Nucleic Acids Research Advance Access published May 28, 2012

Nucleic Acids Research, 2012, 1–10
doi:10.1093/nar/gks468

Folding path of P5abc RNA involves direct coupling of secondary and tertiary structures

Eda Koculi[1], Samuel S. Cho[2], Ravi Desai[1], D. Thirumalai[2,*] and Sarah A. Woodson[1,*]

[1]T.C. Jenkins Department of Biophysics, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218 and [2]Program in Biophysics, Institute for Physical Sciences and Technology and Department of Chemistry, University of Maryland, College Park, MD 20742, USA

Received December 7, 2011; Accepted May 1, 2012

ABSTRACT

Folding mechanisms in which secondary structures are stabilized through the formation of tertiary interactions are well documented in protein folding but challenge the folding hierarchy normally assumed for RNA. However, it is increasingly clear that RNA could fold by a similar mechanism. P5abc, a small independently folding tertiary domain of the *Tetrahymena thermophila* group I ribozyme, is known to fold by a secondary structure rearrangement involving helix P5c. However, the extent of this rearrangement and the precise stage of folding that triggers it are unknown. We use experiments and simulations to show that the P5c helix switches to the native secondary structure late in the folding pathway and is directly coupled to the formation of tertiary interactions in the A-rich bulge. P5c mutations show that the switch in P5c is not rate-determining and suggest that non-native interactions in P5c aid folding rather than impede it. Our study illustrates that despite significant differences in the building blocks of proteins and RNA, there may be common ways in which they self-assemble.

INTRODUCTION

In RNA folding, the stability gap separating the free energy of the native state from other non-native structures is often small (1–3). Thus, misfolded conformations in which even the secondary structure contains errors can form with substantial probability (4). One way to correct errors in the secondary structure is by coupling base pair reorganization with RNA tertiary folding, either by realigning entire helices (5–7) or by shuffling non-Watson–Crick base pairs within a helix (8,9). Although it has been known for some time that certain base pairs are not stable until the RNA tertiary structure forms (10), tertiary interactions may also be used to correct non-native secondary structures that are present in the RNA. Direct coupling between RNA secondary and tertiary structure challenges the conventional view (3,11) that native secondary structures always form before the establishment of tertiary interactions.

The ease with which RNAs form alternative stable structures raises the possibility that tertiary interaction-driven base pair switching, and not merely helix stabilization, might be more widespread than commonly assumed. To explore this possibility, we investigated the refolding kinetics of the P5abc RNA that switches its secondary structure when inter-helix tertiary interactions are formed (Figure 1) (6,12). The P5abc RNA is an independently folding tertiary domain of the *Tetrahymena thermophila* group I ribozyme (13,14) which is one of the fastest regions of the ribozyme tertiary structure to form (15). P5abc is composed of three helices (paired regions P5a, P5b and P5c) that interact via a rich web of tertiary contacts in the folded state (16). One set of tertiary interactions is formed between an A-rich bulge in helix P5a and the loop of helix P5c and another is formed by tandem G-A pairs at the junction between helices P5a, P5b and P5c.

A nuclear magnetic resonance (NMR) structure of a truncated form of the P5abc domain (tP5abc) showed that the P5abc secondary structure in the absence of Mg^{2+} (Figure 1A) differs significantly from that of the Mg^{2+}-bound state (Figure 1B), implying that tertiary folding is coupled to a reorganization of the RNA secondary structure (6). First, the base pairs in P5c change register, expanding the stable GNRA tetraloop in the extended form of P5c to a less stable pentaloop in the native state that positions U168 to interact with G188 in

*To whom correspondence should be addressed. Tel: +1 410 516 2015; Fax: +1 410 516 4118; Email: swoodson@jhu.edu
Correspondence may also be addressed to D. Thirumalai. Tel: +1 301 405 4803; Fax: +1 301 314 9404; Email: dave.thirumalai@gmail.com
Present addresses:
Eda Koculi, Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637, USA.
Samuel S. Cho, Departments of Physics and Computer Science, Wake Forest University, Winston Salem, NC 27109, USA.

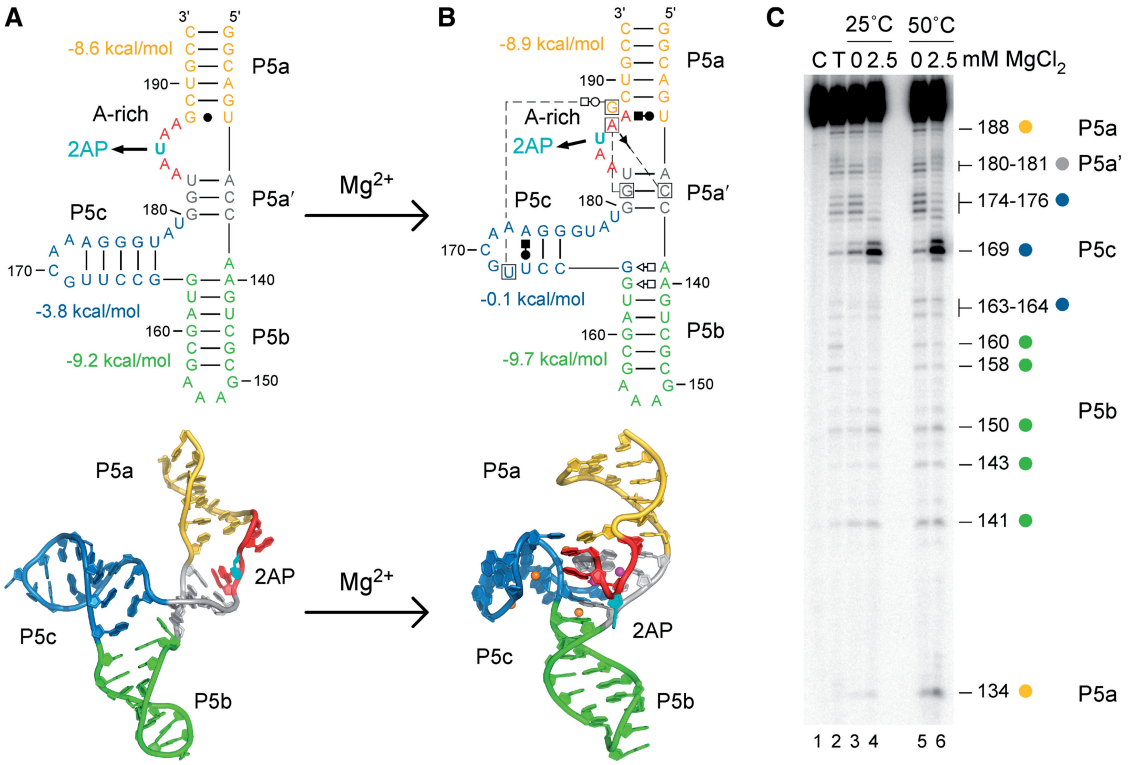

Figure 1. Structures of the tP5abc RNA. (**A**) The secondary structures (top) and ribbon diagrams (bottom) of the extended tP5abc from NMR (6) are compared with (**B**) the folded P5abc structure from X-ray crystallography (22). Mg^{2+} ions visible in the X-ray structure are represented by spheres. Nucleotides are numbered discontinuously because 5 bp of the natural P5b are missing in tP5abc. Non-Watson–Crick base pairs are represented by Leontis/Westhof symbols (49). U185 was replaced with the fluorescent analog 2AP (cyan), which unstacks in the folded RNA. Base-pairing free energies are from MFOLD (35). (**C**) RNase T1 probing of tP5abc secondary structure at 25 and 50°C, with no Mg^{2+} (extended) or 2.5 mM Mg^{2+} (folded). Lane C, no digestion control; lane T, RNase T1 sequencing ladder.

P5a (6,12,17). Second, reorganization of the three-helix junction replaces base pairs A139-U179 and G164-U177 with tandem G-A pairs (A139-G164 and A140-G163) in the native from of P5b.

Although the native secondary structure of P5abc is less stable than its extended secondary structure, this is compensated by the favorable free energy of tertiary interactions in the presence of Mg^{2+} ions (17,18). In addition to non-specific charge stabilization (19,20), Mg^{2+} ions specifically coordinate a 'corkscrew' in the backbone of the A-rich bulge (magenta spheres; Figure 1B), while four additional ions interact with P5c and with tandem G-A pairs in P5b (orange spheres; Figure 1B) (21,22). Evidence for specific binding of Mg^{2+} to P5abc RNA in solution comes from NMR, phosphorothioate substitution-interference (12,21) and refolding experiments (23,24).

Force denaturation experiments confirmed the importance of tertiary interactions for global stability of the P5abc domain, as high forces were required to rupture the P5abc structure in Mg^{2+} (25). More importantly, when the force was relaxed, refolding occurred in two steps by a kinetic partitioning mechanism (26). Of particular relevance is that the single-molecule experiments showed that folding of P5abc, upon release of force, involves intermediates regardless of the Mg^{2+} concentration. However, such experiments provide no information about the structure of the intermediates nor their location along the refolding pathways.

Herein, we provide evidence that helix switching occurs in the same step of folding as the formation of tertiary interactions in P5abc. Using a fluorescence assay, we show that the solution refolding pathway contains at least one intermediate. However, experiments and simulations show that the base pair shift in P5c does not limit the rate of folding. Instead, our results show that switching of the P5c helix occurs in the same folding step as the formation of tertiary interactions in the A-rich bulge and that both of these events follow the slow step of folding. Our results show that the energy landscape of even the rapidly folding P5abc subdomain is rugged, exhibiting many of the complexities found in the folding kinetics of larger ribozymes.

MATERIALS AND METHODS

Coarse-grained dynamics simulations

Coarse-grained simulations used the three interaction site (TIS) model (Supplementary Figure S1) in which each nucleotide is represented by three interactions sites corresponding to the centers of masses of the base, sugar and phosphate (27). The force field for the TIS model includes base-stacking interactions which were taken from

empirical energy parameters (28). G•A pairs were given the same weight as A-U pairs. Replacing G•A parameters with those for G-C pairs resulted in negligible differences to our simulation results (compare Figure 3 and Supplementary Figure S2).

For thermodynamic analyses, multiple trajectories containing numerous folding and unfolding transitions were collected. All the simulations were performed using Langevin dynamics in the low friction limit ($\zeta = 0.05\ \tau_L^{-1}$), which enhances the rate of conformational sampling (29). The unit of time is $\tau_L = \left(ma^2/\varepsilon_h\right)^{1/2}$, in which the mass of the RNA beads is m = 0.1–0.16 kg/mol, the average distance between beads is a = 4.6 Å and the energy scale is $\varepsilon_h \sim 1$ kcal/mol. Profiles of the specific heat with respect to temperature, C_v(T), were obtained using Weighted Histogram Analysis Method (WHAM), and a peak in C_v(T) designates a melting transition temperature. The trajectories at the melting temperature were collected to generate the free energy profile with respect to the reaction coordinate, Q, the fraction of native contacts.

RNA preparation

Truncated P5abc RNA (6) containing 2-aminopurine (2AP) was chemically synthesized (Dharmacon) and purified by denaturing 10% polyacrylamide gel electrophoresis. The RNA solution was desalted and concentrated by ultracentrifugation (Centricon Plus-20; Millipore) and stored at −20°C in water. The typical yield after purification was 10–20%.

UV thermal denaturation

Thermal denaturation of P5abc was monitored at 260 nm at a heating rate of 0.1°C/min on a Cary 400 spectrophotometer. tP5abc RNA (2 μM) was dissolved in 10 mM Na–cacodylate (pH 7.5), 40 mM NaCl plus the desired concentration of $MgCl_2$. Global Melt Fit provided by D. Draper (http://ded.chm.jhu.edu:16080/~draper/) was used to obtain enthalpies and melting temperatures, assuming two transitions and temperature-dependent baselines.

RNase T1 partial digestion

5′-[^{32}P]-labeled tP5abc was prepared with T4 polynucleotide kinase and gel-purified as described previously (30). The RNA was equilibrated 10 min in CE buffer (10 mM cacodylate, pH 7.5; 0.5 mM EDTA, pH 8.5) with or without 2.5 mM $MgCl_2$ before treatment with 0.1 U RNase T1. Products were resolved on a 12% sequencing gel.

Equilibrium Mg^{2+} titrations

For equilibrium folding, 2 ml 0.5 μM RNA in CE and 2 ml 0.5 μM RNA in CE plus $MgCl_2$ were denatured 5 min at 50°C and degassed. These solutions were mixed to produce Mg^{2+} titrations and incubated 5 min at 30°C after each addition before the emission spectra (3–5 scans) were recorded with excitation at 310 nm (Perkin Elmer LS50B). Excitation and emission slit widths were 15 and 10 nm, respectively.

The fluorescence intensity F at 370 nm versus Mg^{2+} concentration, C, was fit to

$$F = \alpha_U + (\alpha_N + \beta_N C - \alpha_U)\left(\frac{C^n}{C^n + C_m^n}\right) \quad (1)$$

in which C_m is the midpoint of the folding transition, n is the Hill coefficient, α_U is the fluorescence intensity at zero Mg^{2+} and a linear baseline for the folded RNA is given by α_N and β_N. The ΔG of folding was calculated from $\Delta G = RT\sqrt{8\Omega_c \ln\left(3+2\sqrt{2}\right)}$, in which $\Omega_c = C_{max}^2(df/dC)_{max}/\Delta Cat1/2max$ is a measure of the cooperativity (31). Unfolding by dilution of Mg^{2+} produced similar results (data not shown).

Urea denaturation

At equilibrium, 0.5 μM RNA in CE plus 1 mM $MgCl_2$ was mixed with the same solution also containing 9 M urea to produce the desired final urea concentration. The fluorescence data were fit to the linear binding model (32,33):

$$F = \frac{(\alpha_N + \beta_N D) + (\alpha_U + \beta_U D)\exp(-m(D_m - D)/RT)}{1 + \exp(-m(D_m - D)/RT)} \quad (2)$$

in which D is the concentration of urea, D_m is the concentration of denaturant at the midpoint of the unfolding transition and m is a constant that reflects the change in folding free energy with denaturant concentration. Linear baselines ($\alpha + \beta D$) corrected for a slight change in the fluorescence of the folded state with urea, and a larger increase in 2AP fluorescence of the extended state with urea as stacking between 2AP and the adjacent bases was disrupted. The data were normalized as in Equation (1).

Folding kinetics

Folding kinetics was measured on an Applied Photophysics SX.18MV stopped-flow spectrometer with a dead time of 1.3 ms. The excitation wavelength was 310 nm, and the emission signal was measured using a 360 nm wide band cut-off filter. Folding reactions were performed by mixing tP5abc RNA prepared as mentioned earlier (1 μM in CE) with an equal volume of CE plus 2 × $MgCl_2$. Unfolding reactions were performed by pre-folding the RNA in CE plus 1 mM $MgCl_2$ and diluting it 26-fold with buffer to reach the desired final $MgCl_2$ concentration. Unfolding rates in urea were obtained by mixing 1 volume folded RNA (CE, 1 mM $MgCl_2$) with 10 volumes urea in CE, 1 mM $MgCl_2$. Folding rates were obtained by diluting RNA in 5.5 M urea 10:1 with buffer. Individual kinetic traces were fit to a single exponential rate equation and the resulting rate constants averaged, except for folding reactions of G174A, C166U RNA which were fit to a double exponential equation.

The observed rate k_{obs} as a function of Mg^{2+} concentration C was fit by a three-state model $E \leftrightarrow I \leftrightarrow N$, in which the forward rate was given by $k_{obs,f} = k_f C^r$, where r is the apparent number of ions linked to the slow step of

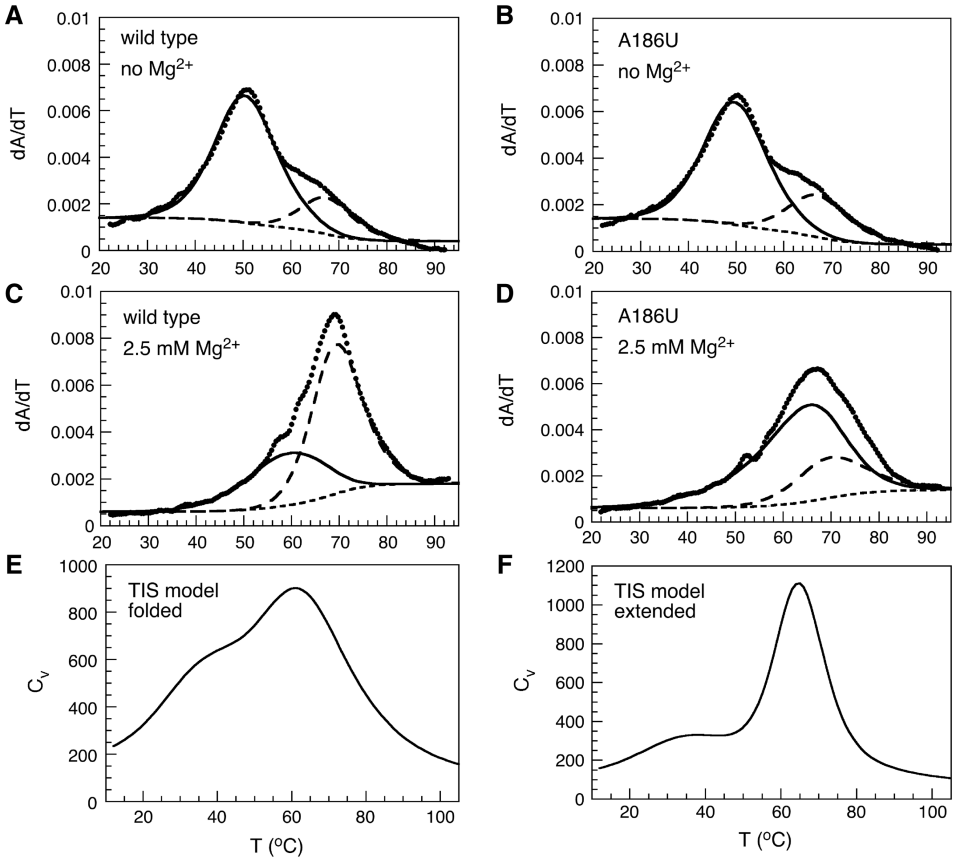

Figure 2. Thermal stability of tP5abc RNA. (**A–D**) The first derivative of absorbance at 260 nm with temperature was fit to two transitions. All experiments were in 40 mM NaCl, 10 mM Na-cacodylate (pH 7.5). Symbols: dotted lines, data; solid lines, first transition; long dashed lines, second transition; short dashed lines, baseline. (**E** and **F**) Simulated thermal denaturation profiles showing the specific heat C_v versus temperature for the folded (E) and extended (F) structures of tP5abc.

the forward folding reaction (7). The unfolding pathway (N to I) was assumed to depend on the stability of N. Expressions for the forward and reverse paths were combined yielding the following equation:

$$k_{obs} = k_u\left(\frac{K_2^m}{C^m+K_2^m}\right)+k_f C^r \quad (3)$$

in which k_f and k_u are the folding and unfolding rate constants, K_2 is the midpoint and m is the slope of the equilibrium between I and N.

β values

The values of β for the Mg^{2+} titrations were calculated from $\beta = n^{\ddagger}/n_H$, in which $n_H = d\ln K/d\ln C$ was taken from the equilibrium folding curves and $n^{\ddagger} = d\ln k_{obs}/d\ln C$ was taken from the folding kinetics over the Mg^{2+} concentration ranges in which the folding and unfolding arms were linear.

RESULTS

Thermal denaturation

For these studies, we used the 56 nt tP5abc, which lacks five base pairs in the center of P5b and three nucleotides at the top of P5a (6) (Figure 1) yet forms the correct tertiary structure around the three-helix junction (12,17). UV thermal denaturation experiments in the absence of Mg^{2+} showed that the extended secondary structure unfolds in two transitions (Figure 2A) with midpoints at 50 and 68°C (Supplementary Table S1). The melting profile in the absence of Mg^{2+} was altered very little by the mutation A186U, which disrupts tertiary interactions between the A-rich bulge and P5c but does not change the predicted secondary structure (Figure 2B). In 2.5 mM Mg^{2+}, the secondary structure opened at a higher T_m, as expected (Figure 2C and D). Moreover, the 'wild type' tP5abc exhibited an additional low temperature unfolding transition not visible for the A186U mutant, which we assigned to denaturation of the tertiary structure.

RNase T1 probing

We next probed the tP5abc secondary structure by RNase T1 partial digestion at 25 and 50°C to verify that Mg^{2+} induces the expected shift in P5c base pairing (Figure 1C). At 25°C, which is below the first thermal transition, Gs in P5a, P5b and at the junction between P5b and P5c (G163 and G164) were protected from cleavage, indicating that these helices are stable at 25°C in the extended structure (lane 3, Figure 1C). As expected, these residues were

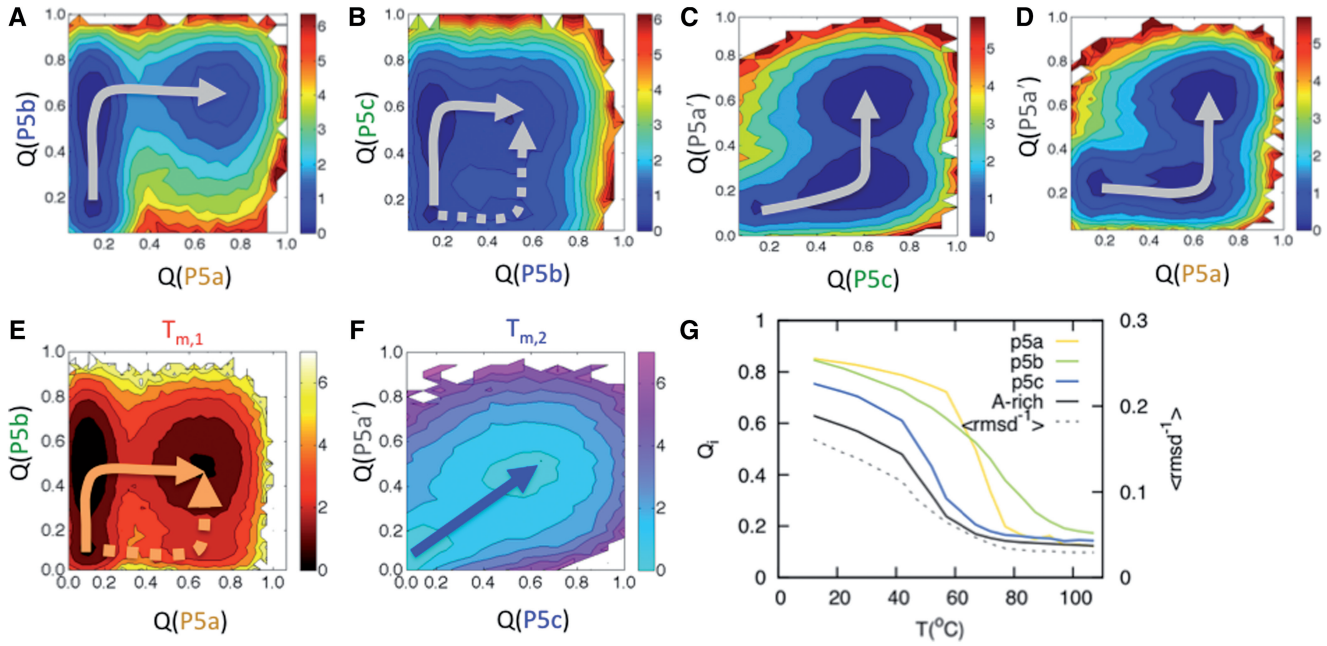

Figure 3. Free energy landscapes of tP5abc from simulation. Free energy profiles were projected to the fraction of native contacts (Q) for each region of tP5abc as defined in Figure 1. The free energy profile is calculated at the melting temperature in units of k_BT. (**A–D**) Energy profiles for the extended RNA. (**E** and **F**) Energy profiles for the folded state, at the high temperature transition (E) or low temperature transition (F). (**G**) Comparison of secondary structures (Q_i) with native tertiary contacts, as reported by the overall root mean square deviation (dashed line).

cleaved more strongly at 50°C, the T_m of the first thermal unfolding transition (lane 5, Figure 1C).

When the tP5abc was folded in 2.5 mM $MgCl_2$, G174-G176 in P5c and G180-G181 in P5a′ also became protected at 25°C, showing that these base pairs near the three-helix junction become more stable in the tertiary structure (lane 4, Figure 1C). Although bases in P5c are predicted to be paired in the extended structure, this helix may breathe or rearrange more than expected in low salt conditions. However, Mg^{2+} strongly enhanced cleavage of G169 in the loop of P5c, confirming the expected shift from a sheared G-A pair in the tetraloop of the extended form to an unpaired and exposed G in the pentaloop of folded tP5abc (lanes 4 and 6, Figure 1C).

Simulations of structural transitions during thermal unfolding

To interpret the unfolding pathway implied by the thermal denaturation curves, we performed coarse-grained molecular dynamics simulations of tP5abc using the TIS model in which each nucleotide is represented by three beads (27). We performed two sets of simulations, in which the TIS Hamiltonian was derived from either the extended tP5abc structure or the folded tP5abc structure (Supplementary Figure S1). In simulations of folded tP5abc, the specific heat (C_v) profile had a well-defined shoulder (Figure 2E), indicating two distinct melting transition temperatures as observed in our experiments (Figure 2C). In contrast, the C_v profile for the extended tP5abc had one peak (Figure 2F), corresponding to a single melting transition very similar to that of the A186U mutant in the presence of Mg^{2+} (Figure 2D). The overall specific heat profiles were in good agreement with the measured profiles, with similar transition temperatures, relative peak heights and overall shape.

To quantify the structural transitions associated with the melting of tP5abc, we calculated a series of 2D, temperature-dependent free energy profiles, $F(Q_i, Q_j)$, in which Q_i and Q_j are the fractions of native contacts of the individual helices in the extended tP5abc or folded tP5abc (Figure 3). Although the single peak in the C_v profile indicated cooperative folding of extended tP5abc (Figure 2F), the free energy profiles revealed intermediates corresponding to the formation of individual helices. The 2D landscapes showed that the internal helices P5b and P5c form independently of each other with similar free energy barriers, in agreement with the RNase T1 data, while P5b is prerequisite for the closing helix, P5a (Figure 3A and B). P5a and P5c in turn stabilize P5a′, completing the extended secondary structure (Figure 3C and D). Because the extended form of P5c is much more stable than the native P5c helix (Figure 1A and B), our simulations predict that it forms early in the folding while native P5c does not.

To compare the folding mechanism of the extended structure to that of the folded structure, we constructed free energy landscapes for folded tP5abc at each of the two melting temperatures. These landscapes showed that the high temperature transition corresponds to sequential unfolding of P5a then P5b (Figure 3E), while the lower melting temperature corresponds to the cooperative unfolding of P5c and P5a′ (Figure 3F). Thus, under native conditions, P5c and P5a′ adjacent to the A-rich bulge fold cooperatively, rather than sequentially as in the extended state.

We next compared formation of tertiary structure, represented by the overall root mean square deviation, with

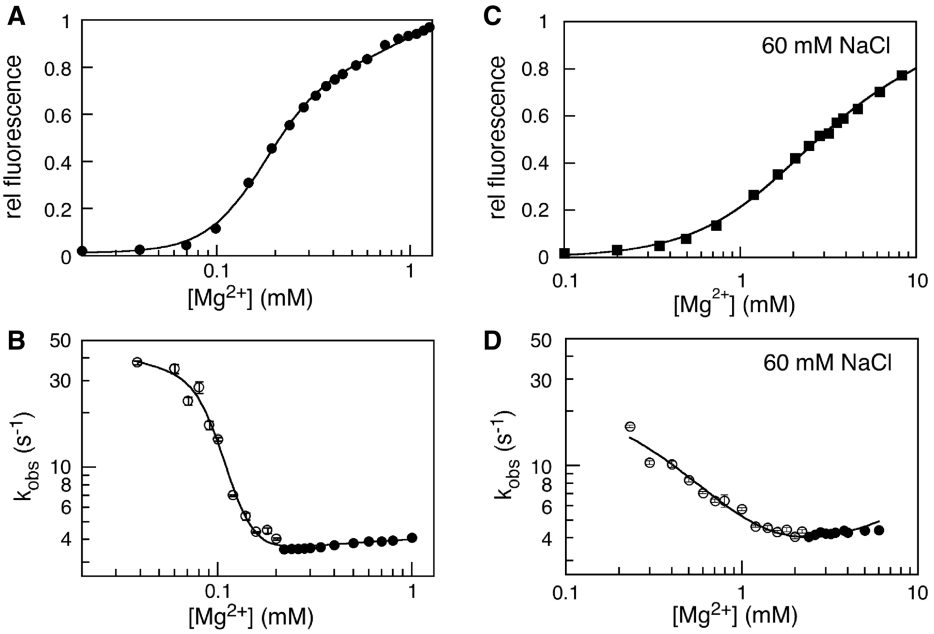

Figure 4. Mg^{2+} dependence of folding at 30°C. (**A**) Cooperative change in the folding equilibrium with Mg^{2+}. The fluorescence intensity at 370 nm versus Mg^{2+} concentration in 10 mM CE was fit to a cooperative model [Equation (1)]; see Supplementary Table S2 for fit parameters. (**B**) Observed folding kinetics versus Mg^{2+} concentration. Errors are the standard deviation of three experiments. The line represents the best fit to the model in Equation (3), with $k_u = 36\,s^{-1}$, $k_f = 4\,s^{-1}$ (Supplementary Table S3). Filled and open symbols represent kinetics from folding and unfolding reactions, respectively (see 'Materials and Methods' section). (**C**) The midpoint of the Mg^{2+} titration was 10-fold higher in 60 mM Na^+ than in CE buffer alone because the Na+ ions compete with Mg^{2+} for access to the RNA ($C_m = 1.8$ mM; Supplementary Table S2). (**D**) Observed folding rate versus Mg^{2+} concentration in 60 mM NaCl, as in (**B**), with $k_u \approx 40\,s^{-1}$ and $k_f = 3.6\,s^{-1}$ (Supplementary Table S3).

secondary structure formation as measured by Q_i (Figure 3G), and found that the tertiary structure forms in the lower temperature transition, concurrently with P5c and P5a′. Overall, the experimental data and the simulations show that native tP5abc unfolds through at least one equilibrium intermediate, corresponding to the loss of tertiary interactions and P5c. Our finding that the native P5c helix forms (or unfolds) in the same thermal transition as the tertiary structure shows that these steps are strongly coupled as previously suggested (6) and in agreement with the experimental RNase T1 results (Figure 1C).

Folding P5abc in Mg^{2+}

We next asked whether similar intermediates were populated when P5abc refolds in solution. To study the Mg^{2+}-dependent folding pathway of tP5abc experimentally, we introduced the fluorescent base analog 2AP at position 185 (cyan; Figure 1A). U185 stacks with adjacent adenosines in the extended conformer but flips out in the native structure. When U185 was replaced by 2AP, the fluorescence intensity of the labeled RNA increased upon titration with Mg^{2+}, consistent with unstacking of this base as the RNA folds (Supplementary Figure S3A).

Several control experiments indicated that the structure of tP5abc was not significantly perturbed by the 2AP substitution. First, the 2AP substitution made almost no difference to the thermal denaturation profiles of the RNA (Supplementary Figure S3B). Second, the folding time ($t_{1/2}$ = 180 ms in 0.5 mM $MgCl_2$ at 30°C; Supplementary Figure S3C) was compatible with previous estimates of the folding kinetics from time-resolved hydroxyl radical footprinting (34) and NMR (12). Finally, a similar Mg^{2+}-dependent folding transition was observed when pyrene was covalently attached to the 2′OH of U142 (Supplementary Figure S4).

The stability of tP5abc was obtained from the increase in fluorescence intensity with Mg^{2+} concentration at equilibrium (Figure 4A). The midpoint of the folding transition (C_m) was 0.18 mM Mg^{2+} at 30°C (in 20 mM Na-cacodylate), reasonably similar to previous results (17). From the steepness of the folding transition with respect to Mg^{2+} concentration (31), we obtained $\Delta G_{UN} = -1.3 \pm 0.5$ kcal/mol (Supplementary Table S2).

Mg^{2+} dependence of the folding kinetics

To probe the folding pathway of tP5abc in solution, the folding and unfolding kinetics was measured by stopped-flow fluorescence in Mg^{2+} concentrations spanning the equilibrium transition (Figure 4B). The observed unfolding rate became faster as the Mg^{2+} was reduced below 0.2 mM, which was expected because the equilibrium no longer favors the folded state (Figure 4B). Surprisingly, however, the observed folding rate did not increase as the Mg^{2+} was raised above 0.2 mM, although the native tertiary structure was becoming more stable. These results showed that the folding rate is determined by a slow step that does not depend on the stability of the native state or the recruitment of additional Mg^{2+} ions to the RNA. Higher salt did not eliminate the

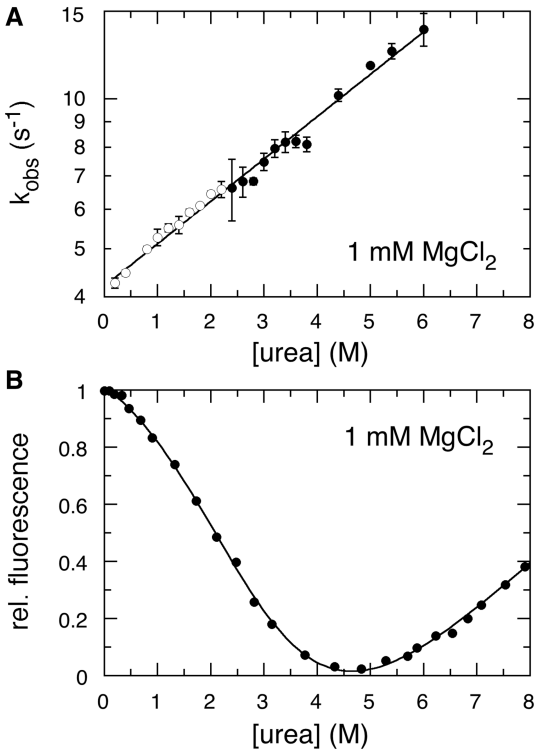

Figure 5. Urea increases the folding rate. (**A**) Urea denaturation of tP5abc in 1 mM Mg^{2+} at 30°C, fit to Equation (2). The increase in fluorescence intensity between 5 and 8 M urea is due to denaturation of the secondary structure and the loss of stacking interactions between 2AP and adjacent bases. (**B**) Observed folding rate at 30°C increases with urea concentration. Errors are the standard deviation for three experiments. Unfolding rates (open circle) were obtained by mixing prefolded RNA (1 mM Mg^{2+}) with urea in CE, 1 mM $MgCl_2$. Folding rates (filled circle) were obtained by diluting denatured RNA in 5.5 M urea with CE plus 1 mM Mg^{2+}.

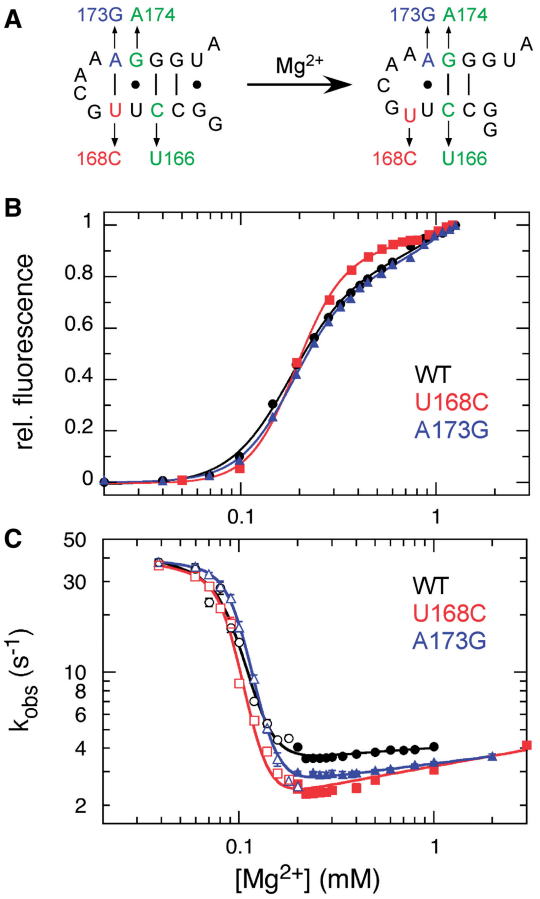

Figure 6. Rearrangement of P5c is not rate-limiting. (**A**) Base substitutions in P5c predicted to destabilize the extended conformation (left) without destabilizing the native structure (right). (**B**) Equilibrium Mg^{2+} titrations of wild-type tP5abc (black, filled circle), U168C (red, filled rhombus), A173G (blue, filled triangle) and G174A, C166U (green). Data were fit to Equation (1), which accounts for changes in cooperativity as well as midpoint (Supplementary Table S2). (**C**) Folding kinetics in Mg^{2+} at 30°C. Folding rates were measured as in Figure 4B and fit to Equation (3) (Supplementary Table S3). Open symbols, unfolding; filled symbols, folding conditions.

Mg^{2+}-independent slow step in the folding pathway, as similar results were obtained when the tP5abc folding kinetics were measured in the presence of 60 mM NaCl (Figure 4C and D and Supplementary Table S3).

Structural reorganization during folding

To estimate how many interactions must reorganize during folding, we used an Arrhenius plot to measure the activation energy for folding in 0.5 mM Mg^{2+}, which was 21.3 kcal/mol (Supplementary Figure S5). This activation energy is comparable to ΔH for opening five base pairs based on empirical free energy parameters (35). Next, we measured the folding equilibrium and kinetics of tP5abc in urea. In 1 mM Mg^{2+}, urea denaturation of native tP5abc at equilibrium had a slope of $m = 0.52$ kcal/mol•M (Figure 5A). Based on studies of RNA duplexes and tRNA, this sensitivity to urea is comparable to the loss of four to five base pairs or the exposure of 2900 Å^2 of buried surface (33). Strikingly, urea strongly increased the folding rate under both native and denaturing conditions (Figure 5B). From these results, we concluded that tP5abc refolds through one or more intermediates and that formation of the intermediate requires the disruption of interactions in the extended state. However, we show below that this reorganization is not due to opening of P5c.

Switching of the P5c helix

Previous studies showed that mutations which favor the extended conformation of P5c inhibit base pair switching and tertiary folding, while mutations that favor the native secondary structure of P5c stabilize the native state (17). We used mutations that favor the native form of P5c to determine whether P5c switching slows down refolding of tP5abc (Figure 6A). These included U168C, A173G and the double mutation C166U, G174A, all of which disrupt base pairs in the extended P5c helix while maintaining the native base pairs. The mutant RNAs were slightly more stable than the parental RNA (Figure 6B and Supplementary Table S2), confirming a very small shift in their folding equilibria toward the native structure.

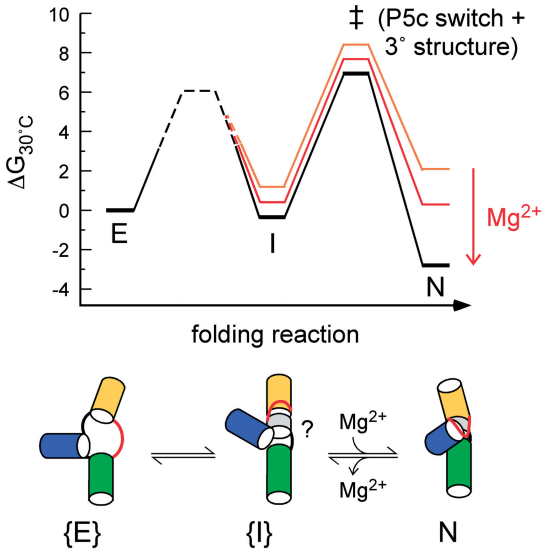

Figure 7. Model for refolding of tP5abc. The folding pathway from the extended state ensemble $\{E\}$ to the folded state N involves at least one metastable intermediate ensemble, $\{I\}$, which requires partial reorganization of interactions in E and a second step that represents the Mg^{2+}-dependent transition state (‡). The switch in P5c occurs in the second step of the folding reaction, together with the formation of tertiary interactions between P5c and the A-rich bulge. Black, native conditions (1 mM Mg^{2+}); red, midpoint (0.18 mM Mg^{2+}); orange, unfolded conditions (0.05 mM Mg^{2+}). The landscape is based on the equilibrium stability of native tP5abc and simulations of experimental folding kinetics (Berkeley Madonna) assuming an intrinsic reaction time $\tau_0 = 1\,\mu s$. The barrier between E and I was poorly constrained (dashed black line).

Although we expected the mutant tP5abc RNAs to fold faster because the native P5c helix is more preferred, they refolded at a similar rate as the parental tP5abc ($3\,s^{-1}$ versus. $4\,s^{-1}$ in 1 mM Mg^{2+}; Supplementary Table S3). About 15% of the C166U, G174A mutant refolded at a much slower rate ($0.64\,s^{-1}$; Supplementary Figure S6). Like the parental tP5abc, the folding rates did not increase with Mg^{2+} (Figure 6C). Thus, mutations that favor the native secondary structure of P5c failed to overcome the rate-limiting step of folding. From these results, we concluded that the overall folding rate is determined by reorganization of other interactions within the extended P5abc structure and not the P5c switch. In the subsequent sections, we discuss how the results of these experiments and our simulations can be rationalized by coupled formation of tertiary interactions and helix switching.

DISCUSSION

How secondary and tertiary interactions are coupled in RNA remains an important question, as many RNAs such as riboswitches and even some ribozymes use helix switching to drive regulation or substrate recognition (36,37). When P5abc RNA folds, its secondary structure rearranges to a base pairing scheme that is less energetically favorable but that positions certain bases to participate in tertiary interactions (6). Mg^{2+} ions stabilize the tertiary interactions both directly and indirectly, compensating for the energetic cost of reorganizing the base pairs and maintaining a net favorable folding free energy (17,18). Thus, remodeling of the RNA secondary structure is also coupled to the specific coordination of metal ions.

Tertiary folding drives helix switching

Remarkably, both the simulated equilibrium folding landscapes and the folding kinetics of P5c mutants suggest that P5c switches to its native form just before or concurrently with formation of tertiary interactions in the A-rich bulge (Figure 7). Not only is the native P5c helix predicted to form in the same thermal transition as the tertiary structure (Figure 3) but mutations that destabilize the extended form of P5c (and thus favor the native P5c) failed to speed folding (Figure 6), as would be expected if this switch happens before the tertiary structure forms. The folding kinetics also suggests that P5c switching occurs after the slow step of folding, which itself may involve local tertiary interactions around the three-helix junction (see below). Both aspects of our model are consistent with the proposal that the tertiary interactions directly promote the secondary structure rearrangement (18): the first step creates an intermediate conducive to switching, while the second step couples switching to tertiary interactions between P5c and the A-rich bulge. Ion concentrations, temperature and mutations, which modulate the folding landscapes of P5abc, can be used to control the extent to which tertiary interactions determine P5c switching.

In our model, the energetically costly P5c switch occurs very late in the folding landscape (Figure 7). As shown in our simulations, the native P5c helix is much less stable than the extended form and thus has little probability of forming in the extended state. Thus, if the P5c switch occurred early, the activation barrier to folding would be very large, and P5abc might never fold. Instead, we suggest earlier secondary and tertiary folding steps couple the change in secondary structure to the formation of tertiary interactions in the A-rich bulge, guiding the RNA through a lower energy path to the native state.

Interestingly, the double mutation G174A, C166U, which stabilizes the native P5c helix relative to the extended form, results in biphasic folding kinetics. This could be due to the presence of an alternative folding path in which P5c switches before the rate-limiting conformational change. Such a possibility is observed in the simulated folding landscapes (Figure 3B).

Folding involves a metastable intermediate

The tP5abc refolding kinetics in Mg^{2+} were simulated by a two-step folding pathway involving at least one intermediate (I), consistent with high thermal activation energy of Mg^{2+}-induced refolding and the acceleration of folding by urea (Figure 7). Small differences in the observed rate constants of folding and unfolding reactions near the Mg^{2+} midpoint (Figure 4B) suggest these pathways sample different transition state ensembles consistent with this model.

If P5c switching is not rate-limiting, folding must involve another structural change leading to the I state. This step could involve proper base pairing of P5a, which our simulations show depends on prior formation of the internal P5b and P5c helices in the extended state.

The *I* state may also involve organization of the three-helix junction, including the tandem G-A pairs in P5b which are critical to the native tertiary structure. Substitution of G163 with 2,6-diaminopurine prevented tP5abc RNA from folding (Eda Koculi unpublished data), consistent with its participation in an extensive tertiary hydrogen bonding network. We considered whether the *I* state and the transition from *I* to *N* arises from reorganization of the tertiary interactions after (rather than before) P5c switching. However, this model does not explain why folding is insensitive to Mg^{2+} concentration and is inconsistent with our simulations and thermal melting results.

An extended transition state

The relative position of the transition state ensemble along the reaction coordinate can be evaluated by comparing the sensitivity of the transition state free energy to Mg^{2+} ($n^{\ddagger}$) to that of the overall reaction free energy (n_H), which is expressed as $\beta = n^{\ddagger}/n_H$ (38,39). In unfolding conditions, $\beta_U = 0.7–1$, consistent with a transition state ensemble similar to the unfolded state, while under folding conditions, $\beta_F = 0$, which is also consistent with a transition state close to the unfolded state (Supplementary Table S3). These observations suggest that the transition states for both folding and unfolding are more open or extended than the native state and do not involve strong, site-specific coordination of metal ions. By contrast, the hairpin ribozyme folds through a compact transition state that is stabilized by Mg^{2+} similar to the native state (40). Thus P5abc, which must find a low energy path for helix switching, appears to adopt a different folding strategy than RNAs which fold via quasi-parallel docking of two helical domains.

Effect on ribozyme folding

Time-resolved footprinting experiments showed that the P5abc subdomain is the fastest region of the *Tetrahymena* ribozyme to fold (15). Therefore, the P5abc region likely follows a similar refolding and helix switching pathway in its native context as it does in isolation. However, interactions with other ribozyme sequences modulate formation of tertiary interactions in P5abc, and thus by extension, the rate of P5c switching. First, although the P5abc domain is marginally stable in isolation, extensive tertiary interactions between P5abc and the core of the *Tetrahymena* ribozyme stabilize the P5abc subdomain in its native context (41). Second, folding of P5abc is retarded in the wild-type ribozyme by premature docking against helices P4 and P6, suggesting that non-specific docking inhibits reorganization of P5c (34). Third, circular permutation of the ribozyme so that P5abc is near the 5′-end of the RNA causes P5abc to fold more rapidly than it does in the wild-type ribozyme (42).

Although P5abc folds more rapidly when it is decoupled from other domains of the ribozyme, the wild-type connections improve folding of the ribozyme core (42,43). Not only does the tertiary structure of P5abc favor the active form of the ribozyme (13,44) but P5c switching is required for the long-range P14 kissing loop interaction between L5c and L2, which stabilizes the ribozyme core (45). As P14 impedes reorganization of the misfolded ribozyme core (46,47), the P5c switch might improve folding overall by delaying formation of P14 (12,48). Our work suggests that the switch kinetics is itself sensitive to alignment of helices in the ribozyme core, perhaps providing feedback on the quality of the fold.

CONCLUSION

In conclusion, our simulations and experiments show that reorganization of the P5abc RNA secondary structure is tightly coupled to the formation of tertiary interactions, which are in turn stabilized by specifically bound Mg^{2+} ions. Our results strongly suggest that these steps are kinetically linked and possibly concomitant, illustrating how coupled formation of secondary and tertiary structure in RNA can open up new, lower energy pathways to the native structure. Although simultaneous secondary and tertiary folding violates the hierarchy often assumed for RNA folding, it is commonly seen among proteins. Our results further imply that non-native interactions in the extended state aid folding rather than impede it. Because helix switching and P5abc tertiary interactions occur in the same step, this refolding pathway may improve coordination between the stabilizing P14 interaction and proper docking of helices in the ribozyme core. Elucidation of the details of the folding mechanism and the extent of alternative folding pathways will require a combination of ensemble and single molecule experiments and well-designed simulations.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–3 and Supplementary Figures 1–6.

ACKNOWLEDGEMENTS

The authors thank M. Deras and A. Werner for initial fluorescence experiments on tP5abc, I. Tinoco for the coordinates of the extended tP5abc, A. Soto and D. Draper for assistance with UV denaturation experiments and D. Barrick for helpful discussions.

FUNDING

National Institutes of Health (NIH) [R01 GM46686]; National Science Foundation [CHE 09-14033]; a Ruth L. Kirschstein National Research Service Award (to S.S.C.). Funding for open access charge: NIH.

Conflict of interest statement. None declared.

REFERENCES

1. Thirumalai,D. and Hyeon,C. (2005) RNA and protein folding: common themes and variations. *Biochemistry*, **44**, 4957–4970.
2. Sosnick,T.R. (2008) Kinetic barriers and the role of topology in protein and RNA folding. *Protein Sci.*, **17**, 1308–1318.

3. Treiber,D.K. and Williamson,J.R. (1999) Exposing the kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.*, **9**, 339–345.
4. Adams,A., Lindahl,T. and Fresco,J.R. (1967) Conformational differences between the biologically active and inactive forms of a transfer ribonucleic acid. *Proc. Natl Acad. Sci. USA*, **57**, 1684–1691.
5. Andersen,A.A. and Collins,R.A. (2001) Intramolecular secondary structure rearrangement by the kissing interaction of the Neurospora VS ribozyme. *Proc. Natl Acad. Sci. USA*, **98**, 7730–7735.
6. Wu,M. and Tinoco,I. Jr (1998) RNA folding causes secondary structure rearrangement. *Proc. Natl Acad. Sci. USA*, **95**, 11555–11560.
7. Gluick,T.C., Gerstner,R.B. and Draper,D.E. (1997) Effects of Mg^{2+}, K^{+}, and H^{+} on an equilibrium between alternative conformations of an RNA pseudoknot. *J. Mol. Biol.*, **270**, 451–463.
8. Puglisi,J.D., Tan,R., Calnan,B.J., Frankel,A.D. and Williamson,J.R. (1992) Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science*, **257**, 76–80.
9. Butcher,S.E., Dieckmann,T. and Feigon,J. (1997) Solution structure of a GAAA tetraloop receptor RNA. *EMBO J.*, **16**, 7490–7499.
10. Crothers,D.M., Cole,P.E., Hilbers,C.W. and Shulman,R.G. (1974) The molecular mechanism of thermal unfolding of *Escherichia coli* formylmethionine transfer RNA. *J. Mol. Biol.*, **87**, 63–88.
11. Tinoco,I.J. and Bustamante,C. (1999) How RNA folds. *J. Mol. Biol.*, **293**, 271–261.
12. 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. USA*, **98**, 3695–3700.
13. van der Horst,G., Christian,A. and Inoue,T. (1991) Reconstitution of a group I intron self-splicing reaction with an activator RNA. *Proc. Natl Acad. Sci. USA*, **88**, 184–188.
14. 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.
15. Sclavi,B., Sullivan,M., Chance,M.R., Brenowitz,M. and Woodson,S.A. (1998) RNA folding at millisecond intervals by synchrotron hydroxyl radical footprinting. *Science*, **279**, 1940–1943.
16. 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.
17. 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.
18. Thirumalai,D. (1998) Native secondary structure formation in RNA may be a slave to tertiary folding. *Proc. Natl Acad. Sci. USA*, **95**, 11506–11508.
19. Woodson,S.A. (2005) Metal ions and RNA folding: a highly charged topic with a dynamic future. *Curr. Opin. Chem. Biol.*, **9**, 104–109.
20. Thirumalai,D., Lee,N., Woodson,S.A. and Klimov,D. (2001) Early events in RNA folding. *Annu. Rev. Phys. Chem.*, **52**, 751–762.
21. 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.
22. Juneau,K., Podell,E., Harrington,D.J. and Cech,T.R. (2001) Structural basis of the enhanced stability of a mutant ribozyme domain and a detailed view of RNA-solvent interactions. *Structure*, **9**, 221–231.
23. Uchida,T., He,Q., Ralston,C.Y., Brenowitz,M. and Chance,M.R. (2002) Linkage of monovalent and divalent ion binding in the folding of the P4-P6 domain of the Tetrahymena ribozyme. *Biochemistry*, **41**, 5799–5806.
24. Das,R., Travers,K.J., Bai,Y. and Herschlag,D. (2005) Determining the Mg2+ stoichiometry for folding an RNA metal ion core. *J. Am. Chem. Soc.*, **127**, 8272–8273.
25. Liphardt,J., Onoa,B., Smith,S.B., Tinoco,I.J. and Bustamante,C. (2001) Reversible unfolding of single RNA molecules by mechanical force. *Science*, **292**, 733–737.
26. Thirumalai,D. and Woodson,S.A. (1996) Kinetics of folding of protein and RNA. *Acc. Chem. Res.*, **29**, 433–439.
27. Hyeon,C. and Thirumalai,D. (2005) Mechanical unfolding of RNA hairpins. *Proc. Natl Acad. Sci. USA*, **102**, 6789–6794.
28. Freier,S.M., Kierzek,R., Jaeger,J.A., Sugimoto,N., Caruthers,M.H., Neilson,T. and Turner,D.H. (1986) Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl Acad. Sci. USA*, **83**, 9373–9377.
29. Honeycutt,J.D. and Thirumalai,D. (1990) Metastability of the folded states of globular proteins. *Proc. Natl Acad. Sci. USA*, **87**, 3526–3529.
30. Chauhan,S. and Woodson,S.A. (2008) Tertiary interactions determine the accuracy of RNA folding. *J. Am. Chem. Soc.*, **130**, 1296–1303.
31. Pan,J., Thirumalai,D. and Woodson,S.A. (1999) Magnesium-dependent folding of self-splicing RNA: exploring the link between cooperativity, thermodynamics, and kinetics. *Proc. Natl Acad. Sci. USA*, **96**, 6149–6154.
32. Pace,C.N. (1975) The stability of globular proteins. *CRC Crit. Rev. Biochem.*, **3**, 1–43.
33. Shelton,V.M., Sosnick,T.R. and Pan,T. (1999) Applicability of urea in the thermodynamic analysis of secondary and tertiary RNA folding. *Biochemistry*, **38**, 16831–16839.
34. Deras,M.L., Brenowitz,M., Ralston,C.Y., Chance,M.R. and Woodson,S.A. (2000) Folding mechanism of the *Tetrahymena* ribozyme P4-P6 domain. *Biochemistry*, **39**, 10975–10985.
35. Mathews,D.H., Sabina,J., Zuker,M. and Turner,D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, **288**, 911–940.
36. Serganov,A. (2009) The long and the short of riboswitches. *Curr. Opin. Struct. Biol.*, **19**, 251–259.
37. Reymond,C., Beaudoin,J.D. and Perreault,J.P. (2009) Modulating RNA structure and catalysis: lessons from small cleaving ribozymes. *Cell Mol. Life Sci.*, **66**, 3937–3950.
38. Leffler,J. (1953) Parameters for the description of transition states. *Science*, **117**, 340–341.
39. Matouschek,A. and Fersht,A.R. (1993) Application of physical organic chemistry to engineered mutants of proteins: Hammond postulate behavior in the transition state of protein folding. *Proc. Natl Acad. Sci. USA*, **90**, 7814–7818.
40. Bokinsky,G., Rueda,D., Misra,V.K., Rhodes,M.M., Gordus,A., Babcock,H.P., Walter,N.G. and Zhuang,X. (2003) Single-molecule transition-state analysis of RNA folding. *Proc. Natl Acad. Sci. USA*, **100**, 9302–9307.
41. Doherty,E.A., Herschlag,D. and Doudna,J.A. (1999) Assembly of an exceptionally stable RNA tertiary interface in a group I ribozyme. *Biochemistry*, **38**, 2982–2990.
42. Lease,R.A., Adilakshmi,T., Heilman-Miller,S. and Woodson,S.A. (2007) Communication between RNA folding domains revealed by folding of circularly permuted ribozymes. *J. Mol. Biol.*, **373**, 197–210.
43. Heilman-Miller,S.L. and Woodson,S.A. (2003) Perturbed folding kinetics of circularly permuted RNAs with altered topology. *J. Mol. Biol.*, **328**, 385–394.
44. Johnson,T.H., Tijerina,P., Chadee,A.B., Herschlag,D. and Russell,R. (2005) Structural specificity conferred by a group I RNA peripheral element. *Proc. Natl Acad. Sci. USA*, **102**, 10176–10181.
45. Lehnert,V., Jaeger,L., Michel,F. and Westhof,E. (1996) New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the Tetrahymena thermophila ribozyme. *Chem. Biol.*, **3**, 993–1009.
46. Pan,J. and Woodson,S.A. (1999) The effect of long-range loop-loop interactions on folding of the Tetrahymena self-splicing RNA. *J. Mol. Biol.*, **294**, 955–965.
47. Russell,R., Zhuang,X., Babcock,H.P., Millett,I.S., Doniach,S., Chu,S. and Herschlag,D. (2002) Exploring the folding landscape of a structured RNA. *Proc. Natl Acad. Sci. USA*, **99**, 155–160.
48. Treiber,D.K. and Williamson,J.R. (2001) Concerted kinetic folding of a multidomain ribozyme with a disrupted loop-receptor interaction. *J. Mol. Biol.*, **305**, 11–21.
49. Leontis,N.B. and Westhof,E. (2001) Geometric nomenclature and classification of RNA base pairs. *RNA*, **7**, 499–512.