Published online 4 January 2013 *Nucleic Acids Research, 2013, Vol. 41, No. 4* **2609–2620**
doi:10.1093/nar/gks1256

Dissecting the *in vivo* assembly of the 30S ribosomal subunit reveals the role of RimM and general features of the assembly process

Qiang Guo[1], Simon Goto[2], Yuling Chen[1], Boya Feng[1], Yanji Xu[1], Akira Muto[2], Hyouta Himeno[2], Haiteng Deng[1], Jianlin Lei[1,*] and Ning Gao[1,*]

[1]Ministry of Education Key Laboratory of Protein Sciences, Center for Structural Biology, School of Life Sciences, Tsinghua University, Beijing 100084, China and [2]Department of Biochemistry and Molecular Biology, Faculty of Agriculture and Life Science, RNA Research Center, Hirosaki University, Hirosaki 036-8561, Japan

Received August 20, 2012; Revised November 2, 2012; Accepted November 3, 2012

ABSTRACT

Ribosome biogenesis is a tightly regulated, multi-stepped process. The assembly of ribosomal subunits is a central step of the complex biogenesis process, involving nearly 30 protein factors *in vivo* in bacteria. Although the assembly process has been extensively studied *in vitro* for over 40 years, very limited information is known for the *in vivo* process and specific roles of assembly factors. Such an example is ribosome maturation factor M (RimM), a factor involved in the late-stage assembly of the 30S subunit. Here, we combined quantitative mass spectrometry and cryo-electron microscopy to characterize the *in vivo* 30S assembly intermediates isolated from mutant *Escherichia coli* strains with genes for assembly factors deleted. Our compositional and structural data show that the assembly of the 3′-domain of the 30S subunit is severely delayed in these intermediates, featured with highly underrepresented 3′-domain proteins and large conformational difference compared with the mature 30S subunit. Further analysis indicates that RimM functions not only to promote the assembly of a few 3′-domain proteins but also to stabilize the rRNA tertiary structure. More importantly, this study reveals intriguing similarities and dissimilarities between the *in vitro* and the *in vivo* assembly pathways, suggesting that they are in general similar but with subtle differences.

INTRODUCTION

Ribosome biogenesis is a tightly regulated multi-stepped process, assisted by a wide variety of protein factors, such as transcription factors, endoribonucleases, rRNA helicases and chaperones, rRNA and ribosomal protein modification enzymes and assembly factors (1). As to the 30S subunit, early *in vitro* reconstitution experiments (2–6) have demonstrated that active 30S subunits could be formed from purified ribosomal proteins and 16S rRNA in the absence of other cellular components. The *in vitro* assembly occurs very slowly and requires non-physiological conditions, such as high Mg^{2+} concentration, high ion strength and heat shock. In contrast, the assembly of the 30S subunit *in vivo* starts with rRNA primary transcripts (7) and occurs co-transcriptionally (8) in a much more efficient way, underscoring the essential contribution of assembly factors. In recent years, application of new techniques, such as pulse-chase monitored by quantitative mass spectrometry (PC/QMS) (9), time-resolved X-ray footprinting (10) and time-resolved electron microscopy (11), has brought our understanding of the *in vitro* assembly process to a new level, providing a large amount of valuable kinetic and structural information. Together with earlier work [reviewed in (12)], these data have established that the *in vitro* 30S subunit assembly starts from multiple sites on the 16S rRNA (10), following parallel pathways (9–11) and the free energy of the assembly can be represented by a complex landscape (9). More importantly, kinetic data revealed that for several subsets of 3′-domain proteins, the thermodynamic interdependence does not align well with measured kinetic cooperativity (11,13), and at these locations, the *in vitro* assembly often encounters kinetic traps,

*To whom correspondence should be addressed. Tel: +86 01062794277; Fax: +86 01062771145; Email: ninggao@tsinghua.edu.cn
Correspondence may also be addressed to Jianlin Lei. Tel: +86 01062797699; Fax: +86 01062771145; Email: jllei@tsinghua.edu.cn

The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors.

suggesting that assembly factors might be involved in subverting kinetic traps in the assembly landscape (9,11,13).

Over the past two decades, accumulating experimental data, mainly through genetic approaches, has implicated a number of factors, including RbfA, RsgA, KsgA, Era, ribosome maturation factor M (RimM), RimP, RimJ [(reviewed in (12)] and YqeH (14,15), in the maturation of the 30S subunit in bacteria. However, the specific molecular roles of most of these factors remain unclear. Among these factors, RimM was first identified as a factor required for a fast growth in rich medium (16). The gene-encoding RimM (*yfjA*) in *Escherichia coli* is co-localized to the *trmD* operon (17) with genes for ribosomal proteins S16 and L19, and a tRNA methyltransferase (TrmD), a hint that RimM might be directly involved in ribosome-related function. Indeed, deletion of RimM confers a slow growth phenotype (18), with accumulation of 16S rRNA precursors and free 30S subunits (19) as well as reduced level of polysomes (20). RimM associates with free 30S subunit *in vivo* (18,20) and also binds to S19 *in vitro* (20,21). Moreover, suppressor mutations to the Δ*rimM* mutant were found on S13 (18) and suppressor mutations to a *rimM*-Y106AY107A mutant were found on S19, helices 31 and 33b of the 16S rRNA (20).

In this study, we characterize the immature 30S subunits purified from an *E. coli* Δ*rimM* strain biochemically and structurally. Our data indicate that the immature 30S subunits are a collection of assembly intermediates, with the 3′-head domain proteins severely underrepresented, such as S10, S14, S13 and S19. Moreover, protein composition analysis of another category of immature 30S subunits from a Δ*rbfA*Δ*rsgA* strain shows a different spectrum, with much enhanced levels for these proteins, suggesting that RimM promotes the assembly of these slow binding proteins *in vivo*. Structural analysis shows that these Δ*rimM* intermediates also differ largely in rRNA conformation, particularly the rotational position of the 3′-head domain relative to the body domain. An incubation of recombinant RimM with the immature 30S subunits significantly reduces the flexibility of the head domain. More importantly, our data also suggest that the *in vivo* assembly process occurs along multiple pathways in a certain degree as well, and the rRNA maturation is tightly coupled with ribosomal protein binding. The functional depiction of RimM thus illustrates that there are possible checkpoints along the *in vivo* assembly pathways where maturation factors come into play to direct the process to more efficient branches.

MATERIALS AND METHODS

Escherichia coli strains

We used *E. coli* A19 (Hfr, *rna-19*, *gdhA2*, *his-95*, *relA1*, *spoT1*, *metB1*) (22) as the source of the 30S subunit. A19Δ*rimM* is an A19 derivative in which the *rimM* gene is replaced by a short peptide gene containing an FRT sequence, constructed as follows. The kanamycin-resistant marker of a *rimM* disruptant from Keio collection (23), in which the *rimM* gene has been substituted by an FRT-flanked kanamycin-resistant cassette, was transduced into A19 using phage P1vir to produce an intermediate strain. Then, the kanamycin-resistant cassette was removed from the intermediate strain using an FLP expression plasmid pCP20 (24) to produce the A19Δ*rimM* strain. A19Δ*rbfA*Δ*rsgA* is an A19 derivative in which both of the *rbfA* and *rsgA* genes are replaced by a short peptide gene containing an FRT sequence, constructed by transducing *rbfA*::FRT-*kan*-FRT into A19 and removing *kan* using pCP20 and then transducing *rsgA*::FRT-*kan*-FRT into the resulted strain and removing *kan* using pCP20. Sources of *rbfA*::FRT-*kan*-FRT and *rsgA*::FRT-*kan*-FRT are intermediate strains produced during the construction of W3110Δ*rbfA* (25) and the *rsgA*-disrupted strain of Keio collection (23), respectively. Both the A19Δ*rimM* and A19Δ*rbfA*Δ*rsgA* strains were confirmed with polymerase chain reaction (PCR).

Spot assay and ribosome profile

A19, A19Δ*rimM* and A19Δ*rbfA*Δ*rsgA* strains were grown in liquid LB at 37°C to OD 0.8 and diluted to a series of concentrations, 10^0, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4} and 10^{-5}. Three microliters of each dilution was dropped to a LB plate and incubated at 37°C overnight. The cell extracts from the A19, A19Δ*rimM* and A19Δ*rbfA*Δ*rsgA* strains were loaded onto a 10–40% sucrose gradient containing 10 mM $Mg(OAc)_2$ and centrifuged for 3.5 h at 39 000 rpm in a SW41 rotor (Beckman Coulter). The gradients were analyzed with A254 absorbance using a Teledyne ISCO fractionation system.

Immature and mature 30S subunit purification

Escherichia coli cells (A19, A19Δ*rimM* and A19Δ*rbfA*Δ*rsgA* strains) grown in LB medium were harvested, lysed and clarified in opening buffer [20 mM Tris–HCl (pH = 7.5), 150 mM NH_4Cl, 10 mM $Mg(OAc)_2$ and 0.5 mM ethylenediaminetetraacetic acid (EDTA)]. The lysate was loaded onto the top of 5 ml sucrose cushion [20 mM Tris–HCl (pH = 7.5),150 mM NH_4Cl, 10 mM $Mg(OAc)_2$, 0.5 mM EDTA and 1.1 M sucrose] and centrifuged for 18 h at 28 000 rpm in a 70Ti rotor (Beckman Coulter). The resulting pellets were resuspended in binding buffer and centrifuged through a 10–40% sucrose gradient with 10 mM $Mg(OAc)_2$ for 7 h at 30 000 rpm in a SW32 rotor (Beckman Coulter). Fractions containing the immature 30S and 70S peaks were pooled separately and concentrated with buffer changed to binding buffer for the 30S fractions and to separation buffer [20 mM Tris–HCl (pH = 7.5), 150 mM NH_4Cl and 2 mM $Mg(OAc)_2$] for the 70S fractions. The 70S fractions were further centrifuged through a 10–40% sucrose gradient with 2 mM $Mg(OAc)_2$ to get the mature 30S and 50S subunits.

RimM preparation, rRNA extraction and identification of the 3′ and 5′ends of the 17S rRNA

Full details are available in the Supplementary Data.

Pelleting assay

Mature or immature 30S subunits (2.5 pmol) were incubated with 30-fold excess of RimM for 15 min at 37°C in binding buffer. The mixture was then layered

onto a 150 µl sucrose cushion and centrifuged at 96409 rpm for 4 h in a TLA-120.1 rotor (Beckman Coulter). The pellets and the supernatants were separated and 1/2 of total pellets and 1/20 of supernatants were resolved by 12% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE).

Quantitative mass spectrometry

For quantitation of targeted protein, samples with same A260 absorption value were separated on 1D Tricine-SDS–PAGE. Among all ribosomal proteins, S1 was not included in the QMS analysis, because it dissociates readily from the 30S subunits during centrifugation-based purification. The gel bands corresponding to the targeted protein were excised from the gel, reduced with 10 mM of Dithiothreitol (DTT) and alkylated with 55 mM iodoacetamide. Then, in-gel digestion was performed with the sequence grade modified trypsin (Promega) in 50 mM ammonium bicarbonate at 37°C overnight. The peptides were extracted twice with 1% trifluoroacetic acid in 50% acetonitrile aqueous solution for 30 min. The extractions were then centrifuged in a speedvac to reduce the volume. Peptides from different samples were labeled with tandem mass tags (TMT) reagents (Thermo, Pierce Biotechnology) according to the manufacturer's instruction (TMT 127, 129 and 130 for the samples from the A19 mature 30S, A19Δ*rimM* and A19Δ*rbfA*Δ*rsgA* samples, respectively). Briefly, the TMT label reagents were dissolved by anhydrous acetonitrile and carefully added to each digestion products. The reaction was performed for 1 h at room temperature, and hydroxylamine was used to quench the reaction. The TMT-labeled peptides were desalted using the stage tips.

For LC-MS/MS analysis, the TMT-labeled peptides were separated by a 65-min gradient elution at a flow rate of 0.250 µl/min with an EASY-nLCII^{TM} integrated nano-HPLC system (Proxeon), which is directly interfaced with a Thermo LTQ-Orbitrap mass spectrometer. The analytical column was a home-made fused silica capillary column (75 µm ID, 150 mm length; Upchurch) packed with C-18 resin (300 A, 5 µm; Varian). Mobile phase A consisted of 0.1% formic acid and mobile phase B consisted of 100% acetonitrile and 0.1% formic acid. The LTQ-Orbitrap mass spectrometer was operated in the data-dependent acquisition mode using the Xcalibur 2.0.7 software and there was a single full-scan mass spectrum in the Orbitrap (400–1800 m/z, 30 000 resolution) followed by three MS/MS scans in the quadrupole collision cell using the higher energy collision dissociation.

The MS/MS spectra from each LC-MS/MS run were searched against the selected database using an in-house Mascot or Proteome Discovery searching algorithm. Peptides that have XCorr/Charge scores >2.75 for 2+ and 3.0 for 3+ were used for protein identification and MS/MS spectra for all matched peptides were manually interpreted and confirmed. The QMS experiments were repeated for three times and similar results were obtained. For TMT quantification of a specific protein, ratios of 129:127 and 130:127 for each of the ribosomal proteins were examined by Grubbs' test to remove outliers. Ratios of two or more tryptic peptides from the same protein were used to calculate the means and the standard deviations (Supplementary Table S1).

Cryo sample preparation and cryo-electron microscopy

Cryo-grids for the immature 30S subunits were prepared as previously described (26). The grids were examined in an FEI Tecnai F20 microscope operated at 200 kV, and images were recorded at a nominal magnification of 80 000× on a Gatan UltraScan 4000 CCD camera, under low-dose conditions (~20 e-/Å^2). The complex of the immature 30S subunit bound with RimM was formed by an incubation of a 40-fold excess of RimM with the immature 30S subunits at 37°C for 15 min. The grids of the 30S complex were examined in an FEI Titan Krios cryo-TEM operated at 300 kV, and images were collected at a nominal magnification of 59 000× on an FEI Eagle 4k × 4k CCD camera, under low-dose condition. Data collection was done with AutoEMation software package (27).

Image processing

All the micrographs were decimated by a factor of 2. Particle picking was performed using the SPIDER package (28) with a method based on a locally normalized cross-correlation function (29). The resulting particles (125 × 125 in window size, 2.76 and 3.0 Å in effective pixel size, for the 30S and the 30S complex samples, respectively) were manually verified using a method based on correspondence analysis (30). To ensure the performance of the 2D and 3D analysis, particles were further subjected to another round of manual screen, which finally rendered 164 368 and 94 535 particles for the 30S and 30S complex, respectively. The parameters of the contrast transfer function (CTF) were estimated using SPIDER at the micrograph level. Particles were then CTF corrected using the phase-flipping method (31).

2D image classification was performed using a maximum-likelihood approach (32) with the XMIPP software package (33). Particles from both samples were classified into 100 groups in 100 iterations, and the performance of the classification was monitored by log-likelihood function. To facilitate further comparison, class average images were subjected to a multi-reference alignment to 83 2D projections generated from a cryo-EM map of the mature 30S subunit (26), at an angular interval of 15° (Supplementary Figure S2).

3D classification was performed using a 3D maximum-likelihood approach with XMIPP (34). The initial model was generated by low-pass filtering (60 Å) of a cryo-EM map of the mature 30S subunit (26). Particles from the both samples were classified into five groups in 50 iterations, at an angular sampling of 10°. Refinements of the class structures were performed with SPIDER, following the standard reference projection matching procedures (31), with a gradual decrease of the angular step from 15° to 1°. Amplitude correction to the density maps was performed as previously described (26,35). The final resolutions of the refined density maps were estimated with a soft Gaussian mask approach (36,37) using 0.5 cutoff

criterion of the Fourier Shell Correlation (Supplementary Table S2).

Atomic model and temperature map building

The head and body domains of a 30S subunit crystal structure [PDB ID: 3OFA, (38)] were docked into the cryo-EM maps as rigid bodies first using Chimera (39), followed by a flexible fitting method based on molecular dynamics simulation (40) *in vacuo* for 1 000 000 steps with a 0.5-kcal mol^{-1} scaling factor using NAMD (41). To avoid overfitting, ribosomal protein S2 and S3 were removed from the initial model before flexible fitting due to their low occupancies. For class No. 5 of the immature 30S subunits (Supplementary Table S2), all proteins in the head domain were removed and only the rRNA structure was refined. For better comparison, after fitting, S2 and S3 proteins were added back to the fitted structure using their contacting rRNA helices as reference. The 10 models were aligned using the 30S body domain as reference and 10 temperature maps were constructed in PyMOL (42) by calculating the deviation of the 16S rRNA in the fitted models from the mature 30S structure. The scripts used for root-mean-square deviation (RMSD) calculation and temperature map visualization were downloaded from http://pldserver1.biochem.queensu.ca/~rlc/work/pymol/. Chimera and PyMOL were used for graphic visualization.

RESULTS

Construction of a series of *E. coli* A19 strains

RNase I is the major non-specific endoribonuclease localized in periplasm and often found to be in 30S subunit fractions in cell extracts (43). To avoid undesired degradation of the rRNA precursors in the immature 30S subunits during the sample preparation, we chose the RNase I defective A19 strain (22) as the source of the 30S subunits. In this genetic background, we further constructed strains with the *rimM* gene deleted and with both *rsgA* and *rbfA* genes deleted. The two resulting strains grow poorly on LB medium and show an accumulation of free 30S subunits (Figure 1). Interestingly, both the cell growth test (Figure 1A) and the ribosome profile analysis (Figure 1B) show that the deletion of *rimM* is more deleterious. As a result, there is an intermediate peak between the 30S and 50S peaks, probably representing immature 50S precursors caused by globally decreased protein production in the Δ*rimM* strain (Figure 1B).

Compositional characterization of the immature 30S subunits from the A19 Δ*rimM* and Δ*rbfA*Δ*rsgA* strains

RNA gel analysis shows that the rRNAs in the 30S fractions from the Δ*rimM* strain and the Δ*rbfA*Δ*rsgA* strain are 16S rRNA precursors (Figure 2A), indicating that these free 30S subunits are indeed immature 30S particles. Identification of the two sets of 16S rRNA precursors by a previously established 5′3′-rapid amplification of complementary DNA ends (RACE) technique (44) reveals that a majority of these precursors are unprocessed at both the 5′- and 3′-ends (Supplementary Figure S1). The

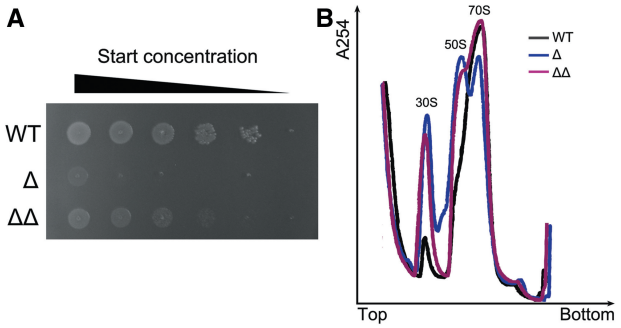

Figure 1. Phenotypes of the Δ*rimM* and Δ*rbfA*Δ*rsgA* strains. (**A**) Spot assay showing that both the Δ*rimM* (Δ) and Δ*rbfA*Δ*rsgA* (ΔΔ) strains grow slowly, compared with the wild-type strain (WT). (**B**) Ribosome profile analysis of the A19, Δ*rimM* and Δ*rbfA*Δ*rsgA* strains. The profile curves of the WT, Δ and ΔΔ strains are colored in black, green and red, respectively. Deletion of RimM or a combination of RbfA and RsgA causes an accumulation of immature 30S subunits. Both experiments indicate that deletion of RimM is more deleterious.

protein gel analysis shows that some ribosomal proteins, e.g. S2 and S3, are underrepresented in the Δ*rimM* sample (Figure 2B). The compositional heterogeneity suggests that the immature 30S subunits from the Δ*rimM* strain are a collection of *in vivo* assembly intermediates that are different in protein composition.

To determine the protein levels, similar to a previously established quantification method (45), we used a QMS technique based on TMT labeling (46). The QMS data reveal that the levels of S21, S10, S14, S13, S19, S3, S2 and S5 are dramatically reduced in the Δ*rimM* sample, <50% of those in the mature 30S subunits (Figure 2C and Supplementary Table S1). Most of them are secondary and tertiary binding proteins from the 3′-head domain of the 30S subunit, except that S21 and S5 are tertiary binder from the central domain and the 5′-domain, respectively. S21 is known to easily dissociate in solution (47) and is therefore not included for further analysis. Thus, these data clearly demonstrate that the deletion of RimM causes a severe delay in the assembly of the 3′-domain of the 30S subunit *in vivo* (Figure 2C and Supplementary Figure S3). Among these 3′-domain proteins, S7 has the highest occupancy (81%) in the Δ*rimM* sample, which is in accordance with the *in vitro* assembly map that S7 is a primary binder and directs the binding of all the rest 3′-domain proteins (48).

In contrast, the protein composition of the immature 30S subunits from the Δ*rbfA*Δ*rsgA* strain shows intriguing difference and similarity (Supplementary Figure S3). The most underrepresented protein in the Δ*rbfA*Δ*rsgA* sample is still S21 (30%), followed by S7, S2, S10, S11 and S19 (49–63%) (Figure 2C and Supplementary Table S1), clearly showing a different pattern. Although many 3′-domain proteins, such as S10, S13, S14, S19 and S3, are also underrepresented, their levels are significantly higher than those in the Δ*rimM* sample (Figure 2D and Supplementary Table S1). In fact, the Δ*rbfA*Δ*rsgA* sample has a higher level for almost all the proteins, compared with the Δ*rimM* sample (Figure 2C and Supplementary Figure S3). For example, S10, S13 and S14 have an over 2-fold increase and S21, S19 and S3 have a moderate increase,

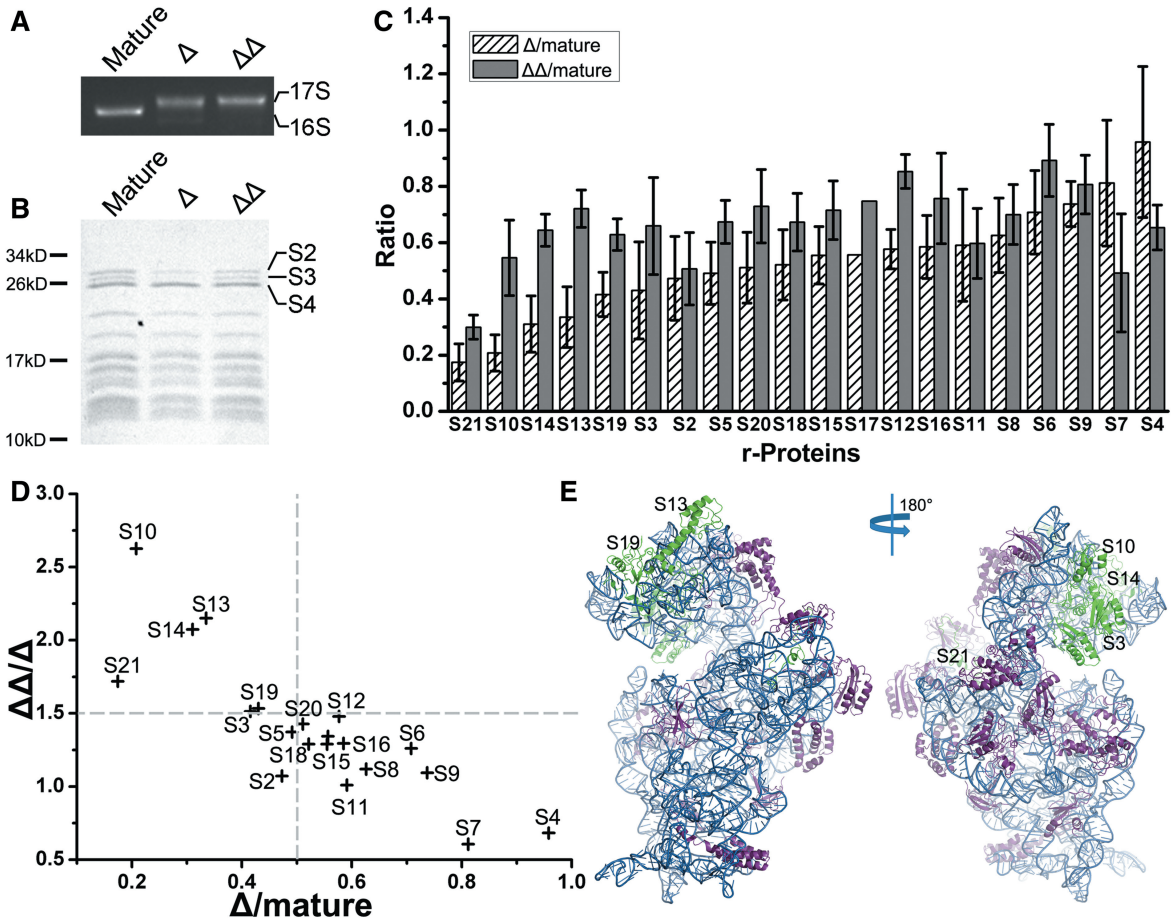

Figure 2. Compositional characterization of the immature 30S subunits. Composition of mature and immature 30S subunits from the A19 Δ*rimM* (Δ) and Δ*rbfA*Δ*rsgA* (ΔΔ) strains was analyzed in both the RNA and protein levels. (**A**) RNA gel analysis of the rRNAs in the Δ and ΔΔ samples. (**B**) Tricine-SDS–PAGE analysis of the protein composition of the immature 30S subunits. (**C**) QMS analysis of the protein composition in the Δ and the ΔΔ samples. Error bars show standard deviations. The difference in protein ratio between the Δ and ΔΔ samples was subjected to a one-tailed *t*-test, which reports a significant difference for S10, S14, S13, S3, S12, S5, S4, S7 ($P < 0.01$), S19 and S6 ($P < 0.05$). (**D**) The relative protein ratios of the ΔΔ sample to the Δ sample (ΔΔ/Δ) are plotted against the ratios of the Δ sample to the mature one (Δ/mature). (**E**) Atomic structure of the mature 30S subunit (38) viewed from the inter-subunit and solvent sides, with proteins in the top-left part of (D) colored in green.

from 1.5- to 2-folds. Interestingly, two primary proteins, S7 and S4, display significantly lower levels in the Δ*rbfA*Δ*rsgA* sample than in the Δ*rimM* sample (Figure 2C and D). Taking together, an evident pattern is that the immature 30S subunits from the Δ*rbfA*Δ*rsgA* strain have significantly higher occupancies for all the secondary and tertiary binding proteins in the 3′-domain (Figure 2E), suggesting that their 3′-head domains are indeed further maturated with more proteins incorporated.

This immediately suggests that a role of RimM *in vivo* is to promote the binding of 3′-domain proteins, since the immature 30S subunits from the Δ*rbfA*Δ*rsgA* strain likely resemble a stage downstream the RimM action. In agreement with this conclusion, the *in vitro* kinetic data show that RimM accelerates the binding of some head domain proteins, S19, S10 and S3 (49).

Structural characterization of the immature 30S subunits from the Δ*rimM* strain

To explore the structural heterogeneity of the immature 30S subunits, we applied the cryo-EM single-particle method to our sample. First, a reference-free 2D image classification technique based on maximum-likelihood optimization (32) was employed to estimate the level of structural variation in the cryo-EM particles. The 2D analysis reveals that a large number of the class average images show smeared densities on the head domain of the 30S subunit. In contrast, densities in these average images corresponding to the body domain are nicely resolved, and the features of the body domain could be easily identified (Supplementary Figure S2). This suggests that the immature 30S subunits are truly composed of multiple assembly intermediates, with a highly flexible head domain and a rather rigid body domain.

Next, a multi-structure refinement method (34) was used to investigate possible metastable structural intermediates at the 3D level. As a result, the particles were grouped into five classes, and as expected, the five cryo-EM maps (at 12–14 Å resolution) display dramatic conformational differences (Figure 3). However, similar to a previous cryo-EM study on the immature 30S subunits from a Δ*rsgA* strain (50), we did not find

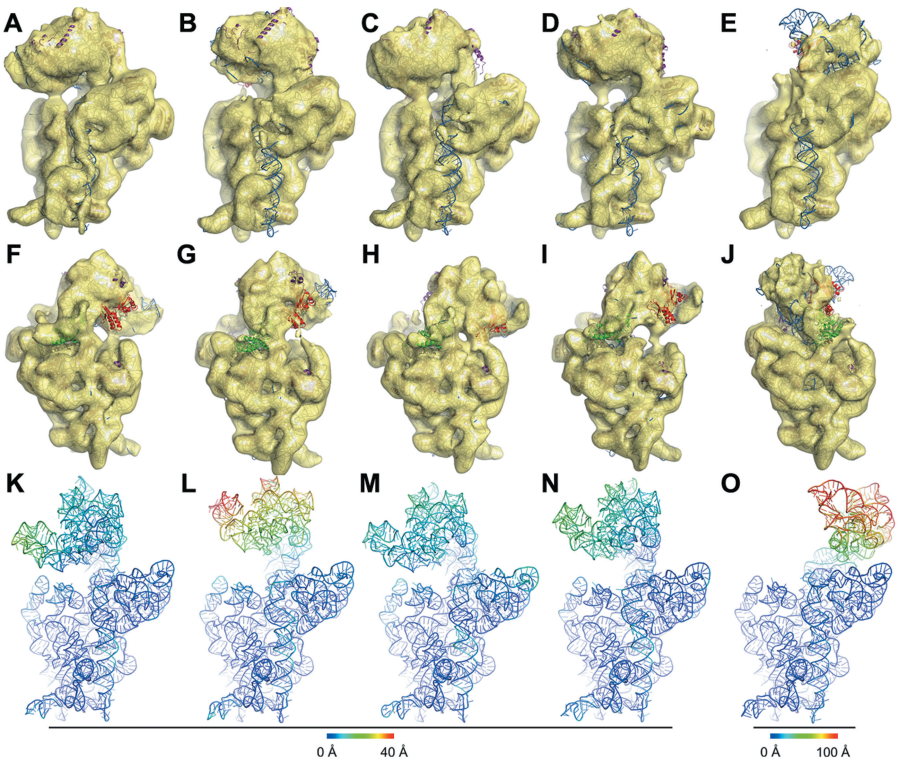

Figure 3. Overview of the five cryo-EM structures of the immature 30S subunits from the Δ*rimM* strain. The five density maps (**A–E** or **F–J**, respectively) are displayed in transparent surface representation, superimposed with flexibly fitted crystal structures in cartoon representation. For each map, both the inter-subunit view (A–E) and the solvent view (F–J) are displayed. The 16S rRNA, S2, S3 and the rest 30S subunit proteins are painted in blue, green, red and purple, respectively. Deviations of the 16S rRNA backbones in the fitted model from that of the mature 30S subunit are colored as indicated by the scale to form the temperature maps (**K–O**).

significant densities that could be attributed to the unprocessed ends of the 17S rRNA.

To facilitate the quantitative comparison of the structural data, we built pseudo-atomic models for the five cryo-EM maps using a flexible fitting technique (40). Based on these models, five temperature maps for the 16S rRNA were then constructed according to their deviations from the structure of the mature 30S subunit (Figure 3K–O). Structural difference can be directly identified from these maps. First, the conformational difference is dominated by a relatively rigid rotational movement of the head domain, which changes the inter-domain orientation between the head and the body domains. Especially, one class has a nearly 60° rotated head domain (Figure 3E, J and O). This rotated structure, derived from nearly one-third of all the particles (Supplementary Table S2), is in fact very similar to one of the Group II *in vitro* assembly intermediates discovered in a time-resolved electron microscopy study (11), which was shown to miss nearly all the 3′-domain proteins. In addition to the rotation, in two classes (Figure 3K and M), the channel between the head and the body domains is closed up, resulting in a narrow down of the mRNA entrance. Second, four of the five maps show very incomplete, fragmented densities for helix 44 of the 16S rRNA, except for one group (Figure 3A), which is close to the conformation of a mature 30S subunit and also has a less rotated head domain. Along with the conformational difference at helix 44, the decoding center is also sharply different among the five maps (Supplementary Figure S4). Third, in agreement with the QMS data, these structures differ in protein composition, as exemplified by S2 and S3 (Figure 3F–J). In fact, none of the structures has a full occupancy for both factors, and interestingly the occupancy of S2 has no correlation with the occupancy of S3 (Supplementary Table S2). This observation appears to align well with the *in vitro* assembly data showing that S2 and S3 could bind in independent order and the prior binding of S2 ahead of S3 leads to kinetically trapped intermediates (11).

In summary, as seen in the temperature maps, the head domain of the 16S rRNA is highly mobile in the immature 30S subunits. It is known that the motion between the head domain and the body domain is intrinsic and is believed to be required for the dynamic interaction with translational components (51). However, the head domain rotation observed in our structures is in a much larger scale, suggesting that hypo-level of proteins in the 3′-domain increases its flexibility. This structural observation demonstrates that the *in vivo* intermediates from the Δ*rimM* strain vary not only in protein composition but also in rRNA conformation.

Structural characterization of the Δ*rimM* immature 30S subunits bound with RimM

Next, we examined the binding preference of RimM to the immature and mature 30S subunits by pelleting assay. While RimM shows almost no binding to the mature

30S subunit, it indeed binds to the immature subunit, with a low affinity (Figure 4). In contrast, both RfbA and RsgA show a considerably higher affinity to the mature 30S subunit containing the 16S rRNA than RimM does, and especially, RsgA displays a strong preference to the mature 30S subunit (25). Therefore, similar to our QMS data, the binding preferences of these factors also suggest that RimM acts, prior to RbfA and RsgA, in the *in vivo* assembly pathway.

We then sought to explore possible structural changes of the immature 30S subunits upon RimM binding. Using the same 2D and 3D image classification techniques, we found that the addition of RimM to the immature 30S subunits seems to stabilize the 30S head domain. At the 2D level, class averages of particles from the RimM-treated sample still show smeared densities in the head region, but the total fraction of the particles with an unstable head domain is significantly smaller (Supplementary Figure S2). About 18% of particles from the untreated sample display apparent instability in the head domain, whereas the percentage in the treated sample is decreased to 11%.

At the 3D level, similarly, we classified the RimM-treated data into five groups, and these cryo-EM maps (at 15–19 Å resolution) also differ in conformation and protein composition (Figure 5). First, the head domain rotation is apparently in a much less scale, as seen in the temperature maps (Figure 5K–O and Supplementary Figure S5). Second, regions, such as the long helix 44 and the decoding center still display a large amount of variation (Supplementary Figure S4), implying the final accommodation of helix 44 is probably a later event, not related to RimM binding. Third, as expected, the occupancies of S2 and S3 are both very low, but surprisingly, the levels of S2 and S3 seem to be even lower than the untreated sample (Supplementary Table S2). This finding suggests that RimM stabilizes the immature 30S subunits in a conformation that disfavors S2 and S3 binding, implying that S2 and S3 binding might be later events in the assembly pathway.

Therefore, our structural analysis of the cryo-EM images from the RimM-treated sample shows that in addition to the role in ribosomal protein assembly, RimM also has a role in stabilizing the rRNA tertiary structure in the 3′-domain.

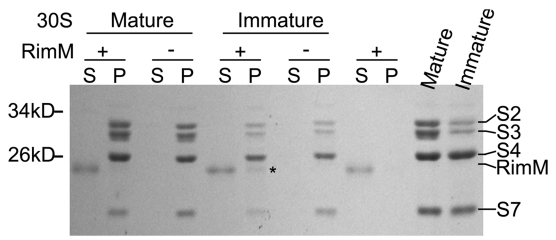

Figure 4. RimM preferentially binds to the immature 30S subunits from the Δ*rimM* strain. Immature 30S subunits from the Δ*rimM* strain and mature 30S subunits from the 70S ribosomes were incubated with or without a 30-fold excess of RimM. The mixtures were pelleted by centrifugation. The pellets (P) and the supernatants (S) were separated and resolved by SDS–PAGE. RimM alone was centrifuged as a control. The asterisk denotes the weak binding of RimM to the immature 30S subunits.

Binding position of RimM on the 30S subunit

The pelleting assay indicates that the affinity of RimM to the immature 30S subunits is very low (Figure 4). This apparently sets an obstacle for us to analyze contact sites of RimM on the 30S subunit in detail. Fortunately, RimM binds to S19 *in vitro* (20,21) and could co-crystalize with S19 (PDB ID: 3A1P). Therefore, the binding position of RimM could be deduced using S19 as a reference (Figure 6), given that RimM does not change its contacts in the context of the immature 30S subunit. In fact, the cryo-EM maps of the RimM-treated immature 30S subunits, although prepared with a 40-fold excess of RimM, have limited densities at locations expected to have RimM bound when the maps are displayed at a 3σ level (Figure 5). Densities corresponding to RimM begin to appear in lower threshold (Supplementary Figure S6). Nevertheless, we could compare the average densities statistically, within a 3D binary mask generated from the aligned RimM crystal structure. Consistently, the densities at RimM-bound region in the cryo-EM maps from the RimM-treated sample are significantly higher than those from the untreated sample (Supplementary Figure S6). This analysis, albeit rather preliminary, proves that RimM is present in these cryo-EM maps.

The structure of RimM is composed of two β-barrels containing domains (21). While the C-terminal domain is shown to interact with S19, the N-terminal domain closely resembles a tRNA-binding domain of EF-Tu (21), suggesting the ability of RimM to bind to the 16S rRNA. Consistently, alignment of the structure of the RimM-S19 complex immediately places the N-terminal domain of RimM at the junction of several helices, such as h29, h30 and h42 (Figure 6). Since, prior to RimM binding, the 30S assembly is in a stage with very limited 3′-domain protein incorporated (Figure 2C), the binding of RimM at this multi-helices interface might stabilize the rRNA conformation globally and therefore allows a faster and/or more stable binding of 3′-domain proteins.

DISCUSSION

The role of RimM in the assembly of the 30S subunit

It is known that disturbance to protein translation might affect the subunit assembly in an indirect way, due to a shortage in the ribosomal protein production. Consequently, the impaired subunit assembly in *E. coli* strains with genes for assembly factors deleted stems not only from the defective assembly process itself but also from a reduced supply of ribosomal proteins. Nevertheless, in this study, the composition of the *in vivo* intermediates from the Δ*rimM* and Δ*rbfA*Δ*rsgA* strains clearly displays a non-uniform level of ribosome proteins, with mostly the 3′-domain proteins significantly underrepresented (Figure 2), suggesting that the secondary effect caused by impaired translation in these strains is negligible and does not over-shadow the assembly defects.

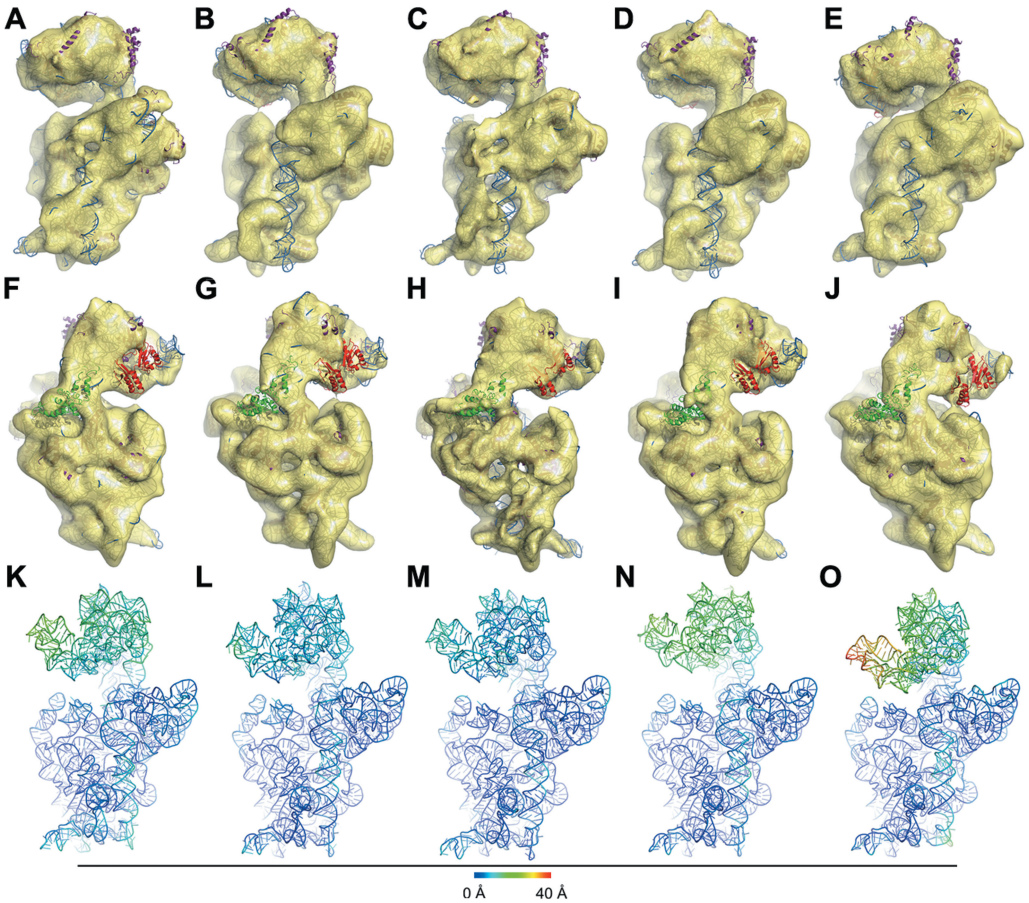

Figure 5. Overview of the five cryo-EM structures of the immature 30S subunits treated with RimM. The five density maps (**A–E** or **F–J**, respectively) are displayed in transparent surface representation, superimposed with flexibly fitted crystal structures in cartoon representation. For each map, both the inter-subunit view (A–E) and the solvent view (F–J) are displayed. The 16S rRNA, S2, S3 and the rest 30S subunit proteins are painted in blue, green, red and purple, respectively. Deviations of the 16S rRNA backbones in the fitted model from that of the mature 30S subunit are colored as indicated by the scale to form the temperature maps (**K–O**).

The role of RimM uncovered in the present work in fact serves as a perfect illustration to the proposed general function of assembly factors, i.e. to subvert possible kinetic traps caused by mis-folded rRNA or rate-limiting binding of certain proteins, during the *in vivo* assembly process (11,13,49). Previous kinetic data showed that the binding of 3′-domain proteins is not obligatory to the prebinding of S7 (13), while in contrast, prebinding of S7 and S19 together dramatically accelerates the binding of the rest 3′-domain proteins (13), indicating the presence of a rate-limiting S7-independent assembly pathway for S19 (13,52). Consistently, our data show that in the Δ*rimM* sample, S19 is among the most underrepresented proteins, whereas in the Δ*rbfA*Δ*rsgA* sample, the level of S19 is dramatically increased (Figure 2). Thus, it is likely that the role of RimM *in vivo* is to counteract the kinetic trap caused by slow binding of S19. In support of this view, both RimM and S19 bind to a multi-helices junction of the 16S 3′-domain (Figure 6), highlighting their potential effect on the global stabilization of the 3′-domain.

Functional interplay of assembly factors

In addition to the two sets of intermediates described in this study, another set of *in vivo* intermediates, isolated from a Δ*rsgA* strain, was also quantitatively analyzed (50). The comparison of these quantitative data from different genetic background would enable us to identify the temporal relationship of assembly factors.

First, assembly intermediates isolated from a Δ*rsgA* strain only have a very small subset of tertiary binding proteins (S21, S2 and S3) largely underrepresented (50), suggesting that RsgA acts at a very late stage when most of the components are already in place. Second, intermediates from the Δ*rimM* strain show underrepresented levels for all secondary and tertiary binding proteins from the 3′-domain. Interestingly, the Δ*rimM* intermediates are to some extent similar in protein composition to a previously identified *in vivo* 21S intermediates (44,53), but differ from the *in vitro* RI intermediates (54). The close resemblance of the Δ*rimM* intermediates to the naturally populated *in vivo* 21S intermediates suggests that they represent an early stage during the assembly of the 3′-domain, likely the entry stage. Last, in contrast, the intermediates from the Δ*rbfA*Δ*rsgA* strain, with only a subset of 3′-domain proteins largely underrepresented, do not resemble any of the known intermediates, indicating that they represent a novel set of intermediates roughly in-between.

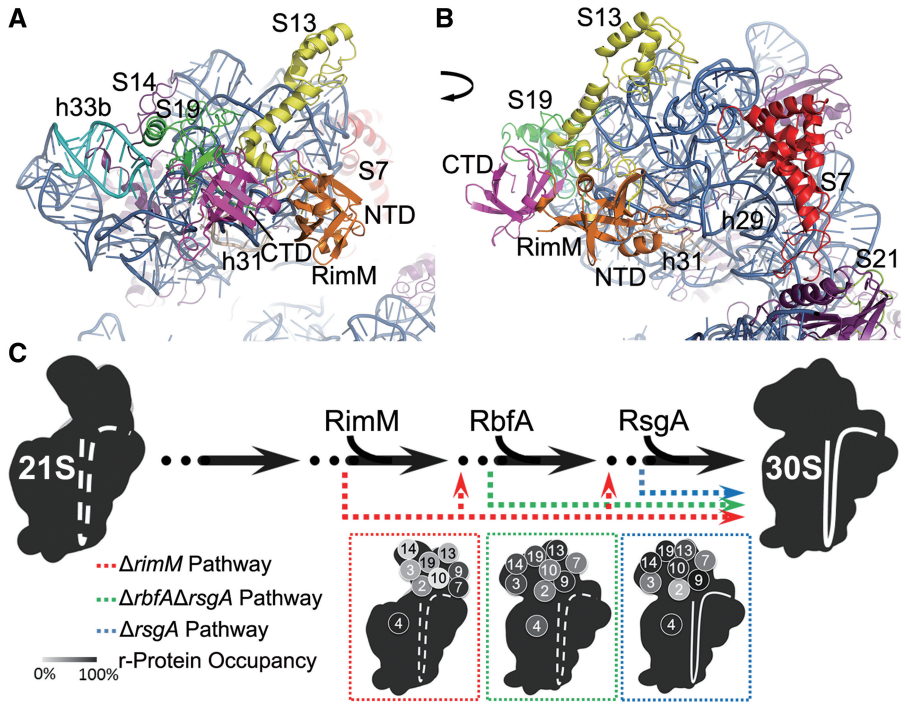

Figure 6. Mechanistic model of the RimM function in the *in vivo* assembly of the 3′-domain. (**A**) The head domain of the 30S subunit viewed from the inter-subunit side. (**B**) Same as (A), with a 70° rotation around *Y*-axis. The h33b, h31 and the rest of the 16S rRNA are painted in cyan, wheat and blue, respectively. The C-terminal domain of RimM (CTD), N-terminal domain of RimM (NTD), S7, S13, S19 and S14 are painted in magenta, orange, red, yellow, green and purple, respectively. (**C**) RimM, RbfA and RsgA act at different checkpoints during the *in vivo* assembly. The deficiency of assembly factors diverts the assembly into less efficient branches (colored dash lines) and causes accumulation of a set of closely related intermediates (colored boxes). The ribosomal protein levels in the three sets of *in vivo* intermediates are displayed in the gray scale. The data of the intermediates from a Δ*rsgA* strain is from a previous study (50). The large conformational differences among the three sets of *in vivo* intermediate were also shown in cartoon: the Δ*rimM* one (red) with a dramatically rotated head domain and a disordered helix 44; the Δ*rbfA*Δ*rsgA* one (green) with a disordered helix 44 only; the Δ*rsgA* one (blue) with a well-resolved helix 44.

Therefore, the protein spectra of the above three sets of intermediates clearly suggest an order for the actions of these factors (Figure 6C), which is consistent with previous genetic and biochemical data (19,25,55). It must be noted that it is difficult to unambiguously timestamp other biogenesis factors in the assembly pathway due to the lack of biochemical data, although genetic data have suggested both functional redundancy and hierarchy for some assembly factors [reviewed in (12)]. Nevertheless, if we view the *in vivo* assembly as a multi-branched process, the seemingly function redundancy among assembly factors is merely a sign of altered contribution of different, inter-connected assembly pathways. As shown in Figure 6C, the late-stage assembly *in vivo* starts with an *in vivo* 21S intermediate (44,53) and proceeds along a highly efficient pathway in the presence of all assembly factors. Assembly factors come in play at different time points to assist certain kinetically disfavored assembly events. The disruption of a factor or a combination of factors would avert the assembly to less efficient branches and cause accumulation of a certain category of kinetically trapped intermediates. Consistent with this view, most of the *E. coli* genes for assembly factors are not essential. The remaining question is whether these kinetically trapped intermediates from various genetic background with different factors disrupted truly represent genuine snapshots of the assembly process in the normal condition, or different 'dead-end' products that are otherwise elusive in the normal condition.

The *in vivo* assembly of the 30S 3′-domain also follows parallel pathways

Early chemical probing of the 16S rRNA conformation (56), as well as recent kinetic measurement of the protein binding (9,11), showed that the 3′-domain assembly is the latest event during the *in vitro* assembly of the 30S subunit, coincident with the 5′- to 3′-transcription order. On the other hand, accumulating evidences (9–11,56) suggest that a major general feature of the *in vitro* assembly of the 30S subunit is that the process proceeds along multiple routes.

In this study, we isolated the *in vivo* assembly intermediates from two genetically modified *E. coli* strains. The most remarkable feature in the protein spectra of these two sets of intermediates is that they both severely lack 3′-domain proteins, suggesting that the maturation of the 3′-domain is also a rate-limiting process *in vivo*. These quantitative data also indicate that the intermediates from both strains are very heterogeneous in ribosomal protein composition, which means they do not represent a single populated assembly intermediate state, but rather a collection of multiple-related intermediates with more than one metastable state enriched. These differently prepared intermediates, although with a recognizable temporal relationship, cannot be easily reconciled by a single continuous assembly pathway.

With respect to the occupancies of individual proteins, there are a number of exceptions to the well-accepted Nomura assembly map (5). To name a few, first, an unexpected observation of our QMS data is that two primary proteins (S4 and S7) in the Δ*rbfA*Δ*rsgA* sample show decreased levels compared with the Δ*rimM* sample. S4 is thermodynamically required for subsequent binding of S16, S12 and S5 to the 5′-domain (5), and more importantly, S4 was shown to have a global stabilization effect on the 5′-domain (57). S7 is the only primary protein in the 3′-head domain and directs the following binding of S9, S13 and S19 (5,48). However, in the Δ*rbfA*Δ*rsgA* sample, the occupancy of S4 and S7 (Figure 2) is even lower than its follower proteins. Similarly, the level of S7 was also reported to be lower than its follower S9, S13 and S19 in the Δ*rsgA* sample (50). Second, levels of S10 and S14 are lower than their follower tertiary proteins S2 and S3 in the Δ*rimM* sample, suggesting that S3 could bind independent of S10. Third, the level of S2 is lower than S3 in the Δ*rbfA*Δ*rsgA* sample, although S3 binding is thermodynamically dependent on the prior binding of S2 in the Nomura map.

The Nomura map was derived by single protein omission reconstitution experiments with fully processed 16S rRNA under equilibrium conditions and therefore does not necessarily reflect the true order of serials of binding events during assembly (52). In addition to our QMS data, deviations from the Nomura map have already been observed from both *in vivo* and *in vitro* studies. Previous genetic data show that S15, a primary protein in the central domain, is dispensable for the 30S assembly *in vivo* (58). Furthermore, *in vitro* kinetic data from Williamson group based on PC-QMS (11,13) or fluorescence triple correlation spectroscopy (52) indicate that S9 and S19 could bind independent of S7 (13,52), and S2 could bind independent of S3 (11). All these data suggest that there are hidden assembly pathways that could not be directly inferred from the Nomura map.

Thus, the seemingly discrepancy between our QMS data and the Nomura map could be easily reconciled if we view the *in vivo* assembly process as a highly branched network. In the presence of not fully processed 17S rRNA and the absence of certain assembly factors, the 30S assembly *in vivo* takes alternative, kinetically inefficient, pathways that are not predicted by the Nomura map. Therefore, the assembly intermediates isolated from these different assembly factor-deficient mutants might represent intermediates kinetically trapped in various parallel branches in the assembly network.

In summary, although the number of possible assembly pathways *in vivo* is limited by the co-transcriptional nature and the presence of assembly factors in the *in vivo* condition, our data provide additional strong evidence to the emerging idea that the *in vivo* assembly also proceeds along parallel pathways in a certain degree (54,58).

The maturation of the 16S rRNA 3′-domain *in vivo* is highly coupled with protein assembly

Our structural data reveal that the *in vivo* assembly intermediates differ largely in rRNA conformation. Through the integration of structural (26,50) and QMS data [the present work and (50)], we could draw a conclusion that the 3′-domain of the 16S rRNA maturates in a progressive manner *in vivo*, paralleling with the ribosomal protein assembly.

First, assembly intermediates from Δ*rimM* cells show dramatic conformational differences in the position of the 3′-domain (Figure 3 and Supplementary Figure S2). In contrast, cryo-EM structures of intermediates from Δ*rsgA* cells (50) also vary in the 3′-domain position, but with a much smaller scale. Second, the long helix 44 of the 3′-minor domain is highly flexible and is almost invisible in the structures of intermediates from Δ*rimM* cells (Figures 3 and 5). However, cryo-EM structures of 30S intermediates from Δ*rsgA* cells (50), which also contain 17S rRNA, display well-resolved densities for helix 44, except for the upper decoding center region. This suggests that helix 44 adapts its mature conformation only at a very late stage. In fact, hydroxyl radical probing data (10,59) already showed that the full accommodation of helix 44 is a late event even when the experiments were performed with the 16S rRNA. Third, further downstream is the cryo-EM structure of the 30S–RsgA complex, which displays an almost identical conformation to the mature 30S subunit (26).

Based on the above structural comparisons, the maturation of the 3′-domain of the 16S rRNA follows the transcription order in a progressive manner, first the 3′-head domain, next the 3′-minor domain (Figure 6C), and more importantly, the conformational maturation is coupled with the gradually increased protein level (Figure 6C). Therefore, our data have revealed another common characteristic shared by the *in vivo* and the *in vitro* processes, i.e. a high cooperativity between protein binding and rRNA folding (10,11,60,61).

ACCESSION NUMBERS

The cryo-EM maps have been deposited in the EMDataBank (EMDB codes 5500-5504 and 5506-5510 for the immature 30S subunits with or without the RimM treatment, respectively). The atomic models of the 16S rRNA have been deposited in the Protein Data Bank (PDB codes 3J28, 3J29, 3J2A, 3J2B, 3J2C, 3J2D, 3J2E, 3J2F, 3J2G, 3J2H).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Materials and Methods, Supplementary Tables 1 and 2, Supplementary Figures 1–7 and Supplementary References [26,32,44].

ACKNOWLEDGEMENTS

We thank Nara Institute of Science and Technology, National BioResource Project and National Institute of Genetics for providing the Δ*rimM* strain from the Keio collection and Dr Krisana Asano for construction of an overproducing system of RimM. We also thank the

Tsinghua National Laboratory for Information Science and Technology for providing computing resource.

FUNDING

Ministry of Science and Technology of China [2010CB912401, 2010CB912402 and 2013CB910404]; the National Natural Science Foundation of China [31170677] and the Japan Society for the promotion of Science [23380054]. Funding for open access charge: Ministry of Science and Technology of China.

Conflict of interest statement. None declared.

REFERENCES

1. Kaczanowska,M. and Ryden-Aulin,M. (2007) Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiol. Mol. Biol. Rev.*, **71**, 477–494.
2. Held,W.A., Mizushima,S. and Nomura,M. (1973) Reconstitution of *Escherichia coli* 30 S ribosomal subunits from purified molecular components. *J. Biol. Chem.*, **248**, 5720–5730.
3. Traub,P. and Nomura,M. (1968) Structure and function of *E. coli* ribosomes. V. Reconstitution of functionally active 30S ribosomal particles from RNA and proteins. Proc. *Natl Acad. Sci. USA*, **59**, 777–784.
4. Culver,G.M. and Noller,H.F. (1999) Efficient reconstitution of functional *Escherichia coli* 30S ribosomal subunits from a complete set of recombinant small subunit ribosomal proteins. *RNA*, **5**, 832–843.
5. Mizushima,S. and Nomura,M. (1970) Assembly mapping of 30S ribosomal proteins from *E. coli*. *Nature*, **226**, 1214.
6. Krzyzosiak,W., Denman,R., Nurse,K., Hellmann,W., Boublik,M., Gehrke,C.W., Agris,P.F. and Ofengand,J. (1987) In vitro synthesis of 16S ribosomal RNA containing single base changes and assembly into a functional 30S ribosome. *Biochemistry*, **26**, 2353–2364.
7. Mangiarotti,G., Turco,E., Perlo,C. and Altruda,F. (1975) Role of precursor 16S RNA in assembly of *E. coli*. 30S ribosomes. *Nature*, **253**, 569–571.
8. de Narvaez,C.C. and Schaup,H.W. (1979) In vivo transcriptionally coupled assembly of *Escherichia coli* ribosomal subunits. *J. Mol. Biol.*, **134**, 1–22.
9. Talkington,M.W., Siuzdak,G. and Williamson,J.R. (2005) An assembly landscape for the 30S ribosomal subunit. *Nature*, **438**, 628–632.
10. Adilakshmi,T., Bellur,D.L. and Woodson,S.A. (2008) Concurrent nucleation of 16S folding and induced fit in 30S ribosome assembly. *Nature*, **455**, 1268–1272.
11. Mulder,A.M., Yoshioka,C., Beck,A.H., Bunner,A.E., Milligan,R.A., Potter,C.S., Carragher,B. and Williamson,J.R. (2010) Visualizing ribosome biogenesis: parallel assembly pathways for the 30S subunit. *Science*, **330**, 673–677.
12. Shajani,Z., Sykes,M.T. and Williamson,J.R. (2011) Assembly of bacterial ribosomes. *Annu. Rev. Biochem.*, **80**, 501–526.
13. Bunner,A.E., Beck,A.H. and Williamson,J.R. (2010) Kinetic cooperativity in *Escherichia coli* 30S ribosomal subunit reconstitution reveals additional complexity in the assembly landscape. *Proc. Natl Acad. Sci. USA*, **107**, 5417–5422.
14. Loh,P.C., Morimoto,T., Matsuo,Y., Oshima,T. and Ogasawara,N. (2007) The GTP-binding protein YqeH participates in biogenesis of the 30S ribosome subunit in *Bacillus subtilis*. *Genes Genet. Syst.*, **82**, 281–289.
15. Uicker,W.C., Schaefer,L., Koenigsknecht,M. and Britton,R.A. (2007) The essential GTPase YqeH is required for proper ribosome assembly in *Bacillus subtilis*. *J. Bacteriol.*, **189**, 2926–2929.
16. Persson,B.C., Bylund,G.O., Berg,D.E. and Wikstrom,P.M. (1995) Functional analysis of the ffh-trmD region of the *Escherichia coli* chromosome by using reverse genetics. *J. Bacteriol.*, **177**, 5554–5560.
17. Bystrom,A.S., Hjalmarsson,K.J., Wikstrom,P.M. and Bjork,G.R. (1983) The nucleotide sequence of an *Escherichia coli* operon containing genes for the tRNA(m1G)methyltransferase, the ribosomal proteins S16 and L19 and a 21-K polypeptide. *EMBO J.*, **2**, 899–905.
18. Bylund,G.O., Persson,B.C., Lundberg,L.A. and Wikstrom,P.M. (1997) A novel ribosome-associated protein is important for efficient translation in *Escherichia coli*. *J. Bacteriol.*, **179**, 4567–4574.
19. Bylund,G.O., Wipemo,L.C., Lundberg,L.A. and Wikstrom,P.M. (1998) RimM and RbfA are essential for efficient processing of 16S rRNA in *Escherichia coli*. *J. Bacteriol.*, **180**, 73–82.
20. Lovgren,J.M., Bylund,G.O., Srivastava,M.K., Lundberg,L.A., Persson,O.P., Wingsle,G. and Wikstrom,P.M. (2004) The PRC-barrel domain of the ribosome maturation protein RimM mediates binding to ribosomal protein S19 in the 30S ribosomal subunits. *RNA*, **10**, 1798–1812.
21. Suzuki,S., Tatsuguchi,A., Matsumoto,E., Kawazoe,M., Kaminishi,T., Shirouzu,M., Muto,Y., Takemoto,C. and Yokoyama,S. (2007) Structural characterization of the ribosome maturation protein, RimM. *J. Bacteriol.*, **189**, 6397–6406.
22. Gesteland,R.F. (1966) Isolation and characterization of ribonuclease I mutants of *Escherichia coli*. *J. Mol. Biol.*, **16**, 67–84.
23. Baba,T., Ara,T., Hasegawa,M., Takai,Y., Okumura,Y., Baba,M., Datsenko,K.A., Tomita,M., Wanner,B.L. and Mori,H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.*, **2**, 2006.0008.
24. Datsenko,K.A. and Wanner,B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA*, **97**, 6640–6645.
25. Goto,S., Kato,S., Kimura,T., Muto,A. and Himeno,H. (2011) RsgA releases RbfA from 30S ribosome during a late stage of ribosome biosynthesis. *EMBO J.*, **30**, 104–114.
26. Guo,Q., Yuan,Y., Xu,Y., Feng,B., Liu,L., Chen,K., Sun,M., Yang,Z., Lei,J. and Gao,N. (2011) Structural basis for the function of a small GTPase RsgA on the 30S ribosomal subunit maturation revealed by cryoelectron microscopy. *Proc. Natl Acad. Sci. USA*, **108**, 13100–13105.
27. Lei,J. and Frank,J. (2005) Automated acquisition of cryo-electron micrographs for single particle reconstruction on an FEI Tecnai electron microscope. *J. Struct. Biol.*, **150**, 69–80.
28. Frank,J., Radermacher,M., Penczek,P., Zhu,J., Li,Y., Ladjadj,M. and Leith,A. (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.*, **116**, 190–199.
29. Rath,B.K. and Frank,J. (2004) Fast automatic particle picking from cryo-electron micrographs using a locally normalized cross-correlation function: a case study. *J. Struct. Biol.*, **145**, 84–90.
30. Shaikh,T.R., Trujillo,R., LeBarron,J.S., Baxter,W.T. and Frank,J. (2008) Particle-verification for single-particle, reference-based reconstruction using multivariate data analysis and classification. *J. Struct. Biol.*, **164**, 41–48.
31. Shaikh,T.R., Gao,H.X., Baxter,W.T., Asturias,F.J., Boisset,N., Leith,A. and Frank,J. (2008) SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nat. Protoc.*, **3**, 1941–1974.
32. Scheres,S.H., Valle,M., Nunez,R., Sorzano,C.O., Marabini,R., Herman,G.T. and Carazo,J.M. (2005) Maximum-likelihood multi-reference refinement for electron microscopy images. *J. Mol. Biol.*, **348**, 139–149.
33. Scheres,S.H., Nunez-Ramirez,R., Sorzano,C.O., Carazo,J.M. and Marabini,R. (2008) Image processing for electron microscopy single-particle analysis using XMIPP. *Nat. Protoc.*, **3**, 977–990.
34. Scheres,S.H., Gao,H., Valle,M., Herman,G.T., Eggermont,P.P., Frank,J. and Carazo,J.M. (2007) Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat. Methods*, **4**, 27–29.
35. Gabashvili,I.S., Agrawal,R.K., Spahn,C.M., Grassucci,R.A., Svergun,D.I., Frank,J. and Penczek,P. (2000) Solution structure

of the *E. coli* 70S ribosome at 11.5 A resolution. *Cell*, **100**, 537–549.
36. LeBarron,J., Grassucci,R.A., Shaikh,T.R., Baxter,W.T., Sengupta,J. and Frank,J. (2008) Exploration of parameters in cryo-EM leading to an improved density map of the *E. coli* ribosome. *J. Struct. Biol.*, **164**, 24–32.
37. Stewart,P.L., Cary,R.B., Peterson,S.R. and Chiu,C.Y. (2000) Digitally collected cryo-electron micrographs for single particle reconstruction. *Microsc. Res. Tech.*, **49**, 224–232.
38. Dunkle,J.A., Xiong,L., Mankin,A.S. and Cate,J.H. (2010) Structures of the *Escherichia coli* ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action. *Proc. Natl Acad. Sci. USA*, **107**, 17152–17157.
39. Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) UCSF chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.
40. Trabuco,L.G., Villa,E., Mitra,K., Frank,J. and Schulten,K. (2008) Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure*, **16**, 673–683.
41. Phillips,J.C., Braun,R., Wang,W., Gumbart,J., Tajkhorshid,E., Villa,E., Chipot,C., Skeel,R.D., Kale,L. and Schulten,K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, **26**, 1781–1802.
42. Delano,W.L. (2010) The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA.
43. Datta,A.K. and Burma,D.P. (1972) Association of ribonuclease I with ribosomes and their subunits. *J. Biol. Chem.*, **247**, 6795–6801.
44. Rene,O. and Alix,J.H. (2011) Late steps of ribosome assembly in *E. coli* are sensitive to a severe heat stress but are assisted by the HSP70 chaperone machine. *Nucleic Acids Res.*, **39**, 1855–1867.
45. Jiang,M., Datta,K., Walker,A., Strahler,J., Bagamasbad,P., Andrews,P.C. and Maddock,J.R. (2006) The *Escherichia coli* GTPase CgtAE is involved in late steps of large ribosome assembly. *J. Bacteriol.*, **188**, 6757–6770.
46. Thompson,A., Schafer,J., Kuhn,K., Kienle,S., Schwarz,J., Schmidt,G., Neumann,T., Johnstone,R., Mohammed,A.K. and Hamon,C. (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.*, **75**, 1895–1904.
47. Bunner,A.E., Trauger,S.A., Siuzdak,G. and Williamson,J.R. (2008) Quantitative ESI-TOF analysis of macromolecular assembly kinetics. *Anal. Chem.*, **80**, 9379–9386.
48. Grondek,J.F. and Culver,G.M. (2004) Assembly of the 30S ribosomal subunit: positioning ribosomal protein S13 in the S7 assembly branch. *RNA*, **10**, 1861–1866.
49. Bunner,A.E., Nord,S., Wikstrom,P.M. and Williamson,J.R. (2010) The effect of ribosome assembly cofactors on in vitro 30S subunit reconstitution. *J. Mol. Biol.*, **398**, 1–7.
50. Jomaa,A., Stewart,G., Martin-Benito,J., Zielke,R., Campbell,T.L., Maddock,J.R., Brown,E.D. and Ortega,J. (2011) Understanding ribosome assembly: the structure of in vivo assembled immature 30S subunits revealed by cryo-electron microscopy. *RNA*, **17**, 697–709.
51. Gabashvili,I.S., Agrawal,R.K., Grassucci,R. and Frank,J. (1999) Structure and structural variations of the *Escherichia coli* 30 S ribosomal subunit as revealed by three-dimensional cryo-electron microscopy. *J. Mol. Biol.*, **286**, 1285–1291.
52. Ridgeway,W.K., Millar,D.P. and Williamson,J.R. (2012) Quantitation of ten 30S ribosomal assembly intermediates using fluorescence triple correlation spectroscopy. *Proc. Natl Acad. Sci. USA*, **109**, 13614–13619.
53. Lindahl,L. (1975) Intermediates and time kinetics of the in vivo assembly of *Escherichia coli* ribosomes. *J. Mol. Biol.*, **92**, 15–37.
54. Sykes,M.T., Shajani,Z., Sperling,E., Beck,A.H. and Williamson,J.R. (2010) Quantitative proteomic analysis of ribosome assembly and turnover in vivo. *J. Mol. Biol.*, **403**, 331–345.
55. Campbell,T.L. and Brown,E.D. (2008) Genetic interaction screens with ordered overexpression and deletion clone sets implicate the *Escherichia coli* GTPase YjeQ in late ribosome biogenesis. *J. Bacteriol.*, **190**, 2537–2545.
56. Powers,T., Daubresse,G. and Noller,H.F. (1993) Dynamics of in vitro assembly of 16 S rRNA into 30 S ribosomal subunits. *J. Mol. Biol.*, **232**, 362–374.
57. Ramaswamy,P. and Woodson,S.A. (2009) Global stabilization of rRNA structure by ribosomal proteins S4, S17, and S20. *J. Mol. Biol.*, **392**, 666–677.
58. Bubunenko,M., Korepanov,A., Court,D.L., Jagannathan,I., Dickinson,D., Chaudhuri,B.R., Garber,M.B. and Culver,G.M. (2006) 30S ribosomal subunits can be assembled in vivo without primary binding ribosomal protein S15. *RNA*, **12**, 1229–1239.
59. Dutca,L.M. and Culver,G.M. (2008) Assembly of the 5′ and 3′ minor domains of 16S ribosomal RNA as monitored by tethered probing from ribosomal protein S20. *J. Mol. Biol.*, **376**, 92–108.
60. Holmes,K.L. and Culver,G.M. (2004) Mapping structural differences between 30S ribosomal subunit assembly intermediates. *Nat. Struct. Mol. Biol.*, **11**, 179–186.
61. Holmes,K.L. and Culver,G.M. (2005) Analysis of conformational changes in 16 S rRNA during the course of 30 S subunit assembly. *J. Mol. Biol.*, **354**, 340–357.