Published in final edited form as:
Science. 2010 October 29; 330(6004): 673–677. doi:10.1126/science.1193220.

Visualizing Ribosome Biogenesis: Parallel Assembly Pathways for 30S Subunit

Anke M. Mulder[1], **Craig Yoshioka**[1], **Andrea H. Beck**[2], **Anne E. Bunner**[2], **Ronald A. Milligan**[1], **Clinton S. Potter**[1], **Bridget Carragher**[1], and **James R. Williamson**[2]

[1]Department of Cell Biology, The Scripps Research Institute, La Jolla, CA 92037

[2]Departments of Molecular Biology and Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037

Abstract

Ribosomes are self-assembling macromolecular machines that translate DNA into proteins, and an understanding of ribosome biogenesis is central to cellular physiology. Previous studies on the *E. coli* 30S subunit suggest that ribosome assembly occurs via multiple, parallel pathways rather than through a single rate-limiting step, but little mechanistic information is known about this process. Discovery Single-particle Profiling (DSP), an application of time resolved electron microscopy, was used to obtain over 1 million snapshots of assembling 30S subunits, identify and visualize the structures of 14 assembly intermediates, and monitor the population flux of these intermediates over time. DSP results were integrated with mass spectrometry data to construct the first ribosome assembly mechanism that incorporates binding dependencies, rate constants, and structural characterization of populated intermediates.

Ribosome biogenesis in bacteria requires the coordinated synthesis and assembly of 55 ribosomal proteins and 3 large ribosomal RNAs (rRNAs) facilitated by approximately 30 assembly cofactors (1). Remarkably, both the 30S and 50S ribosomal subunits can be assembled *in vitro* from purified components and most of the information required to assemble ribosomes is encoded in the sequences of the component RNAs and proteins (2), (3). The 30S ribosomal subunit is composed of a single ~1500 nucleotide 16S RNA component and 20 ribosomal proteins (`r-proteins, Figure S1). In reconstitution experiments under equilibrium conditions, 30S subunit assembly is both parallel and hierarchical with primary-binding proteins binding independently to each domain, followed by secondary- and tertiary-binding proteins (4). As would be expected based on the co-transcriptional assembly pathway *in vivo* (1), the observed kinetic order for protein binding *in vitro* is that 5'-domain proteins bind fastest, followed by the central domain proteins and 3'-domain proteins (5),(6). While there is evidence for parallel assembly pathways and the existence of multiple assembly intermediates, the structures of such intermediates are not known and little mechanistic information is available for the order of their assembly.

Single-particle EM image analysis can resolve heterogeneous populations of molecules and classify them into homogenous sub-populations that can be visualized in two and three dimensions (7,8). Here we have used single-particle reference-free alignment and classification to “discover” sub populations of assembly intermediates within the heterogeneous assembly reaction over a time course (Figure 1A and Supplementary Discussion). After synchronous initiation of 30S ribosome assembly (4,9,10), there is a

Corresponding Author: James R. Williamson 10550 N Torrey Pines Rd MB-33 La Jolla, CA 92037 Phone: (858) 784-8740 Fax: (858) 784-2199 jrwill@scripps.edu.

distribution of assembly intermediates at various stages of RNA folding and protein binding (6). Negative stain sample preparation rapidly traps these intermediate complexes at various time points, and high throughput EM data collection and analysis allows the visualization of these intermediates without the need for biochemical purification. Classification of the entire time course dataset allows determination of the population, conformation, and protein composition of the prevalent assembly intermediates as a function of time; we call this combined approach Discovery Single-particle Profiling (DSP).

Datasets were collected at time points ranging from time 0 to 120 minutes, and the resulting array of subpopulations indicate that, on average, particle sizes start from small structures and become increasingly 30S-like over time (Figure 1B). Notably, sub-populations that are significantly populated at early time points (Figure 1B, 2 min, column 1, yellow) are less populated at later time points (Figure 1B, 60 min, column 9, blue), whereas sub-populations that are significantly populated at later time points (Figure 1B, 120 min, column 1, red) are absent during early time points (Figure 1B, 2 min). Classification of the data (Figures S2–3) revealed four major groups of assembly intermediates and identified 14 distinct assembly intermediates distributed amongst these four groups (Figure 2). Three dimensional structures were visualized for the set of 14 intermediates, and the compositions of individual intermediates were determined by 3D difference mapping, comparison to the x-ray crystal structure, and sub-complex reconstitution experiments (11) (Figures S5–9). Reconstitution experiments, in which known protein components were omitted from the assembly reaction, confirmed that compositional differences between assembly intermediates were not due to negative stain artifacts (also see Supplementary Discussion). The use of random conical tilt (RCT) reconstruction and 3D difference mapping also confirmed that conformational differences between assembly intermediates were real, and not due to misinterpretation of 2D projection views. The structures of the set of 14 assembly intermediates categorized into early (Group I, red), mid (Groups II and III, green and blue), and late (Group IV, purple) stages of assembly are shown in Figure 2.

Using DSP, the entire collection of particles from all time points was classified simultaneously, and the fractional contribution of particles observed at each time point to the four assembly intermediate groups was determined, resulting in single-particle profiles of sub-population flux over time (Figure 3A). Group I classes are primarily populated at early time points, Group II–III classes are populated at mid time points, and Group IV sub-populations are primarily populated at late time points. To draw a correspondence between these population profiles from DSP and protein binding kinetics *in vitro*, the rate of binding of the 30S proteins was measured under identical conditions, using the Pulse-Chase Quantitative Mass Spectrometry (PC-QMS) assay (6) (Figure 3B). The appearance of the Group IV particles exactly tracks binding of the slowest tertiary binding proteins monitored by PC-QMS (Figure 3B, purple circles). Similarly, the depletion over time of Group I sub-populations tracks binding of the fastest primary and secondary binding proteins (Figure 3A–B, red), with Groups II and III tracking binding of the proteins with intermediate binding rates (Figure 3A–B, green and blue). Thus, Group I sub-populations are ultimately converted into Group IV sub-populations, with Groups II and III bridging the assembly landscape between these two extremes. Slow elimination for Group II intermediates and slow accumulation of Group III intermediates is in accordance with the lag observed after assembly of the 5' and central domains *in vitro* (Figure 3B, red vs. green and blue). Thus, these sub-populations, which reflect an assembly landscape of transient intermediates, eventually converge to the native 30S ribosome structure.

Comparison of the 3D assembly intermediate structures to that of the fully formed 30S subunit readily identified known features of small ribosome subunit architecture (Figure 2, inset). Group I fit the body domain of the 30S crystal structure containing primary r-proteins

S4, S17, S20 and secondary r-protein S16. Diffuse density above the structure corresponds to an assembling platform domain in which the r-protein heterodimer S6–S18 is likely transiently bound and central domain rRNA is only partially structured (Figure 2, red, and Figure S8C). Group II intermediates include an assembled body and platform domain with various stages of early head domain assembly (Figure 2, green, and Figure S8C). 3D difference mapping reveals a conformational change in the neck of the 16S rRNA central domain of Group II intermediates relative to Group IV intermediates (Figure 2, arrow and Figure S6B). This conformational change involves a rotation of the body domain relative to the platform domain and reconstitution experiments suggest that this rRNA conformational change is stabilized, but not actively induced, by binding of 3' domain r-proteins (Figure S7). This conformation of the head domain is also seen in volume `c' of Group III intermediates, whereas Group III volumes `a' and `b' have compact shapes that include partially assembled head and body domains with missing density at the spur and the platform domains (Figure 2, blue).

Group IV sub-populations are distinctly 30S-like in appearance, and vary primarily in the second half of the head domain and at the head-body and head-platform interfaces (Figure 2, purple). Fitting of the 30S crystal structure into 3D difference maps suggest that missing density in the platform domain corresponds to S21 and/or S11 and surrounding (H23/H24) rRNA (Figure S6). Notably, the neck of the central domain for Group IV intermediates has converted from that observed for Group II intermediates, thereby allowing room for the additional S21/S11 density. Defined sub-complex reconstitution studies indicate that conversion from the conformation present in Group II to that in Group IV occurs in the presence of S21/S11, but that the associated extra density is not visible until late 3' domain assembly (Figure S7). The structural and compositional diversity of Group IV assembly intermediates is evidence for the existence of parallel assembly pathways in the 3' domain of the 30S subunit.

Decomposition of the kinetic profiles for sub-populations within Group IV show fast accumulation for classes with density corresponding to the S2-H44 region of the 30S subunit, and slow accumulation for classes lacking this density (Figures 3C, S6). In contrast to the observed thermodynamic dependence of S2 binding upon prior binding of S3 (Figure 4A) (4), the S2-H44 region is able to form stably enough for EM visualization before consolidation of the S3-H34/H33 region (Figures S6,8). Kinetic data for S3 binding during 30S assembly indicate a slow-binding constant relative to primary and secondary r-proteins, but association constants for S2 have not been measureable using PC-QMS, presumably due to a weak and transient interaction (6, 12). Fast protections of portions of both the S2 and S3 binding site are observed in time-resolved hydroxyl radical foot printing experiments (13) consistent with formation of transient encounter complexes for these two proteins. The time-resolved EM data suggest independent binding of these two late binding proteins, thus directly illustrating the existence of parallel assembly pathways.

The combination of thermodynamic dependency maps, ensemble kinetic data, and single-particle profiling with structural analysis allows construction of an expanded assembly mechanism that accounts for both thermodynamic and kinetic interdependencies (Figure 4). The most populated branch of assembly starts with Group I intermediates containing fast-binding primary and secondary r-proteins, proceeds to Group II intermediates containing slower binding primary and secondary r-proteins that bind in central and early 3' domains, and culminates in the set of Group IV intermediates resulting from parallel pathways for assembly of slow-binding tertiary r-proteins (Figure 4B). Tertiary rRNA structure in the absence of bound protein is not readily visualized for early assembly intermediates, which is likely due to a combination of unfolded/conformationally heterogeneous rRNA structure and difficulty in directly visualizing RNA using negative stain. Nevertheless, the significant

energetic gains achieved by rRNA folding and the requirement for cooperativity with r-protein binding is illustrated by the rRNA neck conformational change that mediates transition from Group II to Group IV intermediates (Figure 4B `Δ'). Significantly, the neck is the site of interaction for the assembly cofactor *Era*, which has been implicated in facilitation of global rRNA conformational changes during late central and early 3' domain assembly. *Era* is thought to act in concert with cofactor *RimM*, which itself interacts with S19, an r-protein significantly implicated in kinetic cooperativity during 3' domain assembly (14,15). We speculate that this rotated conformation is required for productive binding by slow-binding tertiary proteins such as S10, S14, S2, S3, and S21. The independent order of binding by these proteins in the Group IV intermediates is evidence for parallel assembly in the 3' domain to a much greater extent than previously appreciated. Population flux studies show rapid accumulation of S2-containing intermediates relative to S3-containing intermediates, even though S3 is thermodynamically required for S2 binding. Rapid initial binding by S2 results in kinetically trapped intermediates with slow subsequent binding rates of S3, leading to slow conversion to productively assembled 30S subunits (Figure 4B, dashed arrows). Such kinetic traps are likely avoided *in vivo* by the participation of assembly cofactors such as *RimP*, which is thought to facilitate r-protein binding during the very late stages of assembly (14). Several less populated intermediates in Group III consist of assembled body and head domains with unassembled platform and spur domains, likely resulting from improper assembly initiation due to the presence of the entire 16S rRNA. While it is possible that these intermediates could convert to productive Group IV intermediates through conformational changes or protein association and dissociation, the population profiles suggest that these subpopulations are likely dead-end intermediates in *in vitro* assembly reactions (Figure S4). Another potentially unproductive pathway branches from the major productive pathway at the point of conversion from Group II to Group IV intermediates and culminates in an end-product that is compressed in structure and missing key r-proteins. The co-transcriptional nature of 30S assembly and the presence of cofactors likely minimize the formation of such unproductive intermediates *in vivo*. Nevertheless, our structural and temporal assembly map shows specific evidence for parallel assembly pathways in vitro and likely recapitulates many of the same steps undergone *in vivo*. The Discovery Single-particle Profiling (DSP) method used for this work is a novel approach for understanding structure and dynamics that can be applied to a wide range of large biologically important complexes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health (RR175173, R37-GM-53757, and GM-52468) and the National Science Foundation (Graduate Research Fellowship to AMM). The electron microscopy data collection and analysis were conducted at the National Resource for Automated Molecular Microscopy, which is supported by the National Institutes of Health through the National Center for Research Resources P41 program (RR175173, RR023093). We thank Neil R. Voss and members of the AMI and Williamson laboratories for critical discussion of experiments.

References and Notes

1. Kaczanowska M, Ryden-Aulin M. Microbiology and Molecular Biology Reviews. SEPT.2007 2007.
2. Traub P, Nomura M. Proc Natl Acad Sci U S A. 1968; 59
3. Nierhaus KH, Dohme F. Proc Natl Acad Sci U S A. Dec.1974 71:4713. [PubMed: 4612527]
4. Mizushima S, Nomura M. Nature. Jun 27.1970 226:1214. [PubMed: 4912319]

5. Powers T, Daubresse G, Noller HF. J Mol Biol. Jul 20.1993 232:362. [PubMed: 8345517]
6. Talkington MW, Siuzdak G, Williamson JR. Nature. Dec 1.2005 438:628. [PubMed: 16319883]
7. Frank, J. Three-Dimensional Electron Microscopy of Macromolecular Assemblies. Oxford University Press; New York: 2006.
8. Zhang W, Kimmel M, Spahn CM, Penczek PA. Structure. Dec 10.2008 16:1770. [PubMed: 19081053]
9. Serdyuk I, Agalarov S, Sedelnikova S, Spirin A, May R. Journal of Molecular Biology. 1983; 169:409. [PubMed: 6620384]
10. Nomura M. Science. Mar 2.1973 179:864. [PubMed: 4569247]
11. Mandiyan V, Tumminia SJ, Wall JS, Hainfeld JF, Boublik M. Proc Natl Acad Sci U S A. Sep 15.1991 88:8174. [PubMed: 1896466]
12. Bunner AE, Beck AH, Williamson JR. Proc Natl Acad Sci U S A. Mar 23.107:5417. [PubMed: 20207951]
13. Adilakshmi T, Bellur DL, Woodson SA. Nature. Oct 30.2008 455:1268. [PubMed: 18784650]
14. Bunner AE, Nord S, Wikstrom PM, Williamson JR. J Mol Biol. Apr 23.398:1. [PubMed: 20188109]
15. Sharma MR, et al. Mol Cell. Apr 29.2005 18:319. [PubMed: 15866174]
16. Ruiz T, Radermacher M. Methods Mol Biol. 2006; 319:403. [PubMed: 16719366]

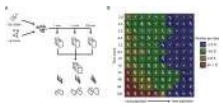

Figure 1. Discovery Single-particle Profiling (DSP)

(A) Assembly of the 30S subunit was initiated as described (9), and aliquots were prepared for negative stain EM (16) at various time points. The DSP method was used to discover sub populations of assembly intermediates from a merged time course dataset of a single assembly reaction (see Supplementary Discussion). (B) Each row of the array results from a single time point, and particle averages are ranked from left to right by decreasing population with a heat map overlaid according to percentage of particles that populate a given class; red: 8–11%, yellow: 6–8%, green: 4–6%, blue: 2–4%.

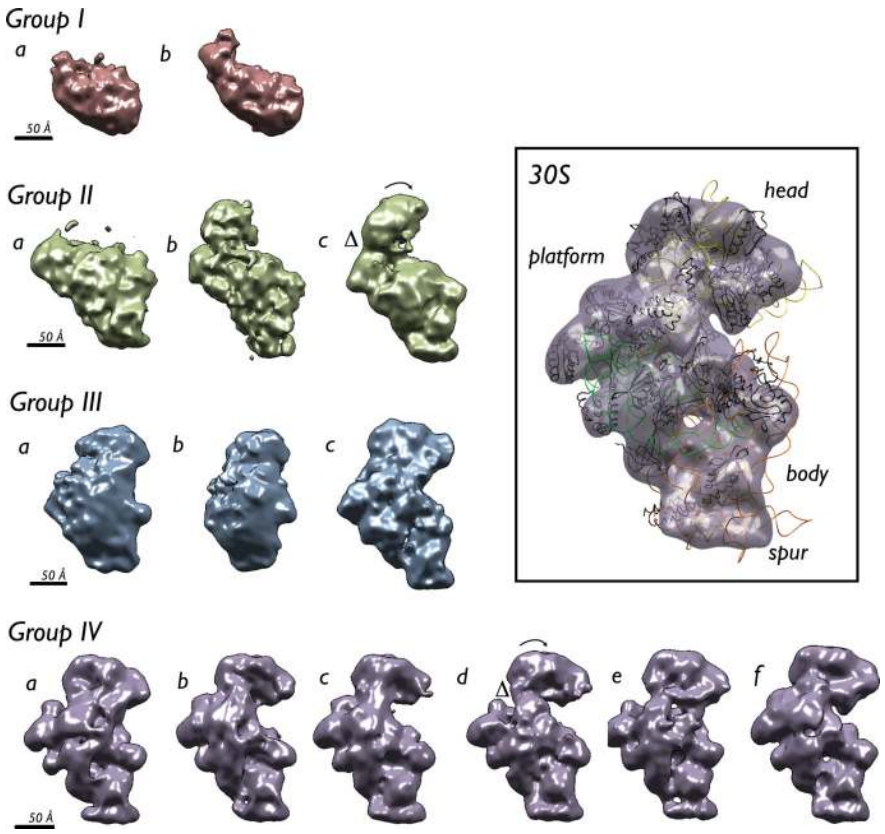

Figure 2. Structural Determination of Assembly Intermediates

Assembly intermediates classify into four groups based on classification of 2D data (Figures S2–3,S5), docking of the 30S subunit crystal structure (Figure S8B), and 3D difference mapping (Figure S6). RCT volumes are color coded in accordance with state of assembly where red < green < blue < purple. A conformational change in the neck of central domain rRNA causing a rotation in the head domain is the major difference between Group II and Group IV intermediates (Δ and arrow).

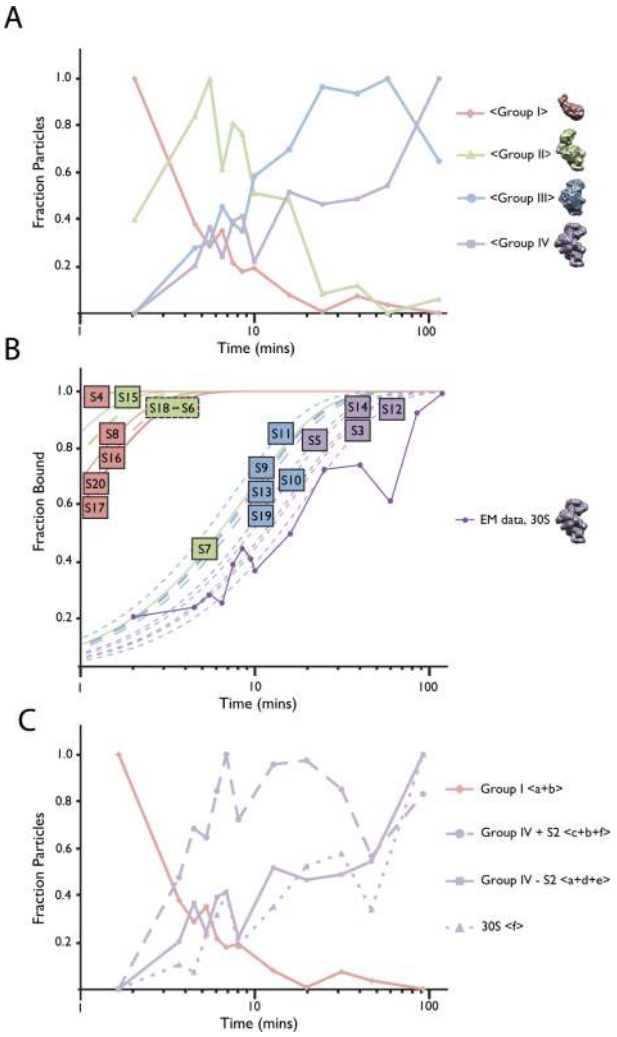

Figure 3. Assembly Intermediate Population Flux over Time

(A) Classification of merged time point datasets allowed determination of the contribution of different time points to assembly intermediate groups I (red), II (green), III (blue) and IV (purple). (B) Measurement of protein binding rates via PC-QMS illustrates the correspondence between population profiles determined via DSP and *in vitro* 30S subunit assembly kinetics. (C) Decomposition of population flux for individual assembly intermediates within group IV (purple) revealed rapid accumulation of S2-containing assembly populations, suggesting the existence of an S2-mediated kinetic trap in the assembly mechanism.

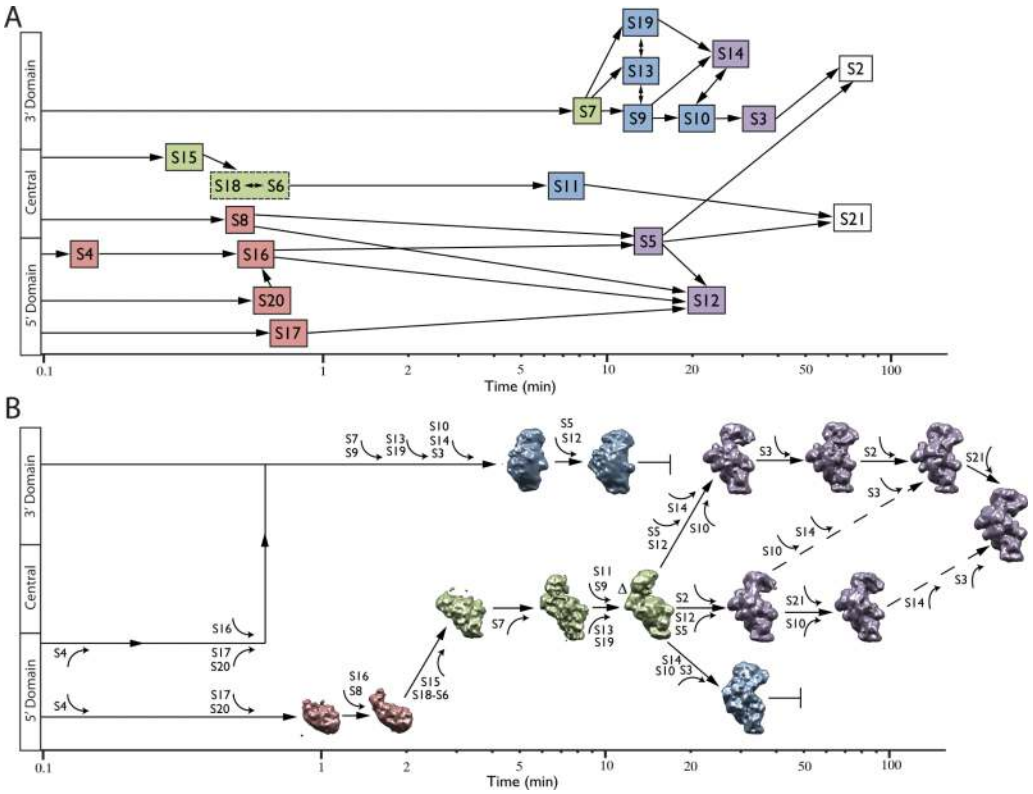

Figure 4. A Parallel Mechanism for 30S Subunit Assembly

(A) A Nomura assembly map on a time axis that takes into account the $t_{0.5}$ of binding for each r-protein as determined via PC-QMS. (B) Combination of kinetic, thermodynamic, and single-particle data allows construction of a mechanism for 30S subunit assembly that accounts for parallel assembly pathways.