

Quantitative tests of a reconstitution model for RNA folding thermodynamics and kinetics

Namita Bisaria^a, Max Greenfeld^{a,b}, Charles Limouse^c, Hideo Mabuchi^c, and Daniel Herschlag^{a,b,d,1}

^aDepartment of Biochemistry, Stanford University, Stanford, CA 94305; ^bDepartment of Chemical Engineering, Stanford University, Stanford, CA 94305; ^cDepartment of Applied Physics, Stanford University, Stanford, CA 94305; and ^dDepartment of Chemistry, Stanford University, Stanford, CA 94305

Edited by Harry F. Noller, University of California, Santa Cruz, CA, and approved July 19, 2017 (received for review March 1, 2017)

Decades of study of the architecture and function of structured RNAs have led to the perspective that RNA tertiary structure is modular, made of locally stable domains that retain their structure across RNAs. We formalize a hypothesis inspired by this modularity—that RNA folding thermodynamics and kinetics can be quantitatively predicted from separable energetic contributions of the individual components of a complex RNA. This reconstitution hypothesis considers RNA tertiary folding in terms of ΔG_{align} , the probability of aligning tertiary contact partners, and ΔG_{tert} , the favorable energetic contribution from the formation of tertiary contacts in an aligned state. This hypothesis predicts that changes in the alignment of tertiary contacts from different connecting helices and junctions (ΔG_{HJH}) or from changes in the electrostatic environment ($\Delta G_{+/-}$) will not affect the energetic perturbation from a mutation in a tertiary contact ($\Delta \Delta G_{\text{tert}}$). Consistent with these predictions, single-molecule FRET measurements of folding of model RNAs revealed constant $\Delta \Delta G_{\text{tert}}$ values for mutations in a tertiary contact embedded in different structural contexts and under different electrostatic conditions. The kinetic effects of these mutations provide further support for modular behavior of RNA elements and suggest that tertiary mutations may be used to identify rate-limiting steps and dissect folding and assembly pathways for complex RNAs. Overall, our model and results are foundational for a predictive understanding of RNA folding that will allow manipulation of RNA folding thermodynamics and kinetics. Conversely, the approaches herein can identify cases where an independent, additive model cannot be applied and so require additional investigation.

RNA folding | single-molecule FRET | RNA tertiary structure | folding kinetics | folding thermodynamics

Structured RNAs are integral to many biological processes, including translation, genome maintenance, and the regulation of gene expression (1, 2). These processes require RNA to fold into intricate 3D structures and to undergo a series of structural transitions (3, 4). As such, an important goal has been to characterize these states, in terms of their conformations and the rates and equilibria that describe the transitions between them (5–8). A more distant but ultimately far-reaching challenge is to quantitatively predict the kinetics and thermodynamics of RNA transitions, an accomplishment that would demonstrate fundamental understanding and effectuate engineering of these systems.

Quantitative analyses of the melting temperatures for nucleic acid duplexes led to predictive rules for RNA secondary structure stability, known as nearest neighbor or Turner rules, such that the energetics of helix formation can be predicted by considering only the identity of each base pair, the identity of its immediate neighbors, and the salt concentration (9–11). In this model, the energetic contribution of each base pair step is additive and can be summed to predict the free energy to an accuracy of 0.2–0.5 kcal/mol for simple helices (12). Although many more thermodynamic measurements will be required to predict the secondary structure stability for large RNAs and RNAs containing any possible noncanonical elements, the current model has been applied with great impact throughout biology and bioengineering (e.g., refs. 13 and 14).

An ongoing major challenge is to develop comparable quantitative and predictive algorithms for RNA tertiary structure formation. Once formulated, this knowledge may be applied to increase our understanding of biological processes that involve structured RNAs, such as RNA conformational transitions in splicing and translation, and to allow engineering of controllable RNA elements in vivo. This challenge would be greatly simplified if it were possible to model RNA tertiary structure as the collective behavior of structural modules such that the energetics of tertiary structure formation could be predicted from the sum of separable terms corresponding to each module, akin to terms used in nearest neighbor rules for predicting secondary structure.

As outlined below, considerable work by many RNA researchers over the past decades has led to a common perspective that RNA's tertiary structure has modular properties (15–19). Based on these observations, we now ask whether this structural modularity extends to energetics, such that folding thermodynamics of complex RNAs can be described in terms of properties of components—a task that has proven elusive for proteins but may be more applicable to RNA given its structural properties. We formulate and test predictions of such a model, and our results provide initial support for energetic additivity in RNA folding thermodynamics and analogous kinetic properties. We discuss our results in terms of comparisons with proteins, noting differences between the properties of these biological macromolecules while taking advantage of the wealth of models and conceptual approaches from the more mature protein-folding field.

Background

RNA tertiary structure can be considered as a collection of helices connected by junctions (i.e., the residues connecting canonical helical elements), with sparsely distributed tertiary contacts that bring together distal parts of the structure (Fig. 1A). Several of

Significance

We propose and test predictions of a thermodynamic and kinetic model for RNA tertiary folding that is based on separable energetic contributions of RNA elements. We define these contributions based on the principle features of RNA, and we test the basic predictions of separability by determining whether the energetic contributions of one component are affected by changes in another component. Our results support energetic separability of RNA elements and suggest that it may be possible to deconstruct RNAs into smaller parts that can be studied in isolation such that the individual folding behaviors of these parts can be used to “reconstitute” the folding of the original RNA.

Author contributions: N.B. and D.H. designed research; N.B. performed research; N.B., M.G., C.L., and H.M. contributed new reagents/analytic tools; N.B. analyzed data; and N.B. and D.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. Email: herschla@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1703507114/-DCSupplemental.

these tertiary interactions have been described as “tertiary motifs” because their sequences are conserved in different structured RNAs and their structures have been shown to closely overlay in three dimensions, suggesting a degree of modularity (15, 16, 18, 20).

As the residues of these tertiary motifs typically do not directly interact with residues outside of the motif, the atomic-level process for forming interactions within a motif may be the same in different RNAs. Thus, there may be a common underlying energetic contribution of the motif to folding. Alternatively, differing connections between tertiary elements could distort the motif's structure, alter the interactions made, and consequently lead to energetic interdependence between the motif and RNA elements that span the contacts. Initial tests of the kinetic behavior of one tertiary contact motif, the tetraloop/tetraloop receptor (TL/TLR), embedded within different structures revealed a common unfolding rate constant in different contexts, providing evidence for a modular, locally defined physical process (21).

With respect to junction elements, classic work by Lilley and coworkers (22–24) on three- and four-way junctions in DNA and RNA established rules for how the sequence and topology of these junctions determine the orientations and stacking of the helices that emanate from these junctions. Subsequent experiments demonstrated conserved orientation preferences for these and other junctions when placed in different RNAs, suggesting the transferability of junction-stacking properties (19, 25). More recently, it has been suggested that RNA substructures of two helices of defined lengths spanned by junctions of various topologies are basic building blocks of RNA tertiary structure; these substructures are referred to generically as helix-junction-helices (HJHs) (26–28).

According to this perspective, HJH elements determine the conformational ensemble explored by unfolded RNA, and ensembles have been demonstrated for highly conserved junctions such as the kink-turn and four-way junction and the biologically important *trans*-activation response element RNA junction (29–31). From this perspective, the probability of forming tertiary interactions depends on the fraction of time that the ensemble aligns two tertiary elements such that contacts form between them. In this way, we hypothesize that the conformational ensemble in isolation of the HJH elements that connect tertiary partners is a separable component in the determination of the probability of tertiary structure formation (27).

Given these observations, we formalized an energetic model for RNA folding consisting of separable contributors that, in principle, are predictable from the properties of isolated components and transferable between systems. We then tested quantitative predictions of this model using single-molecule FRET (smFRET) measurements of RNA folding thermodynamics and kinetics. Our results suggest that the canonical GAAA-11nt TL/TLR tertiary contact is thermodynamically and kinetically modular and that the kinetic and thermodynamic effects of mutations in this tertiary contact are transferable and thus potentially valuable tools in the dissection of folding steps for complex RNAs. Most broadly, these results provide an early step toward a quantitative and predictive understanding of RNA folding.

Results

An RNA Reconstitution Hypothesis. As described in the Introduction, an unfolded RNA exists in an ensemble of states in which the relative locations of tertiary contact partners (multiple ones for complex RNAs) are distributed in space based on the properties of the intervening HJH elements (Fig. 1A). Once aligned, particular tertiary partners can interact and thereby provide stabilization of intermediate and folded states (25–27, 30).

The folding scheme of Fig. 1B illustrates the energetic terms that underlie an RNA folding reconstitution hypothesis, and we use this scheme to describe the hypothesis and to illustrate tests of it. Fig. 1B depicts a simple RNA with a single HJH element

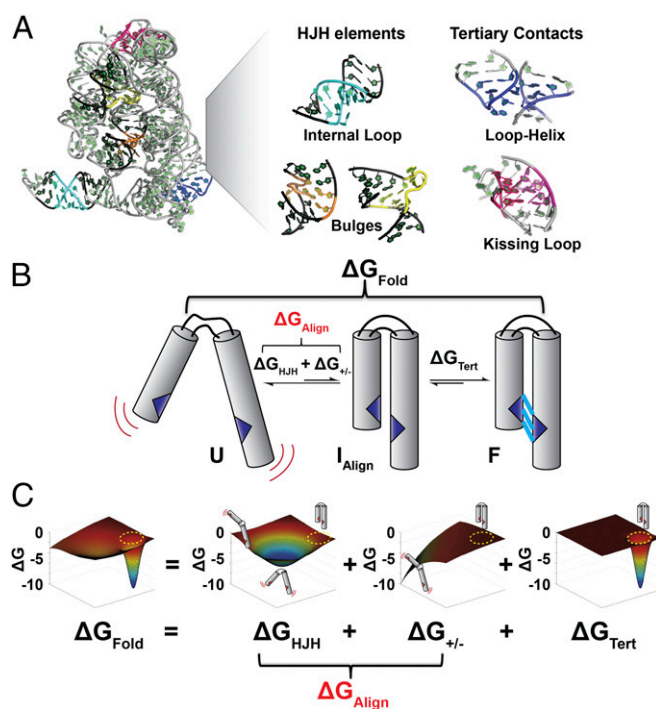


Fig. 1. Building blocks of structured RNAs and the reconstitution hypothesis. (A) Crystal structure of the group II intron with representative RNA tertiary structure building blocks highlighted (73). Tertiary elements include HJH elements, with helices in black and junctions colored: internal loop (cyan), bulges (orange and yellow), and tertiary contacts, loop-helix (blue) and kissing-loop (magenta). (B) The overall folding (ΔG_{Fold}) of a simple HJH element with a single tertiary contact (blue) is shown to illustrate the reconstitution hypothesis. The first folding step involves a search for the conformational states in which tertiary contact partners are aligned and can form (ΔG_{Align}). This step can be further separated into the contributions due to the orientational preferences of the HJH elements connecting the two helices (ΔG_{HJH}) and the electrostatic forces that bias the conformational ensemble of the HJH element to favor more extended conformations (ΔG_{HJH}). The second step in folding involves formation of the tertiary contact (ΔG_{Tert}) that stabilizes the folded state F. (C) The reconstitution hypothesis posits that the overall folding of a HJH element (ΔG_{Fold}) is the sum of the energetic contributions from the HJH elements, electrostatics, and tertiary contact formation ($\Delta G_{\text{Fold}} = \Delta G_{\text{HJH}} + \Delta G_{\text{HJH}} + \Delta G_{\text{Tert}}$). Each energetic contribution is figuratively depicted by a 3D free energy landscape, and these landscapes are summed to give the overall folding energetics. Free energy diagrams were generated in Matlab to qualitatively depict the folding landscape. The x-y plane represents all possible states of the HJH element, and encircled in yellow is the ensemble of states in which the tertiary contact is aligned to form. The value on the z axis is free energy represented by $\Delta G = -RT \ln(K_i)$, where K_i is the ratio of being in the aligned state relative to all other states. Illustrative conformations for the HJH element are shown above each free energy landscape.

and a single set of tertiary contact partners, and an illustrative 2D free energy landscape for folding of this RNA is depicted in Fig. 1C, Left as a distribution of different possible states. The free energy (ΔG_{Fold}) of the overall folding process can be considered in terms of two steps and three energetic terms (Eq. 1): The first step is aligning the tertiary contact partners (ΔG_{Align}), and the second is formation of the tertiary contact (ΔG_{Tert}) (Fig. 1B). [As commonly discussed, the stability of RNA helices allows us to consider tertiary folding starting with secondary structure preformed. The complexity of multiple secondary structures (e.g., for riboswitches and misfolded states) (32) is accommodated in this framework by deriving a value of ΔG_{Align} for each underlying secondary structure.] ΔG_{Align} is made up of two terms, representing the properties of the junction that determine the conformational space explored by the attached helices (ΔG_{HJH})

the electrostatic term ($\Delta G_{+/-}$) do not alter the response of a tertiary element to mutation so that $\Delta\Delta G_{\text{Fold}}$ is the same under different ionic conditions (Eq. 2). To test this prediction, we measured the folding energetics of the TL/TLR tertiary contact motif of the P4–P6 RNA (Fig. 3) using smFRET. P4–P6 RNA has been widely studied as a model for RNA tertiary folding (e.g., refs. 35–39), and our recent thermodynamic and kinetic framework for its folding provides the necessary structural, thermodynamic, and kinetic context for interpretation of the results herein (21). P4–P6 has two tertiary contacts, a canonical GAAA–11nt TL/TLR and a metal core/metal core receptor (MC/MCR), and their formation bring its two helical stacks into a roughly parallel conformation (37, 40–43) (Fig. 3B).

We used Ba^{2+} to vary ionic conditions, rather than Mg^{2+} , as Ba^{2+} does not bind to the MC of P4–P6 (44) and only the TL/TLR forms in the folded state (38) (Fig. 2B). Prior work demonstrated equivalent P4–P6 folding in Mg^{2+} and Ba^{2+} upon ablation of the MC, and independent studies demonstrated equal ion atmosphere occupancy for Mg^{2+} and Ba^{2+} surrounding a nucleic acid (21, 45). Thus, we could change the electrostatic environment without complications from site-specific metal ion binding and formation of an additional tertiary contact (Fig. 3B; see also *Additional Tests of the Potential Interplay Between Structural Context and Tertiary Contact Energetics*).

Fig. 2C shows the values of ΔG_{Fold} for each of four mutants in the TL/TLR of P4–P6 (colored symbols; Fig. 3C) determined using smFRET measurements of folding (21), compared with WT (black symbols, reproduced in each panel for comparison), over a series of Ba^{2+} concentrations. Example traces and data quality analyses for all smFRET measurements can be found in [Dataset S1](#). The range of Ba^{2+} concentrations was defined by conditions that allowed observation of folding transitions for both WT and mutant P4–P6 (*Methods*), and we obtained measured ΔG_{Fold} values over a range of ~ 4 kcal/mol.

Whereas the individual values of ΔG_{Fold} vary across Ba^{2+} concentrations (Fig. 2C), the differences between the WT and mutant folding free energies, $\Delta\Delta G_{\text{Fold}}$, were nearly invariant (Fig. 2D). For each mutant, the variation in $\Delta\Delta G_{\text{Fold, (WT-Mut)}}$ was less than 0.3 kcal/mole (Fig. 2D, green swatches), which corresponds to a less than twofold difference in equilibrium constants and is similar to measurement error ([Table S1](#)). The invariance of $\Delta\Delta G_{\text{Fold, (WT-Mut)}}$ for each of the four mutants supports the first prediction of the reconstitution hypothesis, that tertiary contact energetics (ΔG_{Tert}) are unaffected by changes in electrostatic screening ($\Delta G_{+/-}$) (Fig. 1 and Eq. 2):

$$\Delta\Delta G_{\text{Fold, (WT-Mut)}}^{[\text{Ion}]_1} = \Delta\Delta G_{\text{Fold, (WT-Mut)}}^{[\text{Ion}]_2} \quad [2]$$

From a physical perspective, these data suggest a degree of simplicity to the conformational and energetic landscape of the TL/TLR tertiary contact, specifically that the mutations and varying electrostatic conditions do not substantially alter the interactions that make up this tertiary contact.

Testing the Energetics of Tertiary Contact Formation in Different RNA Contexts. Analogous to the above prediction made for the relationship between $\Delta G_{+/-}$ and ΔG_{Tert} (Eq. 1), the reconstitution hypothesis predicts that the effect of a mutation in a tertiary contact should be the same in different structural contexts (ΔG_{HJH}). Fig. 44 illustrates this prediction, analogous to Fig. 24. To test this prediction, we compared the effects of the TL/TLR point mutations in P4–P6 to the effects of these mutations in an isolated TL/TLR construct (TL/TLR_{iso}), initially developed and studied extensively by Nesbitt and coworkers (46). In this construct, the two tertiary contact partners are connected via a flexible tether and thus lack the orientation biases imposed by P4–P6 (Fig. 4B).

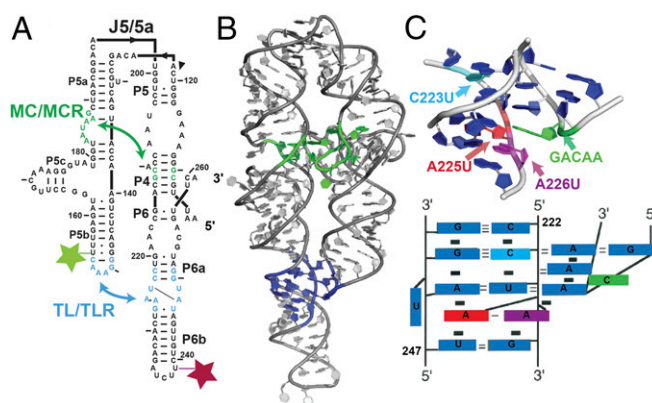


Fig. 3. P4–P6 RNA structure and point mutations. (A) Secondary structure of the P4–P6 RNA. Two tertiary contacts, the MC/MCR (green) and the TL/TLR (blue), are highlighted. (B) Tertiary structure of the P4–P6 RNA. The tertiary contacts are colored as in A [Protein Data Bank (PDB) ID code 1GID] (36). (C) Tertiary structure and schematic of the TL/TLR highlighting the residues that were mutated and studied herein: A225U (red), A226U (magenta), C223U (cyan), and GACAA (green); these colors are used throughout. A225U, A226U, and C223U are point mutations in the receptor, and GACAA is an insertion in the tetraloop.

Fig. 4C shows the folding free energy (ΔG_{Fold}) for each point mutation in TL/TLR_{iso} and in P4–P6, compared with the folding free energy of the corresponding WT molecule across a range of Ba^{2+} concentrations ([Table S2](#)). Because of stability differences, there are only two Ba^{2+} concentrations where we could obtain ΔG_{Fold} values for both constructs and thus directly compare $\Delta\Delta G_{\text{Fold, (WT-Mut)}}$ values for TL/TLR_{iso} and P4–P6 RNA. As shown in Fig. 4D (colored regions), the $\Delta\Delta G_{\text{Fold, (WT-Mut)}}$ values for each mutant at these concentrations are the same for TL/TLR_{iso} and P4–P6. The similar $\Delta\Delta G_{\text{Fold, (WT-Mut)}}$ values in these starkly different structural contexts suggest that the energetic contribution of tertiary contact formation (ΔG_{Tert}) is separable from that of the HJH context (ΔG_{HJH}) and further support the assumptions underlying the reconstitution hypothesis (Fig. 1 and Eq. 1).

Additional Tests of the Potential Interplay Between Structural Context and Tertiary Contact Energetics. In the previous section, we compared the TL/TLR_{iso}, which provides a simple flexible HJH context where the TL/TLR contact partners are able to assume a wide range of conformations in the folded state, to P4–P6, which has a more complex and conformationally restricted set of elements connecting the tertiary contact partners. Functional RNAs often contain multiple constraining tertiary contacts, and the formation of P4–P6's other tertiary contact, the MC/MCR (Fig. 3), would further constrain the encounter angles between tertiary elements. We therefore additionally tested the relationship between the energetics of tertiary contact formation (ΔG_{Tert}) and HJH context (ΔG_{HJH}) with the MC/MCR tertiary contact formed in P4–P6 (Figs. 3 and 54).

The ΔG_{Fold} values for P4–P6 folding with and without MC/MCR are shown in Fig. 5B. Mg^{2+} was used to allow MC/MCR formation and Ba^{2+} to prevent it, as above (44). Also as noted above, P4–P6 MC mutants that prevent Mg^{2+} binding have the same folding kinetics and thermodynamics in Mg^{2+} and Ba^{2+} (21, 47), suggesting that electrostatic screening is highly similar for these divalent metal ions and supporting comparison between Mg^{2+} and Ba^{2+} to isolate and test the effects of MC/MCR formation.

Due to the large stabilization of P4–P6 folding with the MC/MCR tertiary contact formed, we could not obtain folding data with Mg^{2+} and Ba^{2+} at the same concentrations. We therefore used the TL/TLR_{iso} construct to span the divalent metal ion concentration range for both Mg^{2+} and Ba^{2+} used in the P4–P6 measurements

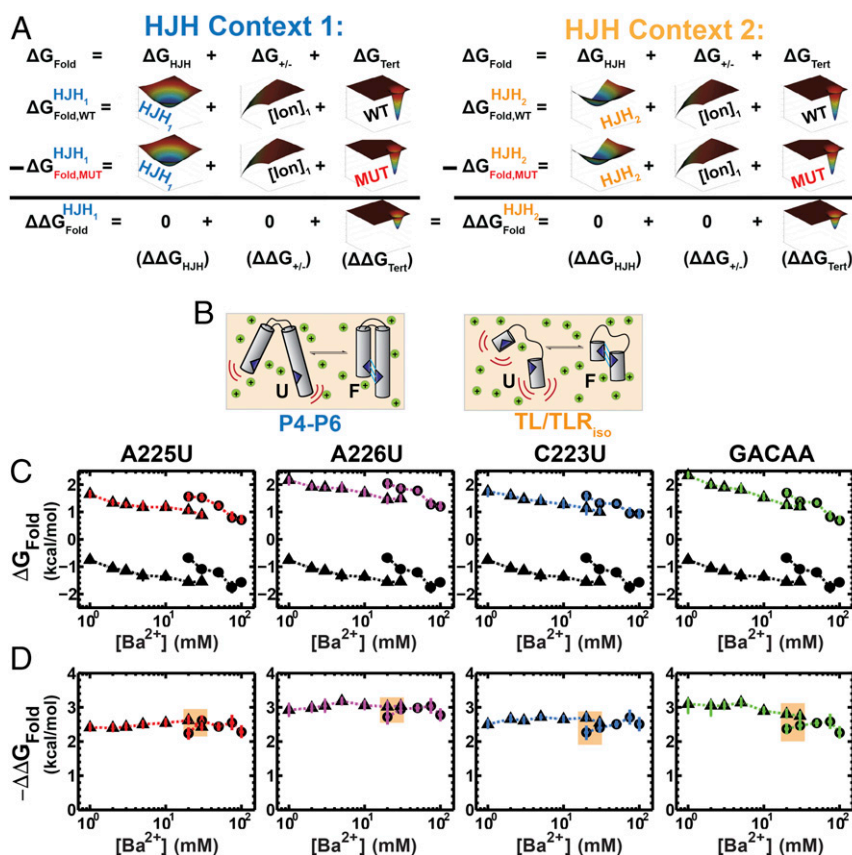


Fig. 4. Comparison of the effect of TLR mutations in P4-P6 RNA and the TL/TLR_{iso} construct. (A) The reconstitution hypothesis predicts that the effects of a mutation in a tertiary contact ($\Delta \Delta G_{\text{Fold}} = \Delta G_{\text{WT}} - \Delta G_{\text{Mut}}$) will be the same in different structural contexts [HJH₁ (Left) vs. HJH₂ (Right)]. Mutation in a tertiary contact leads to a lower probability for formation of the tertiary contact and thus a less negative value of ΔG_{Tert} and weaker overall folding, without affecting the other free energy terms. (B) Depiction of P4-P6 folding (Left) versus the TL/TLR_{iso} (Right) in Ba²⁺. (C) Folding across a range of Ba²⁺ concentrations (1–100 mM) for P4-P6 (circles) and TL/TLR_{iso} (triangles); colors correspond to point mutations in the TL/TLR as in Fig. 3C, and the black points repeated in each panel are for WT P4-P6 and WT TL/TLR_{iso}. Data are provided in Tables S1 and S2. (D) The effect of the mutation relative to WT ($\Delta \Delta G_{\text{Fold,WT-Mut}} = \Delta G_{\text{WT}} - \Delta G_{\text{Mut}}$) versus Ba²⁺ for P4-P6 (circles) and the TL/TLR_{iso} (triangles). Tan overlay designates the overlapping Ba²⁺ concentrations where the $\Delta \Delta G_{\text{Fold,WT-Mut}}$ values can be directly compared. Errors represent the 95% confidence interval of the mean determined by bootstrapping (Methods).

in Fig. 5B. In this way, the $\Delta \Delta G_{\text{Fold,WT-Mut}}$ data for TL/TLR_{iso} serve as a common comparator for $\Delta \Delta G_{\text{Fold,WT-Mut}}$ values between conditions for P4-P6 with and without the MC/MCR, and the shallow dependence of TL/TLR_{iso} folding on metal ion concentration further simplifies this comparison. We therefore plotted the TL/TLR_{iso} $\Delta \Delta G_{\text{Fold,WT-Mut}}$ values from Fig. 4C with the $\Delta \Delta G_{\text{Fold,WT-Mut}}$ values for P4-P6 in Mg²⁺ and Ba²⁺ (Fig. 5C).

For three of the mutants—A225U, A226U, and C223U—there were variations of less than 0.5 kcal/mol between the $\Delta \Delta G_{\text{Fold,WT-Mut}}$ values with and without the MC/MCR, representing differences of less than twofold in equilibrium constant and providing strong support of the predictions from the reconstitution model. Intriguingly, a slightly larger variance of 0.8 kcal/mol was observed for the fourth mutant, GACAA, which is beyond our estimated experimental error (Methods, Fig. 5C, and Tables S1 and S2). Given that this mutant corresponds to an insertion into the GAAA tetraloop to give a GACAA pentaloop, the insertion of the cytosine in the loop may alter the preferred orientation of the loop such that a higher fraction of the conformational states is competent for forming the mutant TL/TLR compared with the WT tetraloop when the MC/MCR is formed; this would cause the insertion mutation to be less deleterious in the most constrained context of P4-P6 with its MC/MCR formed. In principle, comparisons such as those carried out herein can be used to assess differential conformational states of tertiary contacts.

Our results varying ionic conditions and varying the HJH context support the reconstitution hypothesis, as the constant $\Delta \Delta G_{\text{Fold,WT-Mut}}$ values meet the simplest predictions of this model. Our results also suggest the potential for more complex behavior that can arise from the presence of altered states of the tertiary interaction and provide a path to identify such altered states through outlier, nonadditive behavior.

Testing Kinetic Predictions of the Reconstitution Hypothesis. The thermodynamic modularity of the tertiary folding energetics introduces the possibility that there may also be a conservation of the underlying properties that determines the kinetics of tertiary contact formation and dissolution. We can make kinetic predictions based on the diffusion–collision model for RNA (21), whereby folding rate constants are determined by the rate constant at which tertiary elements collide weighted by the probability that the tertiary elements are in a state competent to proceed to form the interactions of the stable contact. To uncover predictions from this model, we break the overall folding process down in this model into three steps (Fig. 6A): (i) local rearrangements of the two tertiary contact elements to form a “tertiary-competent” conformation (K_{Conf}) that has a significant probability of forming the final tertiary contact from an aligned state; (ii) formation of the aligned state from which the tertiary contact can form (K_{Align}), which is determined by the extent to which the intervening helices and junctions favor the aligned versus all other states; and (iii)

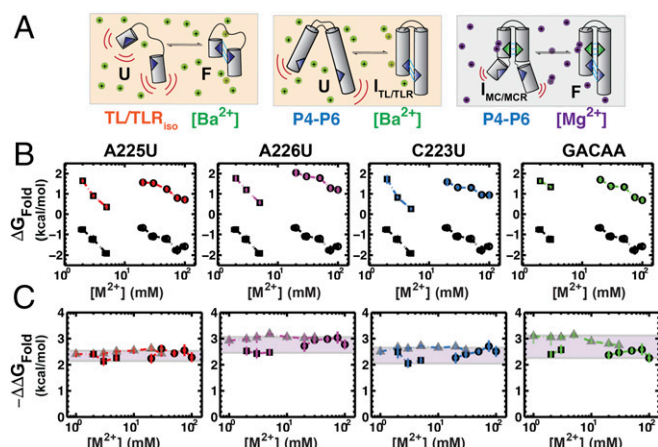


Fig. 5. Effect of HJH changes on the effect of TL/TLR point mutants. (A) Depiction of TL/TLR_{iso} (Left) folding in Ba²⁺, P4-P6 folding in Ba²⁺ (Middle), and P4-P6 folding in Mg²⁺ (Right). In Ba²⁺ only, the TL/TLR forms in P4-P6, whereas both the MC/MCR and the TL/TLR form in Mg²⁺ (38). The formation of the MC/MCR changes the structural connectivity between the TL and TLR and thus the alignment term of the reconstitution hypothesis (ΔG_{align} ; Fig. 2B and Eq. 1). (B) P4-P6 folding across a range of Ba²⁺ concentrations (circles; 20–100 mM) and Mg²⁺ concentrations (squares; 1–3 mM), for point mutations in the TL/TLR; colors are as in Fig. 3. Data are presented in Tables S1 and S2. (C) The effect of each mutation relative to WT ($\Delta\Delta G_{\text{Fold}}$) for P4-P6 in Ba²⁺ (circles) and Mg²⁺ (squares) and for the TL/TLR_{iso} in Ba²⁺ (triangles). Purple overlay designated the variation in the $\Delta\Delta G$ values for each mutant in all three HJH contexts. Errors represent the 95% confidence interval of the mean determined by bootstrapping (Methods).

formation of the interactions, such as hydrogen bonds and stacking interactions, that make up the tertiary contact (k_{Form}). Although in reality there will be a complex landscape of atomic-level rearrangements with many steps and pathways and a range of probabilities that they lead to the final tertiary state, we present a simplified model with three steps that we have evidence to support. In contrast, a two-step model cannot account for the data, as shown below.

RNA tertiary formation, including that for formation of the TL/TLR studied herein, is much slower than expected for diffusional collision; thus, there must be an additional step or steps preceding formation of the contact (21, 27, 48, 49). Although in principle there might only be a need for two steps, tertiary contact rearrangement and alignment (diffusional collision), our results provide evidence for an additional step. Specifically, the initial k_{Conf} rearrangement (Fig. 6A) is not rate limiting; if it were, then changes in the connecting helices and junctions would not alter the overall k_{Fold} for TL/TLR formation, whereas we observe that k_{Fold} is different in different RNAs (see below). Further, an alignment step involving the elements connecting the tertiary contact partners is not rate limiting (k_{Align} , Fig. 6A), because, if it were, mutations that disrupt the tertiary contact would not decrease the overall folding rate (k_{Fold} ; see also below). Thus, a distinct, third step must be involved and rate limiting.

To build a three-step model, we posited that the initial atomic-level rearrangements leading to the “tertiary-competent” conformation (K_{Conf}) would not encompass all of the rearrangements needed to form the full interaction and that additional rearrangements would be required after alignment; these additional rearrangements are encapsulated in k_{Form} (Fig. 6A). As k_{Form} appears to be rate limiting, the initial two steps are expected to equilibrate before the k_{Form} step, so we discuss the kinetics below in terms of equilibria for these steps (K_{Conf} and K_{Align}).

The reconstitution hypothesis and kinetic model of Fig. 6A predict that the rate constants for folding vary as the HJH context is changed, since the alignment probability of tertiary contact partners (K_{Align}) will be different in different RNAs. In contrast, processes

that are inherent to the tertiary interaction (k_{Conf} and k_{Form}) should be the same independent of structural context (i.e., independent of the identity of the helices and junctions in the HJH that link the tertiary contact partners). While we cannot measure the rates of these individual microscopic steps to test whether they are the same, we can test instead the prediction that the effects of tertiary contact mutations on the rate constants ($\Delta\Delta G_{\text{Fold}}^{\ddagger}$) are the same in different contexts. That is, if these individual processes are the same, mutations will have constant effects on the pre-equilibrium of formation of the tertiary-ready state (K_{Conf}) and the rate constant for rearranging to form the tertiary interactions (k_{Form}) and thus have the same effects on overall folding rate constants in different RNAs.

To test this model, we determined the effects on folding rate constants for each of the four tertiary mutations, and we compared the effects in P4-P6 and in TL/TLR_{iso}. Folding rate constants were different between P4-P6 and the TL/TLR_{iso} for all variants, as generally expected, and there were significant effects on k_{Fold} for three of the four mutants (A225U, A226U, and GACAA) of

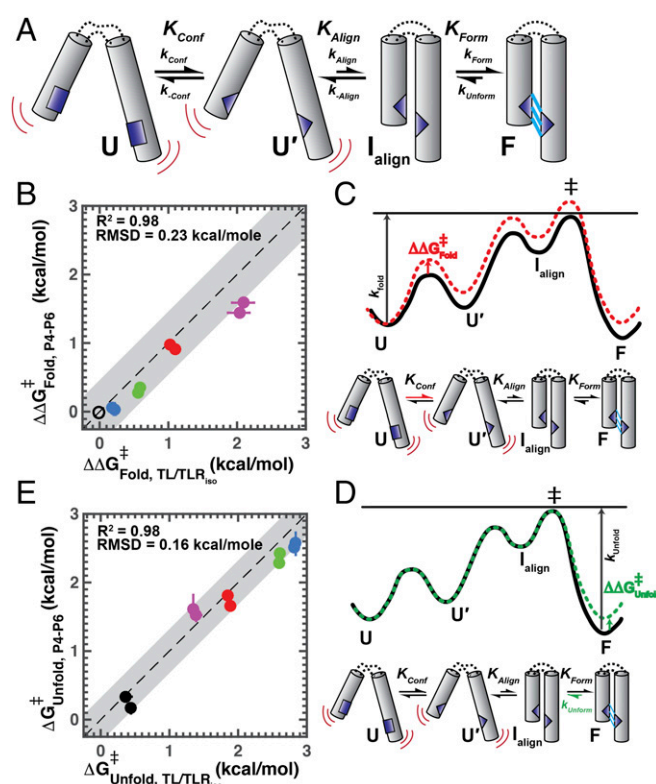


Fig. 6. Kinetic effects of tertiary contact mutations on folding processes. (A) Simple kinetic model for tertiary contact formation based on the diffusion-collision model for RNA folding (21). Folding rates and equilibria for tertiary contact formation are composed of three steps: conformational change of tertiary contact elements to a tertiary-ready state, alignment of tertiary contact elements, and formation of the tertiary interactions between tertiary elements. Comparison of the effects on folding (B) and absolute unfolding (E) rate constants ($\Delta\Delta G_{\text{Fold}}^{\ddagger}$ and $\Delta\Delta G_{\text{Unfold}}^{\ddagger}$; see Methods) for TL/TLR mutations in P4-P6 versus TL/TLR_{iso} (colors as in Fig. 1). Measurements were made at 20 and 30 mM Ba²⁺. Error bars represent propagation of the 95% confidence interval of the mean determined by bootstrapping. Reported is the rmsd to the fit of the data to the model $y = x + b$ (see Methods; data are presented in Tables S1 and S2) and is denoted in gray outline. Free energy diagrams depicting hypothesized effects on overall folding (C) and unfolding (D) rate constants. Effects on k_{Fold} (B) are consistent with effects on the conformational change of tertiary contact partners before formation of the tertiary contact (K_{Conf}), as depicted in the free energy diagram in C. Effects on k_{Unfold} (D) are consistent with tertiary contact breaking as the rate-limiting transition in unfolding (k_{Unform}), as depicted in the free energy diagram in (D).

10–30-fold (Fig. S1A and Tables S1 and S2). The effects on the folding rate constant in each structural context are plotted relative to one another in Fig. 6B in terms of $\Delta\Delta G_{\text{Fold}}^{\ddagger}$ for each (Methods). [These comparisons were made in Ba^{2+} because there are additional folding steps in Mg^{2+} for P4–P6, including rate-limiting formation of the MC/MCR, that would obscure information about the rate of forming the TL/TLR (21).] The congruence of the relationship between these $\Delta\Delta G_{\text{Fold}}^{\ddagger}$ values for TL/TLR_{iso} and P4–P6 is excellent, with an R^2 value of 0.98 and an rmsd of 0.23 kcal/mol for a linear model with the predicted slope of 1, an rmsd value that essentially matches our measurement error (error bars, Fig. 6B). Stated another way, this linear fit suggests that the effects in one context are essentially fully predictive of the effects in the other context. Free energy-reaction profiles illustrate how mutational effect on the tertiary contact conformational change (K_{Conf} ; Fig. 6C) and/or the rate constant for the formation of the tertiary contact (k_{Form} , not depicted) can lead to an overall folding rate constant effect that is independent of the alignment probability (K_{Align}). Fig. 6C shows the free energy diagram depicting this scenario, whereby the conformational change of the receptor (k_{Conf}) is slowed by the mutation, thereby decreasing the overall folding rate constant and adversely affecting the overall folding thermodynamics.

With respect to unfolding rate constants, the reconstitution hypothesis predicts a conservation of the process of tertiary contact dissolution embedded in different structural contexts, as the same physical interactions are being broken regardless of the surrounding helices and junctions. This prediction will hold provided the final step in folding, k_{Form} , is rate limiting (Fig. 6C) and, by microscopic reversibility, the first step in unfolding (Fig. 6D, k_{Unfold}). Fig. 6D shows this scenario in terms of a free energy diagram, whereby the step for dissolution of the tertiary contact (k_{Unfold}) is destabilized by the mutation, leading to an increase in the overall unfolding rate, as this step is rate-limiting in the unfolding pathway.

The unfolding rates for mutant and WT TL/TLRs in the TL/TLR_{iso} and P4–P6 contexts are shown in Fig. S1B, and significant effects on k_{Unfold} of 5–50-fold were observed for each of the four mutations investigated (Tables S1 and S2). As predicted, at equal Ba^{2+} concentrations, the unfolding rate constant for each mutant is the same within error in P4–P6 and in TL/TLR_{iso} (Fig. 6E, k_{Unfold} presented in terms of $\Delta G_{\text{Unfold}}^{\ddagger}$). Thus, the dissociation rate constants highly correlate, with an R^2 value of 0.98 and an rmsd of 0.16 kcal/mol for the model with uniform rate constants in the two contexts.

Discussion

We have formulated and tested predictions from a quantitative thermodynamic and kinetic model for RNA tertiary folding that posits that a complex RNA can be understood from the conformational and energetic behavior of its constituent elements—that is, the helices, junctions, and tertiary motifs that make up the RNA (Fig. 1). This RNA reconstitution hypothesis arises from a general perspective of RNA as modular, which in turn stems from observations that tertiary motifs have similar structures across multiple RNAs, an ability to swap these motifs as modules and maintain function, and conserved conformational properties of junctions embedded in different RNAs (15–17, 19).

Analogously, a longstanding goal within the protein-folding field has been to develop a bottom-up predictive understanding of folding kinetics and thermodynamics. Below we briefly compare and contrast the applicability of additivity models to RNA and proteins, highlighting the expectation that RNA may be more amenable to such approaches.

While hydrogen bonds and individual hydrophobic contacts in protein structures can be readily identified and counted, their energetic behavior is complex. These complexities include the typically coupled formation of secondary and tertiary structure, the differential energetic effects of solvation on buried versus exposed alpha helices and beta sheets (e.g., refs. 50 and 51), and the multiple local interactions and extended interaction networks in which most res-

idues engage. The complications inherent to protein folding are largely absent in RNA folding, due to the simplifying features of its tertiary structure. The stability of base pairs within RNA helices results in preformed RNA secondary structure, allowing, in general, a clean energetic separation of RNA secondary and tertiary structure formation (17). [In cases such as riboswitches, there is coupling between the formation of secondary and tertiary structure, and there may be additional cases where native secondary structure is not stable in the absence of tertiary interactions (e.g., refs. 52 and 53). Nevertheless, these scenarios allow thermodynamic folding predictions to be made via a simple extension of the reconstitution model, using individual energy terms for secondary structure formation (“Turner rules”) and tertiary structure formation (Reconstitution), provided that these individual terms are known to sufficient accuracy.]

Whereas most amino acid side chains make tertiary interactions with multiple additional side chains, most RNA residues are not involved in long-range tertiary interactions. Rather, RNA bases exist in the context of helices and simple junctions that serve as connections between tertiary contacts, and these elements rarely make long-range or tertiary interactions to other parts of the structure. Thus, we might expect that the conformational properties of these helices and junctions to be the same in the structured RNA as when isolated in a much smaller RNA. In this way, the individual conformational preferences of helix and junction elements in isolation can be determined and “added” together to determine the overall conformations of a complex RNA in its unfolded states. Then, the subset of conformational states that allow the tertiary contact to be formed defines the value of ΔG_{Align} for an RNA (Fig. 1B and C). While likely to hold in many cases, the proposed additivity of individual elements may find limitations in contexts where there are more extended networks of interactions such as those found near ribozyme active sites (e.g., ref. 54).

Although the goal of a complete quantitative model has not been attained for protein folding, the theories and formalisms are relevant and in some cases perhaps even more readily applied to studies of RNA folding (e.g., refs. 21, 55, and 56). These formalisms and concepts, such as the diffusion–collision model (57, 58), the contact order model (59), and hydrophobic collapse, have been proposed and tested (40, 60) and have provided useful analogies and led to evaluation of models for RNA folding (21, 55).

Analogous to the studies in RNA presented here, Pace and coworkers (61–63) compared the effects of sequence variants on alpha helix stability in small model peptide systems to the effects of the same variants in proteins with the goal to develop these peptide systems as modular pieces whose behavior could be extrapolated to proteins. Most simply, they asked whether a mutation that affects the propensity of a sequence to form an alpha helix, say decreasing this propensity, destabilizes overall folding of the protein by the same amount. This prediction bore out for mutations in exposed residues on the protein, but this simple rule is not followed for a buried side chain, as desolvation and interactions with neighboring residues affect the energetic contribution of these side chains to folding (51, 64–66). As noted above, this approach and potential modularity may be more applicable to RNA tertiary structures given their limited packing interfaces and the sparse distribution of tertiary contacts throughout a structure.

Indeed, our results support the model that an RNA tertiary motif has the same energetic properties in different contexts. Specifically, we have shown that three out of four point mutations in the TL/TLR have the same energetic effects across different structural contexts and in different ion environments. Most directly, the kinetic and thermodynamic effects of these point mutations can be used to test whether TL/TLR formation is rate limiting in the folding or assembly of structured RNAs, as the rate of overall folding or assembly should be affected by a predictable amount if formation or dissolution of the contact is the rate-determining step in the process.

The principles of the reconstitution hypothesis are readily extendable to account for more complex behaviors such as folding

cooperativity, whereby two contacts are thermodynamically coupled in their formation. When two contacts form cooperatively, like the two tertiary contacts of P4–P6 RNA (67), their successive formation can be accounted for by the reconstitution hypothesis by iteratively calculating the resultant conformational ensembles of the RNA following formation of a contact (represented by a new $\Delta G'_{\text{HH}}$ term). This recalculation will reflect the cooperativity between two contacts, as the conformational ensemble following the formation of a tertiary is more favorable (or less favorable in the case of negative cooperativity) for the next tertiary contact to form. For example, in the case of P4–P6, formation of the MC/MCR contact will predispose the new conformational ensemble ($\Delta G'_{\text{HH}}$) to form the TL/TLR, and vice versa.

Broader characterization of more RNA modules using the approaches used herein will test whether this observed modularity applies to few or many RNA elements and can identify motifs and contexts that do not display energetic additivity. For tertiary motifs that are found to not be energetically modular, these motifs may contain multiple conformational states that could form different interactions depending on orientation or electrostatic requirements of the contact. Moreover, mutations themselves may lead to additional states to be explored, leading to nonadditivity in different contexts. Such a situation is likely for the GACAA mutant in the tetraloop studied herein, which gave a different energetic effect in one of the structural contexts consistent with the larger loop altering the conformational ensemble of the free loop and/or the conformation of the assembled tertiary motif. This result highlights the utility of these quantitative comparisons in uncovering sequence variants and motifs with different behaviors that can be targeted for further structural and energetic studies. Indeed, future work in characterizing motifs will be aided by additivity tests across a broad range of RNA contexts that can provide information about multiple orientations and conformational states. Overall, our results provide a framework to guide future work toward a quantitative and predictive understanding of RNA folding energetics and function.

Methods

Construction of Single-Molecule P4–P6 (smP4–P6) Variants. smP4–P6 variants were prepared using a five-piece splinted ligation protocol, as described previously (47, 67, 68). Variants were prepared with two pieces containing 5-amino-allyl-uridine modification to attach dyes for smFRET (see Table S3 for sequences) and labeled using an NHS-ester reaction (69). All pieces were synthetic RNA and ordered from Integrated DNA Technologies, except for the middle section of the molecule (bases 163–234) due to its length, and were ordered as IDT Ultramer DNAs, PCR amplified, transcribed, and purified. To obtain homogenous 3' ends following transcription, a DNase complementary to the 3' end of the transcript was used during transcription (70) with a polynucleotide kinase treatment following to remove the 2',3'-cyclic phosphate left from DNase cleavage. The ligation reaction included all five RNA pieces, added at ~0.4 μM . The final ligated product was purified using a denaturing polyacrylamide gel [8% PAGE, 8 M urea, 1 \times TBE: 89 mM Tris (pH 7.4), 89 mM Boric Acid, 2 mM EDTA]. Ligated bands corresponding to the full-length ligated molecule were determined by scanning the gel with a Typhoon 9400 (GE Healthcare) at the Cy3 and Cy5 excitation wavelengths. Full-length molecules were gel purified and eluted using three freeze–thaw cycles and evaluated for homogeneity in folding and length on a native (8% acrylamide, 10 mM Mg^{2+}) and denaturing polyacrylamide gel (8% PAGE, 8 M urea, 1 \times TBE).

Preparation of Isolated TL/TLR. RNAs containing a 3' primary amine were ordered from IDT for each WT and mutant TL/TLR_{iso} (Table S3). The con-

structs were labeled with Cy3b/Cy5 with an NHS-ester reaction (69) and gel purified.

SmFRET Experiments. smFRET experiments were performed on a custom-built total internal reflection (TIRF) microscope at room temperature (23 °C). Molecules were excited with diode-pumped solid-state green laser (532 nm; Gem, laser quantum) and their fluorescence captured with an Andor Ixon Ultra Camera with an CAIRN Optosplit II to separate donor and acceptor fluorescence emissions. Before imaging, molecules were annealed to a DNA oligonucleotide containing a 5' biotin (5' biotin AACCAAAATCAACCTAAACTTACACA-3') at 50 °C for 15 min, cooled at 0.1 °C/min to room temperature, and diluted to 10–100 pM for imaging on quartz slides derivatized with biotin-BSA (1 mg/mL; Sigma-Aldrich) and streptavidin (0.1 mg/mL; Sigma-Aldrich). To stabilize dyes from bleaching and blinking, a protocatechuic acid/protocatechuate-3,4-dioxygenase (PCA/PCD) (Sigma-Aldrich) oxygen scavenging system with trolox (6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid; Sigma-Aldrich) was used. Buffers used in smFRET measurements contained 1–5 mM MgCl_2 or 20–100 mM BaCl_2 , 50 mM potassium-Mops, pH 8.0, 100 mM KCl, 2 mM PCA, 0.01 units per mL PCD, and 1 mM trolox. smFRET data were collected at an acquisition rate between 50 and 330 frames per second and analyzed using the SMART data analysis package (71).

Data Analysis. Folding rate and equilibrium constants were inferred from smFRET traces by fitting donor and acceptor intensities using a HMM-based algorithm to a two-state model with a single unfolded and single folded state (21, 71). This analysis extracts transition probabilities between each state while taking into account the noise observed in each intensity channel. Criteria for trace selection were described previously (72); briefly, traces were selected based on exhibiting all of the following: (i) single-step photobleaching, (ii) anticorrelated donor and acceptor channels, (iii) total intensity consistent with a single molecule, and (iv) stable total intensity.

All reported rate and equilibrium constants are the mean of the log distribution of the values for each molecule. Reported errors are the 95% confidence intervals of the mean from bootstrapping parameters from individual molecules. The bootstrapped confidence intervals are asymmetric, so the greater of the two values is reported in the main text for simplicity (Tables S1 and S2). Both errors are reported in Dataset S1. Measurement error was determined to be the average of all confidence intervals for all measurements, which was 0.2 kcal/mol. A cutoff for signal-to-noise (SNR) of 0.75 is imposed for all P4–P6 smFRET data and 0.5 for the isolated TL/TLR, reflecting differences in the change in FRET observed for the two different molecules.

Defining $\Delta\Delta G^{\ddagger}_{\text{Fold}}$. The effect of each mutation on the folding and unfolding rate constant is defined as:

$$\Delta\Delta G^{\ddagger}_{\text{Fold}} = -RT \ln \left[\frac{k_{\text{Mut Fold}}}{k_{\text{WT Fold}}} \right]$$

$$\Delta G^{\ddagger}_{\text{Unfold}} = RT \ln [k_{\text{Unfold}}].$$

The relationships between the rate constants effects between P4–P6 and TL/TLR_{iso} in Fig. 6 were fit to a linear model where the slope is constrained to $m = 1$ (i.e., the model $y = x + b$). The goodness of these fits was determined by the rmsd of the orthogonal distance of the data to the line $y = x$. The coefficient of determination for a linear model (R^2) was computed by the square of a Pearson correlation coefficient for two independent variables:

$$\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

ACKNOWLEDGMENTS. We thank members of the D.H. laboratory for discussion and critical advice throughout this work. This work was supported by the NIH Grant P01GM066275. N.B. was supported by the NSF graduate fellowship. M.G. was supported in part by NIH Training Grant GM008412 in biotechnology.

- Cech TR, Steitz JA (2014) The noncoding RNA revolution-trashing old rules to forge new ones. *Cell* 157:77–94.
- Moore MJ (2005) From birth to death: The complex lives of eukaryotic mRNAs. *Science* 309:1514–1518.
- Staley JP, Guthrie C (1998) Mechanical devices of the spliceosome: Motors, clocks, springs, and things. *Cell* 92:315–326.
- Noller HF (2005) RNA structure: Reading the ribosome. *Science* 309:1508–1514.
- Aitken CE, Puglisi JD (2010) Following the intersubunit conformation of the ribosome during translation in real time. *Nat Struct Mol Biol* 17:793–800.

- Zhang X, Rashid R, Wang K, Shan S-O (2010) Sequential checkpoints govern substrate selection during cotranslational protein targeting. *Science* 328:757–760.
- Parks JW, Stone MD (2014) Coordinated DNA dynamics during the human telomerase catalytic cycle. *Nat Commun* 5:4146.
- Hoskins AA, et al. (2011) Ordered and dynamic assembly of single spliceosomes. *Science* 331:1289–1295.
- Turner DH, Sugimoto N, Freier SM (1988) RNA structure prediction. *Annu Rev Biophys Chem* 17:167–192.

10. Tinoco I, Jr, et al. (1973) Improved estimation of secondary structure in ribonucleic acids. *Nat New Biol* 246:40–41.
11. Anderson CF, Record MT, Jr (1995) Salt-nucleic acid interactions. *Annu Rev Phys Chem* 46:657–700.
12. Xia T, et al. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* 37:14719–14735.
13. Heigwer F, Kerr G, Boutros M (2014) E-CRISPR: Fast CRISPR target site identification. *Nat Methods* 11:122–123.
14. Staley JP, Guthrie C (1999) An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p. *Mol Cell* 3:55–64.
15. Klein DJ, Schmeing TM, Moore PB, Steitz TA (2001) The kink-turn: A new RNA secondary structure motif. *EMBO J* 20:4214–4221.
16. Wu L, Chai D, Fraser ME, Zimmerly S (2012) Structural variation and uniformity among tetraloop-receptor interactions and other loop-helix interactions in RNA crystal structures. *PLoS One* 7:e49225.
17. Tinoco I, Jr, Bustamante C (1999) How RNA folds. *J Mol Biol* 293:271–281.
18. Leontis NB, Westhof E (2003) Analysis of RNA motifs. *Curr Opin Struct Biol* 13: 300–308.
19. Thomson JB, Lilley DM (1999) The influence of junction conformation on RNA cleavage by the hairpin ribozyme in its natural junction form. *RNA* 5:180–187.
20. Nissen P, Ippolito JA, Ban N, Moore PB, Steitz TA (2001) RNA tertiary interactions in the large ribosomal subunit: The A-minor motif. *Proc Natl Acad Sci USA* 98:4899–4903.
21. Bisaria N, et al. (2016) Kinetic and thermodynamic framework for P4-P6 RNA reveals tertiary motif modularity and modulation of the folding preferred pathway. *Proc Natl Acad Sci USA* 113:E4956–E4965.
22. Lilley DM (2000) Structures of helical junctions in nucleic acids. *Q Rev Biophys* 33: 109–159.
23. Tyagi R, Mathews DH (2007) Predicting helical coaxial stacking in RNA multibranch loops. *RNA* 13:939–951.
24. Murchie AI, Thomson JB, Walter F, Lilley DM (1998) Folding of the hairpin ribozyme in its natural conformation achieves close physical proximity of the loops. *Mol Cell* 1: 873–881.
25. Bailor MH, Sun X, Al-Hashimi HM (2010) Topology links RNA secondary structure with global conformation, dynamics, and adaptation. *Science* 327:202–206.
26. Chu VB, et al. (2009) Do conformational biases of simple helical junctions influence RNA folding stability and specificity? *RNA* 15:2195–2205.
27. Herschlag D, Allred BE, Gowrishankar S (2015) From static to dynamic: The need for structural ensembles and a predictive model of RNA folding and function. *Curr Opin Struct Biol* 30:125–133.
28. Al-Hashimi HM, Walter NG (2008) RNA dynamics: It is about time. *Curr Opin Struct Biol* 18:321–329.
29. Hohng S, et al. (2004) Conformational flexibility of four-way junctions in RNA. *J Mol Biol* 336:69–79.
30. Shi X, Huang L, Lilley DMJ, Harbury PB, Herschlag D (2016) The solution structural ensembles of RNA kink-turn motifs and their protein complexes. *Nat Chem Biol* 12:146–152.
31. Salmon L, Bascom G, Andricioaei I, Al-Hashimi HM (2013) A general method for constructing atomic-resolution RNA ensembles using NMR residual dipolar couplings: The basis for interhelical motions revealed. *J Am Chem Soc* 135:5457–5466.
32. Wu M, Tinoco I, Jr (1998) RNA folding causes secondary structure rearrangement. *Proc Natl Acad Sci USA* 95:11555–11560.
33. Klostermeier D, Millar DP (2002) Energetics of hydrogen bond networks in RNA: Hydrogen bonds surrounding G+1 and U42 are the major determinants for the tertiary structure stability of the hairpin ribozyme. *Biochemistry* 41:14095–14102.
34. Silverman SK, Cech TR (1999) Energetics and cooperativity of tertiary hydrogen bonds in RNA structure. *Biochemistry* 38:8691–8702.
35. Doherty EA, Doudna JA (1997) The P4-P6 domain directs higher order folding of the Tetrahymena ribozyme core. *Biochemistry* 36:3159–3169.
36. Cate JH, et al. (1996) Crystal structure of a group I ribozyme domain: Principles of RNA packing. *Science* 273:1678–1685.
37. Das R, Travers KJ, Bai Y, Herschlag D (2005) Determining the Mg²⁺ stoichiometry for folding an RNA metal ion core. *J Am Chem Soc* 127:8272–8273.
38. Takamoto K, et al. (2004) Principles of RNA compaction: Insights from the equilibrium folding pathway of the P4-P6 RNA domain in monovalent cations. *J Mol Biol* 343:1195–1206.
39. Silverman SK, Zheng M, Wu M, Tinoco I, Jr, Cech TR (1999) Quantifying the energetic interplay of RNA tertiary and secondary structure interactions. *RNA* 5:1665–1674.
40. Cate JH, Hanna RL, Doudna JA (1997) A magnesium ion core at the heart of a ribozyme domain. *Nat Struct Biol* 4:553–558.
41. Murphy FL, Cech TR (1994) GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J Mol Biol* 236:49–63.
42. Butcher SE, Dieckmann T, Feigon J (1997) Solution structure of a GAAA tetraloop receptor RNA. *EMBO J* 16:7490–7499.
43. Costa M, Michel F (1997) Rules for RNA recognition of GNRA tetraloops deduced by in vitro selection: Comparison with in vivo evolution. *EMBO J* 16:3289–3302.
44. Travers KJ, Boyd N, Herschlag D (2007) Low specificity of metal ion binding in the metal ion core of a folded RNA. *RNA* 13:1205–1213.
45. Gebala M, Bonilla S, Bisaria N, Herschlag D (2016) Does cation size affect occupancy and electrostatic screening of the nucleic acid ion atmosphere? *J Am Chem Soc* 138: 10925–10934.
46. Hodak JH, Downey CD, Fiore JL, Pardi A, Nesbitt DJ (2005) Docking kinetics and equilibrium of a GAAA tetraloop-receptor motif probed by single-molecule FRET. *Proc Natl Acad Sci USA* 102:10505–10510.
47. Greenfeld M, Herschlag D (2010) Measuring the energetic coupling of tertiary contacts in RNA folding using single molecule fluorescence resonance energy transfer. *Methods Enzymol* 472:205–220.
48. Bartley LE, Zhuang X, Das R, Chu S, Herschlag D (2003) Exploration of the transition state for tertiary structure formation between an RNA helix and a large structured RNA. *J Mol Biol* 328:1011–1026.
49. Downey CD, et al. (2006) Metal ion dependence, thermodynamics, and kinetics for intramolecular docking of a GAAA tetraloop and receptor connected by a flexible linker. *Biochemistry* 45:3664–3673.
50. Serrano L, Sancho J, Hirshberg M, Fersht AR (1992) Alpha-helix stability in proteins. I. Empirical correlations concerning substitution of side-chains at the N and C-caps and the replacement of alanine by glycine or serine at solvent-exposed surfaces. *J Mol Biol* 227:544–559.
51. Minor DL, Jr, Kim PS (1994) Context is a major determinant of beta-sheet propensity. *Nature* 371:264–267.
52. Strulson CA, Boyer JA, Whitman EE, Bevilacqua PC (2014) Molecular crowders and cosolutes promote folding cooperativity of RNA under physiological ionic conditions. *RNA* 20:331–347.
53. Chauhan S, Woodson SA (2008) Tertiary interactions determine the accuracy of RNA folding. *J Am Chem Soc* 130:1296–1303.
54. Forconi M, Sengupta RN, Piccirilli JA, Herschlag D (2010) A rearrangement of the guanosine-binding site establishes an extended network of functional interactions in the Tetrahymena group I ribozyme active site. *Biochemistry* 49:2753–2762.
55. Sosnick TR, Pan T (2004) Reduced contact order and RNA folding rates. *J Mol Biol* 342: 1359–1365.
56. Woodson SA (2002) Folding mechanisms of group I ribozymes: Role of stability and contact order. *Biochem Soc Trans* 30:1166–1169.
57. Karplus M, Weaver DL (1976) Protein-folding dynamics. *Nature* 260:404–406.
58. Karplus M, Weaver DL (1994) Protein folding dynamics: The diffusion-collision model and experimental data. *Protein Sci* 3:650–668.
59. Plaxco KW, Simons KT, Baker D (1998) Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol* 277:985–994.
60. Russell R, Herschlag D (1999) New pathways in folding of the Tetrahymena group I RNA enzyme. *J Mol Biol* 291:1155–1167.
61. Pace CN, Scholtz JM (1998) A helix propensity scale based on experimental studies of peptides and proteins. *Biophys J* 75:422–427.
62. Myers JK, Pace CN, Scholtz JM (1997) A direct comparison of helix propensity in proteins and peptides. *Proc Natl Acad Sci USA* 94:2833–2837.
63. Myers JK, Smith JS, Pace CN, Scholtz JM (1996) The alpha-helix of ribonuclease T1 as an independent stability unit: Direct comparison of peptide and protein stability. *J Mol Biol* 263:390–395.
64. Gao J, Bosco DA, Powers ET, Kelly JW (2009) Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins. *Nat Struct Mol Biol* 16:684–690.
65. Albeck S, Unger R, Schreiber G (2000) Evaluation of direct and cooperative contributions towards the strength of buried hydrogen bonds and salt bridges. *J Mol Biol* 298:503–520.
66. Baase WA, Liu L, Tronrud DE, Matthews BW (2010) Lessons from the lysozyme of phage T4. *Protein Sci* 19:631–641.
67. Sattin BD, Zhao W, Travers K, Chu S, Herschlag D (2008) Direct measurement of tertiary contact cooperativity in RNA folding. *J Am Chem Soc* 130:6085–6087.
68. Solomatin SV, Greenfeld M, Herschlag D (2011) Implications of molecular heterogeneity for the cooperativity of biological macromolecules. *Nat Struct Mol Biol* 18:732–734.
69. Greenfeld M, Herschlag D (2013) Fluorescently labeling synthetic RNAs. *Methods Enzymol* 530:281–297.
70. Schubert S, et al. (2003) RNA cleaving '10-23' DNAs with enhanced stability and activity. *Nucleic Acids Res* 31:5982–5992.
71. Greenfeld M, Pavlichin DS, Mabuchi H, Herschlag D (2012) Single Molecule Analysis Research Tool (SMART): An integrated approach for analyzing single molecule data. *PLoS One* 7:e30024.
72. Shi X, et al. (2014) Roles of long-range tertiary interactions in limiting dynamics of the Tetrahymena group I ribozyme. *J Am Chem Soc* 136:6643–6648.
73. Marcia M, Pyle AM (2012) Visualizing group II intron catalysis through the stages of splicing. *Cell* 151:497–507.