BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

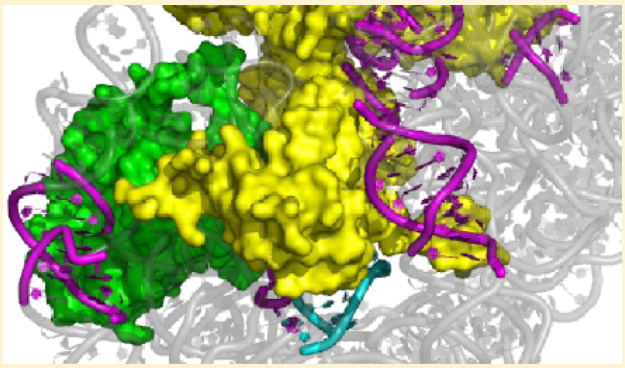
pubs.acs.org/biochemistry

# Ribosome RNA Assembly Intermediates Visualized in Living Cells

Jennifer L. McGinnis and Kevin M. Weeks*

Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290, United States

**ABSTRACT:** In cells, RNAs likely adopt numerous intermediate conformations prior to formation of functional RNA−protein complexes. We used single-nucleotide resolution selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) to probe the structure of *Escherichia coli* 16S rRNA in healthy growing bacteria. SHAPE-directed modeling indicated that the predominant steady-state RNA conformational ensemble in dividing cells had a base-paired structure different from that expected on the basis of comparative sequence analysis and high-resolution studies of the 30S ribosomal subunit. We identified the major cause of these differences by stopping ongoing in-cell transcription (in essence, an in-cell RNA structure pulse−chase experiment) which caused the RNA to chase into a structure that closely resembled the expected one. Most helices that formed alternate RNA conformations under growth conditions interact directly with tertiary-binding ribosomal proteins and form a C-shape that surrounds the mRNA channel and decoding site. These in-cell experiments lead to a model in which ribosome assembly factors function as molecular struts to preorganize this intermediate and emphasize that the final stages of ribonucleoprotein assembly involve extensive protein-facilitated RNA conformational changes.



After transcription, RNA molecules undergo complex processing, folding, and protein and ligand association reactions essential for cellular function. The conformations of many RNAs have been well-defined by biochemical and high-resolution structural determination approaches; however, RNA folding and RNA−protein assembly reactions have only recently been visualized at high resolution in living cells.[1−4] Under optimal growth conditions, cells spend significant energy to assemble the ribosomes that synthesize the proteome necessary for all phases of cellular function.[5] In bacteria, the ribosome is composed of 30S and 50S subunits that ultimately assemble to form the 70S ribosome. The 1542-nucleotide 16S rRNA of *Escherichia coli* is the largest component of the 30S subunit and comprises a large fraction of the total RNA in a bacterium.[6]

Biogenesis of the 30S ribosomal subunit involves a complex series of interrelated events. The precursor 16S rRNA molecule is transcribed, processed via endonucleolytic cleavage to the mature full-length form, and covalently modified at specific, conserved sites. The RNA folds, and ribosomal proteins bind in a semi-ordered fashion to form the 30S subunit. Upon association with initiation factors and the 50S subunit, 30S subunits become part of the 70S ribosome. This large complex associates with elongation factors, tRNA, and mRNA to conduct protein synthesis. Ultimately, ribosomal subunits are disassembled and constituent RNAs are degraded.[6−9]

Significant strides have been made toward understanding ribosome assembly through *in vitro* experiments,[10−12] and recent work has explored the protein and RNA compositions of assembly intermediates in cells.[2−4] *In vitro* and in-cell assembly mechanisms differ. First, *in vitro* assembly experiments are generally performed using full-length rRNA, whereas cellular assembly occurs cotranscriptionally. Second, the rates of 30S assembly and cell-free translation *in vitro* are dramatically slower than the rates *in vivo*.[6,13] Third, 30S assembly *in vitro* requires a sharp temperature increase[6] and specialized buffer systems.[14] Despite their importance, the relationships between in-cell RNA conformation and extensive characterizations *in vitro* are not well understood.

Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) quantitatively measures local nucleotide flexibility at all four nucleotides in an RNA.[15,16] The most useful and structurally informative SHAPE reagent developed to date is 1M7. The reactivity of 1M7 is strongly correlated with model-free measures of molecular motion,[17] and incorporation of 1M7 reactivities as a pseudo-free energy term into structure prediction algorithms yields accurate RNA secondary structure models.[18,19] 1M7 is a small, slightly hydrophobic molecule that readily crosses cell membranes.[1,20] SHAPE reagents are initially suspended in a nonreactive organic solvent, such as dimethyl sulfoxide (DMSO), to prevent premature hydrolysis.[21] Bacterial cells grown in 2−4% DMSO achieve growth rates and final cell densities comparable to those of cells grown in the absence of DMSO.[1] Thus, SHAPE can be performed *in vivo* without impacting bacterial cell viability.

In this work, we use SHAPE to visualize the structure of 16S rRNA in ribosome assembly intermediates in healthy growing

**Received:** February 17, 2014
**Revised:** April 21, 2014
**Published:** May 12, 2014

   dx.doi.org/10.1021/bi500198b | *Biochemistry* 2014, 53, 3237−3247

bacterial cells. We discovered that widespread conformational changes occur in the RNA during late stages of ribosome assembly. Together with other recent work, in-cell SHAPE analysis supports a simple "molecular strut" model for how late-acting ribosome assembly factors facilitate 30S assembly.

## METHODS

**Lifetime of 1M7 in Growth Media.** Hydrolysis of 1M7 was followed by addition of 2.0 mM 1M7 in 100 μL of DMSO to 900 μL of Luria broth (LB) (pH 6.5 or 7.0) pre-equilibrated at 37 °C in a cuvette. Pseudo-first-order rate constants were obtained by monitoring the absorbance of the hydrolysis product (2-methylamino-4-nitrobenzoate) at 430 nm.[1,21]

**SHAPE Analysis of the 16S rRNA Structure *in Vivo*.** DH5α cells (500 mL) were grown in LB to an $OD_{600}$ of ∼0.6 starting from 5 mL of an overnight culture. RNA modification was initiated by adding an aliquot of cell culture (14 mL) to 433 μL of 167 mM 1M7 in DMSO (final 1M7 concentration of 5 mM). The control cell sample was treated with neat DMSO [3% (v/v)]. Samples were allowed to react while being shaken for 5 min at 37 °C. An untreated aliquot of cell culture was obtained for dideoxy sequencing. Cells were collected by centrifugation (4000*g* for 20 min at 4 °C). Cell pellets were resuspended in 200 μL of TE buffer containing 30 mg/mL lysozyme and incubated on ice for 5 min. RNA was purified by affinity chromatography (RNeasy Mini column, Qiagen). RNA concentrations were determined by the absorbance at 260 nm. For all samples, the purified RNA contained total cellular RNA. For experiments in which transcription was inhibited, cell growth was initiated as described above and, at an $OD_{600}$ of ∼0.5, 50 mL of cells was added to either 5.55 mL of rifampicin in $H_2O$ (final concentration of 18.75 μg/mL) or 5.55 mL of $H_2O$. Cells were incubated while being vigorously shaken for 20 min, and in-cell SHAPE was performed as described above.

**Primer Extension.** Five primer sets were used to analyze the 16S rRNA with the 3′ positions of the primers complementary to positions 323, 559, 947, 1112, and 1492.[16] We were ultimately able to examine 94% of the nucleotides in the 16S rRNA but did not analyze ∼65 nucleotides at the very 3′ end of the RNA. Primer extension was performed as described previously[16,22] except that modified total cellular RNA (2.1 μg) and primer (2 pmol) were mixed and diluted to 13 μL with water prior to the primer annealing step. Dideoxy sequencing ladders were produced using unlabeled, unmodified total cellular RNA (in 12 μL) and 1 μL of 2′,3′-dideoxycytidine (5 mM) triphosphate. cDNA fragments were separated by capillary electrophoresis using an Applied Biosystems 3500 capillary electrophoresis instrument.

**Data Analysis.** Capillary electrophoresis traces were processed using QuShape.[23] Data sets were normalized using model-free statistics to a scale of 0 to ∼2. Nucleotides with normalized SHAPE reactivities of 0−0.4, 0.4−0.85, and >0.85 correspond to unreactive, moderately reactive, and highly reactive positions, respectively. Each experiment was performed at least twice, and data were reproducible.

**RNA Structure Modeling.** Alternative secondary structure models were generated using RNAstructure[24] in conjunction with SHAPE pseudo-free energy change constraints.[18,19]

**Solvent Accessibility.** The solvent accessibility surface area (SASA) of the 2′-hydroxyl group was calculated with NACCESS version 2.1.1 (http://www.bioinf.machester.ac.uk/naccess/) using a probe size of 2.8 Å and the atomic coordinates from Protein Data Bank (PDB) entries 3i1m (30S) and 3i1n (50S).[25] Nucleotides were taken to be surface nucleotides if the SASA was greater than 5 $Å^2$. Nucleotides at the subunit interface of the 30S were defined as those that became inaccessible when the 50S subunit was added. Nucleotides in the 16S rRNA in direct contact with proteins were defined as those that became inaccessible when 30S proteins were added. Nucleotides in the 16S rRNA involved in RNA−RNA domain interactions were defined as nucleotides that became inaccessible when the other domains were included. Domains of the 16S rRNA were defined as 5′ (nucleotides 1−566), central (nucleotides 567−912), 3′ major (nucleotides 913−1396), and 3′ minor (nucleotides 1397−1542).

## RESULTS

**Cell Viability, Time Scale, and Quality of *in Vivo* SHAPE.** The hydrolysis half-life of 1M7 in standard growth medium during the midlog phase of cell growth is 24 s (Figure 1A,B); under these conditions, the SHAPE reaction will be complete in approximately 2 min (five hydrolysis half-lives). This is an ideal time scale because the 1M7 reagent is sufficiently stable for straightforward experimental use but makes rapid and facile in-cell structure analysis possible.

*In vivo* SHAPE was performed by growing *E. coli* cell cultures to midlog phase ($OD_{600}$ = 0.6) and adding an aliquot of cells directly to 1M7 in DMSO [3% (v/v), final concentration]. After 2 min, cells were pelleted and lysed, the cellular debris and proteins were removed, and total *E. coli* cellular RNA was purified. Sites of *in vivo* SHAPE adduct formation in the 1542-nucleotide 16S rRNA were detected by primer extension and resolved by capillary gel electrophoresis using five sets of DNA primers.[16] Quantitative SHAPE reactivity data were obtained for 1450 nucleotides (94%) of the 16S rRNA. The quality of the single-nucleotide resolution SHAPE data was very high as evidenced by a low level of background in the no-reagent control reaction and by the perfect agreement between the RNA-based sequencing reaction and the known 16S rRNA sequence (Figure 1C,D). The 1M7 SHAPE reagent thus readily crossed the complex double membrane of the Gram-negative *E. coli* bacterial cell to yield an ∼2 min resolution snapshot of the conformational dynamics of almost all nucleotides in the 16S rRNA.

**The 16S rRNA Structure in Rapidly Growing Cells Differs from the Accepted Structure.** When the *in vivo* SHAPE reactivities were superimposed on the accepted secondary structure model[26,27] of the *E. coli* 16S rRNA, the majority of reactivities showed excellent agreement with the accepted secondary structure (Figure 2, gray highlights). In general, single-stranded nucleotides in the accepted structure showed medium or high reactivities, whereas most nucleotides expected to be base paired were unreactive by SHAPE.

A small number of nucleotides predicted to be single-stranded in the accepted secondary structure model had low SHAPE reactivities. Because SHAPE measures molecular motion,[17] it is likely these nucleotides are involved in higher-order tertiary contacts or interactions with proteins. These unreactive nucleotides in the 16S rRNA were examined in the context of the three-dimensional structure of the 70S ribosome.[25] Almost all of these unreactive nucleotides are indeed involved in RNA tertiary interactions or are located within or near protein-binding sites (Figure 2, black stars). For example, unreactive nucleotides 132−135 and 260−265 interact via direct RNA−RNA contacts. Similarly, the unreactive loop nucleotides at positions 260−265, 321−323, and 326−330 are all located near the binding


dx.doi.org/10.1021/bi500198b | *Biochemistry* 2014, 53, 3237−3247

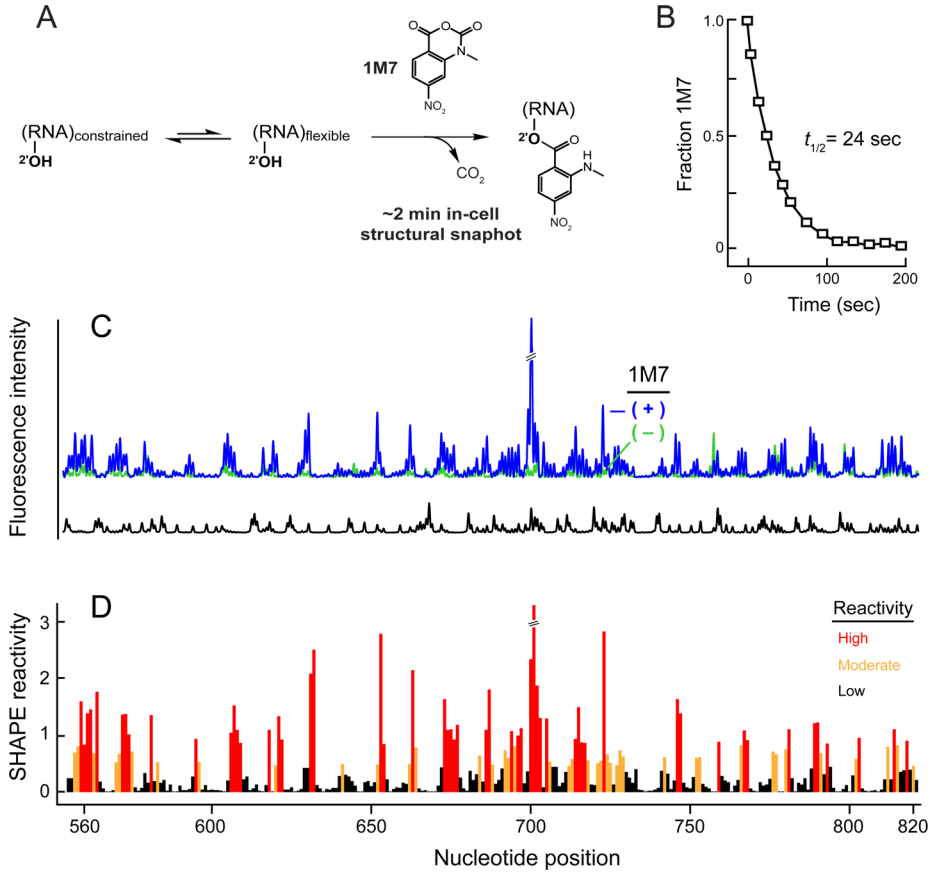



**Figure 1.** *In vivo* SHAPE of 16S rRNA. (A) Schematic for using 1M7 to obtain ~2 min structural snapshots *in vivo*. (B) Pseudo-first-order hydrolysis of 1M7 in Luria broth (37 °C) at pH 7.0, corresponding to the midlog phase of cell growth. This panel was adapted from ref 1. (C) Electropherogram resulting from in-cell SHAPE probing of 16S rRNA. The region shown lies in the central domain. Data for reactions performed in the presence and absence of 1M7 are colored blue and green, respectively; the dideoxy sequencing (ddC) trace is colored black. (D) Histogram of normalized SHAPE reactivities.[23]

site for the S20 protein, one of the first proteins to bind during 30S assembly.[6]

Approximately 10% of 16S rRNA nucleotides predicted to be constrained by base pairing in the accepted structure exhibited medium or high SHAPE reactivities. These nucleotides were not spread uniformly throughout the RNA but were clustered in a small number of helices, most of which were located in irregular or multihelix junctions (Figure 2, magenta and cyan boxes). Of the 14 helices in which these nucleotides are located, 10 have reactive nucleotides on both strands of the helix (H2, H3, H16/H17, H26b, H28, H31, H36, H41, and H44), making it unlikely that these helices are stably formed in the steady-state 16S rRNA structure *in vivo* (Figure 2, magenta boxes). In four of the 14 regions (H12, H33b, H34, and H38), SHAPE-reactive nucleotides are adjacent to unreactive nucleotides that are single-stranded in the accepted secondary structure. In these regions, the predominant pattern of base pairing *in vivo* likely differs from that in the accepted secondary structure model.

We identified plausible alternative structures in these discordant regions by using the in-cell SHAPE data to model local RNA folding[18,19] (Figure 2, cyan boxes, alternate base pairs shown with dark blue lines). For example, in helix 34 in the accepted secondary structure, nucleotides 1055 and 1200−1202 are bulged; these nucleotides were expected to be reactive to the SHAPE reagent but instead are unreactive *in vivo*. In contrast, nucleotides 1053, 1204, and 1205 are base paired in the accepted structure but are reactive *in vivo*. In the alternate SHAPE-directed structure model, nucleotides 1200−1202 pair with nucleotides 1057−1055 (Figure 2, dark blue lines in H34), and nucleotides 1053, 1054, and 1203−1205 are bulged. In-cell SHAPE also suggests local base pairing rearrangements relative to the accepted structure in helices H12, H33, and H38 (Figure 2, cyan boxes). In each case, the alternative base pairing patterns are more consistent with the SHAPE-measured *in vivo* nucleotide dynamics than is the accepted structural model.

At midlog phase, *E. coli* cells are dividing actively, and levels of rRNA synthesis are high. Ribosome components, including the rRNA, are present at multiple stages of biogenesis. Differences observed between the SHAPE-directed model and the accepted secondary structure likely reflect the fact that the snapshot of the 16S rRNA taken by SHAPE in rapidly growing cells reflects the conformations of intermediates in ribosome assembly.

**Halting Transcription Converts the 16S rRNA Structure to Roughly the Accepted One.** Halting transcription will stop production of new 16S rRNA transcripts and should allow the16S rRNA to be incorporated into 30S subunits and 70S ribosomes. To evaluate the effect of transcription inhibition on 16S rRNA structure, we performed a kinetic "chase" experiment at nucleotide resolution *in vivo*. We added the antibiotic rifampicin, which inhibits transcription,[28] to cells at midlog phase ($OD_{600}$ = 0.5). Rifampicin inhibits transcription by binding to the bacterial RNA polymerase but does not bind to the ribosome. Changes in SHAPE reactivity of the 16S rRNA will thus reflect the effects of arrested transcription rather than direct


dx.doi.org/10.1021/bi500198b | *Biochemistry* 2014, 53, 3237−3247

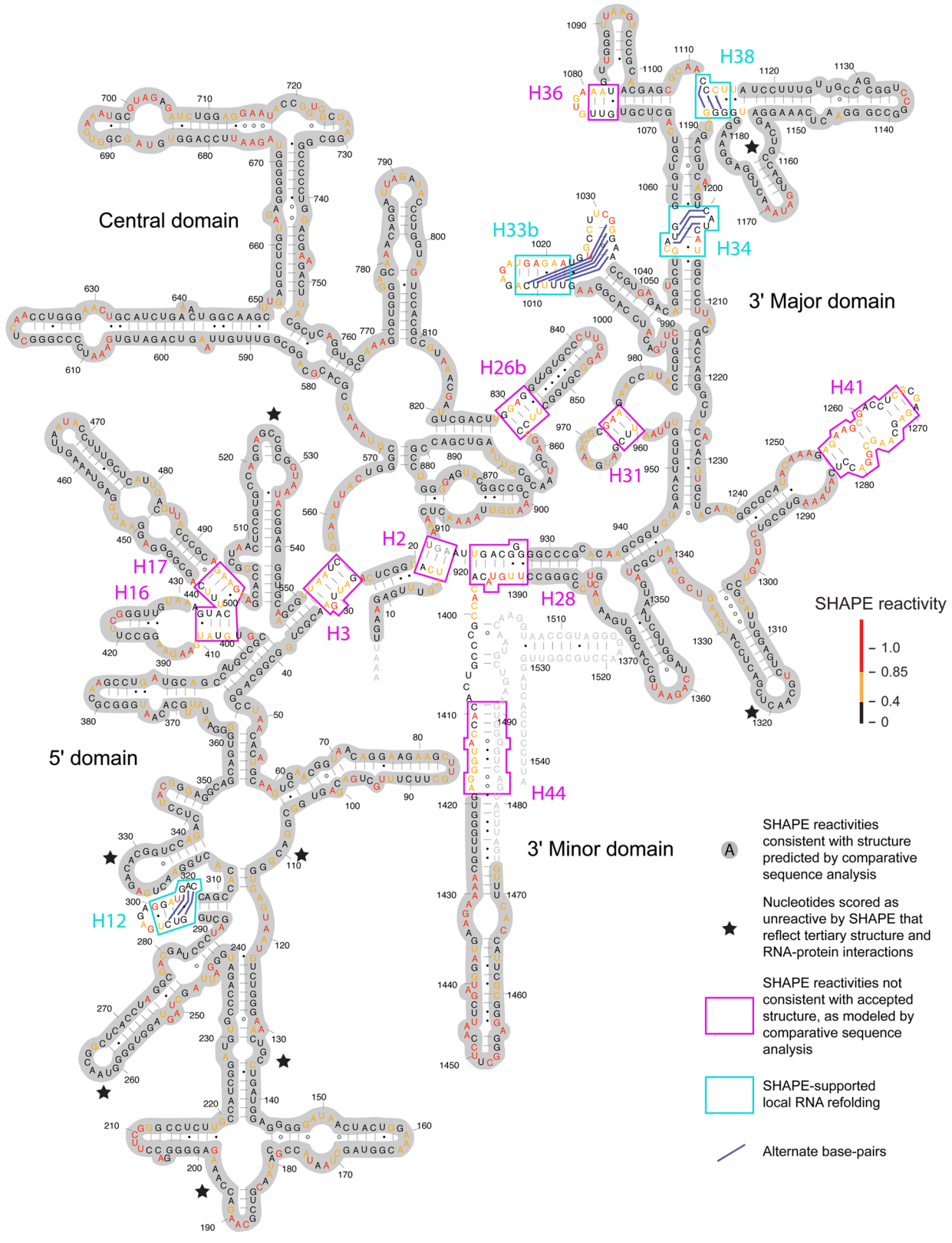



**Figure 2.** *In vivo* SHAPE reactivities superimposed on the *E. coli* 16S secondary structure model determined by comparative sequence analysis. Regions in which SHAPE data are consistent with the phylogenetically[26] and crystallographically[25] supported structures are highlighted in gray. Helices where SHAPE data are not consistent with the secondary structure model are boxed. Alternate base pair pairs, supported by SHAPE probing, are emphasized with dark blue lines. Nucleotide letters are colored by SHAPE reactivity. Nucleotides for which no data were obtained are shown in gray.

effects on rRNA structure. Rifampicin halts transcription initiation within 70 s of addition. It requires approximately 2.2 min to transcribe an rRNA operon[29] and an additional ~2 min for the 16S rRNAs to be processed and incorporated into 30S subunits.[13] When cells were incubated with rifampicin for either 20 or 60 min prior to modification with 1M7, SHAPE reactivity


dx.doi.org/10.1021/bi500198b | *Biochemistry* 2014, 53, 3237−3247

profiles were essentially identical. Therefore, we analyzed SHAPE reactivity profiles for the 16S rRNA 20 min after the addition of rifampicin (approximately five half-lives for rRNA synthesis and assembly).

During rRNA maturation, RNase III cleaves the primary 16S rRNA transcript to generate a form of the 16S rRNA precursor (the 17S rRNA) that contains an additional 115 and 33 nucleotides at the 5′ and 3′ ends, respectively, relative to the mature form.[6] The amount of 17S precursor relative to mature 16S rRNA was quantified by primer extension analysis across the 5′ terminus of the mature rRNA.[30] In log phase cells without the addition of an antibiotic, 16% of the total 16S rRNA contained an unprocessed 5′ end. After incubation with rifampicin for 20 min, the percentage of RNA with an unprocessed 5′ end decreased to 2.5%. Thus, midlog phase *E. coli* cells contain a significant fraction of incompletely processed 16S rRNA, and the addition of rifampicin promotes nearly complete conversion to the fully processed rRNA.

After inhibition of transcription, SHAPE reactivities for many nucleotides in the 16S rRNA differed from those measured in the absence of rifampicin. Of the nucleotides with moderate or high reactivity in the absence of rifampicin (defined as ≥0.4), more than half (273 of 509) exhibited a significant decrease (>0.3 SHAPE unit) in reactivity upon addition of rifampicin (Figure 3, highlighted in blue). Overall, the 16S rRNA becomes much more structured after transcription arrest. Of the 273 nucleotides that decreased by more than 0.3 SHAPE reactivity unit upon inhibition of transcription, 58 (∼20%) changed in a way that resolved discrepancies between the SHAPE-directed model and accepted secondary structure models (Figure 3, magenta and cyan boxes). For example, the mean *in vivo* SHAPE reactivity of nucleotides in the H33b helix decreased from 0.5 to 0.2 when transcription was halted, consistent with formation of the base pairs observed in the accepted secondary structure. Transcription arrest ultimately yielded an in-cell SHAPE profile that was ∼99% consistent with the accepted secondary structure. A small number (17 or ∼1%) of 16S rRNA nucleotides increased by more than 0.3 SHAPE reactivity unit upon addition of rifampicin (Figure 3, circled in orange). All of these nucleotides are in single-stranded or weakly paired regions, suggestive of changes in loop dynamics that accompanied formation of higher levels of overall structure in the 16S rRNA.[16,17]

**Strong 5′ to 3′ Polarity in Rifampicin-Induced Changes in 16S rRNA Structure.** Roughly 20 times more nucleotides showed decreases in SHAPE reactivity versus increases upon addition of rifampicin. There is a clear 5′ to 3′ polarity in the differences between SHAPE reactivity in the absence and presence of rifampicin (Figure 4A). As averaged over 51-nucleotide windows, relatively few nucleotides in the 5′ domain differed in SHAPE reactivity when reactivity profiles in the absence and presence of rifampicin are compared, but the reactivity of 30−35% of the nucleotides in the 3′ major and 3′ minor domains decreased significantly in the absence of transcription. In addition, local peaks in the number of reactivity differences occur precisely at the junctions between domains of the 16S rRNA (Figure 4A, vertical dashed lines). The positions that become less reactive to 1M7 in cells upon rifampicin treatment are superimposed on the three-dimensional structure of the 16S rRNA (Figure 4B). The observed polarity in structural differences is consistent with an ordered ribosomal assembly process and with cotranscriptional folding and assembly into 30S subunits.

We next examined the role that higher-order interactions play in the in-cell assembly process by identifying nucleotides that are accessible to the solvent in the context of each folded domain. SHAPE reactivities are not fundamentally governed by solvent or molecular accessibility: they are instead determined by local nucleotide dynamics.[16,17] However, we used solvent accessibility as a metric to identify nucleotides located on the exterior of each domain in the 16S rRNA. We then identified the nucleotides that differed in reactivity in the presence and absence of rifampicin that became protected from solvent in the context of the crystal structure of the 70S ribosome and assigned these as predominately reflecting interdomain RNA−RNA interactions (Figure 5A, green spheres, 50 total), RNA−protein interactions (yellow spheres, 47 total), and interactions with the 50S subunit (purple spheres, 7 total). Most nucleotides whose protection likely reflects direct RNA−RNA domain interactions are located in the central domain and make contact with the 3′ minor domain (Figure 5A, green spheres). Other nucleotides that differed in reactivity in the presence and absence of rifampicin are located at a compact interface between the 5′ and 3′ domains.

Almost half of the nucleotides protected upon addition of rifampicin are in direct contact with ribosomal proteins, and most of these primarily reflect RNA−protein interactions that occur in the 3′ major domain (Figure 5A, yellow spheres). Thus, association of protein with the 3′ domain appears to be a major rate-determining step for 30S subunit assembly in cells. We identified the small subunit proteins that are in direct contact with the 14 helices that undergo large conformational changes upon addition of rifampicin (Figures 2 and 3). Eight of the 20 small subunit proteins interact with the 14 helices (Figure 5C, circles). Strikingly, six of the eight proteins are those previously identified as tertiary-binding proteins[31] (Figure 5C, emphasized with yellow circles). In addition, S9, although not formally characterized as a tertiary-binding protein, participates in a network involving four of the tertiary-binding proteins (Figure 5C, also emphasized with a yellow circle). The one additional protein−RNA interaction, involving S4 and the H16/17 RNA helix, undergoes a late conformational change distinct from the multiple changes involving the tertiary-binding proteins (Figure 5C, green circle). Overall, the in-cell data imply that binding by the tertiary proteins is accompanied by widespread and previously unobserved structural changes in the 16S rRNA.

## DISCUSSION

**SHAPE Reveals Intermediates in in-Cell Ribosome Assembly.** The 1M7 SHAPE reagent readily traverses the complex cell wall of the *E. coli* bacterium to modify RNA in a structure selective manner. We used this reagent to obtain snapshots of the structure of 16S rRNA in bacteria at the midlog phase of growth, when much of the rRNA has clearly not been fully assembled into ribosomes, and in the presence of a transcription inhibitor when most of the 16S rRNA has been incorporated into mature subunits or ribosomes. 1M7 elegantly balances the needs of experimental ease of use with high accuracy in probing nucleotide-resolution RNA dynamics.[1,19,21]

All experiments probing in-cell states likely report on complex ensembles of steady-state intermediates, including diverse RNA folding states and RNA−protein complexes composed of heterogeneous complements of proteins. We found that 17% of 16S rRNA precursor transcripts have an incompletely processed 5′ end, and thus, the "snapshot" taken using SHAPE in rapidly growing cells reflects 16S rRNA that is in various stages of assembly (Figure 2). Our direct measurement of unprocessed


dx.doi.org/10.1021/bi500198b | *Biochemistry* 2014, 53, 3237−3247

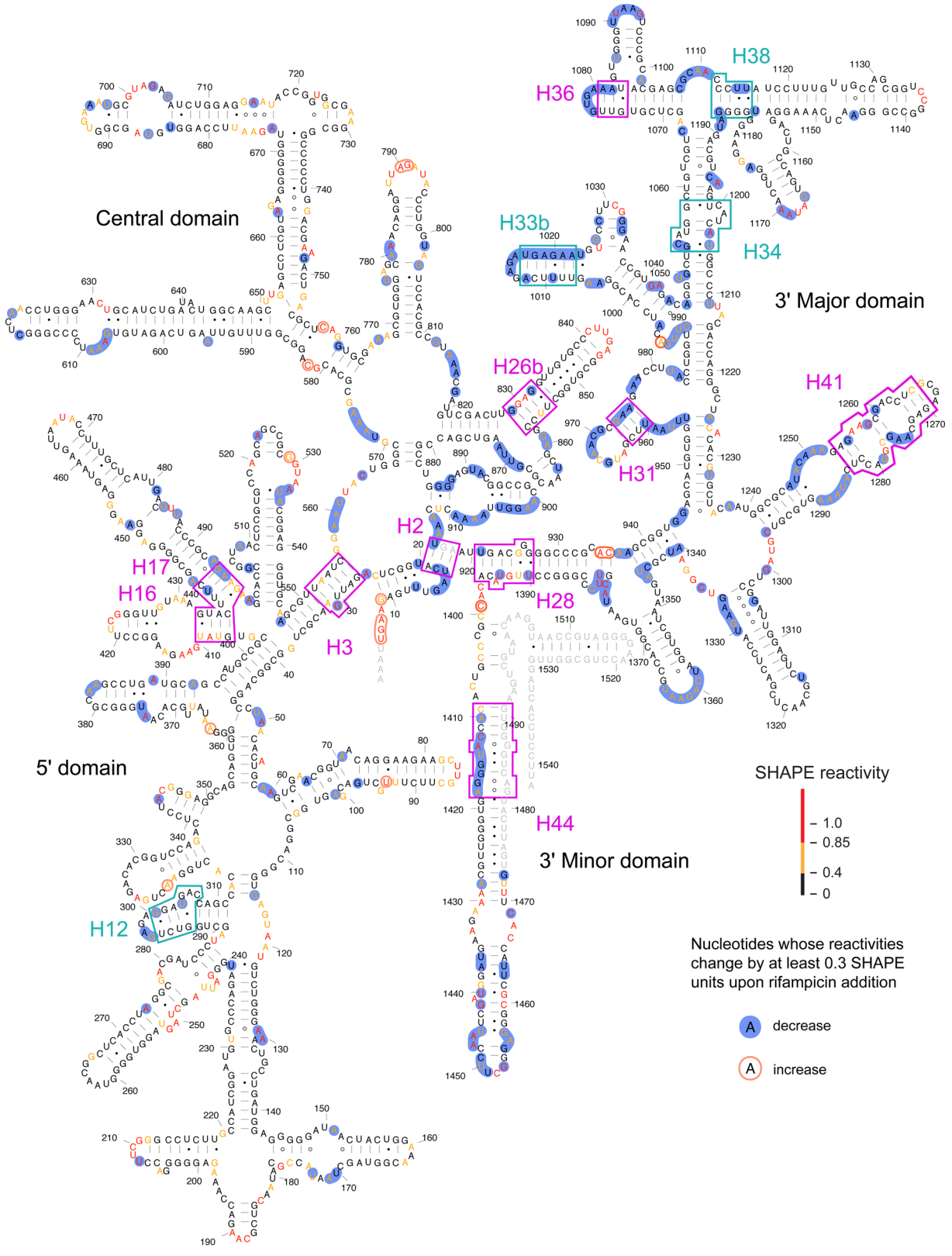



**Figure 3.** *In vivo* SHAPE reactivities for the 16S rRNA after transcription inhibition. Transcription was halted by treatment with rifampicin.[28] Nucleotides with significant decreases or increases in SHAPE reactivity (≥0.3) due to transcription arrest are highlighted in blue and orange, respectively. Boxed helices are those shown in Figure 2.

RNA indicates a much larger fraction of 30S subunits are incompletely assembled in rapidly growing cells than is generally assumed.[13,32] Our results, although surprising, are consistent with recent studies showing high levels of ribosomal particles with incomplete complements of ribosomal proteins[3] and with studies indicating that RNA is generally less structured in dividing cells.[33]

Critically, this work reveals a broad mechanism underlying the observation of increased RNA dynamics in cells. Upon inhibition


dx.doi.org/10.1021/bi500198b | *Biochemistry* 2014, 53, 3237−3247

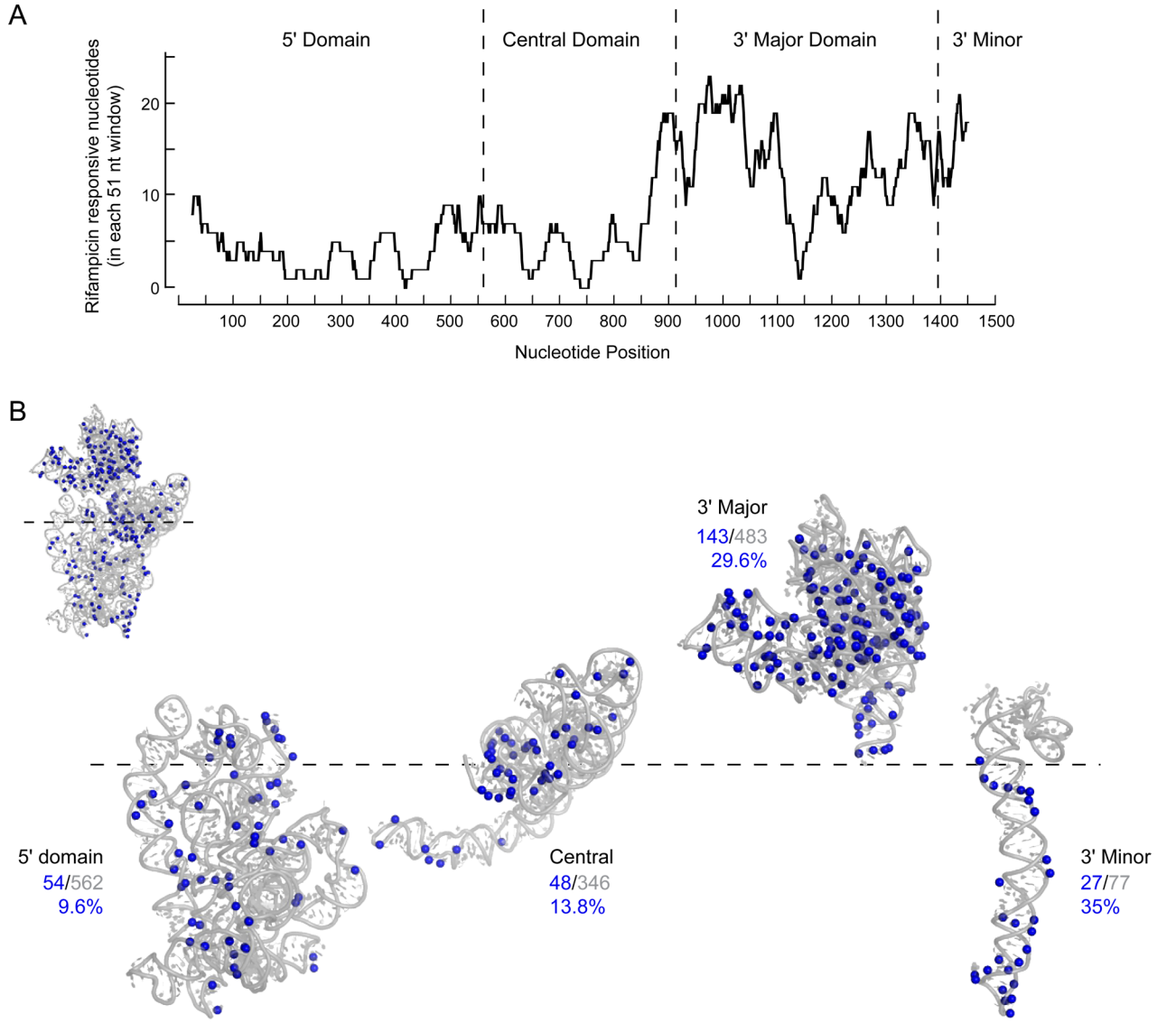



**Figure 4.** Transcription-dependent differences in SHAPE reactivity in the 16S rRNA. (A) Positions with significant differences in SHAPE reactivity in the absence and presence of rifampicin. The numbers of nucleotides with significant differences in reactivity when the two conditions were compared were summed in 51-nucleotide windows. Domain junctions are emphasized with vertical dashed lines. (B) All nucleotides with significant differences in SHAPE reactivity in the absence and presence of rifampicin visualized on the structure of the 16S rRNA (PDB entry 3i1m).[25] Ribosomal domains are expanded along a horizontal axis corresponding to the 5′ to 3′ polarity of their synthesis during transcription (dashed line). The inset shows the complete rRNA. The number and percentage of nucleotides in each 16S rRNA domain whose SHAPE reactivities changed relative to the number of nucleotides in the domain are given.

of transcription using rifampicin, which results in conditions where the vast majority of 16S rRNA is incorporated into 30S subunits, SHAPE reactivities were ∼99% consistent with the accepted secondary[26] and tertiary[25] structure (Figure 3). Thus, many differences observed in growing cells reflect an ongoing ribonucleoprotein assembly process.

There is a strong 5′ to 3′ polarity in the extent to which nucleotide reactivities changed in the absence compared to the presence of rifampicin (Figure 4). This is consistent with cotranscriptional folding and assembly into 30S subunits. In addition, domain junctions were punctuated by significant differences in nucleotide reactivity in the presence and absence of rifampicin, consistent with formation of local interdomain interactions. In-cell SHAPE, coupled with chasing complex steady-state RNA and ribonucleoprotein ensembles into more structurally mature forms by transcription inhibition, is a broadly useful approach for defining RNA−protein assembly intermediates and interdomain junctions in ribonucleoprotein complexes.

**Domain-Specific Assembly Features.** For the 16S rRNA studied here, the differences in nucleotide reactivities in the presence and absence of rifampicin reflect distinct underlying physical processes in each ribosomal domain. In the 5′ domain, all nucleotides whose reactivities were responsive to rifampicin showed decreases in reactivity. Moreover, more than 80% (44 of 54) of these nucleotides are located in RNA elements that form intradomain interactions (Figures 4B and 5A, 5′ domain panels). Thus, the 5′ domain undergoes internal tightening and consolidation in structure in mature ribosomes compared to the structure of less mature forms. In contrast, for the central domain, roughly 65% (31 of 48) of the nucleotides that underwent decreases in reactivity are involved in interdomain or RNA−protein interactions, with the majority reflecting


dx.doi.org/10.1021/bi500198b | *Biochemistry* 2014, 53, 3237−3247

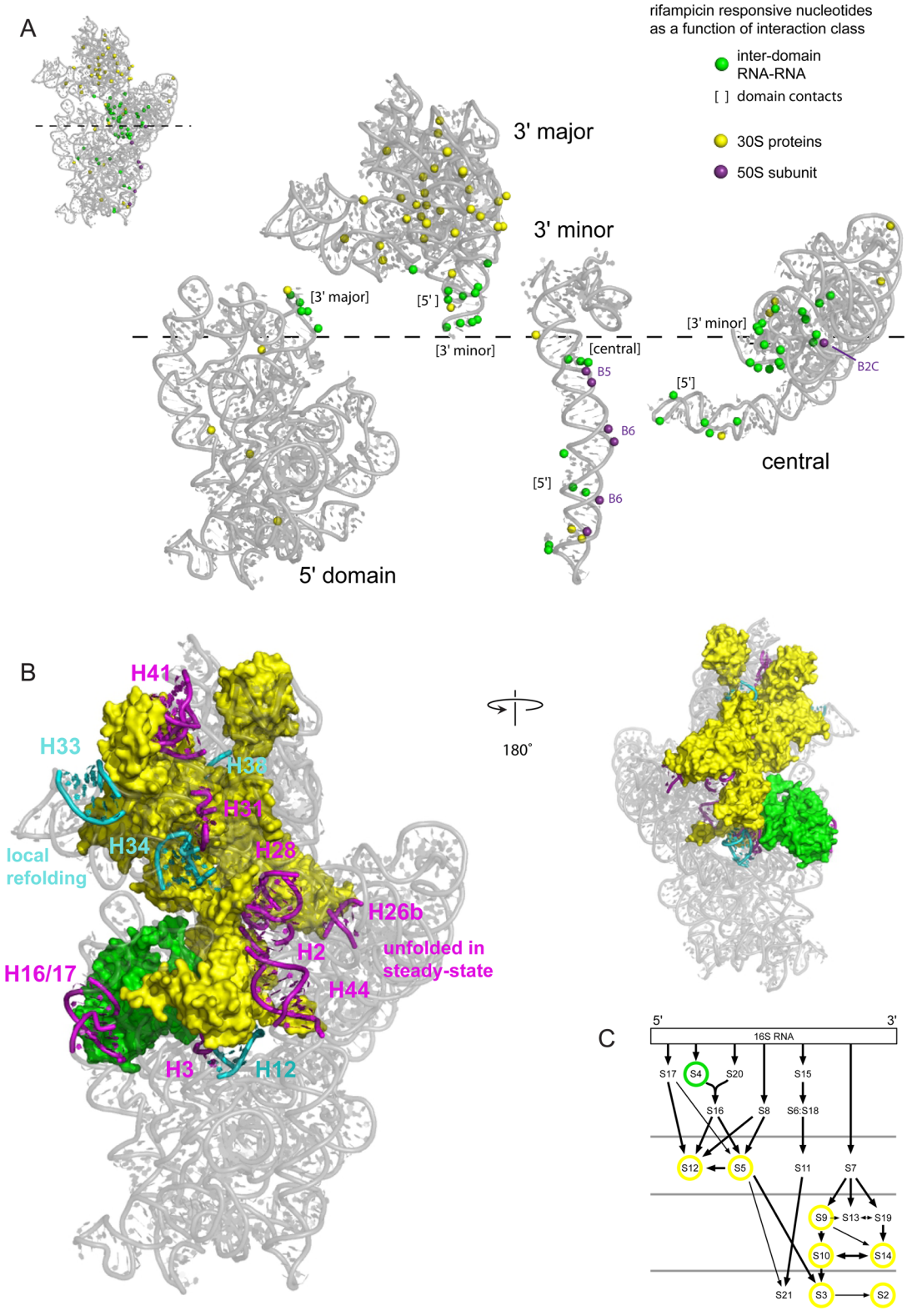



**Figure 5.** RNA–RNA, RNA–protein, and RNA–domain interactions detected by comparison of SHAPE reactivities in the presence and absence of transcription. (A) The 2′-OH groups that show significant changes in SHAPE reactivity upon transcription inhibition and are located at the surface of each rRNA domain are highlighted with spheres. Interacting groups, as inferred from the high-resolution structure, are emphasized by the coloring scheme shown at the right. Domains are expanded in the same left-to-right orientation as in the standard view of the intact 16S rRNA (see inset). Interactions with bridges to the 50S subunit are labeled. (B) Relationship between 16S rRNA helices that do not form stably in rapidly growing cells in the presence of transcription (magenta and cyan, as also shown in Figure 2) juxtaposed with proteins in direct contact with these helices. (C) Positions of ribosomal proteins in direct contact with helices that undergo changes in structure during assembly in a ribosome protein assembly map. Yellow circles emphasize tertiary-binding proteins.[31] The four rows of proteins indicate early to late assembly intermediates, as measured by the size of 30S-containing complexes *in vivo*.[3]


dx.doi.org/10.1021/bi500198b | *Biochemistry* 2014, 53, 3237–3247

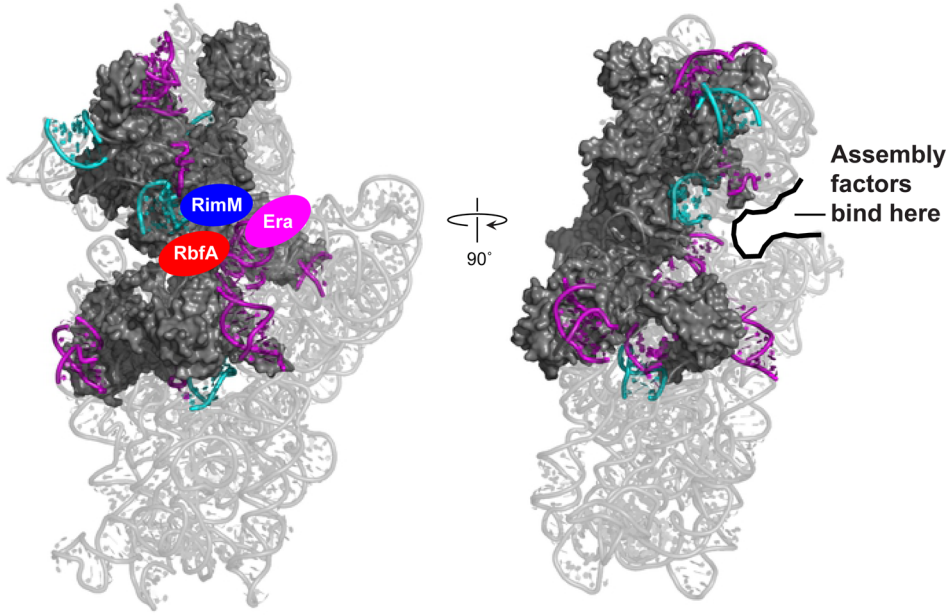



**Figure 6.** Positions of misfolded helices and late-binding proteins relative to late-stage ribosome assembly factors. Positions of RbfA, RimM, and Era are as defined by electron microscopy.[36,41,42] Misfolded helices are colored cyan and purple (see Figure 2), and the proteins that directly contact these RNA elements are shown as gray surfaces.

RNA–RNA interactions (compare Figures 4B and 5A). One nucleotide reflects a direct interaction with the 50S subunit and forms part of the B2C intersubunit bridge (Figure 5A, purple sphere).[34] The pattern is again distinctive in the 3′ major and minor domains. Of the 47 nucleotides that undergo decreases in reactivity in the presence of rifampicin and likely report RNA–protein interactions, 66% (31 of 47) are located in the 3′ major subunit (Figure 5A, yellow spheres). In addition, six nucleotides that underwent decreases in reactivity in the 3′ minor domain reflect direct interactions with the 50S subunit (Figure 5A, purple spheres). Of these six nucleotides, five form the B5 and B6 intersubunit bridges.[34]

In summary, the numbers of rRNA nucleotides with differential reactivity by SHAPE in the absence and presence of rifampicin increase in the 5′ to 3′ direction. These changes show a striking structural polarity reflecting the assembly of the ribosome and begin predominantly with intradomain RNA–RNA interactions (5′ domain), followed by interdomain RNA–RNA interactions (central domain), then intrasubunit (30S) RNA–protein interactions (3′ major domain), and finally intersubunit interactions (3′ minor domain).

**The Predominant 30S State in Growing Bacteria Is a Late Assembly Intermediate.** This work and other recent work[1] emphasize that the cellular environment has large and poorly understood effects on RNA structure. For completely synthesized transcripts, the highly crowded environment in an *E. coli* cell appears to stabilize and favor more organized higher-order RNA structure.[1] This effect undoubtedly influences bacterial ribosome structure, as well. Here we observed an additional effect of the cellular environment, the effect of ongoing cotranscriptional assembly and structural biogenesis. During the snapshot taken during our in-cell probing experiment, the steady-state structure of 16S rRNA in healthy growing *E. coli* cells is heavily influenced by 16S rRNA that has not yet been fully assembled into mature 30S subunits.

Fourteen helical structures do not form stably in the intermediate state that predominates in rapidly growing cells (Figure 2). Current evidence strongly suggests that folding of these helices coincides with a critical, late-stage, rate-determining step in ribosome assembly. First, eight of these helices (H2, H3, H12, H28, H34, H36, H38, and H44) are also unfolded when ribosomes are deproteinized.[18] These unfolded structures thus comprise thermodynamically preferred RNA states. Second, chemical modification of seven of these same helices (H2, H12, H28, H31, H34, H36, and H44) interferes with ribosome assembly.[35] Third, these helices are similar to those that are improperly folded or interact with protein partners that are missing in immature 30S subunits created by deleting the ribosome late assembly factors RbfA and RimM.[4,36] Finally, six of the eight proteins that bind directly to the unfolded helices were defined as late-binding tertiary proteins in the original Nomura maps[31] and have recently been shown to bind during late stages of ribosome assembly *in vivo*[3] (Figure 5B,C, in yellow). Thus, although there are likely to be multiple species present, a significant fraction of the still-maturing interactions detected in the 2 min SHAPE snapshot correspond to events that occur at late stages in 30S subunit assembly.

**A Molecular Strut Model for Ribosomal Assembly Factors.** The 14 incompletely folded helices, seven late assembly proteins, and S4 form a C-shaped motif spanning the 5′ and 3′ domains of the 16S rRNA (Figure 6). This region surrounds the mRNA-binding channel, is important for tRNA binding and codon–anticodon decoding, and must be properly folded and arranged for ribosome function.[37−39] Three factors, RbfA, RimM, and Era, have been shown to facilitate late-stage 30S subunit assembly. These three assembly factors bind at well-defined but distinct and largely nonoverlapping sites and can, in part, suppress deletion of or compensate for each other.[6,40] These proteins all bind in the concave face of the C-shaped structure formed by the helices and proteins that we have shown rearrange in late stages of ribosome assembly (Figure 6). We interpret the partially overlapping functions of these assembly factors in terms of a molecular strut model. Late-stage assembly involves large-scale RNA structural rearrangements (Figure 3)


dx.doi.org/10.1021/bi500198b | *Biochemistry* 2014, 53, 3237−3247

and protein assembly events (Figure 5) that we hypothesize require, or are facilitated by, partial preorganization of the C-shaped interaction network. In this model, assembly factors function in part to promote organization the C-shaped mRNA channel area by binding in the cleft. An individual factor could therefore compensate for contributions by another assembly factor, as long as this structural organization function were provided.

## SUMMARY

In-cell SHAPE probing of growing *E. coli* cells showed that ribosomal assembly intermediates accumulate and may even predominate in rapidly growing cells. The SHAPE-detected transcriptional pulse−chase approach outlined in this work is both straightforward and information-rich and will likely find broad utility in understanding the steady-state RNA dynamics of other RNAs and RNA−protein complexes, and in other cell types.

## AUTHOR INFORMATION

### Corresponding Author

*E-mail: weeks@unc.edu.

### Funding

This work was supported by National Science Foundation Grant MCB-1121024 to K.M.W.

### Notes

The authors declare no competing financial interest.

## REFERENCES

(1) Tyrrell, J., McGinnis, J. L., Weeks, K. M., and Pielak, G. J. (2013) The cellular environment stabilizes adenine riboswitch RNA structure. *Biochemistry 52*, 8777−8785.

(2) Swiatkowska, A., Wlotzka, W., Tuck, A., Barrass, J. D., Beggs, J. D., and Tollervey, D. (2012) Kinetic analysis of pre-ribosome structure in vivo. *RNA 18*, 2187−2200.

(3) Chen, S. S., and Williamson, J. R. (2013) Characterization of the ribosome biogenesis landscape in *E. coli* using quantitative mass spectrometry. *J. Mol. Biol. 425*, 767−779.

(4) Clatterbuck Soper, S. F., Dator, R. P., Limbach, P. A., and Woodson, S. A. (2013) In Vivo X-ray Footprinting of Pre-30S Ribosomes Reveals Chaperone-Dependent Remodeling of Late Assembly Intermediates. *Mol. Cell 52*, 506−516.

(5) Gourse, R. L., Gaal, T., Bartlett, M. S., Appleman, J. A., and Ross, W. (1996) rRNA transcription and growth rate-dependent regulation of ribosome synthesis in *Escherichia coli*. *Annu. Rev. Microbiol. 50*, 645−677.

(6) Shajani, Z., Sykes, M. T., and Williamson, J. R. (2011) Assembly of bacterial ribosomes. *Annu. Rev. Biochem. 80*, 501−526.

(7) Kaczanowska, M., and Ryden-Aulin, M. (2007) Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiol. Mol. Biol. Rev. 71*, 477−494.

(8) Nierhaus, K. H. (1980) The assembly of the prokaryotic ribosome. *BioSystems 12*, 273−282.

(9) Granneman, S., and Baserga, S. J. (2004) Ribosome biogenesis: Of knobs and RNA processing. *Exp. Cell Res. 296*, 43−50.

(10) Calidas, D., and Culver, G. M. (2011) Interdependencies govern multidomain architecture in ribosomal small subunit assembly. *RNA 17*, 263−277.

(11) Woodson, S. A. (2011) RNA folding pathways and the self-assembly of ribosomes. *Acc. Chem. Res. 44*, 1312−1319.

(12) 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.

(13) Lindahl, L. (1975) Intermediates and time kinetics of the in vivo assembly of *Escherichia coli* ribosomes. *J. Mol. Biol. 92*, 15−37.

(14) Blaha, G., Stelzl, U., Spahn, C. M., Agrawal, R. K., Frank, J., and Nierhaus, K. H. (2000) Preparation of functional ribosomal complexes and effect of buffer conditions on tRNA positions observed by cryoelectron microscopy. *Methods Enzymol. 317*, 292−309.

(15) Merino, E. J., Wilkinson, K. A., Coughlan, J. L., and Weeks, K. M. (2005) RNA structure analysis at single nucleotide resolution by selective 2′-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc. 127*, 4223−4231.

(16) McGinnis, J. L., Dunkle, J. A., Cate, J. H., and Weeks, K. M. (2012) The mechanisms of RNA SHAPE chemistry. *J. Am. Chem. Soc. 134*, 6617−6624.

(17) Gherghe, C. M., Shajani, Z., Wilkinson, K. A., Varani, G., and Weeks, K. M. (2008) Strong correlation between SHAPE chemistry and the generalized NMR order parameter ($S^2$) in RNA. *J. Am. Chem. Soc. 130*, 12244−12245.

(18) Deigan, K. E., Li, T. W., Mathews, D. H., and Weeks, K. M. (2009) Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci. U.S.A. 106*, 97−102.

(19) Hajdin, C. E., Bellaousov, S., Huggins, W., Leonard, C. W., Mathews, D. H., and Weeks, K. M. (2013) Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots. *Proc. Natl. Acad. Sci. U.S.A. 110*, 5498−5503.

(20) Wilkinson, K. A., Gorelick, R. J., Vasa, S. M., Guex, N., Rein, A., Mathews, D. H., Giddings, M. C., and Weeks, K. M. (2008) High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biol. 6*, e96.

(21) Mortimer, S. A., and Weeks, K. M. (2007) A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *J. Am. Chem. Soc. 129*, 4144−4145.

(22) Steen, K. A., Rice, G. M., and Weeks, K. M. (2012) Fingerprinting noncanonical and tertiary RNA structures by differential SHAPE reactivity. *J. Am. Chem. Soc. 134*, 13160−13163.

(23) Karabiber, F., McGinnis, J. L., Favorov, O. V., and Weeks, K. M. (2013) QuShape: Rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. *RNA 19*, 63−73.

(24) Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M., and Turner, D. H. (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. U.S.A. 101*, 7287−7292.

(25) Zhang, W., Dunkle, J. A., and Cate, J. H. (2009) Structures of the ribosome in intermediate states of ratcheting. *Science 325*, 1014−1017.

(26) Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Muller, K. M., Pande, N., Shang, Z., Yu, N., and Gutell, R. R. (2002) The comparative RNA web (CRW) site: An online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinf. 3*, 2.

(27) Gutell, R. R., Lee, J. C., and Cannone, J. J. (2002) The accuracy of ribosomal RNA comparative structure models. *Curr. Opin. Struct. Biol. 12*, 301−310.

(28) Hartmann, G., Honikel, K. O., Knusel, F., and Nuesch, J. (1967) The specific inhibition of the DNA-directed RNA synthesis by rifamycin. *Biochim. Biophys. Acta 145*, 843−844.

(29) Gotta, S. L., Miller, O. L., Jr., and French, S. L. (1991) rRNA transcription rate in *Escherichia coli*. *J. Bacteriol. 173*, 6647−6649.

(30) Roy-Chaudhuri, B., Kirthi, N., and Culver, G. M. (2010) Appropriate maturation and folding of 16S rRNA during 30S subunit biogenesis are critical for translational fidelity. *Proc. Natl. Acad. Sci. U.S.A. 107*, 4567−4572.

(31) 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. U.S.A. 59*, 777−784.

(32) Laughrea, M., and Tam, J. (1992) In vivo chemical footprinting of the *Escherichia coli* ribosome. *Biochemistry 31*, 12035−12041.

(33) Rouskin, S., Zubradt, M., Washietl, S., Kellis, M., and Weissman, J. S. (2014) Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature 505*, 701−705.
(34) Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N., and Noller, H. F. (1999) X-ray crystal structures of 70S ribosome functional complexes. *Science 285*, 2095−2104.
(35) Xu, Z., and Culver, G. M. (2010) Differential assembly of 16S rRNA domains during 30S subunit formation. *RNA 16*, 1990−2001.
(36) Guo, Q., Goto, S., Chen, Y., Feng, B., Xu, Y., Muto, A., Himeno, H., Deng, H., Lei, J., and Gao, N. (2013) 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*, 2609−2620.
(37) Guo, Z., and Noller, H. F. (2012) Rotation of the head of the 30S ribosomal subunit during mRNA translocation. *Proc. Natl. Acad. Sci. U.S.A. 109*, 20391−20394.
(38) Noeske, J., and Cate, J. H. (2012) Structural basis for protein synthesis: Snapshots of the ribosome in motion. *Curr. Opin. Struct. Biol. 22*, 743−749.
(39) 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.
(40) Donhofer, A., Sharma, M. R., Datta, P. P., Nierhaus, K. H., Agrawal, R. K., and Wilson, D. N. (2009) Factor-Mediated Ribosome Assembly in Bacteria. In *Encyclopedia of Life Sciences*, John Wiley & Sons, Ltd., Chichester, U.K.
(41) Sharma, M., Barat, C., Wilson, D., Booth, T., Kawazoe, M., Hori-Takemoto, C., Shirouzu, M., Yokoyama, S., Fucini, P., and Agrawal, R. (2005) Interaction of Era with the 30S ribosomal subunit implications for 30S subunit assembly. *Mol. Cell 18*, 319−329.
(42) Datta, P., Wilson, D., Kawazoe, M., Swami, N., Kaminishi, T., Sharma, M., Booth, T., Takemoto, C., Fucini, P., Yokoyama, S., and Agrawal, R. (2007) Structural aspects of RbfA action during small ribosomal subunit assembly. *Mol. Cell 28*, 434−445.