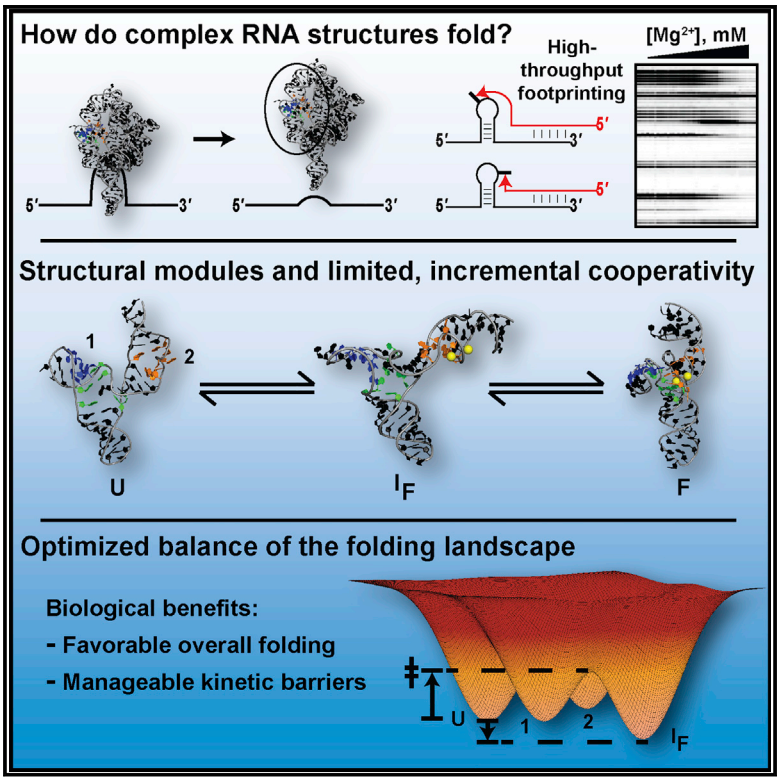
Article

Cell Reports

Hidden Structural Modules in a Cooperative RNA Folding Transition

Graphical Abstract



Authors

Brant Gracia, Hashim M. Al-Hashimi, Namita Bisaria, Rhiju Das, Daniel Herschlag, Rick Russell

Correspondence

rick_russell@cm.utexas.edu

In Brief

Gracia et al. dissect a complex RNA folding transition with changes in secondary and tertiary structure. They show that modular structures fold and interact with limited, incremental cooperativity. This cooperativity is sufficient to give a concerted folding transition and may balance competing biological demands on RNA stability and folding efficiency.

Highlights

- A group I intron domain links changes in secondary and tertiary structure
- Modular structural elements underlie this complex RNA folding transition
- These structural elements fold and interact with limited, incremental cooperativity
- Modular folding optimizes favorable thermodynamics and kinetics

Gracia et al., 2018, Cell Reports 22, 3240–3250
March 20, 2018 © 2018 The Authors.
https://doi.org/10.1016/j.celrep.2018.02.101

CellPress

OPEN ACCESS
CellPress

Cell Reports
Article

Hidden Structural Modules in a Cooperative RNA Folding Transition

Brant Gracia,[1] Hashim M. Al-Hashimi,[2,3] Namita Bisaria,[4,5] Rhiju Das,[4] Daniel Herschlag,[4] and Rick Russell[1,6,*]
[1]Department of Molecular Biosciences and the Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, TX 78712, USA
[2]Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA
[3]Department of Chemistry, Duke University Medical Center, Durham, NC 27710, USA
[4]Department of Biochemistry, Stanford University, Stanford, CA 94305, USA
[5]Present address: Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA
[6]Lead Contact
*Correspondence: rick_russell@cm.utexas.edu
https://doi.org/10.1016/j.celrep.2018.02.101

SUMMARY

Large-scale, cooperative rearrangements underlie the functions of RNA in RNA-protein machines and gene regulation. To understand how such rearrangements are orchestrated, we used high-throughput chemical footprinting to dissect a seemingly concerted rearrangement in P5abc RNA, a paradigm of RNA folding studies. With mutations that systematically disrupt or restore putative structural elements, we found that this transition reflects local folding of structural modules, with modest and incremental cooperativity that results in concerted behavior. First, two distant secondary structure changes are coupled through a bridging three-way junction and Mg^{2+}-dependent tertiary structure. Second, long-range contacts are formed between modules, resulting in additional cooperativity. Given the sparseness of RNA tertiary contacts after secondary structure formation, we expect that modular folding and incremental cooperativity are generally important for specifying functional structures while also providing productive kinetic paths to these structures. Additionally, we expect our approach to be useful for uncovering modularity in other complex RNAs.

INTRODUCTION

Structured RNAs function in myriad cellular processes, including pre-mRNA splicing, protein production, maintenance of chromosome ends, and regulation of gene expression, and many of these RNAs are currently under investigation as therapeutic targets (Bhattacharyya et al., 2015; Burnett and Rossi, 2012; Hille and Charpentier, 2016; Kole et al., 2012; Ng et al., 2006; Yanofsky, 1981). As structural information about functional RNAs has grown, it has become apparent that RNA/protein machines such as the spliceosome, ribosome, and telomerase undergo orchestrated conformational transitions that enable them to switch between functional states as they carry out their reactions (Akiyama et al., 2013; Chen and Moore, 2014; Galej et al., 2014; Parks and Stone, 2017; Yusupova and Yusupov, 2014; Zhu and Meyer, 2015). Large-scale, coupled rearrangements are also key to riboswitch signaling, as binding of a ligand results in an alternative structure and a biological outcome through changes in transcription, translation, or splicing (Mironov et al., 2002; Serganov and Nudler, 2013; Winkler et al., 2002a, 2002b, 2003).

A necessary element in these RNA conformational steps is that changes in multiple regions of the structure are coupled together, allowing regions of the structured RNA to sense and respond to the conformations of other regions through cooperativity. For example, translocation by the ribosome involves large coordinated movements of the large and small subunits that define the stages of each forward step in protein synthesis (Chen et al., 2013; Petrov et al., 2012), and in riboswitches structural changes in the aptamer platform upon ligand binding must be communicated to the expression platform to exert the downstream impact. The same general phenomenon underlies native RNA folding, as cooperativity is necessary for a single global conformation or small ensemble of conformations—the native state—to be favored over the exceedingly large number of partially folded and misfolded intermediates that differ in tertiary structure and/or secondary structure.

Although functional and structural studies have provided valuable descriptions of conformational steps involved with the functions of RNAs and RNA-protein complexes, the mechanisms that drive these folding processes remain poorly understood. An illustrative example comes from the P5abc subdomain of the *Tetrahymena thermophila* group I intron, which undergoes a complex conformational switch upon the addition of Mg^{2+} ion (Figure 1). Initial studies showed that this switch includes changes in the secondary and tertiary structure of the RNA, and subsequent work implied an extremely high degree of cooperativity in the transition (Koculi et al., 2012; Silverman et al., 1999; Wu and Tinoco, 1998; Zheng et al., 2001).

These observations provided valuable descriptions of a complex folding transition but did not reveal the physical basis for the complex rearrangement or show how the changes in secondary and tertiary structure are linked energetically. Despite the

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

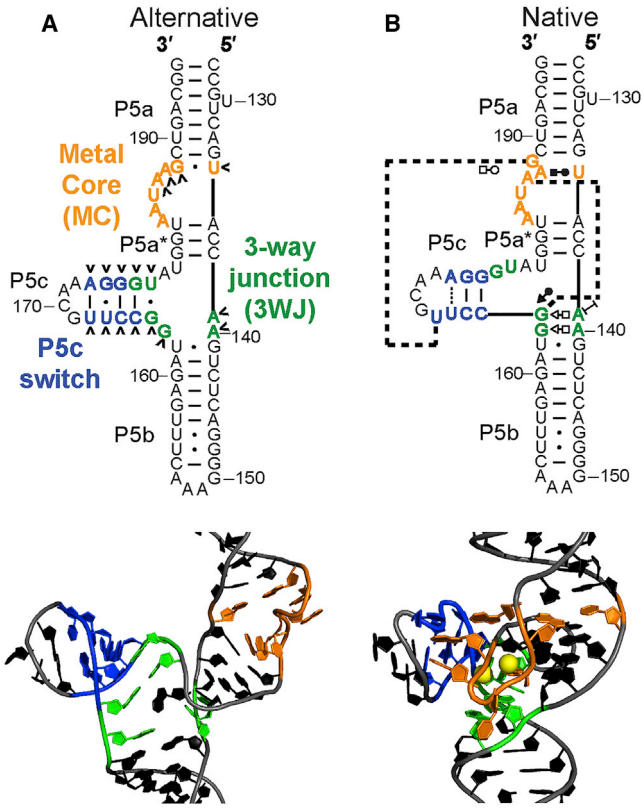


Figure 1. The Complex Folding Transition of P5abc RNA

(A) Alternative structure of P5abc, which represents the unfolded state for folding studies (Cate et al., 1997; Wu and Tinoco, 1998). Nucleotides that change base pairing upon Mg^{2+}-induced folding are marked (′) and regions are colored (P5c, blue; 3WJ, green; MC, orange). A structural model is below the secondary structure (Wu and Tinoco, 1998).

(B) Native P5abc. Long-range base pairs are indicated with dashed lines. Non-canonical base pairs are shown here and elsewhere using Leontis/Westhof nomenclature (Leontis and Westhof, 2001). In the tertiary structure model, two site-bound Mg^{2+} ions are yellow (Das et al., 2005; Frederiksen et al., 2012).

observation of a concerted transition, we wondered whether this complex transition might be mediated by structural modules within P5abc that had remained hidden in previous studies. Supporting the presence of modules, our recent work revealed that P5abc can change some of the secondary structure transiently, without the distal changes in secondary and tertiary structure (Xue et al., 2016), and that it can form at least some of the native tertiary structure without the full complement of secondary structure changes (Bisaria et al., 2016; Gracia et al., 2016).

Here, we introduce a mutation and rescue strategy to test directly for the presence of structural modules in P5abc and to measure cooperativity between them, conceptually analogous to the mutation and rescue cycle that has been instrumental in identifying RNA secondary structures in complex RNAs and ribonucleoproteins (RNPs) (Huang et al., 2013; Nilsen, 1994; Staley and Guthrie, 1998; Tian et al., 2014; Wu and Huang, 1992). We systematically knocked out putative structural modules and then rescued them using mutagenesis, and then we used high-throughput footprinting methods to measure the impacts of these changes on folding of the remaining modules. Using this strategy, we uncovered structural modules in P5abc that had previously been obscured by their folding cooperativity. Furthermore, we measured the cooperativity and established its origins, revealing that the cooperativity emerges in modest increments, from the formation of short-range and long-range contacts, through two folding steps. Our work suggests general roles for modularity in RNA folding and highlights a general strategy to dissect complex RNA folding transitions.

RESULTS

Mg^{2+}-Dependent Folding of P5abc Occurs in Multiple Transitions

P5abc is composed of three helical elements (P5a/P5b/P5c) that converge at a three-way junction (3WJ) (Figure 1) (Lescoute and Westhof, 2006; Wu and Tinoco, 1998). In P5abc folding, two Mg^{2+} ions bind specifically to sites within an A-rich bulge, resulting in a 270° corkscrew in the RNA backbone termed the metal core (MC) (Figure 1B) (Cate et al., 1997; Correll et al., 2003; Frederiksen et al., 2012; Murphy and Cech, 1993; Pabit et al., 2013). The formation of tertiary structure in the MC is accompanied by a change in the base pairing of 17 nucleotides, including a shift in the register of P5c, the formation of non-canonical base pairs in the 3WJ, and a shift in the register of a base pair adjacent to the MC.

To monitor equilibrium folding, we incubated P5abc with various concentrations of Mg^{2+} (0.02–200 mM) and performed quantitative selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) and dimethyl sulfate (DMS) footprinting (Figure 2A) (Cordero et al., 2014; Mortimer and Weeks, 2007; Tijerina et al., 2007). We detected modifications as blockages to reverse transcription and normalized the intensities of each peak to those of reference sequences that flanked P5abc in our RNA construct (Figures 2B and 2C). In contrast to the simplest expectation from early results (Koculi et al., 2012; Silverman et al., 1999; Wu and Tinoco, 1998; Zheng et al., 2001), the Mg^{2+} dependences of modification were complex, with a range of Mg^{2+} midpoints and some nucleotides with multiple phases (Figures 3A and 3B; Tables S1 and S2). These data indicate that P5abc folding is not a two-state process. To derive a minimal model, we fit all of the footprinting data with a global model using KinTek Global Kinetic Explorer (Supplemental Experimental Procedures) (Johnson et al., 2009). The global fitting gave a minimal model with four discrete states and three Mg^{2+}-dependent transitions between them (Figures S1 and S2; Table S3).

At sub-millimolar concentrations of Mg^{2+}, several nucleotides in the 3WJ and the MC displayed modest enhancements in reactivity with increasing Mg^{2+} concentration (Figure 3A, U to U*; Figures S3A–S3C). These enhancements were also observed when Mg^{2+} was replaced with Ba^{2+}, which does not support MC formation (Figure S3D) (Travers et al., 2007), and in the presence of 1 M NaCl (Figure S3E) (Das et al., 2005). These results suggest that the transition reflects alleviation of electrostatic repulsion (Russell et al., 2000; Takamoto et al., 2004) and is not associated with secondary or tertiary structure changes in P5c or the MC. Therefore, it is not discussed further, and U and U* are

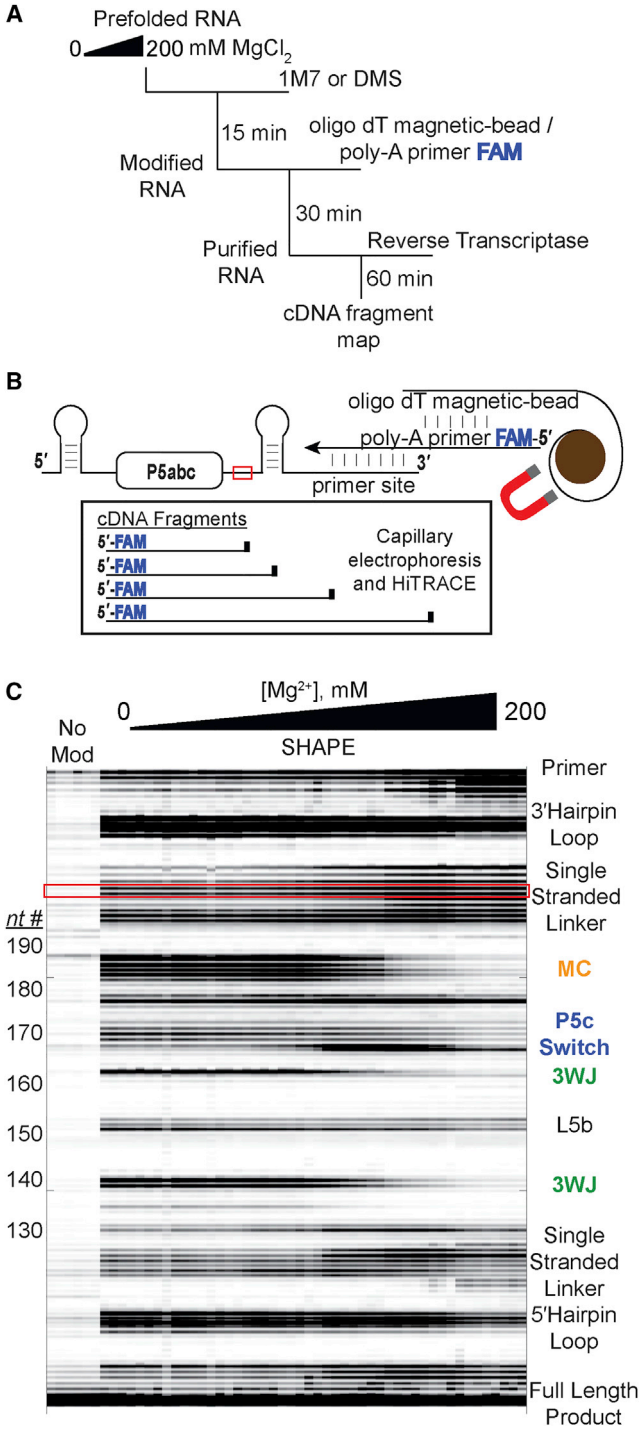


Figure 2. High-Throughput Footprinting to Monitor P5abc Folding

(A) Steps in the high-throughput footprinting approach. The FAM indicates that the primer is fluorescently labeled with fluorescein.

(B) P5abc footprinting cassette. The RNA includes flanking hairpins linked by single-stranded segments used for internal normalization (red box).

(C) Electrophoretic traces of reverse transcription products aligned and quantified by HiTRACE (Yoon et al., 2011). The area under each peak is converted to grayscale and represented as a band (Yoon et al., 2011). The red box indicates the nucleotides used for internal normalization. Eight control reactions without SHAPE reagent are shown at the left (No Mod) and span the Mg^{2+} concentrations used (0–200 mM Mg^{2+}). The nucleotide number (nt #) in the P5abc region is shown on the left edge.

collectively termed "U/Alt" to reflect that they contain the alternative secondary structure and lack stable tertiary interactions.

As the Mg^{2+} concentration was further increased, a second transition was observed, with protections in several regions of P5abc (Figures 3A and 3B, U/Alt to I_F). Nucleotides within the MC and the 3WJ became protected from the SHAPE reagent (Figures 3C and 3D), indicating structure formation, and large enhancements at key positions within P5c indicated its switch to native secondary structure (Figure 3E). These changes occurred with indistinguishable Mg^{2+} dependences and were cooperative with Mg^{2+} concentration ($n = 2.1 \pm 0.1$), as expected for a transition linked to cooperative binding of two Mg^{2+} ions in the MC (Das et al., 2005), and they did not occur in this concentration range when folding was carried out in Ba^{2+} (Figure S3D). We conclude that in the second transition, the MC binds Mg^{2+} specifically and forms tertiary structure, P5c switches to native secondary structure, and the 3WJ forms tertiary structure.

At even higher Mg^{2+} concentrations, a third transition was observed, by both SHAPE and DMS footprinting (Figure 3A, I_F to F). These changes occurred at nucleotides within loop L5c, the 3WJ, and the MC (Figure 3B, dashed boxes, and Figures 3E and 3F; Table S1 and S2). The distributed nature of these changes and their positions suggest that the transition involves the packing of native P5c into a groove formed by the MC and 3WJ to give the conformation observed in crystal structures of P4–P6 (Cate et al., 1996, 1997) and the ribozyme core (Golden et al., 1998) (Figure 1B).

P5c and the MC Can Fold in Isolation and Are Reinforced with Cooperativity

The coincident formation of native P5c secondary structure and MC tertiary structure in the transition from U/Alt to I_F provided support for the previous conclusion that these structural elements fold cooperatively (Koculi et al., 2012; Silverman et al., 1999; Wu and Tinoco, 1998; Zheng et al., 2001). To determine whether each element can fold by itself and to measure the cooperativity, we used mutagenesis to block one of these structural elements from folding, and then we measured folding of the non-mutated element by SHAPE footprinting. A P5abc variant that stabilizes the alternative secondary structure of P5c (U167C) (Gracia et al., 2016; Silverman et al., 1999; Xue et al., 2016) displayed clear protections that indicated formation of the MC, but with an increased $[Mg^{2+}]_{1/2}$ value of 9.8 mM (Figures 4A, 4B, S4A, and S4B; Table S1). The increased Mg^{2+} requirement indicates that cooperativity between the native P5c secondary structure and the MC has been lost in the mutant.

Two additional results indicate that measurements for the U167C mutant reveal the full extent of cooperativity (i.e., that P5c remains in the alternative secondary structure when the MC forms in the mutant at the high Mg^{2+} concentrations). First, there was no measurable Mg^{2+}-dependent transition in P5c up to 50 mM Mg^{2+} (Figure S4B). Second, a double mutant in P5c (U167C/U177C) that further stabilizes the alternative P5c

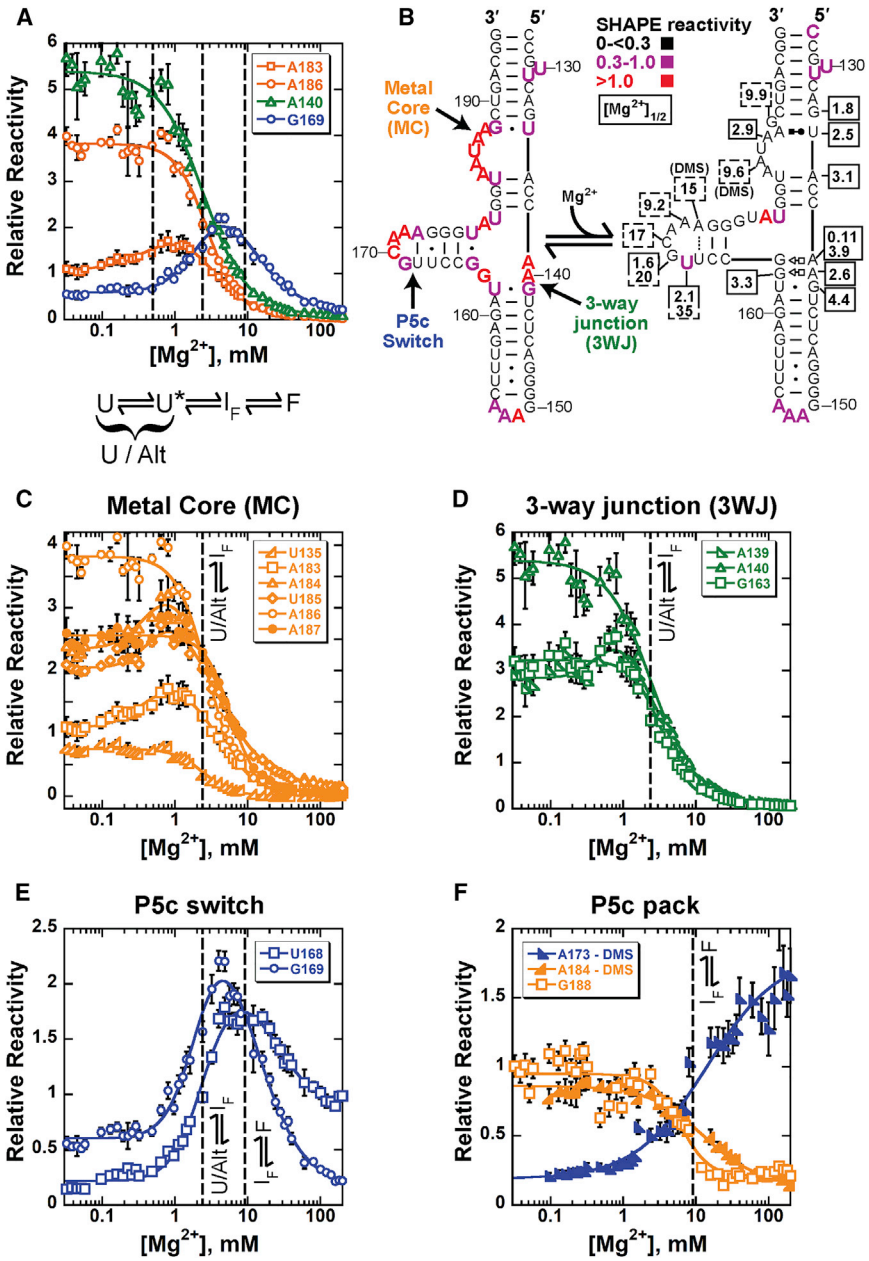


Figure 3. Native Structure Formation by P5abc Occurs in Two Distinct Transitions

(A) Mg^{2+} dependences of SHAPE reactivity for representative nucleotides to illustrate the three transitions. Transitions are indicated with dashed lines and correspond to the minimal model shown below the plot. The coloring is the same as in Figure 1.

(B) Normalized SHAPE reactivity values in the absence of Mg^{2+} (U/Alt state) and the presence of saturating Mg^{2+} (F state). Representative Mg^{2+} midpoints are shown in boxes. Boxes are dashed for the higher Mg^{2+} midpoints, indicating a change in reactivity in the transition from I_F to F. All midpoints are from SHAPE experiments except for A173 and A184, which are from DMS experiments. A complete list of Mg^{2+} midpoints is in Tables S1 and S2.

(C–E) Mg^{2+} dependences of nucleotides that undergo changes in the U/Alt to I_F transition within the MC (C), the three-way junction (3WJ) (D), and P5c (E).

(F) Mg^{2+} dependences of nucleotides that undergo changes in the I_F to F transition.

Here and in subsequent figures, error bars reflect the uncertainty values associated with peak fitting of the raw chromatograms. See also Figures S1–S3 and Table S3.

secondary structure (Gracia et al., 2016) gave the same Mg^{2+} dependence for MC nucleotides as the single mutant (Figures 4A and 4B). Together, these results show that MC formation and P5c secondary structure switching are thermodynamically coupled but that this coupling is sufficiently modest for folding of the MC to be readily observed when P5c is prevented from switching to its native secondary structure.

The observed coupling between P5c switching and MC formation predicts that if MC formation were blocked by mutation, the P5c helix could still switch to the native state, and this transition might be favorable with a higher Mg^{2+} concentration. We therefore carried out this converse experiment, blocking MC formation by replacing A186 with U or the entire A-rich bulge with uridines (Figure 4C) (Das et al., 2005; Murphy and Cech, 1994; Sattin et al., 2008). Indeed, we observed enhancements within P5c by SHAPE footprinting, indicating native secondary structure formation (Figures 4D, S4C, and S4D). Also as expected, the Mg^{2+} requirement was increased, from a $[Mg^{2+}]_{1/2}$ value of 2.1 for wild-type P5abc to 5.4 mM (Table S1), and the lack of a signal at MC nucleotides in these mutants showed that MC structure formation was indeed disrupted (Figure S4D). Thus, P5c can rearrange to form its native secondary structure without MC formation.

To evaluate the cooperativity quantitatively, we focused on footprinting data at 3 mM Mg^{2+}. We chose this Mg^{2+} concentration because it is close enough to the midpoints of folding transitions for the wild-type P5abc and module-knockout mutants that the equilibrium values can be determined directly from the Mg^{2+} dependences, without requiring an extrapolation. Thus, it was possible to determine the level of cooperativity between native folding of P5c and the MC by directly comparing the equilibrium values for one module in the presence or the absence of the other module. This analysis shows that the coupling between P5c and the MC in this transition is modest, with a value of 1.3 kcal/mol (Figure 4E). Equivalent analyses performed from data at 2 or 4 mM Mg^{2+} concentrations gave similar results (1.4 and 1.6 kcal/mol, respectively), indicating that the level of

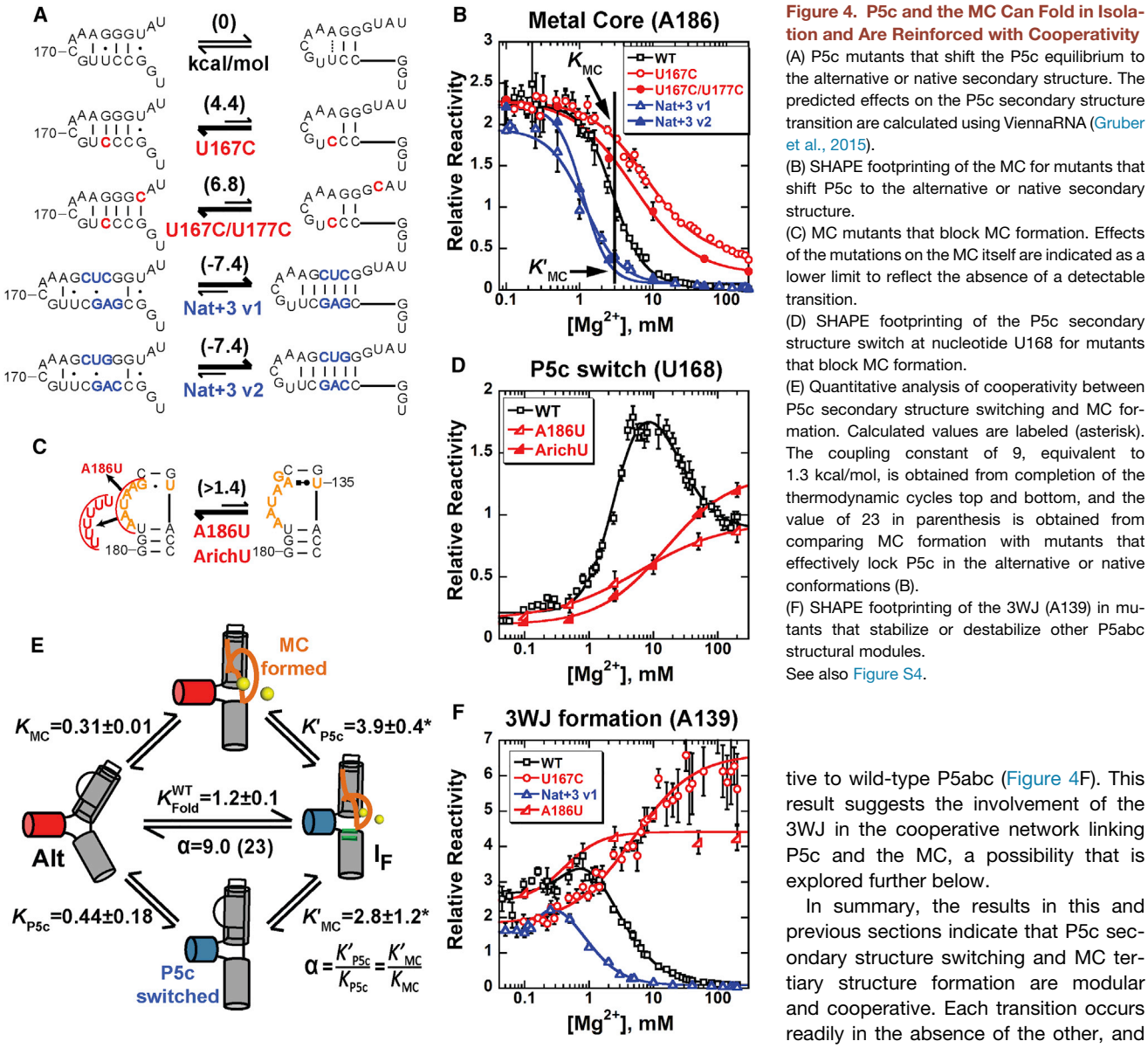


Figure 4. P5c and the MC Can Fold in Isolation and Are Reinforced with Cooperativity

(A) P5c mutants that shift the P5c equilibrium to the alternative or native secondary structure. The predicted effects on the P5c secondary structure transition are calculated using ViennaRNA (Gruber et al., 2015).

(B) SHAPE footprinting of the MC for mutants that shift P5c to the alternative or native secondary structure.

(C) MC mutants that block MC formation. Effects of the mutations on the MC itself are indicated as a lower limit to reflect the absence of a detectable transition.

(D) SHAPE footprinting of the P5c secondary structure switch at nucleotide U168 for mutants that block MC formation.

(E) Quantitative analysis of cooperativity between P5c secondary structure switching and MC formation. Calculated values are labeled (asterisk). The coupling constant of 9, equivalent to 1.3 kcal/mol, is obtained from completion of the thermodynamic cycles top and bottom, and the value of 23 in parenthesis is obtained from comparing MC formation with mutants that effectively lock P5c in the alternative or native conformations (B).

(F) SHAPE footprinting of the 3WJ (A139) in mutants that stabilize or destabilize other P5abc structural modules.

See also Figure S4.

cooperativity between the folding transitions in P5c and the MC does not depend strongly on the Mg^{2+} concentration in this range.

A strong prediction of our results is that a P5abc mutant that favors the native P5c secondary structure would decrease the Mg^{2+} requirement for MC formation, because the wild-type P5abc must "pay" to switch the P5c secondary structure, whereas the mutant would already have native P5c formed (Figure 4E, bottom right equilibrium K'_{MC}). Indeed, a mutant that stabilizes native P5c by inserting three additional Watson-Crick base pairs (Nat+3) displayed a reduced Mg^{2+} requirement for folding the MC (Figures 4A, 4B, and S4E). The equilibrium value at 3 mM Mg^{2+} was 7 ± 1, even a bit larger than the predicted value of 3 ± 1. Interestingly, protections were also observed in the 3WJ, with a reduced Mg^{2+} requirement for this mutant relative to wild-type P5abc (Figure 4F). This result suggests the involvement of the 3WJ in the cooperative network linking P5c and the MC, a possibility that is explored further below.

In summary, the results in this and previous sections indicate that P5c secondary structure switching and MC tertiary structure formation are modular and cooperative. Each transition occurs readily in the absence of the other, and the two transitions are energetically linked, such that each transition is more favorable in the presence of the other, with modest cooperativity comparable with the formation of one or two hydrogen bonds in many systems (Jolley and Znosko, 2017).

A Module-Level Restoration Approach to Uncover the Mechanism of Cooperativity

P5c and the MC are connected by the 3WJ, suggesting that the 3WJ could function as a conduit for transmitting information on the folding of P5c to the MC and vice versa. Furthermore, three nucleotides that are liberated in the transition to native P5c form internal 3WJ contacts and contacts with the MC (Figure 5A).

To test the role of the 3WJ in cooperative folding of P5c and the MC, we took the general strategy of structure restoration mutagenesis a step further. Conceptually analogous to the

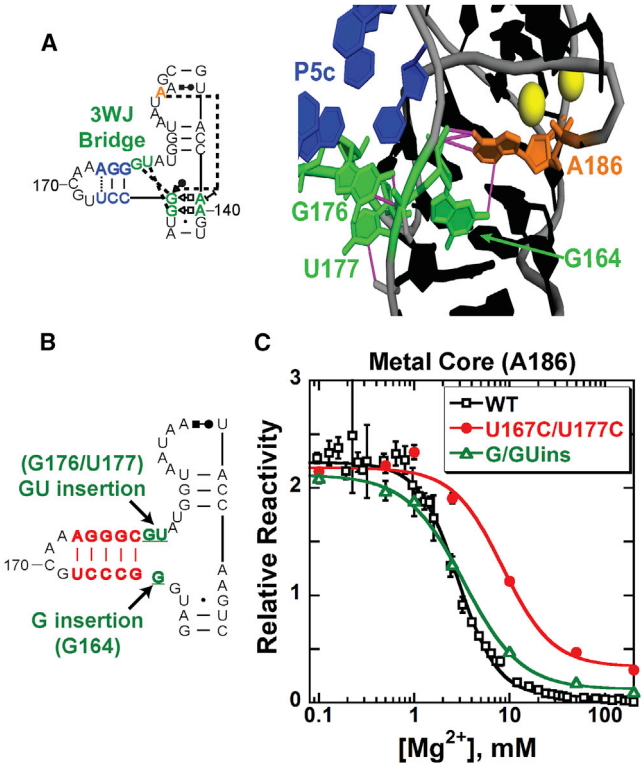


Figure 5. Module-Level Restoration Uncovers the Mechanism of Cooperativity

(A) Contacts of the 3WJ with the MC (A186; orange) and nucleotides liberated when P5c forms native secondary structure (G164, G176, and U177; green). Contacts are indicated by dashed lines in the secondary structure (left) and by magenta lines in the three-dimensional structure image (right).

(B) Schematic of a mutant in which contact-forming nucleotides in the 3WJ are inserted (equivalents of G164, G176, and U177). In this construct, the alternative secondary structure of P5c is stabilized with the substitutions U167C/U177C (alternative P5c base pairs shown in red) to eliminate potential complications from the secondary structure change in P5c. Insertion of each side of the 3WJ individually (G164 or G176/U177) was not sufficient to restore folding to the level of the wild-type P5abc (Figure S5B).

(C) SHAPE footprinting of the mutant shown in (B) (green). Analogous data are shown for wild-type P5abc (black) and the mutant in which P5c secondary structure switching is blocked, but the 3WJ contacts are not restored (red). See also Figure S5.

construct in which native P5c is stable without the need for MC formation, we generated a mutant in which the native 3WJ can form without requiring the native P5c secondary structure by inserting equivalents of these three nucleotides (Figure 5B). In the background of mutations ensuring that P5c remained in the alternative secondary structure (U167C/U177C), we found that the Mg^{2+} requirement for MC and 3WJ formation was restored to the wild-type level for this 3WJ restoration mutant (Figures 5C and S5). Thus, the MC and 3WJ are energetically coupled, which is not surprising, as they contact each other in the folded structure, and the cooperativity between P5c secondary structure switching and MC tertiary structure formation is mediated at least in part through the intervening 3WJ.

A Second Layer of Cooperativity Reinforces the Folded Modules

In the fully folded structure, P5c packs against the 3WJ, forming long-range contacts with the 3WJ and the MC that provide another possible source of cooperativity (Figure 6A) (Wang et al., 2014). Indeed, this packing transition is expected to be cooperative with native structure because a long-range base pair between P5c and the MC (U168-G188) is not possible without native structure, as both nucleotides form alternative local contacts (Figure 1).

To better understand this folding transition, we first tested the role of this base pair by mutating U168 (Koculi et al., 2012) and measuring P5abc folding (Figure 6B). As predicted by the model, the lower Mg^{2+} transition that results in MC formation and P5c secondary structure (U/Alt to I_F) remained intact, but the P5c packing transition at higher Mg^{2+} concentrations (I_F to F) was not detected (Figures 6C and S6). The high Mg^{2+} transition was also absent for the Nat+3 mutant, which blocks the U168-G188 base pair by extending the P5c helix (Figure 6C). Thus, this tertiary folding transition depends on native structure in P5c and the folded MC, providing experimental support for additional cooperativity between these two structural elements.

We next quantitated the cooperativity between P5c and the folded MC in the final folding transition using a module-knockout strategy analogous to that used above to measure cooperativity in the transition from U/Alt to I_F. Thus, we ablated MC folding with the A186U substitution and followed the P5c packing transition against the 3WJ. To eliminate interference from effects of MC formation on P5c secondary structure switching and to ensure that U168 is available for interactions, we used a variant that locks P5c into the native secondary structure (Figure 6D). We found that when MC formation was ablated, signals for P5c packing were also perturbed, indicating cooperativity between MC formation and P5c packing (Figure 6E). Nevertheless, we observed a partial transition at high Mg^{2+} concentrations, which gave a coupling value of ~2.6 kcal/mol between P5c packing and MC formation. Thus, native P5c and the folded MC are mutually stabilized by the long-range contacts that form between them.

DISCUSSION

Cooperativity is a hallmark of biological macromolecules, enabling attainment of a specific functional structure or set of conformations in preference to the enormous number of alternative structures (Creighton, 1995, 1996). Here we developed and used a module-level restoration approach to uncover structural modules in P5abc that had previously remained obscured by their folding cooperativity. In addition, we quantified the cooperativity between the modules and revealed the physical origins of this cooperativity (Figures 7A and S7). Our deconstruction of this complex switch into its components suggests that it may be possible to understand other complex RNA rearrangements quantitatively from the properties of discrete structural modules, with potential utility for understanding biological RNAs and for designing new RNAs.

In one transition, tertiary folding of the MC is linked to a secondary structure change in P5c via structure formed in the

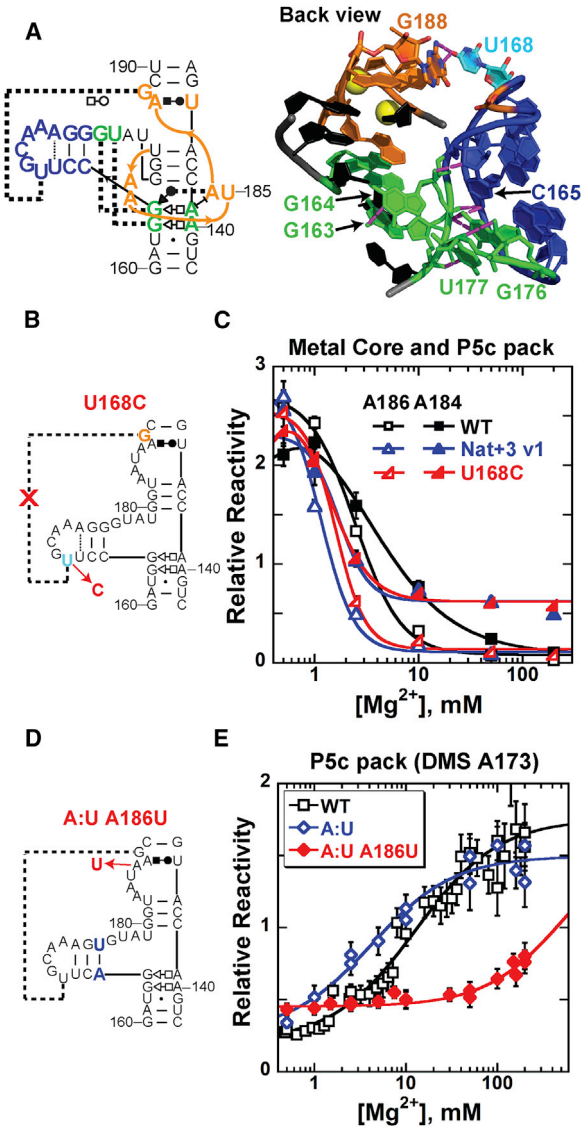


Figure 6. A Second Layer of Cooperativity Reinforces the Folded Modules

(A) Contacts of P5c with the 3WJ and the MC in the native, "packed" conformation. Contacts are shown as dashed lines in the secondary structure diagram (left) and as magenta lines in the three-dimensional structure image (right).

(B) Substitution of U168 to C is expected to perturb the long-range non-canonical base pair of U168 with G188, weakening the packing transition of P5c.

(C) SHAPE footprinting of mutants that weaken or knock out the P5c packing transition. Results are shown for U168C (B) and for the nat+3v1 mutant, which stabilizes the P5c native secondary structure but eliminates the long-range interaction with the MC by extending the P5c helix (Figure 4A). MC formation occurs for both mutants, as reflected by protection of A186 (open symbols) and partial protection of A184 (closed symbols), and P5c packing is detected only for the wild-type P5abc RNA, as reflected in a further protection of A184.

(D) Construct designed to enforce the native P5c secondary structure (C165A/G175U, designated A:U) and knock out MC formation (A186U).

(E) DMS footprinting of the P5abc mutant shown in (D). P5c packing with the 3WJ is monitored by enhanced DMS reactivity of A173 (Supplemental Experimental Procedures).

See also Figure S6.

intervening 3WJ. This linkage provides ~1.3 kcal/mol of cooperativity. The coincident folding of the two structural modules highlights that even modest cooperativity can cause structural rearrangements to appear strictly concerted. In a separate transition, the folded MC and P5c are mutually stabilized by tertiary packing of P5c against the folded MC and 3WJ to provide an additional ~2.6 kcal/mol of cooperativity. The modest cooperativity generated in each step of P5abc folding results in the substantial overall cooperativity of ~4 kcal/mol.

The incremental, stepwise cooperativity we observe here may be a general feature of structured RNAs. Overall, the tertiary packing of structured RNAs is relatively sparse (Vicens and Cech, 2006), such that individual helical segments are extensively solvent exposed and interact with other RNA elements through discrete, modular tertiary contacts. Cooperativity can arise between elements of native secondary and tertiary structure through these discrete contacts, as we have observed here, and it can also arise between tertiary contacts that are distant from each other because of effects on positioning and overcoming electrostatic repulsion (Sattin et al., 2008). Through these mechanisms, there is a general expectation that cooperativity will increase incrementally through the progressive formation of native secondary and tertiary structure.

Cooperativity serves a critical role for the thermodynamics of the folded state by ensuring that partially structured intermediates do not dominate the population. Indeed, a simulated folding landscape for the first P5abc transition in which the free energy for folding of the modules is preserved but the observed cooperativity is removed shows that there would be very little accumulation of the structure with both modules folded even under conditions that permit significant folding of each module individually (Figure 7B). At the other extreme, a very high level of cooperativity would also be detrimental because it would hinder the kinetics of folding. This point is illustrated in a comparison of the free energy surface for the wild-type P5abc (Figure 7C) with a hypothetical surface with higher cooperativity and the same overall free energy of P5abc folding (Figure 7D). In the high-cooperativity landscape, each structural module is unstable individually, so there is very little accumulation of on-pathway intermediates in which one module is formed. Rarer still are intermediates in which both modules are formed but not yet fully reinforced by cooperativity, and as a consequence overall folding is slow (Figure 7D, purple arrow). In contrast, the modular behavior of P5abc, with P5c switching to its native secondary structure without the folded MC and the MC forming its tertiary structure without P5c switching should enhance the kinetics of native P5abc formation by allowing transient formation of these on-pathway intermediates (Figure 7C) (Davis et al., 2016). Indeed, nuclear magnetic resonance (NMR) measurements showed directly that P5c rapidly and reversibly samples the native secondary structure without the folded MC, providing an early folding intermediate on a viable kinetic route to the folded state of P5abc (Xue et al., 2016). These competing effects of cooperativity on the thermodynamics and kinetics of folding lead to the general expectation of modular behavior in biological systems, with limited cooperativity.

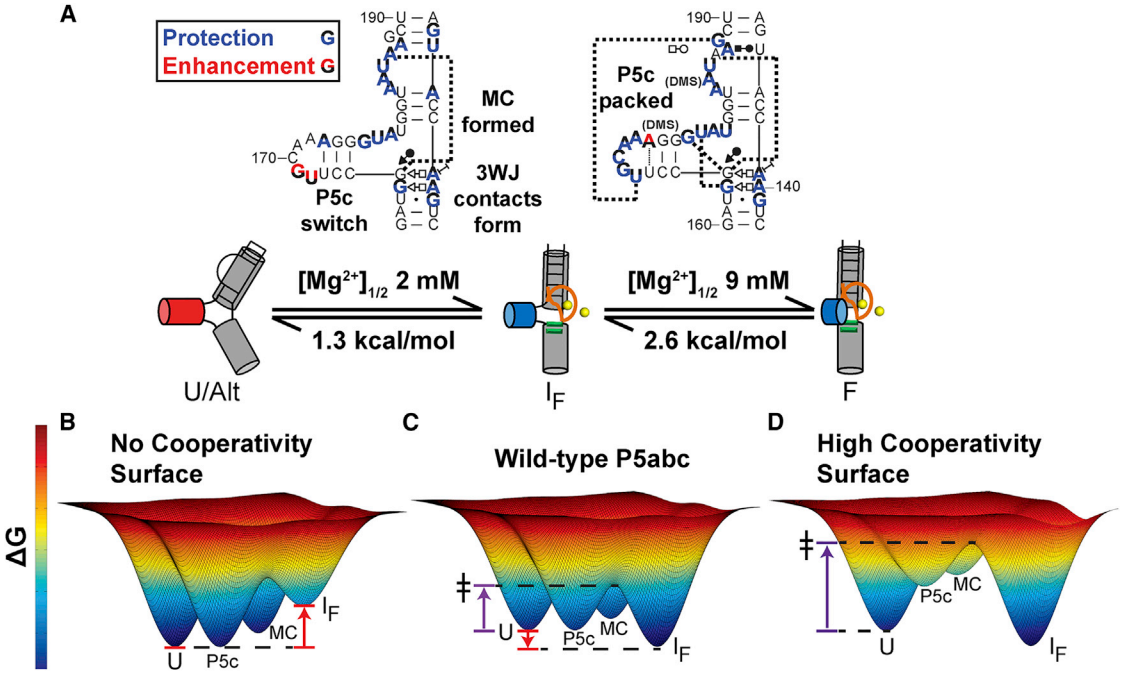


Figure 7. Cooperativity within P5abc Promotes Thermodynamics and Kinetics of Folding

(A) Compiled SHAPE and DMS footprinting signals associated with each transition are indicated with colored letters. Assignments of reactivity changes to one or both transitions are based on global fitting of the footprinting data. The secondary structures illustrate the base pairs formed when the corresponding transition is completed. Cartoon RNA structures are shown as cylinders, with the Mg^{2+} midpoints and cooperativity values shown above and below the arrows, respectively. Cooperativity values are calculated under conditions of 3 mM Mg^{2+}.

(B) A simulated free energy landscape illustrating the poor folding of P5abc (5 mM Mg^{2+}) in a hypothetical landscape in which the free energy of each module is preserved but cooperativity is absent. The red arrow indicates the unfavorable free energy change for the transition from the U/Alt state to the intermediate I_F in which P5c and the MC are in the native states.

(C) Free energy landscape for folding of wild-type P5abc (5 mM Mg^{2+}). The red arrow indicates the favorable free energy change for folding, and the purple arrow indicates the kinetic barrier for folding.

(D) A simulated free energy landscape illustrating the increased transition state barrier (purple arrow) for folding of a hypothetical P5abc RNA (5 mM Mg^{2+}) in which the overall free energy of folding is unchanged from the wild-type P5abc but the cooperativity between structural modules is increased.

See also Figure S7.

Modular structure may also be favored because of its evolvability. In an evolving macromolecule, each new structural element is unlikely to be retained unless it confers a selective advantage on its own, favoring the existence of simple modules that can fold independently. For RNAs, the stability of structural modules means that point mutations outside of the module will typically not disrupt folding, while fortuitous mutations at key positions within or between modules can strengthen global folding or create alternative folded states by establishing progressive layers of cooperative interactions between modules. We are still in the early stages of understanding the complex thermodynamic and kinetic properties of RNA tertiary structure, including its intrinsic features that enhance and limit its cooperativity and its ability to specify conformations and conformational transitions.

EXPERIMENTAL PROCEDURES

RNA Preparation

DNA templates were prepared by assembly PCR with four partially overlapping oligonucleotides (Integrated DNA Technologies). Primer extension reactions were carried out using Phusion High-Fidelity DNA Polymerase (New England Bio-Labs) followed by purification of full-length, double-stranded DNA templates by Magnetic Bead PCR Clean-up (Axygen). PCR-amplified DNA template (0.2 μM) was used in transcription reactions with T7 RNA polymerase, as previously described (Cordero et al., 2014). RNA transcripts were isolated by affinity column (QIAGEN).

Design of P5abc Construct for Footprinting

We used RNA constructs in which P5abc is flanked on both sides by a reference sequence for internal normalization (Figure 2B). The RNA also includes a 3′ primer site. The labeled FAM primer used for reverse transcription includes a 5′ poly-A tail and a 3′ RNA binding site, bridging the RNA to an oligo-dT magnetic bead for purification. After correcting for signal saturation, overmodification, background, and normalization to the internal control (Kladwang et al., 2014), the 5′ and 3′ hairpin loops were comparably reactive at all Mg^{2+} concentrations probed, and independent datasets were highly reproducible, supporting the accuracy of the quantitative analysis. Signals were normalized to the single-stranded linkers such that a reactivity value of 1 indicates that a given nucleotide has the same reactivity as nucleotides within the 3′ single-stranded linker. The results were unchanged when the data were normalized to the 3′ hairpin loop or loop L5b within P5abc.

SHAPE and DMS Footprinting

P5abc RNA (200 nM) was pre-folded in 19 μL at the desired concentration of Mg^{2+} for 15 min at 37°C in a background of 100 mM KCl and 50 mM

K-MOPS (pH 8.0) and then equilibrated at 25°C for 10 min. Longer incubation times up to 120 min did not change the results, suggesting that 10 min is sufficient to reach equilibrium. RNA was chemically modified by adding 1 μL of 1m7 SHAPE reagent (0.42 mg/mL, 5% DMSO final) or DMS (10 mM DMS, 5% DMSO, 0.45% ethanol final) and incubating for 15 min. We found that 5% DMSO decreased the Mg^{2+} requirement for P5c secondary structure switching (Lee et al., 2013) and MC folding, so DMS footprinting reactions were supplemented with 5% DMSO to achieve conditions matching those of the SHAPE footprinting experiments (Table S2) (Lee et al., 2013). The presence of 0.45% ethanol (in the DMS footprinting reactions) had no effect on the SHAPE footprinting profile of wild-type P5abc and was not included in most SHAPE footprinting experiments. A fluorescently labeled DNA primer (FAM) with a 5′ poly-A tail was annealed to oligo-dT magnetic beads (Poly-A Purist kit; Life Technologies), resulting in a 3′ overhang on the DNA primer that is fully complementary to the 3′ end of the RNA. To purify modified RNA, 10 μL of the primer-bead mixture was incubated with RNA for 10 min, washed twice with 200 μL of 70% ethanol, and left to dry at room temperature for 20 min. RNA was reverse transcribed with 5 μL of Superscript III reverse transcriptase mix (Life Technologies) at 42°C for 60 min, followed by RNA degradation with 0.2 M NaOH at 90°C for 3 min. The solution was neutralized with 0.3 M HCl, 0.4 M Na-acetate (pH 4.5), and 700 mM NaCl. cDNAs were purified by oligo-dT magnetic beads and ethanol washed, as described above, for analysis by capillary electrophoresis (ABI 3730).

Data Processing and Statistical Analyses

Footprinting data in the form of electrophoretic traces were aligned and processed using the established HiTRACE pipeline (Kladwang et al., 2014; Yoon et al., 2011). Peaks were quantified by fitting to a sum of Gaussian distributions. Relative uncertainties were generated from the sum of the squares of SDs upon shifting all peak positions by ±0.5 of the mean peak-to-peak spacing. The presence of a transition in Mg^{2+} titrations was assessed by two criteria. First, only nucleotides that undergo a change in reactivity of at least 0.15 reactivity units were considered. Second, only reactivity changes that exceeded the average of the uncertainties of data points within a given transition were considered. Positions that gave significant blocks to reverse transcription in the absence of a footprinting reagent were not analyzed. Changes in reactivity were fit to a sigmoidal binding equation to extract midpoints, amplitudes, and Hill coefficients. Uncertainties reported for folding parameters reflect the SEM from replicate measurements.

Quantitation of Cooperativity between Structural Modules

Cooperativity between native structure formation in the modules P5c and the MC was determined by comparing the equilibrium constants for native folding of one module in P5abc variants that either permit or block folding of the second module. Specifically, we used two methods. In one method, we used the thermodynamic cycle shown in Figure 4E and measurements of the equilibrium constant for the concerted transition in the wild-type P5abc, the MC transition in the mutant that locks P5c in the alternative conformation (U167C), and formation of native P5c secondary structure in the mutant that disables the MC (A186U). This information allows calculation of the remaining equilibrium constants and the degree of cooperativity (Figure 4E; 1.3 kcal/mol, α = 9). In the second method, we compared the equilibrium constants for MC formation with P5c pre-formed in the native conformation (Nat+3) or locked in the alternative conformation (U167C) (Figure 4B). This method yielded a similar, albeit somewhat larger value (Figure 4E; 1.8 kcal/mol, α = 23). Equilibrium constants for each P5abc variant were determined directly from the extent of the change in the footprinting signal at select Mg^{2+} concentrations in the transition range relative to the total change in the transition. With conservative detection limits of 4% in either conformational state, we estimate that the largest coupling value we could measure with this method is approximately 2 kcal/mol without requiring an extrapolation from a lower or higher Mg^{2+} concentration. Consequently, we were able to measure the extent of cooperativity between P5c and the MC without any extrapolation for the transition at lower Mg^{2+} concentration (1.3 kcal/mol) but not for the transition at higher Mg^{2+} concentration (2.6 kcal/mol).

DATA AND SOFTWARE AVAILABILITY

Footprinting data have been deposited in Mendeley, and the dataset DOI is 10.17632/msnb8c6xbf.2.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and three tables and can be found with this article online at https://doi.org/10.1016/j.celrep.2018.02.101.

ACKNOWLEDGMENTS

We thank Portia Saccone, Charlie Gonzales, and Cecil Harkey of the DNA Sequencing Facility at the University of Texas at Austin for processing samples and running capillary electrophoresis with rapid turnaround times. We thank Wipapat Ann Kladwang for teaching us the high-throughput footprinting methodology analyzed by capillary electrophoresis. This work was supported by National Institutes of General Medical Sciences grant P01 GM066275.

AUTHOR CONTRIBUTIONS

Conceptualization, B.G., H.M.A.-H., N.B., R.D., D.H., and R.R.; Methodology, B.G., R.D., D.H., and R.R.; Investigation, B.G.; Writing – Original Draft, B.G. and R.R.; Writing – Review & Editing, B.G., H.M.A.-H., N.B., R.D., D.H., and R.R.; Funding Acquisition, H.M.A.-H., R.D., D.H., and R.R.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 25, 2017
Revised: February 5, 2018
Accepted: February 26, 2018
Published: March 20, 2018

REFERENCES

Akiyama, B.M., Gomez, A., and Stone, M.D. (2013). A conserved motif in Tetrahymena thermophila telomerase reverse transcriptase is proximal to the RNA template and is essential for boundary definition. J. Biol. Chem. *288*, 22141–22149.

Bhattacharyya, D., Diamond, P., and Basu, S. (2015). An Independently folding RNA G-quadruplex domain directly recruits the 40S ribosomal subunit. Biochemistry *54*, 1879–1885.

Bisaria, N., Greenfeld, M., Limouse, C., Pavlichin, D.S., Mabuchi, H., and Herschlag, D. (2016). Kinetic and thermodynamic framework for P4-P6 RNA reveals tertiary motif modularity and modulation of the folding preferred pathway. Proc. Natl. Acad. Sci. U S A *113*, E4956–E4965.

Burnett, J.C., and Rossi, J.J. (2012). RNA-based therapeutics: current progress and future prospects. Chem. Biol. *19*, 60–71.

Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R., and Doudna, J.A. (1996). Crystal structure of a group I ribozyme domain: principles of RNA packing. Science *273*, 1678–1685.

Cate, J.H., Hanna, R.L., and Doudna, J.A. (1997). A magnesium ion core at the heart of a ribozyme domain. Nat. Struct. Biol. *4*, 553–558.

Chen, W., and Moore, M.J. (2014). The spliceosome: disorder and dynamics defined. Curr. Opin. Struct. Biol. *24*, 141–149.

Chen, J., Petrov, A., Tsai, A., O'Leary, S.E., and Puglisi, J.D. (2013). Coordinated conformational and compositional dynamics drive ribosome translocation. Nat. Struct. Mol. Biol. *20*, 718–727.

Cordero, P., Kladwang, W., VanLang, C.C., and Das, R. (2014). The mutate-and-map protocol for inferring base pairs in structured RNA. Methods Mol. Biol. *1086*, 53–77.

Correll, C.C., Beneken, J., Plantinga, M.J., Lubbers, M., and Chan, Y.L. (2003). The common and the distinctive features of the bulged-G motif based on a 1.04 A resolution RNA structure. Nucleic Acids Res. *31*, 6806–6818.

Creighton, T.E. (1995). Protein folding. An unfolding story. Curr. Biol. *5*, 353–356.

Creighton, T.E., Darby, N.J., and Kemmink, J. (1996). The roles of partly folded intermediates in protein folding. FASEB J. *10*, 110–118.

Das, R., Travers, K.J., Bai, Y., and Herschlag, D. (2005). Determining the Mg^{2+} stoichiometry for folding an RNA metal ion core. J. Am. Chem. Soc. *127*, 8272–8273.

Davis, J.H., Tan, Y.Z., Carragher, B., Potter, C.S., Lyumkis, D., and Williamson, J.R. (2016). Modular assembly of the bacterial large ribosomal subunit. Cell *167*, 1610–1622.e15.

Frederiksen, J.K., Li, N.S., Das, R., Herschlag, D., and Piccirilli, J.A. (2012). Metal-ion rescue revisited: biochemical detection of site-bound metal ions important for RNA folding. RNA *18*, 1123–1141.

Galej, W.P., Nguyen, T.H., Newman, A.J., and Nagai, K. (2014). Structural studies of the spliceosome: zooming into the heart of the machine. Curr. Opin. Struct. Biol. *25*, 57–66.

Golden, B.L., Gooding, A.R., Podell, E.R., and Cech, T.R. (1998). A preorganized active site in the crystal structure of the Tetrahymena ribozyme. Science *282*, 259–264.

Gracia, B., Xue, Y., Bisaria, N., Herschlag, D., Al-Hashimi, H.M., and Russell, R. (2016). RNA structural modules control the rate and pathway of RNA folding and assembly. J. Mol. Biol. *428*, 3972–3985.

Gruber, A.R., Bernhart, S.H., and Lorenz, R. (2015). The ViennaRNA web services. Methods Mol. Biol. *1269*, 307–326.

Hille, F., and Charpentier, E. (2016). CRISPR-Cas: biology, mechanisms and relevance. Philos. Trans. R. Soc. Lond. B Biol. Sci. *371*, 20150496.

Huang, Q., Purzycka, K.J., Lusvarghi, S., Li, D., Legrice, S.F., and Boeke, J.D. (2013). Retrotransposon Ty1 RNA contains a 5′-terminal long-range pseudoknot required for efficient reverse transcription. RNA *19*, 320–332.

Johnson, K.A., Simpson, Z.B., and Blom, T. (2009). Global kinetic explorer: a new computer program for dynamic simulation and fitting of kinetic data. Anal. Biochem. *387*, 20–29.

Jolley, E.A., and Znosko, B.M. (2017). The loss of a hydrogen bond: thermodynamic contributions of a non-standard nucleotide. Nucleic Acids Res. *45*, 1479–1487.

Kladwang, W., Mann, T.H., Becka, A., Tian, S., Kim, H., Yoon, S., and Das, R. (2014). Standardization of RNA chemical mapping experiments. Biochemistry *53*, 3063–3065.

Koculi, E., Cho, S.S., Desai, R., Thirumalai, D., and Woodson, S.A. (2012). Folding path of P5abc RNA involves direct coupling of secondary and tertiary structures. Nucleic Acids Res. *40*, 8011–8020.

Kole, R., Krainer, A.R., and Altman, S. (2012). RNA therapeutics: beyond RNA interference and antisense oligonucleotides. Nat. Rev. Drug Discov. *11*, 125–140.

Lee, J., Vogt, C.E., McBrairty, M., and Al-Hashimi, H.M. (2013). Influence of dimethylsulfoxide on RNA structure and ligand binding. Anal. Chem. *85*, 9692–9698.

Leontis, N.B., and Westhof, E. (2001). Geometric nomenclature and classification of RNA base pairs. RNA *7*, 499–512.

Lescoute, A., and Westhof, E. (2006). Topology of three-way junctions in folded RNAs. RNA *12*, 83–93.

Mironov, A.S., Gusarov, I., Rafikov, R., Lopez, L.E., Shatalin, K., Kreneva, R.A., Perumov, D.A., and Nudler, E. (2002). Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. Cell *111*, 747–756.

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

Murphy, F.L., and Cech, T.R. (1993). An independently folding domain of RNA tertiary structure within the Tetrahymena ribozyme. Biochemistry *32*, 5291–5300.

Murphy, F.L., and Cech, T.R. (1994). GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. J. Mol. Biol. *236*, 49–63.

Ng, E.W., Shima, D.T., Calias, P., Cunningham, E.T., Jr., Guyer, D.R., and Adamis, A.P. (2006). Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. Nat. Rev. Drug Discov. *5*, 123–132.

Nilsen, T.W. (1994). RNA-RNA interactions in the spliceosome: unraveling the ties that bind. Cell *78*, 1–4.

Pabit, S.A., Sutton, J.L., Chen, H., and Pollack, L. (2013). Role of ion valence in the submillisecond collapse and folding of a small RNA domain. Biochemistry *52*, 1539–1546.

Parks, J.W., and Stone, M.D. (2017). Single-molecule studies of telomeres and telomerase. Annu. Rev. Biophys. *46*, 357–377.

Petrov, A., Chen, J., O'Leary, S., Tsai, A., and Puglisi, J.D. (2012). Single-molecule analysis of translational dynamics. Cold Spring Harb. Perspect. Biol. *4*, a011551.

Russell, R., Millett, I.S., Doniach, S., and Herschlag, D. (2000). Small angle X-ray scattering reveals a compact intermediate in RNA folding. Nat. Struct. Biol. *7*, 367–370.

Sattin, B.D., Zhao, W., Travers, K., Chu, S., and Herschlag, D. (2008). Direct measurement of tertiary contact cooperativity in RNA folding. J. Am. Chem. Soc. *130*, 6085–6087.

Serganov, A., and Nudler, E. (2013). A decade of riboswitches. Cell *152*, 17–24.

Silverman, S.K., Zheng, M., Wu, M., Tinoco, I., Jr., and Cech, T.R. (1999). Quantifying the energetic interplay of RNA tertiary and secondary structure interactions. RNA *5*, 1665–1674.

Staley, J.P., and Guthrie, C. (1998). Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell *92*, 315–326.

Takamoto, K., Das, R., He, Q., Doniach, S., Brenowitz, M., Herschlag, D., and Chance, M.R. (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.

Tian, S., Cordero, P., Kladwang, W., and Das, R. (2014). High-throughput mutate-map-rescue evaluates SHAPE-directed RNA structure and uncovers excited states. RNA *20*, 1815–1826.

Tijerina, P., Mohr, S., and Russell, R. (2007). DMS footprinting of structured RNAs and RNA-protein complexes. Nat. Protoc. *2*, 2608–2623.

Travers, K.J., Boyd, N., and Herschlag, D. (2007). Low specificity of metal ion binding in the metal ion core of a folded RNA. RNA *13*, 1205–1213.

Vicens, Q., and Cech, T.R. (2006). Atomic level architecture of group I introns revealed. Trends Biochem. Sci. *31*, 41–51.

Wang, J., Daldrop, P., Huang, L., and Lilley, D.M. (2014). The k-junction motif in RNA structure. Nucleic Acids Res. *42*, 5322–5331.

Winkler, W., Nahvi, A., and Breaker, R.R. (2002a). Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. Nature *419*, 952–956.

Winkler, W.C., Cohen-Chalamish, S., and Breaker, R.R. (2002b). An mRNA structure that controls gene expression by binding FMN. Proc. Natl. Acad. Sci. U S A *99*, 15908–15913.

Winkler, W.C., Nahvi, A., Sudarsan, N., Barrick, J.E., and Breaker, R.R. (2003). An mRNA structure that controls gene expression by binding S-adenosylmethionine. Nat. Struct. Biol. *10*, 701–707.

Wu, H.N., and Huang, Z.S. (1992). Mutagenesis analysis of the self-cleavage domain of hepatitis delta virus antigenomic RNA. Nucleic Acids Res. *20*, 5937–5941.

Wu, M., and Tinoco, I., Jr. (1998). RNA folding causes secondary structure rearrangement. Proc. Natl. Acad. Sci. U S A *95*, 11555–11560.

Xue, Y., Gracia, B., Herschlag, D., Russell, R., and Al-Hashimi, H.M. (2016). Visualizing the formation of an RNA folding intermediate through a fast highly modular secondary structure switch. Nat. Commun. *7*, ncomms11768.

Yanofsky, C. (1981). Attenuation in the control of expression of bacterial operons. Nature *289*, 751–758.

Yoon, S., Kim, J., Hum, J., Kim, H., Park, S., Kladwang, W., and Das, R. (2011). HiTRACE: high-throughput robust analysis for capillary electrophoresis. Bioinformatics *27*, 1798–1805.

Yusupova, G., and Yusupov, M. (2014). High-resolution structure of the eukaryotic 80S ribosome. Annu. Rev. Biochem. *83*, 467–486.

Zheng, M., Wu, M., and Tinoco, I., Jr. (2001). Formation of a GNRA tetraloop in P5abc can disrupt an interdomain interaction in the Tetrahymena group I ribozyme. Proc. Natl. Acad. Sci. U S A *98*, 3695–3700.

Zhu, J.Y., and Meyer, I.M. (2015). Four RNA families with functional transient structures. RNA Biol. *12*, 5–20.