichia-Coli*; p. 331–370. *Progress in Nucleic Acid Research and Molecular Biology*, vol 47.]
3. Kaczanowska M, Ryden-Aulin M. Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiology and Molecular Biology Reviews* 2007;71:477. [PubMed: 17804668]
4. Wilson DN, Nierhaus KH. The weird and wonderful world of bacterial ribosome regulation. *Critical Reviews in Biochemistry and Molecular Biology* 2007;42:187–219. [PubMed: 17562451]
5. Chow CS, Larnichhane TN, Mahto SK. Expanding the nucleotide repertoire of the ribosome with post-transcriptional modifications. *Acs Chemical Biology* 2007;2:610–619. [PubMed: 17894445]
6. Polevoda B, Sherman F. Methylation of proteins involved in translation. *Molecular Microbiology* 2007;65:590–606. [PubMed: 17610498]
7. Srivastava AK, Schlessinger D. Mechanism and Regulation of Bacterial Ribosomal-Rna Processing. *Annual Review of Microbiology* 1990;44:105–129.
8. Running WE, Ravipaty S, Karty JA, Reilly JP. A top-down/bottom-up study of the ribosomal proteins of *Caulobacter crescentus*. *Journal of Proteome Research* 2007;6:337–347. [PubMed: 17203977]
9. Jiang MX, Datta K, Walker A, Strahler J, Bagamasbad P, Andrews PC, Maddock JR. The *Escherichia coli* GTPase CgtA(E) is involved in late steps of large ribosome assembly. *Journal of Bacteriology* 2006;188:6757–6770. [PubMed: 16980477]
10. Jiang M, Sullivan SM, Walker AK, Strahler JR, Andrews PC, Maddock JR. Identification of novel *Escherichia coli* ribosome-associated proteins using isobaric tags and multidimensional protein identification techniques. *Journal of Bacteriology* 2007;189:3434–3444. [PubMed: 17337586]
11. Talkington MWT, Siuzdak G, Williamson JR. An assembly landscape for the 30S ribosomal subunit. *Nature* 2005;438:628–632. [PubMed: 16319883]
12. McCloskey JA, Rozenski J. The Small Subunit rRNA Modification Database. *Nucleic Acids Research* 2005;33:D135–D138. [PubMed: 15608163]
13. Guymon R, Pomerantz SC, Crain PF, McCloskey JA. Influence of phylogeny on posttranscriptional modification of rRNA in thermophilic prokaryotes: The complete modification map of 16S rRNA of *Thermus thermophilus*. *Biochemistry* 2006;45:4888–4899. [PubMed: 16605256]
14. Andersen NM, Douthwaite S. YebU is a m(5)C methyltransferase specific for 16S rRNA nucleotide 1407. *Journal of Molecular Biology* 2006;359:777–786. [PubMed: 16678201]
15. Videler H, Ilag LL, McKay ARC, Hanson CL, Robinson CV. Mass spectrometry of intact ribosomes. *Febs Letters* 2005;579:943–947. [PubMed: 15680979]
16. Yamamoto T, Izumi S, Gekko K. Mass spectrometry of hydrogen/deuterium exchange in 70S ribosomal proteins from *E. coli*. *Febs Letters* 2006;580:3638–3642. [PubMed: 16753153]
17. Suh MJ, Pourshahian S, Limbach PA. Developing limited proteolysis and mass spectrometry for the characterization of ribosome topography. *Journal of the American Society for Mass Spectrometry* 2007;18:1304–1317. [PubMed: 17521915]
18. Beardsley RL, Running WE, Reilly JP. Probing the structure of the *Caulobacter crescentus* ribosome with chemical labeling and mass spectrometry. *Journal of Proteome Research* 2006;5:2935–2946. [PubMed: 17081045]
19. Sanbonmatsu KY, Joseph S, Tung CS. Simulating movement of tRNA into the ribosome during decoding. *Proceedings of the National Academy of Sciences of the United States of America* 2005;102:15854–15859. [PubMed: 16249344]
20. Stagg SM, Mears JA, Harvey SC. A structural model for the assembly of the 30S subunit of the ribosome. *Journal of Molecular Biology* 2003;328:49–61. [PubMed: 12683996]

21. Held WA, Ballou B, Mizushima S, Nomura M. Assembly Mapping of 30 S Ribosomal Proteins from *Escherichia coli*. Further Studies. *J Biol Chem* 1974;249:3103–3111. [PubMed: 4598121]
22. Trylska J, McCammon JA, Brooks CL. Exploring assembly energetics of the 30S ribosomal subunit using an implicit solvent approach. *Journal of the American Chemical Society* 2005;127:11125–11133. [PubMed: 16076220]
23. Hamacher K, Trylska J, McCammon JA. Dependency map of proteins in the small ribosomal subunit. *Plos Computational Biology* 2006;2:80–87.
24. Liiv A, Remme J. Importance of transient structures during post-transcriptional refolding of the pre-23 S rRNA and ribosomal large subunit assembly. *Journal of Molecular Biology* 2004;342:725–741. [PubMed: 15342233]
25. Adilakshmi T, Ramaswamy P, Woodson SA. Protein-independent folding pathway of the 16 S rRNA 5' domain. *Journal of Molecular Biology* 2005;351:508–519. [PubMed: 16023137]
26. Holmes KL, Culver GM. Mapping structural differences between 30S ribosomal subunit assembly intermediates. *Nature Structural & Molecular Biology* 2004;11:179–186.
27. Holmes KL, Culver GM. Analysis of conformational changes in 16S rRNA during the course of 30 S subunit assembly. *Journal of Molecular Biology* 2005;354:340–357. [PubMed: 16246364]
28. Dutca LM, Jagannathan I, Grondek JF, Culver GM. Temperature-dependent RNP conformational rearrangements: Analysis of binary complexes of primary binding proteins with 16 S rRNA. *Journal of Molecular Biology* 2007;368:853–869. [PubMed: 17376481]
29. Gutell RR. Collection of small subunit (16S- and 16S-like) ribosomal RNA structures: 1994. *Nucleic Acids Res* 1994;22:3502–3507. [PubMed: 7524024]
30. Wimberly BT, Brodersen DE, Clemons WM, Morgan-Warren RJ, Carter AP, Vonnrhein C, Hartsch T, Ramakrishnan V. Structure of the 30S ribosomal subunit. *Nature* 2000;407:327–339. [PubMed: 11014182]

Figure 1.

Structure and Assembly of the 30S ribosomal subunit. A) Secondary structure of 16S ribosomal RNA [29]. The three major structural domains are color coded: 5'-domain (cyan), central domain (green), 3'-domain (pink), and 3'-minor domain is shown (red). B) Ribbon diagram of 16S rRNA structure from the crystal structure of the 30S ribosomal subunit viewing the subunit interface. [30]. C) Assembly map for the 30S proteins based on the Nomura assembly map [21], reflecting the domain structure of 16S rRNA. The proteins are colored based on the observed rate of protein binding at 15°C using pulse-chase quantitative mass spectrometry [11]. D) Proteins from the 30S subunit, viewed from the solvent face [30], colored according to their binding rates in C) [11].

TABLE I

Processes required for assembly of bacterial ribosomes

1) Transcription of Ribosomal RNA	16S, 23S, 5S rRNA
2) Synthesis of ribosomal proteins	20 "S" proteins, 35 "L" proteins
3) Ribosomal rRNA modification	methylation, pseudouridylation
4) Ribosomal protein modification	methylation, acetylation, glutamylation
5) Ribosomal RNA processing	RNAse III, E, G, P, T
6) Ribosomal RNA folding	
7) Ribosomal protein binding	
8) Assembly factor binding and release	Maturation factors, GTPases, helicases