



Shape changes and cooperativity in the folding of the central domain of the 16S ribosomal RNA

Naoto Hori (堀直人)^{a,b}, Natalia A. Denesyuk^a, and D. Thirumalai^{a,1}

^aDepartment of Chemistry, University of Texas at Austin, Austin, TX 78712; and ^bSchool of Pharmacy, University of Nottingham, Nottingham NG7 2RD, United Kingdom

Edited by B. Montgomery Pettitt, University of Texas Medical Branch, Galveston, TX, and accepted by Editorial Board Member J. A. McCammon December 18, 2020 (received for review October 5, 2020)

Both the small and large subunits of the ribosome, the molecular machine that synthesizes proteins, are complexes of ribosomal RNAs (rRNAs) and a number of proteins. In bacteria, the small subunit has a single 16S rRNA whose folding is the first step in its assembly. The central domain of the 16S rRNA folds independently, driven either by Mg²⁺ ions or by interaction with ribosomal proteins. To provide a quantitative description of ion-induced folding of the ~350-nucleotide rRNA, we carried out extensive coarse-grained molecular simulations spanning Mg²⁺ concentration between 0 and 30 mM. The Mg²⁺ dependence of the radius of gyration shows that globally the rRNA folds cooperatively. Surprisingly, various structural elements order at different Mg²⁺ concentrations, indicative of the heterogeneous assembly even within a single domain of the rRNA. Binding of Mg²⁺ ions is highly specific, with successive ion condensation resulting in nucleation of tertiary structures. We also predict the Mg²⁺-dependent protection factors, measurable in hydroxyl radical footprinting experiments, which corroborate the specificity of Mg²⁺-induced folding. The simulations, which agree quantitatively with several experiments on the folding of a three-way junction, show that its folding is preceded by formation of other tertiary contacts in the central junction. Our work provides a starting point in simulating the early events in the assembly of the small subunit of the ribosome.

RNA folding | divalent ions | ribosome assembly | coarse-grained simulation | three-way junction

The determination of the spectacular ribosome structures has galvanized great interest in dissecting how such a complex structure assembles *in vivo* (1–5). The bacterial ribosome is a complex between the small 30S and the large 50S particles. These subunits themselves are complexes involving the ribosomal RNA (rRNA) and a number of proteins (6, 7). Since ribosomes synthesize proteins in all living organisms, considerable energy and other regulatory mechanisms are used to generate and maintain their homeostasis. Although there have been vast experimental efforts to understand the molecular mechanism of ribosome assembly, which led to important early discoveries, such as Nomura and Nierhaus maps (8, 9), the general principles associated with the assembly process of the ribosome have not been fully resolved (3, 4). To solve the assembly problem, many elegant experiments have been systematically performed initially by investigating how the various domains of the rRNA fold and subsequently by the effect of ribosomal proteins in reshaping the assembly landscape. Along the way it has been pointed out (10) that some of the principles of ribozyme folding might form a useful framework for producing a quantitative theoretical model for ribosome assembly. After all, the rRNA folding problem has to be solved to produce an intact ribosome capable of protein synthesis.

In bacteria, the three rRNA chains (~4,500 nucleotides in total) fold and assemble together with over 50 ribosomal proteins (r-proteins) to build one functional ribosome. The small subunit, the 30S particle, is a large ribonucleoprotein complex consisting of a single 16S rRNA chain (approximately 1,500 nucleotides) and about 20 proteins. The rRNA chain could

be further decomposed into the 5' (~560 nucleotides), central (~350 nucleotides), and 3' (~625 nucleotides) domains. It is known that the three domains fold independently and concurrently in the presence of r-proteins (2, 11–13). Therefore, a logical approach is to understand how the individual components, especially the various rRNA domains, might fold independent of each other.

Previous simulation studies, focusing on the nuances of the protein-induced structural transitions in the 5' domain (14–16), have been most insightful, especially when combined with experiments (16). These studies showed that S4-guided assembly results in the 5' domain reaching the folded state by navigating through multiple metastable states, revealing the rugged nature of the folding landscape of RNA (17, 18). As a prelude toward undertaking computational studies in the early events in the assembly of the 30S particle, here we report our investigations of the Mg²⁺-induced folding of the central domain of the 16S rRNA using an accurate simulation model.

The importance of divalent cations, Mg²⁺ in particular, in stabilizing tertiary structures of RNA cannot be understated. In the context of ribosome assembly, the importance of Mg²⁺ has been recognized from a variety of experiments (19, 20). For instance, it has been recently shown that the entire 23S rRNA from the large subunit of the bacterial ribosome forms a near-native conformation in the presence of Mg²⁺ without any r-proteins (21). In the central domain of the small subunit, *in vitro* studies have shown that a three-way junction (3WJ) (Fig. 1) is in dynamic equilibrium between an open and a closed conformation, whose

Significance

Ribosomes are complexes between ribosomal RNA (rRNA) and a number of proteins. Because ribosome assembly begins with rRNA folding, we simulated the molecular details of Mg²⁺-driven folding of the central domain of the bacterial rRNA. Good agreement with experiments on the folding of the three-way junction in the center of the rRNA validates the model. Coupling of rRNA folding and Mg²⁺ binding shows that ions interact with rRNA segments in a coordinated manner. The shape of rRNA changes from a sphere in the unfolded state to a prolate ellipsoid at high Mg²⁺ concentration, which is the opposite of what transpires when a globular protein folds. Our study provides the needed framework for undertaking ion-driven folding of large RNA molecules.

Author contributions: N.H., N.A.D., and D.T. designed research; N.H. and N.A.D. performed research; N.H. and N.A.D. contributed new reagents/analytic tools; N.H. and D.T. analyzed data; and N.H. and D.T. wrote the paper.

The authors declare no competing interest.

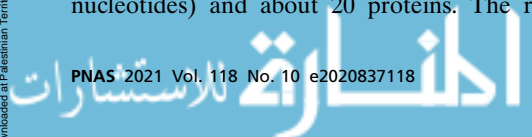
Published under the PNAS license.

This article is a PNAS Direct Submission. B.M.P. is a guest editor invited by the Editorial Board.

¹ To whom correspondence may be addressed. Email: dave.thirumalai@gmail.com.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2020837118/-/DCSupplemental>.

Published March 3, 2021.



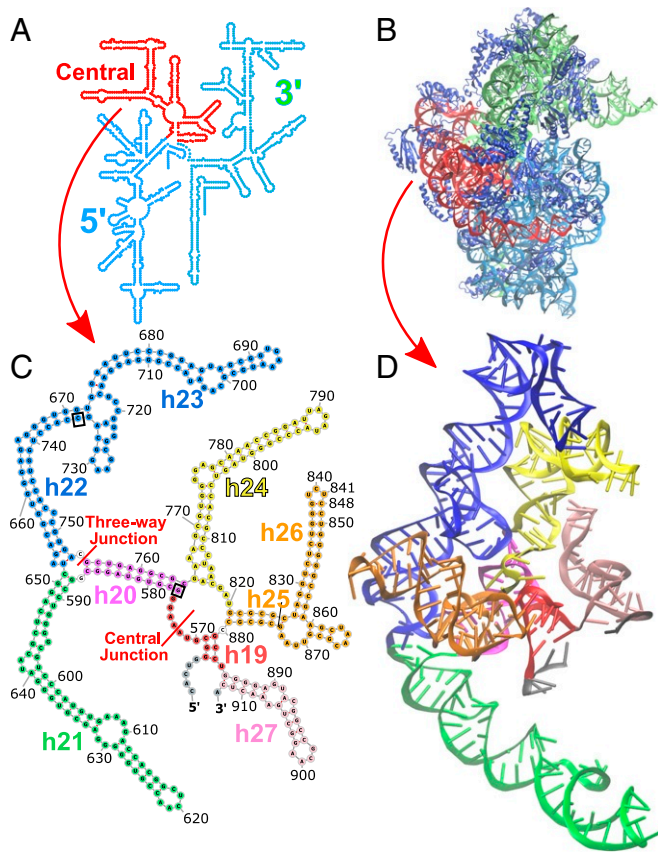


Fig. 1. Sequence and structure of the ribosomal RNA. (A and B) The entire small subunit of *T. thermophilus* ribosome is illustrated as the secondary structure of the 16S rRNA (A) and the crystal structure including ribosomal proteins (B). The central domain simulated is adopted from ref. 30 and the tertiary structure was taken from PDB entry 1J5E (31). (C and D) Secondary and tertiary structures of the 346-nucleotide fragment central domain of the 16S ribosomal RNA.

populations depends on the Mg^{2+} concentration. Experiments have shown that the 3WJ undergoes substantial rearrangements upon addition of Mg^{2+} or S15 r-protein (22–24).

To provide molecular insights into the folding of the central domain of the small subunit of bacterial ribosome, driven by Mg^{2+} , we performed extensive simulations based on an accurate coarse-grained model for RNA (25, 26). We chose the 346-nt RNA fragment as our starting point because it has been used in experimental studies to probe the early stage of RNA assembly involving r-proteins (23, 24, 27–29). In this paper, we focus on the thermodynamics of Mg^{2+} -dependent folding of the isolated RNA. Our simulations reproduced the folding transition of the central domain, from its unfolded state containing only secondary structures to the near-native tertiary structure, as the Mg^{2+} concentration increased. Changes in angle of the three-way junction obtained using simulations accord well with experiments, thus validating the proposed model. Although the folding transition is cooperative, we find that some groups of tertiary contacts form at different Mg^{2+} concentrations. Thus, the components of the domain do not fold simultaneously at a unique midpoint. Importantly, the groups of tertiary contacts are stabilized by specific Mg^{2+} binding at their constituent nucleotides. From analyses of free-energy stabilization by Mg^{2+} binding at a single-nucleotide resolution, we found that the order of folding is dictated by the extent of stabilization by each Mg^{2+} binding.

In the case of the central domain, tertiary folding occurs preferentially around the central junction. We make predictions for Mg^{2+} -dependent nucleotides protection, which are testable by hydroxyl radical footprinting experiments.

Results

Mg^{2+} -Induced Compaction of rRNA. We first focus on global properties that are measurable using small-angle X-ray scattering (SAXS) experiments. From equilibrium simulations of the rRNA fragment at various Mg^{2+} ion concentrations, we calculated the average radius of gyration (R_g) as a function of $[Mg^{2+}]$ (Fig. 2A). rRNA undergoes an apparent two-state folding transition, as indicated by the order parameter R_g , as Mg^{2+} concentration increases from 0 to 30 mM. The maximum decrease in R_g occurs in the range of $[Mg^{2+}] \approx 2$ to 10 mM. In the absence of or at $[Mg^{2+}]$ up to ~ 1 mM, tertiary interactions of the RNA are disrupted, resulting in expanded conformations containing only secondary structures (top left structure in Fig. 2A). After adding an excess amount of Mg^{2+} (>10 mM), rRNA folds to a compact conformation, with an average $R_g \approx 4.3$ nm, that is modestly larger than the R_g of the crystal structure, $R_g^{Native} = 4.0$ nm. This indicates that either or both of the ribosomal proteins and other domains of the rRNA (we simulated only a fragment of the 16S rRNA) may be needed to fully drive the conformation to that found in the crystal structure. Interestingly, Mg^{2+} -driven compaction of rRNA is highly cooperative. The fit of the titration curve to the Hill equation (*Materials and Methods*) yields a high n value ($=2.96$) with an apparent midpoint of $[Mg^{2+}] \approx 3.3$ mM.

To evaluate the similarity between the simulated structures and the crystal structure, we also computed the root-mean-square deviation (rmsd; Fig. 2A, *Inset*). The rmsd of the entire rRNA fragment is 1.5 nm at 30 mM Mg^{2+} , whereas rmsd computed excluding h21 and h27 is 0.7 nm. From these data, we conclude that fluctuations of h21 and h27 contribute to the slightly larger R_g found in the simulations. As illustrated in Fig. 2A, the structure is correctly folded at high $[Mg^{2+}]$.

Fig. 2B and C shows the distance distribution functions at several Mg^{2+} concentrations, which may be obtained as the inverse Fourier transform of the scattering function that is measurable by SAXS experiments, thus serving as a testable prediction. Comparison of the results in Fig. 2B and C shows that the enhanced fluctuations at higher $[Mg^{2+}]$ are due to h21 and h27, which do not (especially h27) engage in extensive tertiary interactions. In the intact 30S subunit, h21 and h27 interact with other ribosomal domains and r-proteins that we did not include in the simulations (*SI Appendix, Fig. S2*).

We also calculated the shape parameters, asphericity Δ ($0 \leq \Delta \leq 1$), and prolateness S ($-0.25 \leq S \leq 2$) (32, 33). Both Δ and S are unity for rods, whereas $\Delta = S = 0$ for spheres. In our previous work (33) we showed, by considering a large number of folded RNA chains, that the distribution of Δ and S is broad with the shapes being highly spherical adopting prolate ellipsoidal shapes. In accord with the expectations, we find that beyond the midpoint, the shape of rRNA does resemble a prolate ellipsoid. Surprisingly, at low Mg^{2+} , rRNA is predicted to be nearly spherical ($\Delta < 0.2$ and $S \approx 0$). Fig. 2C shows that there is a dramatic change in the shape of rRNA (roughly spherical to prolate ellipsoid) as the $[Mg^{2+}]$ concentration increased. This is exactly the opposite of what is typically found as a globular protein folds (34).

Helix h19 Forms Only at High Mg^{2+} . In the rRNA fragment, there are nine distinct helices that are conventionally named h19 through h27 (Fig. 1C). We calculated the probability of helix formation as a function of $[Mg^{2+}]$ (Fig. 3A). All of the helices

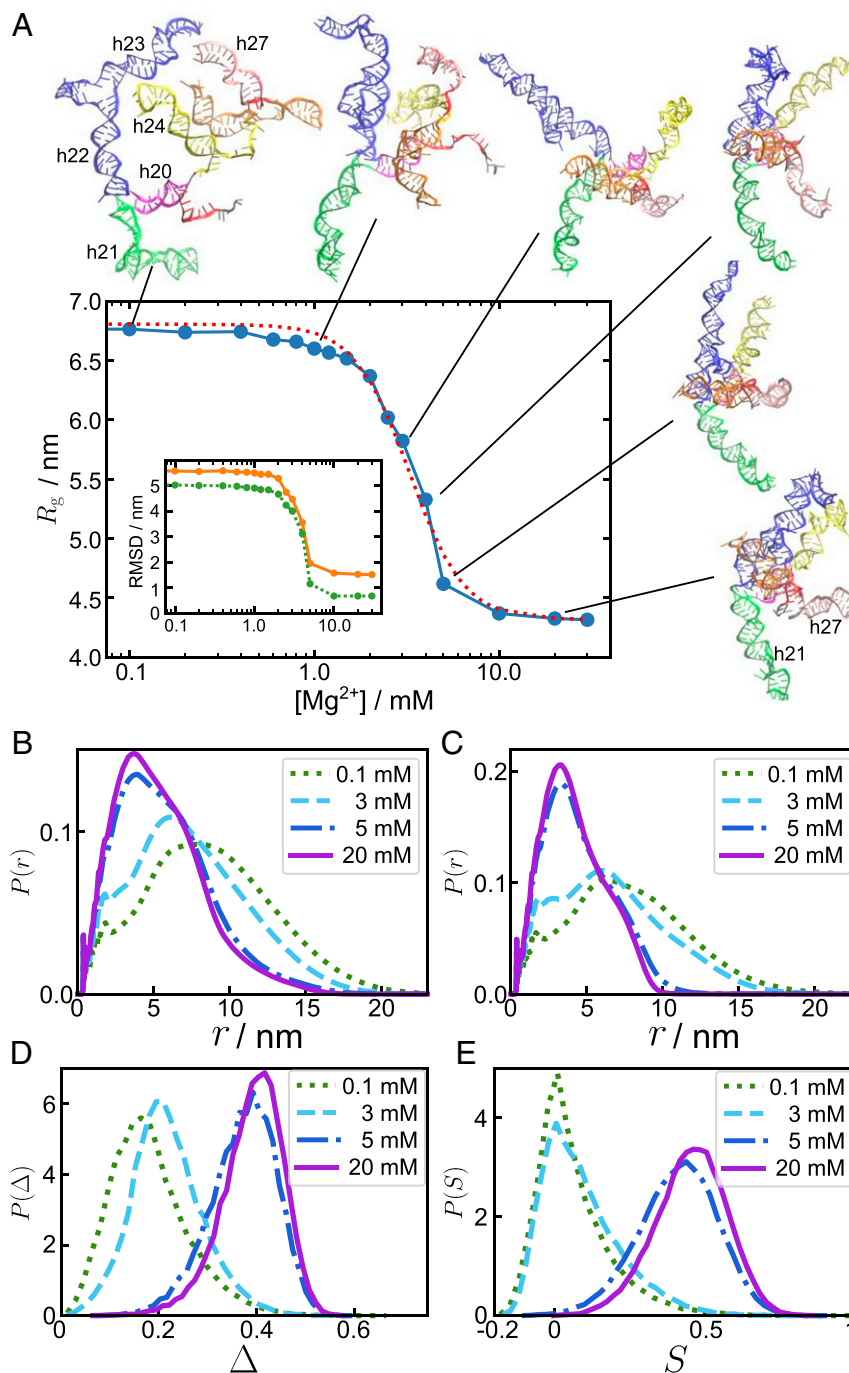


Fig. 2. Global structural transition in Mg²⁺-induced folding. (A) Radius of gyration (R_g) as a function of Mg²⁺ concentration. Representative conformations are shown for $[Mg^{2+}] = 0.1, 1.0, 3.0, 4.0, 5.0,$ and 20.0 mM (left to right; see also *SI Appendix, Fig. S1*). The red dotted line is the fit to the Hill equation with $n = 2.96$ and the midpoint $[Mg^{2+}]_m = 3.3$ mM. *Inset* shows the average rmsd from the crystal structure. The orange solid line shows rmsd for the entire RNA fragment, and the green dotted line is calculated excluding h21 and h27 (helices indicated in the bottom right structure). (B and C) Pair distance distributions at $[Mg^{2+}] = 0.1, 3, 5,$ and 20 mM (labeled in the panel) calculated using (B) the entire fragment and (C) the fragment excluding h21 and h27. (D and E) Probability distributions of the shape parameters, (D) asphericity Δ and (E) prolateness S , at the same concentrations as in B and C.

except the short h19 are stable over the entire range of Mg²⁺ concentrations in the presence of 50 mM KCl at the simulation temperature (37 °C). These helices, which are stable at very low [Mg²⁺], contain at least eight canonical Watson–Crick base pairs. In contrast, h19, connecting the central junction to h27 and the 5' domain, has only three G–C base pairs. Thus, it is likely that the stability of individual helices determines the order of their formation, as previously suggested (35). More impor-

tantly, as shown in Fig. 1C, the two constituent strands of h19 are far separated along the sequence. Consequently, the stability of h19 depends not only on [Mg²⁺] but also on the formation of other tertiary interactions around the central junction as well as the adjacent helix h25. Note that only after h25 fully forms (at $[Mg^{2+}] \approx 1$ mM), which brings the two strands of h19 into proximity (Fig. 1C), does h19 get structured at $[Mg^{2+}] \approx 2$ mM (Fig. 3A). The formation of tertiary interactions is predicated

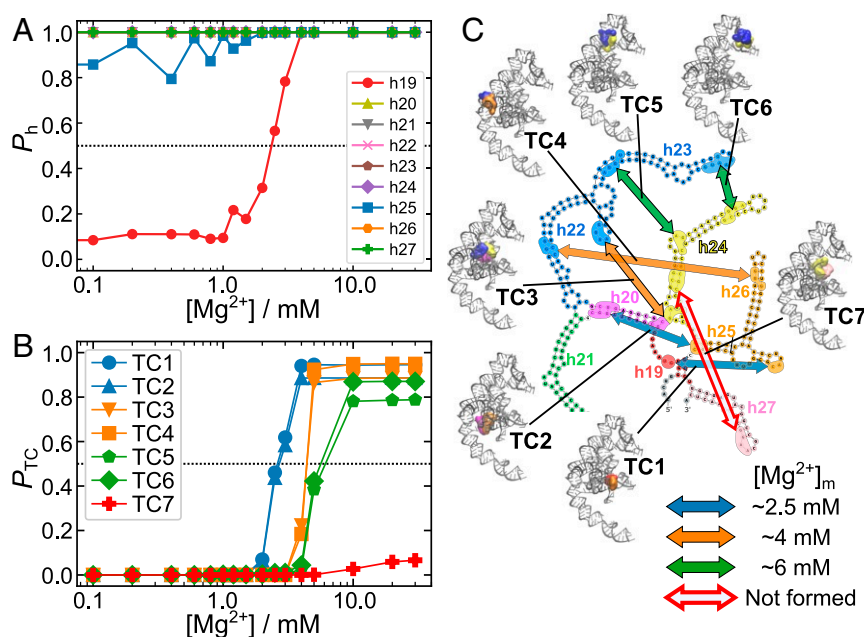


Fig. 3. Mg^{2+} -dependent folding of key structural elements. (A) Probabilities (P_h) of formation of the nine helices as a function of Mg^{2+} concentration. Besides h19 (red) and h25 (blue), all other lines overlap with h27 (green) because these stable helices remain folded even at extremely low Mg^{2+} concentrations. See Fig. 1 for positions of the helices. Both h19 and h25 are near the central junction (Fig. 1C). (B) Fraction of various tertiary contacts (TCs) as a function of Mg^{2+} concentrations. The $[Mg^{2+}]$ midpoints calculated using $P_{TC}([Mg^{2+}]_m) = 0.5$ (shown by the horizontal line) change substantially, depending on the TC. The color code reflects the midpoint of $[Mg^{2+}]$ for each TC as shown in C. (C) Locations of the seven key clusters associated with the TCs are mapped onto the secondary structure diagram. On the periphery, tertiary structures are shown using a surface representation for the corresponding regions.

on the formation of h19 (see below). The formation of h19 is essential for structuring the central junction (the big loop region surrounded by h19, h25, h24, and h20, shown in Fig. 1C). The results in Fig. 3A suggest that h19, whose $[Mg^{2+}]$ midpoint nearly coincides with the folding of the rRNA, must nucleate the formation of the tertiary interactions.

$[Mg^{2+}]$ Midpoints of Tertiary Contact Formation Vary. In the crystal structure of the rRNA fragment, there are seven regions with major tertiary contacts (TCs), where two or more secondary structural elements contact each other (Fig. 3C). These contacts are typically clusters of hydrogen-bonding interactions. We label them as in Fig. 3C, TC1 through TC7. Since we have already shown that all of the secondary structures except h19 are stable even in the near absence of Mg^{2+} , the formation of the TCs must drive the folding and consequently rRNA compaction (decrease in R_g) upon addition of Mg^{2+} ions.

In Fig. 3B, we show the formations of the seven TCs as a function of $[Mg^{2+}]$. There are three inferences that can be drawn from these results: 1) We first note that TC7, involving interactions between h24 and h27, does not form even at the highest $[Mg^{2+}]$. We surmise that the formation of TC7 may need r-proteins or other domains in the rRNA. Indeed, the crystal structure of the mature 30S (31) shows that helix h27 interacts with h44 of the 3' domain and S12 protein (SI Appendix, Fig. S2). 2) The other six TCs exhibit Mg^{2+} -dependent formation. Clearly, the $[Mg^{2+}]$ midpoints at which the TCs order vary substantially, perhaps reflecting the hierarchical structure formation. The fraction of TC formation sharply changes in the $[Mg^{2+}]$ range from 2 to 10 mM. Thus, the hierarchical assembly of the rRNA does not occur cooperatively at a sharp value of $[Mg^{2+}]$, implying that there is no precise midpoint for folding even though global measures indicate otherwise. 3) The range of Mg^{2+} concentrations over which the rRNA orders coincides with the decrease in R_g (Fig. 2). The results confirm that the R_g change, which depends on $[Mg^{2+}]$, is a consequence of the tertiary contact formation.

The six TCs can be classified into three pairs in terms of the Mg^{2+} concentration requirements. From the titration curve (Fig. 3B), the midpoints of $[Mg^{2+}]$ are ~ 2.5 mM for TC1 and TC2, ~ 4 mM for TC3 and TC4, and ~ 5 mM for TC5 and TC6. By mapping the TCs onto the secondary structure (Fig. 3C), we notice that the formations of the TCs occur first at locations near the central junction (TC1 and TC2). Upon further increase in $[Mg^{2+}]$, the two interactions, TC3 and TC4, separated by a moderate distance from the junction, are stabilized. Finally, a much higher concentration of Mg^{2+} (5 to 10 mM) is required to stabilize the long-range contacts between h23 and h24. From this analysis, we conclude that the network of TCs is centered around the central junction. This hierarchy of tertiary contact formation is also revealed in the representative snapshots from the simulations (Fig. 2). Helices h25 and h26 (orange) come close to h20 (magenta) in the early part of the transition at low $[Mg^{2+}]$, and the formation of contacts between h23 (blue) and h24 (yellow) occurs only at higher $[Mg^{2+}]$.

Mg^{2+} Binding to Specific Sites Drives Tertiary Contact Formation.

How do Mg^{2+} ions facilitate the folding of the RNA fragment, in particular, the formation of the TCs? To begin the investigation, we computed the fingerprints of Mg^{2+} binding at the nucleotide resolution as contact Mg^{2+} concentration, c^* (Materials and Methods). In Fig. 4A, the fingerprints are shown at three bulk Mg^{2+} concentrations at which the major conformational changes occur. The data clearly reveal that there are distinct positions where Mg^{2+} ions bind with substantial probability. Although c^* increase as the bulk Mg^{2+} concentration increases from 2.5 to 5.0 mM, many of the distinctive peaks exist even at 2.5 mM. This is a nontrivial prediction because, at $[Mg^{2+}] = 2.5$ mM, only TC1 and TC2 form with $\sim 50\%$ probability but all other contacts are either absent or formed with low probability (Fig. 3). This shows that coordination of Mg^{2+} to rRNA is nucleotide specific and does not occur in a random diffusive manner as is often assumed. The

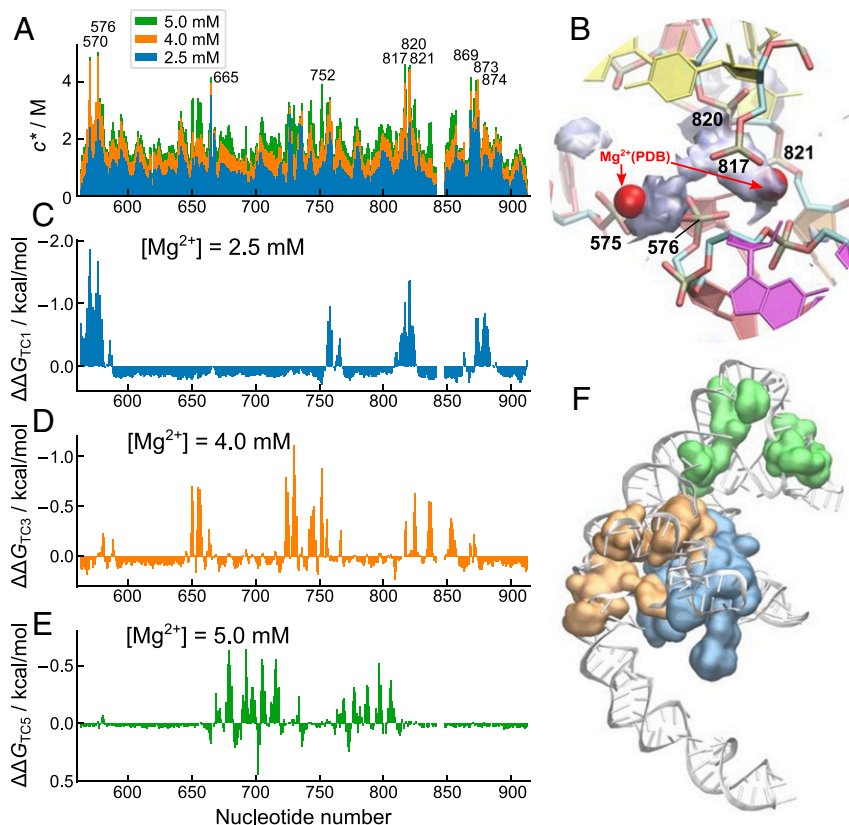


Fig. 4. Fingerprints of Mg^{2+} binding leading to tertiary contact formation. (A) Mg^{2+} binding to each nucleotide is quantified as the contact Mg^{2+} concentrations (c^*) defined in Eq. 1. The data are shown for three Mg^{2+} concentrations that correspond to the midpoints at which different TCs form (Fig. 3B). Some nucleotides that have high c^* are labeled on top. *SI Appendix, Fig. S3* shows similar plots for other solution conditions. (B) Mg^{2+} binding in the central junction. At the center, the filled space in gray represents the region where Mg^{2+} ions were highly localized in the simulations (space filled if more than 50% of the maximum density in the region at $[\text{Mg}^{2+}] = 5 \text{ mM}$). Two red spheres are Mg^{2+} ions solved in a crystal structure (PDB 4Y4O) (37). See *SI Appendix, Fig. S4* for other Mg^{2+} binding sites in h20, h25, and the three-way junction. (C–E) $\Delta\Delta G$ for (C) TC1, (D) TC3, and (E) TC5. The vertical axes are flipped for the ease of comparison to A. Note that part of the data in A–D are not continuous because nucleotides 842 through 847 do not exist in *T. Thermophilus*. (F) Three-dimensional positions of nucleotides with distinct peaks, $\Delta\Delta G < -0.5 k_B T$ ($= -0.31 \text{ kcal/mol}$), are shown using surface representation with the same color code as in C–E.

finding that the specificity of Mg^{2+} binding causes RNA tertiary structure formation appears to be general (26, 36).

To delineate how those Mg^{2+} ions facilitate individual tertiary-contact formation, we further analyzed the three-dimensional positions of Mg^{2+} ions around high- c^* nucleotides revealed in the fingerprint. In Fig. 4B, we present a display of the three-dimensional density of Mg^{2+} ions in the central junction. The highest Mg^{2+} -density region is surrounded by several phosphate groups associated with nucleotides C817, G821, and G576. All of these were detected as nucleotides that have high c^* (Fig. 4A) and thus tend to bind Mg^{2+} preferentially. Another nearby Mg^{2+} ion is located between two phosphate groups of nucleotides G575 and G576. We confirmed that there are precisely the two Mg^{2+} ions resolved in the same region in a high-resolution crystal structure of the same *Thermus thermophilus* (Protein Data Bank [PDB] 4Y4O) (37), which further validates the three-interaction-site (TIS) simulation model. See *SI Appendix, Fig. S4* for other Mg^{2+} binding motifs. It is gratifying that we are able to predict the precise locations of Mg^{2+} ions without adjusting any parameter in the model.

We then analyzed the relationship between TC formations and Mg^{2+} binding to specific positions. The effects of specific binding can be revealed by comparing the free-energy change upon formation of each TC. We define $\Delta\Delta G_\alpha(i)$ (*Materials and Methods* and Eq. 2), which is the difference between bound and unbound

states of Mg^{2+} to the i th nucleotide. The subscript α stands for conformational change considered ($\alpha = \text{TC1, TC2, } \dots, \text{TC6}$). The value of $\Delta\Delta G_\alpha(i)$ would be negative if a specific binding of Mg^{2+} to the i th nucleotide stabilizes the formation of α . Fig. 4 C–E shows $\Delta\Delta G$ for TC1, TC3, and TC5 formations, respectively, at the midpoint Mg^{2+} concentrations. Fig. 4 C–E shows quantitatively the free-energy gain due to binding of Mg^{2+} to specific sites. For example, Fig. 4C shows that TC1 formation stabilizes tertiary contacts involving nucleotides of h19, h20, h25, and a part of h24 (blue surface in Fig. 4F, A563–G587, G756–A767, G809–C826, and A872–U884). Tertiary structure stabilization in other regions occurs at higher Mg^{2+} concentrations. Comparing the three-dimensional positions in Fig. 4F and the locations of TC in Fig. 3C, we conclude that the formation of contacts is associated with specific Mg^{2+} binding to the same nucleotides.

Interestingly, the extent of stabilization by Mg^{2+} (absolute values of $\Delta\Delta G$) is larger in the nucleotides that form contacts at lower $[\text{Mg}^{2+}]$ than in nucleotides that require higher $[\text{Mg}^{2+}]$. For instance, values of $\Delta\Delta G_{\text{TC1}}$ at $[\text{Mg}^{2+}] = 2.5 \text{ mM}$ (Fig. 4C) are more negative than $\Delta\Delta G_{\text{TC3}}$ at 4.0 mM Mg^{2+} (Fig. 4D) or $\Delta\Delta G_{\text{TC6}}$ at 5.0 mM Mg^{2+} (Fig. 4E). This indicates that, at lower Mg^{2+} concentrations, Mg^{2+} ions preferentially bind to nucleotides that provide greater stability by forming Mg^{2+} -driven tertiary interactions. The results in Fig. 4 show clearly that

coordination of Mg^{2+} ions resulting in consolidation of tertiary structure in rRNA occurs in a discrete manner.

Tertiary Stacking Nucleates the Central Junction Assembly. The central junction (Fig. 1C) is stabilized by several tertiary base stacking interactions between nonconsecutive nucleotides (tertiary stacking [TST]). From the crystal structure, we detected six TST interactions around the central junction (SI Appendix, Fig. S5 A and B) that contribute to the correct folding. SI Appendix, Fig. S5C shows that the $[Mg^{2+}]$ -dependent formation of all six

TST interactions occurs cooperatively at the midpoint $[Mg^{2+}] \sim 2.5$ mM. Interestingly, this value corresponds to the midpoint of formation of TC1 and TC2 (compare with Fig. 3B). Thus, the ordering of the central junction takes place cooperatively with the simultaneous formations of TC1 and TC2. These events, resulting in partial folding, as well as the ordering of h19, contribute to the formation of the compact core region, leaving other longer peripheral helices unfolded. Structuring of these helices occurs only at higher $[Mg^{2+}]$.

Three-Way Junction Folds upon Tertiary-Contact Formation of Constituent Helices. The 3WJ consisting of h20 to h22 has been used as a representative folding motif in earlier experimental studies of the rRNA folding (23, 24, 27, 38, 39). In the unfolded state, the three helices are expected to be well separated without having major interactions with each other because of the electrostatic repulsion. Indeed, an experimental study done some time ago using transient electric birefringence indicated that the three angles between the helices are nearly equal (roughly 120°) in the absence of Mg^{2+} and r-protein S15 (23). On the other hand, in the folded state, the two helices h21 and h22 are coaxially stacked, while the other, h20, forms an acute angle with h22 (Fig. 1). Although, in the ribosome, S15 binds to the center of the junction and presumably stabilizes the folded form, experiments demonstrated that Mg^{2+} ions alone are sufficient to stabilize the native form (23, 27). In addition, further analyses showed that the folding of the junction is determined solely by the RNA sequence rather than by binding of S15 (24).

We calculated changes in the three angles between h20, h21, and h22 as a function of Mg^{2+} concentration (Fig. 5A). At Mg^{2+} concentration below 4 mM, the average of the angle $\Theta_{h20-h22}$ is about 80° , whereas the values of the other two angles are around 110° . On an average, the sum of the three angles is smaller than 360° , indicating that the helices are not entirely confined to a plane. All of the three angles fluctuate in the ensemble of conformations, which is shown by the $\approx 20^\circ$ in the standard deviations (Fig. 5A). Nonetheless, the three helices are aligned roughly in a radial manner, in accord with the experiments (23). At $[Mg^{2+}] \sim 5$ mM or above, the angles change dramatically. Two angles, $\Theta_{h21-h22}$ and $\Theta_{h20-h21}$, are around 150° , indicating that the two helices are coaxially stacked. The third, $\Theta_{h20-h22}$, adopts an acute value around 15° , showing that h20 is aligned toward h22, as in the crystal structure. As shown by the shaded regions in Fig. 5A, the fluctuations in the angles are also much less at the higher Mg^{2+} concentrations. The Mg^{2+} concentration at which the 3WJ folds corresponds to the concentration at which TC3 forms, which is reasonable because TC3 is the tertiary interactions between the two ends of h20 and h22.

To illustrate the fluctuations in the 3WJ as function of Mg^{2+} concentration, we calculated the two-dimensional distributions of specific angles between the helices and the distance, which could be used to compare with single-molecule experiments (22, 40) on a related but different 3WJ construct. We find that at low Mg^{2+} concentrations, the distributions are very broad (Fig. 5 B–D) but become narrower when Mg^{2+} increases. The distribution of the distance between G577 and C735 ($d_{h20-h22}$) is unusually broad at low Mg^{2+} concentrations, which is reminiscent of the Förster resonance energy transfer (FRET) efficiency distributions (22). The substantial width of the angle distributions, which are difficult to directly measure experimentally, also shows that upon folding there are substantial conformations in the 3WJ even in the intact central domain. The good comparison between simulations and experiments, obtained without tuning any parameter, again validates the coarse-grained force field.

Hydroxyl Radical Footprinting. Finally, we make predictions for hydroxyl-radical footprinting based on solvent accessible surface area (SASA) of the simulated ensembles at each Mg^{2+}

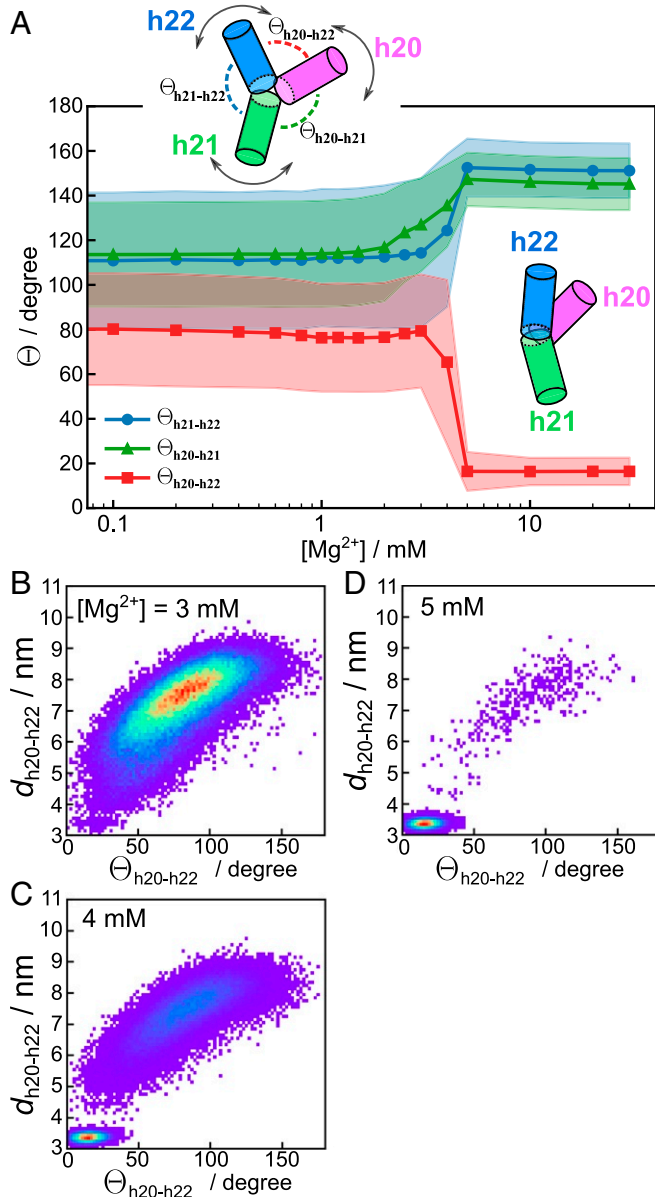


Fig. 5. Folding of the three-way junction. (A) Changes in the angles between helices at the three-way junction as functions of Mg^{2+} concentration. The angle, Θ , between a specific helix pair is shown. The filled regions represent the ranges of standard deviations in Θ . (B–D) Two-dimensional distributions of the distance ($d_{h20-h22}$) and the angle $\Theta_{h20-h22}$ between the helices h20 and h22. The distributions are shown at three Mg^{2+} concentrations around the transition points (B) 3 mM, (C) 4 mM, and (D) 5 mM. The distance $d_{h20-h22}$ was measured using the positions of the sugars of G577 and C735. The red color indicates the highest probability and the purple one is the lowest probability regions. SI Appendix, Fig. S6 shows similar plots for the other two angles, $\Theta_{h20-h21}$ and $\Theta_{h21-h22}$.

concentration. Footprinting is a powerful experimental technique to probe RNA structures in vitro and in vivo (6, 10, 41). Much like hydrogen–deuterium exchange experiments using NMR in the context of protein folding, hydroxyl-radical footprinting is used to assess the extent to which each nucleotide is exposed to the solvent. The solvent exposure has been demonstrated to be highly correlated with the SASA of sugar backbone (42, 43). We first calculated the SASA of each structure generated in the simulations and then determined the protection factor (F_p) of each nucleotide site by averaging over the simulated conformational ensemble at each Mg^{2+} concentration (Materials and Methods).

Figs. 6 A–E shows the nucleotide-dependent protection factors at various Mg^{2+} concentrations. At 0.2 mM Mg^{2+} , no nucleotide exhibits high protection factor, indicating that the RNA is unfolded other than the presence of stable secondary structures. At $[Mg^{2+}] = 1$ mM, only three nucleotides, G666, C726, and G727, show distinct protection ($F_p > 2$). In the crystal structure, G666 is near a bulge of h22 that interacts with a part of h23 (C726 and G727). Our simulation data show that this interaction is formed at relatively low Mg^{2+} concentration (~ 1 mM) and could in principle be detected by footprinting experiments. As Mg^{2+} concentration is increased, additional nucleotides are protected. At $[Mg^{2+}] = 2.5$ mM, another tertiary interaction at the center of h25 results in protections of G869, U870, and G874. Above $Mg^{2+} > 5$ mM, nucleotides around the core region are protected (nucleotides colored in green in Fig. 6 G and H). Many of these nucleotides are located in the cen-

tral junction and are involved in tertiary contacts, as shown in Fig. 3C.

For reference, we also calculated the footprinting profile using the crystal structure coordinates (Fig. 6F). Note that the protection factors calculated from a single-crystal structure should be less accurate and overestimate the protections compared to the values under solution conditions because thermal fluctuations are not included in the calculation. Moreover, the crystal packing reinforces molecular rigidity. Despite these caveats, the positions of the peaks are consistent with the profile we obtained in the simulations at high Mg^{2+} concentration (Fig. 6E; $[Mg^{2+}] = 20$ mM). The predictions in Fig. 6 can be experimentally tested using the footprinting technique.

Discussion

We investigated folding of the central domain of the 16S rRNA with particular emphasis on how Mg^{2+} drives structure formation. To our knowledge, experiments have not investigated the folding of the intact domain in the absence of proteins although the expectation is that the central domain could self-assemble autonomously. Consequently, many of our results are predictions that are amenable to tests using standard experimental techniques. The central domain of the 16S rRNA unfolds at low Mg^{2+} concentrations (roughly below 2 mM in our simulations), where only the secondary structures are intact. As Mg^{2+} concentration is increased, tertiary interactions form in a hierarchical manner in three distinct stages. Considering that the typical Mg^{2+} concentration of bacterial cytoplasm is ~ 1 mM (44), our

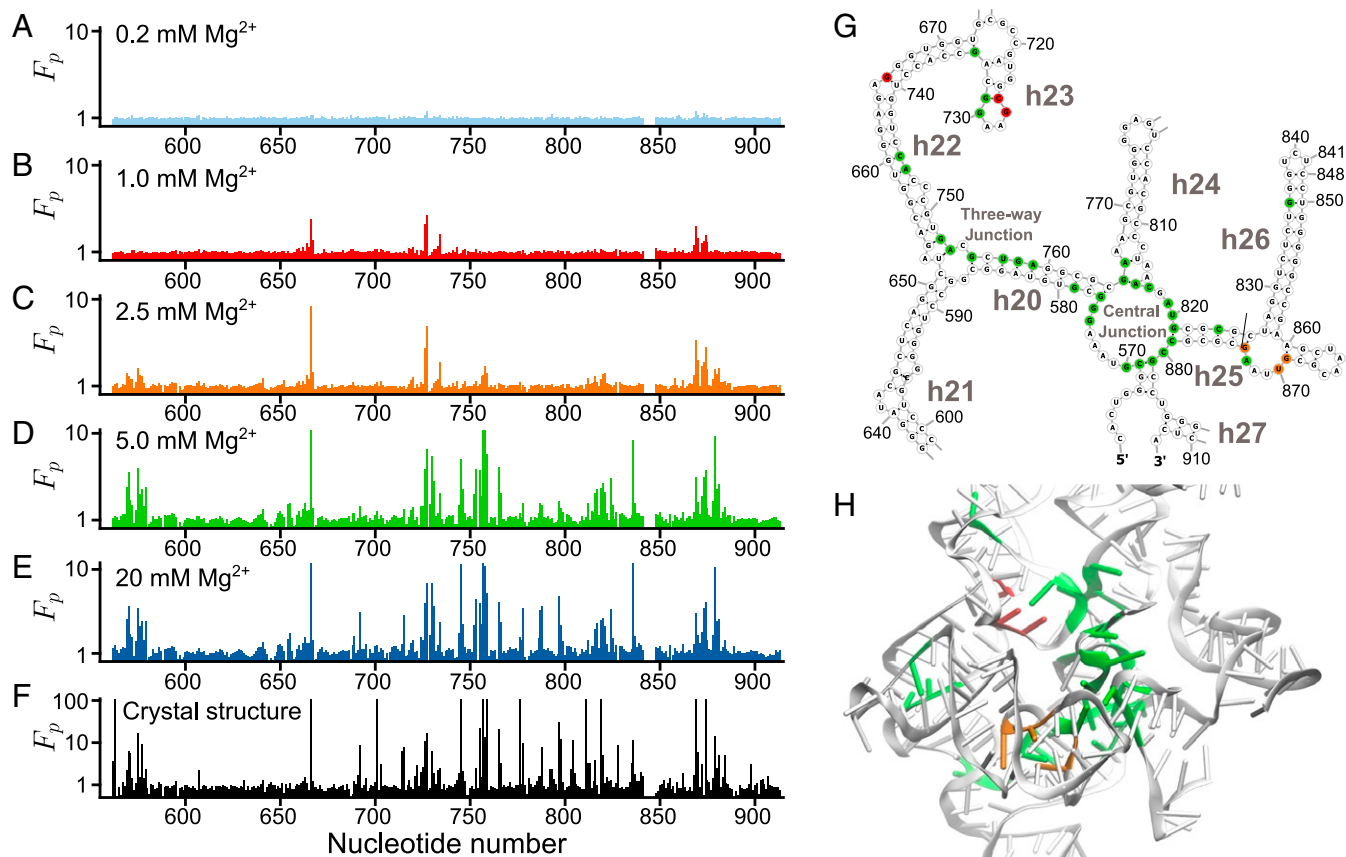


Fig. 6. Prediction of Mg^{2+} -dependent protection factors. (A–E) Footprinting protection factors estimated based on SASA at nucleotide resolution. Mg^{2+} concentrations are labeled on the top of each panel. (F) Protection factors estimated based on SASA calculated for the crystal structure (PDB: 1J5E). A few peaks exceeding $F_p > 100$ were truncated for clarity of presentation. These values are likely to be larger than those under solution conditions. (G and H) Secondary and tertiary structures with protected nucleotides highlighted with the same color scheme as in B–D.

data suggest that the rRNA would not spontaneously form a stable compact structure in the absence of r-proteins (Fig. 2). This conclusion is not definitive because folding *in vivo* occurs in a crowded milieu, which could lower the effective midpoint for folding (45–47).

Fate of the Three-Way Junction. The independent folding of the 3WJ and related constructs in the central domain (Fig. 1) has been studied extensively from over two decades ago by a variety of experimental methods (23, 24, 27). In a quest to understand the influence of r-proteins on the assembly of 30S particle they focused initially on the folding of the 3WJ, which folds in the presence of either S15 or Mg^{2+} (23). Using transient electric birefringence and a model to analyze the data (23), they inferred that the angles between the three helices (h20, h21, and h22) are roughly 120° , which implies that they adopt a planar structure. Upon addition of 1.5 mM Mg^{2+} (or S15) the 3WJ is structured in which h21 is coaxially stacked with h22, and h20 forms an acute angle with h22. The results of our simulations, carried out using the entire central domain, are in excellent agreement with experiments (Fig. 5). We find that the 3WJ folds cooperatively with the formation of tertiary interactions between h20 and h22 (TC3). The good agreement between simulations and experiments not only validates the model but also shows that the folding of the 3WJ does not change significantly when embedded in the intact central domain.

Specificity of Mg^{2+} Binding. The role of Mg^{2+} ions is particularly difficult to probe experimentally because the binding to RNA is likely to be cooperative. Indeed, we find that the Hill coefficient extracted from the dependence of R_g on Mg^{2+} condensation is roughly 3, an indication of cooperative binding (Fig. 2A). Our simulations show in unequivocal terms that Mg^{2+} ion binding is discrete and highly specific. Even at the lowest concentration of Mg^{2+} , there are specific nucleotides where the local concentrations of the divalent ions are significantly higher than the bulk value. As shown in Fig. 4 and *SI Appendix, Fig. S4*, these Mg^{2+} ions are often surrounded by two or more phosphate groups that come into proximity forming tertiary contacts. These Mg^{2+} ions should effectively reduce the negative electrostatic potential. It is possible that these discrete and specific binding events nucleate the folding of the RNA, which was proposed long ago based on the crystal structure of the P4-P6 domain of the *Tetrahymena* ribozyme (48). Indeed, an additional theme that is emerging through simulations is that Mg^{2+} ion binding, which can occur directly or mediated by a single water molecule (49), is specific and dictated by the architecture of the folded state (25, 26). The present study adds one more example of this concept.

Transition Midpoint Is Not Unique. A corollary of specificity of Mg^{2+} ion binding is that the midpoint concentration of Mg^{2+} at which the RNA folds is not unique. This is evident in Fig. 3 which shows that the important tertiary interactions that stabilize RNA order at different Mg^{2+} concentrations. The variations likely reflect the differences in the stabilities in the different regions, with the least stable one requiring higher Mg^{2+} concentration. The nonuniqueness of ordering temperature or denaturation concentration has been previously predicted for proteins (50) in which (typically) the secondary structural elements are not stable, thus making it difficult to validate the predictions experimentally.

Estimates of Mg^{2+} Midpoints. A few remarks on the measurement or calculation of Mg^{2+} midpoints (referred to as c_m), the Mg^{2+} concentrations at which individual secondary or tertiary interactions order, are worth making. The absolute values of

c_m are not highly significant except as a qualitative assessment of the ion-induced folding reaction. There are three principal reasons for making this assertion: 1) The c_m values depend on the RNA concentration, c_{RNA} . For example, experiments show that the value of c_m for *Azoarcus* ribozyme folding is ≈ 0.34 mM at $c_{RNA} = 6.3 \mu M$ (51), increasing to $c_m \approx 0.88$ mM at $c_{RNA} = 15.8 \mu M$ (52). We expect that, at c_m , the concentration of the free ion ($c_{m,0}$) is almost (not exactly) negligible (25), which implies that $c_{m,0} = c_m - m \cdot c_{RNA} \approx 0$ where m is the number of Mg^{2+} ions bound to one RNA molecule. This relationship shows that roughly $c_m \propto c_{RNA}$. 2) In our simulations $c_{RNA} = 38.7 \mu M$, which is larger than the typical value used in experiments by a factor of ~ 2 . In light of the linear dependence of c_m on c_{RNA} noted above, we expect the predicted c_m values to be about a factor of 2 larger than in experiments. 3) The estimate of c_m also depends on the order parameter used to assess the extent of folding. Different probes could and do give different values of c_m . The value of c_m obtained using R_g would be different if native gel assay or average FRET efficiency is used even if c_{RNA} is fixed. Of course, the c_m values would not differ significantly (at most a factor 2 to 3, say) if reasonable values (range in which no inter-RNA interactions occur) of c_{RNA} are used, and an appropriate order parameter is considered. In light of the arguments given here, we conclude that c_m should be treated as a qualitative measure of the efficiency of ions to fold RNA. For instance, for a fixed c_{RNA} , the c_m s are measures of how efficient a particular ion is in folding the RNA.

Concluding Remarks. Currently there is no alternative computational method other than the TIS model with explicit ions for simulating Mg^{2+} -dependent folding of large RNA molecules. The present study on the central domain of rRNA, which yields results that are consistent with experiments on the 3WJ, has produced a number of predictions that could be tested. Based on the results presented here and elsewhere for other RNA constructs, we believe that the proposed coarse-grained model is transferable. The present study sets the stage not only for simulations of the intact rRNA but also for probing RNA–protein interactions by the TIS model and the coarse-grained model for proteins (53, 54).

Materials and Methods

Three-Interaction-Site Model. We used the TIS coarse-grained RNA model with explicit ions (25) to simulate the folding of the ribosomal RNA. In our previous study, we established that the model quantitatively reproduced thermodynamics of folding of RNA hairpins and pseudoknots as well as folding of the 195-nucleotide *Azoarcus* ribozyme (25). Because the details of the model have been reported previously (25), we provide only a brief description here. In the TIS model (55), each nucleotide is represented by three coarse-grained spherical beads corresponding to phosphate, ribose sugar, and a base. Ions in the solution, Mg^{2+} , K^+ , and Cl^- , are explicitly treated whereas water is modeled implicitly with a temperature-dependent dielectric constant. Briefly, the effective potential energy is taken to be $U_{TIS} = U_L + U_{EV} + U_{ST} + U_{HB} + U_{EL}$, where U_L accounts for chain connectivity and bending stiffness of the polynucleic acids, U_{EV} accounts for excluded volume interactions of each chemical group including interactions between RNA sites and ions, and U_{ST} and U_{HB} are the base-stacking and hydrogen-bond interactions, respectively. All of the consecutive bases have the base-stacking interactions, of which the strength depends on the sequence. Any pair of the canonical Watson–Crick base pairs (A–U and G–C) and the Wobble base pair (G–U), separated by at least four nucleotides along the chain, can form hydrogen bonds and thus contribute U_{HB} . In other words, the model accounts for certain nonnative interactions as well. Besides those general stacking and hydrogen-bonding interactions, we generate a list of tertiary stacking and hydrogen bonding that appear in the specific RNA based on the crystal structure (see the next paragraph). The interactions between the charged moieties (phosphate groups and ions), U_{EL} , happen via the standard Coulomb potential. The values of the charges on phosphate groups and ions are -1 (phosphate), $+2$ (Mg), $+1$ (K), and -1 (Cl). All of the force-field parameters used here are the same as in our earlier study (25). Applications to a wide variety of RNA molecules show that

the TIS model is transferable, quantitatively accounting for many aspects of RNA folding, although it would require exhaustive testing to be sure that this is so.

RNA Molecule. We investigated the folding thermodynamics of the central domain of the small subunit (16S) of *T. thermophilus* ribosome. The central domain of 16S rRNA consists of ~350 nucleotides (C562–A914 in *T. thermophilus*, PDB entry 1J5E) (31). The secondary structure map and the tertiary structure are displayed in Fig. 1C. The list of hydrogen bonds was generated by the WHAT-IF server based on the crystal structure (56). The secondary structure diagram was generated using RNApdbee (57) and Fornal (58). All three-dimensional graphics of this paper were generated with VMD (59).

Simulations. In this study, we examined the effects of Mg²⁺ ions by varying the concentration from 0 to 30 mM in the presence of 50 mM K⁺, a typical value contained in the Tris buffer. All simulations were conducted using a periodic cubic box (each side is 35 nm) containing a different number of Mg²⁺, determined by the concentration of divalent ions, and a fixed number of K⁺ ions. An appropriate number of anions (Cl⁻) was added to neutralize the entire system. We performed low-friction Langevin dynamics simulations at 37 °C to sample conformations of the system containing the ribosomal RNA and ions (60). For each condition (Mg²⁺ concentrations), at least 60,000 conformations were collected from the equilibrated trajectories.

Calculation of Contact Mg²⁺ Concentration. To quantify the affinity of Mg²⁺ at each nucleotide site, we computed a contact ion concentration around the *i*th nucleotide using (25)

$$c_i^* = \frac{1}{N_A V_c} \int_0^{r_c} \rho_i(r) 4\pi r^2 dr, \quad [1]$$

where $\rho_i(r)$ is number density of the ion at the distance r from the phosphate of the *i*th nucleotide, V_c is the spherical volume of radius r_c , and N_A is the Avogadro's number to represent c_i^* in molar units. To count only tightly bound Mg²⁺, we used a cutoff distance, $r_c = R_{Mg} + R_p + \Delta r$, where R_{Mg} and R_p are radii of Mg²⁺ and phosphate sites, respectively, and $\Delta r = 0.15$ nm is a margin for contact formation. Since Δr is small, the quantity c_i^* corresponds to the local molar concentration of Mg²⁺ at the surface of the phosphate groups.

Calculation of Mg²⁺ Effects on Conformational Changes. We computed the free-energy contribution of Mg²⁺ binding to conformational changes as $\Delta\Delta G_{\alpha}$, where α represents a certain conformational change (e.g., formation of certain TCs). This quantity can be defined for each Mg²⁺ binding site, namely $\Delta\Delta G_{\alpha}(i)$ for the *i*th nucleotide. The changes in the stability upon specific Mg²⁺ binding are

$$\Delta\Delta G_{\alpha}(i) = \Delta G_{\alpha}^{Mg}(i) - \Delta G_{\alpha}^{\phi}(i), \quad [2]$$

where the superscripts Mg and ϕ indicate that Mg²⁺ is bound and unbound to the *i*th nucleotide, respectively. Each term on the right-hand side gives the stability due to contact formation, given that Mg²⁺ is bound or unbound to the *i*th nucleotide. In other words,

$$\Delta G_{\alpha}^{Mg}(i) = G_{\alpha F}^{Mg}(i) - G_{\alpha U}^{Mg}(i), \quad [3]$$

where $G_{\alpha F}^{Mg}$ and $G_{\alpha U}^{Mg}$ are the free energies of the states where the contact α is formed and disrupted, respectively, given that Mg²⁺ is bound to the *i*th nucleotide. Similarly,

$$\Delta G_{\alpha}^{\phi}(i) = G_{\alpha F}^{\phi}(i) - G_{\alpha U}^{\phi}(i) \quad [4]$$

is the stability of the contact given that Mg²⁺ is unbound. In our simulations, we calculated the difference in the free energy as a combination of the joint probabilities computed from the ensembles of conformations,

$$\Delta\Delta G_{\alpha}(i) = -k_B T \ln \frac{P_{\alpha,i}(F, Mg)P_{\alpha,i}(U, \phi)}{P_{\alpha,i}(U, Mg)P_{\alpha,i}(F, \phi)}, \quad [5]$$

where each $P_{\alpha,i}$ is a joint probability. For instance, $P_{\alpha,i}(F, Mg)$ is the joint probability that the contact α is formed, and Mg²⁺ is bound to the *i*th nucleotide.

Angles between Helices. Our previous work had shown that a key element in the self-assembly of *Azoarcus* ribozyme is the establishment of an angle between certain helices, which due to topological frustration occurs only at high Mg²⁺. To determine whether this is the case for this piece of the rRNA we computed the Mg²⁺-dependent angles between certain helices. In the three-way junction (Fig. 1C), angles (Θ) between helices h20, h21, and h22 were calculated using the axis of each helix. The helix axes were computed using Kahn's algorithm (61) using nucleotides 577 to 586 and 755 to 764 for h20, 588 to 597 and 643 to 651 for h21, and 655 to 672 and 734 to 751 for h22.

Footprinting Protecting Factors (F_p). To calculate the SASA, we first reconstructed atomistic structures based on coarse-grained coordinates using a tool developed in house that employs a fragment-assembly approach and energy minimization by AmberTools (62–64). Using the reconstructed atomically detailed structures, SASA was computed with Free-SASA version 2.0 (65). It is known that experimental footprinting data, obtained using hydroxyl radicals, are highly correlated with SASA of the sugar backbone (42, 43). Considering that hydroxyl radicals preferably cleave C4' and C5' atoms of the RNA backbone (43), we assigned the larger SASA value of C4' and C5' atoms to each nucleotide. From the SASA data, we computed the protection factor of the *i*th nucleotides as

$$F_p(i) = \frac{\langle SASA(i) \rangle_{\text{Unfolded}}}{\langle SASA(i) \rangle}, \quad [6]$$

where the angle bracket indicates an ensemble average (66). We used the conformations obtained in the absence of Mg²⁺ to compute the average SASA of the unfolded state, $\langle SASA \rangle_{\text{Unfolded}}$, as the reference state.

Hill Equation for R_g . The global transition of the rRNA size, measured by R_g as a function of Mg²⁺ concentration ($[Mg^{2+}]$), was fitted to the Hill equation,

$$\theta([Mg^{2+}]) = \frac{[Mg^{2+}]^n}{[Mg^{2+}]_m^n + [Mg^{2+}]^n}. \quad [7]$$

The two parameters, the Hill coefficient (n) and the midpoint of Mg²⁺ concentration ($[Mg^{2+}]_m$), were obtained by fitting the simulation data to Eq. 7. Before fitting, we converted the average R_g values at each $[Mg^{2+}]$ to θ , using

$$\theta = -\frac{R_g - \min(R_g)}{\max(R_g) - \min(R_g)} + 1, \quad [8]$$

where $\max(R_g)$ and $\min(R_g)$ were taken from the average R_g values at the lowest and highest Mg²⁺ concentrations, respectively.

Data Availability. The simulation code and all of the force-field parameters are available at GitHub (<https://github.com/naotohori/16S-central-folding>) (67). The data of three-dimensional distribution of Mg²⁺ are also available at <https://doi.org/10.5281/zenodo.4304537> (68).

ACKNOWLEDGMENTS. Much of this work was carried out while we were at the Institute for Physical Science and Technology at the University of Maryland. Comments from Harold Kim have been most useful. N.H. is grateful to Hung T. Nguyen and Mauro L. Mugnai for valuable discussions. This work was supported in part by a grant from the National Science Foundation (CHE 19-00093) and the Collier-Welch Regents Chair (F-0019) administered through the Welch Foundation.

1. M. W. Talkington, G. Siuzdak, J. R. Williamson, An assembly landscape for the 30S ribosomal subunit. *Nature* **438**, 628–632 (2005).
2. T. Adilakshmi, D. L. Bellur, S. A. Woodson, Concurrent nucleation of 16S folding and induced fit in 30S ribosome assembly. *Nature* **455**, 1268–1272 (2008).
3. S. A. Woodson, RNA folding and ribosome assembly. *Curr. Opin. Chem. Biol.* **12**, 667–673 (2008).
4. Z. Shajani, M. T. Sykes, J. R. Williamson, Assembly of bacterial ribosomes. *Annu. Rev. Biochem.* **80**, 501–526 (2011).
5. J. H. Davis, J. R. Williamson, Structure and dynamics of bacterial ribosome biogenesis. *Philos. Trans. R. Soc. B* **372**, 20160181 (2017).

6. S. Stern, T. Powers, L. Changchien, H. Noller, RNA-protein interactions in 30S ribosomal subunits: Folding and function of 16S rRNA. *Science* **244**, 783–790 (1989).
7. M. M. Yusupov *et al.*, Crystal structure of the ribosome at 5.5 Å resolution. *Science* **292**, 883–896 (2001).
8. S. Mizushima, M. Nomura, Assembly mapping of 30S ribosomal proteins from *E. coli*. *Nature* **226**, 1214–1218 (1970).
9. M. Herold, K. Nierhaus, Incorporation of six additional proteins to complete the assembly map of the 50S subunit from *Escherichia coli* ribosomes. *J. Biol. Chem.* **262**, 8826–8833 (1987).

10. S. A. Woodson, RNA folding pathways and the self-assembly of ribosomes. *Acc. Chem. Res.* **44**, 1312–1319 (2011).
11. R. R. Samaha, B. O'Brien, T. W. O'Brien, H. F. Noller, Independent in vitro assembly of a ribonucleoprotein particle containing the 3' domain of 16S rRNA. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7884–7888 (1994).
12. S. C. Agalarov *et al.*, In vitro assembly of a ribonucleoprotein particle corresponding to the platform domain of the 30S ribosomal subunit. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 999–1003 (1998).
13. C. J. Weitzmann, P. R. Cunningham, K. Nurse, J. Ofengand, Chemical evidence for domain assembly of the Escherichia coli 30S ribosome. *Faseb. J.* **7**, 177–180 (1993).
14. J. Lai, K. Chen, Z. Luthey-Schulten, Structural intermediates and folding events in the early assembly of the ribosomal small subunit. *J. Phys. Chem. B* **117**, 13335–13345 (2013).
15. K. Chen *et al.*, Assembly of the five-way junction in the ribosomal small subunit using hybrid MD–Gö simulations. *J. Phys. Chem. B* **116**, 6819–6831 (2012).
16. H. Kim *et al.*, Protein-guided RNA dynamics during early ribosome assembly. *Nature* **506**, 334–338 (2015).
17. D. Thirumalai, S. Woodson, Kinetics of folding of proteins and RNA. *Acc. Chem. Res.* **29**, 433–439 (1996).
18. D. Thirumalai, C. Hyeon, RNA and protein folding: Common themes and variations. *Biochemistry* **44**, 4957–4970 (2005).
19. D. J. Klein, P. B. Moore, T. A. Steitz, The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA* **10**, 1366–1379 (2004).
20. K. H. Nierhaus, Mg²⁺, K⁺, and the ribosome. *J. Bacteriol.* **196**, 3817–3819 (2014).
21. T. K. Lenz, A. M. Norris, N. V. Hud, L. D. Williams, Protein-free ribosomal RNA folds to a near-native state in the presence of Mg²⁺. *RSC Adv.* **7**, 54674–54681 (2017).
22. T. Ha *et al.*, Ligand-induced conformational changes observed in single RNA molecules. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9077–9082 (1999).
23. J. W. Orr, P. J. Hagerman, J. R. Williamson, Protein and Mg²⁺-induced conformational changes in the S15 binding site of 16S ribosomal RNA. *J. Mol. Biol.* **275**, 453–464 (1998).
24. R. T. Batey, J. R. Williamson, Effects of polyvalent cations on the folding of an rRNA three-way junction and binding of ribosomal protein S15. *RNA* **4**, 984–997 (1998).
25. N. A. Denesyuk, D. Thirumalai, How do metal ions direct ribozyme folding? *Nat. Chem.* **7**, 793–801 (2015).
26. N. Hori, N. A. Denesyuk, D. Thirumalai, Ion condensation onto ribozyme is site specific and fold dependent. *Biophys. J.* **116**, 2400–2410 (2019).
27. S. C. Agalarov, G. Sridhar Prasad, P. M. Funke, C. D. Stout, J. R. Williamson, Structure of the S15, S6, S18–rRNA complex: Assembly of the 30S ribosome central domain. *Science* **288**, 107–113 (2000).
28. M. I. Recht, J. R. Williamson, RNA tertiary structure and cooperative assembly of a large ribonucleoprotein complex. *J. Mol. Biol.* **344**, 395–407 (2004).
29. T. Lavergne *et al.*, FRET characterization of complex conformational changes in a large 16S ribosomal RNA fragment site-specifically labeled using unnatural base pairs. *ACS Chem. Biol.* **11**, 1347–1353 (2016).
30. A. S. Petrov *et al.*, Secondary structures of rRNAs from all three domains of life. *PLoS One* **9**, e88222 (2014).
31. B. T. Wimberly *et al.*, Structure of the 30S ribosomal subunit. *Nature* **407**, 327 (2000).
32. J. A. Aronovitz, D. R. Nelson, Universal features of polymer shapes. *J. Phys.* **47**, 1445–1456 (1986).
33. C. Hyeon, R. I. Dima, D. Thirumalai, Size, shape, and flexibility of RNA structures. *J. Chem. Phys.* **125**, 194905 (2006).
34. R. I. Dima, D. Thirumalai, Asymmetry in the shapes of folded and denatured states of proteins. *J. Phys. Chem. B* **108**, 6564–6570 (2004).
35. S. S. Cho, D. L. Pincus, D. Thirumalai, Assembly mechanisms of RNA pseudoknots are determined by the stabilities of constituent secondary structures. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 17349–17354 (2009).
36. S. J. Chen, Site-specific binding of non-site-specific ions. *Biophys. J.* **116**, 2237–2239 (2019).
37. Y. S. Polikanov, S. V. Melnikov, D. Söll, T. A. Steitz, Structural insights into the role of rRNA modifications in protein synthesis and ribosome assembly. *Nat. Struct. Mol. Biol.* **22**, 342–344 (2015).
38. A. Lescouste, E. Westhof, Topology of three-way junctions in folded RNAs. *RNA* **12**, 83–93 (2006).
39. K. A. Baker, R. Lamichhane, T. Lamichhane, D. Rueda, P. R. Cunningham, Protein–RNA dynamics in the central junction control 30S ribosome assembly. *J. Mol. Biol.* **428**, 3615–3631 (2016).
40. H. D. Kim *et al.*, Mg²⁺-dependent conformational change of RNA studied by fluorescence correlation and FRET on immobilized single molecules. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4284–4289 (2002).
41. R. M. Hulscher *et al.*, Probing the structure of ribosome assembly intermediates in vivo using DMS and hydroxyl radical footprinting. *Methods* **103**, 49–56 (2016).
42. J. H. Cate *et al.*, Crystal structure of a group I ribozyme domain: Principles of RNA packing. *Science* **273**, 1678–1685 (1996).
43. B. Balasubramanian, W. K. Pogozelski, T. D. Tullius, DNA strand breaking by the hydroxyl radical is governed by the accessible surface areas of the hydrogen atoms of the DNA backbone. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9738–9743 (1998).
44. E. A. Groisman *et al.*, Bacterial Mg²⁺ homeostasis, transport, and virulence. *Annu. Rev. Genet.* **47**, 625–646 (2013).
45. D. Kilburn, J. H. Roh, L. Guo, R. M. Briber, S. A. Woodson, Molecular crowding stabilizes folded RNA structure by the excluded volume effect. *J. Am. Chem. Soc.* **132**, 8690–8696 (2010).
46. N. A. Denesyuk, D. Thirumalai, Crowding promotes the switch from hairpin to pseudoknot conformation in human telomerase RNA. *J. Am. Chem. Soc.* **133**, 11858–11861 (2011).
47. N. A. Denesyuk, D. Thirumalai, Entropic stabilization of the folded states of RNA due to macromolecular crowding. *Biophys. Rev.* **5**, 225–232 (2013).
48. J. H. Cate, R. L. Hanna, J. A. Doudna, A magnesium ion core at the heart of a ribozyme domain. *Nat. Struct. Biol.* **4**, 553–558 (1997).
49. H. T. Nguyen, D. Thirumalai, Charge density of cation determines inner versus outer shell coordination to phosphate in RNA. *J. Phys. Chem. B* **124**, 4114–4122 (2020).
50. E. P. O'Brien, B. R. Brooks, D. Thirumalai, Molecular origin of constant m-values, denatured state collapse, and residue-dependent transition midpoints in globular proteins. *Biochemistry* **48**, 3743–3754 (2009).
51. S. Chauhan *et al.*, RNA tertiary interactions mediate native collapse of a bacterial group I ribozyme. *J. Mol. Biol.* **353**, 1199–1209 (2005).
52. J. H. Roh *et al.*, Multistage collapse of a bacterial ribozyme observed by time-resolved small-angle X-ray scattering. *J. Am. Chem. Soc.* **132**, 10148–10154 (2010).
53. Z. Liu, G. Reddy, E. P. O'Brien, D. Thirumalai, Collapse kinetics and chevron plots from simulations of denaturant-dependent folding of globular proteins. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7787–7792 (2011).
54. G. Reddy, Z. Liu, D. Thirumalai, Denaturant-dependent folding of GFP. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17832–17838 (2012).
55. C. Hyeon, D. Thirumalai, Mechanical unfolding of RNA hairpins. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 6789–6794 (2005).
56. G. Vriend, What IF: A molecular modeling and drug design program. *J. Mol. Graph.* **8**, 52–56 (1990).
57. T. Zok *et al.*, RNApDbee 2.0: Multifunctional tool for RNA structure annotation. *Nucleic Acids Res.* **46**, W30–W35 (2018).
58. P. Kerpedjiev, S. Hammer, I. L. Hofacker, Forna (force-directed RNA): Simple and effective online RNA secondary structure diagrams. *Bioinformatics* **31**, 3377–3379 (2015).
59. W. Humphrey, A. Dalke, K. Schulten, VMD – Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
60. J. D. Honeycutt, D. Thirumalai, The nature of folded states of globular proteins. *Biopolymers* **32**, 695–709 (1992).
61. P. C. Kahn, Defining the axis of a helix. *Comput. Chem.* **13**, 185–189 (1989).
62. N. Hori, TIS2AA (2017). <https://zenodo.org/record/581485>. Accessed 19 July 2020.
63. E. Humphris-Narayanan, A. M. Pyle, Discrete RNA libraries from pseudo-torsional space. *J. Mol. Biol.* **421**, 6–26 (2012).
64. D. Case *et al.*, AMBER 2018 (University of California, San Francisco, CA, 2018).
65. S. Mitternacht, FreeSASA: An open source C library for solvent accessible surface area calculations. *F1000 Res.* **5**, 189 (2016).
66. P. L. Adams *et al.*, Crystal structure of a group I intron splicing intermediate. *RNA* **10**, 1867–1887 (2004).
67. N. Hori, N. A. Denesyuk, D. Thirumalai, 16S-central-folding. Github. <https://github.com/naotohori/16S-central-folding>. Deposited 3 December 2020.
68. N. Hori, N. A. Denesyuk, D. Thirumalai, Supplemental Data for "Shape changes and cooperativity in the folding of central domain of the 16S ribosomal RNA." Zenodo. <https://zenodo.org/record/4304537>. Deposited 3 December 2020.