Hidden Structural Modules in a Cooperative RNA Folding Transition

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

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.
Hidden Structural Modules in a Cooperative RNA Folding Transition

Brant Gracia,1 Hashim M. Al-Hashimi,2,3 Namita Bisaria,4,5 Rhiju Das,3 Daniel Herschlag,4 and Rick Russell1,6,*
1Department of Molecular Biosciences and the Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, TX 78712, USA
2Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA
3Department of Chemistry, Duke University Medical Center, Durham, NC 27710, USA
4Department of Biochemistry, Stanford University, Stanford, CA 94305, USA
5Present address: Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA
6Lead 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 Mg2+-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 Nuider, 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 Mg2+ 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
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 or tertiary 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

Mg2+-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 Mg2+ 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 Mg2+ (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 Mg2+ dependencies of modification were complex, with a range of Mg2+ midpoints and some nucleotides with multiple phases (Figures 2A 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 Mg2+-dependent transitions between them (Figures S1 and S2; Table S3).

At sub-millimolar concentrations of Mg2+, several nucleotides in the 3WJ and the MC displayed modest enhancements in reactivity with increasing Mg2+ concentration (Figure 3A, U to U*; Figures S3A–S3C). These enhancements were also observed when Mg2+ was replaced with Ba2+, 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
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 IF). 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, IF 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 IF 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 4E). Second, a double mutant in P5c (U167C/U177C) that further stabilizes the alternative P5c

---

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.
secondary structure (Gracia et al., 2016) gave the same Mg2+ 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 Mg2+ 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 Mg2+ requirement was increased, from a [Mg2+]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 footpring data at 3 mM Mg2+. We chose this Mg2+ 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 Mg2+ 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 Mg2+ concentrations gave similar results (1.4 and 1.6 kcal/mol, respectively), indicating that the level of
cooperativity between the folding transitions in P5c and the MC does not depend strongly on the Mg2+ 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 Mg2+ 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 K0_Mc). Indeed, a mutant that stabilizes native P5c by inserting three additional Watson-Crick base pairs (Nat+3) displayed a reduced Mg2+ requirement for folding the MC (Figures 4A, 4B, and S4E). The equilibrium value at 3 mM Mg2+ 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 Mg2+ 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
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 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 native 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 S3B). Insertion of equivalent G/U pairs in the 3WJ restored the wild-type level for this 3WJ restoration mutant (Figures 5C and S5).

(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.

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 native 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 S3B).

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.

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 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$_{c}$) remained intact, but the P5c packaging transition at higher Mg$^{2+}$ concentrations ($I_{c}$ 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$_{c}$. 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 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 $\sim 2.6$ 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 the secondary structure change in P5c via structure formed in the
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 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.
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), RNA transcripts were isolated by affinity column (QIAGEN).

**SHAPE and DMS Footprinting**

P5abc RNA (200 nM) was pre-folded in 19 mL at the desired concentration of MgCl2 for 15 min at 37°C in a background of 100 mM KCl and 50 mM MgCl2.
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 RNA 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 T4 RNA ligase 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 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/msnb8c6x6bf.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

Akiami, 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., Pavilich, D.S., Mabuchi, H., and Horschlag, 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–E4966.
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.

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).
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.
Supplemental Information

Hidden Structural Modules
in a Cooperative RNA Folding Transition

Brant Gracia, Hashim M. Al-Hashimi, Namita Bisaria, Rhiju Das, Daniel Herschlag, and Rick Russell
Figure S1. Global fitting of footprinting data in Kintek Global Kinetic Explorer. Related to Figure 3
Red dots indicate the footprinting data and the curve shows the best fit by a global model, which included four states and three transitions. See Supplemental Experimental Procedures for further details.
Figure S2. Tests of global fitting using mock data sets with fewer points. Related to Figure 3

We performed a series of tests with subsets of the experimental data to determine whether the same global model would have been reached with data sets consisting of fewer points.

(A) Global fitting was performed using nucleotides with the largest and most descriptive signals for each of the three transitions. This test was performed using SHAPE data from U168, G169, A183, A184, and A186. (Results from A186 are not shown due to space constraints.) When the density of data is reduced by a factor of eight (right panels), the data can be adequately described with a model that includes the two Mg\(^{2+}\)-dependent transitions that are the focus of this work, but the transition at low Mg\(^{2+}\) concentrations is omitted from the model.

(B) Global fitting was performed using nucleotides that give large signals for the two Mg\(^{2+}\)-dependent folding transitions, but with both transitions giving a change in the same direction rather than an increase followed by a decrease in signal at higher Mg\(^{2+}\) concentrations. This test was performed using SHAPE data from U135 and A140, as well as DMS data from A173, A184, and A186. (Results from A186 are not shown due to space constraints.) When these data are reduced in density, they can be adequately described by a model that includes only a single Mg\(^{2+}\)-dependent transition.
Figure S3. Reactivity enhancements at sub-millimolar Mg$^{2+}$ concentrations depend on electrostatic environment and do not reflect formation of specific secondary or tertiary structures. Related to Figure 3

(A) 2D and 3D models indicating nucleotides that undergo SHAPE reactivity enhancements at low Mg$^{2+}$ concentrations (U to U*). MC nucleotides are orange and 3WJ nucleotides are green.

(B–C) SHAPE reactivity profiles for nucleotides within the 3WJ (panel B) and the MC (panel C) that display enhancements in the Mg$^{2+}$-dependent transition from U to U*.

(D) Dependence of SHAPE reactivity of the MC nucleotide A183 in the presence of various concentrations of Ba$^{2+}$, which does not support MC formation (Travers et al., 2007). The enhancement at low Mg$^{2+}$ remains present despite the lack of MC formation in Ba$^{2+}$ for the wild type P5abc (orange circles). In addition, the reactivity profile of A183 is essentially unaffected by a mutation that blocks MC formation (A186U, red triangles), a mutation that stabilizes alternative structure in P5c and thereby weakens MC formation (U167C, red angled triangles), and a mutation that stabilizes the native secondary structure of P5c (Nat+3 v1, blue).

(E) SHAPE footprinting of the MC in the presence of 1 M NaCl (black). This high Na$^+$ concentration results in electrostatic relaxation of the P5abc domain (Das et al., 2005), and we find that it eliminates the observed enhancements at sub-millimolar Mg$^{2+}$ concentrations for A183 (black circles) and A184 (black triangles). Corresponding data under standard conditions are shown in orange for comparison.
Figure S4. Cooperativity between the secondary structure switch in P5c and tertiary structure formation of the MC. Related to Figure 4

(A) SHAPE reactivity heat map of a mutant in which the alternative secondary structure of P5c is stabilized (U167C). Heat maps show reactivity of each nucleotide to SHAPE reagent in the absence and presence of saturating Mg\(^{2+}\) (50–200 mM Mg\(^{2+}\)). Nucleotides that have values between 0.3-1.0 are purple and those with values greater than 1.0 are red. The asterisk at position G169 indicates that this position displayed a substantial signal in the absence of the footprinting reagent, preventing analysis of this position.

(B) Mg\(^{2+}\)-concentration dependence of SHAPE reactivity for nucleotides in the 3WJ and MC (left) and in P5c (right) for the mutant U167C P5abc (red). Reactivity of the same nucleotides is shown also for the double mutant U167C/A186U (magenta) and the wild-type P5abc for reference (black).

(C–D) SHAPE reactivity heat maps and Mg\(^{2+}\)-concentration dependence for a mutant that blocks MC formation (A186U, red). Data are also shown for wild-type P5abc for comparison (black).

(E) SHAPE reactivity heat maps of a mutant that enforces the native P5c (Nat+3 v1).
Figure S5. Module-level restoration of tertiary structure formation. Related to Figure 5

(A) Folding of the construct in which tertiary structure formation in the 3WJ and MC is restored by insertions of G164, G176, and U177 equivalents. The figure shows the SHAPE reactivity pattern in the absence of Mg$^{2+}$ (left) and in the presence of saturating Mg$^{2+}$ (right, 50–200 mM Mg$^{2+}$).

(B) Insertion of nucleotides on one strand or the other of the connection between P5c and the 3WJ (G164 or G176/U177 equivalents) is not sufficient to rescue MC formation, relative to the absence of the insertions in a construct that stabilizes the alternative secondary structure of P5c (U167C).

(C) The G176A substitution destabilizes tertiary structure, supporting involvement of G176 in tertiary structure formation with the 3WJ. The G176A substitution was shown previously to stabilize the native secondary structure (Gracia et al., 2016; Xue et al., 2016) (secondary structures at left). Surprisingly, we found that G176A does not stabilize formation of the MC, as indicated by protection of A186 (purple curve, at right), but instead gives an indistinguishable Mg$^{2+}$ from that of the wild-type P5abc (black). This was surprising because the simplest model would have been that the stabilization of native secondary structure by G176A would have reduced the Mg$^{2+}$ requirement for tertiary structure formation. To test for the possibility of a destabilizing effect of G176A on tertiary structure, we separated out effects on secondary structure by measuring the effect of G176A in the background of a second mutation that enforces the native secondary structure. In this background, the G176A substitution increases the Mg$^{2+}$ requirement for MC formation (red vs blue), indicating that the G176A substitution destabilizes tertiary structure formation in the MC.
Figure S6. Mutation of U168 perturbs the P5c packing transition but does not affect MC formation. Related to Figure 6

(A) The mutation U168C is predicted to favor the native P5c secondary structure by 1.5 kcal/mol relative to the wild-type sequence of P5c.

(B–D) Folding of mutant U168C by SHAPE. Data in panel B show that the metal core forms at least as efficiently as in the wild-type P5abc. Data in panel C show the absence of protections at U168 and G169, relative to wild-type P5abc, suggesting that P5c does not pack efficiently with the MC. Data in panel D show that the 3WJ forms at least as efficiently as in the wild-type P5abc. Thus, the defective folding is specifically in the P5c packing step.

(E) Additional mutations of U168 (U168G and U168A) block the P5c packing transition (A184, closed symbols) but do not affect the lower Mg\textsuperscript{2+} transition (A186, open symbols). Note that A184 also undergoes a protection in the lower Mg\textsuperscript{2+} transition, and the defect in P5c packing is indicated by the lack of further protection in A184 at higher Mg\textsuperscript{2+} concentrations.
Figure S7. Quantitating the cooperativity between P5c packing and the MC. Related to Figure 7
Thermodynamic framework for MC folding, P5c secondary structure switching, and P5c packing. Values are measured at 3 mM Mg$^{2+}$. Calculated values are indicated with an asterisk (*). Note that the species that are populated for equilibrium folding of wild-type P5abc and depicted in Figure 7 are Alt (top left), I$_F$ (middle right) and F (bottom left).
Table S1. Mg$^{2+}$-dependent folding parameters from SHAPE footprinting. Related to Figure 3 (The table is in a separate Excel file due to its large size)
Table S2. Mg^{2+}-dependent folding parameters from DMS footprinting. Related to Figure 3

| Wild-type      | Region | Change    | [Mg^{2+}]_{1/2}, mM | Hill, n   | Amp, A  |
|----------------|--------|-----------|---------------------|-----------|---------|
| A173           | P5c    | Enhancement| 20 ± 10             | 0.92 ± 0.09 | 0.96 ± 0.16 |
| A183           | MC     | Protection | 15 ± 4              | 1.4 ± 0.4  | 1.1 ± 0.2 |
| A184           | MC     | Protection | 14 ± 5              | 1.8 ± 0.4  | 0.73 ± 0.07 |
| A186           | MC     | Protection | 5.9 ± 1.7           | 2.0 ± 0.1  | 1.1 ± 0.2 |
| A187           | MC     | Protection | 6.8 ± 1.7           | 1.6 ± 0.2  | 0.86 ± 0.06 |

| WT, 5% DMSO    | Region | Change    | [Mg^{2+}]_{1/2}, mM | Hill, n   | Amp, A  |
|----------------|--------|-----------|---------------------|-----------|---------|
| A173           | P5c    | Enhancement| 15 ± 1              | 0.78 ± 0.10 | 1.4 ± 0.21 |
| A183           | MC     | Protection | 12 ± 5              | 0.96 ± 0.22 | 1.3 ± 0.3 |
| A184           | MC     | Protection | 9.6 ± 2.8           | 1.0 ± 0.2  | 0.80 ± 0.08 |
| A186           | MC     | Protection | 2.9 ± 0.3           | 1.6 ± 0.4  | 1.1 ± 0.2 |
| A187           | MC     | Protection | 3.9 ± 0.5           | 1.4 ± 0.1  | 0.78 ± 0.06 |

| A:U            | Region | Change    | [Mg^{2+}]_{1/2}, mM | Hill, n   | Amp, A  |
|----------------|--------|-----------|---------------------|-----------|---------|
| A173           | P5c    | Enhancement| 3.4 ± 0.8           | 1.1 ± 0.1  | 1.0 ± 0.2 |
| A183           | MC     | Protection | 77 ± 3              | 2.1 ± 0.3  | 0.69 ± 0.01 |
| A184           | MC     | Protection | 23 ± 8              | 0.81 ± 0.23 | 0.59 ± 0.12 |
| A186           | MC     | Protection | 3.2 ± 0.4           | 0.88 ± 0.02 | 0.71 ± 0.03 |
| A187           | MC     | Protection | 7.3 ± 2.3           | 0.70 ± 0.05 | 0.78 ± 0.16 |

| A:U/A186U      | Region | Change    | [Mg^{2+}]_{1/2}, mM | Hill, n   | Amp, A  |
|----------------|--------|-----------|---------------------|-----------|---------|
| A173           | P5c    | Enhancement| 455 ± 17            | 0.98 ± 0.03 | 0.95 ± 0.05 |
| A183           | MC     | NC        | NC                  | NC        | NC      |
| A184           | MC     | NC        | NC                  | NC        | NC      |
| U186           | MC     | RTS       | RTS                 | RTS       | RTS     |
| A187           | MC     | NC        | NC                  | NC        | NC      |

Nucleotides that do not change are labeled as NC. Positions that have reverse transcription stops are labeled as RTS.
Table S3. Parameters from global modeling. Related to Figure 3

| Model                  | $K_1$ | $K_2$ | $K_3$ | $K_4$ | $\chi^2$ |
|------------------------|-------|-------|-------|-------|----------|
| 4 state, A-B coop      | 0.042 | 11    | 0.98  | NA    | 30.1 ± 0.2 |
| 4 state, B-C coop      | 0.28  | 2.7   | 2.0   | NA    | 63.7 ± 0.2 |
| 5 state, A-B coop      | 0.04  | 1.2   | 2.1   | 12    | 16.0 ± 0.1 |
| 5 state, B-C coop      | 0.5   | 2.4   | 1.9   | 9.1   | 14.7 ± 0.1 |
| 5 state, C-D coop      | 2.8   | 28    | 0.097 | 3.0   | 30.3 ± 0.2 |

See Supplemental Experimental Procedures for the initial constraints of each global fit. NA, not applicable for the model.
Supplemental Experimental Procedures

Global modeling of SHAPE and DMS footprinting data: We imported 29 sets of quantified SHAPE and DMS footprinting data from the wild-type P5abc into Kintek Kinetic Explorer and fit all of the data using one global model (Johnson et al., 2009). The global modeling was evaluated by minimizing the \( \chi^2 \) value of the fit (defined as the sum of the residuals squared, assuming a constant sigma value for all data points). As described in Results, the simplest model that adequately described the data included four P5abc conformations, or states, with three Mg\( ^{2+} \)-dependent transitions between them. From inspection of the data, it was clear that at least one of these transitions, centered at approximately 2 mM Mg\( ^{2+} \), gave steeper dependences than would be possible for binding of a single Mg\( ^{2+} \) ion. To enable one of the transitions to be cooperative with respect to Mg\( ^{2+} \) binding, which is not explicitly allowed in the Kinetic Explorer software, we incorporated an additional state into the model, making a total of five states (Scheme 1). The goal in including this ‘extra’ state was that it would not be populated significantly at any Mg\( ^{2+} \) concentration but would allow a single transition to involve binding of two Mg\( ^{2+} \) ions. Thus, for example state B could bind a Mg\( ^{2+} \) ion to progress to state C, which would not be populated because state C would then bind an additional Mg\( ^{2+} \) ion with higher affinity to progress to state D. Note that this transition likely involves the uptake of two site-bound Mg\( ^{2+} \) ions that are associated with the folded MC (Das et al, 2005), but the dependence indicates only the net uptake of two Mg\( ^{2+} \) ions, which could bind to specific sites or become part of the delocalized ion atmosphere.

\[
A \leftrightarrow B \leftrightarrow C \leftrightarrow D \leftrightarrow E \quad \text{(Scheme 1)}
\]

Reactivity = \( sa_1*A + sb_1*B + sc_1*C + sd_1*D + se_1*E \) \quad \text{(Equation 1)}

As indicated in eq. 1, the observed reactivity for each nucleotide is defined as a linear combination of scaling factors associated with each state. In the global fitting, the scaling factors for each nucleotide were allowed to vary for each state except the non-populated state within the transition that is cooperative with Mg\( ^{2+} \) binding. The scaling factors for the non-populated states were constrained to be equal to those of the state preceding it, and both of these scaling factors were locked to reasonable values prior to the fitting. In the model of Scheme 1, there are three possibilities for this folding step: the transition from A to C, from B to D, or C to E. We tested these possibilities individually by systematically requiring the scaling factors of each nucleotide to be the same for states A and B (\( sa_1 = sb_1 \)), for states B and C (\( sb_1 = sc_1 \)), or for states C and D (\( sc_1 = sd_1 \)). This constraint was applied to mimic the condition of a single transition by preventing the artificial intermediate state from having a unrealistically high scaling factor and consequently contributing to the overall signal in the transition region despite being at most sparsely populated. After global fitting under each constraint, the condition where state C is not significantly populated (five states, with the cooperative transition from state B to state D) produced a global fit that best described the data (Figure S1; Table S3).

We tested whether a simpler model could adequately describe the data by decreasing the number of states to four, one of which would again be a non-populated state included to allow one transition to be cooperative with Mg\( ^{2+} \) concentration. We systematically constrained the scaling factors to allow each transition, in turn, to be the cooperative one, and we performed global fits to the data for the wild-type P5abc as above.

\[
A \leftrightarrow B \leftrightarrow C \leftrightarrow D \quad \text{(Scheme 2)}
\]

React = \( sa_1*A + sb_1*B + sc_1*C + sd_1*D \) \quad \text{(Equation 2)}

Global fitting under constraints to allow the first transition to be cooperative with Mg\( ^{2+} \) concentration (i.e. \( sa_1 = sb_1 \)) converged on a model that fit poorly and in which the first two transitions were non-cooperative, as \( K_1 < K_2 \) (\( K_1 = 0.042 \text{ mM Mg}^{2+} \), \( K_2 = 11 \text{ mM Mg}^{2+} \); Table S3). When the second folding step was allowed to be cooperative with Mg\( ^{2+} \) concentration, the resulting model did include this cooperativity, as indicated by a smaller value of \( K_3 \) than \( K_2 \) (\( K_2 = 2.7 \text{ mM Mg}^{2+} \), \( K_3 = 2.0 \text{ mM Mg}^{2+} \); Table S3), but this model was unable to accommodate the large changes in signal at Mg\( ^{2+} \) concentrations in the range of 10 mM, resulting in a poor overall fit, as indicated by a high \( \chi^2 \) value (Table S3). Thus, the best fit to the data included a total of four ‘real’ states, with an additional fifth state included in the model to allow for one of the transitions to be cooperative with Mg\( ^{2+} \) concentration (Scheme 1).
Assignment of the transition that increases the DMS reactivity of A173: An interesting feature of the footprinting data was the enhanced DMS reactivity of A173, which suggests a tertiary-contact-induced transition from a Watson-Crick pair of U167-A173 to the non-canonical pair observed in the crystal structure. This DMS enhancement at A173 was suggested previously to monitor the secondary structure change of P5c (Silverman et al., 1999), but we conclude that it instead reflects P5c tertiary packing for two reasons. First, its Mg\(^{2+}\) dependence is the same as that for other nucleotides in the P5c packing transition (Figure 2F, A184 and G188) and different from those in the secondary structure change (Figure 2E, U168 and G169 enhancements at 1 mM). Second, the secondary structure change is not expected to enhance the DMS reactivity of A173. A173 remains paired, switching its partner from U168 to U167, and the \(^{15}\)N chemical shift of U167 in the excited state suggests a Watson-Crick pair between U167 and A173 (Xue et al., 2016). DMS footprinting of the P4-P6 domain, the folded and assembled P5abc ribozyme core complex, and the full-length, folded ribozyme show that A173 is strongly reactive to DMS (Gracia et al., 2016; Murphy and Cech, 1993; Tijerina et al., 2007), suggesting that the fully folded and packed state we observe is the native conformation of P5abc present in the intact intron.