Multiple in vivo pathways for Escherichia coli small ribosomal subunit assembly occur on one pre-rRNA

Neha Gupta¹ & Gloria M Culver¹,²

Processing of transcribed precursor ribosomal RNA (pre-rRNA) to a mature state is a conserved aspect of ribosome biogenesis in vivo. We developed an affinity-purification system to isolate and analyze in vivo–formed pre-rRNA–containing ribonucleoprotein (RNP) particles (rRNPs) from wild-type E. coli. We observed that the first processing intermediate of pre–small subunit (pre-SSU) rRNA is a platform for biogenesis. These pre-SSU–containing rRNPs have differing ribosomal-protein and auxiliary factor association and rRNA folding. Each RNP lacks the proper architecture in functional regions, thus suggesting that checkpoints preclude immature subunits from entering the translational cycle. This work offers in vivo snapshots of SSU biogenesis and reveals that multiple pathways exist for the entire SSU biogenesis process in wild-type E. coli. These findings have implications for understanding SSU biogenesis in vivo and offer a general strategy for analysis of RNP biogenesis.

Ribosome biogenesis, the production of mature ribosomal subunits, is essential in all living organisms. Ribosome biogenesis requires proper coordination of transcription, processing, modification and structural rearrangement of ribosomal RNA, ribosomal protein (r protein) synthesis, modification and association with rRNA to form functional ribonucleoprotein complexes (RNPs)¹. The process of forming mature rRNA by the action of several nucleases on precursor rRNA (pre-rRNA) is conserved in all domains of life. Although these events in eukaryotes have been studied in detail², such studies in prokaryotes have been lacking, and this lack of knowledge has hampered understanding of an essential cellular process. Moreover, because changes in bacterial ribosome biogenesis have recently been linked to altered virulence and drug resistance in pathogens³–⁶, understanding this process is of particular importance in the discovery of new antimicrobials⁷.

Like most cellular RNAs, rRNAs are transcribed as primary transcripts, which undergo maturation to a functional state⁸. In all domains of life, a primary rRNA transcript is processed to release fragments containing the individual RNA species. Early pre-rRNA species contain mature rRNA sequences and additional spacer sequences at both ends (5′ and 3′) (ref. 9). In E. coli, seven rDNA operons encode the three rRNAs (16S (SSU), 23S (large subunit (LSU)) and 5S (LSU) rRNAs). Cleavage by RNase III (an endonuclease for double-stranded RNA) is believed to be the first step in the release of pre-rRNAs¹⁰–¹¹. RNase III action on pre-16S rRNA resolves a pre-rRNA product, 17S rRNA, with 115 nucleotides upstream of the mature 5′ end (leader) and 33 nucleotides downstream of the mature 3′ end (trailer) (Fig. 1a)¹². 5′-end maturation involves two cleavage events catalyzed by RNases E and G¹³–¹⁴. Maturation of the 3′ end of 16S rRNA is less well understood because several distinct RNases have been linked to this process¹⁵–¹⁶. Thus, although 16S rRNA–maturation enzymes and pre-rRNA processing intermediates of different lengths have been identified, the substrates, timing of enzyme action and coupling of 16S rRNA maturation to other SSU biogenesis events are not well understood.

An integrated understanding of critical in vivo bacterial ribosome biogenesis events is still lacking after nearly five decades of study. Mutant strains and chemical perturbations have yielded some information on the SSU biogenesis pathway and intermediates in E. coli¹⁷–¹⁹. Small populations of ribosomes undergoing biogenesis (2–5%), rapid kinetics and asynchronous transcription from rRNA loci greatly complicate the identification and characterization of bacterial SSU biogenesis intermediates²⁰. To overcome these deficiencies, we developed an RNA-tagging and affinity-purification system to isolate SSU assembly intermediates from wild-type cells. This system allowed purification of SSU intermediates that were essentially free of contaminating mature SSUs, thus allowing detailed characterization of the pre-SSUs. One pre-SSU rRNA (17S rRNA) was the predominant species associated with all purified assembly intermediates, yet these 17S rRNA–containing RNPs had distinct rRNA architectures and protein compositions. These findings suggest that discrete RNPs form on a 17S rRNA platform during SSU biogenesis in vivo and that ‘snapshots’ of SSU biogenesis can be characterized with this strategy. Although pre-SSU RNPs were distinct, they lacked mature architecture at several functional sites, thus indicating the preclusion of pre-rRNPs from involvement in translation. We observed several independent pathways for maturation of 17S rRNA in these RNPs, results suggesting that flux through SSU biogenesis involves multiple pathways. Additionally, these data suggest that similar analysis would be productive if applied to most dynamic cellular RNPs.

RESULTS

Purification of SSU assembly intermediates formed in vivo

We devised a strategy for purification of pre-16S rRNA–containing SSU assembly intermediates by using an MS2 bacteriophage RNA
The change in the growth rate of tagged plasmids compared to nontagged plasmid is different positions and a nontagged plasmid as the sole source of rRNA. WT, wild-type. Average values from three independent experiments are shown.

chromosomal rDNA copies 24, were able to complement growth at 16S rRNA, and it was previously used to isolate Tag 86M is 86 nucleotides into the mature leader sequence (Fig. 1b).

Tag 20T is 20 nucleotides into the trailer.

Tag 11L is 11 nucleotides into the leader.

for 105L, 11L, 20T and 86M. Tag 105L is 105 nucleotides into the leader and between the trailer base-pair to form a double-helical structure that contains the site for RNase III cleavage. The numbers represent the nucleotide positions from the mature ends.

(b) Sites of MS2-tag (MS2) insertion for 105L, 11L, 20T and 86M. Tag 105L is 105 nucleotides into the leader and between the RNase III and RNase E cleavage sites. Tag 11L is 11 nucleotides into the leader and between the RNase E and RNase G cleavage sites. Tag 20T is 20 nucleotides into the trailer. Tag 86M is 86 nucleotides into the mature 16S rRNA, and it was previously used to isolate mature mutant ribosomes (named Spur)21; it is used here as a positive control, pl., lambda phage promoter.

(c) Growth measurements of the Δ7rrn strain containing an rDNA plasmid with MS2 tags at different positions and a nontagged plasmid as the sole source of rRNA. WT, wild-type. Average values from three independent experiments are shown. The change in the growth rate of tagged plasmids compared to nontagged plasmid is ±5%.

Because disruption of ribosome biogenesis or rRNA maturation can result in slow growth and cold sensitivity18,23, we determined whether the MS2-tag insertions affected SSU biogenesis and therefore growth. The three MS2 rDNA–tagged plasmids (105L, 11L and 20T), each acting as the sole copy of rDNA in an E. coli strain lacking all chromosomal rDNA copies24, were able to complement growth at levels indistinguishable from those of wild-type, nontagged constructs. These results indicate that insertion of these tags does not affect ribosome biogenesis (Fig. 1c).

A second independent in vivo assay demonstrated that expression of 105L, 11L and 20T tagged rDNA plasmids did not affect growth at nonpermissive temperatures; thus, these MS2-tagged rRNAs do not result in the cold sensitivity that is a hallmark of ribosome biogenesis defects (Supplementary Fig. 1b).

In contrast, an rDNA plasmid with an MS2 tag at position 86L exhibited cold sensitivity; this result indicates that defects associated with MS2 tagging of pre-16S RNA can be observed in our assay, but given this defect, we did not undertake further work with this construct. Moreover, we detected 16S rRNA derived from plasmids expressing untagged pre-16S rRNA or pre-rRNA tagged at positions 105L, 11L or 20T in 70S ribosomes at similar levels (Supplementary Fig. 1c and data not shown). These experiments also revealed no discernible change in ribosomal subunit distribution when pre-16S rRNA containing the MS2 tag or mature 16S rRNA was plasmid derived.

Figure 1 Affinity purification of SSU intermediates. (a) Schematic of 16S rRNA maturation. Precursor 16S-rRNA is cleaved by several RNases to form mature 16S rRNA. Cleavage sites and RNases involved are indicated. Asterisk indicates polynucleotide phosphorylase, RNase PH, RNase R, RNase II and YbeY, the RNases implicated in 3′-end maturation of 16S rRNA13–16. The leader and trailer base-pair to form a double-helical structure that contains the site for RNase III cleavage. The numbers represent the nucleotide positions from the mature ends.

(b) Sites of MS2-tag (MS2) insertion for 105L, 11L, 20T and 86M. Tag 105L is 105 nucleotides into the leader and between the RNase III and RNase E cleavage sites. Tag 11L is 11 nucleotides into the leader and between the RNase E and RNase G cleavage sites. Tag 20T is 20 nucleotides into the trailer. Tag 86M is 86 nucleotides into the mature 16S rRNA, and it was previously used to isolate mature mutant ribosomes (named Spur)21; it is used here as a positive control, pl., lambda phage promoter.

(c) Growth measurements of the Δ7rrn strain containing an rDNA plasmid with MS2 tags at different positions and a nontagged plasmid as the sole source of rRNA. WT, wild-type. Average values from three independent experiments are shown. The change in the growth rate of tagged plasmids compared to nontagged plasmid is ±5%.

Figure 2 SSU intermediates purified with all three tags contain 17S rRNA as a major pre-16S-rRNA species. (a) RNA analysis of rRNA purified with tags 105L, 11L and 20T on a 2% denaturing agarose gel stained with ethidium bromide. The three elution fractions with highest amounts of RNA from representative affinity purifications (= 4) are shown. Uncropped images are shown in Supplementary Data Set 3.

(b) Northern blot analysis using probes directed to leader, trailer and mature regions of 16S rRNA for the same fractions as in a. Positions of the probe are indicated in schematic at bottom. Uncropped images are shown in Supplementary Data Set 3. (c) Reverse-transcription (RT)-PCR products of the ligated junctions of pre-16S-rRNA purified with the three tags (modified 3′ RACE; Online Methods), resolved on a 2% agarose gel and stained with ethidium bromide. Only one fraction from the above analysis is shown; similar results were obtained for all elution fractions. † indicates heat treatment of RNA before ligation. The experiments in a–c were carried out at least four times. M, marker.

(d) Modified 3′ RACE products of pre-16S rRNA purified with tag 11L from the Δmrg strain of E. coli, which has been previously shown to accumulate 16.35 rRNA species processed by RNase E at the 5′ end13. 16.35 rRNA is marked. Uncropped image is shown in Supplementary Data Set 3. The identity of the products was confirmed by sequencing (data not shown). The experiments were carried out two times.
**Figure 3** Distinct architecture of the three purified SSU assembly intermediates. (a) Altered nucleotide reactivity of leader in the three purified pre-SSU intermediates, as revealed by kethoxal probing, classified as strong (large circles) and weak (small circles). The hairpin indicates the position of the MS2 tag. (b) Three-dimensional crystal structure (PDB 2AVY) of the SSU, showing nucleotides with altered reactivity in three pre-SSU 17S rRNA-containing complexes purified with different tags and in mature, naked 16S rRNA. (All r proteins are omitted for clarity.) The relative changes in the nucleotide reactivity intensity of the intermediates and naked 16S rRNA (referred to as the starting point in the biogenesis process) are calculated relative to mature SSUs reactivity (as baseline; Online Methods, Supplementary Fig. 3 and Supplementary Data Set 1) and are plotted as circles. Three independent cultures and affinity purification and probing experiments were performed; the nucleotides showing consistent changes in at least two experiments are shown. Residues with >50% changes relative to mature SSUs are shown. The figures were generated with PyMOL (http://www.pymol.org/).

The complementation analysis, lack of cold sensitivity and incorporation of tagged rRNA in mature ribosomes suggest that insertion of the tags at positions 105L, 11L and 20T does not affect ribosome biogenesis. These tagged pre-16S rRNA transcripts allowed affinity purification of pre-SSU RNPs from wild-type *E. coli* with less than 5% contamination with untagged rRNA (Supplementary Fig. 2). Thus, this system offers a robust means to isolate pre-SSU particles from *E. coli*, which are largely unperturbed and growing exponentially.

**17S rRNA is a predominant pre-16S rRNA species in vivo**

Examination of pre-rRNA from the purified RNPs by denaturing gel electrophoresis (Fig. 2a), northern blotting (Fig. 2b), modified 3′ rapid amplification of cDNA ends (RACE) (Fig. 2c) and primer extension (data not shown) revealed that the predominant purified rRNA was 17S rRNA, the pre-SSU species liberated from the primary rRNA transcript by RNase III cleavage. Purification of assembly intermediates tagged at position 11L from an *E. coli* strain lacking RNase G, a 16S rRNA–maturation enzyme, resulted in isolation of substantial amounts of 16.3S rRNA (compared to the 17S rRNA) (Fig. 2d), in agreement with previous reports on pre-rRNA processing in this strain.22,23 These data indicate that affinity purification can yield many pre-16S rRNA–containing RNPs and thus suggest that the abundance of purified 17S rRNA from several tagged positions in wild-type *E. coli* is not due to other species being refractory to purification and that 17S rRNA is a platform for assembly of pre-SSUs in wild-type *E. coli*.

**The 17S RNA-containing RNPs have distinct architecture**

We explored the conformation of 17S rRNA in the three purified RNPs, using chemical probing (kethoxal; Supplementary Data Set 1) and primer extension (data not shown) revealed that the predominant purified rRNA was 17S rRNA, the pre-SSU species liberated from the primary rRNA transcript by RNase III cleavage. Purification of assembly intermediates tagged at position 11L from an *E. coli* strain lacking RNase G, a 16S rRNA–maturation enzyme, resulted in isolation of substantial amounts of 16.3S rRNA (compared to the 17S rRNA) (Fig. 2d), in agreement with previous reports on pre-rRNA processing in this strain.22,23 These data indicate that affinity purification can yield many pre-16S rRNA–containing RNPs and thus suggest that the abundance of purified 17S rRNA from several tagged positions in wild-type *E. coli* is not due to other species being refractory to purification and that 17S rRNA is a platform for assembly of pre-SSUs in wild-type *E. coli*.

**Figure 4** Multiple pathways for ribosomal-protein addition to the three intermediates. (a) Spectral counts of r proteins bound to intermediates compared to mature ribosomes purified with a tag at position 86M (Online Methods and Supplementary Data Set 2). Relative r-protein levels are shown as percentages (two technical replicates from each of three independent cultures and affinity purifications). R-protein percentages of each intermediate are binned into four groups and plotted as heat maps on the crystal structure of SSU (PDB 2AVY).34 The solvent side is shown (interface side in Supplementary Fig. 4a). (b) Hierarchical clustering analysis of r proteins bound to the 105L, 11L and 20T tagged intermediates. Relative protein abundance (RPA) for each r protein present in each of the three SSU intermediates (but not mature SSUs) is calculated as described in Online Methods. Different shades of green indicate the levels of occupancy by a specific r protein among the intermediates, with darker green color indicating more abundant bound r protein in the given intermediate compared to the other intermediates. The r proteins are clustered into three major groups, marked I–III. (c) RPA of r proteins associated with the three intermediates and relative to one another, plotted on the crystal structure of SSU (PDB 2AVY)34. Colors are as in b, and several r proteins are labeled for reference. Solvent surface is shown (interface surface in Supplementary Fig. 4b).
The experiments were carried out three times. (a) Modified 3′5′ RACE products from the three purified intermediates separated on a 2% agarose gel. Arrow, intermediate pre-16S rRNA species observed in all purifications (n = 11). 17S rRNA and intermediate pre-16S RNA are gel purified and PCR amplified, and products (product 1 and product 2) are resolved on a 2% agarose gel. (b) Schematic of sequencing results of the various pre-16S rRNA-processing species associated with the three different purified SSU intermediate products. Each product is designated A–D. The experiments were carried out two times. (c) Modified 3′5′ RACE products resulting from incubation of the purified 105L, 11L and 20T assembly intermediates with wild-type extracts for 0 or 60 min as indicated (in vitro). Plus sign indicates addition of S100. The experiments were carried out three times. Uncropped image is shown in Supplementary Data Set 3. Various pre-16S rRNA species are marked as in b.

intermediates with the same-length pre-rRNA could be due to the availability of MS2 tags for MS2 coat-protein binding as a result of distinct RNP conformations during biogenesis; i.e., the tags could be differentially available during pre-rRNA transcription, folding, processing and incremental protein binding. Notably, leader nucleotides exhibited different reactivity in the three intermediates (Fig. 3a); hence, during biogenesis different conformations exist in intermediates. The reactivity of the MS2-tag nucleotides indicated that stem-loop structure was properly formed, thus suggesting no significant interactions with the pre-rRNA molecule. Moreover, some of these changes are in agreement with prior work showing that the leader (from untagged 17S rRNA) interacts with mature 16S rRNA sequence26,27; this suggests that tagged 17S rRNA can form these conformations and that tagging does not globally perturb rRNA conformation.

Additionally, many other conserved, functional nucleotides of mature 16S rRNA showed altered reactivity in the intermediates compared to mature SSUs, thus indicating that these 17S rRNA–containing intermediates should be unable to enter the translational cycle (Fig. 3b and Supplementary Fig. 3). Some of the observed changes occur at nucleotides involved in pseudoknots28–30 (nucleotides 505, 506, 524, 527, 575 and 917), tRNA interactions31,32 (nucleotides 530, 693, 926 and 1401) and intersubunit bridges33–35 (nucleotides 698, 700, 703, 705, 710, 711, 713, 760, 761, 763, 771, 776, 800, 894, 895, 898, 1415, 1416, 1419 and 1421). In addition, there were many nucleotides that had similar reactivity between the intermediates, and some of these had reactivity similar to that of mature SSUs (representative nucleotides 143–211 (5′ domain of 16S rRNA), 877–885 (central domain of 16S rRNA) and 1220, 1221, 1222, 1278 and 1279 (3′ domain of 16S rRNA)). These results suggest that there are some common folding pathways and that certain regions achieve ‘final’ conformational states early in biogenesis (Supplementary Fig. 3). The probing data suggest that 20T tagged rRNPs most closely resembled mature SSUs, whereas 105L tagged rRNPs were the most distinct from mature SSUs. Additionally, it was clear that the body of the pre-SSUs was more structurally similar to mature SSUs than were the other domains in all three intermediates. These data reveal that the 17S rRNA-containing RNP s are structurally distinct, but all lack the formation of functional sites.

R proteins are differentially associated with the pre-SSUs

To further investigate differences between the 17S rRNA–containing RNPs, we examined r-protein association by label-free liquid chromatography and tandem MS (LC-MS/MS) and spectral counting (Fig. 4 and Supplementary Fig. 4). We used mature ribosomes, purified with an MS2 tag in the mature 16S rRNA (86M), as a normalization standard for r-protein association (Supplementary Data Set 2), to allow direct comparison of r-protein levels in intermediates to mature SSUs. No r protein was present in any of the three 17S rRNA–containing RNPs at levels observed in mature SSUs (Fig. 4a and Supplementary Fig. 4a). R protein S4—which binds to the 5′ domain of 16S rRNA, corresponding to the SSU body—had approximately the same occupancy in all three intermediates and in mature SSUs. Three r proteins, S2, S3 and S21, were highly underrepresented in all intermediates as compared to in mature SSUs (Fig. 4b and Supplementary Fig. 4b). 105L tagged rRNPs had the lowest total r-protein association compared to that in the other two pre-SSUs and mature SSUs. 11L and 20T tagged rRNPs had distinct r-protein association. Thus, distinct r-protein composition corroborates the different conformations observed in probing experiments and further supports that each affinity-purified intermediate is distinct.

To further analyze the differential association of r proteins with purified intermediates and to examine whether r proteins associate in groups with pre-SSUs, we calculated relative protein abundance (RPA; Online Methods) and performed hierarchical clustering analysis (Fig. 4b,c and Supplementary Fig. 4b). This allowed calculation of relative protein association between intermediates (not relative to mature SSUs) in an unbiased, comprehensive manner, so that
r-protein binding to 17S rRNA could be studied in detail. We observed multiple pathways for r-protein association with the body and platform of 11L and 20T tagged rRNPs. R proteins associating with the platform, S8 and S18, were relatively more abundant in 11L tagged RNPs as compared to 20T tagged RNPs, whereas S15 and S11 levels were elevated in 20T tagged RNPs compared to 11L tagged pre-SSUs (Fig. 4b,c). In addition, this analysis revealed three r-protein groups (Fig. 4b and Supplementary Fig. 4c). Group 1 largely included r proteins (with the exception of r protein S4) enriched in 20T tagged pre-SSUs (including r proteins S15 and S11). Group II represented the r proteins that were more abundant in 11L tagged RNPs (including r proteins S8 and S18) than in the other two intermediates. Interestingly, 11L tagged pre-SSUs had increased association with r protein S7 (group II) and several other r proteins (group III) that bind to the SSU head. This analysis reveals multiple pathways of r-protein association and domain assembly on 17S rRNA in wild-type E. coli. Although each intermediate had distinct r-protein association, the clustering analysis (Supplementary Fig. 4c) indicated that overall r-protein addition was in agreement with previous in vitro and in vivo work36–38. Interestingly, late-binding r proteins S5 (group III) and S12 (group II), which play an important part in translational fidelity39 and ribosome assembly40,41, fell into two separate groups and were previously shown to assemble at similar rates40,42. However, S5 aids in the formation of the head region, and it was grouped with most r proteins that bind the head (Fig. 4b), whereas S12 influences SSU body formation. The predominance of 17S rRNA, distinct architecture and r-protein complements in the in vivo–formed intermediates suggest that 17S rRNA can be a major platform for SSU assembly in wild-type E. coli.

Independent 5′ and 3′ end maturation of 16S rRNA in vivo

Our results, and data from other groups, suggest multiple pathways for r-protein association and rRNA folding during SSU assembly in vivo43,44. However, whether further maturation of 17S rRNA to 16S rRNA also follows multiple or obligatory pathways was not known. Modified 3′5′ RACE revealed a predominant 17S rRNA species purified with all three tags; however, minor but detectable products of lengths between 17S rRNA and 16S rRNA were also present (Fig. 2c). These minor products represented additional pre-16S rRNA maturation products (Fig. 5a,b), with either a mature 3′ end (in 105L and 11L tagged intermediates) or a mature 5′ end (in 20T tagged intermediates). These data suggest that further maturation of 17S rRNA can be initiated at either the 5′ or 3′ end of 17S rRNA in wild-type E. coli. Results show that after RNAIII cleavage 17S rRNA processing in vivo can initiate at either end, and, like r-protein addition and rRNA folding, rRNA maturation can follow multiple pathways.

In vitro rRNA maturation of 17S rRNA–containing RNPs

We further evaluated the authenticity of these 17S rRNA–containing intermediates by incubating them with wild-type extracts and examining the resulting rRNA species by modified 3′5′ RACE (Fig. 5c), denaturing agarose gel and northern analysis (data not shown). The 17S rRNA associated with the three different intermediates was processed into several shorter 16S rRNAs with maturation occurring at either the 5′ or 3′ end, or both. The conversion of three different purified 17S rRNA complexes into primarily 16S rRNA suggests that these RNPs are capable of further maturation. The shorter pre-16S rRNA products observed during 17S rRNA maturation to 16S rRNA in vitro were consistent with those observed in purified intermediates in vivo (comparison of Fig. 5c with Fig. 5a, as depicted in Fig. 5b), thus suggesting that 17S rRNA can undergo multiple maturation pathways in vitro, and the purified RNPs are most probably on-path SSU intermediates. Delayed rRNA processing, with accumulation of pre-16S rRNA, upon incubation of these intermediates with cell extracts lacking pre-16S rRNA–processing enzymes, RNase G or RNase E13, indicates that the processing of these purified 17S rRNA–containing RNPs is defined (Supplementary Fig. 5). These data provide evidence that these purified RNPs are bona fide assembly intermediates in wild-type E. coli.

DISCUSSION

We used an RNP affinity-purification strategy to isolate E. coli SSU assembly intermediates formed in a wild-type strain. Our data revealed that 17S pre-rRNA acts as a major scaffold for SSU biogenesis in wild-type E. coli. 17S rRNA–containing RNPs formed in vivo have distinct rRNA folding and r-protein association; these findings are consistent with the multiple SSU biogenesis pathways in vivo. Moreover, we observed additional rRNA maturation events at the 3′ or 5′ ends of 17S rRNA, both in vivo and in vitro, thus indicating that final end maturation can follow multiple pathways. Our results suggest that there are multiple pathways for SSU formation in vivo in wild-type E. coli for all aspects of biogenesis. We have established a system that allows examination of snapshots of RNP formation in vivo and characterization of various biogenesis processes.

A requirement for processing and maturation of cellular RNAs is common throughout all kingdoms of life. Initial processing of double-stranded precursor RNAs is conserved for many cellular RNAs including microRNAs and rRNA6,12,45. For bacterial rRNAs, RNase III, an endonuclease of double-stranded RNA, probably acts early in biogenesis once transcription of pre-16S rRNA is completed and RNase III cleavage sites become available (Fig. 1a); longer pre-rRNA transcripts, uncut by RNase III, are rarely detected in vivo10,11. Our data suggest that this RNase III cleavage product, 17S rRNA, is a platform for biogenesis, and they are consistent with results from mutant bacterial strains in which 17S rRNA accumulation was observed upon disruption of most genes associated with SSU biogenesis17,18. Our data in wild-type E. coli, along with previous data, indicate that 17S rRNA is the major platform for SSU biogenesis.

Maturation of 17S rRNA and stable association of a full complement of r proteins are likely to be rate limiting during SSU biogenesis. In a primary transcript, a double-stranded helix formed by the leader and trailer is required for RNase III cleavage (17S rRNA production). Helical elements might persist in 17S rRNA and could be inhibitory to further 17S rRNA maturation12. Our data suggest that the leader can adopt different conformations at various stages of biogenesis (Fig. 3a). Mutations or deletions in the leader affect SSU formation27,46,47, and the leader can interact with mature 16S rRNA sequences at different stages of biogenesis48. Moreover, cryo-EM studies have failed to detect densities for unprocessed sequences49–51, further suggesting that these sequences might be flexible and multistructured during biogenesis and explaining why there is no defined structure for pre-rRNA elements, unlike for mature 16S rRNA. The different conformations of pre-rRNA could influence the solvent accessibility of the MS2 tag present in this region, thus allowing purification of distinct intermediates with differently positioned tags.

Several lines of evidence suggest that the affinity-purified pre-RNPs are on path. First, insertion of tags in precursor regions of 16S rRNA do not perturb ribosome biogenesis, because these rDNA constructs were able to complement growth at wild-type untagged levels as the sole source of rDNA24 and did not exhibit cold sensitivity. Additionally, rRNA from rDNA plasmids containing tags in precursor regions were incorporated into 70S ribosomes at similar levels as
Untagged plasmid-derived rRNA (Fig. 1c and Supplementary Fig. 1). Moreover, these 17S rRNAs can be converted into 16S rRNA in vitro (Fig. 5). This maturation follows multiple pathways and is dependent on pre-16S rRNA-processing enzymes (Supplementary Fig. 5). 17S rRNA-containing pre-SSU intermediates in single-deletion strains ArgA59, DsrIM50–52 and DebA52 and the double-deletion strain DebA ArgA51 showed several snapshots of SSU assembly with differential levels of r-protein association and distinct architectures suggesting multiple pathways of SSU formation. Our chemical probing and r-protein–addition data are consistent with the changes observed in intermediates that accumulated in these deletion strains. However, our work offers the first analysis, to our knowledge, in wild-type E. coli growing at 37 °C, affording a detailed dissection of SSU biogenesis in vivo.

Chemical probing with kethoxal showed varying degrees of susceptibility of nucleotides to modification (Fig. 3 and Supplementary Fig. 3). One possibility is that each intermediate is composed of a consortium of heterogeneous but related intermediates that yield differential levels of reactivity. Another possibility is that conformational changes occur at different stages yet ultimately achieve a functional structure. This latter possibility is supported by differential r-protein association with 17S rRNA in vitro (Fig. 4 and Supplementary Fig. 4). In addition, strong thermodynamic dependencies observed in vitro and illustrated in the SSU assembly map do not always hold true in vivo (Supplementary Fig. 4c). Several r proteins (S6, S9, S13, S15, S17 and S20) are nonessential in vivo53, whereas they are indispensable for in vitro reconstitution54,55. R protein S15 is essential for assembly of the platform in vitro, but the gene encoding S15 (rpsO) can be deleted in vivo, and SSUs can still appropriately assemble the platform56. Additionally, contrary to predictions from in vitro data, the 5′ domain of 16S rRNA can fold correctly in the absence of r proteins S17 and S20 (ref. 57). The kinetic analysis of r-protein addition in vitro as well as in vivo38,43 and concurrent binding of r proteins to several regions of 16S rRNA during in vitro assembly further suggest the existence of multiple assembly pathways44. We cannot rule out the possibility that some r proteins may be weakly bound in the intermediates and may dissociate during purification. This suggests that these r proteins would exhibit stability differences among intermediates in agreement with induced-fit and multiphasic SSU assembly44. Although it has been shown that r proteins have complex kinetics and that parallel pathways exist, it was unclear whether these parallel pathways were dependent on different pre-rRNA scaffolds or were obligate depending on the pre-rRNA to which they bind during the cascade. Our analysis suggests that 17S rRNA is a major pre-16S rRNA species for these multiple pathways.

We have established that intermediates have distinct conformations and r-protein association, but the general functional regions in the intermediates are not in mature conformations. Sites involved in the process of translation showed altered conformations in isolated intermediates, thus suggesting that 17S rRNA– containing RNPs are not fully functional (Fig. 3b). The nonfunctional state of pre-rRNPs is supported by r-protein data: two translationally important r proteins (S5 and S12) are differentially associated with the distinct pre-rRNPs, thus suggesting that, although distinct pathways are followed, there are similarities in the overall capacity of SSU intermediates. Because bacteria lack nuclear compartmentalization, the retention of the spacer sequences may aid in sequestration of immature SSUs from the translational cycle and may act in SSU quality control, which occurs in the nucleolus in eukaryotes. A U3 boxA-like sequence in the leader has been proposed to act as a chaperone for central pseudoknot formation58. Specific conditions allow 17S rRNA or intermediate pre-16S rRNA to be incorporated into 70S-like ribosomes, and this has detrimental effects on translation19,27. Thus, the presence of 17S rRNA in RNPs with ‘immature’ functional sites in wild-type bacteria suggests a mechanism that allows maturation and function of SSUs in the same cellular environment without deleterious consequences.

Although there are many obvious differences between these purified intermediates, there are also commonalities between them. The length of the purified pre-rRNA is the most obvious (17S rRNA; Fig. 2a). Our chemical-probing data suggest that the SSU body (5′ domain of 16S rRNA) is most similar between all intermediates and mature SSUs (Fig. 3b). A common theme was apparent in which most functional sites are not formed in any of these pre-SSUs. Moreover, r protein S4 is equally represented in all three intermediates, and S4 is generally regarded as an early-binding protein36,55. In contrast, r proteins S2, S3 and S21, which have all been shown to bind late and with slow kinetics38,55,59, are depleted in these intermediates. The relative lack of r protein S21 is of particular note because S21 is important for proper translational initiation19,60. Therefore, although there are apparently multiple pathways for SSU formation in vivo, a common theme has emerged to limit functional site formation until later in biogenesis and thus probably preclude immature subunits from entering the translational cycle.

Our data indicate that ribosome biogenesis occurs via multiple pathways in vivo. Multiple in vivo pathways for r-protein addition and 17S rRNA maturation in wild-type E. coli can be observed (Fig. 5 and Supplementary Fig. 5). Additionally, there are multiple pathways for auxiliary factor action on 17S rRNA– containing RNPs in vivo (N.G. and G.M.C., unpublished data). Thus, several SSU biogenesis processes occur simultaneously in multiple pathways on 17S rRNPs to ultimately form mature, functional subunits. This work highlights the potential of E. coli to adapt under different circumstances and challenges in antimicrobial development. To our knowledge, this is the first work that allows concurrent analysis of RNA processing, r-protein binding and rRNA conformational changes, yielding comprehensive analysis of SSU intermediates and assembly pathways. This work advances understanding of in vivo formation of ribosomal subunits, and this system has the potential to allow study of most other dynamic functional RNPs in vivo.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank D. Ermolenko and E. Phizicky for critical discussions during the preparation of the manuscript as well as the members of Culver laboratory for helpful discussions and technical advice. We thank F. Hagen and the University of Rochester Proteomics Facility for MS analysis. We thank G. Salahura for technical assistance. We thank R. Green (Johns Hopkins University) for 86SM (Spur) and MS2 coat-protein overexpression constructs and M. O’Connor (University of Missouri–Kansas City) for the MC338 strain. This work was supported by US National Institutes of Health grant GM62432 to G.M.C.

AUTHOR CONTRIBUTIONS

N.G. and G.M.C. designed the study and experiments and wrote the manuscript; N.G. performed experiments and analyzed the data.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.
Insertion of MS2 tags in transcribed spacers of 16S rRNA. The pLK35 plasmid containing an tRNA operon under the control of a lambda pL promoter and ampicillin-resistance marker was used for MS2-tag insertion61. A selectable marker C192U (spectinomycin resistance)25,69 in 16S rRNA was introduced by site-directed mutagenesis (QuickChange XL, Stratagene). The spacer sequences of 16S rRNA (leader and trailer) from all seven rRNA operons of E. coli were aligned with ClustalW 2.0 (ref. 64). The variant nucleotides at position 21 in the trailer and 11 in the leader along with additional nonvariant positions in the leader were chosen for MS2-tag insertion (nucleotide numbering as in Fig. 1a). MS2 tagging were placed between nucleotides 105 and 106 (105L), nucleotides 86 and 87 (86L), and nucleotides 11 and 12 (11L) in leader regions and nucleotides 20 and 21 (20T) in trailer according to protocols described previously25. The insertion of the MS2 tag was confirmed by sequencing by GENEWIZ. The plasmid used to transform Escherichia coli (E. coli) strain MRE600 and E. coli strain MRE600 (containing all seven chromosomal rDNA operons) was transformed into the strains with different tagged plasmids was calculated during the exponential phase of growth.

Plasmids containing MS2-tagged RNA operons and pLK35 were transformed in E. coli strain MRE600 (carrying all seven chromosomal rRNA operons). Growth of these strains was monitored and analyzed at 25°C as described above.

Expression and purification of MS2 fusion coat protein. The two mutations V75E and A81G were introduced by site-directed mutagenesis (QuickChange XL, Stratagene) of the MS2 coat protein (MBP-MS2-His6 fusion protein in the pMAL-c2x vector) to prevent oligomerization60. Mutations were confirmed by sequencing (GENEWIZ).

BL21(DE3) carrying pMBP-MS2(V75E A81G)-His6 was grown at 37°C in Terrific Broth with ampicillin (100 µg/ml) until OD600 of 0.6. Overexpression of the fusion protein was induced by 1 mM IPTG, and cells were grown further for 3 h. Cells were pelleted and resuspended in lysis buffer (20 mM NaHPO4, pH 7.0, 500 mM NaCl, and 10 mM imidazole) and lysed with a high-pressure homogenizer. E. coli cells were lysed with 1/100 U per culture volume of amplification-grade DNase I (Sigma Aldrich) for 10 min on ice. The clarified lysate was loaded on a HiTrap chelating column equilibrated with 0.1 M nickel sulfate and purified by fast performance liquid chromatography (GE Healthcare AKTA). The protein was eluted with elution buffer (20 mM NaHPO4, pH 7.0, 500 mM NaCl, and 500 mM imidazole) according to the manufacturer's protocol. The purified MS2 fusion protein was dialyzed overnight in storage buffer (80 mM HEPES, pH 7.4, 1 M KCl, 10 mM MgCl2, and 10% glycerol), diluted to a concentration of 1 mg/ml and stored at -80°C.

Purification of the SSU intermediates. MRE600 cells, carrying plasmids with MS2-tagged RNA operons, were grown at 37°C in LB medium with ampicillin (100 µg/ml) and 0.2% glucose until the OD600 reached 0.6. Cells were cooled, pelleted, resuspended in MBP binding buffer (20 mM Tris-HCl, pH 7.4, 100 mM KCl, 10 mM MgCl2, and 1 mM EDTA) and lysed with a high-pressure homogenizer. E. coli cells were lysed with 1/100 U of amplification-grade DNase I was added, and lysates were incubated on ice for 10 min. The clarified lysates were incubated with amylose beads (New England BioLabs) which were prebound with MS2 fusion protein, for 2 h on a rotating platform at 4°C. The beads were washed with 40 bead volumes of binding buffer and 20 bead volumes of wash buffer (20 mM Tris-HCl, pH 7.4, 200 mM KCl, 10 mM MgCl2, and 1 mM EDTA) and then packed into a column. The beads were washed again with 15 column volumes of wash buffer. The tagged assembly intermediates were isolated with elution buffer (20 mM Tris-HCl, pH 7.4, 500 mM KCl, 10 mM MgCl2, and 1 mM EDTA). Allele-specific primer extension for the C1192U change in plasmid-borne 16S rRNA was performed to check the purity of the isolated intermediates as described previously26,67.

RNA isolation and analysis. The RNA was isolated from the purified MS2-tagged RNPs as described25. 1 µg of the RNA isolated from each elution fraction of the purified RNA-containing assembly intermediates was resolved on a formaldehyde-containing 2% denaturing agarose gel. The gel was visualized with a Bio-Rad VersaDoc. Northern analysis was performed as described previously with the following oligos: mature 16S, 5′-CGCTATCCGCTACGTCG-3′, leader, 5′-CCTGTTGACCTGATGAGCTGG-3′, and trailer, 5′-CAAAGTA CGGTTCTTTTAAG-3′. Modified 3′ RACE was performed as outlined previously49, with the following modifications. 200 ng of purified RNA was used for initial circularization with T4 RNA ligase (New England BioLabs) overnight at 16°C. This step was done with and without a heat-denaturation step. For heat denaturation, RNA was heated at 74°C and cooled quickly before ligation. Reverse transcription was performed with AMV reverse transcriptase (New England BioLabs) for 60 min, and this was followed by PCR amplification. The PCR-amplified products were resolved on a 2% agarose gel and visualized with the Bio-Rad VersaDoc system. To sequence the different 3′ RACE products, the products were gel purified with an Illustra GFX PCR DNA and Gel Band Purification kit (GE Healthcare) and PCR amplified with the same primers as for the initial PCR reaction. The PCR products were resolved on a 2% agarose gel to check the purity and repurified as described above. The purified products were sequenced by GENEWIZ.
the three intermediates were median centered and log$_2$ transformed, subjected to hierarchical clustering analysis by Cluster3.0 (ref. 73) and visualized with Java TreeView74.

**In vitro processing analysis.** The purified intermediates were concentrated and buffer exchanged (20 mM Tris-HCl, pH 7.4, 100 mM KCl, 10 mM MgCl$_2$, 1 mM EDTA, and 6 mM BME). S100 extracts were prepared from MRE600 strain (wild type) along with Δrng and rne$^{de}$ strains of *E. coli*13,14. Purified intermediates were incubated with different S100 extracts (containing 100 µM ATP and GTP) at 37 °C. Aliquots were taken at time points of 0 and 60 min, and the length and identity of pre-16S RNA were analyzed by the methods described above.

**Ribosome sedimentation profile analysis.** Wild-type strains carrying nontagged and tagged plasmids were grown to OD$_{600}$ of 0.5 at 37 °C. The ribosome profiles were obtained as described previously76. Various fractions were collected and ethanol precipitated for RNA extraction and analysis by allele-specific primer extension as described above. Percentages of plasmid-borne rRNA were calculated with ImageJ as described above.

61. Douthwaite, S., Powers, T., Lee, J.Y. & Noller, H.F. Defining the structural requirements for a helix in 23S ribosomal RNA that confers erythromycin resistance. *J. Mol. Biol.* **209**, 655–665 (1989).
62. Sigmund, C.D., Ettayebi, M. & Morgan, E.A. Antibiotic resistance mutations in 16S and 23S ribosomal RNA genes of *Escherichia coli*. *Nucleic Acids Res.* **12**, 4653–4663 (1984).
63. Makosky, P.C. & Dahlberg, A.E. Spectinomycin resistance at site 1192 in 16S ribosomal RNA of *E. coli*: an analysis of three mutants. *Biochimie* **69**, 885–889 (1987).
64. Larkin, M.A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
65. Sun, Q., Vila-Sanjurjo, A. & O’Connor, M. Mutations in the intersubunit bridge regions of 16S rRNA affect decoding and subunit-subunit interactions on the 70S ribosome. *Nucleic Acids Res.* **39**, 3321–3330 (2011).
66. LeCuyer, K.A., Behlen, L.S. & Uhlenbeck, O.C. Mutants of the bacteriophage MS2 coat protein that alter its cooperative binding to RNA. *Biochemistry* **34**, 10600–10606 (1995).
67. Xu, Z. & Culver, G.M. Differential assembly of 16S rRNA domains during 30S subunit formation. *RNA* **16**, 1990–2001 (2010).
68. Al Refaii, A. & Alix, J.H. Ribosome biogenesis is temperature-dependent and delayed in *Escherichia coli* lacking the chaperones DnaK or DnaJ. *Mol. Microbiol.* **71**, 748–762 (2009).
69. Xu, Z. & Culver, G.M. Chemical probing of RNA and RNA/protein complexes. *Methods Enzymol.* **468**, 147–165 (2009).
70. Stern, S., Moazed, D. & Noller, H.F. Structural analysis of RNA using chemical and enzymatic probing monitored by primer extension. *Methods Enzymol.* **164**, 481–489 (1988).
71. Das, R., Laederach, A., Pearlman, S.M., Herschlag, D. & Altman, R.B. SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. *RNA* **11**, 344–354 (2005).
72. Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
73. Eisen, M.B., Spellman, P.T., Brown, P.O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**, 14863–14868 (1998).
74. Saldanha, A.J. Java Treeview-extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248 (2004).
75. Maki, J.A., Schnobrich, D.J. & Culver, G.M. The DnaK chaperone system facilitates 30S subunit assembly. *Mol. Cell* **10**, 129–138 (2002).
76. Connolly, K., Rife, J.P. & Culver, G. Mechanistic insight into the ribosome biogenesis functions of the ancient protein KsgA. *Mol. Microbiol.* **70**, 1062–1075 (2008).