# A conserved rRNA switch is central to decoding site maturation on the small ribosomal subunit

Andreas Schedlbauer[a], Idoia Iturrioz[a], Borja Ochoa-Lizarralde[a], Tammo Diercks[a], Jorge Pedro López-Alonso[a], José Luis Lavin[b], Tatsuya Kaminishi[a,c], Retina Capuni[a], Neha Dhimole[a], Elisa de Astigarraga[a], David Gil-Carton[a], Paola Fucini[a,d*], and Sean R. Connell[a,d*]

[a]Center for Cooperative Research in Biosciences (CIC bioGUNE), Basque Research and Technology Alliance (BRTA), Bizkaia Technology Park, Building 801A, 48160 Derio, Spain.
[b]Neiker–Tecnalia, Berreaga Kalea 1, 48160 Derio, Spain
[c]Division of Medicine, Graduate School of Medicine, Osaka University, Japan.
[d]IKERBASQUE, Basque Foundation for Science, 48013 Bilbao, Spain.
*To whom correspondence should be addressed.

**While a structural description of the molecular mechanisms guiding ribosome assembly in eukaryotic systems is emerging, bacteria employ an unrelated core set of assembly factors for which high-resolution structural information is still missing. To address this, we used single-particle cryo-EM to visualize the effects of bacterial ribosome assembly factors RimP, RbfA, RsmA, and RsgA on the conformational landscape of the 30S ribosomal subunit and obtained eight snapshots representing late steps in the folding of the decoding center. Analysis of these structures identifies a conserved secondary structure switch in the 16S rRNA central to decoding site maturation, and suggests both a sequential order of action and molecular mechanisms for the assembly factors in coordinating and controlling this switch. Structural and mechanistic parallels between bacterial and eukaryotic systems indicate common folding features inherent to all ribosomes.**

**Keywords: Cryo-EM, 30S biogenesis, ribosome assembly, RbfA, RsgA, YjeQ, RimP, KsgA, RsmA, NMR spectroscopy. Abbreviations used: cryo-EM, cryo-electron microscopy.**

Ribosome biogenesis is an essential, energy demanding process in both prokaryotic and eukaryotic cells (1). In humans, defects in ribosome biogenesis are linked to diverse pathologies, called ribosomopathies, that manifest as a broad range of developmental disorders (2, 3). In bacteria, protein factors involved in ribosome biogenesis are critical for cell growth and pathogenesis (4, 5) making them potential anti-microbial targets (6, 7). A detailed knowledge of the molecular mechanisms driving ribosome assembly, as further developed in this work, is essential for understanding the mechanistic basis of ribosomopathies and opening new pharmaceutical approaches to combat multidrug resistance in bacteria. In bacterial systems, like *Escherichia coli*, ribosome biogenesis involves the assembly of two ribosomal subunits, where the smaller 30S subunit is formed by the 16S ribosomal RNA (rRNA) and 21 ribosomal proteins (r-proteins) while the larger 50S subunit comprises the 5S rRNA, the 23S rRNA, and 33 r-proteins. Ribosome assembly occurs co-transcriptionally and RNA folding starts immediately upon synthesis of the 5´-end of the transcript. Moreover, r-proteins associate with the folding rRNA transcript even as it is being processed by RNases (i.e., excised

and trimmed from a longer transcript) and chemically modified by transacting factors, such as the methyl transferase RsmA (also known as KsgA). Despite this complexity, prior biochemical and structural work has shown that assembly of the 30S subunit is a robust process proceeding via multiple redundant parallel pathways, where the 5´-body domain forms first, followed by the central platform and head domains, and finally the 3´minor domain with the functionally important central decoding region (CDR) (8–11). During the later assembly phases including CDR folding, kinetic traps are prevalent. These are local minima in the ribosome's folding landscape and correspond to alternative RNA conformations that result from degenerate local interactions and rearrange slowly into the native structure at physiological conditions (12). Ribosomal assembly factors are presumed to intervene and promote 30S biogenesis by avoiding such kinetic traps, thus allowing the CDR to adopt a functional fold. These factors include RimP, RsmA, RsgA, and RbfA, which play intertwined roles and assist in CDR folding (13) by ensuring the correct placement of the 3´ minor domain (i.e. helices h44 and h45) between the rRNA domains forming the 30S head, body, and platform (see Figure 1, state **M**). The dedication of several assembly factors to CDR folding underscores its central function as an integral part of the 30S A and P-sites where tRNAs recognise and translate the mRNA sequence (14). In the CDR, h44 along with its connecting linker regions participate in mRNA-tRNA binding while the 16S 3´-end harbours the anti-Shine-Dalgarno sequence to recruit mRNA to the 30S subunit (14, 15). While cryo-EM has described the binding positions of the 30S assembly factors RsgA, RbfA, and RsmA at intermediate resolution (13, 16–18), the molecular mechanisms by which they avoid kinetic traps and promote CDR biogenesis remain unclear.

## Overall cryo-EM characterization of 30S complexes with late-stage ribosomal assembly factors

Conformational heterogeneity in isolated ribosomal subunits were first described over 50 years ago by Zamir et al. (1969) and Moazed et al. (1986) (20, 21) when they demonstrated that isolated 30S exist in an "active"
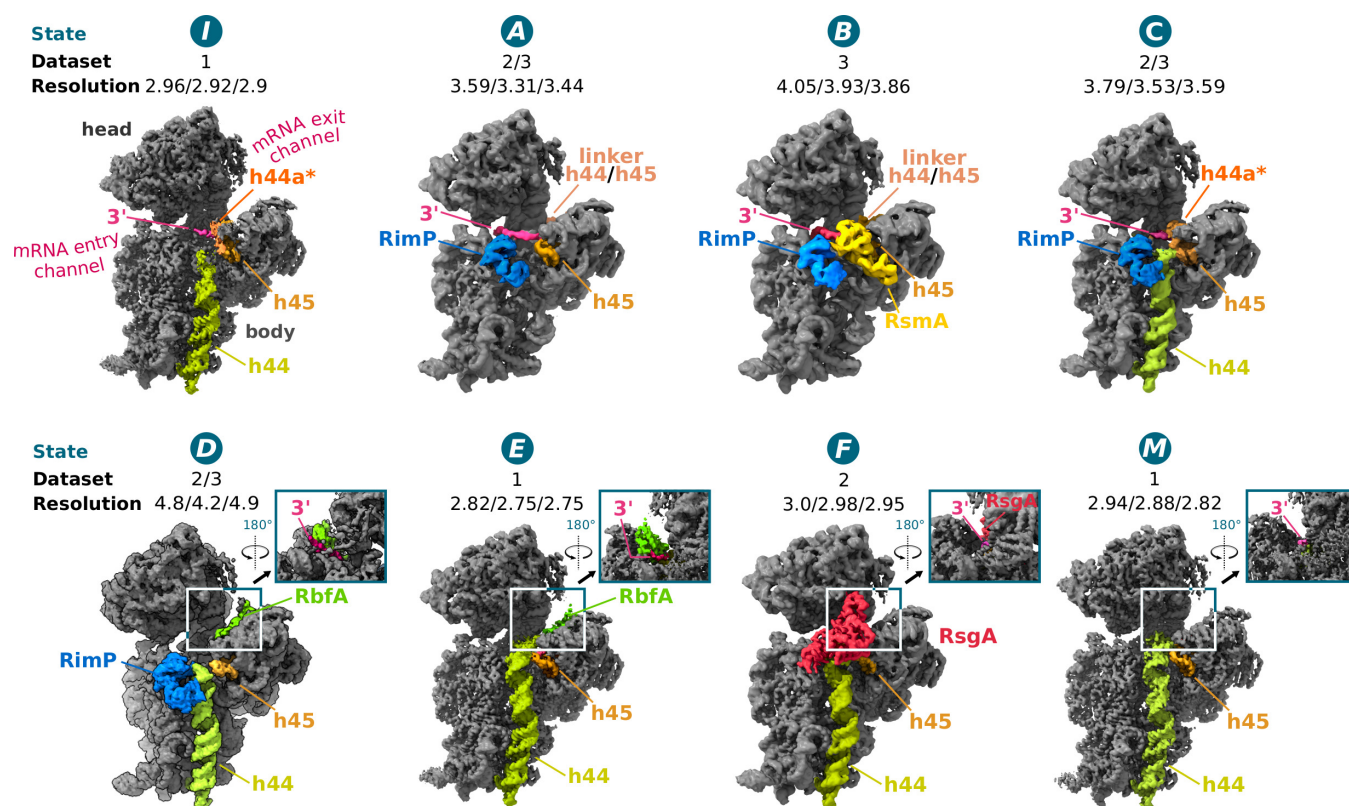
**Figure 1. Cryo-EM structures of 30S subunits in complex with different assembly factors.** Cryo-EM reconstructions (states *I*, *A-F*, and *M*) of the 30S assembly complexes are shown from the front (intersubunit side), with insets showing the backside (solvent side) to highlight the position of RbfA and/or the 16S 3´-end in the mRNA exit channel in states *D-M*. The assigned 30S state is shown above each reconstruction along with the originating dataset, and resolution (overall/body/head). Dataset 1: 30S+RbfA; dataset 2: 30S+RbfA+RsgA+RimP+RimM; dataset 3: 30S+RbfA+RimP+RsmA. Cryo-EM maps are derived from the multibody refinement and shown as composite maps (phenix.combine_focused_maps (19)), and are unsharpened and filtered to their estimated global resolution.

and "inactive" conformation, and more recently by Weeks and colleagues (22) who showed that an "inactive" 30S conformation is a *bonafide in vivo* state. Both *in vitro* and *in vivo* work agree that a conformational change in the CDR differentiates the two 30S conformations. We surmise that the inactive population in the isolated 30S subunits mimics immature states allowing, as previously demonstrated (13, 16, 18), late-stage ribosome assembly factors to interact with purified subunits. Therefore, by using these natively purified mature and immature 30S subunits and different combinations of late-stage ribosomal assembly factors known to influence CDR maturation (RbfA+RsgA+RimP+RimM+RsmA; Figure1, Figure S1), our cryo EM study reveals fundamental molecular mechanisms governing assembly factor-mediated folding of the CDR. In total eight high-resolution structures were determined, including two *apo* 30S structures with no assembly factor bound that show an immature (state *I*) and the mature state (state *M*) consistent with those proposed by the biochemical work half a century ago (Figure 1). The remaining six structures, with up to two different assembly factors bound simultaneously (Figure 1), can be ordered by the increasing maturation, i.e. native structure content, in the CDR (state *A-F*, Figure 1, Table S2), suggesting

a sequence of action for the tested assembly factors and disclosing intermediate conformations in the late stages of CDR folding.

To determine these structures, three reaction mixtures were processed using single-particle analysis (Table S4 and Figures S2 to S4). Refining the entire 30S subunit as a single body (by 3D classification and consensus refinement) led to low local resolution in the head region, indicating its variable orientation relative to the body region (Figures S5 and S6). Therefore, we refined the 30S head and body regions independently using a multibody approach (23), which yielded cryo-EM maps with overall resolutions between 2.75 Å and 4.9 Å. Local resolution in the assembly factor regions ranged between 2.6 Å and > 5 Å (Figure S5) enabling their modelling in the 30S bound state (Tables S4 to S11; Figure 1).

### An rRNA switch in the 16S rRNA delineates the CDR transition into a mature-like state.

Comparing the structures of the immature (state *I*; Figure 2A) and mature 30S subunit (state *M*; Figure 2C) maps the most pronounced differences primarily to the region around rRNA helix h28 that forms the so-called neck

Schedlbauer *et al.* | A conserved rRNA switch is central to decoding site maturation

**Table 1.** Overview of structural features observed in various states during late stage ribosome assembly.

| 30S state | 16S 3´-end position | h28 conformation | h44 position[1] | h44 linker conformation[2] | Assembly factor bound |
|---|---|---|---|---|---|
| I | entry channel | immature | front | h44[a] | - |
| A | entry channel | immature | exit | disordered | RimP |
| B | entry channel | immature | exit | disordered | RimP + RsmA |
| C | entry channel | immature | front | h44[a] | RimP |
| D | exit channel | mature | front | disordered | RimP + RbfA |
| E | exit channel | mature | front | disordered | RbfA |
| F | exit channel | mature | front | mature | RsgA |
| M | exit channel | mature | front | mature | - |

[1] In the front position, h44 localizes on the inter-subunit surface of the 30S subunit. In the exit position, partial cryo-EM density attributed to h44 is observed to project out of the exit channel (see discussion in Figure S8).

[2] Proposed secondary structure and description of disordered/unmodeled 16S rRNA residues in the h44 linkers are shown in Figure S7.

connecting the head and body regions. Here, the most significant change involves the 16S 3´-end (pink) swapping from the mRNA entry channel in state *I* to its canonical position within the mRNA exit channel in state *M* (arrow 1 in Figure 2A-D), by hinging around G1530 that is fixed by stacking onto A1507 in h45.

This 3'-end swap is accompanied by a reorganisation of helix h28: In the immature state *I*, h28 shows a non-native secondary structure, termed h28[immature], where residues U921-C924 base pair with the 16S 3´-end residues U1531-A1534 (Figure 2b), thus stabilizing the latter in the mRNA entry channel. In the mature state *M*, with the 3'-end swapped to the mRNA exit channel, h28 adopts its native secondary structure, termed h28[mature], where residues U921-C924 instead base pair with U1393-A1396 (Figure 2D) that precede the h28/h44 linker (Figure 2F).

This rRNA switch is largely consistent with aforementioned chemical probing studies by the Noller and Weeks groups that showed "a reciprocal interconversion between two differently structured states" with chemical reactivity changes "almost exclusively confined to the decoding site" (20, 22). We, therefore, surmise that the alternative inactive CDR conformation in state *I* is characteristic of the inactive 30S conformation *in vivo* and represents a stable kinetic trap, accessed by interconverting with the mature 30S subunit or during folding of the central decoding region when the assembly factor are observed to play a role in stabilizing specific switch conformations (see below).

Importantly, this RNA switch could widely occur throughout all kingdoms of life, as our analysis of small subunit rRNA sequences indicates that the potential for the 3´-end to base pair and form helix h28[immature] is largely preserved across all phylogenetic groups except in mitochondria (Figure 2G). Furthermore, the rRNA switch provides a mechanism to couple rRNA processing steps with structural maturation during 30S biogenesis. For instance, in the cryo-EM structures of the early states (*I-C*) we observe that both the 3´and 5´ends of the 16S rRNA localise

to the same side of the subunit (within the mRNA entry channel), consistent with prior work showing their interaction during the initial steps of 30S biogenesis and rRNA processing (25). Subsequent swapping of the 3´-end would enable its trimming (i.e. 17S to 16S rRNA conversion) in a process analogous to that observed in eukaryotic ribosomal assembly states, where RNAases are positioned in the exit channel to trim down the 3´-end (26).

Our cryo-EM structures indicate that this rRNA switch has far ranging effects on the 30S subunit. For example, after multibody refinement, a principal component analysis of the 30S head position relative to the body shows that the first principal component separates the dataset in a manner that reflects the h28 conformation (Figure S6). This corroborates the role of h28 as the main connection, or neck, between the 30S head and body regions (Figure 2A) such that the rRNA switch in h28 also influences the head position. Moreover, base pairing in the h28[mature] conformation restrains residues U1393-A1396 that connect h28 to the long decoding helix h44 and can, thus, influence h44 conformation and dynamics. In the most immature states *A* and *B*, for example, residues U1393-A1396 (blue) show no native base pairing interactions in h28[immature] and effectively elongate the h28/h44 linker (C1397-C1399; green), thus conveying more flexibility to h44. Accordingly, in these states, h44 is not observed in its native position on the front of the 30S (Figure 1, Table 1), and additional density in the cryo-EM map instead suggests it is repositioned into the mRNA exit channel (Figure S8). In states *I* and *C*, the elongated h28/h44 linker alternatively forms a labile helix-like structure, termed h44[a], with residues in the h44/h45 linker (A1502-G1505; brown; Figure 2B and 2E). Formation of h44[a] by both h44 linker regions allows the decoding helix to access a native-like position on the front side of the 30S subunit, but with its upper region still displaced by some 20 Å in *I* relative to its position in the mature state *M* (Figure 2E).
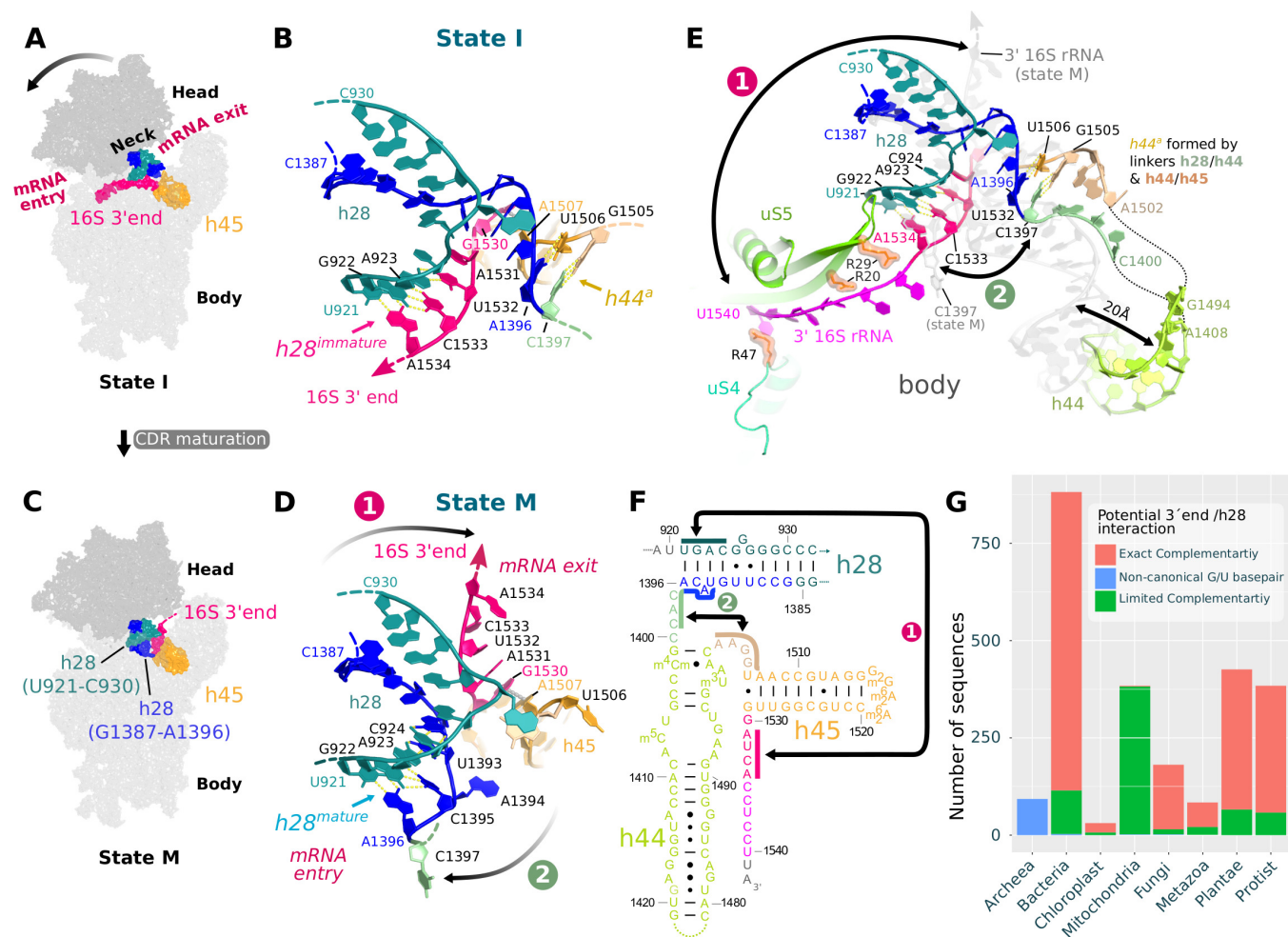
Together this suggests that, during CDR maturation, the

**Figure 2. An rRNA switch is central to CDR folding**. Two conformations of h28 are observed in states *I* (immature 30S, **A**-**B**) and *M* (mature 30S; **C**-**D**) and define an rRNA secondary structure switch that alters the central decoding region (CDR). (**A**, **C**) Schematic view of the CDR. Key rearrangements involve helix h28, which forms the connection (neck) between the 30S head and body regions and alters their relative positioning (arrow), and the 16S 3´-end (pink) that swaps between mRNA entry (state *I*; **A**) and exit (state *M*; **C**) channels. (**B**, **D**) Zoom into the neck region. The swap (arrow 1) of the 16S 3´-end (pink) between mRNA entry (**B**) and exit channels (**D**) entails a reorganisation of h28, where U921-A923 (sea green) base pair with either the 3'-end residues U1532-A1534 (pink) in the immature h28$^{immature}$ (**B**) or residues U1393-U1396 (dark blue) in the mature h28$^{mature}$ conformation (**D**). In state *I* (**B**), U1393-A1396 (dark blue) and subsequent residues C1397-C1399 (green) in the h28/44 linker can interact with A1502-U1506 in the h44/45 linker (brown) to form a labile helix-like h44$^a$ (orange; poorly defined density in the cryo-EM map) with base pairing suggested between A1396:U1506 and C1397:G1505 (Figure S8C). (**E**) Superimposed CDR structures in states *I* (coloured) and *M* (grey) highlight changes outside the h28 region, for example, a displacement of h44. The structure of the CDR in all states is shown in Figure S7, with supporting cryo-EM maps in Figure S8. (**F**) The rRNA sequences that are affected by the rRNA secondary structure switch are indicated by solid coloured bars and complementary sequences are connected with arrows, where arrow 1 indicates the base pairing in h28$^{immature}$ and arrow 2 the formation of h44$^a$. (**G**) Occurrence of small ribosomal subunit rRNA 3'-ends (U1532-A1534) capable of perfect base pairing with h28 (U921-A923) in various phylogenetic groups. Sequences with perfect or limited complementarity (1 or more mismatches) are shown by pink or green bars, respectively. *Archaea* sequences (blue bar) generally have a non-canonical G:U wobble base pair that is common in rRNA and considered isomorphic to Watson-Crick base pairs (24) and could, therefore, be considered as perfectly matching sequences.

secondary structure switch in h28 reins in residues U1393-A1396, thus shortening the h28/h44 linker and confining h44 to a more canonical position on the front of the 30S subunit (Figure 1). Such folding of both h44 linker regions into a helix-like h44$^a$ and displacement of the h44 top end is also observed in eukaryotic 40S assembly complexes (26), further indicating mechanistic parallels between eukaryotic and prokaryotic ribosome assembly and supporting the idea

that the natively isolated 30S inactive state mimics immature assembly intermediates as well.

### Roles for late-stage assembly factors in maturation of the central decoding region

In the cryo-EM reconstructions (Figure 1) we were able to visualize four of the five ribosome assembly factors added to the *in vitro* 30S complex, except for RimM that
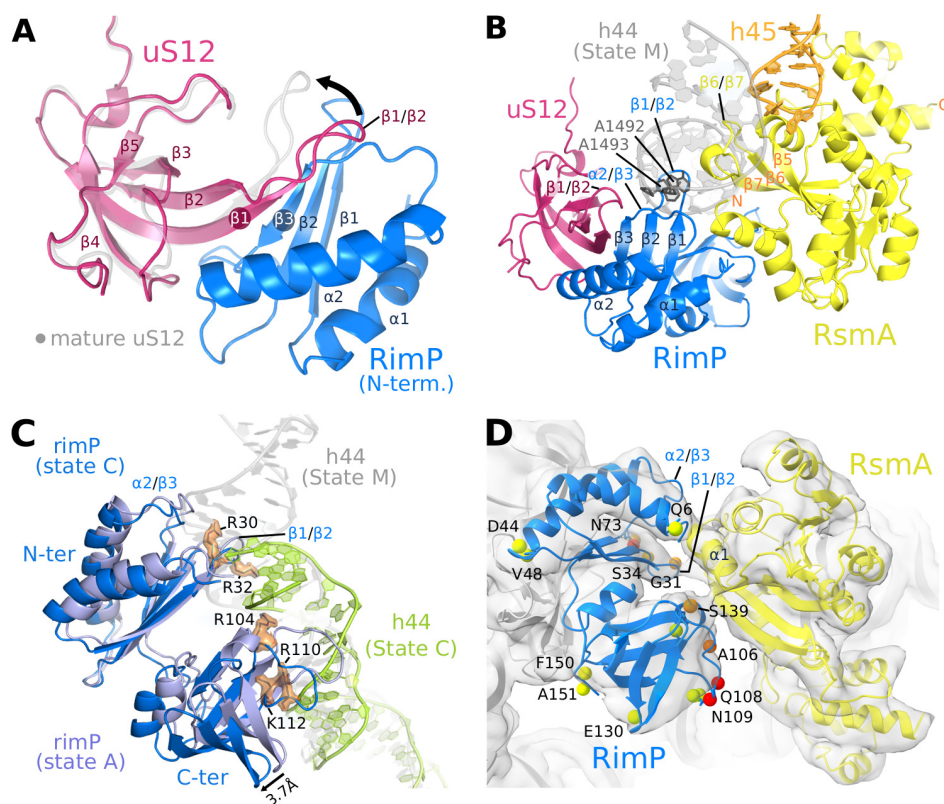
**Figure 3. Roles for the Assembly factors RsmA and RimP in CDR maturation.** (**A**) The intermolecular β-sheet formed by RimP and uS12, as observed in state *A*-*D* (state *C* shown), superimposed on uS12 of the mature 30S subunit (state *M*; grey), to highlight the conformational change in the β1/β2 loop of uS12 (arrow). (**B**) In state *B*, RimP (blue) and RsmA (yellow) bind adjacent on the 30S subunit and occupy the position of h44 in the mature 30S subunit (grey). (**C**) In state *C*, the C-terminal domain of RimP (blue; superposed on state *A* in violet) withdraws from the h44 binding site, such that the lower part of h44 (green; residues C1409 to G1491) repositions as seen in the mature 30S subunit (grey ribbon) while the upper part and linker region are disordered (not visible in the cryo-EM map). (**D**) Zoom on the RimP/RsmA interface seen by cryo-EM (state *B*). Indicated residues (spheres) were observable in the CLEANEX NMR spectrum of free RimP and are colored by the ratio κ of their $H^N/H_2O$ exchange rates in the absence ($k_{ex,free}$) and presence ($k_{ex,+RsmA}$) of RsmA, $\kappa = k_{ex,free}/k_{ex,+RsmA}$: green (κ = 0.8 - 1.2), orange (κ = 1.6 – 2.0), red (κ > 2). All RimP residues with significant solvent protection after RsmA addition (κ ≥ 1.6) localise to the RimP/RsmA interface seen in the cryo-EM structure of their 30S complex. See Figure S12 for RimP CLEANEX spectra and derived $H^N/H_2O$ exchange rates.

binds outside the CDR near uS19 and uS13 in the 30S head (27). Structural models for the 30S bound assembly factors RsmA, RimP, RbfA, and RsgA were generated by refining template coordinates, taken from the PDB (RsmA PDB ID: 1QYR; RsgA PDB ID: 5NO3) or generated *de novo* by solution state NMR (RimP and RbfA; Figure S9 and Table S13), into the cryo-EM map.

**RimP and RsmA delay h44 positioning on the front of the 30S subunit**. In the presented series of ribosomal complexes (Figure 1), the first assembly factors to bind the 30S subunit are RimP and RsmA. While RsmA, from the KsgA/Dim1 protein family, is a universally conserved rRNA methylase that modifies residues A1518 and A1519 in h45 (28), RimP is essential for survival under stress conditions in *Mycobacterium* (5, 29). Among all factors studied here, RimP is the most versatile and is seen in four assembly states (state *A* to *D*) bound on the front of the 30S subunit (blue in Figure 1). During ribosome biogenesis, RimP is known to promote binding of the

r-protein uS12 (29–31). Accordingly, our cryo-EM data shows it to interact directly with uS12 by forming an intermolecular β-sheet via anti-parallel pairing between uS12 strand β1 and RimP strand β3 (Figure 3A). This results in a conformational change in loop β1/β2 of uS12, which contains the universally conserved PNSA motif interacting with the top of h44 during decoding (32). Reorientation of this loop may therefore contribute to the disorder in the top of h44 when RimP is bound (state *A* to *D*). In state *B* (Figure 3B), RimP binds adjacent to RsmA that interacts with h24, h27, and h45, as reported previously (17). Both RimP and RsmA bind such that h44 cannot access its native position (indicated in grey) on the front of the 30S subunit. Specifically, the N-terminal domain of RimP (loop β1/β2) obstructs h44 from approaching uS12 (purple) while loop β6/β7 of RsmA keeps h44 from contacting the h45 loop (orange). In state *C* (Figure 3C), after dissociation of RsmA, the C-terminal RimP domain retracts from the native h44 binding site by

ca. 4 Å (arrow) and the lower part of h44 (green) docks onto the front of the 30S subunit. Yet, the N-terminal RimP domain (loop β1/β2) still prevents the tip of h44 (A1492-A1493) from approaching uS12 and assuming its mature fold, thus keeping the CDR in a non-native conformation. This remaining disorder and flexibility in both h44 linker regions may facilitate the swap of the 16S 3'-end, by keeping an open path for it to leave the mRNA entry channel.

RimP and RsmA were not previously known to directly interact, as suggested by our cryo-EM structures that show them binding adjacently on the ribosome. However, both proteins appear dynamic, showing reduced local resolution in the cryo-EM map (Figure S5) that limits the characterization of their shared interface. We, therefore, employed solution state NMR to probe a direct RimP/RsmA interaction, which was indeed confirmed by several observables. First, the 2D $^1$H,$^{15}$N BEST-TROSY spectrum of [U-$^{15}$N] labeled RimP showed significant global signal broadening and attenuation upon RsmA addition (Figure S10), indicating slowed molecular tumbling of RimP (16.7 kDa) from association with the considerably larger RsmA (30.4 kDa). Yet, the uniformity of signal broadening and absence of discernible chemical shift changes (indicative of a rather weak association with $K_D > 0.1$ mM) precluded a characterization of the binding interface. Second, the NMR derived apparant translational diffusion coefficient of RimP decreased from $1.09\pm0.01\cdot10^{-10}$ m$^2$/s to $0.98\pm0.03\cdot10^{-10}$ m$^2$/s (in aqueous buffer at room temperature) in the presence of 0.25 mole equivalents of RsmA, implying a ca. 11% increase in the averaged hydrodynamic radius of RimP from association with RsmA (Figure S11). Finally, the addition of RsmA also causes a clear reduction of fast H$^N$/H$_2$O exchange rates (by > 30%, i.e. $k_{ex,bound}/k_{ex,free} \leq 0.625$) for 30% (8 of 27) of the solvent exposed amide protons observed in the CLEANEX NMR spectrum (33) of RimP (Figure S11 and Table S1). These significantly solvent shielded amide groups all cluster at the RimP/RsmA interface seen in the cryo-EM model of the ternary RimP/RsmA/30S complex (Figure 3D), suggesting their similar arrangement also in the binary RimP/RsmA complex forming in solution. The interface on RsmA mainly comprises its N-terminal helix 1 and the β6/β7 loop, which may have a functional significance as this loop contains a motif (FXPXPXVXS) common to Erm and KsgA methyltransferases where the conserved phenylalanine stabilizes the substrate base through stacking interactions (28).

Extending the previously described role of RimP in promoting S12 binding (29–31), our results also indicate a broader role in 30S subunit assembly where RimP delays h44 positioning on the front of the 30S subunit, thus exposing the RsmA binding site, and may even pre-associate with RsmA to facilitate its recruitment to the subunit. Moreover, by keeping the upper part of h44 and its linker regions dis-

ordered, RimP maintains an open unstructured channel in the 30S neck region for the 16S 3'-end to swap.

**RbfA promotes the h28$^{immature}$ to h28$^{mature}$ rRNA switch.** While RimP binding is compatible with either position of the 16S 3'-end and binds similarly in states **C** and **D**, RbfA is visualized on the 30S subunit (states **D**-**E**) only after the 3'-end has swapped from the entry to the exit channel. In fact, this swapped 16S 3'-end is an integral part of the RbfA binding site in the mRNA exit channel and specifically interacts with the (type-II) KH-domain of RbfA (Figure 4A): The 16S rRNA residues G1530-A1536 run along the RNA binding surface of RbfA, where A1531-C1533 interact near the GXXG sequence motif (AXG in RbfA; (34)) characteristic of KH-domain proteins while A1534-C1535 extend into a pocket on RbfA and π-stack between the highly conserved Phe78 of RbfA and Arg43 of ribosomal protein bS18 (Figure 4A). Comparing the ribosome-bound (cryo-EM) and free (NMR) solution structures reveals some prominent structural adaptions in RbfA that promote this interaction with the 16S 3'end. Thus helix α1 rotates by about 20° and the subsequent loop α1/β1 rearranges such that the 3$_{10}$ helix present in the *apo* RbfA structure unfolds to reposition near nucleotides A1534-C1536 (Figure 4B).

The interaction of RbfA with the 16S 3'-end in the exit channel is inconsistent with previous cryo-EM reconstructions that localised RbfA on the front of the 30S subunit (16), where it cannot interact with the 16S 3'-end. Therefore, we again employed NMR to confirm this interaction using mimics of the 16S 3'-end to define its binding site on RbfA from induced chemical shift perturbations (CSP) (35) in the well dispersed 2D $^1$H,$^{15}$N HSQC fingerprint spectrum (Figure S12A). Mapping the amide groups with significant CSP on the cryo-EM structure of RbfA in the 30S complex indeed corroborates the rRNA binding surface and structural rearrangement of loop α1/β1 observed in states **D** and **E** (Figure S12C). These NMR titration experiments also show that interactions near the consensus AXG motif in RbfA are unspecific since local CSPs are induced by adding either 3'-end mimics or poly(U) control; in contrast, only 16S 3'-end mimics induce CSPs in loop α1/β1, revealing its interaction with rRNA to be specific (Figure S12B). Our NMR and cryo-EM results, therefore, indicate that RbfA facilitates the RNA secondary structure switch by sequestering residues U1531-C1535, preventing their base pairing in h28$^{immature}$, and holding the 16S 3'-end in the mRNA exit channel. This function might explain the importance of RbfA for cell growth at low temperatures (36, 37), where a h28$^{immature}$ to h28$^{mature}$ conversion may otherwise be kinetically unfavoured.

The RbfA promoted conversion from h28$^{immature}$ to h28$^{mature}$ is not sufficient, however, to complete CDR folding and, in fact, RbfA appears to even delay folding of the h44/45 linker. Besides interacting with the 16S 3'-end, RbfA packs against and distorts the upper part of h28 (arrow in Figure 4C), with its loops β1/β2 and α3/β3 acting

Schedlbauer *et al.* | A conserved rRNA switch is central to decoding site maturation
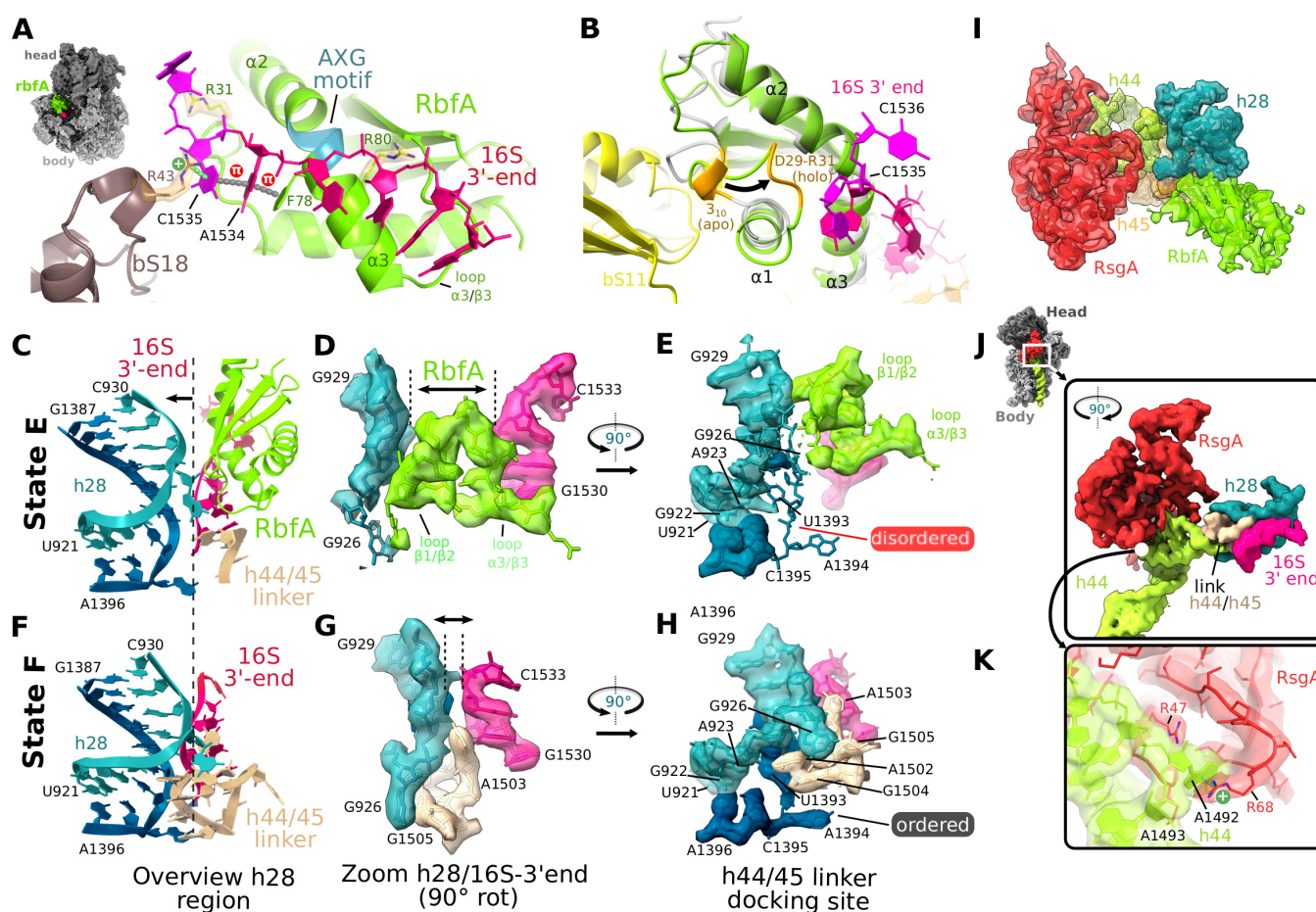
**Figure 4. Effects of RbfA on the conformation of the CDR.** (**A**) RbfA is bound within the mRNA exit channel where it captures the 16S 3′-end, e.g., via a Phe78(RbfA)-A1534-C1535-R43(bS18) π-stacking interaction (dotted lines). (**B**) Structural differences between the NMR solution structure of *apo* RbfA (grey) and 30S bound *holo* RbfA (state **E**; green) include the unfolding of a $3_{10}$ helix (orange) in the α1/β1 loop such that the loop closes around the 16S 3′-end (near A1534-C1535). (**C-H**) The h28 region in the RbfA (state **E**: panels **C**, **D**, and **E**) and RsgA (state **F**: panels **F**, **G**, **H**) bound 30S structures. RbfA inserts between 16S 3'-end and h28, displacing the upper part of h28 (arrow in **C**, compare with **F**) while destabilizing its lower end according to disorder in the cryo-EM map around residues U1393-A1394 (compare map around U1393-A1394 in **E** and **H**), thus keeping h28 from adopting a fully matured conformation. (**I**) The superimposed RbfA (green; state **D**) and RsgA (red; state **F**) cryo-EM maps (segmented density for RsgA, h44, and h28) show no overlap. (**J**) Close-up view of RsgA (red) clamping around the top of h44 in the CDR. (**K**) Corresponding detailed view showing a cation-π interaction that sandwiches A1492 (in h44; green) between Arg47 and Arg68 (in RsgA; red). The cryo-EM density shown in all panels corresponds to the consensus or multibody refinement maps for the body region and was segmented using the underlying model (radius ∼ 2.5 Å).

as a wedge to prevent the tertiary interaction seen in the mature CDR that sandwiches A1503 of the h44/h45 linker between h28 and the 16S 3′-end (compare Figures 4D and 4G). Moreover, in the presence of RbfA, the lower part of h28 around the A923:U1393 base pair is destabilised since the map around U1393-A1394 is weak and fragmented as compared to the same region in the RsgA/30S complex (compare Figures 4E and 4H). These disordered residues play an important role in stabilizing the h44/45 linker in the mature 30S where A1502-G1505 in the linker form a compact turn that packs into the minor groove of h28 around U1393-A1394 (Figure 4H and Figure S13). Finally, the observed interaction between RbfA and the 16S 3′-end is consistent with the finding that mammalian (mitochondrial) RbfA homologues interact with the 3′minor domain

of the small ribosomal subunit rRNA (38). Moreover, the interaction is similar to that observed between the eukaryotic KH-domain containing assembly factor Pno1 and the 3'-end of the 18S rRNA during the late stages of eukaryotic ribosome biogenesis (26), suggesting that RbfA may play a role similar to Pno1 in protecting and positioning the 3' end during rRNA processing. This corroborates our hypothesis that a conserved CDR core folding pathway exists across phylogenetic groups, although the protein factors involved in subunit biogenesis maybe be unrelated by sequence.

**RsgA checks CDR maturation**. RsgA is one of the last assembly factors to interact with the 30S subunit, removing RbfA and acting as a check-point for the entry of the assembling subunit into the pool of translating ribosomes (13, 18, 39). In the RsgA bound state **F** (Figure 1), the
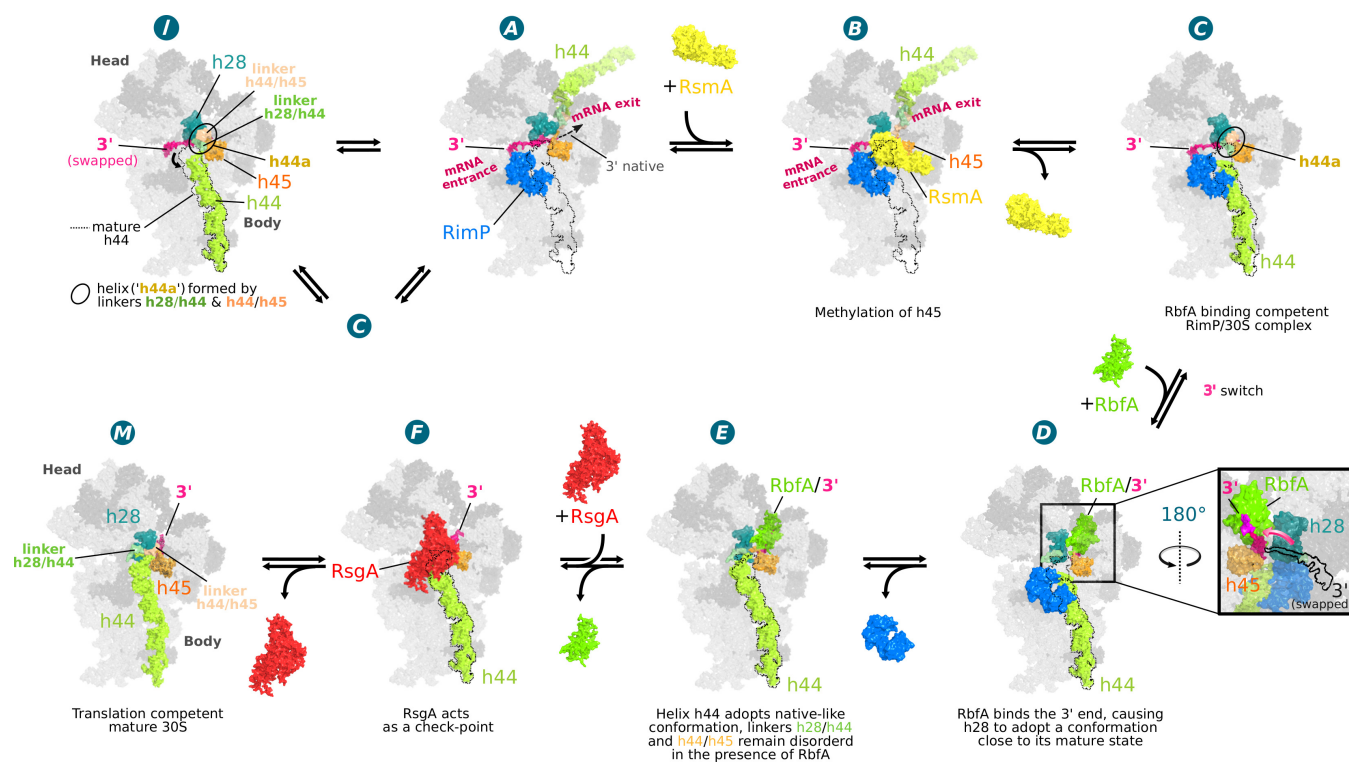
**Figure 5. Schematic model for the stepwise folding of the CDR.** Assembly factor assisted maturation of the CDR starts in state **A** with RimP (blue) bound, the 16S 3'-end (magenta) inside the mRNA entry channel (base paired in h28$^{immature}$), and the decoding helix h44 (light green) positioned in the mRNA exit channel at the back of the subunit (see discussion in Figure S8D), instead of adopting its mature position (dotted outline in all panels). This 30S configuration enables RsmA (yellow) binding in state **B** and exposes residues A1518-A1519 in the h45 loop (orange) for methylation by RsmA (28). Release of RsmA in state **C** partially opens the canonical binding site of h44, allowing it to drop into place while its two linkers form a poorly defined helix similar to h44$^a$ in state **I**. Upon RbfA binding in state **D**, the 3´-end swaps from the mRNA entry to exit channel (see rotated close-up view), where it is stabilized by interactions with the KH-domain of RbfA (green) while h28 (turquoise) adopts its mature conformation (h28$^{mature}$). In state **E**, after RimP dissociation, h44 adopts a more native-like conformation although both linker regions remain largely disordered. In state **F**, after RsgA (red) binding and release of RbfA, the h44 linker regions also fold into a nearby native conformation. Finally, after GTP hydrolysis, RsgA leaves the ribosome in a mature conformation (state **M**). In this schematic model of CDR maturation, the immature state **I** represents an inactive state (or kinetic trap) accessed, for instance, when RimP prematurely dissociates from states **A** or **C** and h44 drops into its mature position before the 3´-end can swap from the mRNA entry into the exit channel.

GTPase domain of RsgA clamps around the CDR while its zinc-binding domain forms a bridge with the 30S head near h29, as seen in previous cryo-EM reconstructions (13, 18). Importantly, the RsgA and RbfA binding sites do not overlap (Figure 4I) indicating that an allosteric effect, rather than steric hindrance, is responsible for the RsgA induced release of RbfA (39). Regarding the nature of this allosteric change, we only observe one RsgA bound state while PCA reveals that RsgA induces a 30S head conformation distinct from all other states (Figures S6E-H). Noller and colleagues showed that 30S head movement is related, in part, to a hinge that lies at a weak point in h28 near the bulged G926 (24). This is the same region pictured in Figure 4C and 4D, where RbfA loops β1/β2 and α3/β3 wedge apart h28 and the 16S 3'-end. Accordingly, RsgA induced allosteric changes in the 30S head could promote an h28 conformation incompatible with RbfA binding, thus leading to its release. The 16S 3´-end would then be free to reposition near h28 and interact with A1503 (Figure 4G) to stabilize the h44/45 linker. Moreover, by clamping around the

CDR (Figure 4J) RsgA constrains its folding landscape and further promotes h44/45 linker maturation via extensive interactions along the minor groove of h44. For instance, Arg47 and Arg68 in the OB-domain of RsgA form a pocket to fix the position of A1492 at the top of h44 presumably via cation-π stacking (Figure 4K). Such direct interactions with the CDR substantiate the previous hypothesis that RsgA acts as a checkpoint and probes the maturation state of the 30S subunit, making its GTPase activity dependent on CDR conformation (13, 18).

## Conclusion

The high-resolution cryo-EM structures presented here delineate a sequence of states (state **A**-**F**; Figure 5) through which the four analysed late-stage assembly factors RimP, RsmA, RbfA, and RsgA guide rRNA folding during CDR maturation, although variations are possible given that assembly is predicted to follow multiple pathways (8–11). A central feature of our model is a conserved rRNA secondary structure switch (h28$^{immature}$ > h28$^{mature}$ conversion) with

Schedlbauer *et al.* | A conserved rRNA switch is central to decoding site maturation

an accompanying swap of the 16S 3´-end from the mRNA entry to the exit channel. This occurs between states **C** and **D** and is assisted by RbfA that stabilizes the 16S 3´-end in its canonical position inside the mRNA exit channel, thus preventing a reversion to the alternative h28[immature] arrangement. RimP and RsmA, in turn, likely facilitate the rRNA switch during the preceding states (state **A**-**C**) by delaying the top of h44 and its linker regions from accessing their canonical position on the front of the 30S subunit. In this regard, we suggest that state **I** represents an inactive, kinetically trapped state where h44 prematurely assumes a non-canonical conformation on the front of the 30S subunit, with its linkers extended or forming a labile, non-native helix-like h44[a]. The structure of this inactive state is consistent with that derived from biochemical analysis half a century ago and, more recently, from studies *in vivo* and on purified subunits (20–22),(40). Finally, there are several apparent parallels between the prokaryotic system studied here and eukaryotic systems. For instance, RbfA has a role analogous to the eukaryotic Pno1 in binding the 3' rRNA end (26), while common rRNA sequence features indicate that the observed reorganisation of h28 can occur in all phylogenetic kingdoms except in the more distinct mitochondria. This suggests that the rRNA switch mechanism observed in our 30S complex structures is inherent to small subunit assembly throughout life.

## ACKNOWLEDGEMENTS

## Data Deposition

All cryo-EM data will be deposited in the appropriate databases including EMPIAR, PDB, EMBD. The coordinates of the isolated assembly factors studied by NMR have been deposited in PDB (7AFQ, 7AFR) and BMRB (34385, 28014).

## Bibliography

1. J. R. Warner. The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci*, 24(11):437–40, 1999. ISSN 0968-0004.
2. M. Hetman and L. P. Slomnicki. Ribosomal biogenesis as an emerging target of neurodevelopmental pathologies. *J Neurochem*, 148(3):325–347, 2019. ISSN 1471-4159. doi: 10.1111/jnc.14576.
3. K. I. Farley and S. J. Baserga. Probing the mechanisms underlying human diseases in making ribosomes. *Biochem Soc Trans*, 44(4):1035–44, 2016. ISSN 1470-8752. doi: 10.1042/BST20160064.
4. Liping Yu, Angelo H. Gunasekera, Jamey Mack, Edward T. Olejniczak, Linda E. Chovan, Xiaoan Ruan, Danli L. Towne, Claude G. Lerner, and Stephen W. Fesik. Solution structure and function of a conserved protein sp14. 3 encoded by an essential streptococcus pneumoniae gene. *Journal of molecular biology*, 311(3):593–604, 2001.
5. Poonam, R. M. Yennamalli, G. S. Bisht, and R. Shrivastava. Ribosomal maturation factor (rimp) is essential for survival of nontuberculous mycobacteria. *3 Biotech*, 9(4):127, 2019. ISSN 2190-572X. doi: 10.1007/s13205-019-1659-y.
6. B. A. Maguire. Inhibition of bacterial ribosome assembly: a suitable drug target? *Microbiol Mol Biol Rev*, 73(1):22–35, 2009. ISSN 1098-5557. doi: 10.1128/MMBR.00030-08.
7. J. M. Stokes and E. D. Brown. Chemical modulators of ribosome biogenesis as biological probes. *Nat Chem Biol*, 11(12):924–32, 2015. ISSN 1552-4469. doi: 10.1038/nchembio.1957.
8. M. W. Talkington, G. Siuzdak, and J. R. Williamson. An assembly landscape for the 30s ribosomal subunit. *Nature*, 438(7068):628–32, 2005. ISSN 1476-4687. doi: 10.1038/nature04261.
9. T. Adilakshmi, D. L. Bellur, and S. A. Woodson. Concurrent nucleation of 16s folding and induced fit in 30s ribosome assembly. *Nature*, 455(7217):1268–72, 2008. ISSN 1476-4687. doi: 10.1038/nature07298.
10. Laura M. Dutca and Gloria M. Culver. Assembly of the 5' and 3' minor domains of 16s ribosomal rna as monitored by tethered probing from ribosomal protein s20. *Journal of molecular biology*, 376:92–108, February 2008. ISSN 1089-8638. doi: 10.1016/j.jmb.2007.10.083.
11. A. M. Mulder, C. Yoshioka, A. H. Beck, A. E. Bunner, R. A. Milligan, C. S. Potter, B. Carragher, and J. R. Williamson. Visualizing ribosome biogenesis: parallel assembly pathways for the 30s subunit. *Science*, 330(6004):673–7, 2010. ISSN 1095-9203. doi: 10.1126/science.1193220.
12. S. A. Woodson. Recent insights on rna folding mechanisms from catalytic rna. *Cell Mol Life Sci*, 57(5):796–808, 2000. ISSN 1420-682X. doi: 10.1007/s000180050042.
13. A. Razi, A. Guarné, and J. Ortega. The cryo-em structure of yjeq bound to the 30s subunit suggests a fidelity checkpoint function for this protein in ribosome assembly. *Proc Natl Acad Sci U S A*, 114(17):E3396–E3403, 2017. ISSN 1091-6490. doi: 10.1073/pnas.1618016114.
14. A. P. Carter, W. M. Clemons, D. E. Brodersen, R. J. Morgan-Warren, B. T. Wimberly, and V. Ramakrishnan. Functional insights from the structure of the 30s ribosomal subunit and its interactions with antibiotics. *Nature*, 407(6802):340–8, 2000. ISSN 0028-0836. doi: 10.1038/35030019.
15. Tatsuya Kaminishi, Daniel N. Wilson, Chie Takemoto, Joerg M. Harms, Masahito Kawazoe, Frank Schluenzen, Kyoko Hanawa-Suetsugu, Mikako Shirouzu, Paola Fucini, and Shigeyuki Yokoyama. A snapshot of the 30s ribosomal subunit capturing mrna via the shine-dalgarno interaction. *Structure (London, England : 1993)*, 15:289–297, March 2007. ISSN 0969-2126. doi: 10.1016/j.str.2006.12.008.
16. Partha P. Datta, Daniel N. Wilson, Masahito Kawazoe, Neil K. Swami, Tatsuya Kaminishi, Manjuli R. Sharma, Timothy M. Booth, Chie Takemoto, Paola Fucini, Shigeyuki Yokoyama, and Rajendra K. Agrawal. Structural aspects of rbfa action during small ribosomal subunit assembly. *Molecular cell*, 28:434–445, November 2007. ISSN 1097-2765. doi: 10.1016/j.molcel.2007.08.026.
17. Daniel Boehringer, Heather C. O'Farrell, Jason P. Rife, and Nenad Ban. Structural insights into methyltransferase ksga function in 30s ribosomal subunit biogenesis. *Journal of Biological Chemistry*, 287(13):10453–10459, 2012. doi: 10.1074/jbc.M111.318121.
18. J. P. López-Alonso, T. Kaminishi, T. Kikuchi, Y. Hirata, I. Iturrioz, N. Dhimole, A. Schedlbauer, Y. Hase, S. Goto, D. Kurita, A. Muto, S. Zhou, C. Naoe, D. J. Mills, D. Gil-Carton, C. Takemoto, H. Himeno, P. Fucini, and S. R. Connell. Rsga couples the maturation state of the 30s ribosomal decoding center to activation of its gtpase pocket. *Nucleic Acids Res*, 45(11):6945–6959, 2017. ISSN 1362-4962. doi: 10.1093/nar/gkx324.
19. D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, and P. D. Adams. Macromolecular structure determination using x-rays, neutrons and electrons: recent developments in phenix. *Acta Crystallogr D Struct Biol*, 75(Pt 10):861–877, 2019. ISSN 2059-7983. doi: 10.1107/S2059798319011471.
20. D. Moazed, B. J. Van Stolk, S. Douthwaite, and H. F. Noller. Interconversion of active and inactive 30 s ribosomal subunits is accompanied by a conformational change in the decoding region of 16 s rrna. *J Mol Biol*, 191(3):483–93, 1986. ISSN 0022-2836. doi: 10.1016/0022-2836(86)90143-9.
21. A. Zamir, R. Miskin, and D. Elson. Interconversions between inactive and active forms of ribosomal subunits. *FEBS Lett*, 3(1):85–88, 1969. ISSN 1873-3468. doi: 10.1016/0014-5793(69)80103-1.
22. J. L. McGinnis, Q. Liu, C. A. Lavender, A. Devaraj, S. P. McClory, K. Fredrick, and K. M. Weeks. In-cell shape reveals that free 30s ribosome subunits are in the inactive state. *Proc Natl Acad Sci U S A*, 112(8):2425–30, 2015. ISSN 1091-6490. doi: 10.1073/pnas.1411514112.
23. T. Nakane, D. Kimanius, E. Lindahl, and S. H. Scheres. Characterisation of molecular motions in cryo-em single-particle data by multi-body refinement in relion. *Elife*, 7, 2018. ISSN 2050-084X. doi: 10.7554/eLife.36861.
24. G. Varani and W. H. McClain. The g x u wobble base pair. a fundamental building block of rna structure crucial to rna function in diverse biological systems. *EMBO Rep*, 1(1):18–23, 2000. ISSN 1469-221X. doi: 10.1093/embo-reports/kvd001.
25. Z. Yang, Q. Guo, S. Goto, Y. Chen, N. Li, K. Yan, Y. Zhang, A. Muto, H. Deng, H. Himeno, J. Lei, and N. Gao. Structural insights into the assembly of the 30s ribosomal subunit in vivo: functional role of s5 and location of the 17s rrna precursor sequence. *Protein Cell*, 5(5):394–407, 2014. ISSN 1674-8018. doi: 10.1007/s13238-014-0044-1.
26. M. Ameismeier, J. Cheng, O. Berninghausen, and R. Beckmann. Visualizing late states of human 40s ribosomal subunit maturation. *Nature*, 558(7709):249–253, 2018. ISSN 1476-4687. doi: 10.1038/s41586-018-0193-0.
27. Q. Guo, S. Goto, Y. Chen, B. Feng, Y. Xu, A. Muto, H. Himeno, H. Deng, J. Lei, and N. Gao. Dissecting the in vivo assembly of the 30s ribosomal subunit reveals the role of rimm and general features of the assembly process. *Nucleic Acids Res*, 41(4):2609–20, 2013. ISSN 1362-4962. doi: 10.1093/nar/gks1256.
28. H. C. O'Farrell, J. N. Scarsdale, and J. P. Rife. Crystal structure of ksga, a universally

conserved rrna adenine dimethyltransferase in escherichia coli. *J Mol Biol*, 339(2): 337–53, 2004. ISSN 0022-2836. doi: 10.1016/j.jmb.2004.02.068.

29. T. Chu, X. Weng, C. O. K. Law, H. K. Kong, J. Lau, S. Li, H. Q. Pham, R. Wang, L. Zhang, R. Y. T. Kao, K. F. Lau, J. C. K. Ngo, and T. C. K. Lau. The ribosomal maturation factor p from. *J Biol Chem*, 294(1):372–378, 2019. ISSN 1083-351X. doi: 10.1074/jbc.RA118.002298.

30. D. G. Sashital, C. A. Greeman, and J. R. Williamson. A combined quantitative mass spectrometry and electron microscopy analysis of ribosomal 30s subunit assembly in e. coli. *eLife*, 3, 2014.

31. Anne E. Bunner, Stefan Nord, P. Mikael Wikström, and James R. Williamson. The effect of ribosome assembly cofactors on in vitro 30s subunit reconstitution. *Journal of molecular biology*, 398:1–7, 2010. ISSN 1089-8638.

32. I. M. Sharma and S. A. Woodson. Rbfa and if3 couple ribosome biogenesis and translation initiation to increase stress tolerance. *Nucleic Acids Res*, 48(1):359–372, 2020. ISSN 1362-4962. doi: 10.1093/nar/gkz1065.

33. T. L. Hwang, P. C. van Zijl, and S. Mori. Accurate quantitation of water-amide proton exchange rates using the phase-modulated clean chemical exchange (cleanex-pm) approach with a fast-hsqc (fhsqc) detection scheme. *J Biomol NMR*, 11(2):221–6, 1998. ISSN 0925-2738. doi: 10.1023/a:1008276004875.

34. Y. J. Huang, G. V. Swapna, P. K. Rajan, H. Ke, B. Xia, K. Shukla, M. Inouye, and G. T. Montelione. Solution nmr structure of ribosome-binding factor a (rbfa), a cold-shock adaptation protein from escherichia coli. *J Mol Biol*, 327(2):521–36, 2003. ISSN 0022-2836. doi: 10.1016/s0022-2836(03)00061-5.

35. M. P. Williamson. Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Magn Reson Spectrosc*, 73:1–16, 2013. ISSN 1873-3301. doi: 10.1016/j.pnmrs.2013.02.001.

36. Pamela G. Jones and Masayori Inouye. Rbfa, a 30s ribosomal binding factor, is a cold-shock protein whose absence triggers the cold-shock response. *Molecular Microbiology*, 21(6):1207–1218, 1996. ISSN 0950-382X. doi: 10.1111/j.1365-2958. 1996.tb02582.x.

37. C. S. Dammel and H. F. Noller. Suppression of a cold-sensitive mutation in 16s rrna by overexpression of a novel ribosome-binding factor, rbfa. *Genes Dev*, 9(5):626–37, 1995. ISSN 0890-9369. doi: 10.1101/gad.9.5.626.

38. A. Rozanska, R. Richter-Dennerlein, J. Rorbach, F. Gao, R. J. Lewis, Z. M. Chrzanowska-Lightowlers, and R. N. Lightowlers. The human rna-binding protein rbfa promotes the maturation of the mitochondrial ribosome. *Biochem J*, 474(13): 2145–2158, 2017. ISSN 1470-8728. doi: 10.1042/BCJ20170256.

39. S. Goto, S. Kato, T. Kimura, A. Muto, and H. Himeno. Rsga releases rbfa from 30s ribosome during a late stage of ribosome biosynthesis. *EMBO J*, 30(1):104–14, 2011. ISSN 1460-2075. doi: 10.1038/emboj.2010.291.

40. Dushyant Jahagirdar, Vikash Jha, Kaustuv Basu, Josue Gomez-Blanco, Javier Vargas, and Joaquin Ortega. Alternative conformations and motions adopted by 30s ribosomal subunits visualized by cryo-electron microscopy. *bioRxiv*, page 2020.03.21.001677, 2020. doi: 10.1101/2020.03.21.001677.