Structural specificity conferred by a group I RNA peripheral element

Travis H. Johnson[†], Pilar Tijerina[†], Amanda B. Chadee[†], Daniel Herschlag[‡], and Rick Russell[†§]

[†]Department of Chemistry and Biochemistry, Institute for Cellular and Molecular Biology, University of Texas, Austin, TX 78712; and
[‡]Department of Biochemistry, B400 Beckman Center, Stanford University, Stanford, CA 94305

Edited by Jennifer A. Doudna, University of California, Berkeley, CA, and approved June 8, 2005 (received for review February 22, 2005)

Like proteins, structured RNAs must specify a native conformation that is more stable than all other possible conformations. Local structure is much more stable for RNA than for protein, so it is likely that the principal challenge for RNA is to stabilize the native structure relative to misfolded and partially folded intermediates rather than unfolded structures. Many structured RNAs contain peripheral structural elements, which surround the core elements. Although it is clear that peripheral elements stabilize structure within RNAs that contain them, it has not yet been explored whether they specifically stabilize the native states relative to alternative folds. A two-piece version of the group I intron RNA from *Tetrahymena* is used here to show that the peripheral element P5abc binds to the native conformation of the rest of the RNA 50,000 times more tightly than it binds to a long-lived misfolded conformation. Thus, P5abc stabilizes the native conformation by ≈6 kcal/mol relative to this misfolded conformation. Further, activity measurements show that for the RNA lacking P5abc, the native conformation is only marginally preferred over the misfolded conformation (<0.5 kcal/mol), indicating that the peripheral structure of this RNA is required to achieve a significant thermodynamic preference for the native state. Such ''structural specificity'' may be a general function of RNA peripheral domains.

folding thermodynamics | misfolded intermediate | RNA folding

Structured RNAs, like proteins, face the fundamental challenge of specifying a native three-dimensional structure that is at least as stable as the combination of all possible unfolded and misfolded structures. Although this basic challenge is shared between these classes of molecules, the specific nature of the challenge differs. Many proteins exist in a two-state equilibrium between the native species and an ensemble of conformations that lack extensive structure, so that the principle challenge for proteins is to stabilize the native structure relative to unfolded conformations. In contrast, because local structure for RNA is independently stable, unfolded species are not significantly populated, and therefore the principal challenge for RNA is to stabilize the native species relative to misfolded and partially folded alternative conformations (1–3).

RNA also differs from protein in the strategies used to generate stable tertiary structure. Whereas proteins derive much of their tertiary stability from the packing of nonpolar side chains within a hydrophobic core, RNA is unable to use this strategy. First, its backbone is charged on every residue; second, although its bases have some hydrophobic character (4), their interactions also involve substantial polar components (5, 6). Further, the bases are largely sequestered within duplexes and unavailable for tertiary interactions. As one alternative strategy, many structured RNAs, including ribosomal RNAs, RNase P RNA, and group I and group II introns, contain "peripheral elements," which lie on the outside of the structure and confer tertiary stability by forming long-range contacts to one another and to the core elements (7–14).

The role of peripheral structure in stability and function has been studied extensively in the archetypal group I RNA from *Tetrahymena thermophila*. In several studies using this RNA and other group I RNAs, deletion of individual peripheral elements gave RNAs that retained some catalytic activity but had increased Mg^{2+} requirements for stable folding (15–18). A particularly well-studied peripheral element that gives such behavior upon deletion is P5abc, which is present in the *Tetrahymena* RNA and conserved within the IC1 subgroup of group I introns (Fig. 1). Further, the P5abc element can be added back to the variant RNA as a separate molecule to restore WT activity (19). It forms an extensive interface with the P4-P6 domain mediated by tertiary contacts (20–22) and forms a loop–loop base-pairing interaction with another peripheral element, P2 (9). Upon binding, P5abc was shown by chemical footprinting to tighten packing of the catalytic core (23), leading to greater preorganization of the active site for substrate binding and catalysis (24).

The nature of RNA folding energetics outlined above raises the question of whether the peripheral architecture stabilizes the native species relative to specific misfolded or partially structured intermediates, because such structural specificity may be critical for RNA (3, 25). Although peripheral structure has been shown clearly to stabilize tertiary structure in a general sense, the specificity of this stabilization has not been explored. Here, we use the two-piece *Tetrahymena* RNA resulting from P5abc deletion to probe whether peripheral architecture stabilizes the native species relative to a long-lived misfolded intermediate, which accumulates during folding but is not populated significantly at equilibrium for the WT ribozyme (26–28). We find that P5abc provides an enormous stabilization of the native state relative to the misfolded state, ≈6 kcal/mol under standard conditions, and, surprisingly, in the absence of P5abc, the native state is only marginally more stable than the misfolded state. Thus, peripheral structure can have a profound effect on structural specificity, stabilizing the native species relative to alternative conformations.

Materials and Methods

Materials. WT (L-21/ScaI) ribozyme, *Tetrahymena* ribozyme variant lacking the P5abc peripheral element ($E^{\Delta P5abc}$), and P5abc RNA were prepared by *in vitro* transcription (29) and purified as described in ref. 30. RNA oligonucleotides (Dharmacon Research, Lafayette, CO) were 5′ end-labeled with [γ-^{32}P]ATP (PerkinElmer) by using T4 polynucleotide kinase (New England Biolabs). P5abc and $E^{\Delta P5abc}$ RNAs were 5′ end-labeled after removal of 5′-phosphoryl groups by shrimp alkaline phosphatase (MBI Fermentas, Amherst, NY). $E^{\Delta P5abc}$ RNA was 3′ end-labeled with [α-^{32}P]dATP (PerkinElmer) by using DNA polymerase I (the exo$^-$ Klenow fragment, New England Biolabs) to extend the ribozyme with a DNA oligonucleotide as a template (31). Labeled RNAs were purified by nondenaturing PAGE (29, 32).

Binding and Release of P5abc. Dissociation of ^{32}P-labeled P5abc (*P5abc) from misfolded and native $E^{\Delta P5abc}$ ribozyme was measured by gel mobility shift using a pulse–chase procedure (33). Unless otherwise indicated, reaction conditions were 50 mM Na-

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: $E^{\Delta P5abc}$, *Tetrahymena* ribozyme variant lacking the P5abc peripheral element; S*, 5′-^{32}P-labeled oligonucleotide substrate CCCUCUA$_5$; *P5abc, ^{32}P-labeled P5abc.

§To whom correspondence should be addressed. E-mail: rick_russell@mail.utexas.edu.

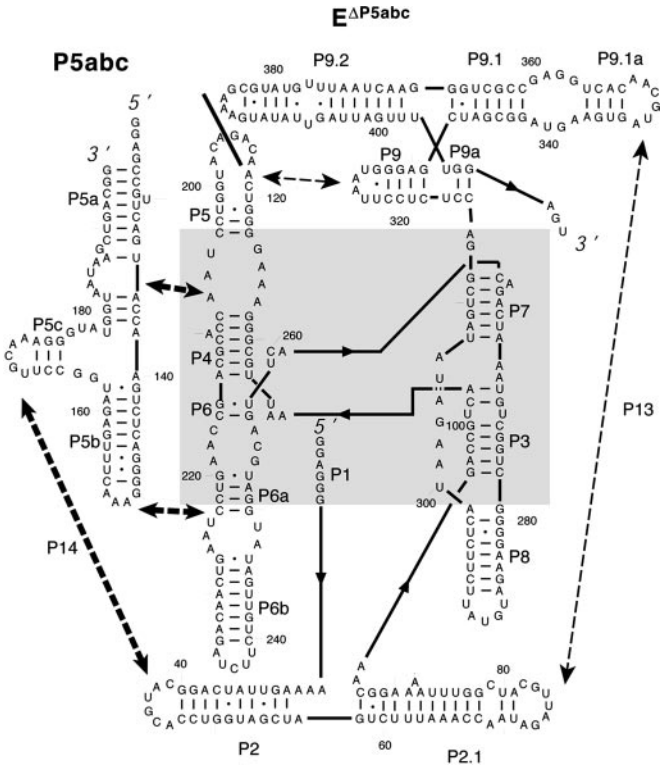

Fig. 1. Two-part ribozyme system consisting of $E^{\Delta P5abc}$ and P5abc. In the WT RNA, P5 and P5a are covalently connected within L5, on either side of the solid line. The core region of the ribozyme is shaded gray. Tertiary contacts of P5abc are shown with thick dashed lines and arrows. Other long-range tertiary contacts are shown with thin dashed lines and arrows.

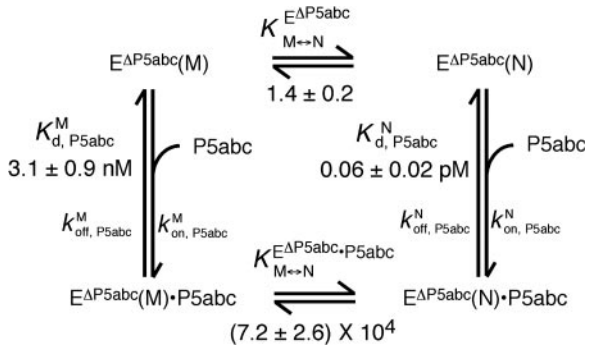

Fig. 2. Thermodynamic cycle of P5abc binding to native and misfolded $E^{\Delta P5abc}$ ribozyme. $E^{\Delta P5abc}$(M) and $E^{\Delta P5abc}$(N) represent misfolded and native $E^{\Delta P5abc}$ ribozyme, respectively. Values of the equilibrium constants $K^{E^{\Delta P5abc}}_{M\leftrightarrow N}$, $K^{M}_{d,\,P5abc}$, and $K^{N}_{d,\,P5abc}$ that were determined in this work are shown (at 25°C, pH 7.0, and 10 mM Mg^{2+}). These values give a calculated value of $(7.2 \pm 2.6) \times 10^4$ for $K^{E^{\Delta P5abc}\cdot P5abc}_{M\leftrightarrow N}$, the equilibrium constant between native and misfolded species in the presence of bound P5abc, which is also shown ($K^{E^{\Delta P5abc}\cdot P5abc}_{M\leftrightarrow N} = K^{E^{\Delta P5abc}}_{M\leftrightarrow N} \times K^{M}_{d,\,P5abc}/K^{N}_{d,\,P5abc}$). The equilibrium constant for conversion between native and misfolded species in the absence of P5abc, $K^{E^{\Delta P5abc}}_{M\leftrightarrow N}$, was calculated by dividing the ''raw'' fraction of native ribozyme (0.52 as determined by enzyme activity; see Fig. 4) by the fraction of misfolded ribozyme, which was inferred to be equal to the remainder of the ribozyme minus a small fraction that is unreactive, presumably from damage (Fig. 4); $(1 - 0.52 - 0.12 = 0.36)$ Thus, $K^{E^{\Delta P5abc}}_{M\leftrightarrow N}$ is calculated as $0.52/0.36 = 1.4$. The equilibrium binding constants, $K^{M}_{d,\,P5abc}$ and $K^{N}_{d,\,P5abc}$, were calculated from rate constants for dissociation and association of P5abc and the $E^{\Delta P5abc}$ ribozyme. Dissociation rate constants were $k^{M}_{off,\,P5abc} = 0.026\ min^{-1}$ (Fig. 3*A*) and $k^{N}_{off,\,P5abc} = 1.1 \times 10^{-6}\ min^{-1}$ (Fig. 3*C*). Association rate constants were measured for populations of ≈60% and ≈15% native ribozyme (Fig. 7). Corrections for the mixed populations gave the following values for P5abc association with the pure native and misfolded species: $k^{N}_{on,\,P5abc} = 1.9 \times 10^7\ M^{-1}\cdot min^{-1}$ and $k^{M}_{on,\,P5abc} = 0.8 \times 10^7\ M^{-1}\cdot min^{-1}$.

Mops (pH 7.0) and 10 mM $MgCl_2$ at 25°C. Ribozyme was briefly incubated (10 mM Mg^{2+} at 25°C) to give predominantly misfolded ribozyme [up to ≈85% (26)] or incubated longer (30 min at 50°C or ≥3 h at 25°C) to give predominantly native ribozyme (ref. 34 and Fig. 4*A*). *P5abc was bound to 100 nM $E^{\Delta P5abc}$ for 5 min, at which time a 10-fold excess of unlabeled P5abc over $E^{\Delta P5abc}$ was added. The fraction of *P5abc bound to $E^{\Delta P5abc}$ over time was determined by 12% nondenaturing PAGE (run at 4°C) and quantitated by using a phosphorimager (Molecular Dynamics).

Rate constants for *P5abc binding to native and misfolded ribozyme were measured by using a similar pulse–chase procedure (34). *P5abc was added to populations of predominantly misfolded or native $E^{\Delta P5abc}$ (50–200 nM, generated as above). At various times, unlabeled P5abc (2.5 μM) was added to block further binding of *P5abc. Samples were immediately loaded on a 12% nondenaturing gel and processed as above. The rate of P5abc binding to native ribozyme was monitored by activity as described in ref. 26.

Activity Assays to Follow Refolding of Misfolded $E^{\Delta P5abc}$ Ribozyme. The fraction of ribozyme present in the native state was determined from the fraction of 5′-^{32}P-labeled oligonucleotide substrate $CCCUCUA_5$ (S*) that was cleaved in a burst, as described in ref. 27. $E^{\Delta P5abc}$ was incubated to generate misfolded ribozyme as above. At various times, aliquots were added to a solution containing guanosine and excess P5abc. [It was necessary to add P5abc so that the fraction of native $E^{\Delta P5abc}$ ribozyme could be determined by catalytic activity (19, 23, 24). P5abc also slows refolding of the misfolded ribozyme (ref. 28 and unpublished data), providing a convenient quench of the refolding reaction.] S* was added for 1 min, at which time quench solution was added (two volumes of 70% formamide/20 mM EDTA/0.1% bromophenol blue/0.1% xylene cyanol). Labeled substrate and product were separated by 20% denaturing PAGE (29) and quantitated by using a phosphorimager. The fraction of S* present as product was plotted against refolding time to give the rate constant and endpoint for refolding to the native state. The fraction of S* that is cleaved rapidly was shown previously to provide a good measure of the fraction of native ribozyme for the WT (27), and we demonstrated this relationship for the $E^{\Delta P5abc}$ ribozyme as well (see *Supporting Text*, which is published as supporting information on the PNAS web site).

Fe(II)-EDTA Footprinting. Footprinting was performed essentially as described in refs. 23 and 35 (see *Supporting Text*). Quantitation was performed by using the software SAFA (SemiAutomated Footprinting Analysis) (36).

Results

P5abc Binds More Tightly to Native Ribozyme than to Misfolded Ribozyme. To probe whether peripheral structure stabilizes the native conformation of the ribozyme relative to the misfolded conformation, we used the two-part RNA composed of the ribozyme lacking P5abc ($E^{\Delta P5abc}$) and the separate P5abc RNA (Fig. 1). We chose this peripheral element because prior work showed that $E^{\Delta P5abc}$ can achieve a native state (17), but, like the WT ribozyme, $E^{\Delta P5abc}$ populates a long-lived misfolded species during folding (26). Further, the ability of P5abc to bind to $E^{\Delta P5abc}$ as a separate molecule (19) would allow us to address whether its presence stabilizes the native state by using the experimental approach depicted in Fig. 2. The two processes depicted (P5abc binding to the ribozyme and refolding between the native and misfolded conformations) constitute a thermodynamic cycle. Because the total free-energy change of the cycle must be zero, if P5abc binds preferentially to either native or misfolded ribozyme, its presence must stabilize that species, and the magnitude of the stabilization must be equal to the difference in the binding free

energy. Therefore, we set out to determine the equilibrium binding constants of P5abc to the native and misfolded conformers of the ribozyme. Because the complex of P5abc with the native ribozyme is extraordinarily tight (23) and the misfolded ribozyme refolds to the native state on the same timescale that would be required for equilibration of P5abc binding (shown here), it is not feasible to measure binding of P5abc to the native and misfolded species directly in equilibrium experiments. We therefore measured the kinetics of P5abc binding and release to the native and misfolded conformers and calculated the equilibrium dissociation constants from these kinetic parameters.

We followed dissociation of P5abc from native and misfolded $E^{\Delta P5abc}$ by using a gel mobility shift assay as described in ref. 23. *P5abc was bound to $E^{\Delta P5abc}$, and then excess "chase" P5abc was added to make dissociation of *P5abc essentially irreversible (33). To distinguish dissociation of *P5abc from native and misfolded ribozyme, we systematically varied the fractions of misfolded and native ribozyme by varying the incubation time of $E^{\Delta P5abc}$ before addition of *P5abc (15 min to 3 h). Because the misfolded ribozyme refolds to the native state on this time scale (see Fig. 4*A*), longer incubations would give larger fractions of native ribozyme.

Dissociation of *P5abc from a population of 80% misfolded ribozyme gave a rate constant of 0.026 min^{-1} for the major phase, suggesting that this rate constant reflected dissociation from misfolded ribozyme (Fig. 3*A*). A fraction of the *P5abc did not dissociate on this time scale, indicating that it was present in a much longer-lived complex. Increasing the preincubation time, and therefore the fraction of native ribozyme, increased the fraction of *P5abc that did not dissociate on the experimental time scale, suggesting that the long-lived complex represented *P5abc bound to the native ribozyme. Further supporting this idea, a plot of the fraction of *P5abc that did not dissociate against preincubation time gave a rate constant of 0.013 min^{-1} (Fig. 3*B*), similar to the rate constant for refolding of misfolded ribozyme to the native state, as measured by enzymatic activity (0.019 min^{-1}; Fig. 4*A*). This result provides strong support for the assignment of the 0.026-min^{-1} phase as representing dissociation from misfolded ribozyme and the endpoint as representing the fraction of *P5abc bound to native ribozyme. Surprisingly, even with preincubation times that should be long enough to allow complete refolding of misfolded $E^{\Delta P5abc}$ to the native state, the early phase of dissociation with a rate constant of ≈0.02 min^{-1} was observed. This finding introduced the possibility that a significant fraction of the $E^{\Delta P5abc}$ ribozyme remains misfolded at equilibrium (see text under the heading "An Equilibrium Mixture of Native and Misfolded $E^{\Delta P5abc}$ Ribozyme").

To determine the rate constant for dissociation of P5abc from the native ribozyme, we sought conditions that would give observable dissociation of P5abc. Although there was a small decrease in the fraction of *P5abc bound under standard conditions at 26 h (0.2%), giving a calculated value of 1.3×10^{-6} min^{-1}, the change was so small that the calculated rate constant had considerable uncertainty. Further, if the true fraction of P5abc remaining bound was only slightly larger than our measured value, it could reflect much slower dissociation than calculated (because it would be possible that essentially no P5abc had dissociated beyond the "background" that was unbound at the start of the experiment). We therefore considered the calculated rate constant to approximate an upper limit for the true rate constant. Increased temperature gave more rapid P5abc dissociation, such that it could be easily measured, and the rate constants gave a linear Arrhenius plot (Fig. 3*C*). Extrapolation from data collected between 30°C and 50°C gave a value of 1.1×10^{-6} min^{-1} at 25°C, in the same range as that obtained from direct measurement. Thus, under these conditions, P5abc dissociates from misfolded ribozyme ≈20,000-fold faster than from native ribozyme (0.026 min^{-1} vs. 1.1×10^{-6} min^{-1}).

To determine whether the slower dissociation of P5abc from native ribozyme reflects stronger equilibrium binding, it was necessary to measure the association rate of *P5abc to the native and misfolded species. Rate constants for P5abc binding to populations of predominantly native or misfolded ribozyme were similar [$(1.5 \pm 0.3) \times 10^7$ M^{-1}·min^{-1} and $(1.0 \pm 0.2) \times 10^7$ M^{-1}·min^{-1}, respectively; Fig. 7, which is published as supporting information on the PNAS web site] and similar to a value determined previously at 37°C using the onset of enzymatic activity (26). Corrections for the mixed populations present in the association rate determinations gave rate constants for P5abc binding to pure populations of native

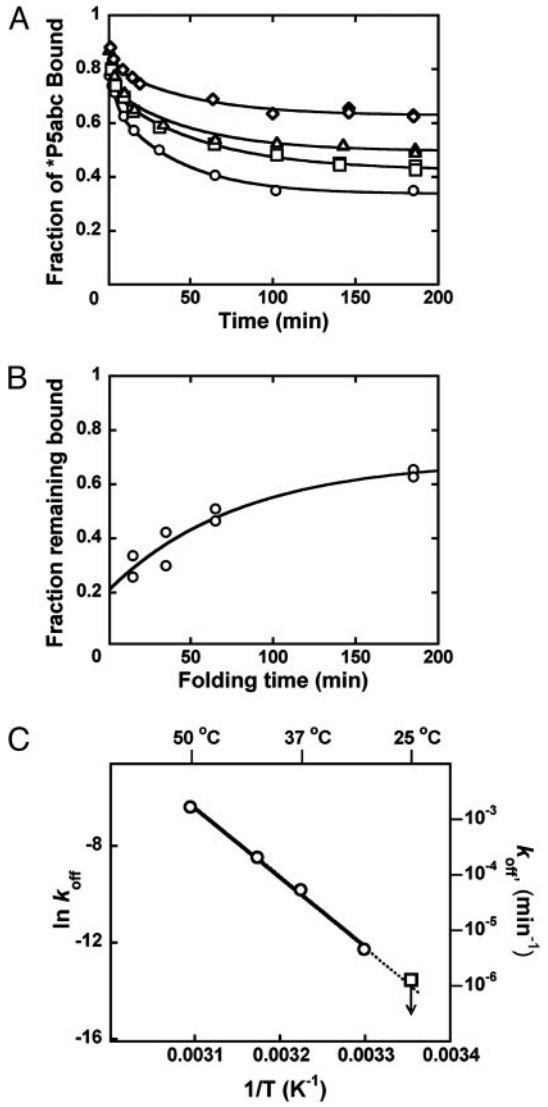

Fig. 3. Kinetics of P5abc dissociation. (*A*) $E^{\Delta P5abc}$ RNA was preincubated with 10 mM Mg^{2+} for 15 (○), 35 (□), 65 (△), or 185 (◇) min to generate mixtures of native and misfolded ribozyme. Rate constants for *P5abc dissociation and endpoints were as follows: 0.025 min^{-1} and 0.34 (○); 0.017 min^{-1} and 0.42 (□); 0.032 min^{-1} and 0.51 (△); and 0.020 min^{-1} and 0.63 (◇). Each curve also displayed a more rapid phase of *P5abc dissociation (0.3 min^{-1}, amplitude 0.1–0.2), which is probably the result of a small fraction of damaged ribozyme or additional folding intermediates. (*B*) Increase in fraction of native ribozyme. The fraction of *P5abc that remained bound at the end of the 0.02-min^{-1} kinetic phase in *A* is plotted against preincubation time (see *Materials and Methods*). The data shown are from the experiment in *A* and from an identical independent experiment. The increase in endpoint gave a rate constant of 0.013 min^{-1}, the same within error as that determined for equilibration of native and misfolded ribozyme by measuring catalytic activity (see Fig. 4). (*C*) Dissociation of *P5abc from native ribozyme. Data between 30°C and 50°C (○) were used to extrapolate to 25°C (dotted line). The value obtained as an upper limit by direct measurement at 25°C (see *Results*) is also shown (□).

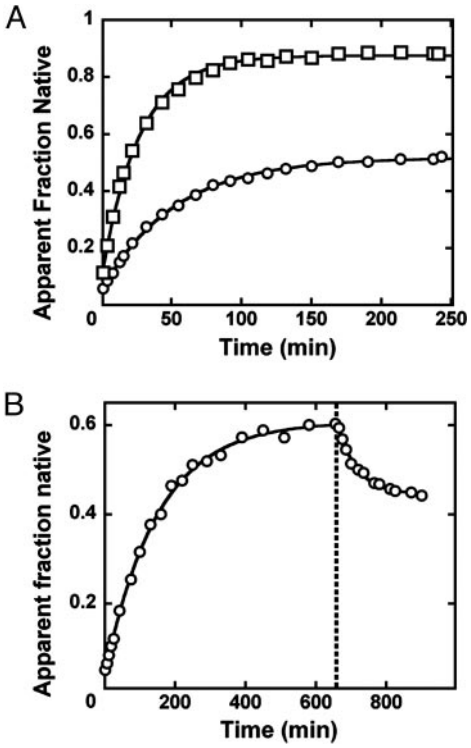

Fig. 4. Approach to equilibrium of native and misfolded ribozyme. (*A*) The fraction of native ribozyme determined by catalytic activity (see *Materials and Methods*) is plotted as a function of folding time in the absence (○) or presence (□) of 2.8 μM P5abc. The reaction in the presence of P5abc also contained 0.2 M Na^+, which accelerates the equilibration of native and misfolded ribozyme, allowing the data to be shown on the same scale. The presence of 0.2 M Na^+ had no effect on the endpoint in the presence or absence of P5abc (data not shown). The data gave rate constants for approach to equilibrium of 0.019 ± 0.003 min^{-1} in the absence of P5abc and 0.037 min^{-1} in the presence of P5abc and 0.2 M Na^+. (*B*) Approach to equilibrium of native and misfolded ribozyme upon Mg^{2+} dilution. After equilibrium was reached in the presence of 50 mM Mg^{2+} (0.007 min^{-1}, 0.61 endpoint), Mg^{2+} was diluted to 10 mM (dotted line), and approach to the new equilibrium was monitored by catalytic activity. The new equilibrium value (0.47) was approached with a rate constant of 0.019 ± 0.001 min^{-1}, the same as that for approach to equilibrium after initial misfolding (0.019 ± 0.003 min^{-1}, *A*).

and misfolded ribozyme of 1.9×10^7 $M^{-1}\cdot min^{-1}$ and 0.8×10^7 $M^{-1}\cdot min^{-1}$, respectively. Combining the slightly faster binding of P5abc to the native ribozyme with the much slower dissociation from native ribozyme leads to the conclusion that P5abc binds ≈50,000-fold more tightly to native than misfolded ribozyme, a free-energy difference of ≈6 kcal/mol.

Kinetic Complexity from Multistep Binding? RNA complexes often form by means of mechanisms described as "induced fit," in which complex formation is accompanied by conformational changes in one or both binding partners (37, 38). Such multistep complex assembly necessitates the formation of assembly intermediates, which in some cases can be long-lived and therefore may accumulate. Accumulation and experimental trapping of such an intermediate can cause the equilibrium dissociation constant calculated from the kinetics of binding and release to be smaller (i.e., indicating tighter binding) than the true value, as has recently been demonstrated for formation of certain RNA–protein complexes (39, 40). This phenomenon may also occur for RNA–RNA complexes, which typically require conformational rearrangements and intermediates (41–43).

To determine whether long-lived intermediates accumulate as P5abc binds the ribozyme, we measured the rate of formation of the active complex of P5abc with the native ribozyme by following the onset of enzymatic activity. If assembly required formation of an intermediate that was long-lived enough to be trapped experimentally, its lifetime would limit the rate of formation of the active complex to $<\approx 0.5$ min^{-1} (corresponding to the time required for the sample to be loaded and to enter the gel matrix). However, the rate constant increased linearly with P5abc concentration to >3 min^{-1}, ruling out the possibility of an intermediate that is trapped experimentally (data not shown). Thus, the 6 kcal/mol tighter binding of P5abc to native ribozyme indicates that P5abc stabilizes the native conformation relative to the misfolded conformation by the same magnitude.

Native Stabilization by P5abc Under a Range of Conditions. We measured the kinetics of P5abc binding and release from the native and misfolded $E^{\Delta P5abc}$ ribozyme under a range of conditions, including varying Mg^{2+} concentration (2–20 mM), Na^+ concentration (20–420 mM in a background of 10 mM Mg^{2+}), and temperature (25–37°C). Under all conditions, P5abc dissociated from the misfolded ribozyme at least 350-fold faster than from the native ribozyme (Table 1, which is published as supporting information on the PNAS web site). The differences were larger with higher Mg^{2+} concentration, lower Na^+ concentration, and lower temperature. Under all conditions that allowed measurement of association rate constants, only small differences were observed between native and misfolded ribozyme (Table 1). Thus, P5abc stabilizes the native conformation by at least 3.5 kcal/mol under a broad range of conditions.

An Equilibrium Mixture of Native and Misfolded $E^{\Delta P5abc}$ Ribozyme. For the WT ribozyme, although the misfolded conformation accumulates during folding, the native state is thermodynamically favored such that refolding to the native state is essentially complete (26, 27). However, the large stabilization of the native conformation by P5abc and the fast phase of *P5abc dissociation from "prefolded" $E^{\Delta P5abc}$ (Fig. 3*B*) introduced the possibility that a fraction of the $E^{\Delta P5abc}$ ribozyme remains misfolded at equilibrium. To explore this possibility, we followed refolding of the misfolded $E^{\Delta P5abc}$ ribozyme by using enzymatic activity (Fig. 4*A*) (27). We measured the fraction of native ribozyme over time by adding S* to aliquots of the reaction (see *Materials and Methods*). The fraction of S* that was rapidly cleaved increased with refolding time (Fig. 4*A*), indicating net formation of native ribozyme. However, instead of proceeding to completion, the refolding reaction gave a maximal fraction of S* cleavage of ≈0.5, suggesting that the equilibrium constant between native and misfolded species is near unity. Nevertheless, there were three alternative possibilities, which are ruled out, in turn, below. First, this low endpoint could have arisen if a substantial fraction of the ribozyme was damaged and therefore nonreactive. Second, the apparent endpoint might not have been a true endpoint; instead, it

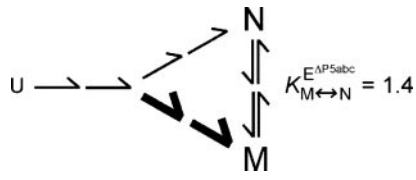

Fig. 5. Minimal folding scheme for the $E^{\Delta P5abc}$ ribozyme. Upon addition of Mg^{2+}, the ribozyme folds from an ensemble of conformations that have secondary structure but lack defined tertiary structure (U). The ribozyme partitions during folding between pathways to the misfolded and native species (M and N), with most of the ribozyme folding to M (indicated by thick arrow). The partitioning is shown as occurring during folding, as shown for the WT ribozyme (27), but the commitment point between pathways has not been determined for the $E^{\Delta P5abc}$ ribozyme. The N and M species slowly reach an equilibrium that slightly favors N (K = 1.4). Folding steps are shown with multiple arrows to emphasize the likelihood of multiple steps and intermediates.

might have reflected the formation of an additional misfolded intermediate that did not refold to the native state significantly on the time scale of the experiment. Third, the endpoint could have represented an equilibrium between native and unfolded species rather than between native and misfolded species.

We addressed the possibility of damaged ribozyme by performing an analogous refolding experiment in the presence of P5abc (Fig. 4*A*). The endpoint of this reaction was substantially higher (0.88), providing strong evidence against the possibility that the lower endpoint above was due to a large fraction of nonreactive ribozyme. Further, this increase in endpoint is predicted from the model that P5abc stabilizes the native state.[¶] Addition of P5abc to ribozyme that had first been folded to the 0.5 endpoint gave a similar increase in the reactive fraction (data not shown). This increase provided a direct visualization of the native stabilization by P5abc and ruled out the possibility that the low endpoint in the absence of P5abc arose because of damage to the RNA that occurred during the refolding reaction.

To determine whether the low endpoint was caused by a longer-lived misfolded species for $E^{\Delta P5abc}$, we took advantage of the observation that increased Mg^{2+} concentration gave a small but reproducible increase in the endpoint accumulation of native ribozyme (Fig. 4*B*). If the low endpoint under standard conditions indeed reflected establishment of an equilibrium, we reasoned that it should be possible to approach that equilibrium from either side. We therefore allowed the ribozyme to fold to its higher endpoint at elevated Mg^{2+} concentration and then diluted back to the standard condition of 10 mM Mg^{2+}. The fraction of native ribozyme decreased to a final value of 0.47. This value is nearly the same as that obtained upon initial folding (Fig. 4*A*), consistent with formation of an equilibrium between species and inconsistent with formation of a kinetically trapped intermediate. A trapped intermediate would be less stable than the native species, by definition, and so would not be expected to increase in concentration relative to the native species. Further, the rate constant for approach to this new equilibrium (0.019 ± 0.001 min^{-1}) was the same as that for initial refolding from the misfolded species, as would be predicted for a simple equilibration between two species (44). A complementary experiment in which Mg^{2+} concentration was increased from 10 to 50 mM gave the expected increase in the fraction of native ribozyme (Fig. 8, which is published as supporting information on the PNAS web site), providing further support for this model.

Finally, we considered the possibility that the low endpoint of refolding for $E^{\Delta P5abc}$ represented establishment of an equilibrium between the native ribozyme and unfolded or partially folded intermediates. It was shown previously by thermal denaturation and Fe(II)-EDTA footprinting that $E^{\Delta P5abc}$ attains its global fold between 5 and 10 mM Mg^{2+} at 37°C (23, 24), providing strong evidence against the possibility of a significant fraction of unfolded or partially folded ribozyme under the conditions of our experiments (25°C and 10 mM Mg^{2+}). Nevertheless, to rule out this possibility, we performed Fe(II)-EDTA footprinting under the conditions of the activity measurements (Fig. 9, which is published as supporting information on the PNAS web site). Several regions of the ribozyme were protected from radical-induced cleavage in the presence of Mg^{2+}; other regions gave enhanced cleavage. All of the observed changes in accessibility occurred with $K_{1/2} < 5$ mM Mg^{2+}, providing further evidence against unfolded or partially folded intermediates accounting for the low folding endpoint.

[¶]The stabilization of the native state predicted from the 50,000-fold tighter binding of P5abc to the native ribozyme would give an endpoint of essentially 1, significantly higher than the observed 0.88 endpoint. This difference presumably arises from a small fraction of damaged ribozyme (≈12%) that binds oligonucleotide substrate $CCCUCUA_5$ but does not give cleavage. We favor this model over the alternative (that this remaining fraction forms an additional long-lived misfolded conformer) because the nonreactive fraction is consistent between experiments for a given ribozyme preparation but varies between experiments from different preparations (from 10% to 25%; unpublished data).

The Same Misfolded Species Accumulates During Folding and Persists at Equilibrium. Together, the data suggested a simple model in which $E^{\Delta P5abc}$ initially partitions between folding to the native and misfolded species, with ≈85–90% misfolding under standard conditions (26), and then slowly refolds to give a nearly equal mixture of the two species (Fig. 5). For this model to hold, however, the misfolded intermediate that accumulates early in folding must be the same species that is populated at equilibrium. Supporting this model, dissociation of P5abc from misfolded ribozyme gave the same rate constant regardless of whether refolding was allowed to proceed long enough to reach equilibrium (Fig. 3*A*), suggesting that these misfolded species were the same. As an independent test of the model, we generated two populations of $E^{\Delta P5abc}$ ribozyme. In one, the ribozyme was folded with Mg^{2+} for a brief time to allow accumulation of the kinetic misfolded intermediate; in the other, the ribozyme was folded long enough to allow accumulation of the equilibrium intermediate. We added excess P5abc to each population and measured rate constants for the formation of additional native ribozyme. Under a wide range of solution conditions, these two ribozyme populations gave the same rate constants for native state formation within error, even as the changes in conditions gave rate constants that ranged nearly three orders of magnitude (Table 2, which is published as supporting information on the PNAS web site). Thus, it is apparently the same misfolded conformer of $E^{\Delta P5abc}$ that accumulates during folding and persists at equilibrium (Fig. 5).

Discussion

To populate a functional structure, an RNA sequence must encode a native architecture that is at least as stable as all other possible architectures combined. Although the strategies that RNAs use to stabilize their functional structures are certainly varied, a recurring theme is the presence of peripheral elements or proteins that interact with core elements. Here, we have shown that peripheral structure within the group I ribozyme from *Tetrahymena* stabilizes the native conformation relative to a particularly long-lived misfolded structure. Although it has been suggested that such "structural specificity" is a critical part of stabilizing RNA structure in general (3), whether peripheral structure can provide such specificity had not been addressed for any RNA.

We found that the P5abc peripheral element binds several orders of magnitude stronger to the native conformer of the ribozyme missing this element than to the misfolded conformer. This result indicates that P5abc stabilizes the native conformer relative to the misfolded one by the same amount, or that it provides 3.5–7 kcal/mol of structural specificity (with the exact amount depending on solution conditions). This energetic effect is enormous, larger than the free-energy change for global folding of many proteins

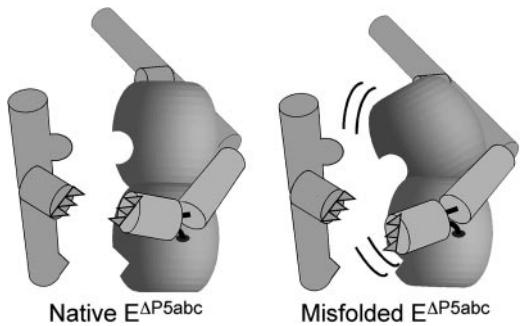

Fig. 6. Tighter binding of P5abc to the native $E^{\Delta P5abc}$ ribozyme than the misfolded ribozyme is suggested to arise from more favorable preorientation of the contact partners for P5abc in the native conformation. Loss of preorientation in the misfolded conformation is depicted schematically as displacement of the contact-forming regions relative to each other, as well as increased mobility of these regions (black curves).

(45), indicating a remarkably high degree of specificity for the formation of RNA tertiary structure.

How does P5abc stabilize the native state? P5abc consists of two coaxially stacked helices and a third helix that originates from a three-helix junction (23, 46, 47). When P5abc binds the ribozyme, it forms an interface in which each helix of P5abc forms a specific set of tertiary interactions with "contact partners" within the ribozyme (Fig. 1) (22). The loop L5b forms a tetraloop–receptor interaction with J6a/6b, and residues within the A-rich bulge of P5a form stacking interactions and hydrogen bonds with residues of P4. In addition, L5c forms a base-pairing interaction with the L2 loop. Because of these extensive interactions, the locations and orientations of diverse parts of the ribozyme are presumably strongly influenced by the structure of P5abc. The stronger binding of P5abc to the native ribozyme likely results from a more favorable prealignment of the contact sites in the native conformation, allowing the contacts to form cooperatively by minimizing the need for unfavorable rearrangements (Fig. 6). There is certainly some rearrangement of the native ribozyme upon binding P5abc, because P5abc binding increases hydroxyl radical protection of residues away from its contact sites and restores catalytic activity to the native ribozyme (23, 24). However, the overall patterns of protection from hydroxyl radicals are similar for the native $E^{\Delta P5abc}$ ribozyme and the native WT ribozyme, and the $E^{\Delta P5abc}$ ribozyme retains some catalytic activity (17, 23, 24), suggesting that $E^{\Delta P5abc}$ adopts a native fold that is similar to that of the WT ribozyme and thus that the contact partners for P5abc binding are likely to be crudely prealigned. In contrast, structural differences in the misfolded ribozyme presumably lead to a lower degree of prealignment, giving less cooperativity for formation of the tertiary contacts with P5abc and therefore weaker binding. It will be interesting to determine whether the stabilization of the native structure is a unique effect of P5abc or whether specificity is lost upon disruption of any tertiary contact, which would suggest that the specificity arises from global cooperativity of tertiary structure formation.

We also found that, in the absence of P5abc, the ribozyme equilibrates nearly equally between the native and misfolded species. Thus, the selective stabilization by P5abc is necessary to generate a significant thermodynamic preference for the native state. On one hand, this finding highlights the ability of RNA to fold to multiple conformations, a property that has been suggested to be crucial for the evolution of functional RNAs (48, 49) and for the ability of RNA to catalyze complex reactions like pre-mRNA splicing and translation, which require extensive conformational rearrangements through steps of the reactions (50, 51). On the other hand, the loss of native stability in the absence of P5abc raises the question of how other group I RNAs stabilize their native folds relative to alternative folds, because many do not contain P5abc, and some lack most of the peripheral structures of the *Tetrahymena* RNA (52). For RNAs that lack extensive peripheral structure, it is possible that features in their cores favor the native conformation or that the limited peripheral contacts they do possess can efficiently constrain the RNA to the native conformation. It is also possible that for many group I RNAs, proteins that bind to the native structures of the RNAs (53, 54) function in part to stabilize the native conformation relative to alternative folds. It will be fascinating to explore the extent to which RNA peripheral elements and RNA-binding proteins, in general, function by stabilizing the functional structures of their cognate RNAs relative to alternative structures.

We thank Rhiju Das and Vesselin Diankov for advice on Fe(II)-EDTA footprinting and 3′ end-labeling of RNA; Alain Laederach, Rhiju Das, Sam Pearlman, and Russ Altman for allowing us to use the SAFA software while it was in development; and Mark Engelhardt for preliminary experiments. This work was funded by Welch Foundation Grant F-1563 (to R.R.) and National Institutes of Health Grants R01-GM070456 (to R.R.) and P01-GM066275 (to D.H.).

1. Sigler, P. B. (1975) *Annu. Rev. Biophys. Bioeng.* **4,** 477–527.
2. Tinoco, I., Jr., & Bustamante, C. (1999) *J. Mol. Biol.* **293,** 271–281.
3. Fang, X. W., Golden, B. L., Littrell, K., Shelton, V., Thiyagarajan, P., Pan, T. & Sosnick, T. R. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 4355–4360.
4. Crothers, D. M. & Ratner, D. I. (1968) *Biochemistry* **7,** 1823–1827.
5. Gellman, S. H., Haque, T. S. & Newcomb, L. F. (1996) *Biophys. J.* **71,** 3523–3526.
6. McKay, S. L., Haptonstall, B. & Gellman, S. H. (2001) *J. Am. Chem. Soc.* **123,** 1244–1245.
7. Michel, F. & Westhof, E. (1990) *J. Mol. Biol.* **216,** 585–610.
8. Michel, F. & Ferat, J. L. (1995) *Annu. Rev. Biochem.* **64,** 435–461.
9. Lehnert, V., Jaeger, L., Michel, F. & Westhof, E. (1996) *Chem. Biol.* **3,** 993–1009.
10. Qin, H., Sosnick, T. R. & Pan, T. (2001) *Biochemistry* **40,** 11202–11210.
11. Swisher, J., Duarte, C. M., Su, L. J. & Pyle, A. M. (2001) *EMBO J.* **20,** 2051–2061.
12. Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. (2000) *Science* **289,** 905–920.
13. Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Jr., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T. & Ramakrishnan, V. (2000) *Nature* **407,** 327–339.
14. Mears, J. A., Cannone, J. J., Stagg, S. M., Gutell, R. R., Agrawal, R. K. & Harvey, S. C. (2002) *J. Mol. Biol.* **321,** 215–234.
15. Price, J. V., Kieft, G. L., Kent, J. R., Sievers, E. L. & Cech, T. R. (1985) *Nucleic Acids Res.* **13,** 1871–1889.
16. Doudna, J. A. & Szostak, J. W. (1989) *Mol. Cell. Biol.* **9,** 5480–5483.
17. Joyce, G. F., van der Horst, G. & Inoue, T. (1989) *Nucleic Acids Res.* **17,** 7879–7889.
18. Salvo, J. L. & Belfort, M. (1992) *J. Biol. Chem.* **267,** 2845–2848.
19. van der Horst, G., Christian, A. & Inoue, T. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 184–188.
20. Murphy, F. L. & Cech, T. R. (1993) *Biochemistry* **32,** 5291–5300.
21. Murphy, F. L. & Cech, T. R. (1994) *J. Mol. Biol.* **236,** 49–63.
22. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996) *Science* **273,** 1678–1685.
23. Doherty, E. A., Herschlag, D. & Doudna, J. A. (1999) *Biochemistry* **38,** 2982–2990.
24. Engelhardt, M. A., Doherty, E. A., Knitt, D. S., Doudna, J. A. & Herschlag, D. (2000) *Biochemistry* **39,** 2639–2651.
25. Fang, X., Pan, T. & Sosnick, T. R. (1999) *Biochemistry* **38,** 16840–16846.
26. Russell, R. & Herschlag, D. (1999) *J. Mol. Biol.* **291,** 1155–1167.
27. Russell, R. & Herschlag, D. (2001) *J. Mol. Biol.* **308,** 839–851.
28. Treiber, D. K. & Williamson, J. R. (2001) *J. Mol. Biol.* **305,** 11–21.
29. Zaug, A. J., Grosshans, C. A. & Cech, T. R. (1988) *Biochemistry* **27,** 8924–8931.
30. Russell, R. & Herschlag, D. (1999) *RNA* **5,** 158–166.
31. Huang, Z. & Szostak, J. W. (1996) *Nucleic Acids Res.* **24,** 4360–4361.
32. Herschlag, D., Eckstein, F. & Cech, T. R. (1993) *Biochemistry* **32,** 8299–8311.
33. Rose, I. A., O'Connell, E. L. & Litwin, S. (1974) *J. Biol. Chem.* **249,** 5163–5168.
34. Herschlag, D. & Cech, T. R. (1990) *Biochemistry* **29,** 10159–10171.
35. Latham, J. A. & Cech, T. R. (1989) *Science* **245,** 276–282.
36. Das, R., Laederach, A., Pearlman, S. M., Herschlag, D. & Altman, R. B. (2005) *RNA* **11,** 344–354.
37. Williamson, J. R. (2000) *Nat. Struct. Biol.* **7,** 834–837.
38. Leulliot, N. & Varani, G. (2001) *Biochemistry* **40,** 7947–7956.
39. Rose, M. A. & Weeks, K. M. (2001) *Nat. Struct. Biol.* **8,** 515–520.
40. Webb, A. E., Rose, M. A., Westhof, E. & Weeks, K. M. (2001) *J. Mol. Biol.* **309,** 1087–1100.
41. Tomizawa, J. (1990) *J. Mol. Biol.* **212,** 683–694.
42. Persson, C., Wagner, E. G. & Nordstrom, K. (1990) *EMBO J.* **9,** 3777–3785.
43. Herschlag, D. (1992) *Biochemistry* **31,** 1386–1399.
44. Fersht, A. (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (Freeman, New York).
45. Plaxco, K. W., Simons, K. T., Ruczinski, I. & Baker, D. (2000) *Biochemistry* **39,** 11177–11183.
46. Wu, M. & Tinoco, I., Jr. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 11555–11560.
47. Deras, M. L., Brenowitz, M., Ralston, C. Y., Chance, M. R. & Woodson, S. A. (2000) *Biochemistry* **39,** 10975–10985.
48. Herschlag, D. (1995) *J. Biol. Chem.* **270,** 20871–20874.
49. Schultes, E. A. & Bartel, D. P. (2000) *Science* **289,** 448–452.
50. Madhani, H. D. & Guthrie, C. (1994) *Annu. Rev. Genet.* **28,** 1–26.
51. Staley, J. P. & Guthrie, C. (1998) *Cell* **92,** 315–326.
52. Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Muller, K. M., *et al.* (2002) *BMC Bioinformatics* **3,** 15.
53. Lambowitz, A. M. & Perlman, P. S. (1990) *Trends Biochem. Sci.* **15,** 440–444.
54. Lambowitz, A. M., Caprara, M. G., Zimmerly, S. & Perlman, P. S. (1999) in *The RNA World*, eds. Gesteland, R. F., Cech, T. R. & Atkins, J. F. (Cold Spring Harbor Lab. Press, Woodbury, NY), pp. 451–485.